

The approach to AI governance in Japanese financial institutions

Identification and management of AI as a "model" with consideration of international trends

In recent years, as AI has been increasingly utilized across various business operations, the development of AI risk management frameworks—AI governance—has gained momentum. Not only is the strategic use of AI ("proactive" perspective) important, but establishing safeguards to manage its risks ("defense" perspective) is equally critical. While companies across industries are exploring both "proactive" strategies and "defensive" structures, financial institutions, which require more robust risk management, must consider how best to establish AI governance. This paper examines international trends indicating that AI is organized and managed as a "model" within a structured risk management framework. Based on these insights, we propose an approach to AI governance for Japanese financial institutions.

Definition of AI

In the operations of financial institutions, AI is utilized almost daily.¹ The areas and tasks where AI is applied are diverse, ranging from chatbot-based Q&A support to image and voice recognition, customer recommendations, risk management and assessment, fraud detection, and compliance. AI is being used for a wide variety of purposes—from improving efficiency to enhancing business practices (from a "proactive" perspective).

At the same time, it is essential to manage AI-specific risks such as bias in input data, hallucinations in output, legal and reputational risks, concentration risks related to third-party providers and cyberattacks (from a "defensive" perspective). Based on this context, governments and financial institutions are actively exploring what kinds of AI risk management frameworks—known as AI governance—should be established.²



Yasuhiro Tanaka
Director
KPMG AZSA LLC
Financial Services
Division FS
Advisory

¹ This paper does not distinguish between conventional AI and generative AI, but rather discusses AI as a whole. For insights on conventional AI and generative AI, FSA's "AI Discussion Paper" defines "Conventional AI in this Paper refers to AI that learns characteristics and trends by being provided with data in advance (for example, machine learning), and obtains answers to input data (including even rule-based models and chatbots that create and operate complex rules from data). Generative AI refers to models with large parameters, such as LLM, that have the function of generating new products such as documents, images, audio and video by using data and content (unstructured data such as text and images) on the Internet for training.

² Regarding AI governance, Ministry of Economy, Trade and Industry (METI) defines it in its publication "AI Governance in Japan Ver. 1.1" as "design and operation of technological, organizational, and social systems by stakeholders for the purpose of managing risks posed by the use of AI at levels acceptable to stakeholders and maximizing their positive impact." In addition, FSA published an "AI Discussion Paper" in 2025, which addresses AI governance in financial institutions and related entities in finance.

To begin with, how is AI defined? Various definitions have been presented by governments, private companies, and international organizations worldwide. Figure 1 highlights the definition of AI as stated in the "AI Guidelines for Business". In the United States, the definition provided by NIST (National Institute of Standards and Technology) is well known, while in Europe, the OECD's definition is widely accepted and is also adopted in documents published by the FSB.

These examples show that, although there are various definitions of AI, a common element among different definitions is that "AI processes some form of input data to generate an output".

Figure 1 : Example of the definition of AI

AI system
A system (such as a machine, robot, and cloud system) that works at various levels of autonomy during the use process and incorporates a software element that has a learning function.
[...]
(For reference, it is defined in the OECD AI Principles overview as follows.) An AI system is a machine-based system that, for explicit or implicit objectives, makes inferences. It generates outputs including predictions, contents, recommendations, decisions and so on to place impact on physical or virtual environments from received data. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.
AI model (ML model)
A model incorporated into an AI system and acquired through machine learning using training data. It produces prediction results in accordance with the input data.

Source: Excerpt from the Ministry of Internal Affairs and Communications and the Ministry of Economy, Trade and Industry's "AI Guidelines for Business (Version 1.1)"
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_14.pdf

Definition of a model

Next, how is a model defined? According to the "Principles for Model Risk Management" published by FSA in 2021 (hereinafter referred to as the "MRM Principles"), a definition of a model is presented in Figure 2. Historically, FRB and OCC have also provided a definition of a model in their 2011 guidance (SR11-7), which is largely consistent with the definition in the MRM Principles.³

Figure 2 : Definition of a model in the MRM Principles

The term "model" refers to a quantitative process or a system of quantitative processes that apply theories and assumptions to process data into an output(s) such as estimates, forecasts, scores or classification. Models include a quantitative process whose inputs or outputs are wholly or partially qualitative or whose inputs are based on expert judgements.

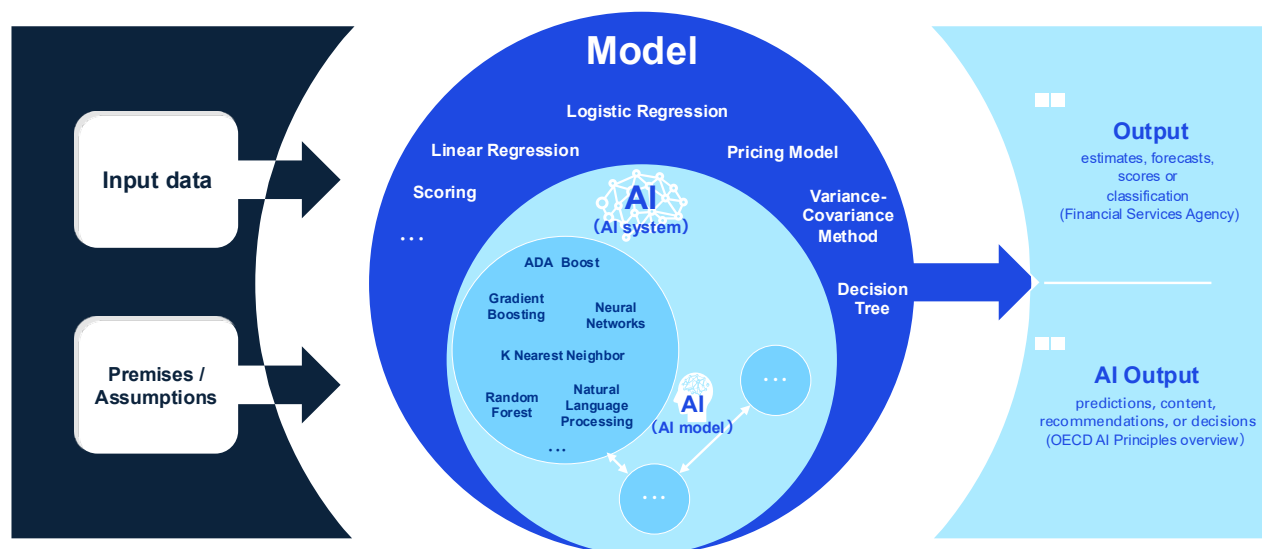
Source: Excerpt from Financial Services Agency's "Principles for Model Risk Management"
https://www.fsa.go.jp/news/r3/ginkou/20211112/pdf_03.pdf

3 [The Fed - Supervisory Letter SR 11-7 on guidance on Model Risk Management -- April 4, 2011](#)

However, AI possesses characteristics that are not typically found in conventional models. For example, one notable feature of AI is its ability to automatically perform computational processing using vast amounts of (often unstructured) data during the output generation process. Furthermore, it is important to recognize that AI carries unique risks, such as biases in input data, hallucinations in output, legal and reputational risks, concentration risks related to third-party dependencies, and vulnerability to cyberattacks.

Nevertheless, despite these differences, the definitions of AI and models discussed so far are largely similar. If AI is considered as one of many modeling techniques, it can reasonably be viewed as falling under the broader category of “models,” and it is unlikely that this perspective will be met with significant objection (see Figure 3).⁴ In particular, overseas financial institutions appear to share this general understanding.

Figure 3 : Conceptual relationship between models and AI



Note: This is for illustrative purposes only. Please be aware that other interpretations regarding the relationship between models and AI are possible.

Source: Created by KPMG in Japan

Lifecycle management

Since the publication of SR11-7, financial institutions in the United States have rigorously managed their models. Within this context, there is a prevailing view that “AI is a model under the definition provided in SR11-7.” As a result, U.S. financial institutions widely adopt the perspective that “since AI is a model, it should be governed under the Model Risk Management (MRM) framework.”

In MRM, models are incorporated into a lifecycle and managed accordingly. Specifically, computational processes and methodologies developed by the 1st line of defense that meet the definition of a model are identified as such, registered in an inventory (a comprehensive list), and then managed from that point onward. Subsequently, the model's risk is assessed based on factors such as its intended use, significance, and complexity. Following this assessment, the model undergoes validation and approval by the 2nd line of defense. Even while a model is in use, its performance is continuously monitored, and its continued use is permitted after revalidation. This lifecycle process applies fundamentally in the same way under MRM, whether the model is AI-based or not.

⁴ In this paper, the inclusion relationship between AI and models is illustrated by focusing on their definitions and methodologies. Another possible approach is to examine the risks associated with AI and models, and to organize their inclusion relationship based on commonalities in those risks.

When it comes to AI, there are issues that are difficult to address within the traditional MRM lifecycle. For example, regarding the identification of models, there are points of discussion such as: "To what extent should AI be identified and managed as a model?", and "Should AI used by individuals for tasks such as searching the web for information also be subject to management? (Managing all AI is not a realistic approach)." In addition, regarding model development and validation, there are concerns such as: "Since MRM guidance like SR11-7 does not focus on the development and validation of AI, it is unclear what aspects should be emphasized in its management."⁵ Furthermore, regarding the involvement of MRM departments, there are questions such as: "Given the unique characteristics of AI, should departments other than MRM also be involved? Is MRM alone sufficient?" Another perspective is: "Rather than leaving everything to AI, humans should ultimately be involved, and a framework to ensure such involvement should be incorporated."

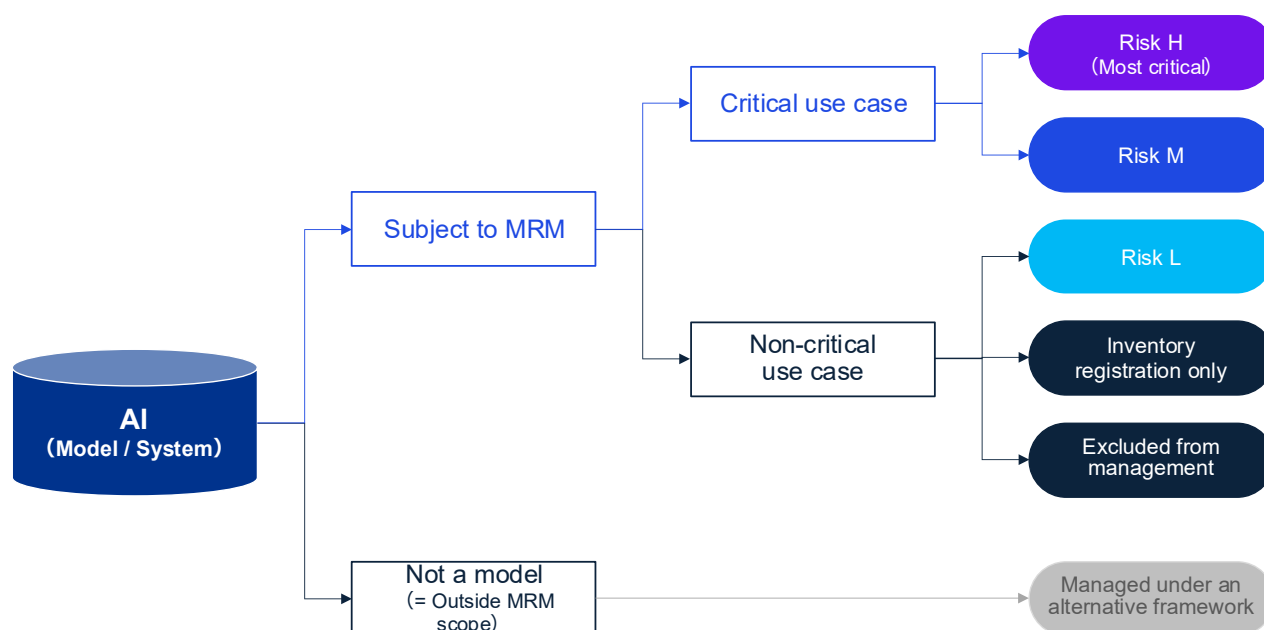
Global trends

For each of these issues, it is possible to find solutions within the framework of MRM.

(1) Model identification

First, regarding model identification, it would be reasonable to manage AI that is not considered critical—based on their intended use and users—as low-risk, register them in the inventory only, or even exclude them entirely from the MRM framework. As company-specific AI trained on proprietary data becomes more widely used, the number of AI is expected to grow significantly. In this context, it is extremely important to tailor management approaches based on the intended use of each AI (see Figure 4). In addition, to visualize what types of AI exist within the organization and who is using them for what purposes, registering AI in an inventory serves as the starting point for effective AI management.

Figure 4 : Example of classification in AI management



Source: Created by KPMG in Japan

⁵ In the aforementioned "AI Discussion Paper" by FSA, comments were made suggesting that "MRM Principles should be clarified".

(2) Model development and validation

Regarding model development and validation, it is important to incorporate AI-specific risks—such as biases in input data and hallucinations in output—into the development and validation process, while also considering the key focus areas outlined in guidance such as SR11-7.⁶

(3) Involvement of departments other than MRM

While the MRM department leads the management of models, as described later, it is essential for other departments—such as data, legal, compliance, and IT systems—to be involved in managing AI-specific risks collaboratively. Engaging these departments during the development and validation of AI is particularly important. Establishing and involving a committee responsible for approving AI usage (e.g., an AI Ethics Committee) could also be a viable approach.

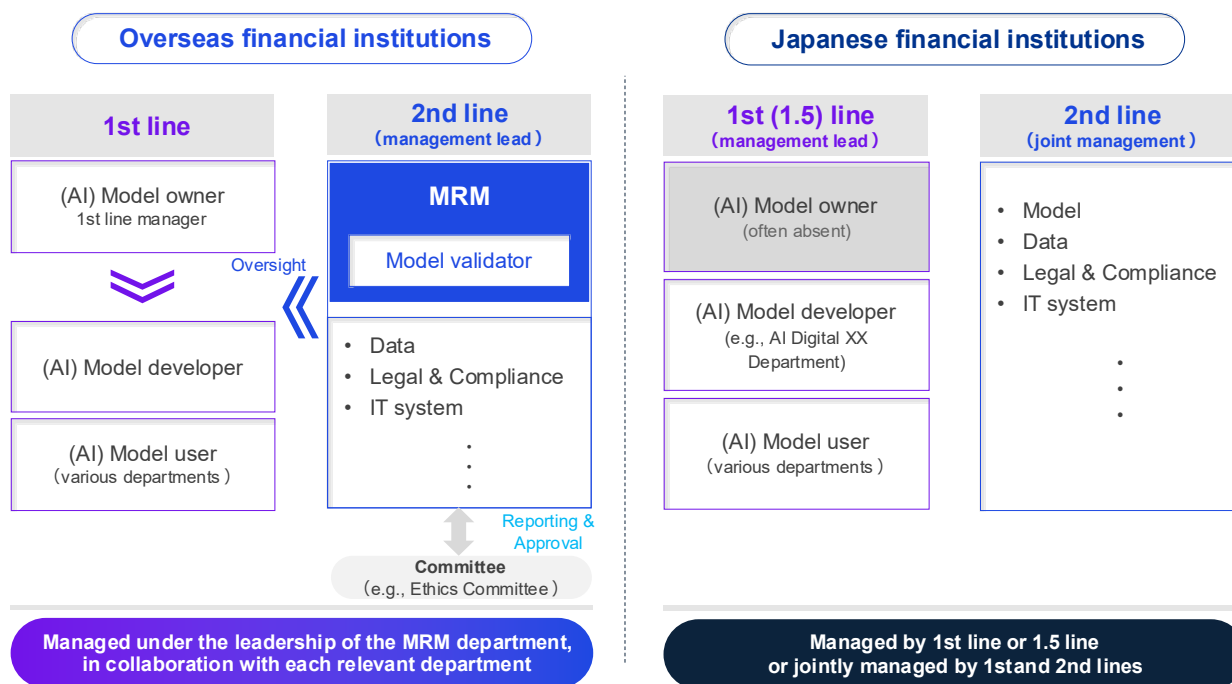
(4) Human involvement

Finally, regarding human involvement, it is important to reduce potential risks by incorporating a "Human in the Loop" framework—where humans are involved in oversight and corrective actions—within the governance of MRM, and by establishing and involving the aforementioned committees.

Based on the above, AI governance in overseas financial institutions (particularly in the U.S.) can be illustrated as shown in Figure 5. For reference, the governance structure of Japanese financial institutions is also presented. In the U.S., it is common for financial institutions of all sizes—from globally significant institutions (G-SIBs) to regional banks—to build governance frameworks similar to Figure 5 to manage AI. The MRM department serves as the foundation for AI management.

6 The KPMG Trusted AI Framework (<https://kpmg.com/jp/ja/home/services/advisory/kpmg-trusted/trusted-ai.html>) emphasizes the importance of managing AI with respect to fairness, transparency, explainability, accountability, data integrity, reliability, security, safety, privacy, and sustainability. Each of these perspectives can generally be organized within the MRM framework. Specifically, it involves identifying, assessing, and controlling AI-related risks across five key areas that are critical during model development and validation: (A) data, (B) methodology (including the model's concept and logic), (C) testing (outcome analysis), (D) implementation, and (E) governance. For example, the bias in input data discussed in this paper (corresponding to fairness in the above perspectives) falls under category (A); hallucinations in output (reliability) fall under categories (B) and (C); legal and reputational risks (privacy and safety) and concentration risks related to third parties (security and safety) fall under category (E); and cyberattacks (security and safety) fall under categories (D) and (E), where risks can be assessed and mitigated. However, especially in the case of generative AI, even among U.S. financial institutions that are advanced in MRM, there is ongoing discussion about how to conduct effective model validation—so-called "effective challenge"—in a meaningful way.

Figure 5: Conceptual image of AI governance in overseas financial institutions



Source: Created by KPMG in Japan

The approach to AI governance in Japanese financial institutions

Meanwhile, in Japanese financial institutions, it appears that departments such as the AI Digital XX Division or the AI Strategy XX Division, which are 1st line (or 1.5 line) AI departments, are taking the lead in management.

There are several factors behind this, such as situations where AI is not regarded as a model (and therefore, it is considered appropriate for a department other than MRM to manage it), or where AI is recognized as a model and should be managed by the MRM department, but in reality, the MRM function within the institution is still under development and thus unable to handle it effectively. Alternatively, there may be a perspective that introducing MRM to AI would increase the intensity of management and hinder the advancement of AI utilization, (and therefore, the involvement of the MRM department is intentionally limited).

However, considering the global trends, it may be worth exploring the involvement (or increased involvement) of the MRM department (or a department with similar functions) in Japanese financial institutions, beginning with the identification of AI as a model. Below, we present an approach for such a case.

(1) Large financial institutions such as G-SIBs and D-SIBs

For large financial institutions, such as G-SIBs and D-SIBs, which are subject to the MRM Principles, it may be appropriate for the MRM department to take the lead in managing AI, given that such institutions typically have a reasonably established MRM function. However, applying MRM may increase the intensity of management, potentially putting a sudden halt to AI utilization. Striking a balance between “proactive” and “defense” perspective is crucial. For example, as shown in Figure 4, a key point may be to tailor the level of management depending on the purpose of AI use when identifying it as a model.

(2) Financial institutions other than (1)

Next, for small and medium-sized financial institutions that are not subject to the MRM Principles, it is likely that their MRM departments are still under development or not yet established. In such institutions, it would be difficult for the MRM department to take the lead, and the departments that develop and use AI 1st line (or 1.5 line) will likely take the primary role in managing AI. At the same time, gradually building up the 2nd line functions—such as the MRM department or departments with similar roles—and increasing their involvement would also be meaningful. At some point, transferring the AI management and approval functions to the 2nd line may be an ideal goal. However, as previously mentioned, if MRM is to be incorporated, it is necessary to carefully consider the balance.

(3) Companies outside the financial sector

Lastly, for companies outside the financial sector, establishing a dedicated MRM department is generally difficult, and the concept of “AI as a model” or “managing models” may not be widely recognized. Therefore, it is reasonable for the departments that develop and use AI to take the lead in its management. However, collaboration with other relevant departments remains an important consideration, just as it is in financial institutions. Moreover, the model management and MRM concepts used in financial institutions can also be valuable. By referring to applicable elements of MRM and gradually building a governance framework, companies can establish more robust AI governance.

Conclusion

This paper has presented an approach to AI governance in Japanese financial institutions, based on global trends. However, it is also true that simply adopting overseas practices may not be appropriate—especially for Japanese financial institutions, where MRM is still developing. It is essential to consider the resources, workload, and skill levels of MRM departments, and it would be undesirable if incorporating MRM were to impose significant restrictions on AI utilization. Even in cases where the MRM department is not involved, it may be possible to adopt certain aspects of lifecycle management similar to MRM. Ultimately, the approach to AI governance in Japanese financial institutions should be flexible as there is “no one size fits all” solution.

As was once the case with Japanese mobile phones (the so-called “Galápagos phones”), there are examples of technologies that, despite having excellent features, failed to align with global trends—such as the rise of smartphones—and were ultimately left behind by the times. Regardless of the AI governance framework Japanese financial institutions choose to build, it is important to consider global trends and the principles of MRM as reference points, while designing a governance structure that best suits their own organization.

In recent years, even in Asia—where MRM has not traditionally been emphasized—there has been growing focus on MRM. This is partly due to increased recognition by Asian regulators of the importance of model management, and to the spreading view that “AI is a model and should be managed under MRM.” In fact, some Asian authorities have published guidance clarifying the relationship between MRM and AI.⁷ We hope that Japanese financial institutions will take these global trends into account and develop AI governance frameworks that best suit their own organizations.

⁷ The Securities and Futures Commission (SFC) of the Hong Kong Special Administrative Region (SAR) of the People's Republic of China has issued the “Circular to licensed corporations – Use of generative AI language models”, while the Monetary Authority of Singapore (MAS) has published “ARTIFICIAL INTELLIGENCE MODEL RISK MANAGEMENT : OBSERVATIONS FROM A THEMATIC REVIEW”.

KPMG AZSA LLC
Financial Services Division
FS Advisory

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

© 2025 KPMG AZSA LLC, a limited liability audit corporation incorporated under the Japanese Certified Public Accountants Law and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization.