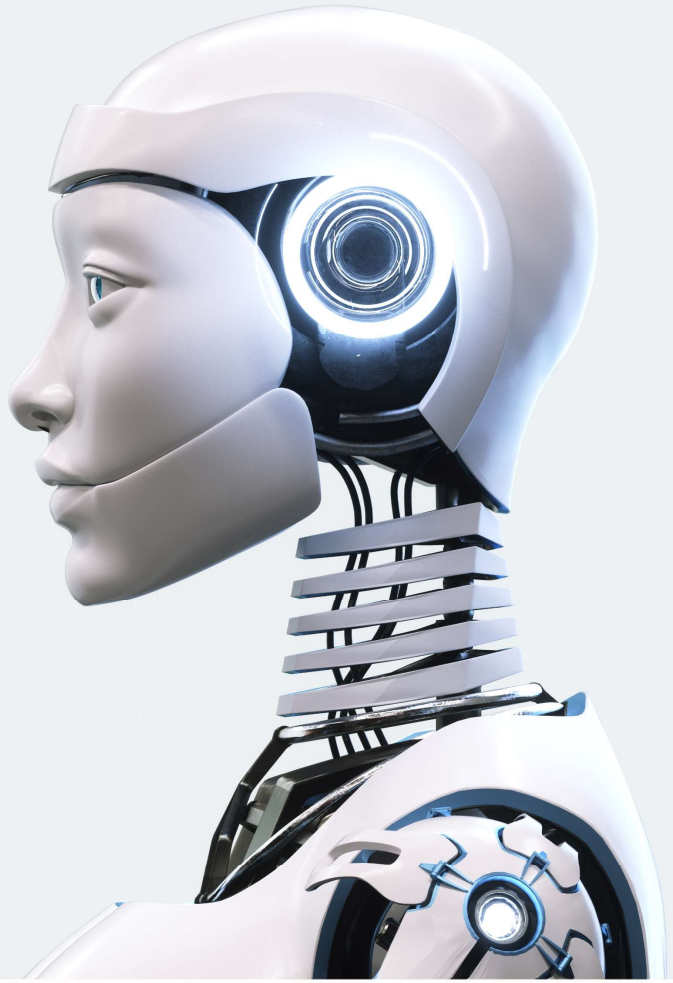
個人資料私隱專員公署
Office of the Privacy Commissioner for Personal Data

HKCS Hong Kong International Computer Conference 2024

Safeguarding Personal Data Privacy in the Age of AI: Governance Recommendations

5 November 2024



Ada CHUNG Lai-ling
Privacy Commissioner for Personal Data

Risks

AI poses privacy risks

Risk	Explanation	Illustration
Data Breach	AI systems, like chatbots, may **retain extensive user records**, making them **a target of hackers a**nd leading to **potential data breach**.	In March 2023, **ChatGPT** suffered a **major data breach,** revealing users' **conversation titles, names, email addresses, and the last four digits of their credit card numbers**.
Use of data	AI models can be **so advanced** that people find it **hard to understand how their personal data would be used.**	Some AI models can **identify the race** of some patients even **if that is not the purpose of the models.**
Excessive data collection	AI applications tend **to collect and retain as much data as possible**, including personal data.	OpenAI **reportedly scraped 300 billion words online** to train ChatGPT.
Data accuracy	Training AI models requires lots of data. But when **the quality and accuracy of that data are suboptimal**, the **AI system risk delivering incorrect analyses.**	An AI recruitment system of a multinational company was **trained with biased data** and **favoured male over female applicants.**

Deepfake
Millions could be lost from deepfake

HK$200 million scam

‘Everyone looked real’: multinational firm’s Hong Kong office loses HK$200 million after scammers stage deepfake video meeting

Employee fooled after seeing digitally recreated versions of company’s chief financial officer and others in video call

Deepfake technology has been in the spotlight after fake explicit images of pop superstar Taylor Swift spread on social media sites

Reading Time: **3 minutes**

Why you can trust SCMP

- In early 2024, an employee of a multinational company was **tricked by fraudsters using deepfake technology to impersonate the CFO** in an online meeting and **order money transfers**
- **HK$200 million** was transferred to the fraudsters

Source: SCMP

Officials & celebrities

- From Nov 2023 to May 2024, 21 **online deepfake video clips** involving **impersonation of government officials** or **celebrities** were identified by or reported to Police
- In Jan 2024, a deepfake video impersonated Chief Executive promoting an investment program with high returns

Source: HKSAR Government; SCMP

Deepfake

This demonstration shows how AI could easily be deployed for improper use of data

AI's risks vis-à-vis Data Protection Principles (DPP)

DPP1 — PURPOSE AND MANNER OF COLLECTION

- Large amount of personal data collected
- Disclose little about collection

DPP2 — ACCURACY AND RETENTION

- Outdated/incorrect data becomes part of training data and is kept longer than necessary

DPP3 — USE OF DATA

- User conversations become new training data and may be reproduced for another purpose

DPP4 — DATA SECURITY

- Security risks of storing large amount of conversations

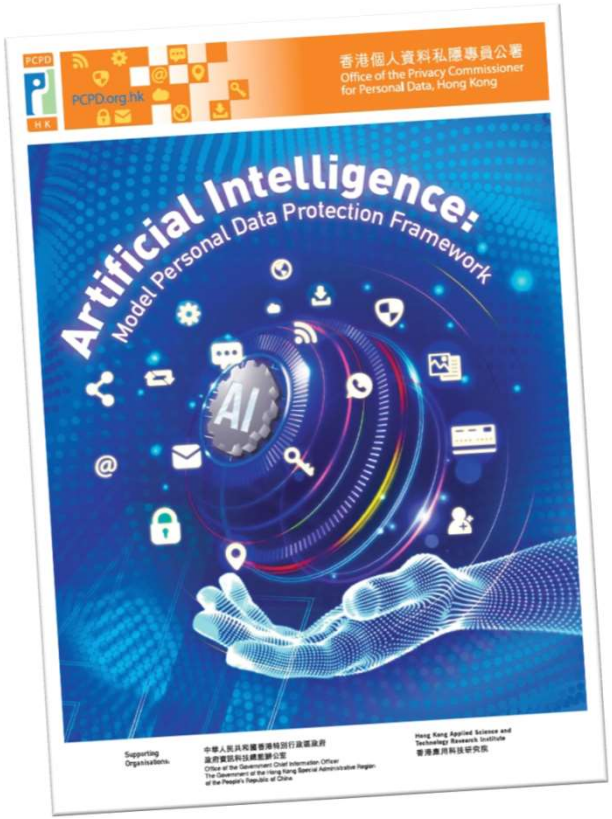
DPP5 — OPENESS AND TRANSPARENCY

- Data subjects are not fully informed of what personal data is held or how personal data is used

DPP6 — ACCESS AND CORRECTION

- Outdated/incorrect data that is part of training data is hard to be accessed or corrected

Artificial Intelligence: Model Personal Data Protection Framework



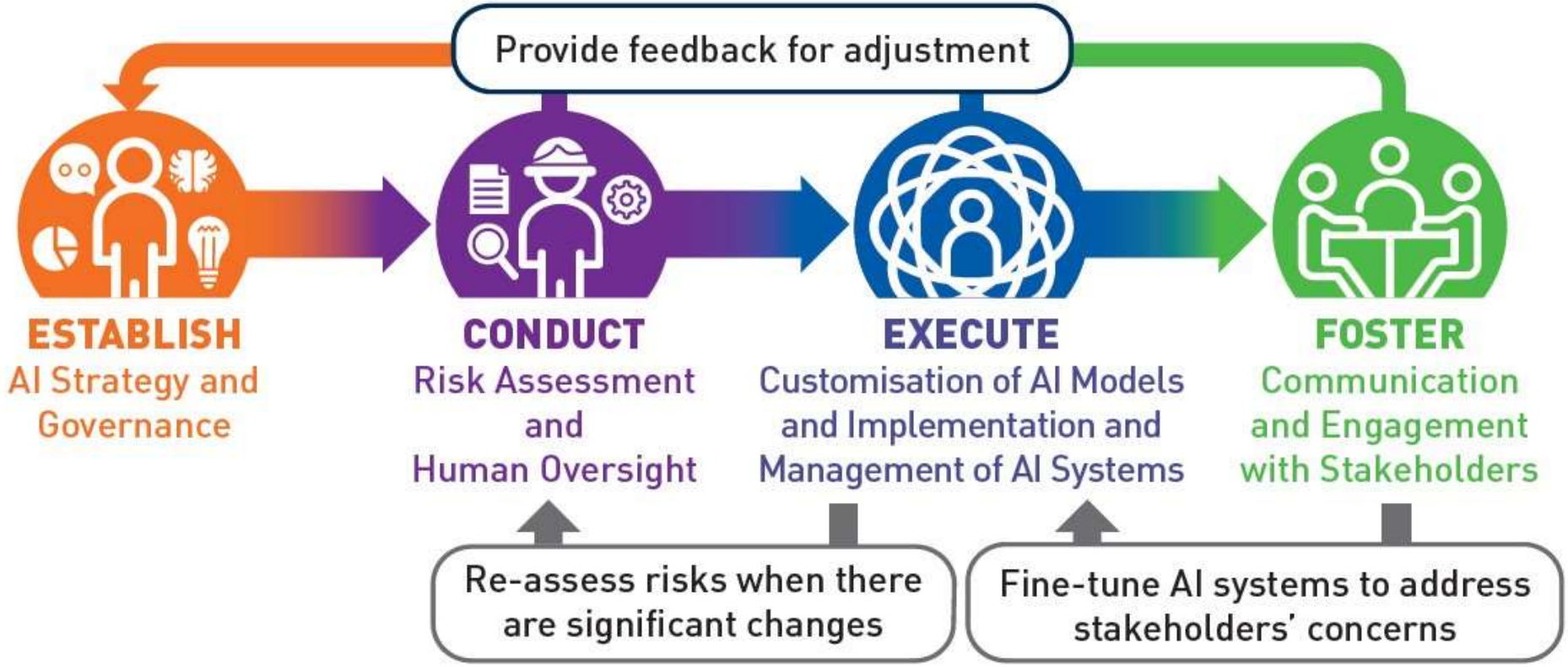
✨ Feature

- **Support Global AI Governance Initiative** of the Country
- **AI security** is one of the major areas of **national security**
- **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**

Model Personal Data Protection Framework



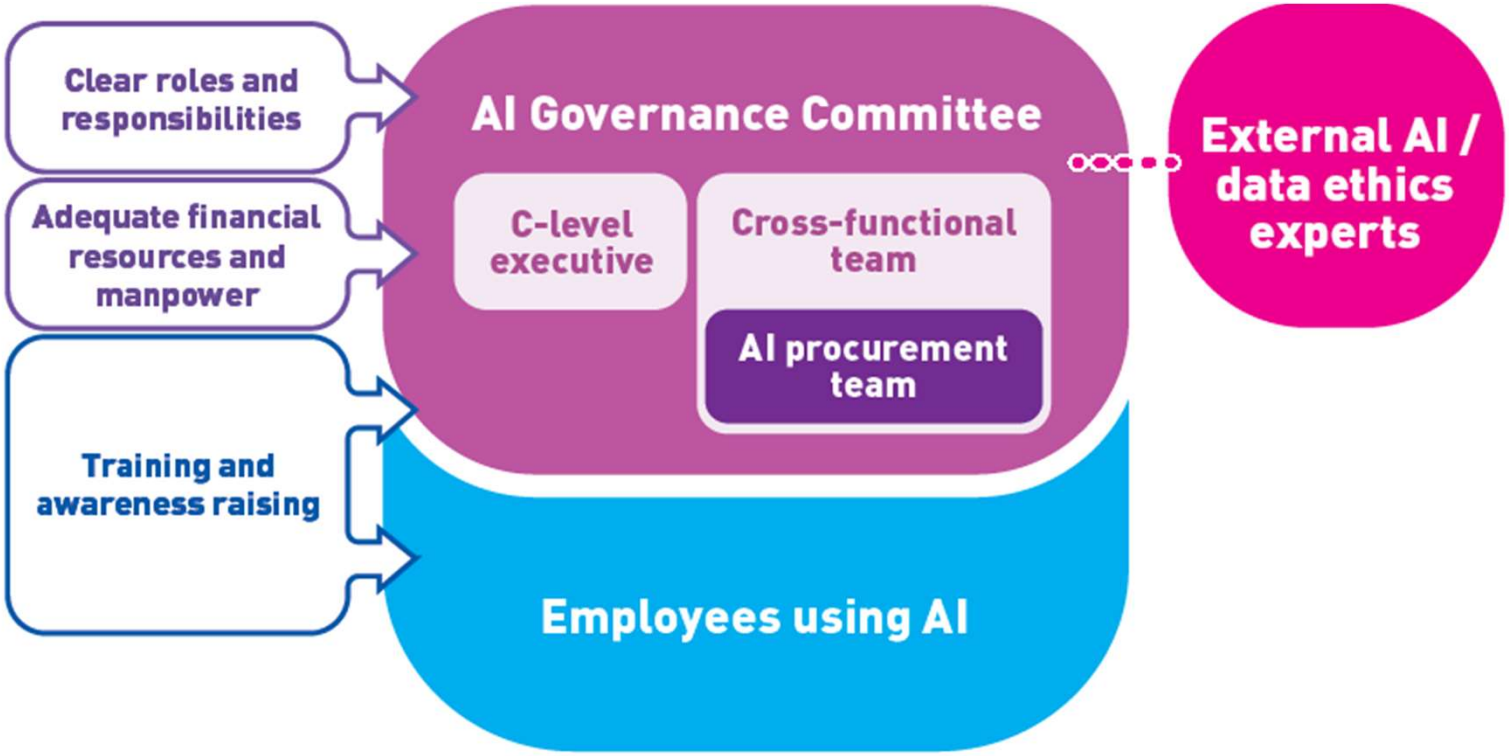
Governance Considerations

An organisation intending to invest in AI solutions may consider:

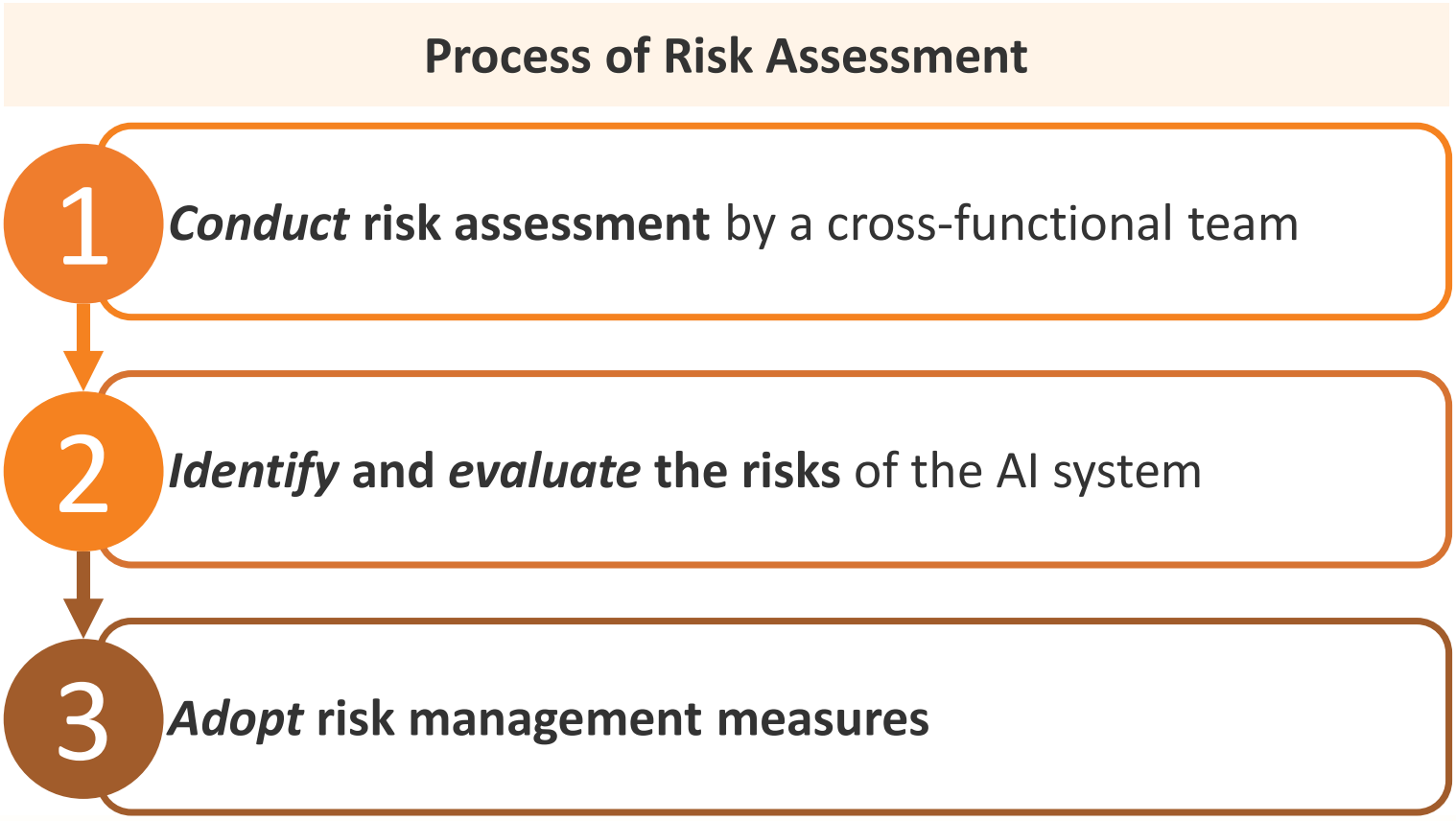
- Purpose(s) of using AI
- Privacy and security obligations and ethical requirements
- International technical and governance standards
- Criteria and procedures for reviewing AI solutions
- Data processor agreements
- Policy on handling output generated by the AI system
- Plan for continuously scrutinising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

Governance Structure

An internal governance structure with sufficient resources, expertise and authority should be established

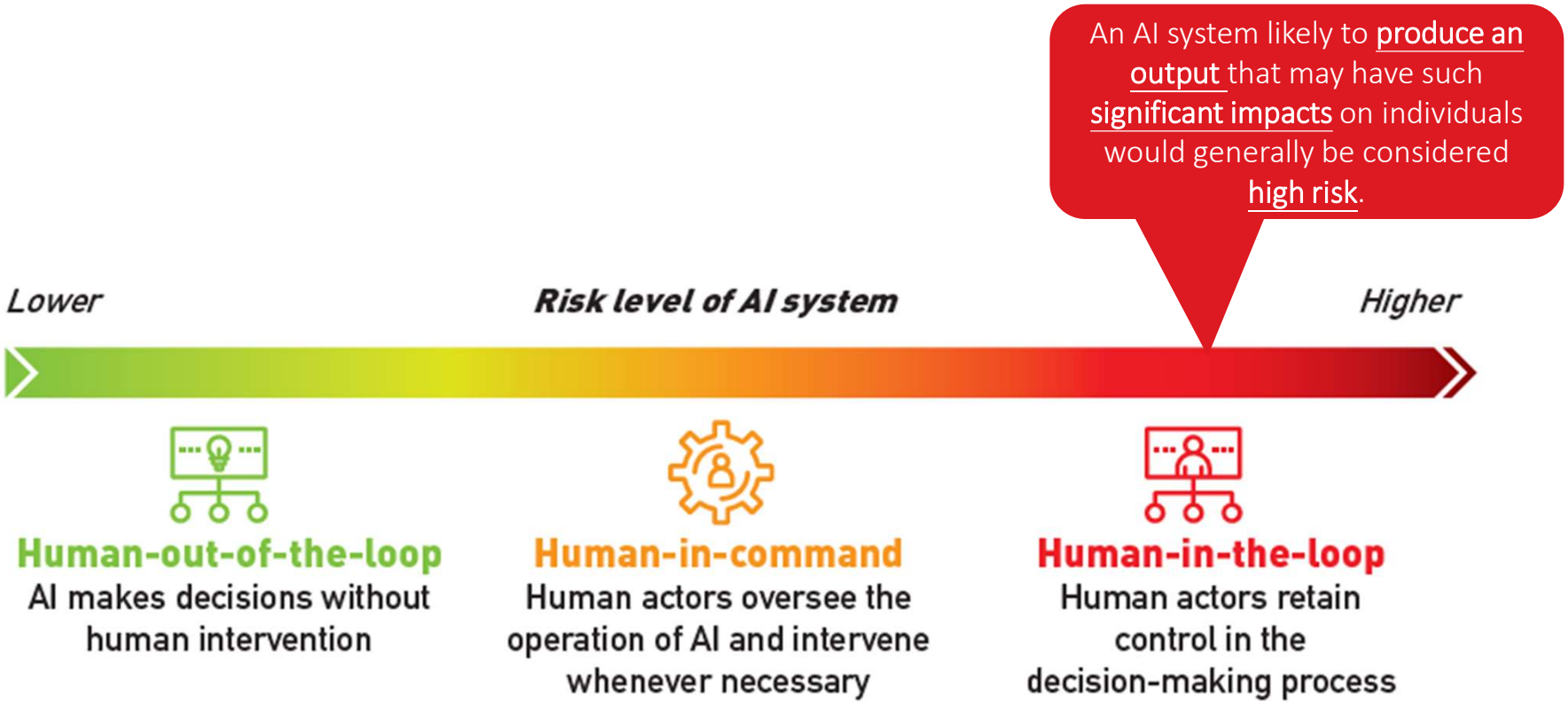


Conduct
Risk assessment and human oversight



Conduct Risk Assessment

The level of human oversight should correspond with the risks identified



Examples

The below use cases may incur higher risks

- Real-time identification of individuals using biometric data
- Evaluation of individuals' eligibility for social welfare or public services
- Assessment of job applicants, evaluation of job performance or termination of employment contracts
- Evaluation of the creditworthiness of individuals for making automated financial decisions
- AI-assisted medical imaging analytics or therapies

Cryptocurrency project
World ID

Execute: Data Preparation

Compliance, data minimisation, quality management, data handling

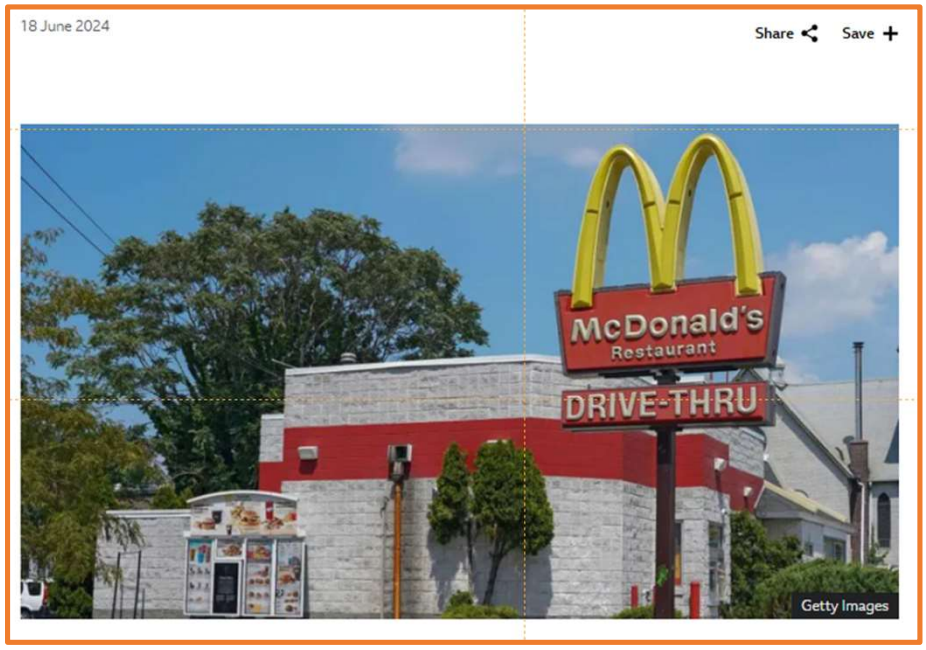
Selected Recommendations	Example
Ensure compliance with privacy law **Minimise the amount of personal data involved** **Manage data quality** **Document data handling**	• A **fashion retail platform** is **purchasing a third-party developed AI chatbot** that it will customise to provide **fashion recommendations** to its customers • The company may find it **necessary** to use the **past purchases** and **browsing histories** of **different segments** of its customer groups to fine-tune the chatbot • However, the use of **personal data**, such as customers' names, contact details and certain demographic characteristics, would **not be necessary**

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

When you order an ice cream

A restaurant chain stopped its voice ordering system after blunders



Test run of AI-powered voice ordering systems for customers

Multiple problems reported on social media

Test run discontinued in mid-2024

Source: BBC

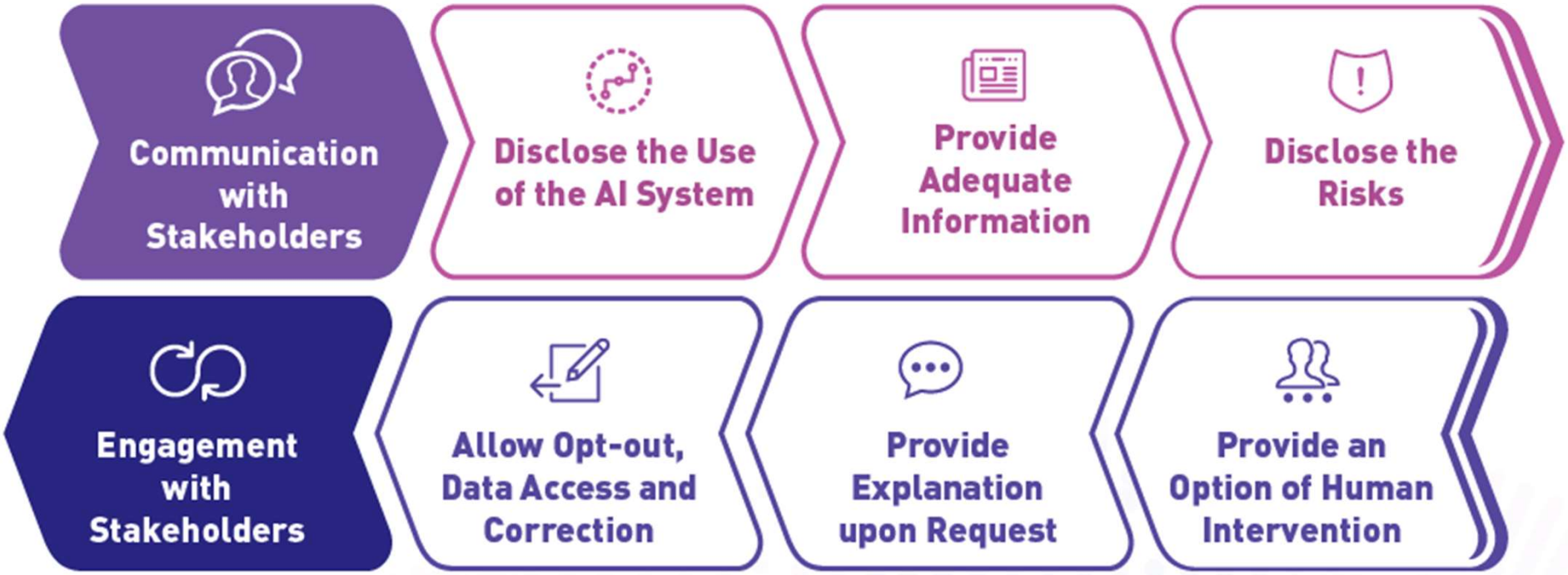
AI Incident Response Plan

The plan may encompass the below six elements

1. Defining an AI Incident
2. Monitoring for AI Incidents
3. Reporting an AI Incident
4. Containing an AI Incident
5. Investigating an AI Incident
6. Recovering from an AI Incident

Foster

Communication and engagement with stakeholders



Thank you!

www.pcpd.org.hk

communications@pcpd.org.hk

Please follow us!