MOORE

September 2024

MOORE NEWSLETTER

AI: Model Personal Data Protection Framework

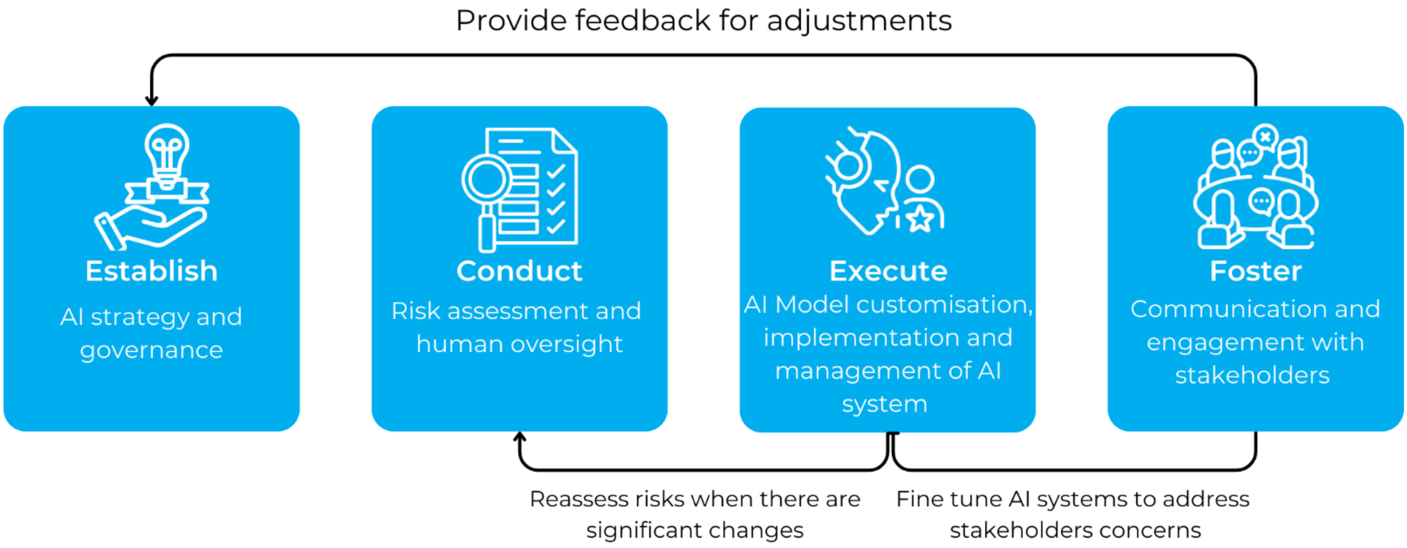
INTRODUCTION

Did you know that over 70% of organisations have experienced a data breach involving AI technologies? With the evolution of AI technologies such as chatbots, deepfake technology and IoT integration, how can we ensure that our data remains secure while harnessing its transformative potential?

In the age of AI, concerns about privacy risks are escalating, making personal data protection increasingly critical for safeguarding privacy, fostering trust and ensuring compliance with regulations. More importantly, organisations must implement robust data governance strategies that emphasise transparency and accountability, ensuring that individuals are well informed about how their data is collected, utilised and safeguarded.

In June 2024, the Office of the Privacy Commissioner for Personal Data has published the Artificial Intelligence: Model Personal Data Protection Framework with the aim of guiding Hong Kong enterprises in capturing the benefits of AI technology while brushing up on personal data privacy protection.

Overview of the model personal data protection framework



Provide feedback for adjustments

Establish
AI strategy and governance

Conduct
Risk assessment and human oversight

Execute
AI Model customisation, implementation and management of AI system

Foster
Communication and engagement with stakeholders

Reassess risks when there are significant changes

Fine tune AI systems to address stakeholders concerns

The main objective of the framework is to provide a set of recommendations on the best practices for any organisations procuring, implementing and using any type of AI systems that involve the use of personal data, including predictive AI and generative AI.

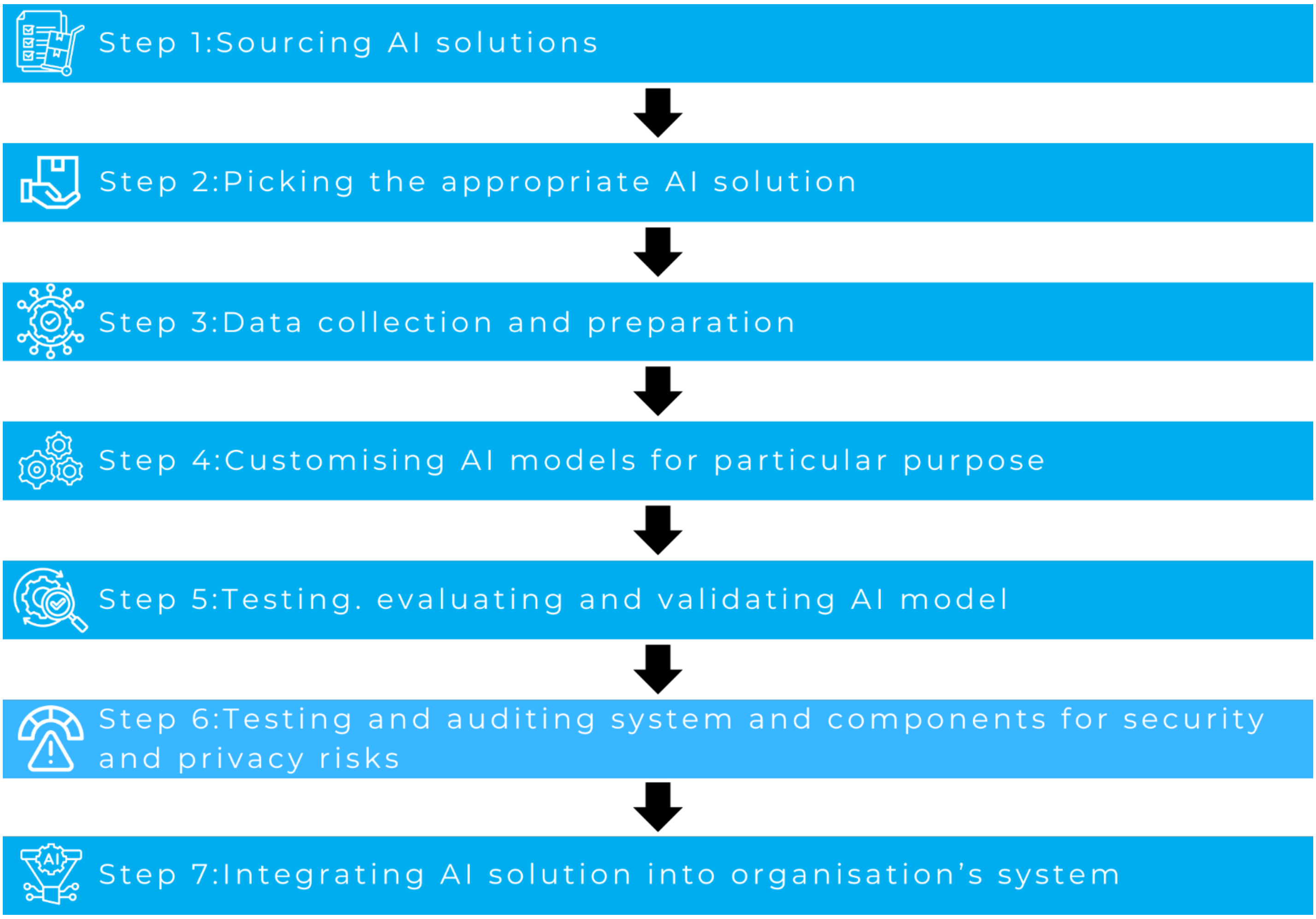
In short, the model personal data protection framework not only assists organisations in complying with requirements under the personal data privacy ordinance (PDPO), but also ensures that the three data stewardship values and the seven ethical principles for AI are adhered to and implemented:

	Data stewardship values	Ethical principles for AI
1	Being respectful	• Accountability • Human oversight • Transparency and interpretability • Data privacy
2	Being beneficial	• Beneficial AI • Reliability, robustness and security
3	Being fair	• Fairness

KEY PRINCIPLES OF THE FRAMEWORK

Phase 1: AI strategy & governance (Establish)

AI model procurement and implementation process



Step 1:Sourcing AI solutions

Step 2:Picking the appropriate AI solution

Step 3:Data collection and preparation

Step 4:Customising AI models for particular purpose

Step 5:Testing. evaluating and validating AI model

Step 6:Testing and auditing system and components for security and privacy risks

Step 7:Integrating AI solution into organisation’s system

According to the World Economic Forum’s Insight Report on Adopting AI Responsibly, the growth of the global AI market is staggering and is estimated to expand at a compound annual growth rate of nearly 40% from 2023 to 2030. This highlights the greater need to establish standards for responsible AI practices and procurement. The AI procurement process shown above assists organisations in managing and allocating different degrees of organisational involvement during each step, with the aim of engaging third parties effectively in terms of procuring AI solutions.

In addition, nine governance considerations in the procurement of AI solutions have been outlined to ensure an organisation’s adherence to regulations, risk mitigation and accomplishing strategic goals.

Purposes of using AI

Privacy and security obligations and ethical requirements

International technical and governance standards

Criteria and procedures for reviewing AI solutions

Data processor agreements

Policy on handling output generated by the AI system

Plan for continuiusly scrutinising changing landscape

Plan for monitoring, managing and maintaining AI solution

Evaluation of AI suppliers

Phase 2: Risk assessment & human oversight (Conduct)

Following an organisation’s establishment of its AI strategy and governance, this phase focuses on identifying and assessing the risks associated with AI systems, along with implementing appropriate mitigation activities.

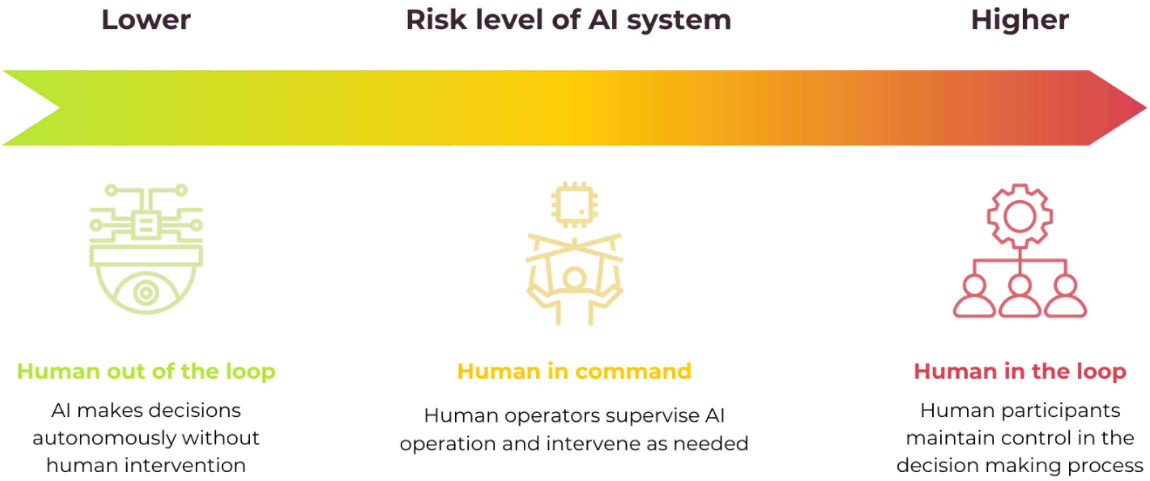
Six non-exhaustive risk factors regarding risk assessment of AI Systems have been outlined as follows:

1. PDPO requirements
2. Volume, sensitivity and quality of data
3. Data security
4. Potential impact on individuals, the organisation and the community
5. Probability, severity and duration of impact
6. Mitigation measures

The model framework also underlines the importance of human oversight as a key measure for mitigating AI-related risks. In other words, humans should be responsible for the decisions and results produced by AI.

The different risk level of AI systems corresponds to different human approaches, for instance, a high-risk AI system should adopt a ‘human in the loop” approach”, where human actors maintain control in the decision-making process to prevent and address errors by AI.

Risk based approach to human oversight



Lower

Risk level of AI system

Higher

Human out of the loop

AI makes decisions autonomously without human intervention

Human in command

Human operators supervise AI operation and intervene as needed

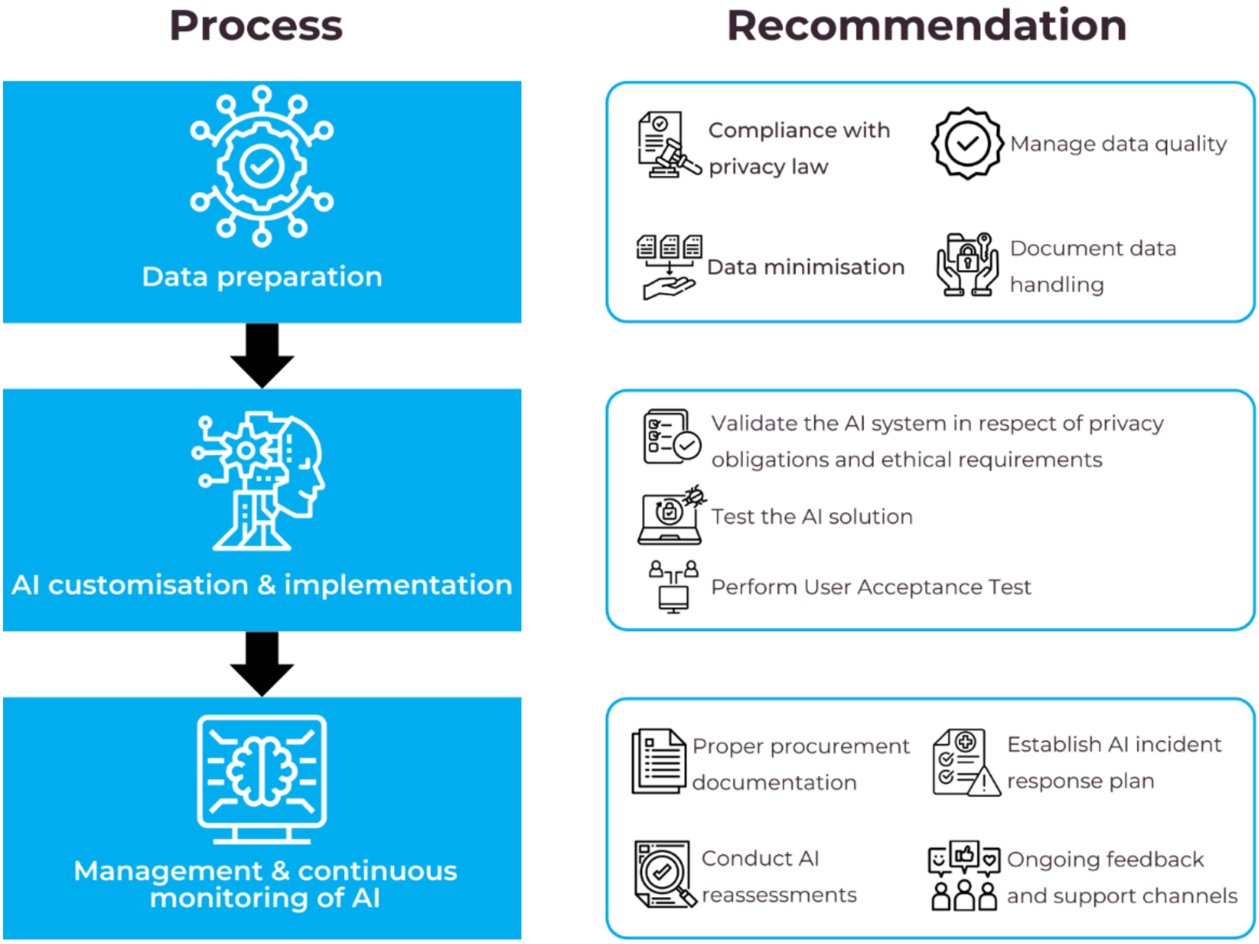
Human in the loop

Human participants maintain control in the decision making process

Examples of AI use cases that may incur higher risk:

- Real-time biometric identification
- Job applicant assessment, performance evaluation, or contract termination
- AI-assisted medical imaging analytics or therapies
- Evaluation of individual’s eligibility for social welfare or public services
- Evaluation of the creditworthiness of individuals for making automated financial decisions

Phase 3: AI model customisation, AI system implementation & management (Execute)



Process

Recommendation

Data preparation

- Compliance with privacy law
- Manage data quality
- Data minimisation
- Document data handling

AI customisation & implementation

- Validate the AI system in respect of privacy obligations and ethical requirements
- Test the AI solution
- Perform User Acceptance Test

Management & continuous monitoring of AI

- Proper procurement documentation
- Establish AI incident response plan
- Conduct AI reassessments
- Ongoing feedback and support channels

This phase emphasises the preparation of the AI solution for a specific business purpose in terms of data preparation, AI customisation & implementation, as well as management and continuous monitoring.

For data preparation of datasets for the customisation and use of AI, organisations should take

the following steps:

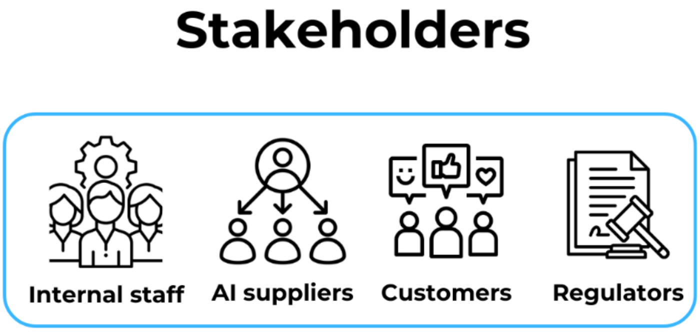
1. Adopted measure must comply with PDPO requirements
2. Data minimisation to ensure the privacy of individual's personal data is protected
3. Data quality ensures fair and unbiased results
4. Proper documentation of data handling

Following data preparation, rigorous testing and validation of AI models should be conducted with the measures below:

1. Validate the AI system in respect of privacy obligations and ethical requirements
2. Test the AI solution for errors to ensure reliability, robustness and fairness
3. Perform User Acceptance Test

Additionally, the Model Framework outlines several review mechanisms for continuous management and monitoring of AI systems, including the documentation of responses to anomalies in the datasets, risk reassessments and periodic review of the AI models.

Phase 4: Communication and engagement with stakeholders (Foster)



Organisations should maintain effective and regular communication with stakeholders, particularly internal staff, AI suppliers, individual customers, and regulators, to enhance the level of transparency and build trust.

In addition, organisations that use personal data to customise and train AI solutions should not only adhere to the PDPO but also notify data subjects:

- Purpose for which the personal data are used (e.g. for AI training or facilitating automated decision-making)
- Classes of persons to whom the data may be transferred (e.g. AI supplier)
- Organisation's policies and practices in relation to personal data in the context of customisation
- and use of AI.

The practice of "Explainable AI" has been highlighted to ensure the decisions and output of AI systems are explainable to stakeholders and build up their trust. When AI systems have the potential to significantly impact individuals, the explanations should encompass:

1. AI's role in the decision-making process, including key tasks for which it is responsible and involvement of human actors.
2. Relevance and necessity of personal data in automated or AI-assisted decision-making.
3. Key factors leading to AI system's overall and individual decisions. If explaining is not feasible, that should be clearly stated.

It is also important to note that communication with stakeholders should be in plain language that is clear and understandable and be drawn to the attention of stakeholders.

CONCLUSION

Organisations that procure, implement, and utilise AI systems handling personal data should refer to the Model Framework and its recommendations to foster trust with stakeholders and ensure compliance with the PDPO in their AI deployments. As Hong Kong moves towards stronger regulations in AI and cybersecurity, it is crucial for these organisations to stay informed about any changes. We expect the Office of the Privacy Commissioner for Personal Data to consistently monitor and update the framework in light of advancements in AI technologies and regulations. Additionally, the PCPD should actively collaborate with diverse stakeholders to encourage the ethical and responsible use of AI in Hong Kong.

REFERENCES

Paper submitted to the Office of the Privacy Commissioner for Personal Data:

https://www.pcpd.org.hk/english/resources_centre/publications/files/ai_protection_framework.pdf

Moore IT & Cybersecurity Services

Why work with us?

We are a global advisory network, with offices and member firms across the globe. Backed by our international network, we provide clients with comprehensive cybersecurity solutions and expertise from security vulnerability assessment to system penetration testing to protect them from modern cyber threats.

Our team is composed of professionals with practical and solid knowledge and experience. They are certified holders or members of professional bodies such as CISA, CISM, CRISC, CISSP, CIPP(A), CEH, OSCP and GPEN.

Our clients range from SMEs to listed companies from a wide variety of industry, and public sectors including government bureaus and authorities. Through our extensive sector knowledge, we provide comprehensive advice to suit each client’s goals.

To find out more, please contact one of our experts below:

PATRICK ROZARIO
Managing Director
T +852 2738 7769
E patrickrozario@moore.hk

KEVIN LAU
Principal
T +852 2738 4631
E kevinlau@moore.hk

The information provided is for general guidance only and should not be treated as professional advice. While we believe the content is correct at the time of publication, we cannot accept responsibility for any loss or inconvenience caused by actions taken based on the information herein. We recommend consulting a qualified advisor to ensure your specific circumstances are properly evaluated before making any decisions.