Guidelines for Artificial Intelligence (AI) Applications in the Financial Industry

Financial Supervisory Commission

June 2024

Table of Contents

Preface

The use of AI has become more commonplace in the financial market. Although AI has the ability to increase the efficiency of the financial sector, reduce costs, enhance the customer experience, manage risks, promote compliance, prevent financial crimes, prevent information security incidents, and promote sustainable development, the introduction of AI without careful planning, review, or failure to calibrate the use of AI in accordance with technological developments or its actual effectiveness may be detrimental to the original purpose of the introduction and result in losses to financial consumers or the financial industry, heighten risks in the industry, and reduce public confidence in the financial market. These Guidelines are established in accordance with the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" to help the financial sector identify and pay attention to key points of concern in the life cycle of AI systems. The Guidelines also referenced regulations or manuals of the Bank for International Settlements (BIS), International Organization of Securities Commissions (IOSCO), European Union, Singapore, and the United States to provide the financial sector with guidelines for the application of AI.

These Guidelines include the General Provisions and six main chapters. The General Provisions explain common issues such as AI-related definitions, the AI life cycle, risk assessment factors, risk-based implementation of core principles, and supervision and management of third-party operators. Chapters 1 to 6 describe the key concerns and measures that the financial industry can take when implementing AI Core Principles 1 to 6 based on the AI life cycle and the assessed risks.

These guidelines are administrative and non-binding in nature and are formulated to encourage the financial industry to introduce, use, and manage AI while controlling risks. The examples cited in this document are provided for reference and financial institutions are encouraged to adopt them based on their respective conditions. Due to the high degree of correlation between the core principles, the financial industry is advised to interactively assess the feasibility of each key point or measure when introducing and using AI systems when referencing these Guidelines. They should avoid focusing on a single core principle and failing to comprehensively control risks. In addition, there are many ways to achieve the goal of "adequate AI risk management." These Guidelines implement the core principles in a risk-based manner, while making reference to best

practices in other countries and in Taiwan to provide internal rules[1]. Depending on the risks associated with the use of AI systems, financial institutions may choose risk mitigation mechanisms and implementation methods in accordance with the core principles, including adopting more cost-effective methods to achieve the objectives.

The financial institutions specified in the chapters in these Guidelines refer to financial holding companies, banks, credit cooperatives, bills finance companies, credit card companies, trust enterprises, electronic payment institutions, post offices engaging in postal remittances or simple life insurance business, securities firms, securities investment trust enterprises, securities finance enterprises, securities investment consulting enterprises, futures commission merchants, leverage transaction merchants, futures trust enterprises, managed futures enterprise, futures consultant companies, insurance companies, insurance cooperatives, insurance agents, insurance brokers, and insurance surveyors. If relevant financial industry associations have set up self-regulatory guidelines on the application of AI, they may reference these Guidelines and incorporate relevant key points and measures. If they have not set up relevant self-regulatory guidelines, financial institutions are advised to reference these Guidelines for the introduction, use, and management of AI systems. When financial institutions apply AI systems for financial innovation businesses, the FSC encourages financial institutions to conduct tests through mechanisms such as FinTech Innovation Experimentation or financial business trials, if deemed necessary.

[1] If branch offices of foreign groups in Taiwan are provided with AI applications by the group, they may be processed in accordance with group regulations.

General Provisions Common Issues

I. Definitions of Artificial Intelligence (AI)[2]

(I) Definitions of AI systems: They refer to systems that simulate human learning, thinking, and response patterns by learning vast amounts of data and using machine learning or related modeling algorithms for perception, prediction, decision-making, planning, reasoning, and communication.

(II) Definition of generative AI: Generative AI refers to AI systems that can generate content with simulated human intelligence and creativity in the forms that include but are not limited to text, images, audio, video, and coding.

II. AI System Life Cycle

The life cycle of an AI system consists of the four stages set forth below:

(I) System planning and design: Set clear system objectives and requirements.

(II) Data collection and input: Data collection, processing, and input into the database.

(III) Model building and validation: Select and build model algorithms and training models, and validate the models to ensure model performance, security, and confidentiality.

(IV) System deployment and monitoring: Apply the system in real-world environments, verify the integrity of the model, and implement continuous monitoring to verify the potential impact of the system.

AI systems applied by financial institutions may be self-developed[3] for use and therefore include the four stages described above. Financial institutions may also commission a third party to develop or purchase an AI system and then deploy the system and implement monitoring. Therefore, not all financial institutions will go through the four stages described above. When applying AI systems, financial institutions are advised to identify the extent to which they can monitor and control the

[2] The "Operating Regulations for the Use of Artificial Intelligence Technologies by Financial Institutions" of the Bankers Association are used as the reference for the definitions of artificial intelligence.

[3] Self-developed systems include those that are developed entirely by financial institutions and AI systems that are cooperatively or jointly developed with other financial institutions, as well as pre-existing AI systems that are independently trained or fine-tuned using existing open-source AI models.

risks in the four stages, and to define the division of responsibilities for risk monitoring and control with vendors through contracts or other means for parts or matters over which they have less control. For the purpose of simplifying text, these Guidelines adopted "introduce" AI to represent stages (I), (II) and (III), and "use" AI to represent stage (IV). The concept of "applying" AI in these Guidelines is an overall framework that includes the four stages described above.

III. Factors in Risk Assessment

When financial institutions apply AI systems, they are advised to evaluate the risks associated with individual usage scenarios. They are also advised to allocate more resources to high-risk AI systems to ensure effective risk management. The factors to be considered in risk assessments are as follows: (The following are examples that help illustrate the risk assessment scenarios and they are not intended to be used for regulating the risk level of the relevant usage scenarios. The level of risk involved in using AI systems must still be determined by the financial institutions after considering all factors in the risk assessment.)

(I) Whether it is used to provide direct customer service or has a significant impact on operations

1. AI systems that provide customer service (customer-facing): AI systems with AI decision-making results that have a significant impact on customers' rights and interests or operations usually incur higher risks, and examples include those used in credit scoring and robotic financial management systems. AI systems with AI decision-making results that only enhance the quality of customer service may incur lower risks, and examples include intelligent customer service systems.

2. AI systems used for internal operations (not customer-facing): AI systems with AI decision-making results that involve supervisory regulations usually incur higher risks, and examples include those used in legal capital adequacy assessment and money laundering prevention systems. AI systems with AI decision-making results that do not involve supervisory regulations may incur lower risks, and examples include systems used to improve the efficiency of internal administrative operations.

(II) Extent of use of personal data: A higher extent of the use of raw personal data[4] or sensitive personal information in an AI system may

[4] Raw personal data refers to personal data that has not been de-identified, processed by privacy-enhancing technologies, or processed through other means.

incur higher risks.

(III) Extent of AI autonomy: AI systems that replace human decision-making to a higher extent or have a higher extent of automated learning may have a higher risk of an increase in unanticipated negative systemic impacts or decreased chance of immediate human intervention.

(IV) Complexity of the AI system: AI systems with a more complex computational model or a higher number and type of parameters used may reduce explainability and incur higher risks.

(V) Extent and reach of impact on different stakeholders: AI system decision-making results may incur higher risks if they have a more profound impact on internal and external stakeholders, or if they affect more categories and a higher number of stakeholders.

(VI) Integrity of relief[5] options: AI system decision-making results that do not provide stakeholder relief options or provide incomplete relief options may incur higher risks.

IV. Risk-Based Implementation of Core Principles

Financial institutions are advised to determine the risk control measures to be adopted and the extent of such measures based on the results of the risk assessment of AI systems, and ensure that they are in line with their current practices. For AI systems with higher risks, in addition to paying attention to the key points and measures listed in Chapters 1 to 6 during the introduction and use of the system, financial institutions must also evaluate whether to adopt the following measures:

(I) Records: The application of high-risk systems should require more comprehensive written or digital records.

(II) Monitoring and control mechanisms: The application of high-risk systems should require the establishment of a higher frequency of monitoring and control and a wider range of monitoring and control mechanisms.

(III) Review and approval: The application of high-risk systems should require a more stringent review and approval process and an elevated decision-making level.

(IV) Audit or evaluation mechanism: After evaluating AI system risks, internal resources, and requirements for professional services, a third-

[5] Remedies may include channels for complaints or remedies and dispute resolution mechanisms.

party audit or evaluation unit may conduct independent verification if necessary. If the introduced AI system is developed or managed by the same group (including its affiliates), relevant audits may be replaced by information provided by the group.

V. Supervision and Management of Third-Party Operators

When financial institutions commission a third-party operator to introduce AI system-related operations, it is advisable to take the following supervisory and management measures:

(I) Financial institutions should first conduct a review to assess whether the third-party operator has the relevant knowledge, expertise, and experience, and determine the concentration risk associated with the commission of the third-party operator (the concentration risk of the financial institution's own commission of the third-party operator). They must then adopt appropriate supervisory strategies and management actions based on the results of the review to prevent any possible risks or issues.

(II) Financial institutions are advised to sign written contracts with third-party operators to specify the scope of the introduction, the third-party operator's scope of responsibilities, and recourse for its failure to meet performance targets or the occurrence of adverse events.

(III) When a financial institution commissions a third-party operator to introduce relevant operations, if it involves the transmission of customer data to a third-party operator for processing, it is advisable to sign an agreement with the third-party operator containing the terms for data protection, which clearly require encrypted data transmission, secure storage, and data disposal upon termination of services.

(IV) When financial institutions commission a third-party operator to introduce an AI system, or when financial institutions conduct testing and monitoring operations, they must pay attention to the third-party operator's outsourcing agreements and they should clarify the division of responsibilities. They must also establish an appropriate data or system migration mechanism in case of the termination of the outsourced operation.

(V) Financial institutions shall require third-party operators to keep written or digital operation records of the execution of the commissioned tasks to facilitate subsequent tracking, verification, and management.

(VI) If a third-party provider is commissioned by a financial institution to

carry out affairs involving the outsourcing of financial industry operations, it should comply with the relevant outsourcing regulations of each industry.

Chapter 1 Establishing Governance and Accountability Mechanisms

Core Principle 1: Establishing governance and accountability mechanisms

(I) 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.

(II) 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.

(III) 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.

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

Financial institutions may apply multiple AI systems. Therefore, the Guidelines advise financial institutions to set up a clear framework and risk management policy for managing AI systems, understand the purpose of the AI system, the applicable business or operation, and the responsible personnel, and ensure that they are able to explain the system's logic internally, explain the overall policy externally, and explain matters that consumers need to know about individual AI systems. They must also set comprehensive procedures for dealing with errors or unanticipated events. Financial institutions are also advised to continue to enhance their employees' understanding of and ability to introduce, use, and manage AI systems in order to adapt to the rapid development and changes in AI technologies.

II. Main Concepts

(I) Internal and external responsibilities of financial institutions for the application of AI systems

1. Internal responsibility refers to the definitions of the authority and responsibility of each unit within the organization, including a clear internal governance framework, supervision and management by senior executives or designated committees that can oversee cross-departmental duties, definitions of the functions and responsibilities of each department or line of business, and implementation of hierarchical management mechanisms.
2. External responsibility refers to the ability of the organization to communicate externally about the organization's actions, including channels or communication mechanisms that allow external parties to inquire about or review relevant information on matters affected by decisions, and to ensure that the use of the AI system is consistent with the planned objectives.

(II) When implementing the principles of governance and accountability, financial institutions are advised to document or digitize relevant mechanisms and operations and establish appropriate oversight mechanisms to the best of their abilities.

(III) Financial institutions should implement all core principles of AI for the financial industry and should not treat any of the principles as a one-time or independent task.

III. Organizational Structure and Accountability Mechanisms

(I) Organizational structure and roles and responsibilities for AI systems: Financial institutions are advised to establish an organizational structure for the application of AI systems in their organization, including whether to designate a department or team responsible for the entire AI system. Each department or team as well as business lines that apply the AI system and units assigned for operations and activities in each phase of the AI life cycle are advised to clearly define their responsibilities as well the roles, functions, and responsibilities of their personnel.

(II) Designate senior executives or committees for supervision and coordination: Financial institutions may designate a senior executive or committee capable of overseeing inter-departmental businesses and take charge of the overall supervision and management of the use of AI systems. The senior executive or committee and the departments or teams they lead should formulate AI policies and take charge of supervising the use of AI systems. If a financial institution introduces an AI system on its own, it should supervise the planning and design of the system, data collection and input, and modeling and validation

of the AI system throughout its life cycle, and ensure compliance with laws and regulations in the process of application.

IV. Risk Management Mechanisms

(I) Set clear risk management policies or risk-based integration into existing mechanisms

1. Financial institutions may set clear risk management policies and guidelines for the application of AI systems that include items such as risk management, data collection, security control, compliance requirements, monitoring, and assessment. Financial institutions are also advised to create an organizational culture conducive to AI development, encourage the implementation of AI core principles and the realization of responsible AI, and set mechanisms to allow employees to raise questions or concerns if they express concern about the AI system.
2. Integrate AI risk management into the existing risk management and internal control framework: Integrated items include model risk management, information security, data protection, and measures for treating customers fairly. If there are still deficiencies, additional risk management and internal control frameworks may be added to comply with the core principles of AI.

(II) Risk management of AI models

1. Pre-deployment management: Financial institutions are advised to understand and document the purpose and intended use of the AI model as well as the methodology and concepts used in the AI model. Financial institutions are advised test the model before deployment to ensure that the results meet the objectives.
2. Continuous validation: Financial institutions are advised to conduct continuous validation of the AI model wherever possible. However, the frequency of validation may vary based on the complexity of the model and the performance of the tools used for periodic reviews. Financial institutions may evaluate the reliability, ability to facilitate recognition, and error correction of the model, and check for any sign of deterioration in the quality of the results produced by the model.
3. Create model inventory: Financial institutions are advised to create and maintain an inventory of AI models that includes information on past and current versions of each AI model and those under development[6]. The inventory should include the type and source of data input, the output and intended use of the model, and an assessment of whether the model is operating as intended.

[6] Information on versions of models in the development process shall be retained for a period of time determined by the financial institution based on the importance of the model.

(III) Continuous monitoring and improvements: Financial institutions are advised to maintain, monitor, document, and review deployed AI systems and provide appropriate resources based on factors in risk assessment, and help the management understand the performance of deployed AI systems and other relevant matters. Under suitable conditions, the monitoring may include autonomous monitoring with AI systems designed to automatically report on the level of confidence in their predictions.

(IV) Regular review of risk management mechanisms to increase effectiveness

1. Establish internal review and monitoring mechanisms: Financial institutions are advised to establish risk-based internal review and monitoring mechanisms to regularly assess whether AI systems meet the original objectives of application and the level of risks to ensure that AI systems are in line with policies and guidelines and to resolve any potential problems in a timely manner. Financial institutions may, where necessary, invite specialists from different fields to participate in the AI evaluation process. They may include specialists in human resources, behavioral science, law, ethics, and sustainable development who can help provide insights for AI development.
2. Establishment of independent third-party review and communication mechanisms: After evaluating the risk, internal resources, and requirements for professional services of AI systems with higher risks, financial institutions may, if necessary, establish mechanisms for an independent third-party with AI expertise to carry out the review and assessment. They may also set up channels or communication mechanisms to allow external parties to inquire or review the relevant information on the matters that are affected by the decision-making process to facilitate feedback from external parties and help financial institutions comply with the core principles in their application of AI systems.

V. Personnel Training

(I) Financial institutions are advised to provide training and resources for the departments, teams, and relevant personnel responsible for AI systems to enhance personnel's understanding of and ability to introduce, use, and manage AI systems, adapt to the rapid development and changes in AI technologies, and implement suitable risk-based decision-making and supervision. These personnel include senior management responsible for the overall AI system, project managers (e.g. development, testing, supervision, compliance, risk control, and internal audit personnel), supervisors, execution teams, and other relevant employees. Financial institutions are also advised to verify that

their board of directors and management have knowledge of the AI system they use.

(II) Financial institutions are advised to identify new and changing roles, assess the skills that must be upgraded or reacquired, and the characteristics of new employees that they need to recruit to help the organization adapt to new ways of working and achieve effective human-machine collaboration.

(III) Financial institutions are also advised to establish communication and interaction channels with stakeholders, so that AI system appliers can easily incorporate feedback from all sectors into each stage of the life cycle.

Chapter 2 Emphasizing Fairness and Human-Centered Values

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

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

Due to the automated nature of AI systems, if operators fail to pay close attention to the design, data collection, and modeling, it may result in discrimination or unfairness, or conditions beyond human control. Therefore, when financial institutions apply AI systems, they are advised to assess their fairness, enhance the reasonableness and accuracy of their decisions, pay attention to the biases, and avoid discrimination wherever possible. When collecting data, it is advisable to pay attention to its source and possible biases, use personal data with care, and regularly review and validate the results of AI models to enhance the accuracy of the AI system and achieve its purpose and to avoid discrimination wherever possible. In addition, AI systems should be applied in a manner that supports human autonomy, respect fundamental human rights, and permits human oversight.

II. Main Concepts

(I) Fairness: The decisions produced by AI systems applied by financial institutions must not result in discriminatory outcomes for specific groups. It means that decisions must be reasonable, accurate, and as non-discriminatory as possible.

1. Reasonableness of decisions: (1) If personal attributes are utilized as factors in the decision-making of the AI model, there must be reasonable justification; (2) If there is no reasonable justification, the

decision made by the AI system must not be systematically unfavorable to a particular group of people (e.g. no unreasonable loan conditions based on the borrower's religion, race, gender, disability, sexual orientation, place of residence, political affiliation, age, nationality, or ethnicity).

2. Accuracy of decisions and avoidance of discrimination where possible: AI decision-making models and data should be regularly reviewed and validated to increase their accuracy and avoid discrimination as much as possible. Financial institutions are also advised to regularly review the resulting decisions generated by the models to ensure that the results of the model calculations are in line with the purpose of the design.

(II) Human-centered: The AI system should be designed to support human autonomy, respect basic human rights, and allow human oversight throughout its life cycle in order to realize human values and achieve the goal of improving human well-being.

(III) For consumers who are affected by adverse outcomes, financial institutions are encouraged to provide remedies, which may include existing remedies already provided by the financial institutions. However, if the AI system applied by a financial institution for money laundering prevention or fraud detection and the supply of remedial options is not suitable, it may be withheld.

(IV) The supervision mechanisms of humans in AI systems' decision-making process employ human-in-command (HIC), human-in-the-loop (HITL), and human-over-the-loop (HOTL) approaches, which are described as follows:

1. "Human-in-command": It refers to the ability of humans to command and supervise the overall activities of an AI system (including its broader economic, social, legal, and ethical impacts), and to decide when and how to use the AI system under any condition.

2. "Human-in-the-loop": It means that humans actively participate in the supervision and retain full control, and the AI system only provides suggestions or information. The AI system cannot make decisions unless the human issues a command to the AI system to do so.

3. "Human-over-the-loop": Humans only take over control and adjust parameters during the computing process when the AI model encounters accidents or undesirable events (e.g. model failure).

III. Implementation of Fairness

Financial institutions are advised to pay attention to the following issues at each phase of the AI system life cycle, assess the fairness risk rating, and take appropriate risk-based measures:

(I) "System planning and design" phase

1. Establish the purpose and identify groups that may be adversely affected: Financial institutions are advised to verify the purpose of the planned AI system and identify groups that may be subject to disadvantageous treatment in the application of the AI system (hereinafter referred to as "adversely affected groups") and the likelihood of adverse impact, and retain written or digital records.
2. Inclusion of professionals: Financial institutions may, where necessary, invite professionals to participate in the design and implementation of the AI system. With their professional advice and guidance, financial institutions ensure that the design and implementation of the AI system do not discriminate against any individual or group.
3. Provide remedial options: Financial institutions are advised to include or provide remedial options in the AI system for feedback from adversely affected groups, and collect opinions or suggestions from such groups for reference by financial institutions when they carry out subsequent calibrations.

(II) "Data collection and input" phase

1. Financial institutions are advised to review the data they plan to collect, the method of collection, and the source of the data for possible bias, and verify whether such bias is unfair or discriminatory.
2. Financial institutions are encouraged to use data from diverse sources that represent a range of backgrounds and characteristics instead of relying solely on data from a single category or group to minimize bias and discrimination against certain groups.
3. Financial institutions should exercise caution when using personal attributes (e.g. age, gender, place of residence, ethnic group, religion, and nationality).

   (1) Before deciding to collect or use personal attributes, it is advisable to determine whether the use of such attributes is consistent with the planned purpose of AI system.

(2) If the adoption of certain personal attributes may adversely affect certain groups, it is advisable to evaluate the necessity of using such attributes and the possibility of using alternative attributes or alternative methods. Financial institutions should also keep the evaluation results, the methods employed, and the reasons in written or digital format.

(III) "Modeling and validation" phase

1. Self-examination of the results of models on different groups: If a financial institution independently develops an AI system or commissions a third party to do so, it can test and validate the AI model's predictions and decisions on different groups to verify that its operations are fair and unbiased. If bias is found, the financial institution should assess and take appropriate adjustments and improvements to avoid unfairness or discrimination as much as possible.

2. Submission of results to independent and qualified external professionals for review and validation: If necessary, the financial institution may submit the results of the AI system to independent and qualified external professionals for review and validation to confirm that decisions made by the AI system are as reasonable, accurate, and non-discriminatory as possible, and make any necessary corrections or improvements. Financial institutions are also advised to conduct due diligence to prevent possible conflicts of interest.

3. Retention of written or digital records: To verify the fairness of the AI model, financial institutions are advised to retain easy-to-understand written or digital records of the purpose of the AI model design, computing logic, and the decision-making process. This ensures that stakeholders can understand the operating principles of the model and the decision-making method, helps determine the fairness of the AI system for different groups, and facilitates tracing and interpretation of the relevant results.

(IV) "System deployment and monitoring" phase

1. Financial institutions are advised to regularly review and analyze the results produced by the AI system for discrimination, and use the information collected through the remedial options to determine whether there is unequal treatment of different groups by the AI model. If an issue involving discrimination is found, the financial institution should swiftly implement adjustments for improvement.

2. Financial institutions are advised to identify any correlation between the application of the AI system and a particular group of individuals who are subject to adverse systemic treatment. If any is found, financial institutions are advised to adopt measures that reduce the impact on that particular group of individuals.

IV. Implementation of the Principles of Human-Centered and Human-Controlled AI

(I) Before a financial institution applies an AI system, it is advised to determine whether the system complies with laws and regulations and determine whether it would affect the autonomy or basic human rights of its customers. If it has not done so, it is advised to discontinue the application of the AI system and implement improvement measures until such concerns are eliminated.

(II) With respect to the scope of application of AI systems, when assessing the degree of human participation required for the use of AI-assisted decision-making, a financial institution is advised to consider the degree of impact that AI decisions may have on its customers or financial institutions, and adopt different levels of supervisory mechanisms or to plan other possible risk mitigation, transfer, or avoidance measures. With regard to the aforementioned applications with a higher degree of impact, it is advisable to adopt human-in-command or human-in-the-loop supervisory mechanisms.

(III) For important critical systems, a financial institution is advised to retain personnel access so that it can review, approve, or make final decisions on the AI system, including ensuring that its personnel can intervene to take over control, and that the AI system can provide sufficient information for personnel to make meaningful decisions; or in the event that personnel cannot control or intervene in the decision-making process, the AI system can be securely shut down by personnel.

V. Risk Management of Output Information from Generative AI

(I) When a financial institution introduces generative AI, it is advised to assess whether it is biased or discriminatory against specific groups, and to reduce possible unfairness according to the above points.

(II) When a financial institution uses generative AI developed by a third party and is unable to control the training process and ensure that its data or the results of its calculations meet fairness requirements, risks

associated with information produced by generative AI must be objectively and professionally controlled by its personnel to avoid unfairness to clients or financial consumers.

Chapter 3 Protecting Privacy and Customer Interests

Core Principle 3: Protecting privacy and customer interests
(I) Financial institutions should fully respect and protect the interests of consumers, and manage and use customer data properly.
(II) Financial institutions applying AI systems to provide financial services should respect the customer's right to choose, and should remind customers of available alternatives.

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

To ensure accuracy, AI systems may need to collect vast amounts of information from consumers or customers and collect and process information when interacting with customers. Financial institutions are therefore advised pay attention to customer privacy protection throughout the life cycle of the AI system, adequately process customer information, avoid risks of data leakage, and adopt the principle of data minimization[7] to avoid the excessive or unnecessary collection of sensitive information. Financial institutions are also advised to respect the customer's right to choose whether to use AI services, and determine whether to provide alternatives based on the risks, feasibility, and costs of the alternatives for the customer and the institution.

II. Main Concepts

(I) Due to the development of big data and AI technology, customers' personal information is often collected in large quantities and used to train AI, which may threaten customers' privacy and affect the public trust in financial institutions and service satisfaction. Financial institutions are therefore advised to pay attention to the protection of customers' privacy and the proper collection and handling of customers' information to avoid the risk of data leakage in the application of AI systems.

(II) Financial institutions are advised to collect and process necessary customer information based on the principle of data minimization and

[7] The principle of data minimization means that the collection of personal data must be appropriate, relevant, and within the scope necessary for the purpose of data processing.

avoid excessive or unnecessary collection of sensitive information.

(III) When applying AI systems in services for customers, financial institutions are advised to inform customers and respect their right to choose whether to use AI services and remind them of the availability of alternatives to protect customer rights and interests.

(IV) When applying AI systems, financial institutions are advised to pay attention to customer privacy protection, including personal information, intellectual property rights, and trade secrets.

III. Privacy Protection and Data Governance

(I) For the implementation of customer privacy protection and data governance, a financial institution is advised to pay attention to the following issues at each phase of the AI system life cycle:

1. "System planning and design" phase

(1) A financial institution is advised to assess whether the design of the AI system complies with the Personal Data Protection Act and other relevant regulations and its internal data governance policies.

(2) A financial institution is advised to evaluate the information to be collected and determine the feasibility of data acquisition channels based on the intended use of the AI system, and follow the principle of data minimization for collection and processing. If the collected public information satisfies the purpose of AI system design, the collection non-public information is not required.

(3) A financial institution is advised to set up mechanisms to protect personal information from unauthorized access, damage, loss, or leakage. It should also process matters in accordance with the "Regulations for the Security Maintenance of Personal Data Files by Non-Governmental Agencies Supervised and Designated by the Financial Supervisory Commission." Financial institutions can take measures such as encrypting sensitive information, setting up access control mechanisms, conducting regular security monitoring and audits, and ensuring that employees receive appropriate confidentiality training.

2. "Data collection and input" phase

(1) A financial institution is advised to record the source of information collection and ensure that information is obtained through legal and reliable channels.

(2) A financial institution is advised to verify the accuracy and completeness of the information and examine the information for errors.

(3) If the data collected by a financial institution contains personal data, the financial institution should verify that it has obtained the consent of the customer or is in compliance with relevant laws and regulations, and implement risk-based evaluation on the necessity of additional privacy protection for the data[8].

3. "Modeling and validation" phase

(1) A financial institution is advised to ensure that the information used to train AI models and the information generated by AI systems do not violate personal data protection laws and related regulations.

(2) A financial institution is advised to ensure that partners and vendors also comply with privacy and security standards.

4. "System deployment and monitoring" phase

(1) A financial institution is advised to regularly monitor AI systems, partners, and suppliers to ensure their continuous compliance with relevant privacy and security standards after deployment, and check for irregularities that may adversely affect the protection of financial consumers' privacy or their rights and interests.

(2) If there is any data leakage or violation of the Personal Data Protection Act in its AI system, a financial institution shall notify and handle the matter through the existing mechanisms and implement necessary adjustments to its AI system.

(II) A financial institution must pay attention to the risk of customer data leakage that may result from the application of its AI system. For instance, in the application of generative AI by its employees, unless appropriate control mechanisms are implemented, generative AI should not be provided with information that has not been provided with the consent of the customer. Examples of appropriate control mechanisms include the use of closed-source generative AI models and verification of the security of the system environment. This requirement does not apply if the information does not involve personal data and cannot be directly or indirectly used to identify the behavior of specific individuals as customers.

[8] Examples include pseudonymization, anonymization, or the use of privacy-enhancing technologies (PETs).

IV. Respect for Customers' Right to Choose and Alternatives

(I) A financial institution applying an AI system to provide financial services should pay attention to the following matters:

1. Should inform financial consumers that the financial service is provided by an AI system.
2. Should provide information to help financial consumers understand how the AI system functions and how AI-assisted decision-making may affect them in their regular use of the system.
3. Should remind financial consumers of the availability of alternatives to the AI system so that they can choose whether to use the services provided by the AI system.

(II) A financial institution may review the following factors when deciding whether to offer alternatives when a customer discontinues the use of services provided by the AI system:

1. The risks and degree of harm to the financial institution.
2. The risks and degree of harm to the customer.
3. The likelihood that the customer will resume use of the AI system after choosing an alternative.
4. Feasibility and cost of the alternatives.
5. The complexity and efficiency of the simultaneous appliction of AI systems and alternatives.
6. Technical feasibility.

(III) If a financial institution decides not to provide an alternative after considering the aforementioned factors, it should evaluate whether to provide remedial measures to the customer.

Chapter 4 Ensuring System Robustness and Security

Core Principle 4: Ensuring system robustness and security
(I) When applying AI systems, financial institutions must ensure system robustness and security to avoid causing harm to consumers or the financial system.
(II) 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.

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

As financial institutions increase the application of AI systems, system stability and security gain prominence in the normal operation of financial institutions. Financial institutions should therefore define the purpose of the system and conduct risk assessment. They should choose a resilient AI model, focus on the quality of the data when collecting data, conduct appropriate interaction verification and resistance testing of the model, and deploy the AI system in an appropriate environment. They should also set up security measures to ensure protection against security threats and attacks, and implement continuous monitoring to ensure the overall security and stable operation of the AI system.

II. Main Concepts

(I) System robustness: This refers to the methods set up in the AI system for preventing the occurrence of risks, which not only ensure reliable performance of its intended purpose, but also minimize unintended or unforeseen adverse effects and prevent unacceptable adverse effects. System robustness includes the following concepts:

1. Stability: A stable AI system means that the computer system adequately addresses errors in execution and erroneous inputs, and that the system or its components function correctly even when subjected to invalid inputs or stressful environmental conditions.

2. Accuracy: An AI system with high accuracy means that the system has the ability to make correct judgments to achieve its intended

purpose, such as correctly dividing information into appropriate categories or making predictions, recommendations, or decisions based on data and models that meet the intended purpose. Well planned and developed AI systems can minimize and correct the unintended risks associated with inaccurate predictions. Even if an AI system occasionally makes inaccurate predictions, the error rate can be determined in tests.

3. Reproducibility: A reproducible AI system means that repeated tests of the AI system under the same conditions yield similar output.

(II) System security: A secure AI system refers to a system with strong defense against external security threats, attacks, or malicious misuse of information security, and complies with the requirements of the financial industry's information security regulations, and ensures that its system operates in accordance with its intended functions.

III. Implementation of System Robustness

Financial institutions are advised to pay attention to the following issues at each phase of the AI system life cycle:

(I) "System planning and design" phase

1. Financial institutions are advised to define the purpose of the AI system and, based on that purpose, determine the indicators to be used to measure system robustness and the thresholds (standards) for such indicators.

2. A financial institution is advised to assess the risk of failure of its AI system to achieve its intended purpose and plan for possible risk reduction, transfer, or avoidance.

(II) "Data collection and input" phase

1. Data management: High-quality data is the basis for ensuring the stability, accuracy, and reproducibility of AI models. Financial institutions are therefore advised to process data appropriately based on the quality of the data and the intended purpose of the AI system, such as replacing missing values, implement data coding, and data standardization.

2. Financial institutions can also use automation tools to ensure data quality.

(III) "Modeling and validation" phase

1. A financial institution is advised to choose a model with greater

resilience[9] and ensure that the model meets its objectives.

2. If a financial institution develops its own AI system or commissions a third party to do so, it can enhance the robustness and reliability of the AI model with cross-validation and calibration.

3. A financial institution can use resilience tests to assess the resilience of AI models in response to unanticipated inputs and make necessary adjustments.

4. A financial institution is advised to conduct effectiveness tests to verify that the AI model meets the minimum system stability indicators in initial settings.

5. For an AI system with higher risk or greater impact, a financial institution should consider whether to test it first and separate the testing environment from the regular operation environment. A financial institution can also test the performance of an AI system under stressful market conditions.

(IV) "System deployment and monitoring" phase

1. A financial institution is advised to deploy an AI system in an appropriate environment to minimize the impact of external factors (e.g. power stability and network bandwidth) on the performance of the AI system.

2. A financial institution is advised to establish appropriate monitoring mechanisms to regularly check whether the AI model exhibits deviation in its effectiveness, and to immediately address any decline in the accuracy of the model or other problems.

IV. Implementation of System Security

(I) A financial institution is advised to comply with system security regulations, set up suitable information security or management measures, and ensure protection against security threats and attacks such as hackers and malicious software. It should also implement continuous monitoring to ensure the security of the AI system.

(II) A financial institution is advised to take control measures to avoid risk of any leak of model parameters or data due to inappropriate operations or human negligence by third-party operators during model training.

(III) A financial institution is advised to pay attention to the following

[9] Resilience refers to the ability of an AI model to appropriately process unanticipated data entries or decline to make predictions when the confidence level is too low.

issues at each phase of the AI system life cycle:

1. "System planning and design" phase

   (1) Enhance employee awareness of security threats and risks and assist employees in planning appropriate methods to mitigate risks.

   (2) Assess potential threats to the system.

   (3) When selecting AI models, it is advised to consider security in addition to functionality and performance.

2."Data collection and input" phase: Strengthen data security control and reduce the risks of data leakage.

3. "Modeling and validation" phase

   (1) Assess or periodically review the security of suppliers of AI technologies and require suppliers to comply with security standards.

   (2) Identify, track, and protect AI-related assets (e.g. models, data, prompts, software, record files, and internal assessments).

   (3) Retain written or digital records of models, data, and prompts.

4. "System deployment and monitoring" phase

   (1) Protect infrastructure with measures including adequate control over the use of APIs, models, and data, and proactively prevent cyberattacks that lead to the theft of models or damage to performance.

   (2) Protect models and data with measures such as implementing information security practices or managing user interfaces.

   (3) Conduct appropriate and effective security assessments before the deployment of AI models.

   (4) Monitor the output and performance of models and systems to observe changes that may jeopardize security.

   (5) While taking care to satisfy the prerequisite condition of privacy and data protection, record and appropriately monitor the contents entered into the system.

   (6) Use secure and modularized processes for updates.

Chapter 5 Ensuring Transparency and Explainability

Core Principle 5: Ensuring transparency and explainability
(I) Financial institutions should ensure the transparency and explainability of their operations.
(II) Financial institutions should make proper disclosure when their AI system interacts directly with consumers.

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

When financial institutions use AI to interact with consumers, their decisions or interactions have significant effects on consumers. Financial institutions are therefore advised to disclose relevant information to consumers in an appropriate manner. If a financial institution develops its own AI system or commissions a third party to do so, it should ensure that their personnel can clearly explain the logic of the AI system's operation internally and to the auditors when necessary, so that when it is necessary to modify or review the AI system, it can be done under appropriate conditions.

II. Main Concepts

(I) Transparency: Transparency refers to the provision of information about AI systems to external stakeholders to help them understand the impact on their rights and interests as well as the limitations and risks of such AI systems.

(II) Explainability: This refers to the possibility to clearly explain the operation of the AI system developed or commissioned by the financial institution for use, as well as the logic behind its prediction or decision-making process. It facilitates the financial institution's assessment of its compliance with internal policies, operational procedures, and supervisory requirements.

(III) A financial institution should learn about how the AI system it applies makes decisions and should enhance the explainability of the AI system to ensure effective management of AI system operations.

(IV) When applying an AI system, a financial institution should proactively

disclose relevant information to stakeholders. If stakeholders request additional explanation, they should provide appropriate explanation of the information used, how the information affects decision-making, and the impact of decisions on stakeholders to increase public trust. However, if the AI system applied by the financial institution is for money laundering prevention, information security, or fraud detection, or if it involves the trade secrets of the company, or if excessive disclosure of information may lead to other risks, it is advisable to carefully review the necessity and extent of disclosure of relevant information to individuals other than the competent authority.

(V) Financial institutions are advised to set common principles for transparency and explainability at each stage of the AI system life cycle:

1. Transparency

(1) A financial institution should establish common principles for assessing the required transparency of AI systems, such as the extent, timing, and form of explanation provided to customers when using an AI system to generate loan decision recommendations.

(2) A financial institution should identify items that they may need to inform customers of at different stages of the customer service life cycle, and prepare sample notifications in advance.

2. Explainability

(1) A financial institution should establish common principles for assessing the explainability of AI systems, including how to assess the degree of explainability required and the intended recipients of relevant information.

(2) A financial institution should select a suitable method of explanation for the AI system and specify the basic requirements of the aforementioned method of explanation.

III. Implementation of Transparency and Explainability

Financial institutions are advised to pay attention to the following issues at each phase of the AI system life cycle:

(I) "System planning and design" phase

1. Transparency

(1) A financial institution should determine the degree of transparency of its AI system in accordance with the principle of commonality

of transparency, and identify the active and passive notification items and their forms required for each stage of the customer life cycle.

(2) When a financial institution plans to use an AI system to interact with financial consumers, it should adopt a plan for appropriate disclosure of relevant information in plain language. Examples include:

- Help financial consumers understand that the financial product or service they are considering is provided by AI.
- Help financial consumers understand the functions and expected performance of the AI system in regular use, and inform financial consumers of the possible positive and negative impacts.
- If a financial institution applies the AI system to assist in making decisions, it is advisable to consider providing additional information to financial consumers so that they understand how these decisions are made.

2. Explainability

(1) A financial institution should determine the required level of explainability and the recipients of relevant information based on the commonality principle of explainability assessed during preparation.

(2) A financial institution should select an appropriate method of explanation based on the degree of explainability mentioned above, and evaluate whether the method of explanation meets the basic requirements.

(3) If a financial institution develops its own AI system or commissions a third party to do so, it is advisable to plan for the preparation of relevant documents or technical reports on the structure of the AI system, and plan for providing supervisory authorities with access to and understanding of the AI system and the data used. It will enable them find information and understand the operations of the financial institution's AI system and the appropriateness of the data used in the future.

(II) "Data collection and input" phase: To meet future transparency or explainability requirements, a financial institution should ensure that information related to the training of the AI system (e.g. data sources, time points, consumer consent, data dictionaries, and known

limitations of the data) is retained written or digital form.

(III) "Modeling and validation" phase

1. Transparency

(1) A financial institution is advised to test and validate the corresponding functions in accordance with transparency requirements.

(2) A financial institution is advised to adopt a plan for appropriate adjustments of operating procedures (e.g. customer service and complaint handling procedures) to enhance transparency to customers and train employees to respond to customer inquiries.

(3) A financial institution is advised to appropriately disclose and update the terms of service for customers or websites. It should, for example, explain to customers how the AI system uses customer information, the benefits and risks to customers, and how customers can participate, withdraw, and raise questions.

2. Explainability

(1) If a financial institution develops its own AI system or commissions a third party to do so, it should submit an explainability report in accordance with the explainability requirements.

(2) The financial institution shall review and confirm whether the explanation is appropriate and ensure the level of explanation is commensurate with the importance of the AI system application. If the explainability of the AI system does not meet the financial institution's expectations, it is advisable to plan other measures (e.g. switching to simpler models, removing difficult-to-explain functions, or introducing more human supervision).

(3) After building a model, the financial institution should verify that its employees understand the architecture, algorithms, functions, and decision-making factors of the AI system, and that the operating procedures of the AI system can be understood and explained. If employees fail to clearly explain such information, it is advisable to review whether the financial institution should avoid the use of architecture, algorithms, or functions that are too complex or unexplainable.

(IV) "System deployment and monitoring" phase:

1. Before a financial institution uses AI to interact directly with a

consumer, it is advisable to proactively inform the consumer that the interaction or service is provided with automated response from an AI system.

2. A financial institution os advised to provide appropriate description and explanation based on customers' needs. For customers who may be subject to adverse and unfair treatment in the system, a financial institution may assess the logic of the predictions, recommendations, or decision-making in the AI system in a simple and easy-to-understand manner. However, it must remain vigilant regarding risks of excessive information disclosure.

3. A financial institution should periodically monitor its AI system to ensure the attainment of transparency and explainability.

4. To increase market trust in AI, a financial institution may, if necessary, proactively inform stakeholders of its AI practices through the publication of reports, technical documents, or the disclosure of relevant information on its website, including describing the purpose of its AI system, its principle of operations, data and algorithms used, and the possible impacts of its AI system.

Chapter 6 Promoting Sustainable Development

Core Principle 6: Promoting sustainable development

(I) 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 wellness.

(II) 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.

***This core principle is based on the "Core Principles and Policies for Artificial Intelligence (AI) Applications in the Financial Industry" published by the FSC on October 17, 2023.**

I. Purpose

The operation of AI systems may consume energy and water, and may deprive or threaten the work of current employees while exacerbating the digital divide. Financial institutions should therefore enhance their social and environmental responsibilities, such as using emerging technologies to reduce the consumption of resources and facilitating inclusive finance in the digital transformation. Financial institutions should also consider the transformation of their regular employees during the digital transformation to pursue sustainable and stable development.

II. Main Concepts

(I) When financial institutions apply AI systems, they should consider the society and the environment as stakeholders and support social equity and ecological responsibility, such as promoting inclusive finance in the digital transformation, reducing the digital gap, and minimizing the consumption of water, electricity, and other energy in the process of using AI.

(II) The strategy for the application of AI systems by financial institutions and implementation strategies should be based on international sustainability goals as well as their own sustainability principles. They should also include appropriate sustainability indicators.

III. Implementation of Sustainable Development

(I) Identify the impact: A financial institution should establish mechanisms to identify and assess the impact or risks of its AI system on the environment and society.

(II) Optimize hardware facilities: A financial institution should choose hardware facilities with higher energy efficiency to reduce energy consumption, such as using energy-efficient servers, low-power processors, and high-efficiency data center equipment. It should also optimize the configuration and management of hardware facilities to improve energy efficiency.

(III) Sharing resources and virtualization: A financial institution can use virtualization technology and resource sharing to centralize the management of computing resources and data storage. It can also reduce the duplication of hardware to reduce energy consumption.

(IV) Improve models and algorithms: A financial institution can optimize the algorithms of AI systems and reduce the complexity of models and computing requirements to increase computing efficiency and resource utilization.

(V) Pre-processing data: A financial institution can pre-process data to reduce unnecessary data transmission and improve data quality to reduce the energy consumed by repeated computing for the purpose of improving accuracy.

(VI) Intelligent energy efficiency management: A financial institution can maximize the use of energy management systems to monitor the energy consumption and performance of the AI system in real time, promptly identify and resolve energy waste issues, and continuously improve the energy efficiency of the system.

(VII) Recycling and reuse of resources: To reduce the overuse of global resources, a financial institution can recycle and reuse old hardware equipment to reduce the creation of electronic waste. A financial institution can also utilize right-of-use hardware equipment to reduce the use of raw materials.

(VIII) Reducing digital anxiety and the digital gap: Financial institutions can provide services that meet the needs of financial consumers based on their attributes to reduce any possible digital anxiety or digital gap. For example, they can provide financial consumers with the opportunity to opt for face-to-face services to reduce customer anxiety about AI systems, or provide a digital experience with incentives for disadvantaged groups to adopt digital services for digital applications.

IV. Employee Education and Training

(I) Financial institutions must respect and protect the rights and interests of regular employees, including providing appropriate education and training to help them adapt to the new work environment and minimize the risk of unemployment during the digital transformation.

(II) When using AI systems, financial institutions should provide relevant training to their employees, including the basic concepts and operation of AI and how AI affects individual departments and work processes. They should set up task forces as necessary to monitor the impact of the AI system and the adaptation of their employees. They should also adjust the training programs based on actual requirements.

(III) Financial institutions may also raise the awareness of their employees in terms of energy conservation, reducing the overuse of resources, and care for the digitally disadvantaged, and provide corresponding training and guidance to enhance their implementation of sustainable development.

References

1. Board of Governors of the Federal Reserve System, and Office of the Comptroller of the Currency (OCC), 2011, "Supervisory Guidance on Model Risk Management."

2. Department for Science, Innovation and Technology, 2023, "Capabilities and risks from frontier AI."

3. European Commission (EU), 2019, "Ethics Guidelines for Trustworthy AI", High-Level Expert Group on Artificial Intelligence.

4. Financial Stability Institute (FSI), 2021, "Humans Keeping AI in check—emerging regulatory expectations in the financial sector", FSI Insights on policy implications No. 35.

5. International Organization of Securities Commissions (IOSCO), 2021, "The Use of Artificial Intelligence and Machine Learning by Market Intermediaries and Asset Managers."

6. Monetary Authority of Singapore (MAS), 2022, "Veritas Document 3: FEAT Principles Assessment Methodology."

7. Monetary Authority of Singapore (MAS), 2022, "Veritas Document 3A: FEAT Fairness Principles Assessment Methodology."

8. Monetary Authority of Singapore (MAS), 2022, "Veritas Document 3B - FEAT Ethics and Accountability Principles Assessment Methodology."

9. Monetary Authority of Singapore (MAS), 2022, "Veritas Document 3C - FEAT Transparency Principles Assessment Methodology."

10. National Cyber Security Centre (NCSC) et al., 2023, "Guidelines for secure AI system development."

11. National Institute of Standards and Technology (NIST), 2023, "AI Risk Management Framework (AI RMF 1.0)."

12. The Organization for Economic Co-operation and Development (OECD), 2021, "OECD Business and Finance Outlook 2021."

13. The Personal Data Protection Commission of Singapore (PDPC), 2020, "Model Artificial Intelligence Governance Framework."

14. Bankers Association (research unit: EY Advisory Services Inc.), 2023, "Outsourced Research Project on Artificial Intelligence Information Security Protection."

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

1. Foreword

The increasing applications of AI[1] 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 AI[2] 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 AI[3] 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' Meeting[5] 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 2020[6] 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 Rights[7], 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 Act[8] 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 Technology[9] 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

[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 States[13], Singapore[14], South Korea[15], 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 personnel[16].

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 consumers[17].

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.