

Overview of the Artificial Intelligence Policy

Overview of the Artificial Intelligence Policy

Artificial Intelligence (AI) has become an integral part of the banking sector's digital transformation, enhancing efficiency, productivity, and service quality while supporting risk management and compliance. However, AI adoption also presents risks, particularly in relation to privacy, data security, algorithmic bias, and legal compliance. AI must therefore be used in a safe, ethical, and responsible manner.

Bank Mandiri's AI governance framework refers to the AI Governance Guidelines (Pedoman Tata Kelola AI) issued by the Financial Services Authority (Otoritas Jasa Keuangan or OJK) and the National Strategy for Artificial Intelligence 2020–2045 (Strategi Nasional Kecerdasan Artifisial or Stranas KA) issued by the National Research and Innovation Agency (Badan Riset dan Inovasi Nasional or BRIN). Both frameworks emphasize ethics, security, transparency, accountability, and data protection in the application of AI.

Bank Mandiri's AI-related policy has been approved and signed by the Board of Directors, the highest decision-making body providing endorsement. The policy provides a foundation for strengthening the governance and oversight of responsible AI adoption and for delivering sustainable benefits to customers and all stakeholders.

The Bank is currently preparing for external certification of its AI Management System in accordance with ISO/IEC 42001. The planned certification scope covers the governance and implementation of AI for customer experience, including MITA WhatsApp Chatbot and Call Center Agent Voicebot solutions. The certification is intended to provide independent external verification of the Bank's AI management practices and further strengthen AI governance, risk management, Responsible AI, and AI lifecycle management.

AI Foundation

The AI Foundation comprises the fundamental components required to support compliance, risk management, and a robust AI Lifecycle. The AI Foundation comprises the fundamental pillars that underpin effective AI governance, risk management, and the AI Lifecycle. It is supported by five key foundation areas: Data Governance, Data Privacy & Protection, Technology & Infrastructure, Cybersecurity, and AI Competency to enable reliable, compliant and responsible AI.

1. Data Governance

The Bank ensures that data used in AI systems is available in adequate quantities, of appropriate quality, integrated, and secure. Data must meet the principles of availability, integrity, and integration and be accurate, complete, consistent, reasonable, current, and unique to produce accurate, relevant, and high-quality AI outputs. Governance mechanisms, data quality and quantity requirements, and data security measures supporting AI performance are governed by the Bank's internal provisions on Data Management.

2. Data Privacy & Protection

AI systems must respect data privacy in their use and/or development and, where personal data is processed, apply the privacy-by-design principle, integrating personal data protection from the planning stage through system implementation. Each personal data processing activity must be

identified in a Record of Processing Activities (RoPA). Where processing may present a high risk, a Data Protection Impact Assessment (DPIA) must be conducted to identify, assess, and mitigate privacy risks at an early stage and to ensure compliance with applicable personal data protection regulations and the ethical principles governing AI use.

3. Technology & Infrastructure

The Bank establishes policies and manages technology and infrastructure to ensure that AI is implemented securely, efficiently, reliably, and sustainably, whether developed internally or in collaboration with third parties. AI technology and infrastructure must:

- a. apply encryption, multi-factor authentication, and role-based access controls to prevent unauthorized access and personal data leakage;
- b. provide scalable capacity to accommodate the growth of AI operational requirements and maintain high system availability;
- c. ensure that third parties meet security, privacy, and compliance standards equivalent to the Bank's internal standards;
- d. support large-scale data storage and processing for AI model training and inference;
- e. support computing technologies such as GPUs, TPUs, or other accelerators for complex AI models; and
- f. provide real-time performance monitoring, *anomaly* detection, and proactive adjustment mechanisms to maintain system stability and efficiency.

4. Cybersecurity

The Bank applies comprehensive cybersecurity principles to AI systems, protecting the cybersecurity of systems in the use and/or development of AI, to safeguard them against cyberattacks and ensure that they operate securely and reliably. Cybersecurity measures include infrastructure and application security, access and identity controls, threat monitoring and detection, incident response, and periodic security testing of AI systems. Implementation refers to the Bank's internal provisions on information technology.

5. AI Competency

AI Competency is required to establish clear roles and responsibilities and segregation of *duties across* the relevant work units. It supports effective AI oversight and management, business enhancement, and regulatory compliance, and covers both organizational and employee capabilities. The Bank develops employee capabilities in a structured and sustainable manner through training and periodic evaluation. Capability development covers the technical and non-technical aspects relevant to the AI Lifecycle, AI technologies, and their use in supporting business objectives, in accordance with the duties and responsibilities of each work unit. The Bank also ensures the availability of competencies in data governance, data analytics, cybersecurity, data privacy, audit, and risk management. In addition, training of employees on the ethical use and/or security of AI is provided through awareness programmes and/or other relevant training to strengthen employees' understanding and promote the responsible use of AI in line with the needs of their respective work units.

AI Ethics Principles

AI models and systems implemented by the Bank must observe ethical principles to ensure that AI is used reliably, transparently, securely, accountably, fairly, and responsibly.

1. Reliability

AI systems and models are designed, tested, validated, and monitored periodically to ensure the reliability, consistency, and accuracy of outputs under varying conditions of use and their alignment with established business and operational objectives. This principle is intended to minimize potential adverse impacts on users or customers and to provide a human intervention mechanism where the model cannot automatically detect or correct errors.

2. Transparency & Explainability

The Bank promotes transparency and explainability in the use of AI through its AI governance framework and AI Ethics Impact Assessment. For customer-facing applications that utilize AI solutions, such as AI-powered chatbots and voicebots, AI-generated content and interactions are appropriately labelled to ensure customers are aware when they are interacting with an AI system. As part of the AI Ethics Impact Assessment, the Bank assesses the appropriate level of transparency and explainability based on the characteristics and risk profile of each AI use case, including how AI outputs or decisions can be understood and communicated where relevant. Customers are also provided with mechanisms to escalate or continue their interaction with a human employee or agent where the AI is unable to address their needs or does not provide a satisfactory response, supporting meaningful human oversight and intervention.

3. Security & Resilience

AI systems are designed and operated to be secure, reliable, and resilient to disruptions and cyber threats, safeguarding system confidentiality, integrity, and availability. Measures include security controls, periodic testing and updates, third-party risk management, incident response, and disaster recovery mechanisms to maintain the operational continuity of AI systems.

4. Accountability

AI system governance must establish clear roles and responsibilities, including establishing clear accountability for outcomes produced by AI models/tools, and be supported by adequate documentation and audit trails throughout the AI Lifecycle. This ensures traceability of AI processes, decisions, and impacts and enables human oversight and intervention in decisions with significant consequences.

5. Data Privacy & Protection

AI systems prioritize personal data protection through privacy-by-design and privacy-by-default principles. Safeguards include lawful and limited data processing, encryption, access controls, anonymization, and protection against data leakage or misuse. The Bank also complies with applicable personal data protection laws and regulations, including the fulfilment of data subject rights and the completion of RoPA and DPIA for high-risk processing activities.

6. Inclusivity, Ethics & Fairness

AI must be applied fairly, ethically, and inclusively by avoiding potential bias in the use and/or development of AI, preventing discrimination and unreasonable bias in AI processes and decision outcomes, and not using/deploying AI systems that engage in manipulative behavior. Training data must be representative and non-discriminatory, including in relation to ethnicity, religion, race, and inter-group status, as well as individual relationships or conflicts of interest. Ethical evaluation, oversight, and human intervention are conducted on an ongoing basis to prevent adverse impacts or infringements of individual rights.

7. Human Oversight

Final decisions that have a significant impact on individuals or the Bank remain subject to human oversight. AI systems must incorporate *Human-in-the-Loop*, *Human-on-the-Loop*, and *Human-in-Command mechanisms* so that humans can monitor, evaluate, and, where necessary, intervene in AI processes.

8. Sustainability

The Bank incorporates sustainability considerations into the development and implementation of AI through its AI governance framework. As part of the AI Ethics Impact Assessment conducted during AI project development, the Bank assesses relevant sustainability considerations, including the potential long-term impacts of AI on customers, operations, and stakeholders, and its alignment with the Bank's sustainability commitments. These considerations are integrated throughout the AI lifecycle to support sustainable value creation. From a technology perspective, the Bank promotes efficient and maintainable technologies, continuous monitoring and improvement, and responsible management of technology resources to support stable and sustainable AI operations over the long term.

AI systems must support the Bank's long-term sustainability objectives by considering energy efficiency, reduced environmental impacts, social sustainability, and the prudent use of resources. Measures may include digital document processing to reduce paper consumption and the automation of administrative processes to reduce device usage time and electricity consumption. Sustainability risks may be mitigated by using efficient or optimized models, selecting cloud providers committed to green energy, monitoring and reporting the carbon footprint of AI projects, and selecting third parties committed to green computing.

AI Lifecycle

The AI Lifecycle covers planning, development, validation, implementation, monitoring, updating, and the eventual termination of AI systems. This approach is intended to maintain quality, security, accountability, and alignment with the Bank's ethical values and business objectives.

AI governance does not replace the Bank's existing model development, software development, and risk management lifecycles. Instead, it introduces additional review points and control parameters at the relevant stages.

1. AI Strategy, Development & Deployment

AI Strategy, Development & Deployment comprises a structured set of activities to ensure that AI systems are developed and deployed ethically and in accordance with applicable regulations.

The stages include:

- a. *Plan & Design;*
- b. *Collect & Process Data;*
- c. *Build Model;*
- d. *Verify & Validate;* and
- e. *Deploy.*

2. AI Operation, Monitoring & Evaluation

AI systems are evaluated and monitored periodically to ensure performance, security, data protection, regulatory compliance, and alignment with ethical principles. Monitoring covers anomaly detection, model accuracy, data protection, and human oversight. The monitoring and control process also includes defining clear boundaries for what the AI can and cannot do and limiting access to sensitive AI capabilities, based on the characteristics and risk level of each AI use case.

The Bank incorporates fairness and bias considerations into its AI governance framework through the implementation of an AI Ethics Impact Assessment during the development phase. Following deployment to production, the Bank regularly evaluates and monitors AI models and systems to ensure they remain reliable, accurate, fair, and fit for their intended purpose throughout their

lifecycle. Such evaluations consider representative data, user feedback, model performance, data quality and integrity, interpretability, and relevant fairness and bias indicators, including, where applicable, false positive and false negative rates.

Monitoring is conducted against defined acceptable thresholds and at a frequency in accordance with the Bank's Model Management requirements, complemented by continuous monitoring to identify performance degradation, anomalies, systematic failures, discriminatory outcomes, and emerging risks. Model performance is monitored against acceptable thresholds, including model accuracy, false positive rates, and false negative rates. These thresholds are set based on the characteristics of the use case, the model's risk level, and applicable model management requirements.

Where material deviations or emerging risks are identified, or where monitoring indicates that model performance has exceeded the established thresholds, including increases in false positive or false negative rates, the AI Development Unit and Product Owner may evaluate and improve the model or implement other mitigation measures in accordance with applicable requirements. The Bank also applies human oversight and intervention in accordance with the AI system's risk level, based on the relevant AI Risk Assessment, thereby keeping humans "in the loop" for critical decisions and allowing human intervention. Oversight mechanisms include Human-in-the-Loop (HiTL), Human-on-the-Loop (HoTL), and Human-in-Command (HiC).

The Bank maintains mechanisms for human intervention, escalation, and appeal for AI-enabled customer-facing solutions, allowing customers or affected parties to seek assistance, raise objections, submit complaints, or report issues where an AI-generated response or outcome is inadequate, inappropriate, or requires further review. The channels and procedures for submitting objections, complaints, or issues related to the use of AI are governed by the respective internal policies and procedures applicable to each AI-enabled product or service, including available external reporting channels such as the Mandiri Contact Center.

3. AI Improvement or Termination

Based on monitoring and evaluation results, Mechanisms to detect and correct drift or degradation of AI models over time AI models and systems may be enhanced or terminated. Termination may occur during development or operation where an AI system can no longer be improved and continues to produce outputs that do not meet the applicable requirements.