PCPD.org.hk

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

政府資助機構
Government-funded organisation

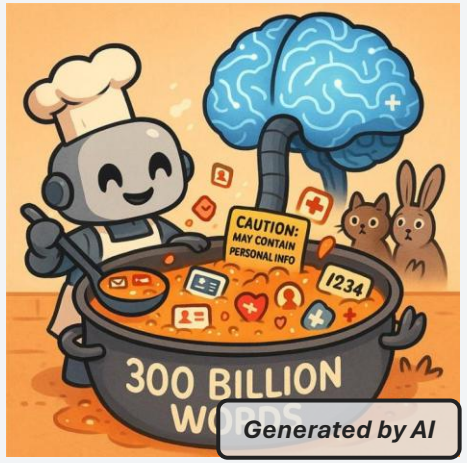
Navigating Data Privacy Risks in the Use of AI in Higher Education

Alex Chan
Assistant Privacy Commissioner for Personal Data (Compliance, Global Affairs and Research)

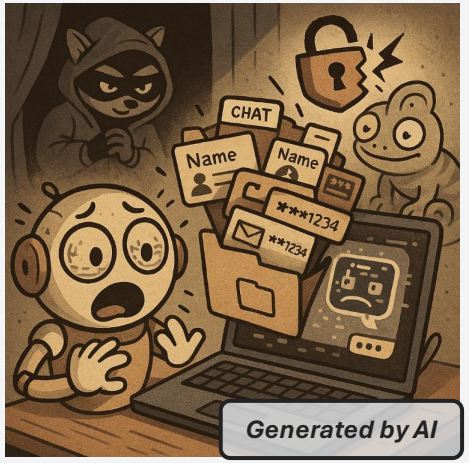
30-June-2026

守護私隱・改革創新
Protecting Privacy · Embracing Innovation

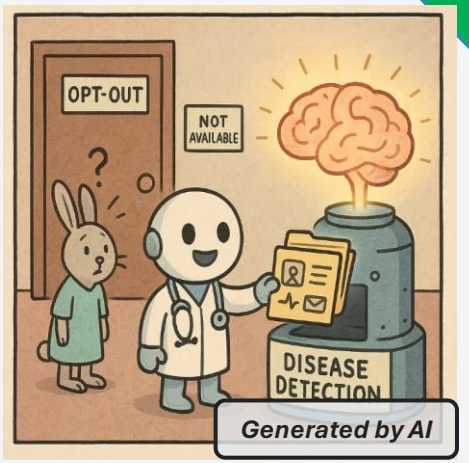
Privacy Risks related to the use of AI In Education



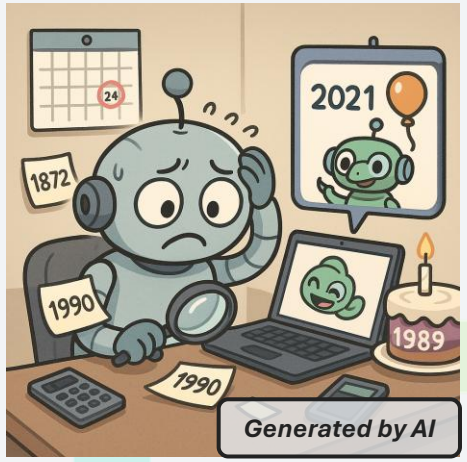
1) Excessive Data Collection



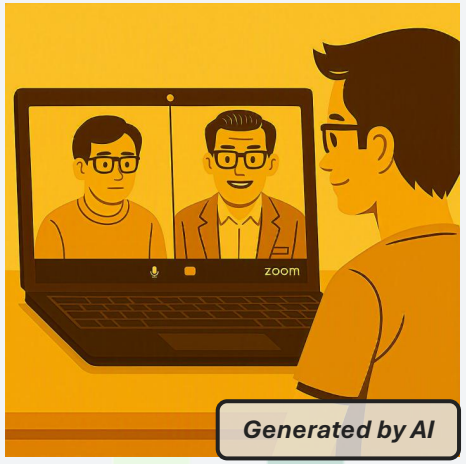
2) Data Breach



3) Use of Data



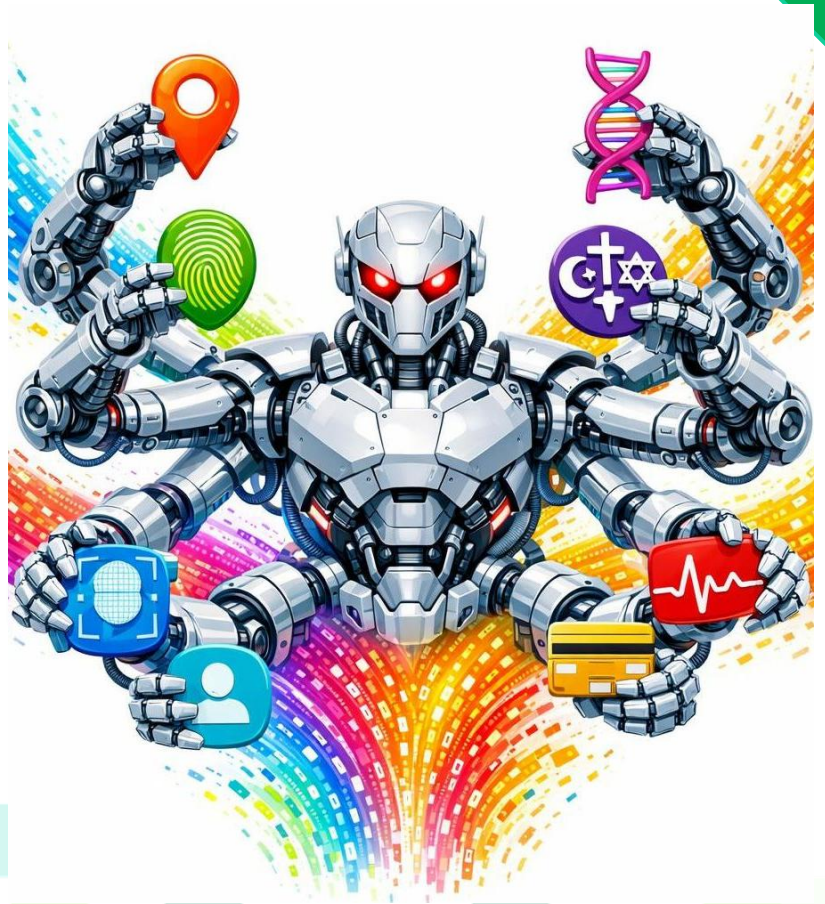
4) Data Accuracy



5) Deepfake / AI Scam

1) Excessive Data Collection

- AI chatbot applications all **collect some form** of user data.
- Most AI chatbots collect users' **locations** and even other **sensitive information**, such as **ethnicity, religion, or biometric data**.



See how well your AI chatbot knows about you

Try this Prompt: Please describe what you know about me in 100 words

Preparing a concise description of what I know about you

You are Alex Chan, a cybersecurity and data privacy leader based in Hong Kong. You are a bilingual professional who blends technical expertise with public policy and advocacy. `Memory`

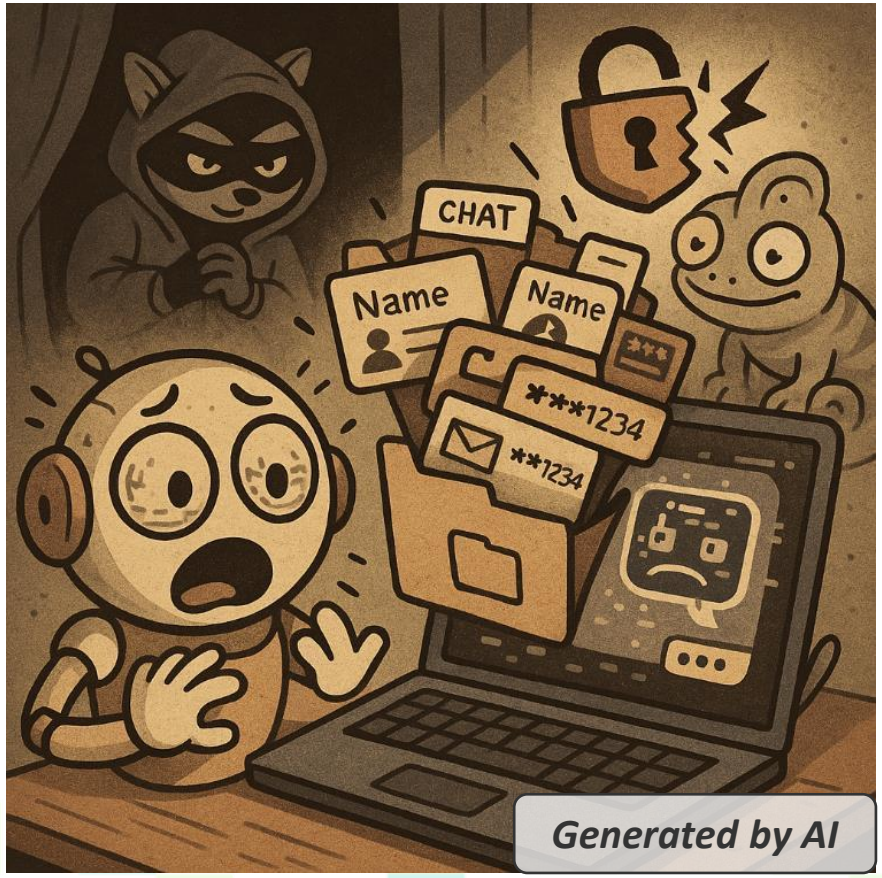
Professional Expertise

You specialize in data protection, AI governance, and incident response, navigating frameworks like the PDPO. You are highly technical, utilizing Python for security automation and actively managing advanced self-hosted home networks, including Wi-Fi 7 mesh setups `Memory +1`

Personal Interests

Outside of work, you are a dedicated Tottenham Hotspur football supporter. You also enjoy Cantopop music and are a price-conscious consumer who carefully researches the best tech deals, particularly within the Apple ecosystem. `Memory +1`

2) Data Breach



資料來源: Search Engine Journal

Overview

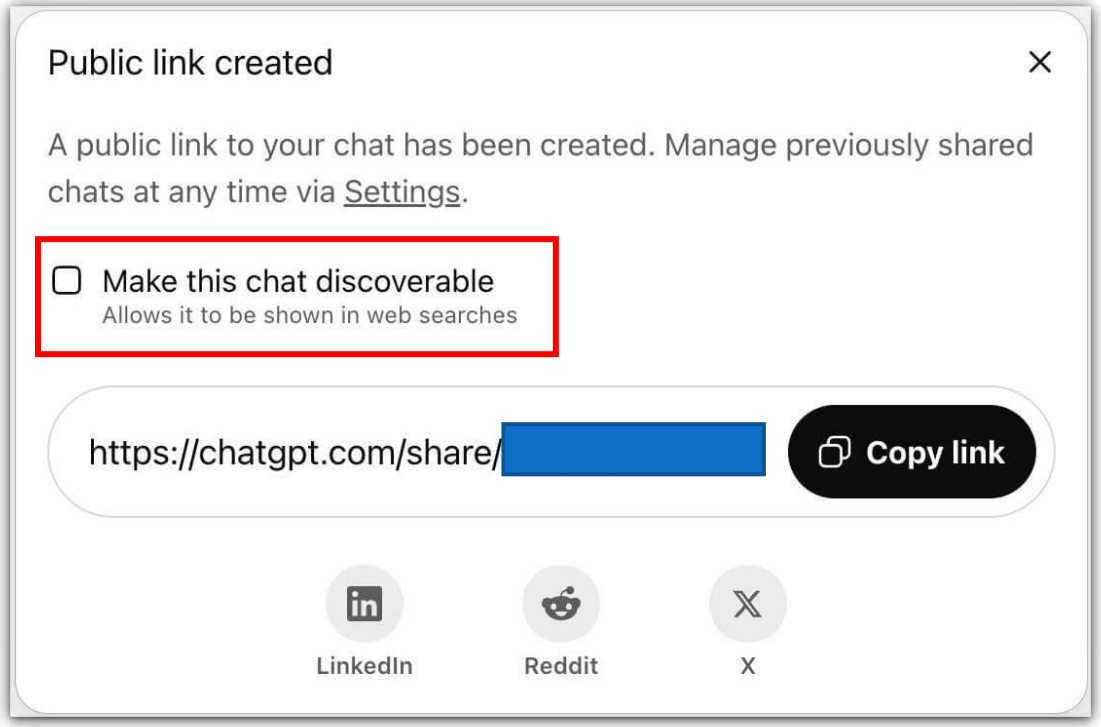
- AI systems, such as chatbots, are commonly used to handle personal and sensitive data. If they are **poorly configured or designed**, this could result in **data disclosure**.

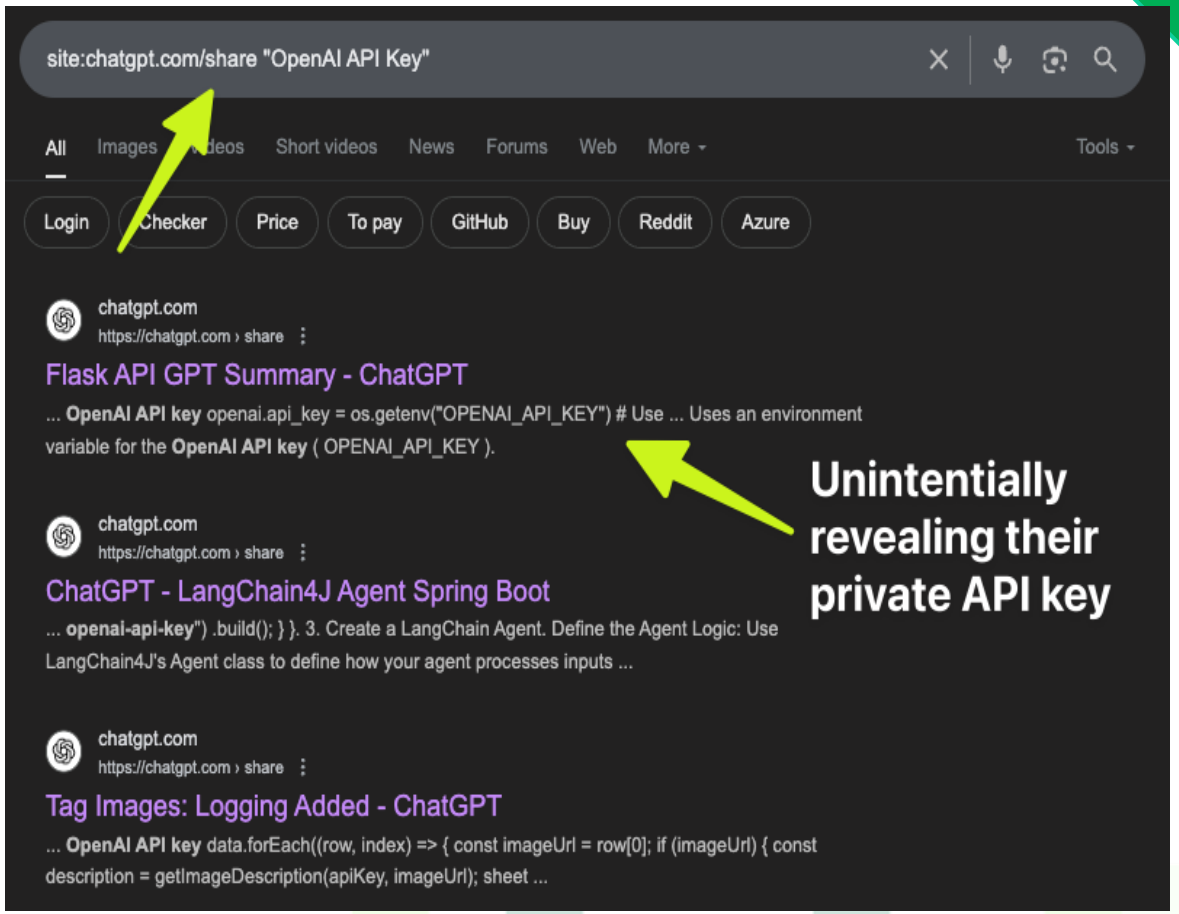
Example

- At the beginning of 2025, chatbots began allowing users to choose to make their chat **content indexable by search engines**
- As a result, in July, media found thousands of chatbot conversations indexed by online search engines, including **names, bios, user emotional states, confidential work content, and more**

Added

2) Data Breach





Source: x.com

Added

2) Data Breach

BBC

Register Sign In

Hundreds of thousands of Grok chats exposed in Google results

21 August 2025

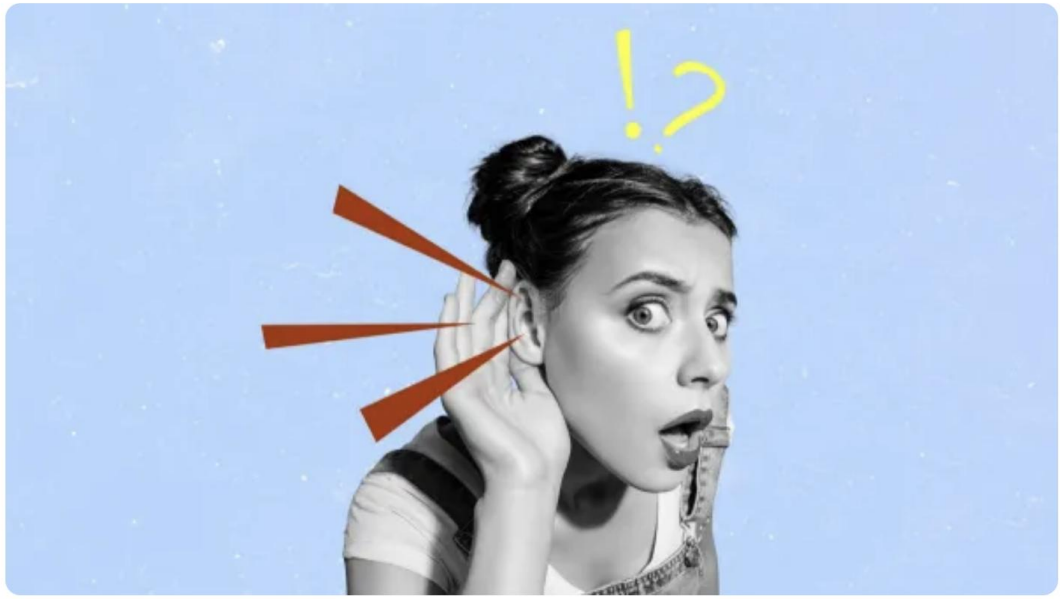
Share Save Add as preferred on Google

Liv McMahon
Technology reporter

Malwarebytes | Labs

AI chat app leak exposes 300 million messages tied to 25 million users

by Pieter Arntz | February 9, 2026



3) Use of Data

- Using a data subject's personal information to train AI **without their knowledge or without obtaining their consent**.
- American artist **Taylor Swift** has filed three trademark applications in the United States covering **her voice and appearance** to address challenges posed by **AI imitation**. This has become an example of celebrities using trademarks to protect their rights and guard against the misuse of AI.

Changed

4) Data Accuracy

The Guardian

Int

News | Opinion | Sport | Culture | Lifestyle

World | Europe | US news | Americas | Asia | Australia | Middle East | Africa | Inequality | Global development

ChatGPT

This article is more than 1 year old

Norwegian files complaint after ChatGPT falsely said he had murdered his children

Arve Hjalmar Holmen, who has never been accused of or convicted of a crime, says chatbot's response to prompt was defamatory

Arve Hjalmar Holmen said ChatGPT gave a 'completely false' response to his prompt: 'Who is Arve Hjalmar Holmen?' Photograph: Arve Hjalmar Holmen

Source: Guardian

Overview

AI chatbots are prone to producing responses **containing false information** because they are built on models that predict the next most likely word in a sentence. This can result in factual errors and wild assertions, but the plausible nature of the responses can **trick users** into thinking that what they are reading is **100% correct**

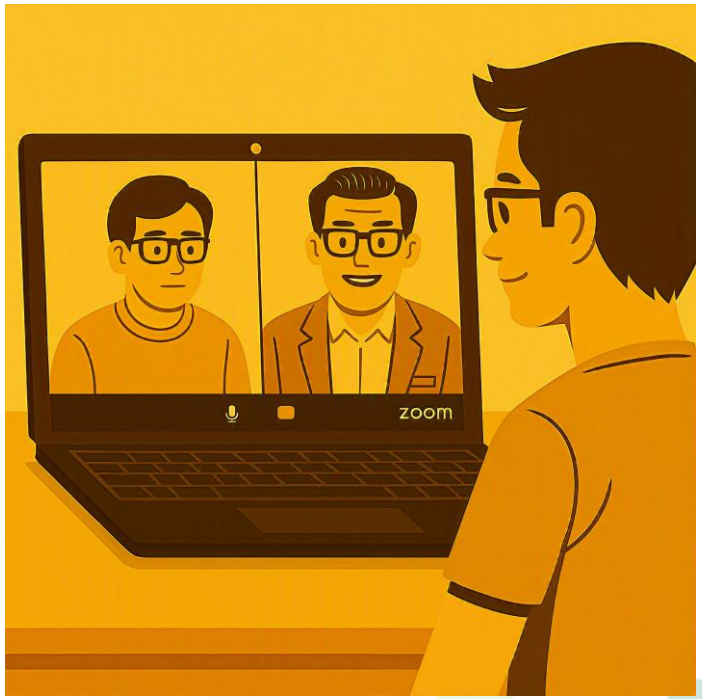
Example

- A Norwegian man, Arve Hjalmar Holmen, filed a complaint after ChatGPT **falsely said** he had murdered two of his children and **been jailed for 21 years**.
- The chatbot **mixed real personal details with a fabricated murder allegation**, making the output look credible but false.
- The complaint argues this violated **GDPR accuracy rules** and **filed a complaint** to the **Norwegian Data Protection Authority**.

5) Deepfake / AI Scam

Main causes of the growing trend in Deepfake fraud:

- **Technology Advancement**
  - ❑ Generative AI tools can now produce high-quality videos, audio, and voice clone using very minimal input data (e.g., just a few seconds of audio or video).
- **Low entry barriers (technical and financial)**
  - ❑ Even individuals without technical expertise can create deepfake videos or voice recordings at low cost.
- **Challenges in detection**
  - ❑ The accuracy of humans in detecting high-quality deepfake videos is approximately 60%.
  - ❑ Source (PNAS, Arxiv)
- **Organisations lack training or measures to address deepfake-related fraud.**



Voice Cloning Demo

Authentic human recording: "Good morning, welcome to today seminar" (3 Sec)

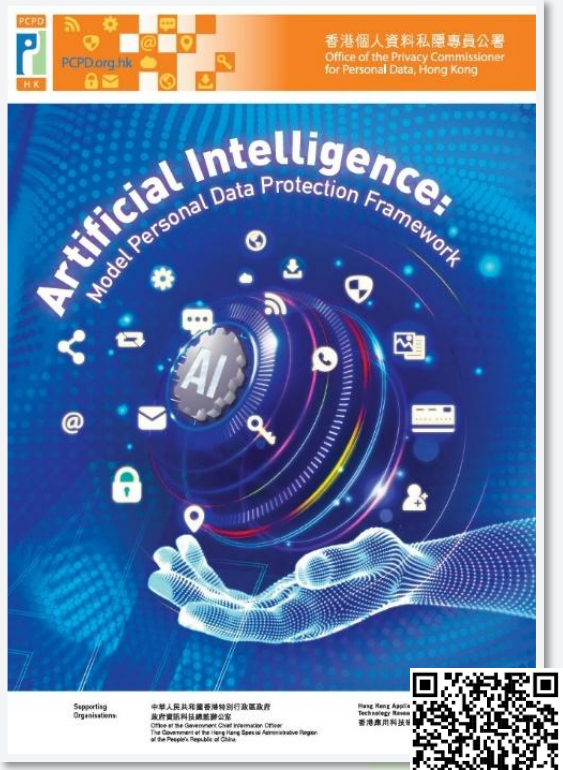
AI-generated voice: "I am going to talk about the latest privacy risks related to AI and share with you real live examples. I hope you could learn and share with your family and friends" (11 Sec)

With just about a 3-second recording, the above synthetic voice can be generated in under a minute.

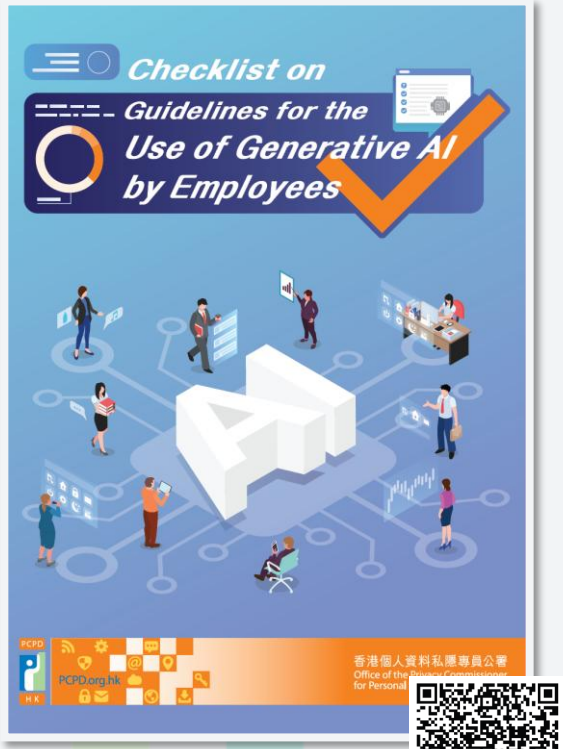
PCPD's Guidance

Organisations

Aug 2021

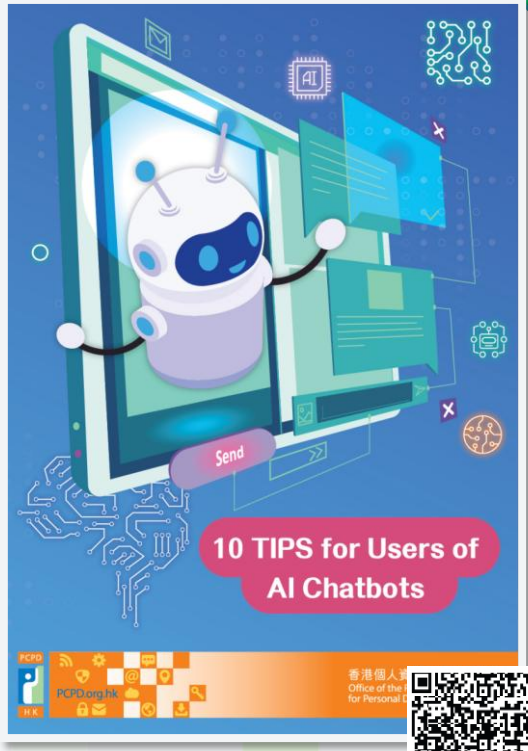


Jun 2024



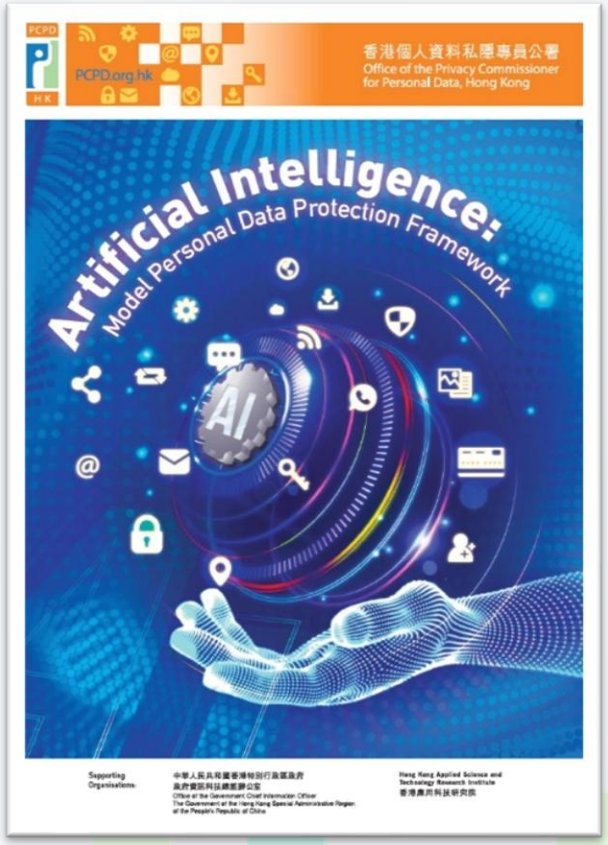
Mar 2025

Public



Sep 2023

Artificial Intelligence: Model Personal Data Protection Framework



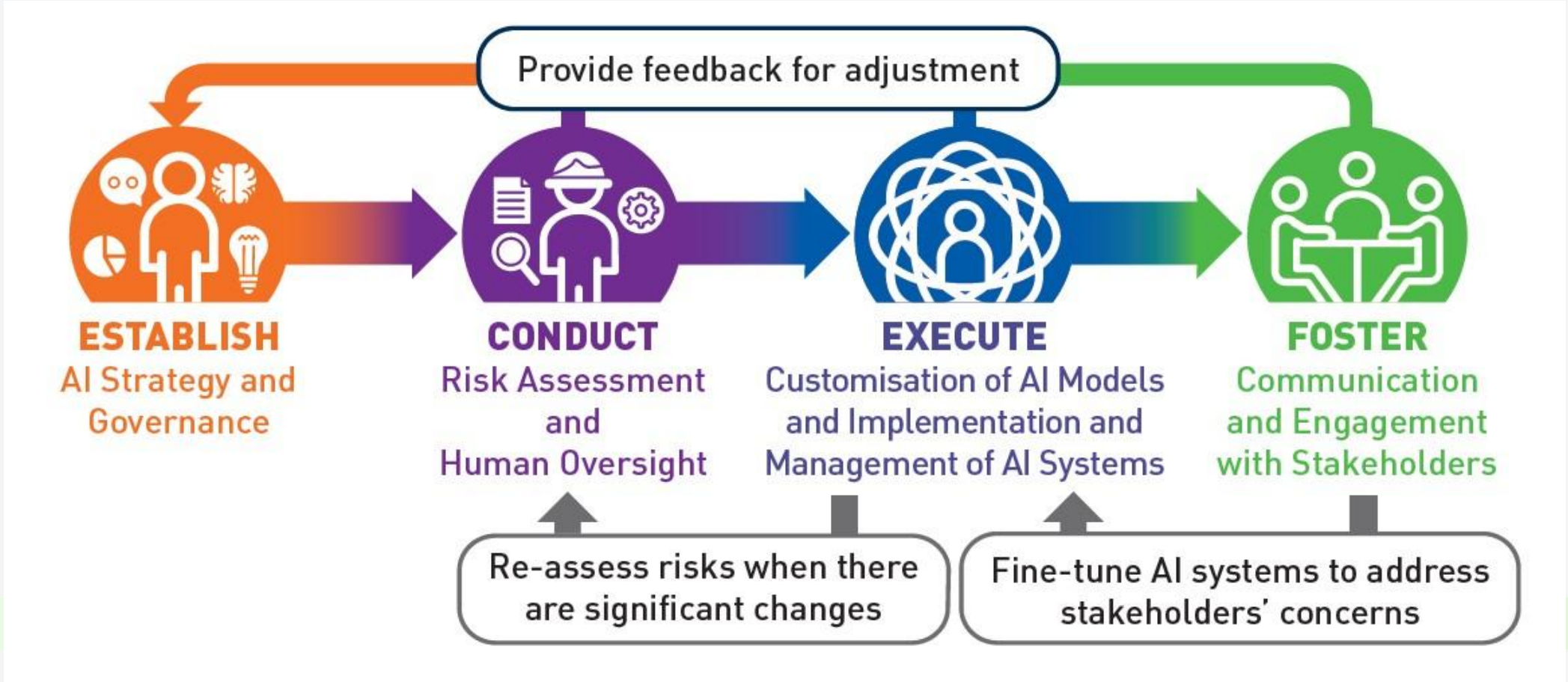
Feature

A set of recommendations on the best practices for organisations **procuring, implementing and using any type of AI systems**, including generative AI, that involve the use of personal data

Benefits

- Assist organisations in complying with the requirements of the Personal Data (Privacy) Ordinance
- Nurture the healthy development of AI in Hong Kong
- Facilitate Hong Kong's development into an innovation & technology hub
- Propel the expansion of the digital economy not only in HK but also GBA

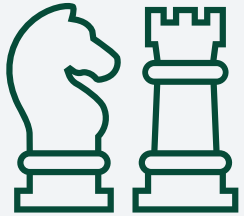
Artificial Intelligence: Model Personal Data Protection Framework



Establish
AI Strategy and Governance

ESTABLISH
AI Strategy and Governance

1



AI Strategy

2

Governance Structure

3

Governance Considerations for Procuring AI Solutions

4

Training and Awareness Raising

AI Strategy

An AI strategy shows management's commitment

ESTABLISH AI Strategy and Governance

AI Strategy

Functions

Demonstrate the commitment of top management to the ethical and responsible procurement, implementation and use of AI

Provide directions on the purposes for which AI solutions may be procured, and how AI systems should be implemented and used

Elements that may be included

Setting out **ethical principles**

Establishing **specific internal policies and procedures**

Determining the **unacceptable uses** of AI systems

Regularly **communicating the AI strategy**, policies and procedures

Establishing an **AI inventory**

Considering **emerging laws and regulations** that may be applicable

Conduct
Risk assessment and human oversight

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* risk management measures**

Risk-based approach

The level of human oversight should correspond with the risks identified

CONDUCT
Risk Assessment
and
Human Oversight

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

Execute
Customisation of AI Models and implementation and management of AI systems

1. **Data preparation**
2. **Customisation and implementation**
3. **Management and Continuous Monitoring**

Data Preparation

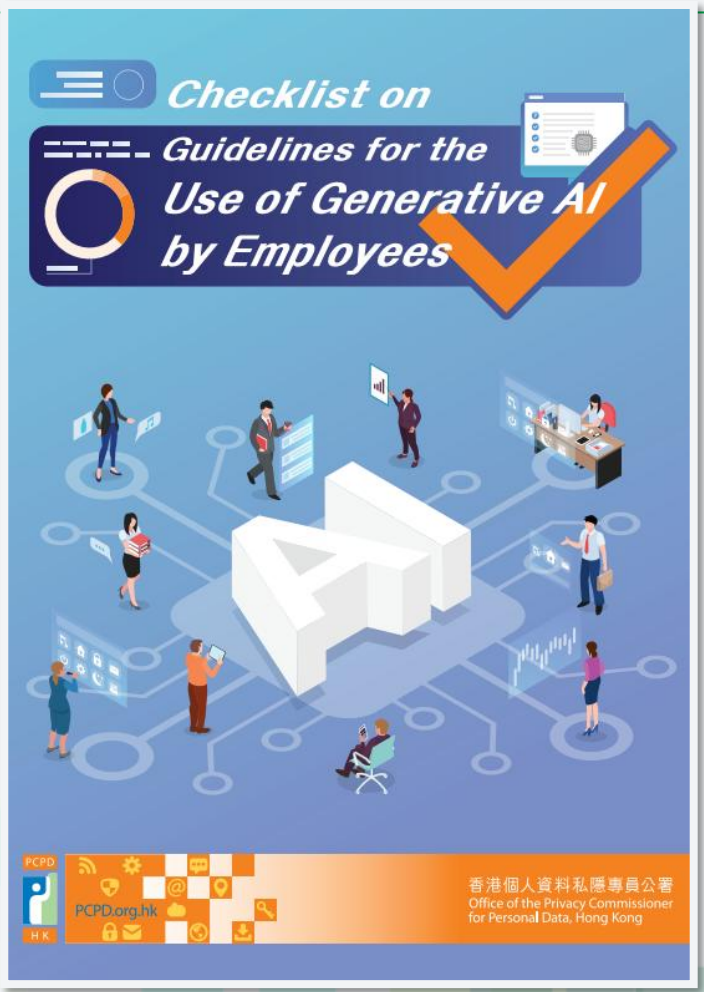
Compliance, data minimization, quality management, data handling

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

Process	Selected Recommendations
Data Preparation	Ensure compliance with privacy law Minimise the amount of personal data involved Manage data quality Document data handling
Customisation and Implementation of AI	Conduct rigorous testing and validation of reliability, robustness and fairness 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
Management and Continuous Monitoring of AI	Maintain proper documentation Establish an AI Incident Response Plan Conduct periodic audits Consider incorporating review mechanisms as risk factors evolve

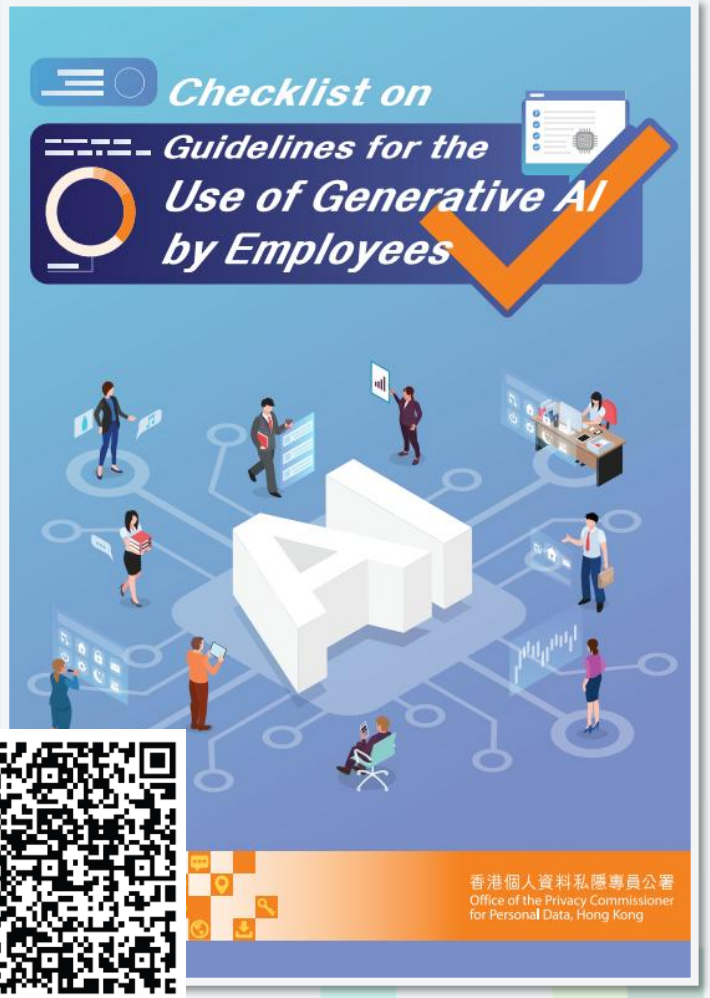
Foster
Communication and Engagement with Stakeholders

1. **Information Provision**
2. **Data Subject Rights and Feedback**
3. **Explainable AI**
4. **Language and Manner**



Overview – “Checklist on Guidelines for the Use of Generative AI by Employees”

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

01 Scope

Scope	Details
Permitted tools	**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 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 that the information generated by Gen AI tools be stored according to the organisation's **information management policy** and deleted according to its **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 Lawful and ethical use and prevention of bias

Unlawful activities

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 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

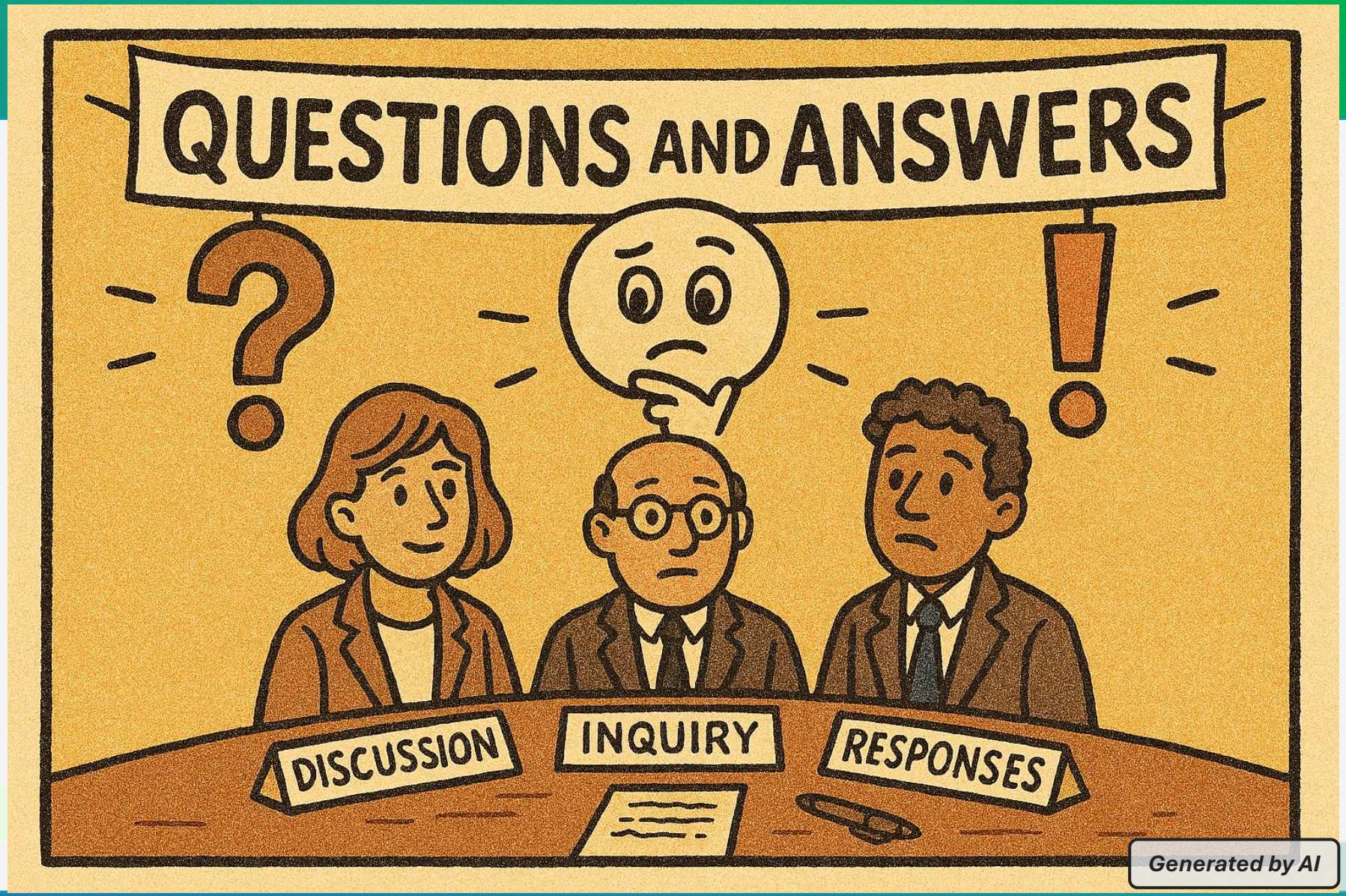
Require employees to use **unique and strong passwords** along with **multi-factor authentication**

Security settings

Require employees to maintain **stringent security settings**

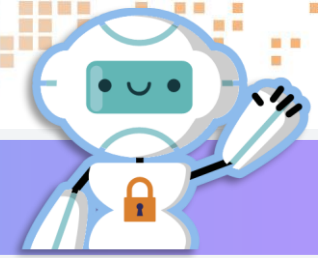
Response to AI incident and data breach incident

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



Generated by AI

守護私隱·改革創新
Protecting Privacy · Embracing Innovation



聯絡我們

電話: 2827 2827
傳真: 2877 7026
網站: www.pcpd.org.hk
電郵: communications@pcpd.org.hk
地址: 香港灣仔皇后大道東248號
大新金融中心13樓1303室

追蹤我們

私隱公署PCPD
香港个人资料私隐专员公署
Office of the Privacy Commissioner for Personal Data, Hong Kong

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