

Testimony before US House of Representatives, Financial Services Committee, the Task Force on Artificial Intelligence

on *Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age*

by Jeffery Yong, Principal Advisor, Financial Stability Institute, Bank for International Settlements

13 October 2021

Good afternoon Chair Foster, Ranking Member Gonzalez and distinguished members of the Task Force. My name is Jeffery Yong and I am a Principal Advisor at the Financial Stability Institute of the Bank for International Settlements (BIS). I offer my remarks today entirely in my personal capacity based on a publication that I co-authored with my colleague Jermy Prenio entitled *FSI Insights no 35, Humans keeping AI in check – emerging regulatory expectations in the financial sector*.¹ The views expressed in that paper are our own and do not necessarily represent those of the BIS, its members or the Basel-based committees. I am appearing before the Task Force voluntarily and would like to note that my statements here today are similarly my personal views, and they do not represent the official views of the BIS, its members or the Basel-based committees.

By way of background, the Financial Stability Institute (FSI)² is a unit within the BIS with a mandate to support implementation of global regulatory standards and sound supervisory practices by central banks and financial sector regulatory and supervisory authorities worldwide. One of the ways the FSI carries out this mandate is through its policy implementation work, which involves publishing FSI Insights papers. The papers aim to contribute to international discussions on a range of contemporary regulatory and supervisory policy issues and implementation challenges faced by financial sector authorities.

In preparing FSI Insights no 35, my co-author and I found that regulatory expectations on the use of artificial intelligence (AI) in financial services were at a nascent stage. Accordingly, we drafted the paper with four key objectives:

1. to identify emerging common financial regulatory themes surrounding AI governance;
2. to assess how similar or different these common regulatory themes are viewed in the context of AI vis-à-vis that of traditional financial models;
3. to explore how existing international financial regulatory standards may be applied in the context of AI governance; and
4. to examine challenges in implementing the common regulatory themes.

To this end, we canvassed a selection of policy documents on AI governance issued by financial authorities or groups formed by them, as well as other cross-industry AI governance guidance that apply to the financial sector. In total, we examined 19 policy documents issued by 16 national or regional authorities and two international organisations. Most of these documents are either discussion papers or high-level principles, which underscores the fact that financial regulatory thinking in this area is at a very early stage.

We identified five common themes that recur in the policy documents that we examined. These are reliability, accountability, transparency, fairness and ethics.

¹ See [FSI Insights, no 35, Humans keeping AI in check – emerging regulatory expectations in the financial sector](#).

² See [Financial Stability Institute](#).

On the theme of reliability, emerging supervisory expectations for AI and traditional models appear to be similar. What seems to be different is that the reliability of AI models is viewed from the perspective of avoiding harm to data subjects, for example through discrimination.

On the theme of accountability, it is acknowledged that both traditional and AI models require human intervention. In the case of AI, however, this requirement is motivated by the need to make sure that decisions based on AI models do not result in unfair or unethical outcomes. Moreover, external accountability is emphasised in the case of AI models so that data subjects are aware of AI-driven decisions and have channels for recourse.

On the theme of transparency, supervisory expectations related to explainability and auditability are similar for AI and traditional models. However, expectations on external disclosure are unique to AI models. This refers to expectations that firms using AI models should make data subjects aware of AI-driven decisions that impact them, including how their data is used.

On the theme of fairness, there is a distinct and strong emphasis in emerging supervisory expectations on this aspect in the case of AI models. Fairness is commonly described in the policy documents we reviewed in terms of avoiding discriminatory outcomes.

Similarly on the theme of ethics, there is also a distinct and strong emphasis on this aspect in AI models. Ethics expectations are broader than fairness and relate to ascertaining that customers will not be exploited or harmed, either through discrimination or other causes (eg AI using illegally obtained information).

Given the similarities of these themes in the context of AI and traditional financial models, existing financial regulatory standards that govern the use of traditional models may be applied in the context of AI. However, there may be scope to do more in defining financial regulatory expectations related to fairness and ethics. These could supplement consumer protection laws that cover non-discrimination clauses, which could also apply in the context of the use of AI in the financial sector.

The use of AI in the financial sector, however, presents certain challenges in a direct application of existing financial regulatory requirements. A key challenge is due to the level of complexity and lack of explainability that characterise AI models. These limit the transparency of the models, which in turn makes it challenging for financial supervisors to assess the reliability, accountability, fairness and ethics in the use of AI in the financial services industry.

A way to overcome these challenges is to consider a tailored and coordinated regulatory and supervisory approach. This means differentiating the regulatory and supervisory treatment on the use of AI models, depending on the conduct and prudential risks that they pose. In addition, coordination between conduct, prudential, as well as data protection authorities will help address the cross-cutting implications of the use of AI models in the financial services industry.

Thank you very much.