Part 1 on the Future of AI Regulation—The RFI on AI from U.S. Banking Regulators

April 21, 2021

Several recent developments provide new insight into the future of artificial intelligence ("AI") regulation. First, on March 29, 2021, five U.S. federal regulators published a request for information ("RFI") seeking comments on the use of AI by financial institutions. Second, on April 19, the FTC issued a document entitled "*Aiming for truth, fairness, and equity in your company's use of AI*," which provides seven lessons on what the FTC views as responsible AI use. Third, on April 21, the European Commission released their much-anticipated draft regulation on AI, which is widely viewed as the first step in establishing a GDPR-like comprehensive EU law on automated decision making. In this series on the future of AI regulation, we will examine each of these developments, what they mean for the future of AI regulation, and what companies can do now to prepare for the coming AI regulatory landscape.

The U.S. banking regulators' RFI on AI, which was issued jointly by the FRB, OCC, FDIC, CFPB, and the National Credit Union Administration ("NCUA," and collectively, the "Joint Regulators"), provides several important details for our understanding of the coming AI regulation in the United States, which we discuss in detail below, including:

- Listing several examples of how AI is being used by financial institutions;
- Flagging what the Joint Regulators view as the most significant risks of AI; and
- Identifying existing laws that the Joint Regulators view as applicable to AI.

Examples of AI Use Cases

The RFI includes the following uses of AI by financial institutions:

- Fraud Detection: Using AI to identify suspicious transactions for detecting money laundering and improper employee practices.

- Customer Service: Using voice recognition and natural language processing ("NLP") to automate routine customer interactions (*e.g.*, chatbots), triage customer calls, provide targeted marketing, and customizing trade recommendations.
- Data and Text Analysis: Using AI to analyze regulations, news flow, earnings reports, consumer complaints, analyst ratings changes, and legal documents.
- Credit Decisions: Utilizing AI to employ traditional and alternative data sources to inform, enhance, or supplement credit decisions.
- Risk Management: Using AI to manage liquidity risk through monitoring of market conditions and management of collateral.
- Cybersecurity: Using AI to detect cyber threats and identify compromised systems.

What is perhaps most interesting about this list is the variability of the use cases, and the challenge that poses for any comprehensive AI regulation. Indeed, some applications of AI are relatively low risk, and therefore can be significantly impaired by overregulation. For example, if a company is facing a new dangerous cyber threat, it may need to quickly deploy an AI tool to defend against that attack. Any AI regulation that unnecessarily delays that deployment through onerous testing, governance, training, or policy requirements, poses a significant risk to the company.

AI Risks

The Joint Regulators begin the risk section of the RFI by noting that financial institutions should have processes in place for identifying potential AI risks, and that many of these risks are not unique to AI (*e.g.*, operational vulnerabilities, cyber threats, model risk, vendor risk, unlawful discrimination, unfair practices, etc.). But, the RFI also recognizes that AI does present some unique challenges in the areas of explainability, data usage, and dynamic updating.

Explainability

The RFI defines explainability as how an AI application uses inputs to produce outputs, and the RFI separates the explainability of an AI model's overall functions ("global explainability") from how the model arrives as a particular outcome ("local explainability"). The general view of regulators is that when AI is used to make lending decisions, the lenders must disclose the key data inputs for the model, and which of those inputs were critical in any adverse decision. But the details of that framework remain unclear, including whether the lender needs to disclose:

- All of the inputs, or just the ones that significantly impacted the decision.

- The weight of each input on the decision.
- What the applicant could do to achieve a more favorable result.
- Whether the inputs affect the results independently or synergistically, and how.

The RFI will hopefully provide the Joint Regulators with information that will help provide concrete guidance on these issues, by requesting among other things, which AI use cases pose the greatest explainability challenges, and how financial institutions account for and manage the varied challenges and risks across use cases.

Cybersecurity

The Joint Regulators note that AI can be subject to cyber threats, and in particular, "data poisoning attacks," which attempt to corrupt the AI system to compromise or manipulate its performance. To learn more about these threats and whether specific recommendations should be made to reduce their risks, the Joint Regulators' RFI requests that financial institutions identify cybersecurity risks that are specific to AI, and any effective measures to mitigate those risks.

Dynamic Updating and Model Drift

Many forms of AI involve the ability of the model to learn or adapt as it ingests and processes new data. This can result in the AI model behaving differently over time with out any human involvement ("drift"), which can present significant challenges for testing and validation of the model. This risk arises most acutely when macro changes, such as economic downturns or pandemics, cause significant changes to the inputs of operational AI models.

While the Food and Drug Administration recently published an action plan to deal with the unique challenges of dynamically updating AI models, the Joint Regulators have not yet spoken on this issue. To better understand the risk of AI drift in dynamically updating models, the RFI asks financial institutions how they are addressing this risk, and how they measure whether an AI model that produces different results over time is operating as intended.

Oversight of Third Parties and Alternative Data

Many financial institutions use AI models or data that were developed or provided by third parties, which presents questions as to what level of diligence is required when using an AI tool that was created (in whole or in part) by another company. We have also previously written about creating an AI risk framework for critical vendors, based on prior guidance from the OCC, FRB, and FDIC. The RFI notes that existing agency guidance already contains principles for validation of third-party vendors, and asks

financial institutions to describe the challenges or impediments they are facing in using AI provided by third parties, and how they are managing those risks.

Fair Lending

The Joint Regulators express two concerns relating specifically to loans. First, using AI that lacks transparency or explainability for lending decisions may not be compliant with applicable regulations, such as fair lending laws. Second, and relatedly, if unrepresentative data sets are used to train models, or if the models are poorly designed, AI that is used for loans may contain bias and result in discrimination against protected classes.

Against this backdrop, companies and academics have increasingly focused on deploying tools aimed at reducing AI bias, including by decreasing disparities across protected classes. The Joint Regulators' RFI seeks further information on how companies are addressing bias risk or ensuring the compliance of their AI models with fair lending laws, including by asking:

- What techniques are available to evaluate the compliance of AI-based credit determinations with fair lending laws, or mitigate risks of noncompliance?
- What are the risks that AI can be biased and/or result in discrimination based on protected classes, and are there ways to reduce these risks?
- To what extent do model risk management practices aid or inhibit evaluations of AI-based credit determinations for compliance with fair lending laws?
- Do companies face challenges in applying model risk management practices to the development or use of new models designed to evaluate fair lending risks?
- Does the Equal Credit Opportunity Act Regulation B provide sufficient clarity around the statement of reasons for an adverse credit action when AI is used?

Existing Laws Applicable to AI

The RFI includes an appendix of laws, regulations, and supervisory guidance that the Joint Regulators identify as relevant to AI, including the Fair Credit Reporting Act, the Equal Credit Opportunity Act and its implementing Regulation B, the Fair Housing Act, Section 5 of the Federal Trade Commission Act (prohibiting unfair or deceptive acts or practices), Sections 1031 and 1036 of the Dodd-Frank Act (prohibiting unfair, deceptive, or abusive acts or practices), and the SR 11-7 Supervisory Guidance on Model Risk Management. This should serve as reminder that the Joint Regulators view AI as already

within the scope of their regulatory and supervisory authority and governed by existing laws.

Take-Aways

The RFI is the latest step by the Joint Regulators in assessing the regulatory landscape around AI, but is unlikely to result in new regulations in the near term. It does, however, provide several signals that forward-thinking companies should consider in designing their compliance programs around AI, in order to reduce reputational risk, comply with existing regulations that govern automated decision-making, and prepare for the AI-specific regulation that is likely coming in the European Union, the United States., and elsewhere.

In our next installments in this series on the Future of AI Regulation, we will provide a list of steps that companies can take now to limit the risks of developing AI tools that will be viewed as noncompliant with the AI regulatory landscape that is likely to take shape over the next few years. Those steps will cover:

- accountability
- appeal rights
- bias testing
- board reporting
- business continuity
- cybersecurity
- documentation
- guardrails
- human oversight
- AI inventories
- ongoing monitoring
- privacy protection
- regulatory disclosures
- risk assessments
- training
- transparency
- explainability
- vendor management

* * *

Comments to the RFI may be submitted up until June 1, 2021, and should be submitted to each of the Joint Regulators. Please do not hesitate to contact us with any questions.

NEW YORK

Avi Gesser
agesser@debevoise.com

Anna R. Gressel
argressel@debevoise.com

Amy Aixi Zhang
aazhang@debevoise.com