BANK POLICY INSTITUTE

September 2, 2021

Via Electronic Mail

Dr. James Olthoff
Acting NIST Director and Undersecretary of Commerce for Standards and Technology
U.S. National Institute of Standards and Technology
100 Bureau Drive (MS 20899)
Gaithersburg, MD 20899-2000

Re: Request for Information on the Artificial Intelligence Risk Management Framework (Docket No. [210726-0151])

Dear Dr. Olthoff:

The Bank Policy Institute[1] appreciates the opportunity to comment on the National Institute of Standards and Technology's Request for Information on the Artificial Intelligence Risk Management Framework. The development of the AI RMF is an important effort that will heighten awareness and help organizations across all industries understand and manage AI risks.

Banks have a strong history and culture of risk management and have decades of experience in designing processes to manage risks related to emerging technologies, including AI. The financial services sector is unique and stands out from other industries in that banks are subject to extensive regulatory requirements which provide a comprehensive framework to manage the implementation of AI across various banking use cases. We believe NIST's AI RMF will help other organizations not already subject to similar requirements improve their awareness of AI risks and implement a governance structure to ensure AI is used in a responsible and trustworthy manner.

In this letter, BPI provides high-level comments and recommendations for NIST's consideration in developing the AI RMF, based upon the financial sectors' experience adopting AI and expanding

[1] The Bank Policy Institute is a nonpartisan public policy, research and advocacy group, representing the nation's leading banks and their customers. Our members include universal banks, regional banks and the major foreign banks doing business in the United States. Collectively, they employ almost two million Americans, make nearly half of the nation's small business loans, and are an engine for financial innovation and economic growth.

existing model governance frameworks to AI. For additional information, especially to address NIST's second goal to gain a greater awareness about how organizations are managing AI risk and have incorporated risk management standards into policies and practices, we encourage NIST to reference BPI's response to the joint financial regulators' Request for Information on Financial Institutions' Use of AI, including Machine Learning,[2] provided as an Appendix to this letter. In addition to providing an overview of how banks are currently utilizing AI and managing any associated risks, BPI detailed the existing laws, regulations and guidance that provide a comprehensive risk management framework for AI in the financial sector, all of which relates directly to several of the questions posed in NIST's RFI.

BPI and its member banks agree with and support the eight proposed attributes of the AI RMF. In particular, BPI appreciates that the Framework is intended to be voluntary, risk-based, useful to a variety of stakeholders and adaptable over time as the technology and AI applications continue to evolve. Our suggestions, therefore, are modest in scope, and intend to emphasize the importance of certain attributes as NIST continues to develop the AI RMF.

I. Ensure the AI RMF is Consistent with Existing Regulatory Requirements

While we appreciate NIST's effort to remain agnostic to any specific law or regulation, we emphasize that NIST's AI RMF should be interoperable and consistent with existing regulatory requirements across various industries. As previously noted, banks are subject to extensive regulatory requirements, many of which align with the characteristics of trustworthiness outlined in the RFI, specifically privacy, security and mitigation of bias.[3] For example, the Gramm-Leach-Bliley Act (GLBA) requires financial institutions to protect the privacy and security of personally identifiable financial information related to individuals, and the federal banking agencies' Interagency Guidance Establishing Security Standards, implementing guidance of the GLBA, requires banks to develop, implement and maintain an information security program.[4] Additionally, the Equal Credit Opportunity Act, along with its implementing regulation, Regulation B, prohibits discrimination in credit transactions.[5] While not specific to AI, these laws and regulations provide the foundation for a comprehensive framework for managing AI risks over the specific activities in which AI is being used in financial services. The AI RMF should be designed in a way that voluntary adoption of NIST's framework would not conflict with banks' existing regulatory obligations.

Various U.S. federal regulators are currently considering laws, regulations and guidance around the use of AI. As previously referenced, the financial regulators recently issued an RFI on Financial Institutions' Use of AI, including Machine Learning. The Federal Trade Commission (FTC) provided

---

[2] Office of the Comptroller of the Currency, Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, Bureau of Consumer Financial Protection, National Credit Union Administration; Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning, 86 Fed. Reg. 16837, (Mar. 31, 2021).

[3] For a detailed list of financial laws and regulations that may be relevant to AI, refer to the appendix contained within the financial agencies' Request for Information on Artificial Intelligence, including Machine Learning, available at https://www.federalregister.gov/documents/2021/03/31/2021-06607/request-for-information-and-comment-on-financial-institutions-use-of-artificial-intelligence.

[4] 12 C.F.R. pt. 30, App. B

[5] 15 U.S.C. §1691 *et seq.*; 12 C.F.R. pt. 1002

guidelines on using AI and algorithms in April 2020, and more recently issued a blog post on truth, fairness and equity in AI in April 2021.[6] In its most recent publication, the FTC encouraged organizations to embrace transparency and independence by using transparency frameworks and independent standards, among other methods, and defined unfairness as any act that "causes more harm than good".[7] In developing the AI RMF, NIST has an opportunity to harmonize frameworks and definitions across various industries and should work with federal regulators to ensure consistency with existing requirements, and where possible, remove any ambiguity. NIST may also consider the analysis from a recent Bank for International Settlements paper, published in August 2021, which outlined a role for standards-setting bodies to develop international guidance or standards for using AI in the financial sector.[8]

II. Ensure the AI RMF is Risk-Based and Compatible with Existing Risk Management Frameworks

BPI agrees that the AI RMF should be risk-based and non-prescriptive, and appreciates that the Framework is intended to be consistent with other approaches to managing AI risk. Banks currently implement comprehensive risk management and corporate governance processes over AI models, as outlined by Supervisory Guidance on Model Risk Management ("Model Risk Management Guidance" or "Guidance"). The Guidance, issued jointly by the Office of the Comptroller of the Currency and the Federal Reserve Board and subsequently adopted by the Federal Deposit Insurance Corporation, provides a flexible, risk-based approach to managing AI risks, where controls can be scaled or enhanced based on the overall risk entailed. Specifically, the Guidance requires banks to develop effective model risk management frameworks, including robust model development, implementation and use; effective validation; and sound governance, policies and controls.[9]

The Model Risk Management Guidance is unique to the financial sector; other industries do not have such comprehensive guidance for managing AI models. In addition to ensuring that the AI RMF is compatible with banks' existing risk management frameworks, NIST should consider using the Model Risk Management Guidance as a resource for developing the AI RMF. The core concepts outlined in the Guidance provide an effective framework for managing AI risks in the financial sector that could be applied more broadly to other industries. For example, the Guidance emphasizes the importance of sound model development and validation, and risk management processes commensurate with the materiality of the model. It also outlines processes for effective challenge to ensure models are fit for

---

6 Andrew Smith, "Using Artificial Intelligence and Algorithms" (Apr. 8, 2020), Federal Trade Commission, https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms; and Elisa Jillson, "Aiming for truth, fairness, and equity in your company's use of AI" (Apr. 19, 2021), Federal Trade Commission, https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai.

7 *Id.*

8 Prenio, Jermy; Young, Jeffery, "Humans keeling AI in check – emerging regulatory expectations in the financial sector," Bank for International Settlements (Aug. 2021), https://www.bis.org/fsi/publ/insights35.pdf.

9 FRB, SR 11-7, Supervisory Guidance on Model Risk Management (Apr. 4, 2011), https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf; OCC, Bulletin 2011-12, Supervisory Guidance on Model Risk Management (Apr. 4, 2011), https://occ.gov/news-issuances/bulletins/2011/bulletin-2011-12a.pdf ; FDIC, FIL-22-2017, Adoption of Supervisory Guidance on Model Risk Management (June 7, 2017), https://www.fdic.gov/news/financial-institution-letters/2017/fil17022.pdf.

purpose and ongoing monitoring so that models are working as intended. Additionally, the Guidance provides a template for organizational structure, outlining roles for ownership and accountability of risks, indicating that the board and senior management are ultimately responsible for maintaining an effective model risk management framework.[10] Accountability, highlighted as a key principle of the AI RMF in the RFI, is particularly important in managing AI risks. NIST should consider incorporating expectations for establishing clear lines of responsibility and accountability for AI, aligning with traditional enterprise risk management frameworks. These expectations, however, should be non-prescriptive in nature, given that appropriate governance structures may differ across industries and specific organizations.

We appreciate that NIST is evaluating how organizations take into account both the benefits and challenges related to inclusiveness in AI design, development, use and evaluation, as detailed in question 8 of the RFI. It is important to recognize that in some cases, while AI may present certain risks, it may nonetheless be significantly safer and sounder, or more effective, than non-AI approaches currently used. For example, dynamic updating models or reinforced learning techniques may introduce challenges as they evolve over time, but may also provide important benefits, such as by improving a banks' ability to detect fraud and ultimately better protect its' customers. Banks are currently exploring the potential benefits of dynamic updating models and other advanced techniques to institutions and customers, and are doing so in a responsible, risk-based manner by scaling and adapting controls under their risk-management frameworks. The AI RMF should be designed in a way that enables organizations to effectively manage the tradeoff between opportunities and risks in AI applications, not losing sight of the benefits that AI may deliver.

lll. Establish Common Definitions for Aspects of AI Risk

BPI supports NIST's efforts to develop common definitions and an AI risk taxonomy, and believes NIST is appropriately suited to undertake this effort. Establishing common definitions related to AI has been an ongoing challenge in the financial services sector, due to the evolving nature of the technology and increasing implementation of AI across various use cases. For this reason, we appreciate that NIST is focusing its efforts on defining AI characteristics, such as trust and trustworthiness, rather than developing specific definitions for the technical aspects of AI. We believe that a common taxonomy for discussing AI risks will enable firms across all industries communicate AI risks in a consistent manner for stakeholders to understand. Additional efforts similar to NIST's draft four principles of explainability, which provided a valuable starting point for how we think about and discuss explainability of AI systems, will help contribute to the broader goal of creating an AI risk taxonomy.[11]

lV. Ensure the AI RMF is Forward-Looking and Adaptable Over Time

BPI appreciates that the AI RMF is intended to be a living document and capable of being readily updated over time. The financial sector is only just scratching the surface of the potential benefits that AI may provide for financial institutions and consumers, and the AI RMF should be designed in a way

---

10 *Id.*

11 National Institute of Standards and Technology, "Four Principles of Explainable Artificial Intelligence", Draft NISTIR 8312, (August 2020) *available at* https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8312-draft.pdf, and BPI's comment letter response, *available at* https://bpi.com/wp-content/uploads/2021/01/2020.10.15-BPI-Comments-on-the-Four-Principles-of-Explainable-Artificial-Intelligence-NISTIR-8312.pdf.

that focuses on how to manage emerging risks as new AI technologies and applications are uncovered. To that end, we believe that a risk-based and non-prescriptive approach, as previously discussed, is the appropriate way to ensure the Framework is adaptable over time. For example, dynamic updating models, which have the capacity to update on their own sometimes without human interaction, are currently being applied to relatively limited applications in financial services as banks assess potential risks related to model drift. However, the application of dynamic updating models may increase over time as we continue to learn how to responsibly manage related risks. The AI RMF should be technology-agnostic and respective of the future, and refrain from creating both overly narrow and overly broad requirements over technologies as our understanding of AI-related risks evolves over time.

* * * * *

BPI appreciates NIST's efforts to develop the AI RMF in an open, transparent process involving various stakeholders. Thank you for the opportunity to respond to the request for information and we look forward to future engagement with NIST on this subject. If you have any questions, please contact the undersigned by phone at 202-589-2432 or by email at *Stephanie.Wake@bpi.com.*

Respectfully submitted,

Stephanie Wake
Vice President, BITS
Bank Policy Institute

APPENDIX

June 25, 2021

Via Electronic Mail

Chief Counsel's Office
Attention: Comment Processing
Office of the Comptroller of the Currency
400 7th Street SW
Suite 3E-218
Washington, DC 20219

Ann E. Misback
Secretary
Board of Governors of the Federal Reserve System
20th Street & Constitution Avenue NW
Washington, DC 20551

James P. Sheesley
Assistant Executive Secretary
Federal Deposit Insurance Corporation
550 17th Street NW
Washington, DC 20429

Comment Intake
Bureau of Consumer Financial Protection
1700 G Street NW
Washington, DC 20552

Melane Conyers-Ausbrooks
Secretary of the Board
National Credit Union Administration
1775 Duke Street, Alexandria, VA 22314-3428

Re: Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning (Docket No. OCC-2020-0049; OP-1743; RIN 3064-ZA24; CFPB 2021-0004; NCUA 2021-0023)

To Whom It May Concern:

The Bank Policy Institute[1] appreciates the opportunity to respond to the request for information and comment relating to financial institutions' use of artificial intelligence, including machine learning.[2]

---

[1] The Bank Policy Institute is a nonpartisan public policy, research, and advocacy group, representing the nation's leading banks and their customers. Our members include universal banks, regional banks, and the major foreign banks doing business in the United States. Collectively, they employ almost 2 million Americans, make nearly half of the nation's small business loans, and are an engine for financial innovation and economic growth.

[2] Office of the Comptroller of the Currency, Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, Bureau of Consumer Financial Protection, National Credit Union

Artificial intelligence (AI) includes a family of technologies capable of performing tasks that traditionally would have required human cognitive intelligence, such as thinking and decision-making. Machine learning (ML) is a subset of AI that generally refers to the ability of a software algorithm to identify patterns and automatically optimize and refine performance from processing large data sets.[3] Traditional, statistical methods, such as regression-based models and simple, rules-based automation that replicates human actions are not typically considered AI. For purposes of this comment letter, we refer to AI generally throughout our response.

The banking sector is strongly committed to promoting the responsible use of AI given the potential long-term benefits for consumers and the future of financial products. The adoption of AI throughout financial services varies by institution and will continue to evolve as we learn more about the benefits of AI's capabilities. To that end, BPI supports the Agencies' coordinated efforts to gain more information on the use of AI in financial services, including how financial institutions ensure the utility of AI outputs and manage model risk. We emphasize, however, that AI is a technology like any other, and the risks posed by AI as outlined in the RFI can be managed within existing laws and regulations on the activities in which AI is applied across the financial industry. BPI believes that new regulations are not necessary, and that the Agencies should apply a flexible, principled, and risk-based approach for the risk assessment, implementation and oversight of AI. Through this approach, the Agencies have an opportunity to encourage the responsible use of AI in financial services consistent with safety and soundness standards, consumer protection and principles of fairness.

In Part I of this letter, BPI proposes principles for the Agencies to consider in evaluating the current regulatory framework surrounding AI. Part II of this letter responds to the specific questions raised in the RFI, elaborating on these principles, and identifying areas where clarification by the Agencies would be useful to facilitate the responsible use of AI within financial services. BPI looks forward to further engaging with the Agencies on this subject. The evolution of the capabilities of AI, and of the compliance efforts that reduce the risks of AI, require an ongoing dialogue between banks and the Agencies on the issues presented in this RFI.

I. Guiding Principles for the Agencies' Regulatory Review

BPI agrees that an assessment of risk management practices related to the use of AI is an important step in evaluating this innovative technology. However, the RFI focuses on specific risks over limited business applications, and the Agencies should not lose sight of the broad nature of the technology and the bigger-picture view of how AI is being implemented across banks. AI, like human intelligence, is a critical resource in improving financial services. Uses of new technologies in banking have at times been met with initial resistance and regulatory uncertainty, only to ultimately become essential components of the financial system. The Agencies should also recognize that banks are already subject to applicable, comprehensive regulatory requirements that can be applied to AI. BPI's response here is designed to help the Agencies understand the benefits of AI's capabilities, while ensuring safety

---

Administration (collectively, the "Agencies"); Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning, 86 Fed. Reg. 16837 (Mar. 31, 2021).

[3] For additional detail on the definitions of AI and ML, see Artificial Intelligence: Recommendations for Principled Modernization of the Regulatory Framework, Bank Policy Institute and Covington & Burling LLP, (September 14, 2020), https://bpi.com/wp-content/uploads/2020/10/Artificial-Intelligence-Recommendations-for-Principled-Modernization.pdf.

and soundness and appropriate consumer protections. The following principles should guide the Agencies' regulatory review.

A. The Agencies should conduct a balanced assessment of the benefits and risks of AI in financial services.

The adoption of AI throughout the financial sector has the potential to significantly improve financial outcomes for both businesses and consumers. AI can capture and process broader and deeper data sets and can use both more sophisticated analytical tools and powerful new computing capabilities to enhance bank processes and operations. By embracing AI, banks are able to make more informed decisions, optimize back-office operations, reduce compliance and operational risk and provide personalized customer experiences, transforming business functions and resulting in cost efficiencies. AI has the potential to provide advantages to consumers, such as by improving customer communications or providing customized financial products and services that empower consumers to better their financial lives.

While the RFI recognizes that AI has benefits, the questions for comment focus on a limited set of prospective challenges and risks that AI may present. The regulation and supervision of the use of AI should not focus on the potential risks in a vacuum. Just as the Agencies assess the risks of AI, and steps that can be taken to mitigate those risks, they should also assess the potential incremental benefits of AI, and how to facilitate the use of AI and other innovative technologies to benefit consumers and financial institutions. This approach is consistent with the Office of Management and Budget's (OMB) November 2020 Guidance for Regulation of Artificial Intelligence Applications, which notes, "Agencies should, when consistent with law, carefully consider the full societal costs, benefits and distributional effects when considering regulations related to the development and deployment of AI applications."[4] BPI encourages the Agencies to follow the principles set forth by the OMB and publicly post their plans for achieving consistency with the Guidance.

Banks currently utilize AI and ML in a wide variety of operations, including but not limited to fraud detection and prevention, marketing, customer service, cybersecurity, anti-money laundering, credit underwriting and back-office processing. The current application of AI within financial services varies by institution, and continues to evolve as we learn more about possible applications of AI and the interactions between artificial and human intelligence both within organizations and externally with customers. The below examples illustrate some of the bank functions where AI holds promise in improving bank operations:

- Fraud Detection and Prevention: As payments fraud has increased in volume and complexity, AI models have become increasingly important to fraud detection. Changes in digital footprints and patterns have made fraudulent attacks difficult to detect using rules-based logic. AI models using predictive analytics can find anomalies in transactions, proactively identifying outliers that do not conform with clients' past patterns or payment activity. Certain ML models can identify relationships in activity using historical data, which can be used to identify transactions that are most likely to be fraudulent, allowing human investigations to focus on high-risk cases. These models not only improve the performance

[4] Office of Management and Budget, Guidance for Regulation of Artificial Intelligence Applications, (Nov. 17, 2020), https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf.

of banks' fraud detection capabilities, but also help catch fraudulent activity before it impacts customers.

- Customer Service: Shifts in consumer demand for more digital and interactive financial products and services have dramatically changed the financial marketplace, which now includes an increasing number of nonbank fintechs and other companies facilitating access to consumer data to provide such products and services. AI plays an increasing role in customer services due to its ability to create a more personalized experience, as banks can utilize AI to gain a better understanding of a consumer's needs to provide customized products and communications. For example, dynamic updating models used in marketing can capture real-time feedback from customers and provide content most relevant to prior customer responses. By understanding a consumer's needs more precisely, banks can offer more personalized experiences and tailored products that allow customers to more effectively manage their financial livelihoods and enable banks to compete in an increasingly digitized marketplace.

- Cybersecurity: The financial services sector continues to be a primary target for cybersecurity attacks, which aim to cause devastating financial losses affecting individuals, organizations, and potentially the entire financial sector. Banks may use AI to detect and respond to cyberattacks more quickly and efficiently than human intelligence alone. For example, banks can utilize natural language processing for email monitoring to detect and identify threats such as phishing attacks. Additionally, AI-based network security software can monitor incoming and outgoing network traffic to identify suspicious patterns in the data traffic. As cyber criminals continue to exploit vulnerabilities with more sophisticated cyberattacks, AI can be particularly helpful in enhancing cybersecurity activities.

- Anti-Money Laundering: Banks devote significant resources to the detection and reporting of suspicious activity in compliance with long-established anti-money laundering (AML) and countering the financing of terrorism (CFT) regulatory requirements. AI models have the potential to improve detection of suspicious activity as they can understand complex patterns in data, resulting in the detection of unusual activity and reduction of false positives. In addition, banks may utilize natural language processing applications to automatically generate Suspicious Activity Reports by evaluating large volumes of unstructured data and converting such data into text, replacing work typically conducted by investigators and allowing investigators to focus efforts on responding to a smaller number of higher-risk activities. Given the potential of AI to improve AML/CFT detection and reporting, the recent Anti-Money Laundering Act of 2020, embedded within the annual National Defense Authorization Act, commissions the Financial Crimes Enforcement Network (FinCEN) to evaluate the implementation of AI and other emerging technologies to improve U.S. AML/CFT efforts.[5]

- Credit Underwriting: AI credit underwriting systems represent a new type of automated credit underwriting that may be better in evaluating creditworthiness and provide opportunities to enhance fair, unbiased, and more accurate lending. With the help of AI, banks may be able to process more data, including potentially alternative or nontraditional

[5] Anti-Money Laundering Act of 2020, §6211(f).

data, within existing credit decisioning engines. Such processes have the potential to reduce underwriting times and delinquency rates.

- Back-Office Management: Some banks are beginning to utilize AI models to replace or supplement processes and workflows for several back-office purposes typically consisting of repetitive, routine and clerical tasks. For example, banks may use AI to resolve IT issues and provide bandwidth back to service desks to support customers, or to identify common themes in customer complaints to improve customer service. For these types of tasks, manual processing can be slow and costly and can lead to inconsistent results. AI can improve and expedite these business processes and make employees more efficient.

The above examples demonstrate the promise that AI-based applications hold for both organizations and consumers. Indeed, it is hard to predict where AI will not be used in the future. As with any new tool or technology used by banks, the implementation of AI in financial services may encompass risks that must be managed in an appropriate, risk-based manner. Further, the implementation and maintenance of AI requires significant investment by banks to both build the AI model and maintain appropriate guardrails to ensure the model does not overstep its bounds. As a result, banks are taking a measured approach to implementing AI that includes an assessment of internal operational costs and potential return on investment in addition to an evaluation of the benefits and risks.

Regulatory expectations around the use of AI in financial services should be based on a balanced analysis and understanding of both the potential risks and potential rewards of AI, whether as an alternative or supplement to existing non-AI approaches. It is important for the Agencies to consider whether both the risks and the benefits of AI-based approaches are greater or lesser than the non-AI-based approach it would supplement or replace. In some cases, while innovation may present certain risks, it may nonetheless be significantly safer and sounder, or more effective, than non-AI approaches currently used. This risk-based approach is consistent with the approach currently taken by the Agencies in assessing non-AI functions of banks. BPI elaborates on the tradeoffs between opportunities and risks based on the particular application of AI, and how banks manage such risks, in response to the specific RFI questions below.

B. The Agencies should avoid creating or applying new regulatory expectations that may hinder progress in using this evolving technology.

Consumers are best served by regulatory approaches that are not static or rigid, but are sufficiently flexible and adaptable to the emergence of new technologies and methods of providing financial products and services. This is particularly true of AI, as one of its strengths is that it is constantly evolving and improving. As noted in the 2020 OMB AI Guidance, "Rigid, design-based regulations that attempt to prescribe the technical specifications of AI applications will in most cases be impractical and ineffective, given the anticipated pace with which AI will evolve and resulting need for agencies to react to new information and evidence."[6]

Existing banking regulations and guidance provide a comprehensive framework to manage the implementation of AI across various banking use cases. Banks are subject to extensive regulatory

[6] Office of Management and Budget, Guidance for Regulation of Artificial Intelligence Applications, (Nov. 17, 2020), https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf.

requirements, including with regard to how they manage model risk, store and protect sensitive information, ensure consumer privacy, defend against cyberattacks, manage third parties and engage in fair lending. To ensure appropriate compliance, banks are subject to regular supervision by prudential banking agencies and the CFPB. The Agencies detailed many of these laws and regulations as well as supervisory guidance and other statements relating to safety and soundness and consumer protection as an Appendix to the RFI.

Regulations should remain technology-neutral and focus on the activities that take place, rather than the technology itself. New specific regulations related to AI would be counterproductive given the heterogeneity in which AI is used, as detailed in the above examples, the evolving nature of the technology and the complexity in specifically defining aspects of AI. AI is just another technology, and does not pose unique risks that cannot be managed within existing regulations and risk frameworks over the specific activities in which AI is being used. Additionally, the adoption of AI in financial services is likely to evolve as we learn more about the blend of AI and human intelligence and the various AI applications. The Agencies should avoid creating a new, prescriptive framework around AI that may prevent the financial industry and its customers from realizing future benefits of AI. The balance between regulations and innovation is significant, and financial institutions may risk missing out on applying or developing innovative solutions if regulatory burdens become too restrictive.

C. The Agencies should apply a principled, risk-based approach for the risk assessment, implementation and oversight of AI.

Banks have a strong history and culture of risk management and have decades of experience in designing processes to manage risks inherent to banking operations and to ensure consumer protections. Banks are attentive to the importance of AI and the corresponding need to manage any financial, reputational or legal risks posed by AI. To that end, institutions are engaged in an ongoing, extensive process to evaluate the capabilities of AI to benefit business operations, consumers and financial services as a whole, while implementing existing and developing new oversight processes to manage any associated risks. For example, the introduction of AI may change the risks that banks manage today from operational risk driven by manual execution of processes to an increase in data, model and technology risk driven by increased automation. In evaluating AI capabilities and risks, one of the primary guidance documents that banks utilize to ensure risks are appropriately managed is the Supervisory Guidance on Model Risk Management (hereafter, "Model Risk Management Guidance" or "Guidance").[7]

The Model Risk Management Guidance is generally principles-based and flexible enough to cover risks related to AI. The Guidance requires banks to develop effective model risk management frameworks, including: robust model development, implementation and use; effective validation; and sound governance, policies and controls. These principles are being applied to address risks related to AI, such as those presented within the RFI, including explainability, data quality and data processing, overfitting and dynamic updating. Further, controls to mitigate these risks can be scaled or enhanced appropriately depending on the complexity, materiality and application of AI models and the overall risk

---

[7] FRB, SR 11-7, Supervisory Guidance on Model Risk Management (Apr. 4, 2011), https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf; OCC, Bulletin 2011-12, Supervisory Guidance on Model Risk Management (Apr. 4, 2011), https://occ.gov/news-issuances/bulletins/2011/bulletin-2011-12a.pdf; FDIC, FIL-22-2017, Adoption of Supervisory Guidance on Model Risk Management (June 7, 2017), https://www.fdic.gov/news/financial-institution-letters/2017/fil17022.pdf.

entailed. BPI agrees with Federal Reserve Governor Lael Brainard's assessment that, "Guidance has to be read in the context of the relative risk and importance of the specific use-case in question" and "the level of scrutiny should be commensurate with the potential risk posed by the approach, tool, model or process used."[8]

BPI recommends that the Agencies further emphasize the flexibility that banks have in determining how they assess the risk of AI models within their model risk management frameworks. Banks have deep experience and expertise in assessing and addressing risks within the existing regulatory framework. Expressly recognizing that banks have flexibility in how they apply the Model Risk Management Guidance to AI models based on perceived risk would encourage the responsible adoption of AI. BPI appreciated the recent statement by the federal banking agencies addressing how the Model Risk Management Guidance relates to systems used in complying with Bank Secrecy Act (BSA)/AML requirements. Specifically, in addition to emphasizing that the Guidance does not have the force and effect of law, the interagency statement provides helpful clarifications that will better facilitate flexibility in the extent to which, and how, institutions apply the principles of the Guidance to BSA/AML tools, including those determined to be models.[9]

BPI also recommends that the Agencies take a flexible and risk-based approach in overseeing banks' implementation of AI. From a supervisory perspective, the Agencies should not place a greater burden on AI models merely because they are characterized as such without evaluating whether the AI model in fact introduces greater risk compared to a traditional model. Instead, oversight of banks' use of AI should be appropriately tailored to the risk entailed by the business application of AI and the Agencies should expect a similar level of diligence as other models applied to similar contexts. BPI member banks do not currently believe there are risks unique to AI that cannot be controlled for within existing risk management frameworks and regulatory guidance. However, as the application of AI evolves across the financial industry, BPI encourages the Agencies to engage in ongoing dialogue with the banks to determine whether there are distinctive features of AI models that should be handled differently than traditional models, based on the underlying risk and benefits of the technology and application.

D. The Agencies should ensure that existing regulations and guidance are applied consistently across banks and nonbanks engaged in financial services.

Consumers should be equally protected regardless of what kind of entity they engage with for financial services. Banks have a long history of compliance with the Model Risk Management Guidance and are regularly supervised by prudential regulators. In contrast, nonbanks engaged in financial services are not subject to model risk management or regulatory compliance standards and are not regularly or consistently supervised by prudential regulators. In the fair lending context, while nonbank lenders using AI credit underwriting models or alternative data are required to comply with fair lending laws, they have no obligation to follow the Model Risk Management Guidance or answer to regulators through supervisory examinations. BPI believes that a lack of consistent standards between banks and nonbanks, and regulators overseeing these entities, puts consumers at risk by not affording equal

---

8 Governor Lael Brainard, "What Are We Learning about Artificial Intelligence in Financial Services?", Speech at Fintech and the New Financial Landscape, Philadelphia, Pennsylvania (Nov. 13, 2018), https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm.

9 FRB, FDIC, OCC, "Interagency Statement on Model Risk Management for Bank Systems Supporting Bank Secrecy Act/Anti-Money Laundering Compliance," (April 9, 2021), *available at* https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20210409a2.pdf.

protections to all consumers, especially those engaging outside the well-regulated banking industry. BPI recommends the Agencies coordinate to apply consistent standards for model risk management and oversight of AI across banks and nonbanks that provide financial services.

E. The Agencies should continue to coordinate to ensure AI is used in a safe and sound manner that protects consumers.

BPI applauds the Agencies for acting collectively in issuing the RFI to assess more broadly the regulatory framework governing AI. A coordinated approach is crucial to providing banks with clarity in the expectations around how to implement new technologies and the confidence to innovate accordingly. BPI encourages the Agencies to continue to coordinate to promote the responsible adoption of AI and ensure principles are applied consistently across the financial industry.

II. Comments on Specific Questions in the RFI

A. Explainability

The RFI poses a series of questions (Questions 1-3) related to risks from a lack of explainability. Before addressing the specific questions, BPI provides certain general observations to frame this topic.

Banks strive for an appropriate level of transparency in all business processes and applications and recognize that explainability is a key aspect to enabling trust, understanding and adoption of AI technologies. However, a one-size-fits-all approach to explainability does not exist. As Governor Brainard noted in January 2021, "Explanations serve a variety of purposes, and what makes a good explanation depends on the context."[10] The concept of AI explainability varies from the perspective of users, developers and owners. To that end, it is important to recognize the difference between (1) "explainability" as concerns a bank's ability to interpret and evaluate the system's efficacy, and (2) "explainability" as concerns the bank's ability to explain to an affected person, and for that affected person to accurately interpret, how and why the bank made a decision. The expectation in these two contexts involves different policy objectives and considerations across use cases. For example, banks may hold AI models that are customer-facing, such as those used for consumer lending, to a higher standard of external explainability than AI models that are used for internal operational processes, such as processing documents.

There should be different expectations of explainability based on the context in which the AI model is being used as well as the recipient of the explanation, which banks evaluate and determine within their risk management frameworks. This approach is consistent with Governor Brainard's comment that "Not all contexts require the same level of understanding of how ML models work."[11] Banks recognize that there are predefined areas where explainability is necessary, and in some instances, required by statute. However, there may also be instances where some models with low risk do not require any level of explainability. The Agencies should be careful not to extrapolate those requirements more broadly to impose explainability requirements where not necessary. Any

---

[10] Governor Lael Brainard, "Supporting Responsible Use of AI and Equitable Outcomes in Financial Services," speech at the AI Academic Symposium hosted by the Board of Governors of the Federal Reserve System, Washington, D.C. (Jan. 21, 2021), https://www.federalreserve.gov/newsevents/speech/brainard20210112a.htm.

[11] *Id.*

overarching requirement for explainability of AI models would significantly stifle innovation in financial services.

Further, the Agencies should consider the advantages posed by the complex reasoning and deeper use of data by AI against any explainability risks. Specifically, the utility of AI algorithms should be evaluated against the system they are replacing. For example, an AI algorithm may be more accurate than doctors in detecting cancer in patients but lack a specific explanation for why. [12] In this scenario, it would be a lost opportunity to disregard the results due to the complexity of the explanation.

Question 1: *How do financial institutions identify and manage risks relating to AI explainability? What barriers or challenges for explainability exist for developing, adopting, and managing AI?*

Banks are aware that a lack of explainability may pose risks. However, issues related to explainability are present in all models – not just AI models – and banks have various tools at their disposal to manage such risks. As previously indicated, banks have a strong culture of risk management, and long history in complying with the Model Risk Management Guidance. The evaluation and management of risks related to explainability of AI models is built into banks' model risk management framework and overall governance processes, and builds upon the evaluation of risk of traditional models. Banks dedicate significant attention to assessing risk and building a governance framework around the entire AI lifecycle, from development of AI models, to implementation and use of models, to continued oversight of outputs and performance. Explainability is one of several components within the model risk management framework and overall governance process to determine whether to use (or continue to use) an AI model.

Banks utilize a principled and risk-based approach to address explainability, rather than a prescriptive technical approach, which allows banks to appropriately tailor and evaluate explainability based on the type of model and context in which it is being used. During the model development process, model developers strive to assess the importance of the inputs, understand the inner workings of the algorithm in question and determine how the algorithm produces outputs form these inputs. For example, model developers demonstrate the intuition for each selected feature, linking it to the problem the model is trying to solve. Additionally, banks review explainability as part of the risk assessment process for the use case of every model, which drives the depth and scope of the independent validation process. Model validators are responsible for determining whether a model or use case raises any concerns on AI explainability. The level of documentation, testing and validation required increases with the complexity of the algorithm or model. Models designated as higher risk are subject to an increased level of scrutiny by modelers, model validators, compliance teams, legal divisions and other relevant parties.

Further, we must bear in mind that humans are involved in all aspects of decisions and provide the final determinations of whether to use an AI model or the outputs of an AI model. There are very few instances where AI models will be fully autonomously making decisions. Instead, AI models are more likely to be used to support or inform decision making, as opposed to making decisions without human review and control. Within an institution, managing explainability risks entails an ongoing discussion between model developers, model validators and business units.

---

[12] Svoboda, Elizabeth, "Artificial intelligence is improving the detection of lung cancer," (Nov. 18, 2020), https://www.nature.com/articles/d41586-020-03157-9.

While banks have developed appropriate controls to manage risks related to explainability within their model risk frameworks, there are certain challenges that stand out related to AI explainability. First, the expectations for explainability of AI models are often held to a higher standard than traditional models. AI is simply another technology and is being treated as such by banks from a risk perspective. While explainability is a key aspect of fostering trust in AI outputs, we should not hold AI to higher standards for explainability than warranted based on banks' risk-based model frameworks, as doing so may prevent banks and consumers from realizing the potential benefits of AI applications.

Second, as described above, different stakeholders require different types of explanations based on context. As such, model explanations need to satisfy a broad spectrum of constituents, including model developers, validators and reviewers, internal governance, regulators and consumers. Further, explanations come with the risk of misinterpretation. If an individual does not properly understand the explanation techniques and underlying assumptions, that individual may incorrectly assess the explanation. To help mitigate risks related to interpretability, some banks have designed communication strategies or educated relevant parties. Additional detail or criteria on how explanations should differ based on the audience or end user may help clarify expectations for explainability across use cases within the financial sector. To this end, BPI supports work being conducted by the National Institute of Standards and Technology (NIST) to identify benchmarks for explainability and types of explanations expected based on context.[13]

> ***Question 2:*** *How do financial institutions use post-hoc methods to assist in evaluating conceptual soundness? How common are these methods? Are there limitations of these methods (whether to explain an AI approach's overall operation or to explain a specific prediction or categorization)? If so, please provide details on such limitations.*

As noted in response to question 1, banks have mature, risk-based processes in place to address explainability and evaluate conceptual soundness on a holistic basis, and do so at all stages of the model risk management process. Evaluation of conceptual soundness, for both AI and non-AI models, relies most on an intuitive understanding of whether the methods used by the model are appropriate for the model and business case at hand. Explainability techniques that may assist in evaluating conceptual soundness can be applied before developing the model (i.e., through exploratory data analysis), by building explainability within the model (i.e., building explainable/interpretable models), or after the model has been developed (i.e., post-hoc methods to extract explanations). Banks may rely on a combination of these techniques to evaluate conceptual soundness and mitigate the risks of opaque models or decision making.

Several post-hoc explainability methods are emerging that can be used to assess business knowledge against model mechanics, in the same manner as is done for traditional models. For example, methods such as Local Interpretable Model-agnostic Explanation (LIME), SHapley Additive exPlanations (SHAP), partial dependency plots, Anchors, counterfactual methods and others may be useful in providing simplified intuition on the relationships of complex model inputs and outputs. However, there is no single explainability technique that works for all use cases; each have benefits and limitations, and post-hoc methods for explainability are an active area of research in the data science, ML and statistics community. Banks are in the process of evaluating the performance of these tools and

---

[13] National Institute of Standards and Technology, "Four Principles of Explainable Artificial Intelligence", Draft NISTIR 8312, (August 2020), https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8312-draft.pdf and BPI's response, https://bpi.com/bpi-submits-comment-letter-to-nist-on-the-four-principles-of-explainable-artificial-intelligence/.

practicality of using post-hoc explainability methods depending on context. The Agencies should take a cautious approach in considering these post-hoc explainability methods and avoid establishing expectations, requirements or endorsements of specific tools in evaluating conceptual soundness. As post-hoc methods continue to emerge, open dialogue between the Agencies and industry would promote further understanding of the various methods and their effectiveness.

> ***Question 3:*** *For which uses of AI is lack of explainability more of a challenge? Please describe those challenges in detail. How do financial institutions account for and manage the varied challenges and risks posed by different uses?*

Reduced explainability is more of a challenge for use cases that have a specific consumer impact, where model outputs affect individuals or use personal data and explanations are crucial and required by law. At the same time, these are areas where AI provides a new perspective that may deliver important benefits, such as identifying potential fraud or providing consumers more suitable financial products. For example, AI fraud models can be instrumental in detecting and preventing suspicious payments activity, but may result in transactions being rejected or accounts frozen with limited transparency into the decision. Additionally, AI models for marketing new products may provide consumers with more targeted and beneficial financial products, but require some level of explainability to assess fairness and potential correlation with protected or socioeconomic classes.

Banks take seriously their responsibility to protect consumer information and are uniquely suited to managing customer-facing risks from lack of explainability, given decades of experience in applying sound risk management practices throughout the organization. The Model Risk Management Guidance and existing laws and regulations related specifically to the use case (i.e., credit underwriting) provide banks with sufficient tools and processes to adequately manage risks, commensurate with the use case and complexity of the model. For example, within the model risk management framework, some banks evaluate risks of models through tiering and classify risks into different categories, including accuracy risk, stability risk and potential for misuse. Depending on the model's tier, risk category and impact, the bank may require compensatory actions to control for the risk. Compensatory actions may include, for example, more frequent monitoring of the model or human review for possible intervention. If the bank determines that risks cannot be controlled, the model will be rejected. BPI recommends that the Agencies continue to apply a flexible, risk-based approach to evaluating risks related to explainability.

There are certain critical use cases within banking where aversion is high and resulting exploration and use of AI in these areas is limited. For example, for models with a customer impact, even if an institution can explain the most important drivers, if an AI model is overly complex and not easily understood by consumers or regulators, banks may choose to not implement such models. This aversion is appropriate given banks' commitment to use models in an ethical and responsible manner. BPI recommends that the Agencies work with the financial industry to identify solutions or risk mitigants to these barriers to further encourage AI applications in these areas that provide important benefits to consumers.

B. Risks from Broader or More Intensive Data Processing and Usage

> ***Question 4:*** *How do financial institutions using AI manage risks related to data quality and data processing? How, if at all, have control processes or automated data quality routines changed to address the data quality needs of AI? How does risk management for alternative data compare*

to that of traditional data? Are there any barriers or challenges that data quality and data processing pose for developing, adopting, and managing AI? If so, please provide details on those barriers or challenges.

The unprecedented proliferation and availability of data has enabled significant innovations in financial services and products that benefit consumers. As detailed in Part I of this comment letter, AI systems can sort through and analyze large volumes of data, including payment transactions, email communications, network traffic and trading data, resulting in business efficiencies and advantages for consumers. Given the importance of data in enabling such innovations, banks dedicate significant attention to ensuring that the quality of data utilized in all models and analytical tools employed by banks is accurate, complete and suitable for the context in which it is being used.

The risks of poor data quality or processing are by no means unique to AI. The use of inaccurate, incomplete, or unsuitable data may result in erroneous or biased predictions, regardless of the type of model or tool being used. However, data quality and data processing risks may be heightened for AI models due to the volume of data used. Additionally, data quality risks for externally sourced data, such as that provided by third-party vendors or utilized in vendor models, is elevated as compared to internal bank data. Banks utilize their overall risk management and control frameworks to manage risks related to data quality and data processing. The Model Risk Management Guidance appropriately details the importance of data used to develop a model and outlines steps to evaluate data quality. Examples of data quality and data processing controls that banks use include the following:

- Assessment of training data for data quality and potential biases;
- Automated testing at the data source prior to entering an algorithm to identify missing data, data errors or abnormalities;
- Upstream monitoring of the distribution of raw data inputs to identify inappropriate predictions;
- Continuous monitoring of AI models for algorithm effectiveness and accuracy; and
- Regular updates to confirm reliable data and fast processing of data.

These specific control processes are strengthened and intensified based on perceived risk of the AI model, type of data and/or the use case. As more data is used to feed into AI models, banks scale data quality routines accordingly. This is consistent with the risk-based approach banks utilize to validate all models under their model risk management frameworks. For example, banks may implement testing alongside a dynamic updating model to operate as a part of the model in order to identify abnormalities in real time. In this instance, the actual testing may not be different from a traditional model, but the application of the testing is elevated in accordance with risk. Additionally, banks may enhance continuous monitoring activities for AI models based on risk of the technical nature of the algorithm itself or use case the system is being applied to.

The type of data being used may also impact the risk and corresponding data quality and processing controls in place, regardless of whether the bank is using an AI model or traditional model. Alternative data, such as utility payment history or rental payment history, may offer important new perspectives that have the potential to improve accuracy and access to credit. Such alternative data may

also generate heightened concerns relative to traditional data. To address these concerns, banks employ substantial exploratory analysis prior to structuring and using alternative data in the model development process, including an intensified focus on monitoring data input properties. A key validation step is to assess how the data was collected, and what evidence exists that the training data properties match those in production.

Given the increase in the volume and importance of data, and the use of data to fuel AI, it is critical for banks to maintain efficient and up-to-date data systems. AI models can be particularly valuable in managing data quality risks and optimizing bank infrastructure. Specifically, AI can be used to automate testing and controls to identify and reduce risks of bad data or missing variables. Routine data quality checks within banks have become more efficient and thorough, as AI models can implement multiple control flags at once. As the Agencies assess the risks related to data quality and processing of AI, they should not lose sight of these benefits that AI can provide as the availability of data increases across the financial sector.

Question 5: *Are there specific uses of AI for which alternative data are particularly effective?*

Just as the proliferation of data generally has encouraged advancements in AI and financial services, the introduction of alternative data may provide benefits in various use cases. The RFI defines alternative data as "information not typically found in the consumer's credit files of the nationwide consumer reporting agencies or customarily provided by consumers as part of applications for credit."[14] This broad definition may include utility or rental payment history, other cash-flow transactional information from a bank account, education history, employment history and other data sources. The federal banking agencies and the CFPB jointly recognized the benefits of using alternative data in credit underwriting in their Interagency Statement on the Use of Alternative Data in Credit Underwriting, issued in December 2019. They found that the use of alternative data may improve the speed and accuracy of credit decisions, help firms evaluate the creditworthiness of consumers who may not be able to obtain credit in the mainstream credit system, and enable consumers to obtain additional products or more favorable pricing or terms based on enhanced assessments of repayment capacity.[15]

Alternative data may be helpful in a variety of use cases in addition to credit underwriting, such as in identifying fraud or evaluating customer complaints. For example, new sources or patterns of deposit account and other transaction data may be useful for improving fraud detection. Alternative data that is orthogonal to currently utilized data has the most potential, as it may yield new insights and relationships to the outcome because it represents a different dimension.

Banks are still in the process of evaluating the utility of alternative data across various use cases. There are many unknowns about how alternative data impacts specific outputs and whether alternative data inadvertently introduces biases or other unfair outcomes, which is why banks conduct significant analysis of the alternative data before use, as noted in response to question 4 above. The Federal Trade Commission (FTC) emphasized the importance of not exaggerating what an algorithm can do or whether

---

[14] Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning. 86 Fed. Reg. 16837 (March 31, 2021).

[15] FRB, CFPB, FDIC, NCUA, and OCC, Interagency Statement on the Use of Alternative Data in Credit Underwriting (Dec. 3, 2019) https://www.consumerfinance.gov/documents/8242/cfpb_interagency-statement_alternative-data.pdf.

it can deliver fair or unbiased results in an April 2021 blog post.[16] Banks are carefully analyzing the effectiveness of alternative data to ensure it is used responsibly and in accordance with relevant laws and regulations.

To that end, BPI provides the following recommendations to encourage the use of alternative data while ensuring consumers receive effective protections on their data. First, the Agencies should consider ways in which banks can explore the use of alternative data without facing fair lending and other consumer compliance regulatory consequences, such as through the use of safe harbors or pilot programs. Second, the Agencies should continue to assess and provide clarity into the types of alternative data suitable across various use cases, building upon the broad definition provided in the RFI. Third, the Agencies should ensure nonbank lenders using alternative data are held to the same rigorous standards applied to banks, by subjecting all lenders using alternative data in credit underwriting to model risk management requirements and by providing consistent examination across lenders to detect misuses of alternative data. These steps would help ensure alternative data and models using alternative data are appropriate, consistent and in the best interest of consumers.

C. Overfitting

Question 6: *How do financial institutions manage AI risks relating to overfitting? What barriers or challenges, if any, does overfitting pose for developing, adopting, and managing AI? How do financial institutions develop their AI so that it will adapt to new and potentially different populations (outside of the test and training data)?*

Overfitting is not unique to AI or ML and is appropriately managed as a part of well-established model risk management procedures. The data-driven nature of ML may elevate the inherent risk of overfitting. Indeed, this is likely commensurate with the potential presented by ML. However, banks are not simply accepting an elevated risk of overfitting as the cost of ML's potential. Banks instead have enhanced their development testing, validation performance monitoring and other controls under their model risk management frameworks to render a residual risk of overfitting that is muted relative to the potential.

Overfitting often arises when the complexity of the model exceeds the complexity of phenomena the model is supposed to tackle. Banks' risk-based model frameworks discourage the development of models where the complexity cannot be justified. During the model development process, banks may control for overfitting by using a newly built model to make predictions on out-of-sample and out-of-time data to see how the new model performs on the data the model has never seen. Specific practices for overfitting that banks may use include: (1) benchmarking or simulated scenario analysis to test performance against specific variables; (2) sensitivity analysis to identify changes in variables; (3) cross-validation to split data into model calibration and model evaluation segments; (4) hyperparameters, such as learning rates and tree depths; (5) robustness testing; and (6) examination of bias-variance plots. Additionally, to address risks of overfitting the model on tenuously or spuriously correlated variables, banks can conduct tests to ensure appropriate feature selection when building the model.

---

[16] Elisa Jillson, "Aiming for truth, fairness, and equity in your company's use of AI" (Apr. 19, 2021), Federal Trade Commission, https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai.

To supplement these robust model development and validation controls, banks enhance continuous monitoring activities to manage risks related to overfitting. For example, banks may define and apply metrics during the model development phase to track performance for each model, which are continuously monitored after models are put into production. Risks from overfitting will always be present; however, the rigorous methods banks employ to mitigate these risks as a part of their existing model risk management processes are appropriate.

D. Cybersecurity Risk

Question 7: *Have financial institutions identified particular cybersecurity risks or experienced such incidents with respect to AI? If so, what practices are financial institutions using to manage cybersecurity risks related to AI? Please describe any barriers or challenges to the use of AI associated with cybersecurity risks. Are there specific information security or cybersecurity controls that can be applied to AI?*

Cybersecurity risks posed by AI are not currently unique or pronounced, and should not be treated differently than cybersecurity risks posed by any other forms of technology. Banks implement and maintain strong information security programs designed to protect the bank and its clients, meet regulatory requirements and adjust to the risks presented by an evolving threat landscape. The existing laws and regulatory guidance provide an appropriate framework for banks to implement controls to mitigate cybersecurity risks, including any risks related to the usage of AI. Specifically, the federal banking agencies' Interagency Guidance Establishing Security Standards ("Interagency Guidance"), implementing guidance of the Gramm-Leach-Bliley-Act (GLBA), requires banks to develop, implement and maintain an information security program to identify and control risks to customer information and systems.[17] Guidance to the industry from the Federal Financial institutions Examination Council (FFIEC) and NIST provide additional direction to help financial institutions establish effective security measures and address cybersecurity risks.[18] Further, industry efforts such as the Financial Services Sector Cybersecurity Profile[19] and tools such as the Microsoft/MITRE Adversarial ML Threat Matrix[20] demonstrate the proactive leadership shown by the financial sector and private sector generally to enhance cybersecurity and resiliency through standardization.

---

17 12 C.F.R. pt. 30, App. B.

18 See FFIEC IT Examination Handbooks, Information Security, available at https://ithandbook.ffiec.gov/it-booklets/information-security.aspx. See also NIST Cybersecurity Framework, available at https://www.nist.gov/cyberframework.

19 The Financial Services Sector Cybersecurity Profile was developed under the leadership of the Financial Services Sector Coordinating Council and is now maintained through an open, consensus-driven process run by the Cyber Risk Institute. The Profile serves and is supported by every level of the financial sector to reduce an organization's cyber compliance burden, allowing them to return resources to frontline defense. For more information, see https://cyberriskinstitute.org/the-profile/.

20 The Adversarial ML Threat Matrix is an industry-focused open framework that empowers security analysts to detect, respond to and remediate threats against ML systems. Created by Microsoft and MITRE, in collaboration with 11 other organizations, the Adversarial ML Threat Matrix is intended to bolster monitoring strategies around organizations' mission critical ML systems. For more information, see https://github.com/mitre/advmlthreatmatrix.

Under this comprehensive framework, banks develop strong disaster recovery and continuity plans, business service resiliency plans and network design patterns, which can be applied to cover AI and used to mitigate risks of adversarial attacks using AI. Specific IT security controls that may be effective in mitigating AI cybersecurity risks include but are not limited to: implementation of strong data quality controls to identify suspicious data; enhanced model performance monitoring to identify anomalous performance; encryption of data at rest; and open-source software controls. Additionally, banks ensure that controls are coordinated across various business lines involved, including the model risk management teams, cybersecurity teams and others.

The primary risks banks face are the continued threats of integrity and availability specific to the data sources used by their AI models and the potential for adversarial attacks to penetrate and overwhelm bank defenses. Hackers and fraudsters are using more sophisticated methods, including the use of AI in some cases, to bypass detection systems and gain access to financial and personal information. While not specific to AI models, banks are aware of and monitor the following groups of threats that adversaries may use to exploit security weaknesses:

- Data poisoning: Adversaries may contaminate the data used for training models to impair the overall solution performance, negatively affecting its learning processes or outputs, and potentially resulting in bad behavior or inserted backdoors.

- Data privacy attacks: Adversaries may be able to retrieve sensitive/confidential information from the model, potentially compromising the privacy of the data. Data privacy attacks may occur through model inference, where the adversary infers information from training data by querying the models, or through model inversion, where the adversary extracts training data from the model directly.

- Evasion attacks: Adversaries may manipulate inputs or introduce perturbed inputs that appear normal but cause the model to misclassify the output.

- Model extraction: Adversaries may attempt to steal the model itself.

The controls described above are currently adequate to manage these risks, and banks devote consistent and considerable resources towards continual improvements and vigilance that have helped keep institutions, their customers and the broader economy safe. As AI continues to be leveraged by cyber threat actors in cyberattacks, a broader set of defenses across the ecosystem may be necessary. The Agencies should continue to collaborate with industry partners to identify where AI poses specific risks related to cybersecurity. For example, BPI and its members appreciated the taxonomy and terminology of adversarial machine learning published by NIST, which provided an understanding of the key types of attacks, defenses and consequences from adversarial machine learning.[21] Continued discussions and enhanced partnership between industry and government stakeholders will better prepare the financial sector and decrease the likelihood of AI cybersecurity attacks in the future. Additionally, as noted in Part I of this comment letter, banks are increasingly adopting AI in the field of cybersecurity and beginning to realize the potential benefit of AI to prevent cyber threats and protect consumers. Given the resources entailed to ensure banks have effective cybersecurity programs, the use of AI may result in cost savings, reduced time to identify specific incidents or threats, and other

[21] National Institute of Standards and Technology, A Taxonomy and Terminology of Adversarial Machine Learning, Draft NISTIR 8269, (Oct. 2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8269-draft.pdf.

efficiencies. The Agencies should consider ways to encourage the use of AI in cybersecurity, given that cyber threats are unceasing and the potential for a breach that impacts the industry is always present.

E. Dynamic Updating

Question 8: *How do financial institutions manage AI risks relating to dynamic updating? Describe any barriers or challenges that may impede the use of AI that involve dynamic updating. How do financial institutions gain an understanding of whether AI approaches producing different outputs over time based on the same inputs are operating as intended?*

As indicated in the RFI, some AI approaches have the capacity to update on their own, sometimes without human interaction, often known as dynamic updating. It is worth noting that banks' implementation of dynamic updating models is relatively limited, as there are few applications that would benefit from having models automatically retrain and dynamic updating models are typically only applicable to environments that are extremely dynamic. Moreover, the use case and context in which dynamic updating models are applied is highly relevant. Banks recognize the risks related to models evolving over time as they learn from new data, potentially resulting in model drift. Consistent with banks' risk-based approaches to model deployment, banks primarily implement dynamic updating models for lower risk use cases, such as marketing or customer interaction, and do not implement dynamic updating models for high-risk use cases where explainability and understanding of inputs and outputs are required, such as credit origination.

The Model Risk Management Guidance provides banks with the flexibility to scale and adapt techniques to manage risks from dynamic updating models, tailored to the distinct attributes of the model and context in which it is being used. Indeed, these enhanced risk management practices are comparable to how banks scale techniques for other risks highlighted in the RFI, including those from explainability, overfitting, and data processing. Despite the currently limited use of dynamic updating models in banking, the below examples highlight specific steps that banks may take to manage risks from dynamic updating:

- Banks may implement standards requiring simulation results to determine the expected drift of model parameters, hyperparameters and outputs in order to determine the monitoring requirements and thresholds at which to trigger an alert to review the model.

- Banks may implement data collection requirements to ensure training data continues to be appropriately diverse and sufficiently representative of the population on which model is applied.

- Banks may implement internal standards requiring model development teams to specify the planned re-training schedule in advance, which may include changes in techniques, data sources, time horizons, hyperparameters, and reselection of features. These specific plans are documented and reviewed during the model validation process.

- Banks may utilize a risk control matrix to track, monitor, regulate and take action against model changes at any stage in the process. The risk control matrix may ensure that any changes to the model are directly linked to changes in the data (i.e., population, variable relationships, new events, etc.).

- Banks employ periodic testing of dynamic updating models to determine whether unreported model updates have occurred. Bank also utilize testing tools, such as Champion Challenger, to compute the effectiveness of one or more data attributes used in a model.

- Banks may require dynamic updating models to have mechanisms in place to roll back to previously approved states.

- Banks perform ongoing monitoring of performance to identify and control for model drift. Ongoing performance monitoring ensures that AI models are producing outputs as expected within specific parameters.

Banks recognize that models with dynamic or real-time updating features may introduce additional challenges and complexity, and thus may require changes to model governance processes. Many of the challenges surrounding dynamic updating models are associated with demanding internal standards to manage associated risks. First, internal standards may require modeling teams to maintain a log of all changes to the model, which can be burdensome given the dynamic nature of updates. Second, it may be challenging to establish internal expectations for "materiality" of changes to models that require additional validation. Third, operations teams require solid software engineering frameworks to ensure automated processes, testing and explainability checks, among other methods, are working properly and not introducing incremental risks of their own. Finally, frequent updates to dynamic updating models may not be meaningful, and may miss relationships that evolve over time. Certain banks may have the resources and subject matter expertise to manage these challenges related to dynamic updating models, and are thus more comfortable with deploying them based on risk.

There are certain contexts where dynamic updating models have proven to be extremely useful. With the ability to update on their own, sometimes without human interaction, dynamic updating models may be useful in addressing challenges in environments that are time-sensitive and truly dynamic. For example, we are aware of banks using dynamic updating models to enhance customer experience by using real-time feedback on click data to provide customers with appropriate website pages based on the goal or application. Banks have also used dynamic updating models when launching a new digital product, where the model is able to quickly capture customer opinions and learn from such opinions to anticipate how a customer may respond in the future.

BPI urges the Agencies to continue to allow banks the flexibility to tailor risk management practices based on the risk and use case of the model, particularly as it relates to the dynamic and constantly evolving aspects of AI models. Banks are only scratching the surface of the potential that dynamic updating models may provide institutions and consumers, and are exploring this potential in a responsible, risk-based manner, as evidenced by the currently limited use of dynamic updating models in financial services. The Agencies should also consider ways to encourage the use of dynamic updating models, given that dynamic updating models may lead to advances that open the door to tackling entirely new problems that previously may have seemed unreachable.

F. Oversight of Third Parties

Question 10: *Please describe any particular challenges or impediments financial institutions face in using AI developed or provided by third parties and a description of how financial institutions*

manage the associated risks. Please provide detail on any challenges or impediments. How do those challenges or impediments vary by financial institution size and complexity?

Over the last several years, the financial services industry has experienced a rapid emergence of third-party vendors providing AI products. Banks use third-party vendor applications and/or data to scale specific processes and optimize operations where they may not have the resources or skillset available to do so internally. Both the types of AI vendor products and application of these products vary across banks. For example, banks may utilize vendors that provide simple AI models for lower risk use cases, such as portfolio analysis, or more complex AI models for higher risk use cases, such as fraud and AML monitoring or cybersecurity. The risks associated with using AI provided by third parties are primarily associated with the proprietary nature of the vendor product and ability to obtain sufficient information to understand and measure the risk associated with using this type of product or service. These risks may include:

- The level of maturity in the development, testing, transparency and governance of vendor-owned AI models varies.
- Banks may not be exposed to underlying algorithms or source codes, making it difficult to investigate "under the hood," and resulting in limited transparency.
- Vendors may only provide the model-based output, making it difficult to determine the reach of model governance.
- Vendors may not provide adequate monitoring data for model risk management teams to evaluate.
- Banks may have a lack of visibility into fourth party models (i.e., externally developed models or vendors used within a vendor's product), limiting knowledge of systemic risks that may be introduced from the combination of techniques.

However, as with all types of risks banks face, banks have developed and continue to develop ways to manage these risks through various methods. Generally, banks work closely with potential vendors to ensure that they have enough information and documentation to develop an internal comfort with the model. While some vendors resist explaining how their AI works, there are now many AI vendors, and banks thus have the benefit of choice and look to partner with those that are more transparent and willing to share information that the bank needs to evaluate the third-party model's benefits and potential risks. Banks also may consider including contractual requirements regarding the AI models' testing, methodology, explainability of the results generated by the system, and/or intellectual property rights which may be derived from the use of the system. Overall, banks make a conscious effort to decide whether to use a vendor based on their willingness to cooperate with the banks' risk control measures, transparency of model methodologies and underlying risk profile of the vendor.

In addition to these overarching methods to manage risks from vendors providing AI, banks typically conduct due diligence and perform model validation activities over vendor models in a similar manner as any model developed in-house. Specifically, prior to purchasing a vendor model, banks conduct due diligence on key data, methodologies and performance of the AI model. Banks' validation activities may focus on example-based testing with a review of the evidence provided on the

performance and maintenance of the vendor model. Banks may also conduct fair lending and consumer compliance assessments of vendor models, and may require vendors to sign a fair lending affirmation representing that their models do not contain prohibited bases or proxies and have been tested to confirm compliance with fair lending laws. Further, banks perform ongoing monitoring of vendor model performance, and implement contingency plans to address potential failovers.

These activities to manage potential risks posed by vendor AI models are consistent with both the Third-Party Risk Management Guidance[22] and the Model Risk Management Guidance. As noted in the RFI, existing guidance on third-party risk management describes information and risks that may be relevant to financial institutions when selecting third-party vendors for AI. The OCC clarified how bank management should address third-party risk management when using a third-party model in a set of frequently asked questions (FAQ) issued in March 2020. The FAQs specify that "third-party models should be incorporated into the bank's third-party risk management and model risk management processes" and "bank management should conduct appropriate due diligence on the third-party relationship and on the model itself."[23]

The Third-Party Risk Management Guidance and Model Risk Management Guidance are generally principles-based and flexible enough to cover the risks related to using vendor-provided AI products. Certain types of vendor-provided AI products and/or the use cases in which vendor products are applied may raise greater risks than others, and banks thus manage those risks differently. For example, an AI product provided by a third-party travel agency for a bank's travel needs may not be subject to the same level of risk assessment and model validation as a vendor's AI product used in fraud detection or credit underwriting. The Model Risk Management Guidance provides for such flexibility, noting "the rigor and sophistication of validation should be commensurate with the bank's overall use of the models, the complexity and materiality of its models, and the size and complexity of the bank's operations."[24]

However, one of the key challenges that banks face in managing risks from AI provided by third parties and validating vendor AI models is the general perception by regulators and examiners that models labeled as "AI" or "ML" entail higher risk. Banks apply a risk-based approach in evaluating risks from third parties, including their AI models, depending on the context in which a vendor's AI model is deployed, consistent with the Model Risk Management Guidance and general bank risk-management practices. The Agencies' expectations of banks' approaches to managing vendor risk with respect to AI should also reflect this risk-based approach. Specifically, the Agencies should ensure that the context in which a vendor's AI model is deployed is considered by regulators and examiners when evaluating banks' risk management practices in this regard. Clearer expectations with respect to required due diligence on different types of models provided by vendors would benefit both banks and their third parties. Further, it may be useful for the Agencies, in collaboration with banks, to identify certain types of commoditized vendor-provided AI products used in low-risk applications that may benefit from lower

---

[22] FDIC: Guidance for Managing Third-Party Risk (FIL)-44-2008, https://www.fdic.gov/news/financial-institution-letters/2008/fil08044.html; OCC Bulletin 2013-29, OCC Bulletin 2020-10; NCUA: Evaluating Third Party Relationships, Supervisory Letter (SL) 07-01 (Oct. 2007); and FRB: Guidance on Outsourcing Risk (SR 13-19), https://www.federalreserve.gov/supervisionreg/srletters/srletters.htm.

[23] Third-Party Relationships: Frequently Asked Questions to Supplement OCC Bulletin 2013-29, (OCC Bulletin 2020-10), https://www.occ.gov/news-issuances/bulletins/2020/bulletin-2020-10.html.

[24] FRB, SR 11-7, Supervisory Guidance on Model Risk Management, page 9.

model validation expectations and/or a potentially standardized review. These tools are so widely used and commonplace today that, besides not being able to access to the model mechanics, there is little to no value added by each bank putting these models through similarly rigorous model risk management processes. Examples of lower-risk and mature vendor AI solutions that would benefit from such an approach include:

- Optical character recognition (OCR) for standardized scanning of forms;
- Fingerprint verification for mobile phone logins;
- Machine translation for text; and
- Text-to-speech transcription if used on websites for accessibility purposes (i.e., reading out text).

While banks recognize the importance of model validation and due diligence, a standardized approach commensurate with the low-risk nature of the product and application would benefit both banks and their third parties. This list could be updated on a continuous basis as additional low-risk applications are identified by banks and regulators.

G. Fair Lending

BPI appreciates the guidance and other clarifications the Agencies individually or collectively have issued in the past few years to provide industry with guidance on using AI credit underwriting systems and alternative data.[25] These releases, along with the demonstrated commitment of the Agencies to pursue policies to promote financial innovation, including AI innovation, have provided real value to industry and consumers.[26] The RFI poses a series of fair lending questions (Questions 11-15) relating to the use of AI in credit underwriting. Before addressing the specific questions, BPI has certain general observations to frame this important topic of AI and fair lending.

---

[25] *See* Patrice Alexander Ficklin, Tom Pahl, Paul Watkins, "Innovation Spotlight: Providing adverse action notices when using AI/ML models" (July 7, 2020), https://www.consumerfinance.gov/about-us/blog/innovation-spotlight-providing-adverse-action-notices-when-using-ai-ml-models/; *see also* Interagency Statement on the Use of Alternative Data in Credit Underwriting (Dec. 3, 2019); Governor Lael Brainard, "Supporting Responsible Use of AI and Equitable Outcomes in Financial Services," speech at the AI Academic Symposium hosted by the Board of Governors of the Federal Reserve System, Washington, D.C. (Virtual Event) (Jan. 12, 2021), Speech by Governor Brainard on supporting responsible use of AI and equitable outcomes in financial services - Federal Reserve Board.

[26] The CFPB's No-Action Letters to Upstart represent a good example of using innovation policies to allow controlled experiments that foster greater understanding of the use of AI. CFPB, Letter from Edward Blatnick, Acting Assistant Director, Office of Innovation to Alison Nichol, General Counsel, Upstart Network, Inc. (Nov. 30, 2020), *available at* https://files.consumerfinance.gov/f/documents/cfpb_upstart-network-inc_no-action-letter_2020-11.pdf; CFPB, Letter from Christopher M. D'Angelo, Associate Director for Supervision, Enforcement, and Fair Lending to Thomas P. Brown, Paul Hastings, LLP (Sept. 14, 2017), *available at* https://files.consumerfinance.gov/f/documents/201709_cfpb_upstart-no-action-letter.pdf.

First, the use of AI credit underwriting systems is subject to the fair lending laws and regulations to the same extent as the use of conventional underwriting systems.[27] The Equal Credit Opportunity Act (ECOA) along with its implementing regulation, Regulation B, is the primary federal law prohibiting discrimination in credit transactions.[28] ECOA and Regulation B prohibit creditors from discriminating against an applicant in any aspect of a credit transaction on a prohibited basis, including race, gender, national origin and age, among certain other prohibited bases.[29] ECOA and Regulation B also require a creditor to provide an adverse action notice to an applicant when a creditor denies an application for credit or takes other adverse action against an applicant.[30] The Fair Housing Act (FHA) prohibits discrimination in any aspect of a residential real estate transaction, including credit, and enumerates disability and familial status as additional prohibited bases.[31] Existing fair lending laws and regulations contain flexible requirements that enable creditors to use AI for credit decisioning through risk-based business determinations, and such flexibility should be preserved. These existing regulations and guidance were mostly written in an era of conventional and judgmental underwriting systems. Therefore, it is appropriate for the Agencies to seek public input through this RFI on whether the development of AI necessitates any regulatory updates or innovations. BPI provided initial recommendations for regulatory modernization to foster the responsible use of AI and alternative data in credit underwriting in its white paper, *Artificial Intelligence: Recommendations for Principled Modernization of the Regulatory Framework*.[32]

Second, innovation in credit underwriting, specifically the use of AI credit underwriting and reducing reliance on human judgment, may promote fair lending and reduce the potential for discrimination and bias in credit decisions.[33] Further, BPI member experience indicates that federal regulators generally prefer for creditors to make credit decisions using empirical, automated underwriting systems, rather than judgmental systems that potentially could introduce human biases. BPI urges the Agencies to consider the opportunities and benefits offered by AI credit underwriting

---

27 The same safety and soundness considerations also apply to all forms of credit underwriting, including AI and conventional underwriting systems. AI systems are designed to improve the accuracy of underwriting decisions, and so may provide a net benefit to a bank's safety and soundness.

28 15 U.S.C. § 1691 *et seq.*; 12 C.F.R. pt. 1002.

29 15 U.S.C. § 1691(a); 12 C.F.R. §§ 1002.2(z) and .4(a). The Fair Housing Act also prohibits discrimination in the sale or rental of housing on the basis of certain prohibited characteristics similar to the ECOA prohibited bases. *See* 42 U.S.C. §§ 3601-3619.

30 12 C.F.R. § 1002.9(a), (b).

31 42 U.S.C. § 3601 *et seq.*

32 Artificial Intelligence: Recommendations for Principled Modernization of the Regulatory Framework, Bank Policy Institute and Covington & Burling LLP (Sept. 14, 2020), https://bpi.com/wp-content/uploads/2020/10/Artificial-Intelligence-Recommendations-for-Principled-Modernization.pdf.

33 CFPB Examination Manual, ECOA 6 (Oct. 2015), https://files.consumerfinance.gov/f/documents/201510_cfpb_ecoa-narrative-and-procedures.pdf (comparing the use of "judgmental systems that rely on a credit officer's subjective evaluation of an applicant's creditworthiness" with "more-objective, statistically developed techniques such as credit scoring.").

systems and encourage flexible, risk-based approaches that allow the use of AI credit underwriting systems while appropriately managing fair lending and related risks.[34] In applying this flexible, risk-based approach, the Agencies should clarify that banks are expected to exercise their judgment – without second guessing by examiners – when navigating the evolving complexities and risks of using such systems in a responsible manner consistent with fair lending laws and the Model Risk Management Guidance. BPI further notes that regulators generally have welcomed the introduction of AI models in other contexts, for example, as automated tools to facilitate BSA/AML monitoring,[35] and should similarly encourage the use of AI models in credit underwriting for the benefits those models offer.

Third, BPI member banks have been carefully assessing the opportunities provided by the use of alternative data in credit underwriting against the potential fair lending risks. Although a broad spectrum of alternative data may be available, our members have been diligently evaluating which types of alternative data are appropriate for use in AI credit underwriting models. For example, cash flow and bill payment data provide insight into a consumer's overall financial health and profile and generally are considered appropriate and unbiased data points to consider in credit underwriting, and the Agencies have previously acknowledged the beneficial uses of cash flow data.[36] In contrast, banks are less likely to include other data points in AI credit underwriting systems, especially if the data point lacks a clear nexus to a consumer's financial well-being or may inadvertently introduce fair lending risk into the underwriting process. Our members also are evaluating how AI credit underwriting models may perform during an economic downturn. BPI welcomes the opportunity to engage with the Agencies to identify practical solutions to enable banks to use alternative data in credit underwriting for the benefit of consumers consistent with fair lending laws.

Finally, BPI believes that relevant laws, regulations and guidance are not applied equally to bank and nonbank creditors alike, which results in an un-level playing field for banks and nonbanks using AI credit underwriting systems and less robust consumer protection for customers of nonbanks. The Agencies should coordinate to apply consistent standards for model risk management and oversight of AI across banks and nonbanks for model risk management and fair lending purposes.

Question 11: *What techniques are available to facilitate or evaluate the compliance of AI-based credit determination approaches with fair lending laws or mitigate risks of non-compliance?*

---

[34] *See generally* Prepared Remarks of CFPB Director Richard Cordray at the Alternative Data Field Hearing, Charleston, West Virginia (Feb. 16, 2017), *available at* https://www.consumerfinance.gov/about-us/newsroom/prepared-remarks-cfpb-director-richard-cordray-alternative-data-field-hearing/?_gl=1*1m7qhns*_ga*NzY1MjExNDAyLjE0OTI1MTkxNDU.*_ga_DBYJL30CHS*MTYyMjU1MjIxNi4xNy4xLjE2MjI1NTM2MTYuMA ("[I]f fair lending concerns cast a large enough shadow, they prevent people from considering and using alternative data that might open up more credit for minority and underserved consumers. This could interfere with progress for the very people these laws are intended to protect.").

[35] Joint Statement on Innovative Efforts to Combat Money Laundering and Terrorist Financing (Dec. 3, 2018), https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20181203a1.pdf.

[36] Interagency Statement on the Use of Alternative Data in Credit Underwriting at 2 (Dec. 3, 2019) ("[T]he agencies are aware that the use of certain alternative data [specifically, cash flow data] may present no greater risks than data traditionally used in the credit evaluation process."), https://files.consumerfinance.gov/f/documents/cfpb_interagency-statement_alternative-data.pdf.

Please explain these techniques and their objectives, limitations of those techniques, and how those techniques relate to fair lending legal requirements.

Banks generally follow the same or similar approaches for evaluating the compliance of AI credit underwriting systems with fair lending laws that they use with conventional automated underwriting systems for such purposes. Building on a fair lending risk assessment, banks may evaluate and mitigate fair lending risk through: (1) pre-implementation or front-end model development, input and design choices; (2) pre-implementation fair lending testing; (3) ongoing monitoring; and (4) periodic back-testing of model outcomes and trend analyses. Banks choose and shape their fair lending risk management approaches to fit the relevant characteristics of the institution, the risks to the institution and the attributes of the credit underwriting system. For example, banks review variables to ensure that models do not consider prohibited bases or close proxies for prohibited bases to mitigate disparate treatment risk. Banks also consider whether each variable has a close nexus to creditworthiness and, if not, whether the variable might result in additional fair lending risk. Banks also conduct statistical testing of model outcomes to assess whether facially neutral models pose disparate impact risk and whether model changes would produce less impact on a protected class without undermining model performance. Although there is no regulatory or industry-standard best practice for statistical testing, such testing may involve various techniques and metrics to calculate potential disparities, such as: (i) adverse impact ratios; (ii) standardized mean differences; (iii) marginal effects measures; and (iv) odds ratios.[37] Fair lending testing methodologies used by industry rely on various assumptions or proxies for protected classes, such as Bayesian Improved Surname Geocoding or BISG method, and typically focus just on one or two protected classes.

Because AI credit underwriting systems are more complex than conventional underwriting systems, substantially more time, attention and effort may be appropriate to apply existing techniques to AI models. For example, at the pre-implementation phase, the inclusion of more variables and the evaluation of relationships between different variables may require additional work to identify and exclude potential proxies for discrimination and the use of more complex statistical analyses. In addition, banks are modifying or adopting existing techniques to address new challenges. For example, some banks and technology vendors are using various post-hoc explanation methods to identify ECOA adverse action reason codes and explain model outcomes. Further, banks may use a variety of approaches to prevent AI models from generating outcomes inconsistent with fair lending requirements. Implementation of AI models is usually undertaken in conjunction with extensive human training, decision making, validation and/or testing.

The fair lending laws and regulations and Model Risk Management Guidance provide banks with flexibility to implement and adapt various approaches for evaluating models to facilitate compliance with fair lending laws in a risk-based manner, tailored to the distinct attributes of their use of AI credit underwriting models. The Agencies should preserve this flexibility. In this regard, BPI encourages the CFPB to reiterate in more formal guidance the staff observations about the flexibility of ECOA and FCRA

---

[37] For descriptions of these techniques and metrics, *see* CFPB Supervisory Highlights, Issue 9, at 28-30 (Fall 2015), and Navdeep Gill, Patrick Hall, Kim Montgomery, and Nicholas Schmidt, "A Responsible Machine Learning Workflow with Focus on Interpretable Models, Post-hoc Explanation, and Discrimination Testing," at 5 (2020), https://www.bldsllc.com/publications/20200229_A_Responsible_Machine_Learning_Workflow.pdf.

regarding explainability and identifying the principal reasons for adverse action set forth in the July 2020 blog post published by CFPB staff.[38]

At the same time, the Agencies should recognize that whenever a model is subject to fair lending review – whether the model is an AI model or a traditional model – fair lending professionals are required to exercise risk-based judgment on difficult practical issues for which there is no regulatory guidance. Some of the questions fair lending officers regularly face include making tradeoffs between reduced model performance and reduced disparities, balancing the acceptable level of reduced performance against incremental reductions in disparities, and selecting among competing models that may have different disparities for different protected classes. In addition, some methodologies that could help reduce bias are not used by banks because they would require consideration of prohibited bases. Given these challenges, the Agencies should apply the same flexible standards to both AI models and traditional models, and should not second-guess the judgment of fair lending professionals in the absence of published guidance on these types of questions.

Accordingly, the Agencies should set clear expectations for examiners to apply a flexible, risk-based approach when evaluating how banks, informed by fair lending risk assessments, evaluate fair lending and model risk management compliance for AI credit underwriting models. These expectations should clarify that: (a) the fair lending and model risk management standards for reviewing AI models and conventional models are the same, and there is not a more rigorous standard for AI models; (2) fair lending model review requires the exercise of risk-based judgment by lenders based on the facts specific to the model and its alternatives; (3) the risk-based judgments made by fair lending officers should not be second-guessed in the absence of published regulatory guidance; and (4) the distinctive features of AI credit underwriting systems may be reasonably expected to result in some modification or adaptation of certain model development and testing practices. The Agencies should consider explaining these points in examination manuals and examiner training materials to ensure that examiners, in fact, provide flexibility to banks developing or using AI credit underwriting models.

> ***Question 12:*** *What are the risks that AI can be biased and/or result in discrimination on prohibited bases? Are there effective ways to reduce risk of discrimination, whether during development, validation, revision, and/or use? What are some of the barriers to or limitations of those methods?*

As noted previously, AI credit underwriting systems may create opportunities to prevent prohibited basis discrimination and reduce or eliminate bias in credit decisions. These opportunities stem from: (1) the automation of credit decision making and reduced reliance on human judgment; (2) the consideration of alternative data, such as cash flow and bill payment data, that may help recent immigrants, younger consumers and other consumers qualify for credit who may otherwise be deemed credit invisible by conventional underwriting models; and (3) the use of broader data sets to promote more accurate credit decisions.

At the same time, there also is a risk that the use of AI credit underwriting models could be susceptible to biases in the underlying data – a problem that is not unique to AI, but that exists for both conventional and AI credit underwriting models – or in the rules applied to that data, potentially resulting in discrimination on a prohibited basis. First, the data sets used to train AI algorithms may not

---

[38] Patrice Alexander Ficklin, Tom Pahl, Paul Watkins, "Innovation Spotlight: Providing adverse action notices when using AI/ML models" (July 7, 2020), https://www.consumerfinance.gov/about-us/blog/innovation-spotlight-providing-adverse-action-notices-when-using-ai-ml-models/.

be accurate, reliable or representative of the population as a whole, may not be tested for the anticipated use case or may contain other biases.[39] Biases may be embedded in underlying data generating processes reflecting historical or social inequities, created inadvertently during data collection and sampling, or introduced by combining AI outputs with other data for decision making purposes. In addition, data preparation and feature engineering decisions may inadvertently prompt AI algorithms to reinforce or amplify biased data patterns. Banks, however, understand these risks and the importance of building models using unbiased data sets, and have long experience testing data and building models that comport with fair lending compliance obligations.

Second, the use of alternative data and the assessment of relationships between different data points could introduce unrecognized proxies for prohibited bases. For example, social media and marketing data may be correlated with socioeconomic class, cultural or group identification. By contrast, cash flow and bill payment data pose less risk of introducing discriminatory bias. In addition, the flexible model structure of AI could combine seemingly unrelated and innocuous data in unintended, non-linear ways that lead to biased outputs. Based on their experience developing models that satisfy fair lending compliance obligations, banks understand the importance of critically examining specific data points, excluding data from underwriting models that potentially could result in discriminatory outcomes, and implementing guardrails and controls established by human operators to prevent discriminatory outcomes.

Third, the rules applied to the data may result in discriminatory outcomes. The complexity of AI credit underwriting decision-making calls for commensurately sophisticated fair lending strategies. The extrapolation of AI rules to data that is difficult to use for predictive purposes may introduce unintentional bias. Banks are applying their extensive experience developing and testing rules for conventional underwriting systems to the new challenges of AI models to ensure that they monitor for, identify and prevent discriminatory outcomes.

There are many effective ways to reduce the risk of discrimination when using AI credit underwriting systems. Depending on the context, specific steps undertaken to mitigate the risk of banks or any other creditors using AI credit underwriting systems in a discriminatory manner may include:

---

[39] *See, e.g.,* Governor Lael Brainard, "Supporting Responsible Use of AI and Equitable Outcomes in Financial Services," speech at the AI Academic Symposium hosted by the Board of Governors of the Federal Reserve System, Washington, D.C. (Virtual Event) (Jan. 12, 2021), Speech by Governor Brainard on supporting responsible use of AI and equitable outcomes in financial services - Federal Reserve Board ("Unfortunately, we have seen the potential for AI models to operate in unanticipated ways and reflect or amplify bias in society. . . Thus, it is critical to be vigilant for the racial and other biases that may be embedded in data sources. It is also possible for the complex data interactions that are emblematic of AI—a key strength when properly managed—to create proxies for race or other protected characteristics, leading to biased algorithms that discriminate."); See also Carol Evans, Associate Director, Division of Consumer and Community Affairs, the Board of Governors of the Federal Reserve System, "Keeping Fintech Fair: Thinking About Fair Lending and UDAP Risks," Consumer Compliance Outlook (Second Issue 2017), https://www.consumercomplianceoutlook.org/2017/second-issue/keeping-fintech-fair-thinking-about-fair-lending-and-udap-risks/ ("[I]t is important to consider whether the data are accurate, reliable, and representative of a broad range of consumers. . . [I]t is important to ask if the data have been validated and tested for the specific uses [because f]air lending risk can arise in many aspects of a credit transaction.").

- Filtering data sets so that AI credit underwriting systems do not consider prohibited bases, known proxies for discrimination, or data that could be a proxy for discrimination, alone or in combination with other data.

- Using representative data sets to build and test AI credit underwriting systems, consistent with expected uses of the systems.

- Including fair lending considerations in front-end testing of systems by reviewing each variable and, if appropriate, the overall system for any prohibited bases or proxies for prohibited bases.

- Programming AI credit underwriting systems and/or restricting variables to prevent systems from considering prohibited bases or proxies for discrimination, such as narrow geographic areas or social media data that may reflect socioeconomic status or race.

- Monitoring AI credit underwriting systems to check for potentially discriminatory decision-making or unforeseen outcomes, which may include automated techniques for detection and mitigation of algorithmic bias or conventional human oversight by diverse human decision makers.

- Validating that AI credit underwriting systems are not making decisions on a discriminatory basis by conducting periodic testing of models and their results (including assessments of group-equality) and trend analysis of those systems, supplemented by file reviews when warranted.

This list is neither exhaustive nor mandatory. It merely illustrates some of the effective ways to mitigate discrimination risk when using AI credit underwriting models.

In addition to the above, banks perform fair lending compliance assessments, and determine the nature, depth and breadth of the assessment based on the potential risk of disparate impact, the complexity of the model, intended model use and model risk ratings. Fair lending compliance assessments can be qualitative or quantitative depending on the context of model use, materiality, model structure and availability of data. Qualitative analysis may consider mitigating controls from structural design or control processes. Quantitative analysis may involve independent fair lending testing and review by fair lending professionals. Of course, fair lending testing faces certain material constraints, such as the regulatory expectation that such testing be performed independent of the development process and the lack of access to protected characteristic data for testing purposes.

Finally, banks recognize the importance of having diverse teams involved in developing, testing, validating and monitoring AI models to reduce the possibility of introducing bias in algorithms and generating biased outcomes.[40] Such diversity includes diversity of background, as well as diversity of roles, responsibilities and experiences.

---

[40] Testimony of Brian Moynihan, CEO, Bank of America, Before the Senate Banking Committee at 15 (May 26, 2021), https://www.banking.senate.gov/imo/media/doc/MoynihanTestimony%205-26-21.pdf ("Importantly, we take measures to ensure we have a diverse team in place to build, test and refine our AI capabilities. This helps remove the potential bias in algorithms. Ultimately, we understand that members

Question 13: *To what extent do model risk management principles and practices aid or inhibit evaluations of AI-based credit determination approaches for compliance with fair lending laws?*

Historically, model risk management principles and practices do not apply directly to questions of fair lending compliance, except to the extent that the Regulation B criteria for qualifying as an empirically derived, demonstrably, and statistically sound credit scoring system incorporates certain high-level model risk management principles. The Regulation B criteria focus on the quality of the data used to build the model, purpose of the model, use of accepted statistical principles and methodologies in model development and validation, and periodic revalidation of the model to maintain predictive ability.[41]

Nonetheless, existing model risk management principles and practices, including the Model Risk Management Guidance, provide a useful framework for developing and evaluating models, including AI credit underwriting models, that complements fair lending compliance. Specifically, the Model Risk Management Guidance describes aspects of an effective model risk management framework of general applicability to a broad range of models, including AI credit underwriting models, that can be used in fair lending compliance and risk monitoring, for example, by helping banks identify models that may pose fair lending risk and should be subject to fair lending reviews. Although fair lending and model risk management activities are distinct, they work in tandem as banks consider, monitor and test for potential discrimination during model development and independent model validation. While AI models may present more complexity than conventional models, the same model risk management principles and practices apply to both types of models and the same relation between fair lending and model risk management remains in effect.

In applying the Model Risk Management Guidance, BPI urges the Agencies to apply a flexible, risk-based approach that ensures consistent application of the guidance by examiners at all Agencies working across all institutions to –

- Recognize the distinctive features of AI models, specifically the dynamic, constantly evolving aspect of AI model algorithms;
- Accommodate the use of risk management techniques modified, adapted, or targeted specifically for AI models and their complexity;
- Avoid the rigid or prescriptive application of methods and approaches developed and used regularly with conventional underwriting models; and
- Provide the same degree of flexibility to the validation of AI and conventional models developed by third-party vendors.

---

of our team must be held accountable for the output of our AI. Human oversight is a critical factor in AI success.")

[41] 12 C.F.R. § 1002.2(p)(1). The ability to assess fair lending risks in model development frequently is limited by the lack of data regarding which credit applicants belong to protected classes. Regulation B generally prohibits creditors from collecting most types of protected class data, except in the context of mortgage lending. 12 C.F.R. §§ 1002.5(b) and .13.

A significant limitation of the Model Risk Management Guidance is that it applies only to banks, not to nonbank lenders. While banks face intense scrutiny from regulators in complying with the Model Risk Management Guidance, nonbank lenders may utilize AI credit underwriting models with no obligation to follow the Model Risk Management Guidance or answer to regulators through supervisory examinations. The result is uneven and unequal protection for consumers. BPI believes that nonbank use of AI credit underwriting models poses equally significant model and fair lending risks and that banks and nonbanks therefore should be subject to the same standards for reviewing and implementing AI credit underwriting models. For this reason, BPI encourages the CFPB to scrutinize and supervise nonbank lenders for adherence to the model risk criteria found in Regulation B that a model must satisfy to qualify as an empirically derived, demonstrably and statistically sound credit scoring system.

> ***Question 14:*** *As part of their compliance management systems, financial institutions may conduct fair lending risk assessments by using models designed to evaluate fair lending risks ("fair lending risk assessment models"). What challenges, if any, do financial institutions face when applying internal model risk management principles and practices to the development, validation, or use of fair lending risk assessment models based on AI?*

At this time, banks typically use traditional modeling techniques, not AI techniques, to develop fair lending risk assessment models. Therefore, banks do not have significant experience with AI-based fair lending risk assessment models and cannot address the question. As a general matter, BPI member banks note that model risk management principles and practices are not well-suited to evaluating fair lending risk assessment models and, if applied to such models, could be counterproductive and result in less effective fair lending risk assessment models.

> ***Question 15:*** *The Equal Credit Opportunity Act (ECOA), which is implemented by Regulation B, requires creditors to notify an applicant of the principal reasons for taking adverse action for credit or to provide an applicant a disclosure of the right to request those reasons. What approaches can be used to identify the reasons for taking adverse action on a credit application, when AI is employed? Does Regulation B provide sufficient clarity for the statement of reasons for adverse action when AI is used? If not, please describe in detail any opportunities for clarity.*

Regulation B sets forth flexible standards for providing applicants with a statement of the "specific" and "principal" reasons for taking adverse action with supplemental guidance in the official commentary.[42] Appendix C to Regulation B provides creditors with sample adverse action reasons along with broad flexibility to "add or substitute" reasons that reflect the basis of its credit decision making.[43] In 2020, CFPB staff issued a blog post that addressed industry concerns about the challenges in generating explanations for the outcomes of AI credit underwriting models in the form of adverse action reasons.[44] The nature of AI models often requires the use of indirect methods of extracting adverse action reasons from the model. For example, industry is exploring a variety of methods, such as SHAP

---

42 12 C.F.R. § 1002.9(b)(2); 12 C.F.R. part 1002, supplement I, § 1002.9(b)(2).

43 12 C.F.R. part 1002, Appendix C.

44 Patrice Alexander Ficklin, Tom Pahl, Paul Watkins, "Innovation Spotlight: Providing adverse action notices when using AI/ML models" (July 7, 2020), https://www.consumerfinance.gov/about-us/blog/innovation-spotlight-providing-adverse-action-notices-when-using-ai-ml-models/, (**"**The existing regulatory framework has built-in flexibility that can be compatible with AI algorithms.").

and LIME, to identify, isolate and weigh the relative importance of the factors or combinations of factors that most impacted decisions to take adverse action.

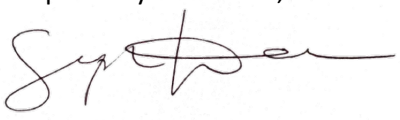
The blog post clarified the flexibility afforded creditors using AI credit underwriting systems to provide adverse action reasons not listed in Appendix C and consistent with the creditor's innovative use of AI for underwriting.[45] Without endorsing any specific tool or method, the blog post accurately noted that "[i]ndustry continues to develop tools to accurately explain complex AI decisions, and we expect more methods will emerge. These developments hold great promise to enhance the explainability of AI and facilitate use of AI for credit underwriting compatible with adverse action notice requirements."[46] The CFPB's blog post also encouraged industry to use existing innovation policy tools "to address areas of regulatory uncertainty" and noted that the CFPB "intends to leverage experiences gained through the innovation policies to inform policy."[47]

BPI and its members very much appreciated these clarifications, as well as CFPB staff's recognition of the challenges and methods under consideration, and believe the blog post provides helpful guidance with appropriate flexibility to facilitate AI innovation. Our suggestions are, therefore, modest in scope. First, the CFPB and the Agencies collectively should consider restating the substance of the blog post in more formal agency guidance with express recognition of the challenges – and the promise – of using indirect methods to extract adverse action reasons. Second, the CFPB should make clear that use of its innovation policy tools is an option, but not a prerequisite, to industry's pursuit of AI innovations in credit underwriting, including evolving methods of explainability.

* * * * *

BPI appreciates the opportunity to respond to the request for information. If you have any questions, please contact the undersigned by phone at 202-589-2432 or by email at *Stephanie.Wake@bpi.com.*

Respectfully submitted,



Stephanie Wake
Vice President, BITS
Bank Policy Institute

---

45 *Id.*

46 *Id.*

47 *Id.*