IEEE SA STANDARDS ASSOCIATION

July 1, 2021

Office of the Comptroller of the Currency (OCC): Docket ID OCC-2020-0049
US Federal Reserve: Docket No. OP-1743
Federal Deposit Insurance Corporation (FDIC): RIN 3064-ZA24
Bureau of Consumer Financial Protection (CFPB): Docket No. CFPB-2021-0004
National Credit Union Administration (NCUA): Docket No. NCUA-2021-0023

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

To the Agencies:

The IEEE Standards Association (IEEE SA) acknowledges the Office of the Comptroller of the Currency (OCC), the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation (FDIC), the Bureau of Consumer Financial Protection (CFPB) and the National Credit Union Administration (NCUA) for their efforts to gather input through their Request for Information and Comment on Financial Institutions Use of AI, Including Machine Learning. We are pleased to provide comments in response to the request. Below please find our general comments along with responses to the specific questions posed.

As background, the IEEE SA, is a globally recognized standards-setting body within IEEE. We develop consensus standards through an open process that engages industry and brings together a broad stakeholder community. IEEE standards set specifications and best practices based on current scientific and technological knowledge. IEEE SA has a portfolio of over 1,500 active standards and over 650 standards under development, including technical and impact standards relating to IoT and Cybersecurity.

Thank you for the opportunity to contribute to your process.

Best regards,

Kristin Little
Senior Public Affairs Manager, IEEE SA

IEEE SA suggests that the use of AI in a safe and sound manner, compliant with applicable laws and regulations, is not enough in itself to ensure that there will be added value for the client (be it a legal person or a natural person). We recommend that discussions on this topic start by focussing on the client (human/end user in the loop). Doing so will make it apparent that the use of AI directly, or indirectly, results in a customer-facing output that will create an imbalance between the financial institutions and its client, ultimately hurting the client. This is especially the case when using self-learning algorithms, with or without the use of hybrid/synthetic data.

The focus therefore, when deciding to make use of AI, would need to be to restore the balance. Thus, rather than concentrating on the AI itself, the USE of AI should be looked at by means of a risk-based approach. Within Europe, the European Commission has taken this philosophy on board in its recent draft regulation for AI.

IEEE has also taken this risk-based approach in its ECPAIS certification scheme which allows for a thorough assessment of the risks of a particular use of AI, with special attention for bias, accountability and transparency. In the future, other elements will be added such as privacy and governance, among others. The value of the IEEE scheme is that it allows for a holistic 360 degree assessment, is documented, audited, and can be certified, thus providing trust in the use of AI.

The ECPAIS scheme takes into account the following. Some human-centered basics when assessing the use of AI:

- Data minimization
- Right to be informed about the use of an AI and automated decisions
- Right to request a human interface
- Right to appeal an automated decision (similar to the European GDPR in article 22)
- Limits on re-use of data for other purposes
- Data sets used for building any algorithm must be relevant for the purpose
- Data sets must be checked for bias
- Signing of an account agreement cannot include agreement to be subjected to automated decisions
- Cyber security is paramount, also to avoid secondary use of data illegally obtained

For more information on the ECPAIS certification scheme please contact Alpesh Shah at a.shah@ieee.org.

Questions posed by the US Department of Treasury:

Explainability

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?

Most AI technologies in the banking world will not be 'explainable'. Thus, the emphasis should be on 'transparency' and 'accountability'. This implies, for example, the availability of documented risk assessment procedures, and ensuring a human in the loop for AI 'decisions' in client facing processes.

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.

The May 2020 UK Information Commissioner's Office report on "Explaining decisions made with AI" identified almost all approaches to explanation as post hoc. The post hoc approach describes a plausible process by which an AI system could have come to a decision, not necessarily the actual one that the AI system used. An AI may game the system (e.g., identify horses based on metadata about where the images came from). Causation and correlation may be confused. Post hoc reasoning is inductive or abductive, with the limitations those bring. Most importantly, a post hoc method cannot be used as evidence of what happened.

Having said that, there is no indication that an inability to "open" the black box (the techno-philosophical challenge of knowing the unknowable) will stop the growth of AI systems, including AI systems designed by AI systems. Given that continued development, the most important requirement in using AI systems is an acknowledgement that the process of the AI system is not known, and claims may be contested. In the Australian Evidence Act 1995 (Cth) there can be an expectation of some technology, e.g., a photocopier, that it "ordinarily produces that outcome". No such assumption should be made for AI systems explained through post hoc reasoning.

We suggest that the emphasis on 'explainability' should be left and replaced by transparency and accountability through well documented risk assessment, complemented with (well documented) governance procedures. Certification of the result would be very beneficial.

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?

Lack of explainability is more of a challenge where the AI decision may impact the opportunity or wellbeing of a person. In these cases, the system should require a "human in the loop."

IEEE suggests that every customer-facing decision should allow for clear processes to ensure a human in the loop and a corresponding right for re-assessment of the decision impacting the customer. As a rule, clients should not be subjected to automated decisions without the right for a reassessment. Note the following from the GDPR:

1. *The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.*
2. *Paragraph 1 shall not apply if the decision:*
    1. *is necessary for entering into, or performance of, a contract between the data subject and a data controller;*
    2. *is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or*
    3. *is based on the data subject's explicit consent.*
3. *In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.*[1]

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.

Hybrid or synthetic data as an output of AI will greatly complicate data quality and data processing. There is no guarantee that the outcome is 'right'. Hence, the type of AI used should be appropriate for the particular process to ensure the right type of data management for that specific process or service.

[1] GDPR Article 22

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?

Whether the introduction of AI systems most benefits attackers or the attacked is a subject of considerable debate within the cybersecurity community. Vendor literature generally focuses on benefits of protection rather than attack.

One risk is that the introduction of AI into a system already vulnerable to cybersecurity breaches will only magnify the problem. To manage cybersecurity risks related to AI, banks can demonstrate that sufficient measures have been taken to combat cybersecurity using a certification scheme like ECPAIS-- IEEE's Ethics Certification Program for Autonomous and Intelligent Systems.

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?

Dynamic updating, or any mechanism which draws on data from the analogue world, introduces a second level of uncertainty. In particular, these systems will not be deterministic, and those which digitize analogue values (e.g., age) will always have quantization error. Thus, dynamic updating may not be appropriate for all applications. Different applications warrant different types of AI.

AI Use by Community Institutions

Question 9: Do community institutions face particular challenges in developing, adopting, and using AI? If so, please provide detail about such challenges. What practices are employed to address those impediments or challenges?

We suggest that a (documented) risk assessment should be carried out also when relying on third party services to address the challenges faced by community institutions. This risk assessment should be mandatory for both in-house and third party work.

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?

Commercial secrecy of the vendor might be an issue. See Houston Federation of Teachers Local 2415 v Houston Independent School District (2017) 251 F.Supp.3d 1168. This case involved "the use of privately developed algorithms to terminate public school teachers for ineffective performance."

Fair Lending

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

Ensuring a human in the loop will greatly help to address fairness and transparency. Also a (documented) risk assessment looking at transparency, accountability, and bias will positively contribute (e.g., IEEE's ECPAIS)

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 mentioned in our general comments, ensuring a human in the loop will greatly help to address discrimination. It should always start with a documented risk assessment looking at transparency, accountability and bias (e.g., IEEE's ECPAIS).

Question 13: To what extent do model risk management principles and practices aid or inhibit

Model risk management principles are challenged by the inability to develop application computer controls to test models in a black box AI system.

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.

The requirement for a statement of specific reasons for the action taken is a challenge when depending on post hoc explanation. A further problem arises if the AI system generates the explanation, as this may encourage it to propose an acceptable explanation rather than a true explanation. In this case if a white box (a system mimicking the black box but with known characteristics) can be shown to come to the same conclusion, the explanation from the white box would suffice.

Additional Considerations

Question 17: To the extent not already discussed, please identify any benefits or risks to financial institutions' customers or prospective customers from the use of AI by those financial institutions. Please provide any suggestions on how to maximize benefits or address any identified risks.

The increasing use of AI creates a sizable imbalance between the client and the bank. Because of the opaque/automated nature of the decision making, the client has very little or no means to object to decisions he is faced with. We recommend addressing this imbalance to avoid unfair outcomes and to ensure access to bank services for all.

Also, we recommend that the type of data that can be used by financial institutions should be

limited to and grounded in the service they are needed for. The client should be given clear rights regarding access to their data, right to rectify, etc. A data privacy law at the federal level would be very helpful to restore part of the balance and avoid a sectoral approach.

From a risk to reputation perspective, the use of AI systems to automate existing jobs has a strong parallel to business process outsourcing/offshoring, and it is recommended that they be considered in any AI risk framework.

General Comments

In 2020, as a component of our Ethics Certification Program for AIS (ECPAIS), IEEE began an investigation into the adoption of AI in Financial Services and Instruments -- *Economic Franchise in Under-regulated Financial Services & Instruments employing Emerging Technologies (AIS)* (fair, transparent and secure access to credit, receipt & making of payments, goods and services and accountable supporting ecosystem that uphold users' privacy). The program scrutinized the role of AI and emerging instruments in a poorly regulated financial environment partly driven by the pace of innovation that considered a number of major areas of concern:

1- Accessibility to Fair banking, use/saving of money, Credit and demonetization
2- Freedom from Unacceptable Bias in Financial Services
3- Transparency & Accountability of the packaging and Delivery of Financial Instruments and Services
4- Fair Accessibility to Capital and the Supporting Ecosystem
5- Franchised Participation in the Financial Ecosystems

A comprehensive suite of technical and governance criteria under each topic above are currently being formalized and can be shared as appropriate.