

Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry

1. Foreword

The increasing applications of AI¹ in financial services have benefitted financial institutions in the provision of customer service. However new risks, problems, and supervision challenges have also arisen. To help financial institutions make the most of AI technology while effectively managing risks, ensuring fairness, protecting consumer interests, maintaining system security, and achieving sustainability, the Financial Supervisory Commission (FSC), based on the Executive Yuan's "AI Taiwan Action Plan 2.0" and promotion strategies determined in the "Special Meeting for Coordination of Digital Policies and Legislation," and in reference to the relevant guiding principles issued by the financial supervisory authorities of major countries and international organizations, and taking into consideration the development status of our financial market and the directions of FSC supervisory policy, has formulated six core principles for AI applications tailored to our financial sector. These principles seek to steer financial institutions toward actively investing in technological innovation and bringing about financial service upgrade under the premise that consumer interests, the order of the financial market, and social responsibility are also taken into account.

Below are descriptions of the impacts of AI and the positions and rules of international organizations or major countries on AI, the current status of AI application in Taiwan's financial sector, the necessity of formulating AI principles and policies, six core principles for AI applications in our financial sector, and the FSC's supporting policies for the development and promotion of AI.

2.Impacts of AI and the positions and rules of international organizations or major countries on AI

2.1 Advantages of AI technology and potential issues

¹ Given the continuous advancements in AI, international organizations and governments around the world have not defined AI, but rather focus on "AI systems." This paper will use the term "AI" and "AI system" alternately. According to the definition of The Organization for Economic Cooperation and Development (OECD), "AI system" is a machine-based system that can, for a given set of objectives, generate outputs (e.g. predictions, recommendations or decisions) to influence the environments. It uses machine- or human- based data and inputs to (1) perceive real or virtual environment; (2) abstract such perceptions into models through automated analysis (e.g. use machine learning) or human analysis; and (3) use model inference to formulate options for information or action. Different AI systems are designed with different levels of autonomy after deployment.

The development of AI technology is expected to bring enormous economic and social benefits, which will be demonstrated in many different areas, including the environment, health, the public sector, finance, transportation, domestic affairs, and agriculture. As AI utilizes massive data and computing power, it is particularly useful for improving prediction, optimizing operations and resource allocation, and customizing services, thereby benefiting mankind tremendously.

However, AI systems also raise new ethical issues that may influence the value and lives of mankind. Thus AI should be used with caution, for issues arising from the use of AI may have implications regarding decision-making, labor and employment, social interactions, healthcare, education, media, access to information, the digital divide, personal data and consumer protection, the environment, democracy, rule of law, security and law enforcement, and human rights and fundamental freedoms (including freedom of expression, privacy, and non-discrimination). In addition, AI may accentuate those ethical challenges. That is because AI algorithms can replicate and amplify existing biases, thereby exacerbating existing forms of discrimination, prejudice, and stereotyping. Previously when certain problems were handled by humans, it was easier to control situations. But an AI system is capable of executing tasks rapidly. In the absence of proper design and control, it could be perilous when AI performs certain tasks for humans and unforeseen circumstances arise.

Moreover, the rapid evolution of generative AI² in recent years and its extensive impact on many fields are regarded as a major breakthrough for AI. The introduction of generative AI may help boost production efficiency and provide a wide variety of functions and services. But it may also involve personal data leaks, privacy issues, information security risks, and other legal risks. The capability of generative AI to create new content rapidly might create massive amounts of questionable or false information, thereby raising concern about the spread of false information.

In a long-term perspective, the use of AI systems, in particular generative AI that enables the creation of customized texts and images based on personal preferences, could challenge humans' special sense of experience and agency, thereby raising additional concerns about human self-understanding, social, cultural, and environmental interaction, autonomy, agency, worth, and dignity.

² According to Article 28b(4) of EU's Artificial Intelligence Act proposal, generative AI is defined as "AI systems specifically intended to generate with varying levels of autonomy, content such as complex text, images, audio or video."

2.2 Positions of international organizations and major countries on AI

In light of the rapid development of AI in recent years, major organizations are paying increasing attention to AI. Those organizations also call on members to include their recommendations in their policies, regulations, and measures. In the Recommendation on the Ethics of AI³ released by the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 2021, it identifies "respecting, protecting and promoting human rights and fundamental freedoms, and human dignity" as core values and recommends ten principles - "proportionality and do not harm," "safety and security," "fairness and non-discrimination," "sustainability," "right to privacy and data protection," "human oversight and determination," "transparency and explainability," "responsibility and accountability," "awareness and literacy," and "multi-stakeholder and adaptive governance and collaboration" for reference in every policy area.

The "AI Principles"⁴ published by The Organisation for Economic Co-operation and Development (OECD) in 2019 promote the use of AI that is "innovative and trustworthy and that respects human rights and democratic values," and recommend five values-based principles – "inclusive growth, sustainable development and well-being," "human-centric values and fairness," "transparency and explainability," "robustness, security and safety," and "accountability" for adoption by policymakers.

The G7 Digital and Tech Ministers' Meeting⁵ held in April 2023 declares an agreement on five principles for developing emerging technologies – "rule of law," "due process," "democracy," "respect for human rights," and "utilizing opportunities for innovation."

The United States passed the National Artificial Intelligence Initiative Act of 2020⁶ in 2020, and enacted it in 2021. The Act seeks to ensure continued US leadership in AI research and development, lead the world in the use of trustworthy AI systems in the public and private sectors, prepare the present and future US workforce for the integration of AI systems, and coordinate ongoing AI activities among departments of the federal government. The White House published the Blueprint for an AI Bill of Rights⁷, which outlines five principles – "safe and effective systems," "algorithmic discrimination protections," "data privacy," "notice and explanation," and "human alternatives, consideration and fallback" to

³ Please refer to <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

⁴ Please refer to <https://oecd.ai/en/ai-principles>

⁵ Please refer to <https://asia.nikkei.com/Business/Technology/G-7-ministers-agree-to-five-principles-for-assessing-AI-risks>

⁶ Please refer to <https://www.congress.gov/116/crpt/hrpt625/CRPT-116hrpt617.pdf#page=1210>

⁷ Please refer to <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>

guide the design, use, and deployment of AI automated systems to protect the American public.

The European Council passed the draft Artificial Intelligence Act⁸ in June 2023 to ensure that Europe develops and uses AI in a manner consistent with the rights and values of the EU. The Act classifies the risks of AI systems into four categories - "unacceptable risk," "high risk," "limited risk," and "minimal risk," and adopts corresponding measures of "prohibited AI practices," "regulatory control," "transparency," and "no legal obligations" based on the level of risk. For instance, AI practices that are considered to be a clear threat to humans' livelihoods and rights, such as social scoring, which present "unacceptable risk," are banned. For AI systems that might create adverse impact on humans' fundamental rights or safety, such as access to private essential services, which present "high risk," such a service provider must first undergo a conformity assessment test and register with the EU before it is put on the market (existing service providers are regulated by existing law), and must comply with a range of requirements on risk management, testing, technical robustness, data training and data governance, transparency, human oversight, and cybersecurity. AI systems presenting limited risk, such as systems that interact with humans (i.e. chatbots), emotion recognition systems, biometric categorization systems, and AI systems that generate or manipulate images and audio or video content (i.e. deepfakes) would be subject to a set of transparency regulations. AI systems presenting minimal risk, such as email filtering software, could be developed and used in the EU without conforming to any additional legal obligations. However, the EU will create codes of conduct for the reference of businesses.

Taiwan's Ministry of Science and Technology⁹ drafted the "AI Technology R&D Guidelines" in September 2019. These Guidelines identify the core values of "human-centered values," "sustainable development," and "diversity and inclusion," and take into consideration the academic freedom of AI researchers to encourage AI innovation, uphold human rights and universal values, and perfect the domestic AI R&D environment. The Guidelines set out eight principles, including "Common Good and Well-being," "Fairness and Non-discrimination," "Autonomy and Control," "Safety," "Privacy and Data Governance," "Transparency and Traceability," "Explainability," and "Accountability and Communication."

The Executive Yuan approved and published the "AI Taiwan Action Plan

⁸ Please refer to

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)

⁹ Please refer to <https://www.nstc.gov.tw/nstc/attachments/53491881-eb0d-443f-9169-1f434f7d33c7>

2.0" in April 2023, which declares the visions of "Use AI to drive industrial transformation and upgrading, use AI to help improve social well-being, and turn Taiwan a global vanguard of AI." The Action Plan 2.0 also lays out the tasks of starting with individual industries to bring about the transformation and upgrade of all industries, building a trustworthy AI environment that puts equal emphasis on technological innovation and risk governance, being prepared for social impacts brought about by AI, developing an inclusive digital economy with AI, and addressing major challenges faced by society to enhance the well-being of all humans.

2.3 Supervisory recommendations of international organizations and major countries on AI applications in the financial sector

AI has come into increasing use in financial services in recent years. While the implementation of AI has created benefits for the financial sector, including better service efficiency, lower cost of services, and an enhanced customer experience, it has also created new risks and supervisory challenges regarding such possibilities as personal data leaks, privacy breaches, ethical concerns, and workforce transformation. Thus, the question of how to make sure financial institutions can effectively manage risks, ensure fairness, protect consumer interests, maintain system security, and achieve sustainable development while using AI is an issue that financial supervisory authorities and financial institutions must address.

The Financial Stability Board (FSB) published the "Artificial intelligence and machine learning in financial services—Market developments and financial stability implications"¹⁰ in 2017, which relates that AI and machine learning (ML) can help financial institutions process information and data more efficiently, and the adoption of regulatory technology (RegTech) and supervisory technology (SupTech) can help improve regulatory compliance and increase supervisory effectiveness. The FSB also points out the following risks associated with AI and ML:

- (1) Network effects and scalability of new technologies may in the future give rise to third-party dependencies. This could in turn lead to the emergence of new systemically important players that could fall outside the regulatory perimeter.
- (2) Applications of AI and ML could result in new and unexpected forms of interconnectedness between financial markets and institutions, because of the use by various institutions of previously unrelated data sources.
- (3) The lack of interpretability or auditability of AI and ML methods

¹⁰ Please refer to <https://www.fsb.org/wp-content/uploads/P011117.pdf>

could become a macro-level risk. Similarly, a widespread use of opaque models may result in unintended consequences.

The FSB advises that it is important to assess uses of AI and ML in view of their risks, including adherence to relevant protocols on data privacy, conduct risks, and cybersecurity. Adequate testing and training of tools with unbiased data and feedback mechanisms is important to ensure that applications can function as expected.

The Financial Stability Institute (FSI) under the Bank for International Settlements (BIS) released "Humans Keeping AI in check—emerging regulatory expectations in the financial sector"¹¹ in 2021. The paper classifies AI systems used by financial institutions into two categories – customer-facing and non customer-facing. In the customer-facing category, AI systems are further divided into two types based on their impact on customers – low impact (e.g. chatbots) and high impact (e.g. credit scoring); in the non customer-facing category, AI systems are also further divided into two types, depending on whether supervisory approval is required - supervisory approval not required (e.g. internal operating processes) and supervisory approval required (e.g. regulatory capital adequacy assessment). It is hoped that financial supervisory authorities will set forth supervisory measures based on four principles – "transparency," "reliability and soundness," "accountability," and "fairness and ethics," and at the same time, address possible challenges through proportionality.

The International Organization of Securities Commissions (IOSCO) released "The Use of Artificial Intelligence and Machine Learning by Market Intermediaries and Asset Managers"¹² in 2021, which proposes six measures that financial supervisory authorities can adopt, including requiring financial institutions to: (1) have appropriate governance, controls and oversight frameworks; (2) monitor the development, testing, use and performance monitoring of AI and ML on a continuous basis; (3) make sure that staff have adequate knowledge, skills and experience to implement, oversee, and challenge the outcomes of AI and ML; (4) understand their reliance on the third-party providers of AI and ML services, and establish sound management and oversight mechanisms; (5) provide proper transparency and disclose information to investors, regulators and stakeholders; and (6) have appropriate controls in place to ensure that the data and the performance of the AI and ML can minimize bias.

¹¹ Please refer to <https://www.bis.org/fsi/publ/insights35.pdf>

¹² Please refer to <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf>

The financial supervisory authorities of the United States¹³, Singapore¹⁴, South Korea¹⁵, and many other countries have set out relevant guidelines and principles, which mostly emphasize "fairness and ethics," "transparency," "accountability," "consumer interests and privacy protection," and "security and robustness."

3.Current status of AI adoption in Taiwan's financial sector, the necessity of formulating AI principles and policies

3.1 AI or generative AI applications in Taiwan's financial sector

According to a survey of 175 financial institutions in Taiwan conducted by the FSC in May 2023, 63 of them (36%) have adopted AI technology for use in (1) customer relationship management, such as smart customer service and robo-advisor services; (2) risk management and compliance, including money laundering prevention, analysis of suspicious transactions, and account opening review; (3) process improvement, including optical character recognition (OCR) and backend process automation; (4) data analysis, such as customer attributes and consumption behavior analysis, market trend analysis, etc.; and (5) others, such as using threat intelligence to analyze security scenarios.

In addition, to understand the use of ChatGPT or other generative AI by Taiwan's financial institutions, the FSC completed a questionnaire survey in April 2023. The survey finds that currently no financial institutions use generative AI in financial services or internal operations. However two institutions are planning the implementation of AI, while 80 financial institutions and self-regulatory organizations are evaluating the introduction of AI.

Another 95 financial institutions and self-regulatory organizations have implemented control over the use of generative AI by employees, including the setup of firewalls to block intranet connection to ChatGPT and putting in place internal rules to prevent the leakage of customer data.

In the future, the FSC will continue to communicate with financial institutions regarding the use of generative AI, and evaluate and adjust relevant rules in view of industry's development status and applications to facilitate compliance by financial institutions.

3.2 FSC legislation for implementation of emerging technologies by

¹³ Please refer to https://content.naic.org/sites/default/files/inline-files/AI%20principles%20as%20Adopted%20by%20the%20TF_0807.pdf

¹⁴ Please refer to https://www.mas.gov.sg/-/media/mas/resource/news_room/press_releases/2018/annex-a-summary-of-the-feat-principles.pdf

¹⁵ Please refer to <https://www.fsc.go.kr/eng/pr010101/76209>

financial institutions

3.2.1 Banking industry:

- (1) In response to the development of new technologies, the FSC has approved the amended Operating Rules for the Use of Emerging Technologies by Financial Institutions in April 2020 to help banks manage risks associated with new technologies and promote the sound operation of banking business, including the security management of cloud services, social media management procedures, security management of BYOD (Bring Your Own Device), and security management of biometric data.
- (2) With regard to the prevention of AI deepfakes, the FSC has instructed the Bankers Association in its Standards for the Security Management of Electronic Banking Business of Financial Institutions to require that when a bank uses video conference in its electronic banking business, it shall make sure it is the customer, a real person, who appears in the video to prevent fake identity through the use of pre-recorded film, masking, simulated imaging, or deepfake, and retain an authentication record and transaction trail to facilitate verification. The FSC has also instructed the Bankers Association to include AI use in self-regulatory rules for banks.

3.2.2 Securities and futures industry:

- (1) Preventing the use of AI deepfake technology in the creation of fake identities: The FSC in September 2022 approved amended information security self-regulatory rules for emerging technologies drafted by the Securities Association and Futures Association that includes the prevention of deepfakes. The amended rules require that when a securities/futures enterprise uses video imaging for identity verification, it shall also use a one-time password (OTP), a phone call, or an in-person visit and check the photograph on identity documents for enhanced verification. Securities/futures enterprises are also advised to carry out education and training on deepfake awareness and information security to shore up their risk management of emerging technologies. The existing rules will be reviewed in a timely manner in view of AI development.
- (2) Enhancing the supervision of automated investment services to safeguard the interests of investors: The current regulations require that if an investment consulting firm offers automated investment services, it should put in place effective supervision and management with regard to whether the algorithmic design is logical and whether

the computing results are accurate, conduct preliminary evaluation and periodic review, and assign personnel to monitor the system operation to safeguard the interests of investors. The regulations also require investment consulting firms or their group enterprises to set up a committee in charge of supervising and managing the development and modification of algorithms, or participating in the review and onsite investigation of external software developers so as to evaluate the adequacy of system design and ensure that the firm has constructed comprehensive prevention, detection, and handling measures for network security.

3.2.3 Insurance industry:

In response to the development of Fintech, the FSC has been promoting the application of new technologies (e.g. AI, blockchain, biometrics) in insurance products and services and has drafted related supervisory regulations, such as the Directions for Insurance Enterprises Engaging in Electronic Commerce Business, Directions for Insurance Companies Providing Distance Insurance Underwriting and Services, and compliance matters and self-regulatory rules for all kinds of businesses, which require insurance enterprises to implement information security (including obtaining ISO 27001 certification for their Information Security Management System and Personal Information Management System), money laundering control, personal data protection, protection of customer interests, and including those activities in their internal control systems to facilitate compliance by insurance enterprises.

3.3 The necessity for the FSC to draft AI principles and guidance

From a tech-neutral standpoint, financial institutions still need to comply with the prevailing financial regulations and market self-regulatory rules when using AI in their operations. However issues such as ethics, fairness, privacy, and transparency may arise in the application of AI technology. To ensure that associated risks are fully addressed in the prevailing regulations, it is necessary to draft a set of principles and policies in conformance with the recommendations of international organizations so as to ensure the use of AI will bring only positive benefits to society, the economy, and consumers.

In reference to the supervisory principles and guidelines of international organizations and major countries for AI and ML, and in consideration of AI practices in finance, possible issues arising therefrom, and the development status of our financial market, the FSC has drafted six core principles for use of AI in the financial sector. It is hoped that those six principles can guide financial institutions in the innovation and practice of

AI applications while ensuring that AI can be beneficial to maintaining the stability of financial systems and safeguarding the interests of consumers. The FSC will subsequently draft guidelines based on those six core principles for the reference of and compliance by financial institutions. In the future, the FSC will carry out rolling evaluations and adjustments in view of technological developments and the use status of AI systems in our financial market to keep abreast of the times and meet practical development needs.

4. Six core principles for AI applications in the financial sector

In reference to the guiding principles developed by international organizations or the supervisory authorities of major countries for the use of AI in financial services, and taking into account the supervisory principles of "responsible innovation," "enhancing compliance," "treating customers fairly," "financial inclusion," "information security," "information disclosure," "sustainable finance," and "caring for employees," the FSC proposes the following six core principles for use of AI in the financial sector:

4.1 Core Principle 1: Establishing governance and accountability mechanisms (corresponding to the supervisory principle of "responsible innovation")

- (1) Financial institutions should bear internal and external responsibilities corresponding to the AI systems they use. Internal responsibilities include assigning a senior executive to be responsible for related oversight and management, and establishing an internal governance framework; external responsibilities involve responsibilities to consumers and the society, including protecting consumer privacy and data security.
- (2) Financial institutions should establish a comprehensive and effective AI risk management mechanism, integrate it into prevailing risk management and internal control operations or processes, and conduct periodic evaluation and testing.
- (3) Financial institutions should ensure that their employees have adequate knowledge and skills to work with AI, and carry out risk-based decision making and supervision.

Description:

The FSC has adopted the strategy of promoting "responsible innovation" to encourage the use of technology to develop innovative financial products or services. On the other hand, the FSC emphasizes the

importance of financial market order and consumer protection to make sure financial innovation poses controllable risks and will have a positive effect on financial market stability. Against this backdrop, the FSC has proposed the principle of "establishing governance and accountability mechanisms," stressing that financial institutions should be responsible for internal governance and consumer protection when applying AI systems, and should oversee the risk management and use of AI systems. This is to ensure that while pursuing innovation, financial institutions should also fulfill their commitment to social responsibility and strive to build a stable and fair financial market environment.

The principle of "establishing governance and accountability mechanisms" is of utmost importance to the use of AI systems by financial institutions. The spirit of this principle lies in the conviction that financial institutions should be responsible, both internally and externally, for the use of AI systems. As such, financial institutions must establish a sound internal governance framework in documented form, assign a senior officer to take charge of related supervision and management, explicitly define the duties of respective lines of business, and ensure the retention of records. These measures help ensure the normal operation of AI systems and discover and resolve potential problems in a timely manner.

The responsibilities of financial institutions with regard to the use of AI systems also cover consumers and social responsibilities. That is, the use of AI systems must fully respect and protect the privacy of consumers and information security. This is not only basic protection of consumer interests, but also paramount to maintaining the business reputation and stable operations of the financial institution.

Financial institutions must establish an effective AI-related risk management mechanism using a risk-based approach, and integrate it into their overall risk management and internal control operations and processes. In addition, after its AI system goes online, a financial institution should conduct regular evaluation and testing, including documenting and monitoring situations unforeseen at the time of system development, and if necessary, should modify the system to make sure the AI system is secure and effective.

Lastly, this principle also emphasizes the importance of cultivating and enhancing employees' AI knowledge and risk identification and management skills, and taking a risk-based approach to decision making and supervision. As AI technology advances, the knowledge and skills of employees (including those engaged in development, testing, supervision, compliance, risk management, and internal auditing) will have a direct bearing on whether the financial institution is able to make proper

decisions and effectively supervise the operations of AI systems. Thus financial institutions should conduct education and training on an ongoing basis to make sure employees are capable of adapting to the rapid growth and evolution of AI technology and appropriately addressing related risks and challenges.

4.2 Core Principle 2: Emphasizing fairness and human-centered values (corresponding to the supervisory principle of "treating customers fairly and improving financial inclusion")

- (1) Financial institutions should, in the process of utilizing AI systems, try their best to avoid unfairness resulting from algorithmic bias.
- (2) The application of AI systems should be human-centric and controllable by humans, and should respect the rule of law and democratic values.
- (3) Risks associated with information produced by generative AI must be objectively and professionally controlled by financial institution personnel¹⁶.

Description:

The principle of "emphasizing fairness and human-centered values" is vital during the process of AI system implementation. Financial institutions that introduce AI systems should recognize fully and try their best to avoid potential algorithmic bias. Especially when providing customer-facing AI services, financial institutions must ensure the appropriateness of data sources and the quality of data, and conduct testing and validate the algorithm in an independent environment prior to formal launch to avoid unintended consequences. Financial institutions should also strive to provide all customers with fair and non-discriminatory financial services to achieve financial inclusion. In addition, the data, database, and models of an AI system should be subject to regular review and accuracy verification to reduce bias.

In addition, human-centered and human-controllable values and applications are facets that must be taken into consideration when financial institutions use AI systems. Aside from respecting consumer privacy during the use of AI systems, financial institutions must make sure universal values such as the rule of law and democracy are also respected, and must put into place mechanisms to ensure that AI system evolve in such a way as to maintain the initial intent of system creation, that is, to assist humans without causing harm to humans while ensuring human

¹⁶ In reference to Point 2 of the draft of "Reference Guidelines for the Use of Generative AI by Executive Yuan and Subordinate Agencies (Institutions)."

autonomy and control. Risks associated with information created by generative AI must be subject to the objective and professional control of financial institution personnel.

By stressing the principle of "emphasizing fairness and human-centered values," the FSC urges financial institutions to pay attention to the values of diversity and inclusion when first designing AI systems, provide fair and inclusive financial services during the use of AI systems, and conform to social values and public expectations consistently throughout the entire process.

4.3 Core Principle 3: Protecting privacy and customer interests (corresponding to the supervisory principle of "protecting financial consumers")

- (1) Financial institutions should fully respect and protect the interests of consumers, and manage and use customer data properly.
- (2) Financial institutions applying AI systems to provide financial services should respect the customer's right to choose, and should remind the customer of available alternatives.

Description:

"Protecting privacy and customer interests" is another important element financial institutions must consider when applying AI systems. With the development of big data and AI technology, the personal data of customers are often massively collected and used for the training of AI to increase its accuracy. But this practice may impact the privacy of customers, negatively affect public trust toward financial institutions, and lessen the degree of service satisfaction.

When using customer data, financial institutions must fully respect and protect customer privacy and properly manage and utilize related data to avoid any risk of data leakage. For example, in the absence of a proper control mechanism (an example of a proper control mechanism is a generative AI model deployed in a closed, on-premises environment in a way that ensures the security of the system environment), a financial institution should not input data without customer consent for disclosure into generative AI. Financial institutions should also respect the right of customers to choose AI services or not, and remind customers whether there are alternatives available. This is a way to protect customers' right to choose, and is also a way to safeguard customer interests.

By implementing the principle of "protecting privacy and customer interests," financial institutions are able to meet regulatory requirements,

enhance consumer confidence and satisfaction, and promote sound business development.

4.4 Core Principle 4: Ensuring system robustness and security (corresponding to the supervisory principle of "enhancing information security")

- (1) When applying AI systems, financial institutions must ensure system robustness, security, and safety to avoid causing harm to consumers or the financial system.
- (2) Financial institutions that outsource the development or operation of AI systems for financial services should conduct appropriate risk management and oversight of the third-party providers.

Description:

Digital technology is evolving rapidly, and AI has become an important tool for the financial industry to promote innovation and improve service quality. However, risks associated with the spread and use of AI are real. In particular, due to the special nature of the financial services sector, any system error or data leak could cause immense harm to consumers or even the entire financial system. Those risks could arise from technological malfunctions, malicious attacks, or wrong decisions made by an AI system.

Therefore, it is of utmost importance for financial institutions to "ensure system robustness, security, and safety" in the process of applying AI systems. A financial institution must have adequate ability to develop and maintain secure AI systems, and the ability to conduct ongoing monitoring of the operational results of AI systems, and take appropriate control measures where necessary.

To successfully "ensure system robustness, security and safety," risk management and oversight are indispensable when a financial institution provides financial services applying AI systems developed or operated by third-party service providers. Prior to engaging the service of such a third-party service provider, a financial institution must perform due diligence to determine whether the service provider possesses related knowledge, expertise, and experience, and must fully understand and evaluate the operating modes of its AI system and potential risks (including the handling of data privacy, operating risks, information security, and concentration risk), and based on the evaluation results, must implement a proper oversight strategy and management practices to prevent possible risks or problems. The financial institution should also clearly define the responsibilities of the service provider and plan ahead-of-time solutions for problems that may arise in case of adverse events. In addition, financial

institutions should have AI related algorithmic rules and save the trails and records to facilitate subsequent verification and management.

Based on the principle of "ensuring system robustness, security and safety," the FSC hopes that while maintaining financial stability and protecting consumer interests, financial institutions are able to use AI systems in an innovative manner, with confidence and robustness, so as to provide consumers with better financial services.

4.5 Core Principle 5: Ensuring transparency and explainability (corresponding to the supervisory principle of "information disclosure")

- (1) Financial institutions should ensure the transparency and explainability of their operations.
- (2) Financial institutions should make proper disclosure when their AI system interacts directly with consumers¹⁷.

Description:

In the field of artificial intelligence, the importance of "ensuring transparency and explainability" is widely acknowledged. It denotes that how an AI system works should be comprehensible and explainable. This principle is all the more imperative when applying AI in financial services.

For financial institutions, explainable AI systems enable them to effectively assess and manage the risks of implementing AI and determine the effectiveness of AI systems. Financial institutions must understand how AI makes decisions so as to detect potential problems and make necessary adjustments to ensure effective management of AI operations.

For stakeholders, the principle of "ensuring transparency and explainability" implies that financial institutions should evaluate how to make proper disclosure to stakeholders. For instance, a financial institution should make proper disclosure when using AI to interact directly with consumers; as for AI systems relating to money laundering prevention, information security, or fraud prevention, or involving trade secrets, risk may arise from over-disclosure of information. In such event, a financial institution should be cautious about the contents of disclosure or opt not to disclose. Financial institutions should also make sure the degree of explainability matches the importance of an AI system. In particular, humans who are adversely affected by an AI system should be able to understand or challenge the results of the AI system based on plain and

¹⁷ In reference to Point 6 of the draft of "Reference Guidelines for the Use of Generative AI by Executive Yuan and Subordinate Agencies (Institutions)."

easy-to-understand information on the factors, and the logic that serves as the basis for predictions, recommendations, or decisions.

4.6 Core Principle 6: Promoting sustainable development (corresponding to the supervisory principle of " Sustainable finance and employee care")

- (1) Financial institutions should, when applying AI systems, ensure that their AI development strategies and implementation are commensurate with the principle of sustainable development, including reducing economic and social inequality, and protecting the natural environment, thereby promoting inclusive growth, sustainable development, and social well-being.
- (2) Financial institutions should, in the process of applying AI systems, provide employees with proper education and training that could help them adapt to changes brought about by AI, and make efforts to protect employees' work rights.

Description:

The purpose of formulating the principle of "promoting sustainable development" is to remind financial institutions that AI is used not only to improve efficiency and profitability, but also to promote the inclusive growth and sustainable development of the entire society.

Financial institutions should ensure that their AI strategy and implementation methodology match the principle of sustainable development. This means that, while offering innovative financial services, a financial institution should also take into consideration how to increase the financial participation of disadvantaged groups, how to reduce economic, social, gender, and age inequalities, and how to promote environmental protection through financial technology. In the aspect of sustainable development, financial institutions should be aware that massive amounts of energy used by AI systems could produce a negative impact on the environment and ecosystems.

In addition, financial institutions must respect and protect the work rights of regular employees, including by providing proper education and training to employees during the course of AI transformation to help them adapt to the new working environment. This helps safeguard employees' work rights and cultivates professionals with the skills to operate an AI system that will enhance the institution's competitive edge.

Overall, this principle emphasizes that while using AI to pursue economic benefits, a financial institution should give equal emphasis to the

ideals of sustainability, caring for the disadvantaged, reducing inequality, and protecting worker rights to promote the well-being and progress of the society as a whole.

5. FSC's supporting policies for promoting AI development

- (1) On the basis of the aforementioned principles, the FSC will publish the "Guidelines for AI Applications in the Financial Industry". Taking into account factors such as the extent to which financial institutions use AI systems, institutions' reliance on data, AI functions, autonomous behaviors, and whether AI systems are customer-facing or non-customer-facing, the guidelines will set out matters to comply with based on the principle of proportionality, using a risk-based approach. The guidelines will seek to bring about the sound development of the financial sector, maintain financial stability, and safeguard consumer interests.
- (2) Continue to examine relevant FSC regulation and make regulatory adjustments in a timely manner to build a sound regulatory environment for the use of AI systems in the financial sector.
- (3) Use AI technology to develop SupTech to enhance the efficiency and effectiveness of financial supervision.
- (4) Engage in exchanges and collaborations with international organizations and foreign financial supervisory agencies to ensure that our supervisory policies are consistent with mainstream international trends.
- (5) Continue to encourage financial institutions to actively participate in AI research, development and application to provide customers with better financial services or develop RegTech. The FSC will also hold seminars and workshops to assist financial institutions in implementing best practices.
- (6) Continue to grasp and inspect the actual status of AI use in the financial sector, and where necessary, conduct special financial examinations to make sure financial institutions comply with applicable regulations and practice risk management in the use of AI systems, and use AI to increase public trust and social well-being.
- (7) Instruct financial industry associations to draft self-regulatory standards and best practices regarding the use of AI systems by financial institutions to enhance their information security, internal control, and fair customer treatment, and make sure financial institutions continue to assist their employees in AI transformation

and pay attention to worker rights in the course of AI implementation.

- (8) Continue to oversee financial institutions' compliance with fair customer treatment and friendly financial guidelines, and conduct financial literacy programs to enhance public awareness of the use of digital financial tools, thereby reducing the digital divide and ensuring the fair transformation of digital finance.

6. Conclusion

As AI advances rapidly and its applications continue to expand, trustworthy AI will play an increasingly important role in the field of finance. AI offers wide-ranging potential, in matters ranging from improving the efficiency of financial services to deepening customer relations. However, to make the most of this new technology, one must seek a balance between innovation and responsibility in the use of AI systems. One must ensure fairness and transparency, and be fully prepared for possible risks and challenges.

The introduction of principles and policies for the use of AI in the financial sector aims to guide financial institutions to create value when applying AI systems, and moreover, to protect consumer interests, maintain financial stability, and realize inclusive sustainability. The FSC hopes that through the implementation of related principles and policies, financial institutions can maximize their functions and bring greater benefits to customers and society in the upcoming era of AI.

Attachment: Six Core Principles for AI applications in the Financial Sector

1.Establishing governance and accountability mechanisms

- (1)Financial institutions should bear internal and external responsibilities corresponding to the AI systems they use. Internal responsibilities include assigning a senior executive to be responsible for related oversight and management, and establishing an internal governance framework; external responsibilities involve responsibilities to consumers and the society, including protecting consumer privacy and data security.
- (2)Financial institutions should establish a comprehensive and effective AI risk management mechanism, integrate it into prevailing risk management and internal control operations or processes, and conduct periodic evaluation and testing.
- (3)Financial institutions should ensure that their employees have adequate knowledge and skills to work with AI, and carry out risk-based decision making and supervision.

2.Emphasizing fairness and human-centered values

- (1)Financial institutions should, in the process of utilizing AI systems, try their best to avoid unfairness resulting from algorithmic bias.
- (2)The application of AI systems should be human-centric and controllable by humans, and should respect the rule of law and democratic values.
- (3)Risks associated with information produced by generative AI must be objectively and professionally controlled by financial institution personnel.

3.Protecting privacy and customer interests

- (1)Financial institutions should fully respect and protect the interests of consumers, and manage and use customer data properly.
- (2)Financial institutions applying AI systems to provide financial services should respect the customer's right to choose, and should remind the customer of available alternatives.

4.Ensuring system robustness and security

- (1)When applying AI systems, financial institutions must ensure system robustness, security, and safety to avoid causing harm to consumers or the financial system.

- (2) Financial institutions that outsource the development or operation of AI systems for financial services should conduct appropriate risk management and oversight of the third-party providers.

5. Ensuring transparency and explainability

- (1) Financial institutions should ensure the transparency and explainability of their operations.
- (2) Financial institutions should make proper disclosure when their AI system interacts directly with consumers.

6. Promoting sustainable development

- (1) Financial institutions should, when applying AI systems, ensure that their AI development strategies and implementation are commensurate with the principle of sustainable development, including reducing economic and social inequality, and protecting natural environment, thereby promoting inclusive growth, sustainable development and social well-being.
- (2) Financial institutions should, in the process of applying AI systems, provide employees with proper education and training that could help them adapt to changes brought about by AI, and make efforts to protect employees' work rights.