香港個人資料私隱專員公署
Office of the Privacy Commissioner
for Personal Data, Hong Kong

AI Driven Education & Cybersecurity Challenges in AI: Balancing Innovation and Data Protection

Highlights of AI Guidelines & Governance

Ms Joanne Wong
Assistant Privacy Commissioner for Personal Data (Compliance, Global Affairs and Research)

27 June 2025

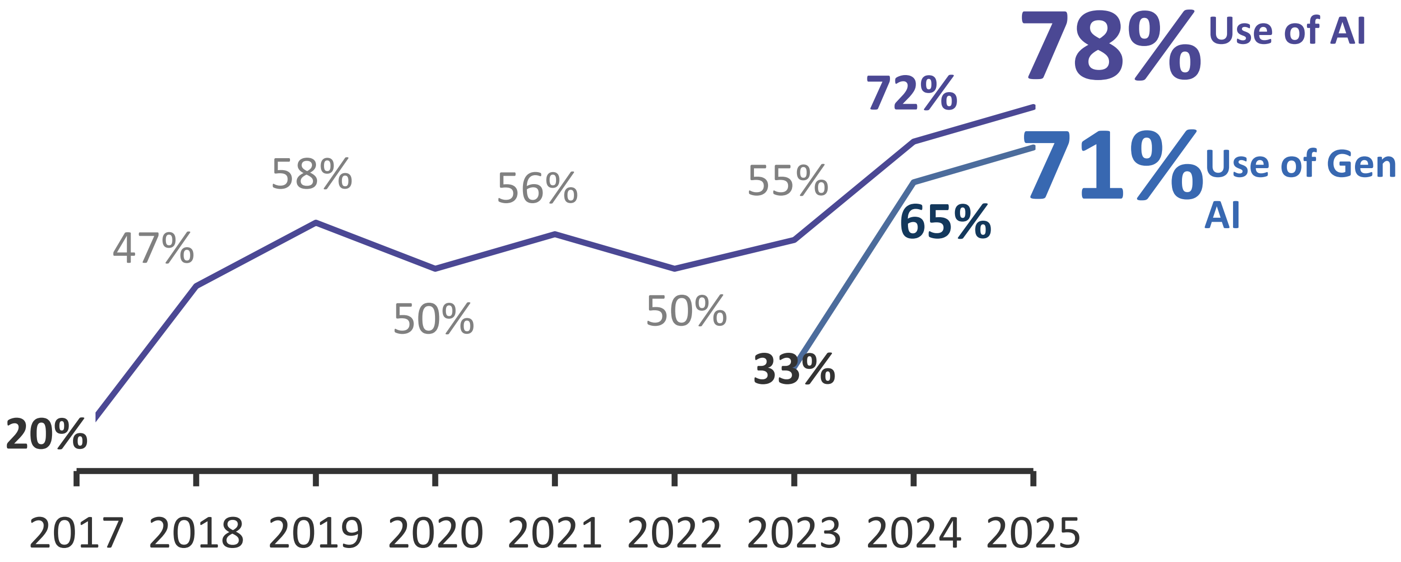
Statistics

Organisations across the globe and university students are actively using AI

📈 Global AI (including Gen AI) adoption rate has soared

Organisations that use AI in at least 1 business function

Organisations from >100 countries, 2017-2025

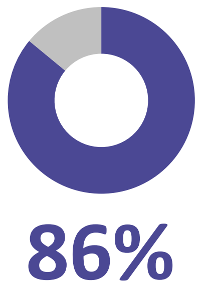


Source: McKinsey

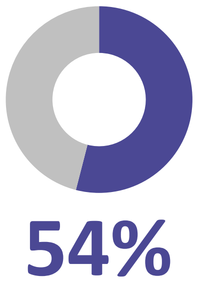
🏛 University students actively use AI

Usage of AI by university students

University students from 16 countries, 2024



Percentage of **students using AI in their studies**

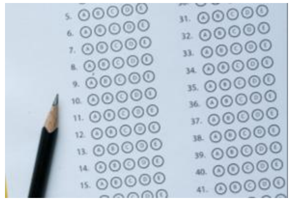


Percentage of **students using AI at least on a weekly basis**

Source: Digital Education Council

1st Area of Use – Assistance with Teaching

Examples of AI-assisted teaching



Generating educational resources

- **Course module structure** and **quiz generation**
- "**AI virtual patient**" with distinct personalities and medical histories
- **Textbook creation** using transcripts and slides from online classes

AI-powered lectures

- **Digital instructors** to deliver lectures in virtual reality and on-screen formats

Examples of privacy and ethical risks

Data breach

Teachers may **input students' personal data** into AI systems **without the knowledge of students**. Such data may be stored in **insecure systems** or used to **train AI models**, potentially leading to **unintended disclosure** in other users' conversations

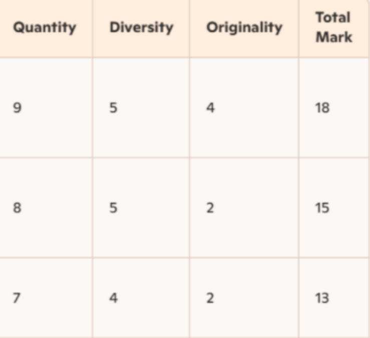
Excessive data collection

AI applications tend **to collect and retain as much data as possible**, including personal data

Source: HKU Med School; Microsoft(1); PolyU; Microsoft (2); HKUST; People.cn

2nd Area of Use – Learning Assessment

Examples of AI-assisted assessment of students' assignments



Quantity	Diversity	Originality	Total Mark
9	5	4	18
8	5	2	15
7	4	2	13

Automatic assignment evaluation

- **Grade** students' answers based on **pre-defined marking rubric**
- Provide **feedback based on** the course's **objectives and requirements**

Analysis of student performance

- Provide **insights on students' answers to enhance learning**
- Provide **personalised feedback reports** for students

Source: Microsoft; Bau; PolyU; HK01

Examples of privacy and ethical risks

Black box problem

AI models can be **so advanced** that people find it **hard to understand how their personal data would be used**, including how AI grades the answers of students

Informed and express consent

Universities should determine whether students are **informed of**, and have **expressly consented to**, the **use of AI in handling personal data**

3rd Area of Use – Personalised Learning

🏛 Examples of how AI enables personalised learning experiences for students

24/7 AI Tutor

- **Pre-class** preparation with **AI tutoring**
- **Post-class note generation** & content discussion
- **Turn academic content into podcasts**
- **Mock practice questions**
- Some AI tutors are **even accessible on instant messaging platforms**

Source: Sing Tao; HKUST; Bau; PolyU

Examples of privacy and ethical risks

Purpose of data use

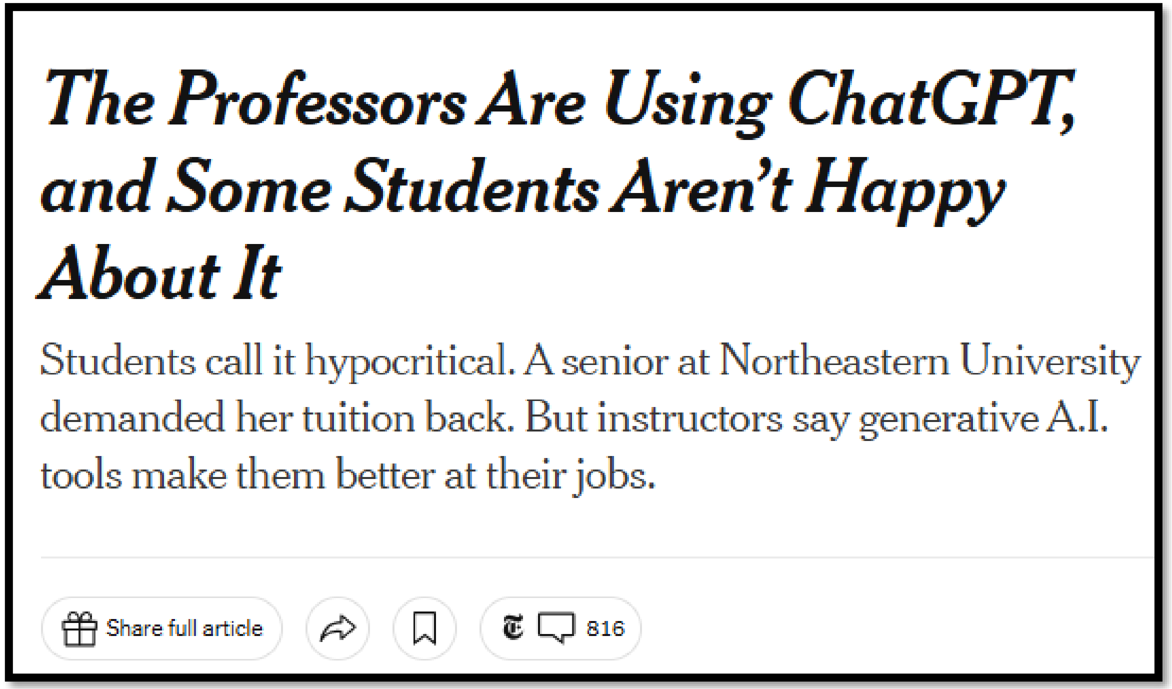
AI platforms often collect **data about students' engagement, academic performance, and even behavioural patterns.** It is important to check whether personal data is **used for any "new purpose"**

Data security

AI platforms may **store lots of data** related to students, including **learning progress, audio recordings**, and **AI conversations containing personal data**, making them **attractive targets for cyberattacks**

Controversial Example

Complaint against a professor for producing course content with Gen AI



The Professors Are Using ChatGPT, and Some Students Aren't Happy About It

Students call it hypocritical. A senior at Northeastern University demanded her tuition back. But instructors say generative A.I. tools make them better at their jobs.

Source: New York Times

Development

A student found anomalies in teaching materials

- Lecture notes contain alleged AI prompts
- Photos of unnatural body features

Accusation of double standards

- Students' use of AI banned
- Students think they paid for tuition for human teaching

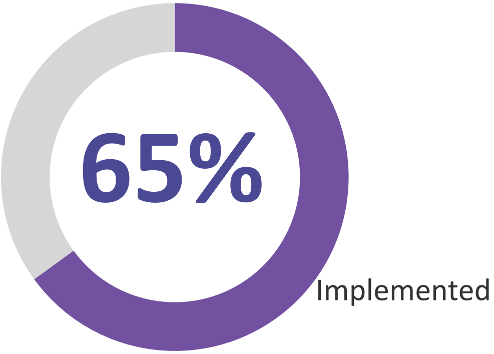
Handling by the university

- Rejected request for tuition reimbursement
- **University subsequently issued a formal AI policy**

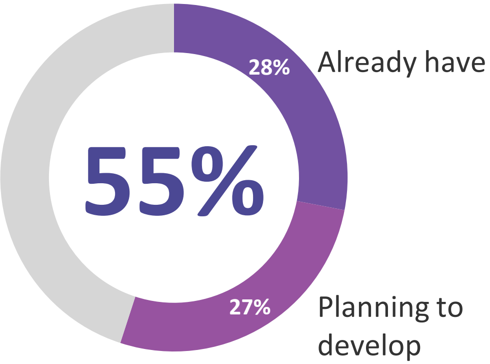
Measures

Organisations mitigate AI-related risks and refer to PCPD’s guidance

👍 Among organisations using AI in operations (2024)



Implemented data security measures



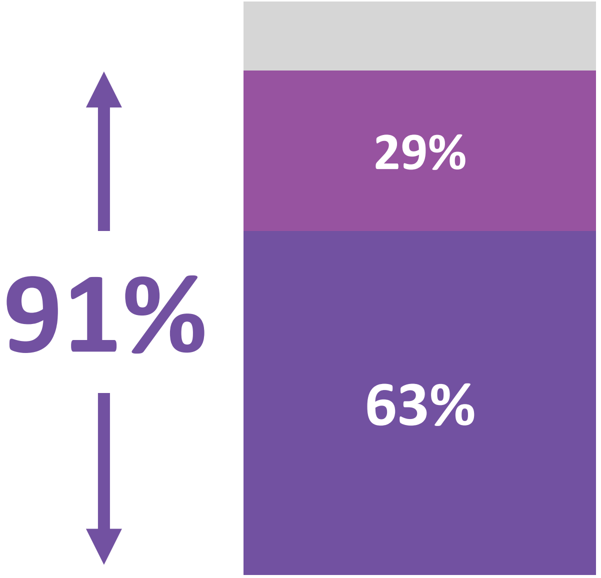
Already have or planning to develop an AI security policy

Source: PCPD-HKPC survey (2024)

📑 Many organisations collecting / using personal data through AI refer to PCPD’s guidance/advice

Reference to PCPD’s guidelines/advice on AI

Organisations that collected and/or used personal data through AI systems in PCPD’s compliance checks, 2025



Planning to make reference to PCPD’s AI guidance/advice

Have made reference

Source: PCPD

PCPD’s Guidance

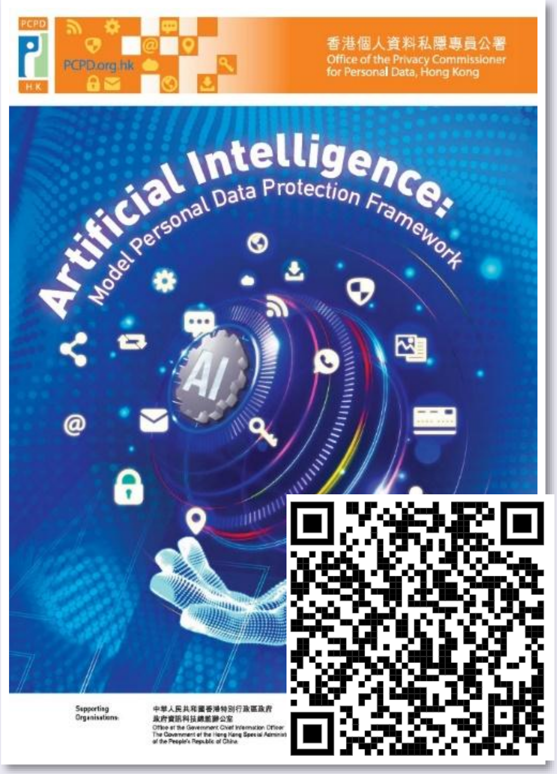
The PCPD has published different guidance in response to AI development

Organisations

Guidance on the Ethical Development and Use of
Artificial Intelligence

PCPD.org.hk
香港個人資料
Office of the Priv
for Personal Dat

Aug 2021

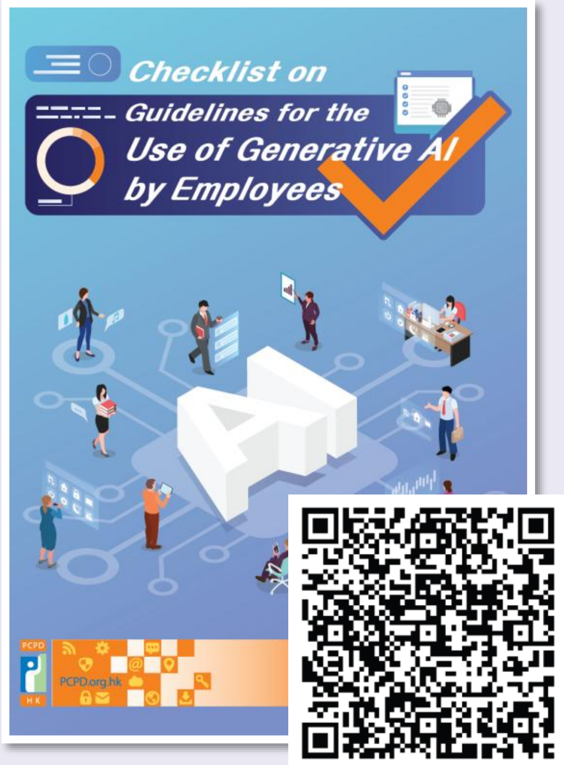

香港個人資料私隱專員公署
Office of the Privacy Commissioner
for Personal Data, Hong Kong

PCPD.org.hk

Artificial Intelligence:
Model Personal Data Protection Framework

Supporting Organisations:
中華人民共和國香港特別行政區政府
政府資訊科技總監辦公室
Office of the Government Chief Information Officer
The Government of the Hong Kong Special Administ
of the People's Republic of China

Jun 2024


Checklist on
Guidelines for the
Use of Generative AI
by Employees

PCPD.org.hk

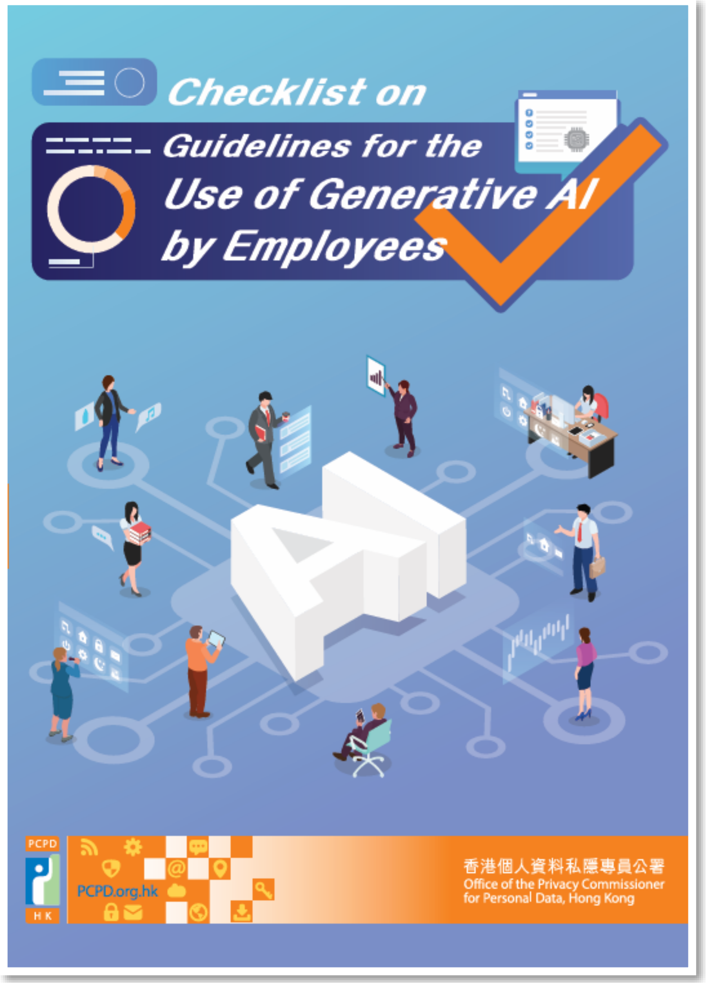
Mar 2025

Public

PCPD.org.hk

Sep 2023

Checklist on Guidelines for the Use of Generative AI by Employees



Objectives

To assist organisations **in developing internal policies or guidelines on the use of Gen AI by employees at work** while **complying with the requirements of the PDPO**

Features

Presented in the form of a **checklist**

As a matter of good practice, organisations should devise their own policies and guidelines **in alignment with their values and mission**

Recommended Coverage of Policies or Guidelines on the Use of Gen AI by Employees

01 Scope

02 Protection of personal data privacy

03 Lawful and ethical use and prevention of bias

04 Data security

05 Violations of policies or guidelines

01

Recommended Coverage of Policies or Guidelines on the Use of Gen AI by Employees – Scope

Scope	Details
Permitted tools	**Clearly specify the Gen AI tools and applications that are permitted within the organisation,** for example: - Publicly available Gen AI tools or applications - Internally developed Gen AI tools or applications
Permissible use	**Clearly specify the tasks or activities for which employees can use Gen AI tools,** for example: - Drafting - Summarising information - Creating textual, audio and/or visual content
Policy applicability	Specify if the policy applies to the **whole organisation**; **specific departments**; **specific ranks**; and/or **specific employees**

02

Recommended Coverage of Policies or Guidelines on the Use of Gen AI by Employees – Protection of personal data privacy

Permissible types and amounts of input information

Provide clear instructions on:

✓ **The types and amounts of information that can be inputted into the Gen AI tools**

✗ **The types of information that cannot be inputted**

Permissible use of output information

Provide clear instructions on the **permissible purposes** for using the information (including personal data) generated by Gen AI tools, and whether, when and how such personal data should be anonymised before further use

Permissible storage of output information

Require employees that the information generated by Gen AI tools be deleted according to the organisation's **information management policy** and **data retention policy**

Compliance with other relevant internal policies

Ensure that **the policy on the use of Gen AI is aligned with the organisation's other relevant internal policies**

03

Recommended Coverage of Policies or Guidelines on the Use of Gen AI by Employees – Lawful and ethical use and prevention of bias

Unlawful acts

Specify that employees shall not use Gen AI tools for unlawful or harmful activities

Emphasise the importance of employees acting as human reviewers

Accuracy and verification

Emphasise the need for employees to **verify the information provided by AI**

Prevention of bias and discrimination

Alert employees to the possibility that AI-generated output can be **biased and discriminatory**

Set out the **correction and reporting mechanisms**

Watermarking / labelling

Provide clear instructions on **when and how** AI-generated output should be **watermarked or labelled**

04

Recommended Coverage of Policies or Guidelines on the Use of Gen AI by Employees – Data security

Permitted devices

Specify **the devices** on which employees are permitted to **access Gen AI tools**

Permitted users

Specify the **permitted employees** of Gen AI tools

User credentials

Require employees to use **unique and strong passwords** along with **multi-factor authentication**

Security settings

Require employees to maintain **stringent security settings**

Response to AI incident and data breach incident

Require employees to **report AI incidents according to the organisation's AI Incident Response Plan**

05 Violations of policies or guidelines and practical tips

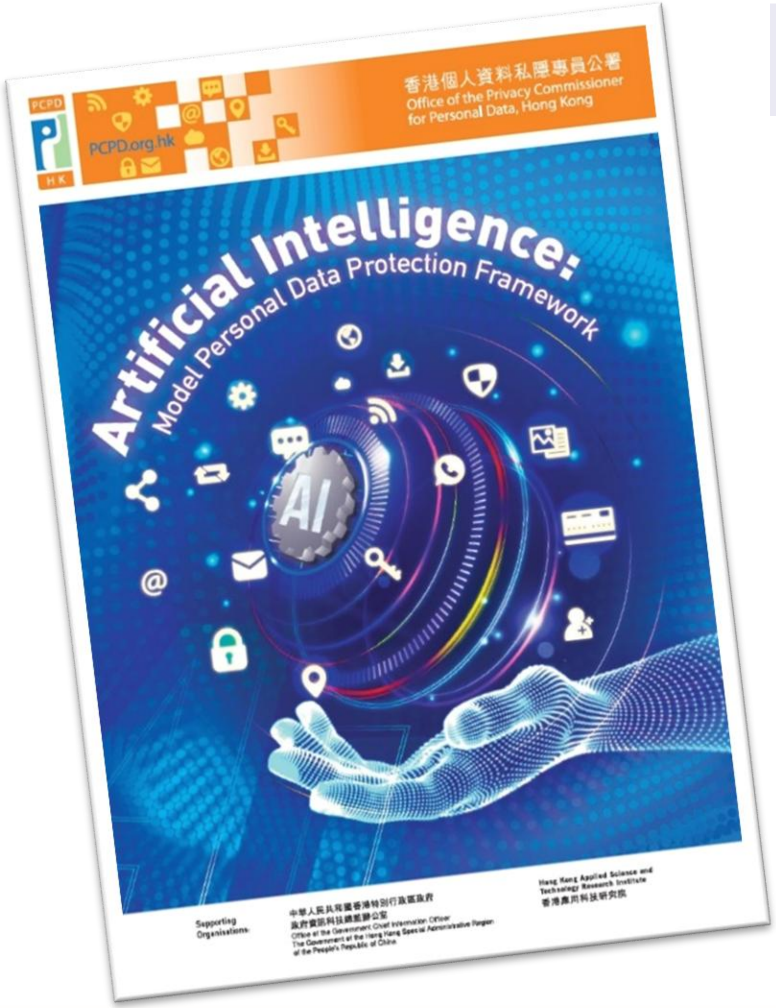
Consequences of violation

- **Specify** the possible **consequences** of employees' violation of the policies or guidelines on the use of Gen AI
- Refer to the PCPD's **Model Framework for recommendations** on establishing **Gen AI governance structure and measures**

Practical tips on supporting employees in using Gen AI tools

- Transparency
- Training and Resources
- Provide Support Team
- Feedback Mechanism

Artificial Intelligence: Model Personal Data Protection Framework



PCPD HK PCPD.org.hk 香港個人資料私隱專員公署 Office of the Privacy Commissioner for Personal Data, Hong Kong

Artificial Intelligence: Model Personal Data Protection Framework

Supporting Organisations:

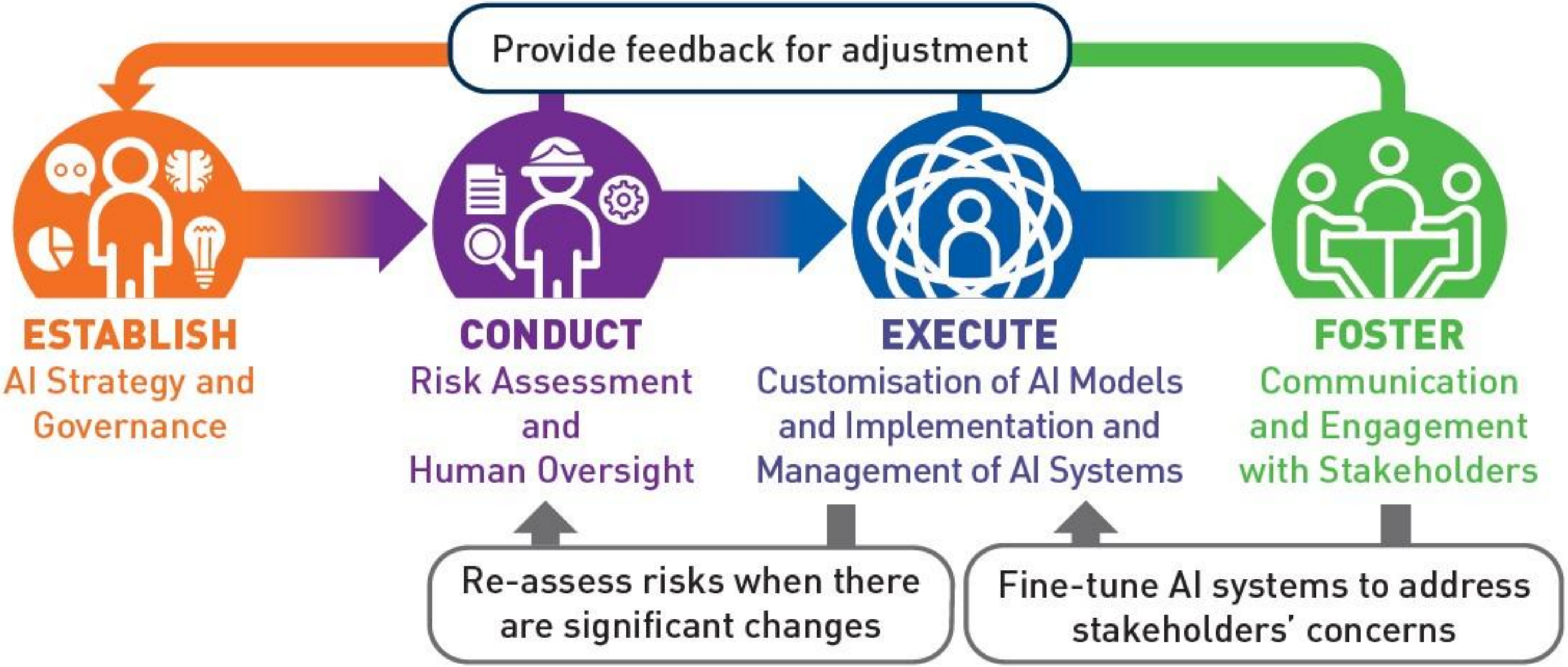
中華人民共和國香港特別行政區政府 政府資訊科技總監辦公室
Office of the Government Chief Information Officer
The Government of the Hong Kong Special Administrative Region of the People's Republic of China

Hong Kong Applied Science and Technology Research Institute
香港應用科技研究院

✅ Benefits

- Assist organisations in complying with the requirements of the Personal Data (Privacy) Ordinance
- Ensure **AI Security**
- Increase **competitiveness**
- Provide a set of **recommendations on AI governance** and the **best practices** for organisations **procuring, implementing and using any type of AI systems, including generative AI**, that involve the **protection of personal data privacy**

Model Personal Data Protection Framework



Provide feedback for adjustment

ESTABLISH
AI Strategy and Governance

CONDUCT
Risk Assessment and Human Oversight

EXECUTE
Customisation of AI Models and Implementation and Management of AI Systems

FOSTER
Communication and Engagement with Stakeholders

Re-assess risks when there are significant changes

Fine-tune AI systems to address stakeholders' concerns

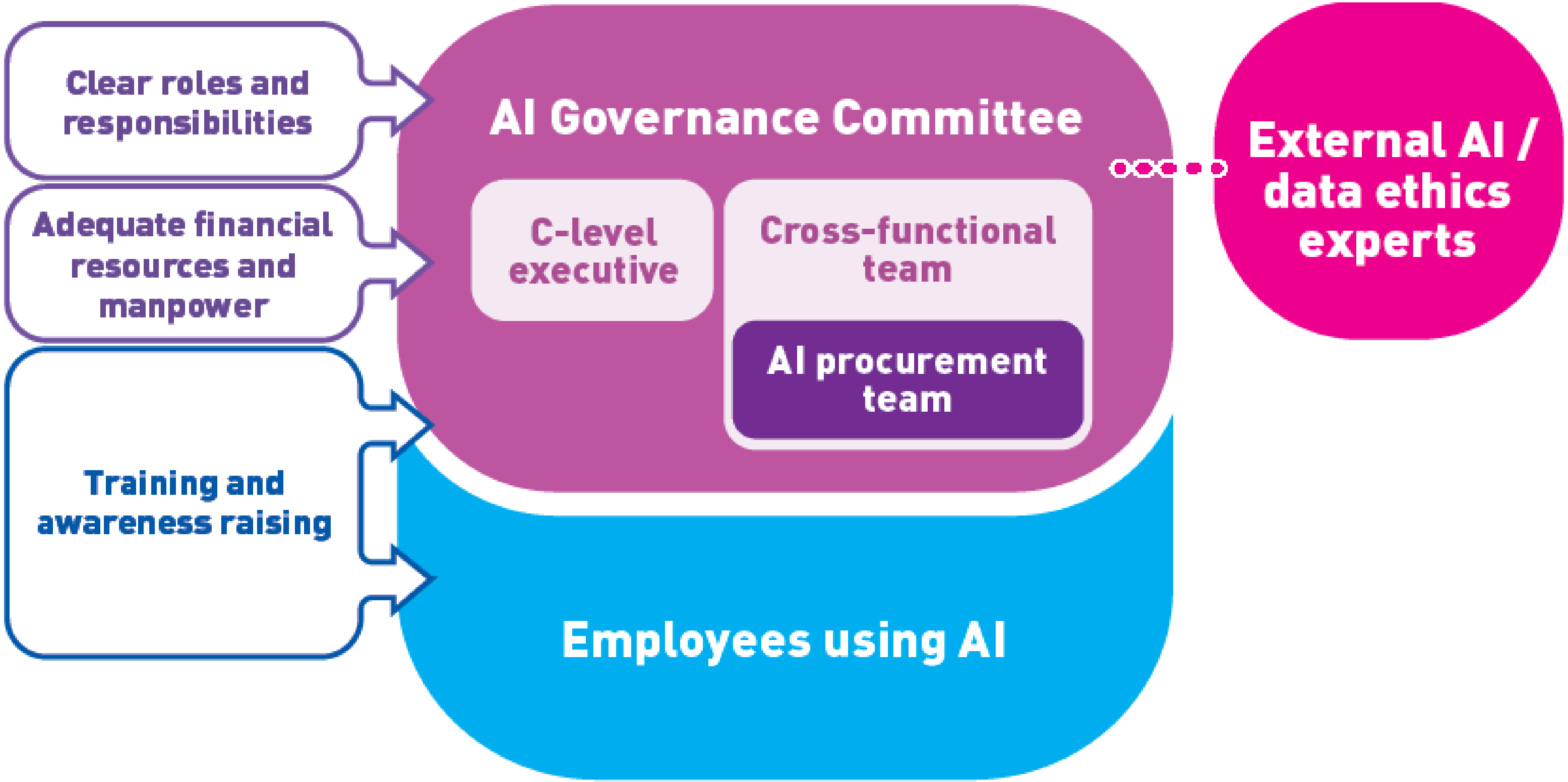
Formulate AI Strategy and Governance

9 governance considerations

- Purpose(s) of using AI
- Criteria and procedures for reviewing AI solutions
- Plan for continuously scrutinising changing landscape
- Privacy and security obligations and ethical requirements
- Data processor agreements
- Plan for continuously monitoring, managing and maintaining AI solution
- International technical and governance standards
- Policy on handling output generated by AI system
- Evaluation of AI supplier

Formulate AI Strategy and Governance

Governance structure



Clear roles and responsibilities
Adequate financial resources and manpower
Training and awareness raising
AI Governance Committee
C-level executive
Cross-functional team
AI procurement team
External AI / data ethics experts
Employees using AI

Conduct Risk Assessment and Human Oversight

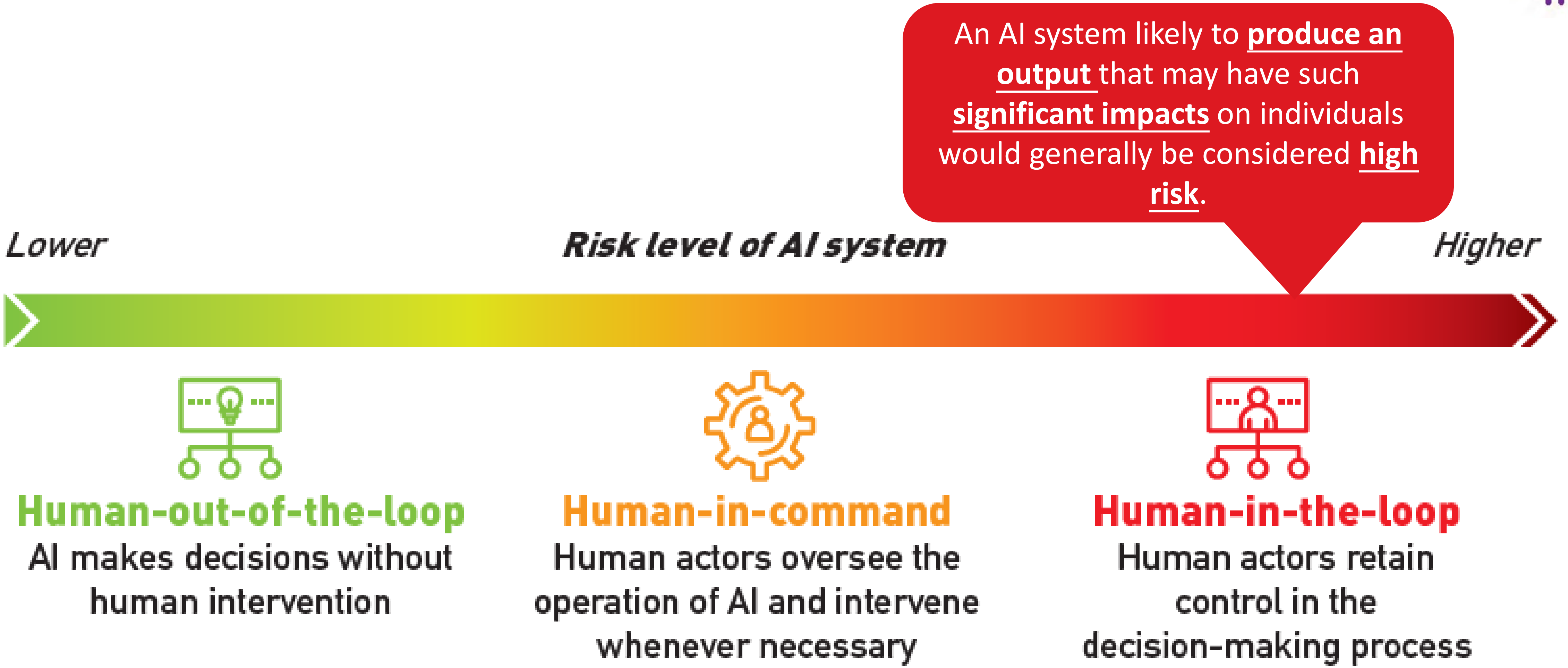
Process of risk assessment

1. **Conduct risk assessment** by a cross-functional team
2. **Identify and evaluate the risks** of the AI system
3. **Adopt appropriate risk management measures** commensurate with the risks

Risk type	Factors to be considered
Privacy Risks	• The **allowable uses** of the data for customising procured AI solutions and / or to be fed into AI systems to make decisions • **Volume** of personal data • **Sensitivity** of data involved • The **security** of personal data used in an AI system
Ethical Risks	• The **potential impacts** of the AI system on the affected individuals, the organisation and the wider community • The **probability that the impacts** of the AI system on individuals **will occur**, as well as the **severity** and **duration** of the impacts

Conduct Risk Assessment and Human Oversight

Risk-based approach to human oversight



An AI system likely to **produce an output** that may have such **significant impacts** on individuals would generally be considered **high risk**.

Lower ***Risk level of AI system*** *Higher*

Human-out-of-the-loop
AI makes decisions without human intervention

Human-in-command
Human actors oversee the operation of AI and intervene whenever necessary

Human-in-the-loop
Human actors retain control in the decision-making process

Execute Customisation of AI Models and Implementation and Management of AI Systems

Process	Selected Recommendations	Example
1 Data Preparation	Compliance with the requirements of privacy law Minimisation of personal data involved Management of data quality Proper documentation of handling of data	• A **school** is liaising to purchase **a third-party developed Gen AI tool** which will be **customised** to help teachers prepare teaching materials, homework and exam questions, etc. • The school may find it necessary to use its internal materials to customise the tool • However, the school should note that the use of personal data, such as **students' names, class and class number**, may not be necessary for customising the AI tool

Execute Customisation of AI Models and Implementation and Management of AI Systems

Process	Selected Recommendations	Examples
2 Customisation and implementation of AI	☑ Conduct rigorous testing and validation of reliability, robustness and fairness on the AI model Consider compliance issues based on the hosting of AI solution (‘on-premise’ or on a third-party cloud) prior to integration → Ensure system security and data security	- A **school** is **procuring** a **third-party developed AI chatbot** to help its employees **draft school notices, teaching materials** and handle **clerical works** - Considering whether the chatbot is hosted on-premise or on cloud, **employees of the school should be cautioned of the risks of entering personal data** and/or **other confidential data** into the chatbot

Execute Customisation of AI Models and Implementation and Management of AI Systems

Process

3

Management and Continuous Monitoring of AI

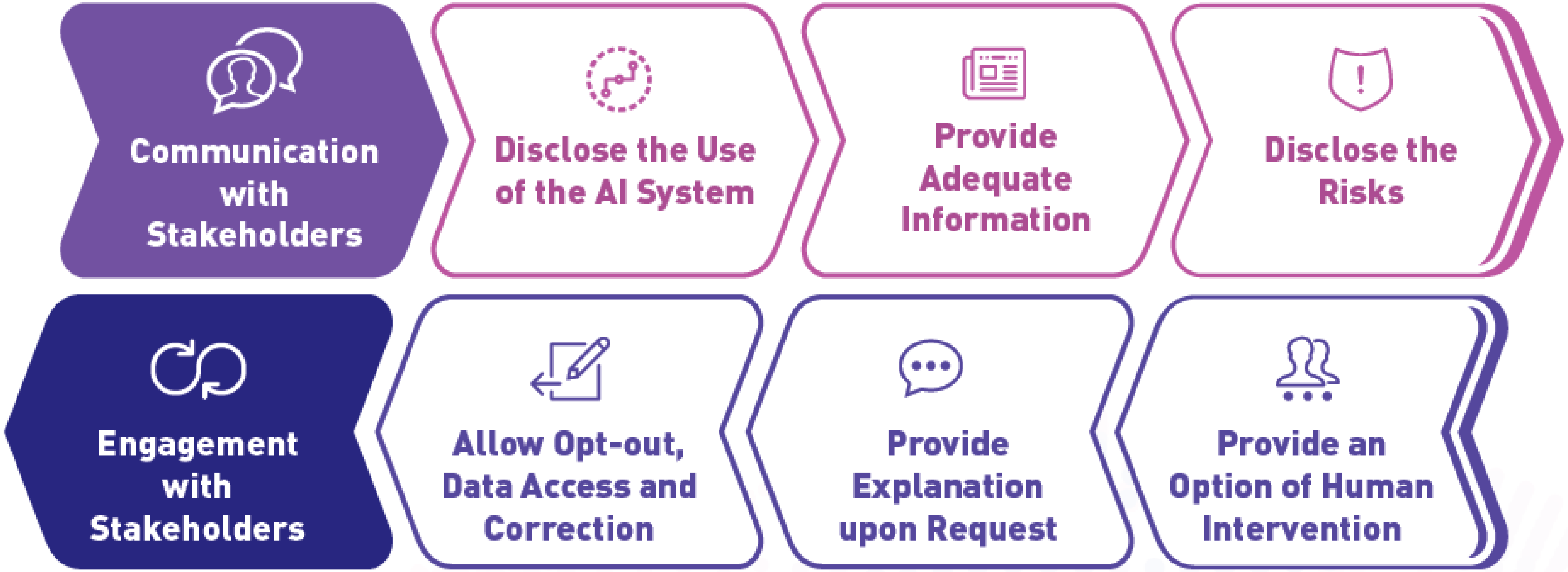
Selected Recommendations

- Maintain proper documentation
- Conduct periodic audits
- Establish an AI Incident Response Plan
- Consider incorporating review mechanisms as risk factors evolve →

Examples

- **Human oversight** should aim to **prevent and minimise the risks posed by AI to individuals**. Personnel who exercise human oversight should:
  - **Understand the capacities and limitations of the AI system**, to the extent possible;
  - **Avoid the tendency to over-rely** on the **output** produced by AI;
  - **Correctly interpret and assess** the output **produced by AI**; and
  - **Flag and, where appropriate, disregard, override or reverse the output** produced by AI if it is **abnormal**

Foster Communication and Engagement with Stakeholders



Communication with Stakeholders

Disclose the Use of the AI System

Provide Adequate Information

Disclose the Risks

Engagement with Stakeholders

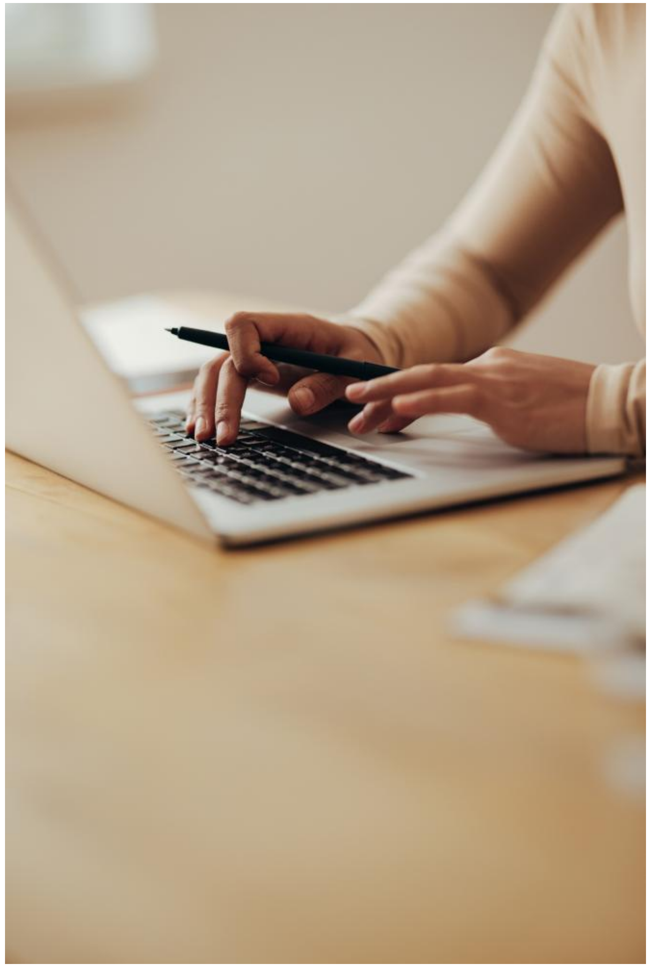
Allow Opt-out, Data Access and Correction

Provide Explanation upon Request

Provide an Option of Human Intervention

Enhance AI Security = Enhance Competitiveness

Universities can refer to the steps below to enhance AI security



- Download AI-related guidelines from the PCPD
- Evaluate organisation’s strategies on the use of AI
- Develop relevant governance strategies and framework, draft relevant internal policies or guidelines
- For enquiries, please contact PCPD’s “AI Security Hotline” (2110 1155)
- Join PCPD’s seminars and request internal training seminars

Thank you!

www.pcpd.org.hk

communications@pcpd.org.hk

Please follow us!