Box 2
Potential risks and supervision trends in the use of artificial intelligence (AI) technology by financial institutions

In recent years, financial institutions have increasingly used artificial intelligence (AI). AI technology can effectively improve operational efficiency and customer service experience and thereby benefit financial institutions and consumers, but it may also pose problems such as financial exclusion,[1] privacy infringement, black box operations, high outsourcing concentration and herd behavior, which could affect financial stability. Therefore, while AI technology innovation is encouraged, how to appropriately regulate AI to ensure consumer rights and stability of the financial system has become an important issue for supervisory authorities across the world. The following section briefly describes the benefits and potential risks of AI, introduces the international trends in the supervision of the use of AI in the financial industry, and gives account of Taiwan's policies and guidelines for AI usage in financial institutions devised by referencing international supervision trends.

1. Benefits and potential risks of AI

At present, there is no consistent definition of AI accepted by supervisory agencies around the world. The most commonly cited one is from the Financial Stability Board (FSB):[2] "the theory and development of computer systems able to perform tasks that traditionally required human intelligence." Generative Artificial Intelligence (GenAI), which has been booming recently, refers to related AI systems that can generate content simulating human intelligence. The content in GenAI includes articles, images, audio, videos, and program codes, but it is not limited to the above-mentioned applications.

AI technology can process large amounts of data quickly through strong computing power and produce great benefits such as improving forecasting capabilities, optimizing operations and customizing services. Related applications include automating internal processes, analyzing customer information to provide customized suggestions, and streamlining the processes of customer services through technologies such as facial recognition and image recognition. GenAI technology, which has emerged in recent years, has also brought substantial changes to human life. People are able to obtain results in a very short time by simply inputting their needs and related data into the GenAI system, significantly reducing manual work time.

Although the use of AI technology has many benefits, its applications in the financial

industry may bring about the following potential risks, which, if not properly supervised and managed, could harm financial consumer rights and financial stability.

(1) For financial consumers: There are concerns such as breach of personal privacy, bias or discrimination in prediction results.

(2) For financial institutions: They face risks such as AI black-box operations, unclear responsibilities, and concentration in operations outsourced to a limited number of third-party providers.

(3) For financial markets: AI applications such as high-frequency program trading may cause herd behavior or increase market connectivity.

In addition, "The Global Risk Report 2024"[3] (hereinafter referred to as the Report) released by the World Economic Forum (WEF) in January 2024 listed "AI-generated misinformation and disinformation" as the second largest risk in the world in 2024 and the top risk in the next two years. The Report pointed out that if AI is not properly managed and thus misused, it might cause concerns such as hate crimes and terrorism, loss of job opportunities, crime and cyber-attacks, prejudice and discrimination, and even impact the global political system, economic markets and national security.

2. International supervisory trends in the use of AI in the financial industry

As the use of AI in the financial industry becomes increasingly popular, how to regulate it appropriately to safeguard consumer rights and financial stability has caused greater attention from international organizations and financial authorities around the world. In 2019, the Organization for Economic Cooperation and Development (OECD) first proposed the "Recommendation of the Council on Artificial Intelligence,"[4] listing five important principles, including: (1) inclusive growth, sustainable development and well-being, (2) human-centered values and fairness, (3) transparency and explainability, (4) robustness, security and safety, and (5) accountability. The five principles have been adopted by the G20 members. Since then, international financial organizations have successively issued supervisory recommendations for AI usage by financial institutions, and the EU has passed the *Artificial Intelligence Act*. Major countries have also successively proposed supervisory principles or guidelines for AI applications by taking reference from the recommendations of international organizations.

2.1 International financial organizations proposed principles or recommendations for the use of AI by financial institutions

The International Organization of Securities Regulators (IOSCO) issued a guidance document in September 2021,[5] proposing six supervisory measures for the use of AI by market intermediaries and asset management institutions. The six measures included requiring financial institutions to: (1) establish appropriate governance, control and supervision structures; (2) continuously monitor the development, testing, operation and performance of AI; (3) ensure that personnel have sufficient knowledge, skills and experience to use and supervise the outputs from AI systems; (4) understand their dependence on third-party AI service providers and establish appropriate management and supervision mechanisms; (5) provide sufficient transparency and information disclosure to investors, competent authorities and stakeholders; and (6) establish appropriate control mechanisms to ensure that bias in data and system performance is minimized.

Moreover, the Financial Stability Institute (FSI) under the BIS released an AI supervision report in August 2021,[6] recommending that financial supervisory agencies should adopt AI-related supervision measures based on four principles: (1) transparency, (2) trustworthiness and soundness, (3) accountability, and (4) fairness and ethics, and should consider proportionality. The report also recommended that the use of AI in the financial sector should be divided into two categories based on whether it interacts with customers. For AI systems that face customers, supervisory agencies should adjust the intensity of supervision based on the systems' impacts on consumers (for instance, chatbots have a lower impact and credit scores have a higher impact). If AI systems do not face customers, supervisory agencies should strengthen supervision of systems requiring approval (such as statutory capital adequacy assessment), while adopting moderate supervision of those without requirement (such as internal operations).

2.2 Major countries/regions published supervision principles or guidelines for AI applications in the financial industry

In December 2023, the EU passed the *Artificial Intelligence Act*,[7] dividing AI systems into four supervision levels according to risk levels. The government can supervise AI systems with specific risk levels when necessary and should appropriately retain space for technological innovation. The four supervision levels are as follows: (1) unacceptable risk level, (2) high risk level, (3) limited risk level, and (4) low risk level: not subject to mandatory regulation.

Many developed countries or regions, such as the US, the UK, Singapore, Hong Kong, France, and the Netherlands, have issued relevant principles or guidelines for AI applications in the financial sector, taking into account the recommendations of

international organizations. These guidance documents include five common principles: (1) reliability and robustness, (2) accountability, (3) transparency, (4) fairness, and (5) ethics. The first three principles are similar to the traditional model supervision concepts, allowing supervisory agencies to fine-tune based on the standards of traditional models. The principles of "fairness" and "ethics," whose concepts are to prevent AI models from producing discriminatory or biased results, may require newly-formulated standards. Moreover, "data privacy," "third-party dependency," and "operational resilience" are also key concerns in many guidance documents.

In addition, in order to establish a risk management framework for the use of GenAI in the financial industry, the Monetary Authority of Singapore (MAS) released an executive summary of the "Emerging Risks and Opportunities of GenAI for Banks" whitepaper in November 2023,[8] which is the world's first guidance document for GenAI applications in the financial industry. The whitepaper covered seven dimensions of risks while using GenAI in the financial industry: (1) fairness and bias, (2) ethics and impact, (3) accountability and governance, (4) transparency and explainability, (5) legal and regulatory, (6) monitoring and stability, and (7) cyber and data security, aiming to enable the banking industry to use GenAI in a responsible manner. The MAS will gradually apply the seven-dimension risk framework to the entire financial system in the future.

3. Taiwan's FSC also looked at related international principles and published the core principles and guidelines for AI applications in the financial industry

According to a survey conducted by the FSC in May 2023, about 36% (63 financial institutions) of the 175 financial institutions surveyed have adopted AI technology. The application fields included customer relationship management (such as intelligent customer service), risk management and legal compliance (such as suspicious transaction analysis), process optimization (such as back-office process automation), and data analysis. As for GenAI applications, most financial institutions and related units were in the evaluation stage. Only a few planned to introduce GenAI into their financial business or internal operations, but these plans have not yet been officially executed.

In order to assist financial institutions leverage the advantages of AI technology and effectively manage potential risks therefrom, the FSC, taking into account the recommendations of international organizations such as the OECD, released the *Core Principles and Policies for AI Applications in the Financial Industry* in October 2023,[9] which outlined six core principles for the use of AI in the financial industry, including: (1) establishing governance and accountability mechanisms; (2) emphasizing fairness and

human-centric values; (3) safeguarding privacy and customer rights; (4) ensuring system robustness and security; (5) emphasizing transparency and explainability; and (6) promoting sustainable development. The document also included eight supporting policies, such as formulating guidelines, adjusting regulations, and supervising the development of self-regulatory norms.

In addition, the FSC further issued a draft of the *Guidelines for AI Applications in the Financial Industry* (hereinafter referred to as the *Guidelines*) in accordance with the six core principles mentioned above in December 2023. Based on the AI life cycle[10] and the assessed risks, the *Guidelines* proposed key concerns and feasible measures so as to encourage financial institutions to introduce, use, and manage AI systems under the premise of controllable risks. The regulatory direction of the *Guidelines* is broadly the same as the FSI and IOSCO recommendations and the practices of major countries. The FSC also called for relevant associations in the financial industry to formulate new self-disciplinary rules for AI by consulting the FSC's *Guidelines*, or to incorporate the concept into their existing rules.

4. Conclusion

AI technology has great potential in improving the efficiency of financial services, promoting financial inclusion, and deepening customer relationships. However, it should be used properly, and potential risks need to be sufficiently addressed to safeguard consumer rights and financial stability. In response to the increasing influence of AI on the financial system, the FSC has taken into account the recommendations of international organizations and the practices of major countries to gradually strengthen the supervision of AI usage in the financial industry. Based on macro-prudential supervision purposes, the Bank will continue to pay close attention to the development of international supervision, and study and analyze the application and possible impact of AI in the domestic financial industry to ensure a balance between the benefits and risks of AI applications so as to maintain sound development of the financial sector.

Notes: 1. Financial exclusion refers to the phenomenon of economically disadvantaged groups being unable to access mainstream financial products and services. Take credit granting for example: groups that are underrepresented in the AI model dataset may find it difficult to obtain a favorable credit score because the model has learned that these applicants did not obtain enough loans in the past.

2. FSB (2017), "Artificial Intelligence and Machine Learning in Financial Services," November.

3. WEF (2024), "The Global Risks Report 2024," January.

4. See OECD (2019), "Recommendation of the Council on Artificial Intelligence," May. In May

2024, the OECD proposed an updated version. See OECD (2024), "Recommendation of the Council on Artificial Intelligence," May.

5. IOSCO (2021), "The use of artificial intelligence and machine learning by market intermediaries and asset managers," September.
6. Jermy Prenio and Jeffery Yong (2021), "Humans keeping AI in check – emerging regulatory expectations in the financial sector," *FSI Insights on policy implementation No. 35*, BIS, August.
7. Council of the European Union (2024), "Artificial Intelligence Act," January.
8. MAS (2023), "Emerging Risks and Opportunities of Generative AI for Banks – Executive Summary," November.
9. FSC (2023), "Core Principles and Policies for AI Applications in the Financial Industry," October.
10. The life cycle of an AI system can be divided into four stages: (1) system planning and design, (2) data collection and input, (3) model establishment and verification, and (4) system deployment and monitoring.