



BANGKO SENTRAL NG PILIPINAS

OFFICE OF THE DEPUTY GOVERNOR | FINANCIAL SUPERVISION SECTOR

MEMORANDUM NO. M-2026-031

To : **All BSP-supervised Financial Institutions (BSFIs)**

Subject : **Guidance Paper Entitled “Governance Principles for Artificial Intelligence (AI) in Financial Services”**

As Artificial Intelligence (AI) continues to advance and integrate into the financial sector, financial institutions must establish effective controls and safeguards against the attendant risks of AI adoption, such as data privacy concerns, bias leading to unfair and discriminatory practices, and misuse of technology, among others.

The BSP hereby issues the guidance paper entitled “Governance Principles for Artificial Intelligence (AI) in Financial Services” that is intended to guide financial institutions in formulating their own AI governance policies and risk management frameworks. It is recommended that financial institutions formally develop their own AI Governance Framework, proportionate to the nature, extent, scale, complexity, and materiality of their AI systems, as well as the institution’s overall operational complexity and risk profile, following the principles put forth in the guidance paper.

By ensuring that these principles are embedded in the overall governance frameworks of BSFIs, AI can be harnessed efficiently and safely for the benefit of various stakeholders. Ultimately, the ethical and responsible use of AI can foster trust, strengthen resilience, and promote sustainable innovation within the financial ecosystem.



LYN I. JAVIER
Deputy Governor

24 June 2026

GOVERNANCE PRINCIPLES FOR ARTIFICIAL INTELLIGENCE (AI) IN FINANCIAL SERVICES

June 2026



Introduction	1
Definition of Terms	2
Scope	2
STARS of AI Governance in Financial Services	3
AI System Lifecycle.....	5
References	8
Pertinent Laws, Rules, Regulations and Other Frameworks	9

Introduction

The Bangko Sentral ng Pilipinas (BSP) promotes innovation in the financial industry through the ethical and responsible use of emerging technologies. In recent years, Artificial Intelligence (AI), Machine Learning (ML), and data science, among others, have captured the attention of both the technology and financial sectors. While the BSP recognizes the significant potential of these technologies, it emphasizes the importance of upholding the highest standards of ethics, transparency, and accountability across the AI system lifecycle, particularly when AI is integrated into the strategic operations of financial institutions.

The BSP acknowledges that financial institutions are at different stages of AI maturity. In this regard, governance and risk management should be proportionate to the nature, extent, scale, complexity, and materiality¹ of their AI systems, as well as the institution's overall operational complexity and risk profile. The absence of a comprehensive AI governance framework introduces multiple layers of risk, including operational risk, information technology (IT), model risks, market conduct risk, reputational risk, strategic risk, and legal risk. Moreover, without clear guidelines and safeguards, AI systems may perpetuate biases, leading to unfair practices and the exclusion of individuals from access to financial products and services.

Given that AI applications transcend sectoral boundaries, a holistic and principles-based regulatory approach is essential. For the financial sector, the BSP has developed a Guidance Document that sets out the Governance Principles for the Use of AI in the Financial Sector in line with the BSP's mandate to promote financial stability. These principles provide minimum supervisory expectations for AI adoption and are intended to guide financial institutions in formulating their own AI governance policies and risk management frameworks. The principles presented in this document are non-binding and compliance is voluntary in nature. This policy shall be implemented in a manner consistent with, and without prejudice to, existing laws and regulations.

This document is structured as follows: the first section outlines the definition of terms and scope of the issuance; the second section discusses the STARS of AI governance in the financial sector; the third section maps the AI system lifecycle and identifies the risks and mitigating controls at each stage; and the fourth section presents recommended actions for financial institutions.

¹ Refer to the BSP Model Risk Management Framework for model risk tiering criteria considering factors such as materiality, impact, and sophistication following the concept of proportionality to the institution's size, operational complexity, risk profile, and extent of model use.

Definition of Terms

Artificial Intelligence (AI) is the theory and development of computer systems able to perform tasks that traditionally have required human intelligence.²

AI Policy Framework refers to an enterprise-wide policy that entities should observe and implement for activities involving AI systems.

AI System is an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.³

Ethical AI ensures compliance with ethical norms with similar ideas such as i) “do good”; ii) “be fair”, iii) “do no harm”, iv) accountability; v) competence, and vi) governance.⁴

Human Oversight is a technique used to combine AI/ML output with human decision/interaction to prevent or minimize risks arising from using AI systems, commensurate with the risks.⁵

Outsourced Service Providers (OSPs) refers to vendors with contractual arrangement with financial institution to perform designated activities on behalf of the institution on a continuing basis.⁶

Scope

Emerging technologies are covered by the existing IT Risk Management (ITRM) framework, focusing on information security, outsourcing, and project management. The forthcoming Model Risk Management (MRM) framework will tackle algorithmic fairness and model risks throughout the entire model lifecycle using a proportionate and risk-based approach. As for the broader AI risks, the Governance Principles for AI in Financial Services is a principle-based approach that institutions should consider in developing their own AI governance framework which are operationalized through the full AI system lifecycle.

The Governance Principles for the Use of AI in the Financial Sector cover financial institutions supervised by or registered with the BSP. These principles also extend to vendors or OSPs that support or engage in AI-related activities of BSP-Supervised Financial Institutions, implementing the shared responsibility model based on what each can control.⁷

These principles are not intended to limit or restrict technological advancement in the field. Rather, they aim to encourage ethical and responsible use of AI in the industry. The BSP does not endorse or prohibit specific technologies or tools. The adoption of these governance principles must align with full compliance with existing laws, rules, and regulations.

² Adapted from: Financial Stability Board, “Artificial Intelligence and Machine Learning in Financial Services,” Fsb.org, November 2017, <https://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/>.

³ Adapted from: OECD Recommendation on AI:2019; ISO/IEC 22989:2022. The definition is function-based, covering narrow AI, generative AI, and agentic AI.

⁴ Lambert Hogenhout, “A Framework for Ethical AI at the United Nations,” 2021, <https://arxiv.org/pdf/2104.12547>.

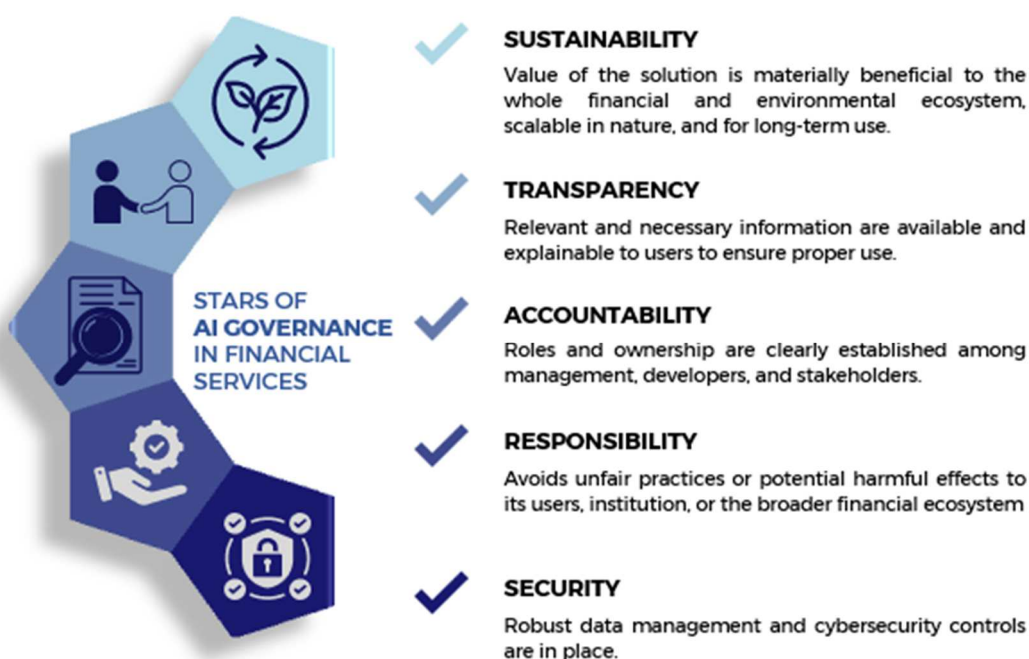
⁵ Adapted from European Union (EU) AI Act. Article 14: Human oversight. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

⁶ Refer to BSP Circular No. 899 and Circular No. 1137 for the guidelines on outsourcing.

⁷ Refer to the Financial Stability Institute (FSI) Insights on policy implementation No 63 Regulating AI in the financial sector: recent developments and main challenges. 2024. <https://www.bis.org/fsi/publ/insights63.pdf>

STARS of AI Governance in Financial Services

AI in the financial system should be governed by ethical and responsible use of technology. The BSP aims to promote sound AI governance across the financial sector through the STARS principles:



1) **Sustainability.** Ensure that the value of the solution is materially beneficial to the whole financial ecosystem and sustainable in nature. Ensuring that AI systems are designed and implemented with environmental and societal consciousness not only mitigates adverse ecological impacts but also promotes long-term economic benefits.

- a. **Environmental** – Development of AI systems should aim to minimize carbon footprints and should opt for energy-efficient implementation, which include considering power consumption (e.g., runtime metrics) and long-term usability (e.g., predicted cost over time) in the conceptualization and development phase. Environmental impact of AI can likewise be minimized through reduction of electronic waste, utilization of sustainable materials, and prioritization of green computing.⁸
- b. **Social** – Development of AI systems should focus on building a human-centered AI in which human values are at the core. Workforce development and capacity building activities should be conducted focusing on the responsible and ethical use of AI.

Activities related to AI systems should be aligned with the broader sustainability efforts of the financial sector. Financial institutions must remain agile and continuously re-evaluate the sustainability principle as AI technologies and system mature.

⁸ Also known as sustainable computing involving both computer hardware and software design and use. Source: Merritt, R., 2022. NVIDIA. <https://resources.nvidia.com/en-us-sustainable-computing/green-computing>

2) **Transparency.** Relevant and necessary information must be readily available and tailored to the knowledge and expertise of intended users or stakeholders. Ensuring transparency enables users to critically evaluate or challenge the AI's output. To avoid blind reliance on AI system recommendations, users must be equipped to answer the questions such as "how" the output was derived and "why" a recommendation was made.⁹ The following should be observed:

- a. Maintaining and updating a **centralized AI inventory** for a clear overview and organized management of AI systems;
- b. Ensuring that **users are notified** when the AI's output is being used in the product or is included in the process by incorporating this information into the product's terms and conditions or other relevant documentation. All relevant **caveats¹⁰ must be disclosed** particularly if there is a risk of misinterpretation or misuse.
- c. Ensuring **auditability¹¹ by maintaining sufficient documentation** of the design process, decision-making process, error tracking and resolution, and options in algorithm build and data sources.
- d. Ensuring that **identified risks, errors, and uncertainties have clearly defined mitigation procedures**. Regular review of all AI systems should be done for appropriate monitoring and assessment of stakeholders.
- e. Ensuring that activities, products, and/or services provided by **OSPs adhere** to the same level of transparency defined above through documentations or system performance reports.

3) **Accountability.** Accountability must be clearly defined among the management, developers, and all relevant stakeholders. While AI systems provide recommendations, humans are ultimately accountable for decisions made. The output of AI systems should not replace or diminish human responsibility. It is crucial to:

- a. Establish robust guardrails by **integrating human oversight** across the AI system lifecycle, reducing risks and enhancing accountability.
- b. Ensure proper **segregation of roles and ownership** of each component within the AI lifecycle. This includes (i) designation of roles for technology recovery and resilience, (ii) oversight functions and effective challenge of the board of directors (BOD), (iii) proper execution from senior management; and iv) clearly established mechanisms for issue identification, escalation, and remediation.
- c. Ensure that the head of the team responsible for the development and deployment of AI has enough **experience and competence** in the field of AI, data science, and finance. Likewise, ensure that individuals undertaking specific roles and tasks have the necessary qualifications, skills, and expertise.
- d. Enable the **BSP to monitor, oversee, and review** the development and deployment of the AI algorithm and technology in line with its mandate to promote financial stability.

An example of accountability is assigning specific roles for monitoring AI system outputs and building checkpoints to ensure compliance with regulatory standards.

⁹ Transparency, explainability, and interpretability are distinct characteristics with different perspectives, but ultimately support the main goal of understanding the AI system as discussed in the National Institute of Standards and Technology (NIST) AI Risk Management Framework. Source: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

¹⁰ Caveats such as what data is being used or collected, system limitations, or scope of interpretation.

¹¹ **Auditability** refers to the readiness of an AI system to undergo an assessment of its algorithms, data and design processes. Adapted from Personal Data Protection Commission (PDPC), & Info-communications Media Development Authority (IMDA). (2020). Artificial Intelligence Governance Framework Model Second Edition. <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf>

4) **Responsibility (Social Fairness).** AI system must be designed and deployed in a way that **avoids unfair practices or potential harm** to its users, institution, or the broader financial ecosystem through the following:

- a. Ensure that the training data is well-prepared, well-represented, and free from unauthorized interference. The system must **avoid harmful effects** on any demographic, particularly on minority or vulnerable groups.
- b. Ensure that **individuals can exercise their rights over their personal data** in accordance with existing regulations and laws.¹² AI systems must incorporate clear and transparent opt-in and opt-out mechanisms, as applicable, which should be available and clear to the data owners.¹³
- c. Ensure **strict compliance** with pertinent laws and regulations concerning the protection and safety of users, stakeholders, and the institutions.
- d. Ensure that the main objective of the AI system is **to contribute positively** to the individual's well-being and support the greater good of the financial system.

5) **Security.** Rigorous cybersecurity and data quality controls must be in place. Institutions must:

- a. Conduct **regular risk-based assessment** to identify potential vulnerabilities in AI systems and implement commensurate governance and security protocols.
- b. Ensure effectiveness of **cybersecurity defense mechanisms** against AI-specific threats such as adversarial attacks and data poisoning. Risk mitigation controls, systems testing, and incident response procedures should extend to AI models and systems.
- c. Regularly monitor and evaluate **data quality and model performance** to detect anomalies (such as bias and hallucination) early on. AI system's overall accuracy, reliability, and effectiveness hinge on the quality of its data sets.

¹² In accordance with the Data Privacy Act of 2012 (Republic Act 10173) and the Consumer Act of the Philippines (Republic Act 7394)

¹³ In accordance with the National Privacy Commission (NPC)'s Advisory No. 2023-01 which include guidelines on deceptive design patterns in both analog and digital interfaces, particularly in relation to data privacy and consent. This includes deception or confusion to get the data subject's consent, hidden or hard-to-access options, misleading or ambiguous language, and providing false or incomplete information.

AI System Lifecycle

The AI lifecycle includes, at the minimum, the checkpoints defined in this section.¹⁴ The AI system lifecycle is an iterative process. Risks associated with the use of AI might materialize at any point in its system lifecycle. Hence, interventions should also be implemented at any stage to mitigate or control associated risks. The outlined procedures and controls in each stage of the AI lifecycle may not be universally applicable to all use cases (e.g., evaluation procedures for Generative AI are different from predictive models). However, it is considered the best practice to identify and implement relevant controls.



PLAN

- Implement risk-based assessment considering the alignment of purpose and the AI system's intended use. It is crucial to determine whether the model is suitable for the specific application.
- Clearly delineate roles, responsibilities, and accountabilities of AI systems' owners, users, developers, validators, reviewers, deployer, senior management, and BOD.
- Ensure data availability, data quality, and feature relevance.
- Ensure transparency by maintaining accurate and complete documentation with version control.

DEVELOP

- Maintain adequate documentation of options and decisions made throughout the AI system lifecycle. Methodology should be well documented to justify the soundness of each step.
- Ensure that the AI system validation, testing, and monitoring framework facilitates identification of errors at an early stage, while maintaining fairness, usability, and relevance.
- Perform sufficient testing (e.g., back-testing¹⁵, out-of-sample¹⁶ test) when choosing the final model, version, or AI system.

¹⁴ Refer to:

i) BSP Circular No. 808 dated 22 August 2013 on the Guidelines on Information Technology Risk Management for All Banks and Other BSP Supervised institutions

ii) BSP Circular No. 900 dated 18 January 2016 on the Guidelines on Operational Risk Management

iii) BSP Circular No. 971 dated 22 August 2017 on the Guidelines on Risk Governance on the enhanced requirements on risk management systems of institutions.

iv) BSP Circular No. 989 dated 04 January 2018 on the Guidelines on the Conduct of Stress Testing Exercises provides the overarching governance standards and risk management expectations on stress testing practices in the banking industry, including the supervisory expectations on the use of models for stress testing.

¹⁵ Back-testing is a technique used to compare model or system output versus the actual historical outcomes usually used in predictive analytics.

¹⁶ Out-of-sample (OOS) is a type of back-testing which compares model or system output versus historical outcomes not used in the training data or window.

VALIDATE

- Promote integrity of results by ensuring that the validation process, aside from the usual developer validation, is conducted by a team independent from those that developed the model.
- Ensure that the results of an algorithm are reproducible, when applicable. Aside from ensuring robustness and accuracy, reproducibility also ensures transparency of AI systems.
- Challenge any decision, output, or result.
- Evaluate stakeholder and business metrics and success criteria.

DEPLOY

- Ensure transparency of deployment through official channels upon obtaining the necessary approvals, in line with accountability principle, with sufficient documentation of each step.
- Conduct sufficient testing in production or live environment, in any appropriate form.

MONITOR

- Ensure Business Continuity Plan (BCP) is kept updated and tested. This should include plans for AI-related disruptions.
- Conduct regular monitoring by documenting identified business and stakeholder metrics, continuously assessing signs that may signal the need for adjustment, redevelopment, decommissioning, or replacement.
- Prepare regression process in case of system or output irregularity. When an AI system appears not to be working as expected, a well-defined process for regression or resolution should be implemented.

Recommendation

It is recommended that financial institutions formally develop their own AI Governance Framework following the guidelines and principles put forth in this document. By ensuring that these principles are intact, AI can be harnessed efficiently and safely for the benefit of various stakeholders and relevant users. Ultimately, institutions can enhance public trust and confidence in financial products and services through the responsible and ethical use of AI.

The BSP will continue to monitor the developments in the field and, whenever necessary, issue appropriate regulations or policies to foster innovation and preserve the stability and competitiveness of the financial system.

References

- Association of Southeast Asian Nations. (2024). *ASEAN guide on AI governance and ethics*. https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf
- Financial Stability Board. (2017, November). *Artificial Intelligence and Machine Learning in Financial Services*. Retrieved from Financial Stability Board: <https://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/>
- Crisanto, J. C., Leuterio, C. B., Prenio, J., & Yong, J. (2024). Regulating artificial intelligence in the financial sector: Recent developments and main challenges (FSI Insights on policy implementation No. 63). Bank for International Settlements. <https://www.bis.org/fsi/publ/insights63.pdf>
- European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, L 1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- National Institute of Standards and Technology (NIST). 2023. Artificial Intelligence Risk Management Framework (AI RMF 1.0). <https://doi.org/10.6028/nist.ai.100-1>.
- Organisation for Economic Co-operation and Development (OECD). (2019). *Recommendation of the Council on Artificial Intelligence*. Retrieved from OECD/LEGAL/0449: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>
- Organisation for Economic Co-operation and Development (OECD). (2024, September). *Regulatory Approaches to Artificial Intelligence in Finance*. Retrieved from OECD iLibrary: <https://doi.org/10.1787/f1498c02-en>
- Personal Data Protection Commission (PDPC) & Info-communications Media Development Authority (IMDA). (2020). *Model Artificial Intelligence Governance Framework Model Second Edition*. Retrieved from <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation>
- United Nations Industrial Development Organization. (2021). *Empowering SMES through 4IR Technologies: Artificial Intelligence*. Retrieved from UNIDO Knowledge Hub: https://hub.unido.org/sites/default/files/publications/Empowering%20SMEs%20through%204IR%20Technologies_0.pdf

Pertinent Laws, Rules, Regulations and Other Frameworks

1. Republic Act (RA) 10173 or the Data Privacy Act of 2012 (DPA)
2. RA 8293 or the Intellectual Property Code (IPC)
3. RA 7394 or the Consumer Act of the Philippines
4. RA 11765 or the Financial Products and Services Consumer Protection Act (FCPA)
5. BSP Circular No. 808 – Guidelines on Information Technology Risk Management for All Banks and Other BSP Supervised Institutions
6. BSP Circular No. 982 – Enhanced Guidelines on Information Security Management
7. BSP Circular No. 1160 – Regulations on Financial Consumer Protection to implement Republic Act No. 11765, otherwise known as the "Financial Products and Services Consumer Protection Act"
8. BSP Circular No. 1140 – Fraud Management Systems and Consumer Education and Awareness
9. BSP Circular No. 989 – Guidelines on the Conduct of Stress Testing Exercises
10. BSP Circular No. 971 – Guidelines on Risk Governance
11. BSP Circular No. 900 – Guidelines on Operational Risk Management
12. BSP Circular No. 855 – Guidelines on Sound Credit Risk Management Practices
13. Department of Trade and Industry's (DTI) Artificial Intelligence Roadmap 2.0¹⁷.
14. Department Of Information and Communications Technology (DICT) and the Philippine Civil Service Commission (PSC) draft joint memorandum circular on the ethical and trustworthy use of AI in the government¹⁸
15. Project Sapiens: Thematic Review on the Use of Artificial Intelligence (AI) and Machine Learning (ML) in the Philippine Financial Services 2024

¹⁷ Department of Trade and Industry. 2024. "DTI Launches National AI Strategy Roadmap 2.0 and Center for AI Research." Department of Trade and Industry Philippines. July 3, 2024. <https://www.dti.gov.ph/archives/news-archives/dti-launches-national-ai-strategy-roadmap-2-0-center-ai-research-positioning-philippines-center-excellence-ai-rampd/>.

¹⁸ Department of Information and Communications Technology. 2024. "Principles and Guidelines for an Ethical And Trustworthy Use of Artificial Intelligence (Ai) in the Government." April 2024. https://dict.gov.ph/ictstatistics/wp-content/uploads/2024/05/DRAFT-COPY_JMC-on-the-Principles-for-an-Ethical-and-Trustworthy-Use-of-Artificial-Intelligence-AI-in-Government.pdf.