Deloitte.
デロイト トーマツ

AI regulation in the financial sector

How to ensure financial institutions' accountability

September 2023

MAKING AN
IMPACT THAT
MATTERS
since 1845

Contents

1. Introduction

The adoption of artificial intelligence (AI) systems and techniques has grown considerably in the financial services sector, driven by the increase in available data and the improvement of computing capacity.[1] AI has the potential to transform traditional business models in the sector by contributing to greater efficiency and profitability through the reduction of friction costs and improvements in productivity. On the other hand, there are concerns that AI could amplify existing risks and/or give rise to new risks. It is therefore considered that AI needs to be subject to the appropriate regulations to ensure that its 'invisible hand'[2] is used for the benefit of our society.

The rest of this article discusses how financial institutions should address AI risks and how AI needs to be regulated, introducing developments on AI regulation at the international and financial sector levels.

It should be noted that this article is intended to be read in conjunction with the report 'Artificial Intelligence (AI) state of play in insurance regulation' (Center for Regulatory Strategy US, Deloitte). By doing so, readers will be able to understand the overall global AI regulation landscape more comprehensively.

2. International initiatives

One of the first international agreements related to AI was the 'Recommendation of the Council on Artificial Intelligence - Principles for responsible stewardship of trustworthy AI', which was adopted by **the OECD** in May 2019 (hereinafter referred to as 'OECD AI Principles'). The OECD AI Principles present five principles (Table 1), acknowledging in its preamble that 'AI has the potential to transform societies and economic sectors by, for example, contributing to positive sustainable global economic activity, increasing innovation and productivity and helping respond to key global challenges. On the other hand, these transformations may have disparate effects within/between societies and economies. Therefore, the trustworthiness of AI systems is a key factor for the diffusion and adoption of AI'[3]. The OECD AI Principles were subsequently integrated into the annex of the G20 Leaders' Declaration in June 2019 as the 'G20 AI Principles'[4].

Table 1. Outline of the OECD AI Principles

1. **Inclusive growth, sustainable development and well-being**: Stakeholders should proactively engage in responsible stewardship of trustworthy AI in pursuit of beneficial outcomes for people and the planet.
2. **Human-centred values and fairness**: AI actors should respect the rule of law, human rights and democratic values throughout the AI system lifecycle and, to that end, implement the necessary mechanisms and safeguards.
3. **Transparency and explainability**: AI actors should commit to transparency and responsible disclosure regarding AI systems and, to that end, provide meaningful information to (i) foster a general understanding of AI systems, (ii) make stakeholders aware of their interactions with AI

---

[1] OECD (2021), 'AI in finance', in OECD Business and Finance Outlook 2021: AI in Business and Finance, OECD Publishing, Paris, https://doi.org/10.1787/39b6299a-en.

[2] IMF (2023) 'The Power and Perils of the "Artificial Hand": Considering AI Through the Ideas of Adam Smith' by Gita Gopinath, IMF Speech to commemorate the 300th anniversary of Adam Smith's birth on 5 June 2023, https://www.imf.org/en/News/Articles/2023/06/05/sp060523-fdmd-ai-adamsmith.

[3] OECD (2019) 'Recommendation of the Council on Artificial Intelligence', https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[4] Ministry of Foreign Affairs of Japan (2019) 'G20 OSAKA LEADERS' DECLARATION', https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/en/documents/final_g20_osaka_leaders_declaration.html.

systems, (iii) enable those affected by an AI system to understand the outcome and (iv) enable those adversely affected by an AI system to challenge its outcome.

4. **Robustness, security and safety**: AI systems should be robust, secure and safe throughout their entire lifecycle and, to that end, ensure traceability that enables analysis of the AI system's outcomes.

5. **Accountability**: AI actors should be accountable for the proper functioning of AI systems.

In light of the ethical use of AI, one global initiative was the adoption of the 'Recommendation on the Ethics of Artificial Intelligence' by **the United Nations Educational, Scientific and Cultural Organization** (UNESCO) in November 2021 [5] . The UNESCO Recommendations present five objectives, which include the following.

■   To provide a universal framework of values, principles and actions to guide jurisdictions in the formulation of their legislation, policies or other instruments regarding AI, consistent with international law

■   To guide the actions of individuals, groups, communities, institutions and private sector companies to ensure the embedding of ethics in all stages of the AI system life cycle

■   To protect, promote and respect human rights, fundamental freedoms and human dignity and equality; to safeguard the interests of present and future generations; to preserve the environment, biodiversity and ecosystems; and to respect cultural

diversity in all stages of the AI system life cycle

The UNESCO Recommendations highlight 11 areas that require policy actions. One of these areas is 'Ethical Governance and Stewardship', in which jurisdictions are required, for example, to ensure that AI governance mechanisms are inclusive, transparent, multidisciplinary, multilateral and multi-stakeholder.

An agreement by G7 Leaders in May 2023 can be regarded as a more recent milestone. During **the G7 Summit** held in Hiroshima, Japan, the leaders agreed on advancing international discussions on inclusive AI governance and interoperability to achieve their common vision and goal of trustworthy AI. They committed in the Summit's Communiqué to continue to work on this agenda item in cooperation with other organisations, including the OECD, recognising the importance of procedures that advance transparency, openness, fair processes, impartiality, privacy and inclusiveness in promoting responsible AI as well as the importance of international discussions on AI governance and interoperability between AI governance frameworks[6]. UNESCO subsequently released a statement in support of the G7 initiatives, underscoring the necessity of 'ethical guardrails' for the safe development of AI[7].

---

[5] UNESCO (2021) 'Recommendation on the Ethics of Artificial Intelligence', https://unesdoc.unesco.org/ark:/48223/pf0000381137.

[6] Ministry of Foreign Affairs of Japan (2023) 'G7 Hiroshima Leaders' Communiqué', https://www.mofa.go.jp/files/100506878.pdf.

[7] UNESCO (2023) 'UNESCO supports G7 leaders calling for 'AI guardrails'', https://www.unesco.org/en/articles/unesco-supports-g7-leaders-calling-ai-guardrails.

3. Developments in major jurisdictions

At jurisdictional levels, **the European Union** (EU) is one of the jurisdictions that have made significant progress in establishing a regulatory framework for AI. On 14 June 2023, the European Parliament adopted the Artificial Intelligence Act (hereinafter referred to as 'EU AI Act'). The objective of this act is to 'promote the uptake of human centric and trustworthy artificial intelligence and to ensure a high level of protection of health, safety, fundamental rights, democracy and rule of law and the environment from harmful effects of artificial intelligence systems in the Union while supporting innovation and improving the functioning of the internal market'[8].

The EU AI Act follows a risk-based approach, differentiating between uses of AI that create (i) an unacceptable risk, (ii) a high risk or (iii) a low or minimal risk. AI practices that create unacceptable risks are prohibited and AI systems that create high risks are subject to the prescribed regulatory requirements (Tables 2 and 3). It is noteworthy that the use of AI systems for decision-making in life and health insurance contracts was added as a use case for high-risk AI systems in the text adopted by the European Parliament.

The EU AI Act will undergo further consideration and finalisation through a process known as a 'trilogue' that involves the European Parliament, the Council of the European Union and the European Commission.

Table 2. Overview of prohibited AI practices and high-risk AI systems[9]

Prohibited AI practices
• **Use of AI systems that deploy subliminal techniques** beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective of materially distorting the person's behaviour by appreciably impairing his/her ability to make an informed decision
• **Use of AI systems that exploit vulnerabilities of a person**, such as characteristics of his/her known or predicted personality traits, social or economic situation, age, physical or mental ability, with the objective or to the effect of materially distorting his/her behaviour
• **Use of biometric categorisation systems** that categorise natural persons according to sensitive or protected attributes, etc.
• **Use of AI systems for the social scoring** of natural persons over a certain period of time, with the social score leading to detrimental or unfavourable treatment, etc. for this person
• **Use of 'real-time' remote biometric identification systems** in publicly accessible spaces
• **Use of AI systems for assessing the risk of a natural person for offending** or for predicting the occurrence of criminal or administrative offences

---

based on his/her profiling, etc.

- **Use of AI systems that create facial recognition databases** through the untargeted scraping of facial images from the internet or CCTV footage

- **Use of AI systems that infer the emotions of a natural person** in the areas of law enforcement and border management, or in the workplace and educational institutions

- **Use of AI systems for the analysis of recorded footage** from publicly accessible spaces through 'post' remote biometric identification systems

High-risk AI systems

- **An AI system that is intended to be used as a safety component of a product** (or an AI system itself is a product) covered by the legislation set separately **and the product**, whose safety component is the AI system, **that is required to undergo a third-party conformity assessment related to risks for health and safety**.

- AI systems falling under one or more of the critical areas and use cases listed below (in cases where these systems pose a significant risk of harm to the health, safety or fundamental rights of natural persons)

  a. **Biometric identification and categorisation of natural persons**

  b. **Management and operation of critical infrastructure** (e.g., road, rail and air traffic as well as water, gas, heating, electricity and critical digital infrastructure)

  c. **Education and vocational training**

  d. **Employment, workers management and access to self-employment**

  e. **Access to and enjoyment of essential private services and public services and benefits** (including AI systems intended to be

used for making decisions or materially influencing decisions on the eligibility of natural persons for health and life insurance)

  f. **Law enforcement**

  g. **Migration, asylum and border control management**

  h. **Administration of justice and democratic processes**

Table 3. Requirements associated with high-risk AI systems

- ■ **Regulatory requirements for high-risk AI systems**

  - Risk management system

  - Data governance

  - Technical documentation of AI systems

  - Record-keeping

  - Transparency and provision of information to users

  - Human oversight

  - Accuracy, robustness and cybersecurity

- ■ **Obligations of providers, users, etc. of high-risk AI systems**

  - Quality management system

  - Automatically generated logs

  - Corrective actions

  - Supervisory reporting, etc.

The regulatory trends in **the U.S.**, particularly those within the insurance sector, are covered in detail in the report 'Artificial Intelligence (AI) state of play in insurance regulation' mentioned at the beginning of this article. In recent developments, the Colorado Division of Insurance published in May 2023 an amended version of their AI bill,[10] which was initially proposed in February 2023. The bill establishes governance and risk management requirements for life insurers that use external consumer data and information sources (ECDIS) as well as algorithms and predictive models that use ECDIS.

One of the most material changes from the initial proposal seems to be the removal of the definition of 'disproportionately Negative Outcome' that was defined as 'a result or effect that has been found to have a detrimental impact on a group as defined by race, color, national or ethnic origin, religion, sex, sexual orientation, disability, gender identify, or gender expression, and that impact is material even after accounting for factors that define similarly situated consumers'. Instead, the revised bill explicitly requires insurers to establish a 'risk-based' framework for governance and risk management. Nevertheless, insurers are still required within their governance and risk management framework to, for example, have in place written policies and processes for the design, development, testing, deployment, use and on-going monitoring of algorithms and predictive models that use ECDIS as well as for the selection and oversight of vendors.

In **the U.K.**, the Department for Science, Innovation and Technology published a draft AI regulatory framework in March 2023[11]. The draft framework is built on essential characteristics, which include the following: (i) pro-innovation to enable responsible innovation; (ii) proportionate to avoid unnecessary burdens for businesses and regulators; and (iii) trustworthy to address real risks and foster public trust.

The key concept of the draft framework is 'regulating the use, not the technology'. The proposed approach is to base regulations on the outcomes that AI is likely to generate in accordance with the following five principles: (i) safety, security and robustness; (ii) appropriate transparency and explainability; (iii) fairness; (iv) accountability and governance and (v) contestability and redress[12].

In **Japan**, the AI Strategy Council of the Cabinet Office published a document titled 'Tentative summary of AI issues' in May 2023[13], in which the council presented three basic guiding principles as follows.

■ Leadership: Japan will play a leading role towards developing international rules on AI.

■ Addressing concerns and risks: Japan will seek to adequately address concerns and risks associated with AI to support the development, provision and use of AI.

■ Swift and flexible responses: Japan plans to involve diverse stakeholders and respond rapidly and flexibly to solve AI-related challenges.

---

[10] Colorado Division of Insurance (2023) 'Draft proposed new regulation xx-xx-xx: Governance and risk management framework requirements for life insurance carriers' use of external consumer data and information sources, algorithms, and predictive models', https://communications.willkie.com/125/2137/uploads-(icalendars-pdf-documents)/draft-proposed-algorithm-and-predictive-model-governance-regulation-version-5.26.23-redlined.pdf.

[11] U.K. Department for Science, Innovation and Technology (2023) 'Policy paper: A pro-innovation approach to AI regulation',

https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (updated on 22 June 2023).

[12] For an overview of the draft framework, see 'Regulatory developments in the global insurance sector (Vol. 33, March to April 2023)' (Deloitte Tohmatsu): https://www2.deloitte.com/content/dam/Deloitte/jp/Documents/financial-services/ins/202304_ins_regulation_eng.pdf.

[13] Cabinet Office, Government of Japan (2023) 'Tentative summary of AI issues', https://www8.cao.go.jp/cstp/ai/ronten_youshi_yaku.pdf.

Regarding the approach to addressing risks, a basis policy proposed in the council's document is that AI developers, providers, users, etc. primarily assess the risks of AI by themselves. Subsequently, as necessary, the government, etc. will consider having in place frameworks to address those risks. In addition, the Financial Services Agency (FSA) has indicated its intention to consider viewpoints related to concerns and risks associated with AI based on the AI Strategy Council's document[14].

4. Regulatory developments in the financial sector

Discussions regarding the development of regulatory and supervisory frameworks for AI in the financial sector are underway. **The Basel Committee on Banking Supervision** (BCBS) mentioned the use of AI and machine learning (ML) as one of the priorities in their work programme for 2021 to 2022 published in April 2021[15]. Subsequently, in their March 2022 newsletter, the BCBS presented their intention to continue discussions focusing on (i) the extent and degree to which the outcomes of models can be understood and explained, (ii) AI/ML model governance structures and (iii) the potential implications of broader usage of AI/ML models for the resilience of individual banks and financial stability[16].

In the UK, **the Bank of England, Prudential Regulation Authority** (PRA) **and Financial Conduct Authority** (FCA) published a discussion paper titled 'Artificial Intelligence and Machine Learning' in October 2022. This paper addresses several key aspects, including (i) potential benefits, risks and harms from the use of AI in financial services, (ii) applicability of the current legal requirements and guidance to risks associated with AI and (iii) necessity of exploring additional policy actions, etc.[17] Feedback from stakeholders to this discussion paper is expected to be published by the end of 2023[18].

In the insurance sector, **the European Insurance and Occupational Pensions Authority** (EIOPA) published their AI governance principles titled 'Artificial intelligence

---

[14] Financial Services Agency (2023) 'Responses to the comments for the public consultation document, "A summary of issues and practices for supervisory dialogues with regard to financial institutions' IT governance" (updated version)', https://www.fsa.go.jp/news/r4/sonota/20230630/20230630.html.

[15] Basel Committee on Banking Supervision (2021) 'Basel Committee publishes work programme and strategic priorities for 2021-22', https://www.bis.org/press/p210416.htm.

[16] Basel Committee on Banking Supervision (2022) 'Newsletter on artificial intelligence and machine learning', https://www.bis.org/publ/bcbs_nl27.htm.

[17] Bank of England (2022) 'DP5/22 – Artificial Intelligence and Machine Learning', https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence.

[18] Financial Conduct Authority (2023) 'DP22/4: Artificial Intelligence' (last updated on 25 July 2023), https://www.fca.org.uk/publications/discussion-papers/dp22-4-artificial-intelligence.

governance principles: Towards ethical and trustworthy artificial intelligence in the European Insurance Sector' in June 2021[19]. Six principles presented are: (i) principle of proportionality; (ii) principle of fairness and non-discrimination; (iii) principle of transparency; (iv) principle of human oversight; (v) principle of data governance of record keeping and (vi) principle of robustness and performance[20].

One of the most recent developments in the insurance sector is the exposure draft for 'Model Bulletin: Use of algorithms, predictive models, and artificial intelligence systems by insurers,' which was published by **the National Association of Insurance Commissioners** (NAIC) in July 2023. This draft model bulletin provides regulatory guidance and expectations on the use of AI systems by insurers, covering governance, risk management, internal controls and third-party risk management[21].

In the securities sector, **the International Organization of Securities Commissions** (IOSCO) published a report titled 'The use of artificial intelligence and machine learning by market intermediaries and asset managers: Final report' in September 2021[22]. The IOSCO sets forth guidance for supervisory authorities regarding (i) governance, (ii) testing and monitoring of algorithms, (iii) compliance and risk management, (iv) management of third-party service providers, (v) disclosures and (vi) data governance.

5. Discussion: Addressing the risks of AI

AI technology is evolving rapidly, and it would not, therefore, be an overstatement at this juncture to say that envisioning its ultimate destination is nearly impossible. This means that there are not yet optimal solutions and/or best practices not only in the deployment of AI by financial institutions but also in the regulatory and supervisory approaches used by regulatory authorities to address the risks associated with AI. Amidst this landscape, how should financial institutions strategise to effectively navigate the risks associated with AI in order to harness the opportunities (to be) presented by AI?

One approach is to enhance AI literacy. By deepening people's understanding of AI use cases and its associated risks, a foundation can be built for the effective implementation of AI and the pragmatic management of its risks. Moreover, it is crucial to establish internal governance frameworks, risk management structures and operational rules for AI, all while keeping abreast of the latest regulatory developments. Financial institutions can then leverage AI to achieve, for example, further efficiency in their operations.

In the pursuit of such initiatives, financial institutions must always be mindful of 'principles' from a risk management perspective. These principles include the OECD AI Principles mentioned at the beginning of this article. As financial institutions strive to achieve business sustainability and customer-centric operations, financial institutions must be fully accountable for the outcomes of

[19] EIOPA (2021) 'Artificial intelligence governance principles: Towards ethical and trustworthy artificial intelligence in the European insurance sector: A report from EIOPA's Consultative Expert Group on Digital Ethics in insurance', https://www.eiopa.europa.eu/system/files/2021-06/eiopa-ai-governance-principles-june-2021.pdf.

[20] For an overview of the EIOPA's AI principles, see 'Regulatory developments in the global insurance sector (Vol. 12, June to July 2021)' (Deloitte Tohmatsu): https://www2.deloitte.com/content/dam/Deloitte/jp/Documents/financial-services/ins/202107_ins_regulation.pdf.

[21] NAIC (2023) 'Exposure Draft of the Model Bulletin on the Use of Algorithms, Predictive Models, and Artificial Intelligence Systems by Insurers 7/17/2023', https://content.naic.org/sites/default/files/07172023-exposure-draft-ai-model-bulletin.docx.

[22] IOSCO (2021) 'The use of artificial intelligence and machine learning by market intermediaries and asset managers: Final Report', https://www.iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf.

AI when they use AI through, in particular, business processes that could impact their customers, such as assessing credit risks, underwriting insurance or evaluating insurance claims. In such cases, maintaining accountability for the outcomes of AI becomes paramount. To that end, these institutions need to have in place governance frameworks and internal controls that enable them to maintain their accountability.

From the standpoint of regulation and supervision, financial regulators and supervisors are strongly encouraged to regulate and supervise the use of AI by financial institutions appropriately based on those principles for mitigating risks that can arise from AI usage, particularly with a focus on customer protection. Further policy actions are expected in the financial sector.

End of article

Note: The opinions expressed in this article are those of the author and do not represent the official views of the organisation to which the author belongs.
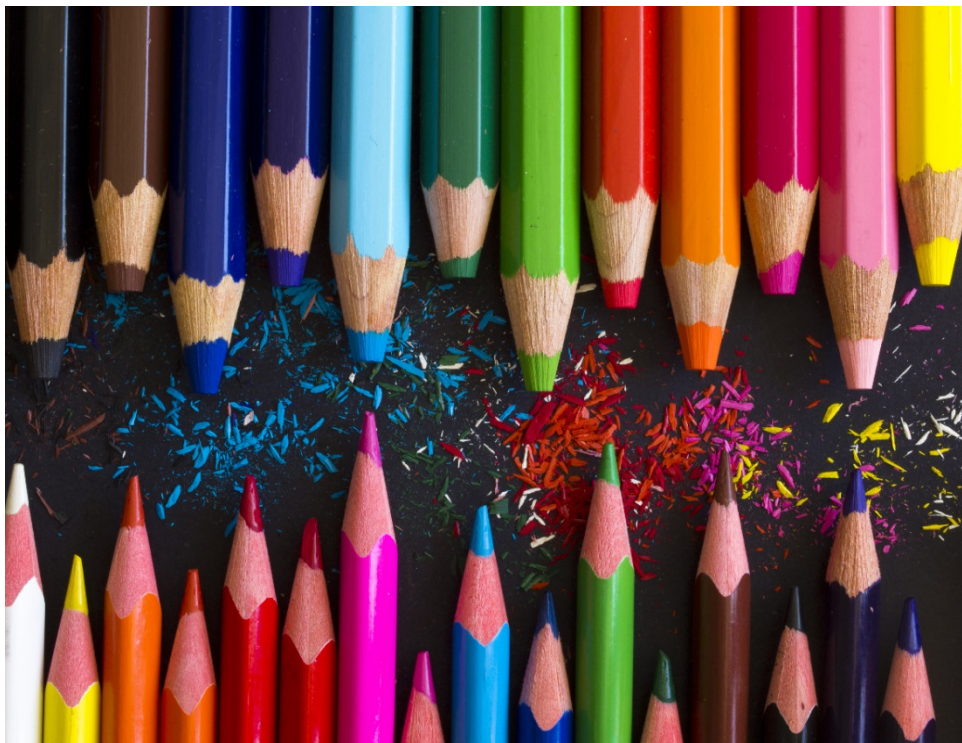
Author

Shinya Kobayashi

Managing Director

Center for Risk Management Strategy (CRMS)

Risk Advisory

Deloitte Touche Tohmatsu LLC

Deloitte.
デロイト トーマツ

IS 669126 / ISO 27001

BCMS 764479 / ISO 22301