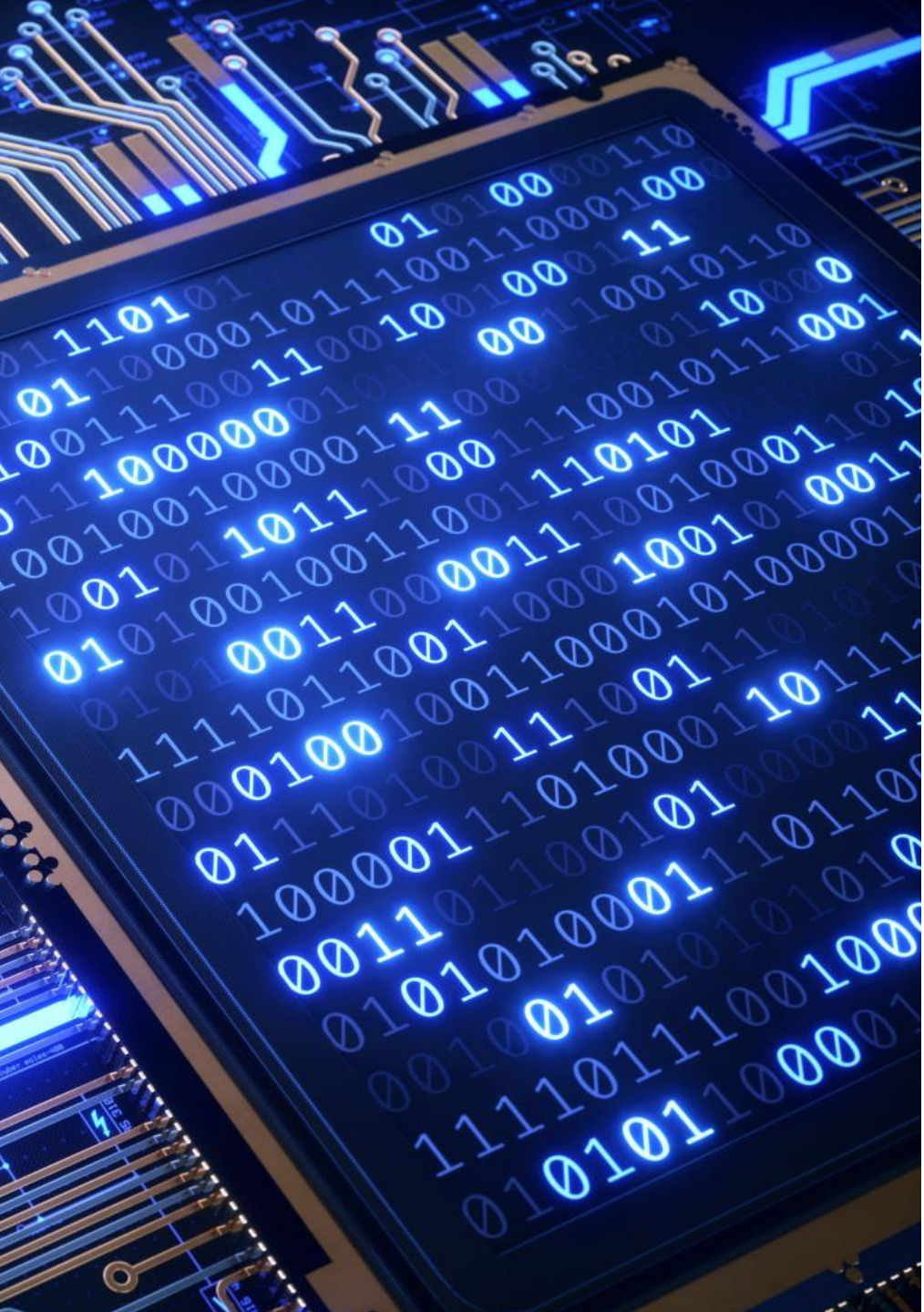




PCPD



H K

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

香港AI治理研討會

生成式AI時代的 私隱保障新篇章

個人資料私隱專員
鍾麗玲女士

2025年5月30日

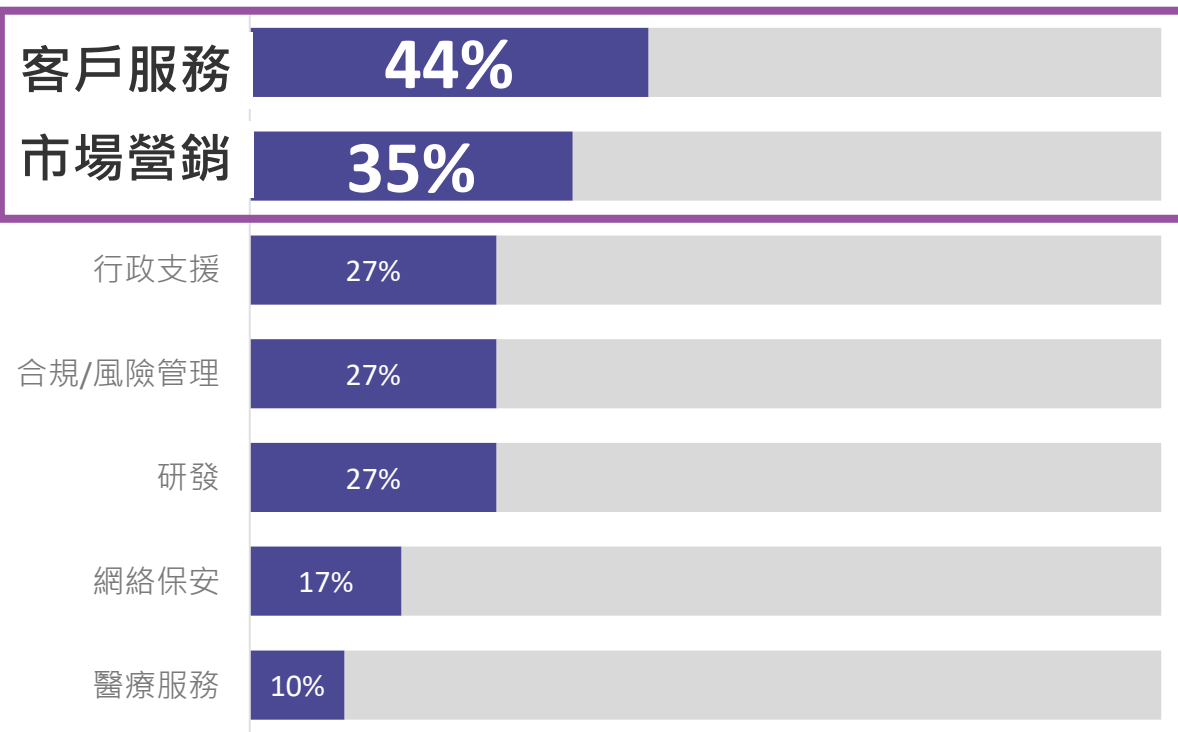


用例

公署循規審查顯示，不少企業使用AI協助客戶服務和市場營銷

AI系統應用領域

公署循規審查中有使用AI的香港企業，2025



資料來源: 私隱專員公署



企業使用生成式AI的例子



自動化內容創作

- 創建廣告文案和圖像
- 按客戶的數據，創作個人化的廣告內容



數據分析和預測

- 協助企業洞察市場趨勢、消費者行為及競爭對手業務狀況



提升客戶服務質素

- 透過聊天機械人作售前、購物過程中，和售後服務，提升客戶購買時的體驗

風險

AI 對個人資料私隱構成風險

 香港企業：生成式AI在新興技術中
存在最高的私隱風險

新興技術中的私隱風險

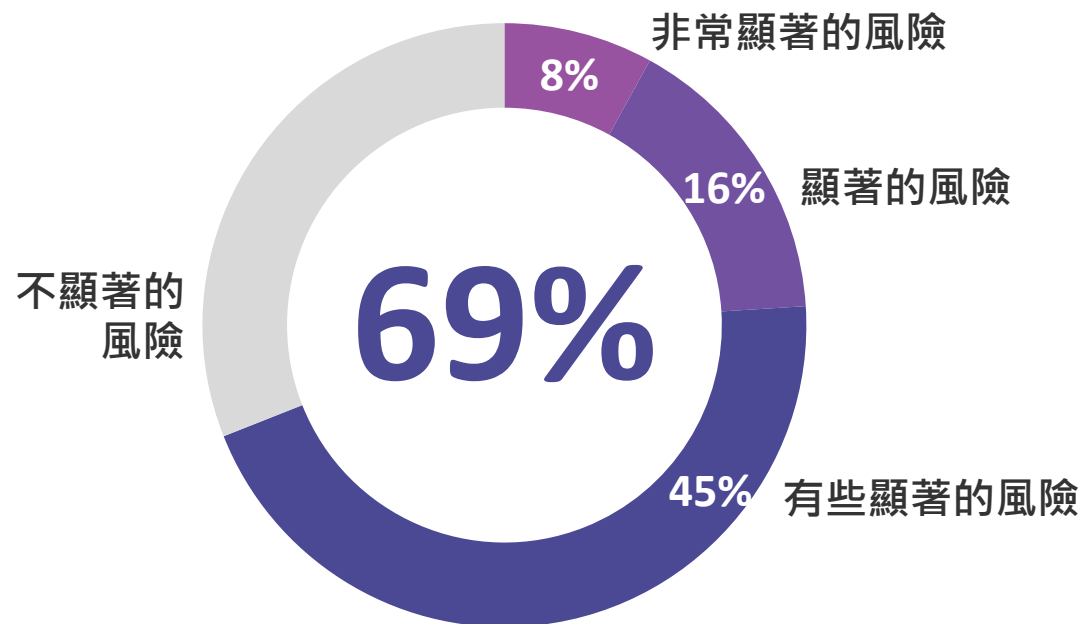
香港企業，2023

- 1  生成式AI
- 2  Cookies 和其他在綫追蹤器科技
- 3  雲端計算
- 4  物聯網
- 5  區塊鏈相關技術
- 6  數據分析及營運流程自動化

資料來源：私隱專員公署和生產力局

 企業認為於營運中使用AI
帶來顯著私隱風險

香港企業，2024



資料來源：私隱專員公署和生產力局

風險

AI 可能對個人資料私隱構成多重風險

風險

解釋

例子



資料外洩

AI系統（如AI聊天機械人）可能儲存大量用戶對話資料，容易成為黑客攻擊的目標，引致資料外洩

2023年3月，**ChatGPT**發生了一次嚴重的資料外洩事故，洩露了部分用戶過往對話的標題、用戶的姓名、電郵地址和信用卡號碼的最後四位數字



資料收集過量

AI 傾向於收集和保留盡可能多的數據，包括個人資料

OpenAI被指擷取了3000億個網上詞彙來訓練ChatGPT



資料的使用

「黑盒」難題：AI使用者無法得知AI系統的內部運作邏輯

AI模型能在醫療影像沒提及病人的種族的情況下也準確推測到病人的種族



資料準確性

若資料的質素和準確性有偏差，AI可能因數據不準確而做出錯誤分析，從而影響制定決策

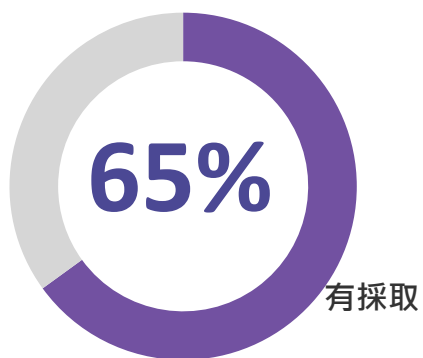
跨國科技公司用於審視應徵者履歷的AI系統，因AI系統的訓練數據缺乏女性應徵者資料，不當地傾向挑選男性應徵者，對女性應徵者造成不公

資料來源: [Fortune](#); [刺針](#), [Patterns](#); [MIT Technology Review](#)

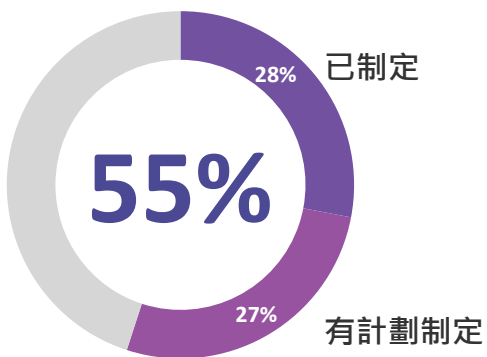
措施

機構有緩減AI相關風險，並參考公署指引

👍 於營運中有使用AI的企業(2024)：



有採用
數據安全防
護措施

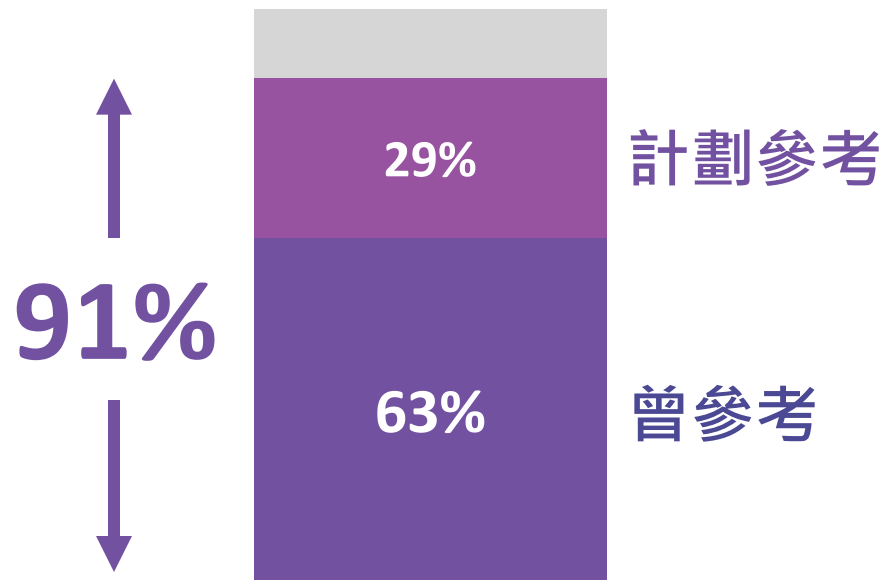


已制定或計劃
制定
AI 安全
風險政策

資料來源：私隱專員公署和生產力局

👍 公署循規審查顯示逾九成使用AI收集個人資料的機構有參考或會參考公署指引

公署循規審查中有使用AI收集、使用及處理個人資料的香港企業，2025



資料來源：私隱專員公署

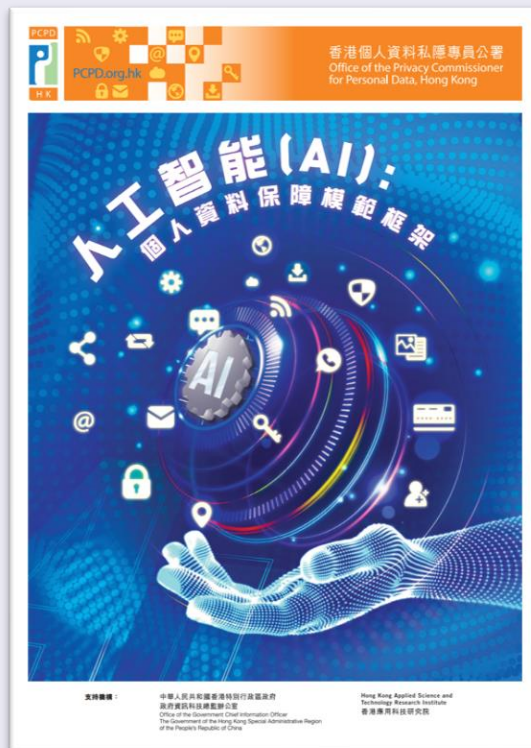
公署的指引

公署因應AI的發展發布了不同指引

機構



(2021年8月)



(2024年6月)



(2025年3月)

公眾



(2023年9月)

《僱員使用生成式AI的指引清單》



目的

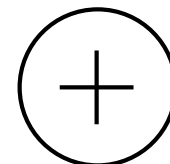
協助機構制定僱員
在工作時使用生成
式AI的內部政策或
指引，以及遵從
《私隱條例》有關
處理個人資料的相
關規定



特色



以清單形式呈現



作為良好的行
事方式、機構
可以制定與其
價值觀及使命
一致的內部政
策或指引

僱員使用生成式AI的政策或指引的建議內容

01

範圍

02

保障個人資料私隱

03

合法及合乎道德的
使用及預防偏見

04

數據安全

05

違反政策或指引

僱員使用生成式AI的政策或指引的建議內容 – 範圍

方面

內容



獲准使用的工具

清晰訂明准許使用的生成式AI工具及應用程式, 例如：

- 公眾可用的AI工具或應用程式
- 內部開發的AI工具或應用程式



獲准許的用途

清晰指明僱員可以使用生成式AI工具處理甚麼工作或活動，例如：

- 起草
- 總結資訊
- 生成文本、音頻及 / 或視像內容



政策適用性

訂明政策是否適用於**整個機構**；**指定部門**；**指定職級**；及 / 或**指定僱員**



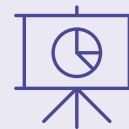
僱員使用生成式AI的政策或指引的建議內容 – 保障個人資料私隱



獲准輸入的資訊種類及數量

提供清晰指示，說明：

- ✓ 可輸入至生成式AI工具的資訊種類及數量
- ✗ 禁止輸入的資訊種類



輸出資訊的獲准許用途

提供清晰指示，說明生成式AI工具所生成的資訊（包括個人資料）的**獲准許用途**，以及僱員應否、何時及如何在進一步使用這些個人資料前將其匿名化



輸出資訊的獲准許儲存方式

要求僱員根據機構的**資訊管理政策**儲存資訊和**資料保留政策**刪除生成式AI工具所生成的資訊



遵從其他相關內部政策

確保**使用生成式AI的政策**與機構的**其他相關內部政策**一致

僱員使用生成式AI的政策或指引的建議內容 – 合法及合乎道德的使用及預防偏見

違法行為



僱員不能為進行非法
或有害的活動
使用生成式AI工具

強調僱員有責任擔當審查員



準確度及核實

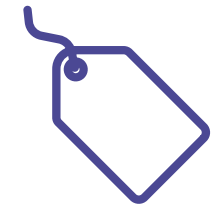
強調僱員需要核
實AI所提供的資
訊



預防偏見及歧視

提醒僱員AI生成的結
果可能帶有偏見及歧
視

訂明更正及報告機制



加上水印 / 標籤

說明應何時及如
何在AI生成結果
上加上水印或標
籤

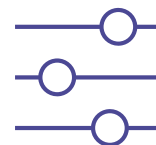
僱員使用生成式AI的政策或指引的建議內容 – 數據安全

獲准許裝置



訂明准許僱員可用**哪些裝置**來取用生成式AI工具

保安設定



要求僱員保持**嚴格的保安設定**

獲准許使用者



訂明**可以使用生成式AI工具的僱員**

AI事故及資料外洩事故應變



要求僱員根據機構的**AI事故應變計劃報告AI事故**

用戶憑證



要求使用**獨特且高強度的密碼及多重認證**

《人工智能 (AI): 個人資料保障模範框架》

✓ 好處

i 協助機構遵從《私隱條例》的規定

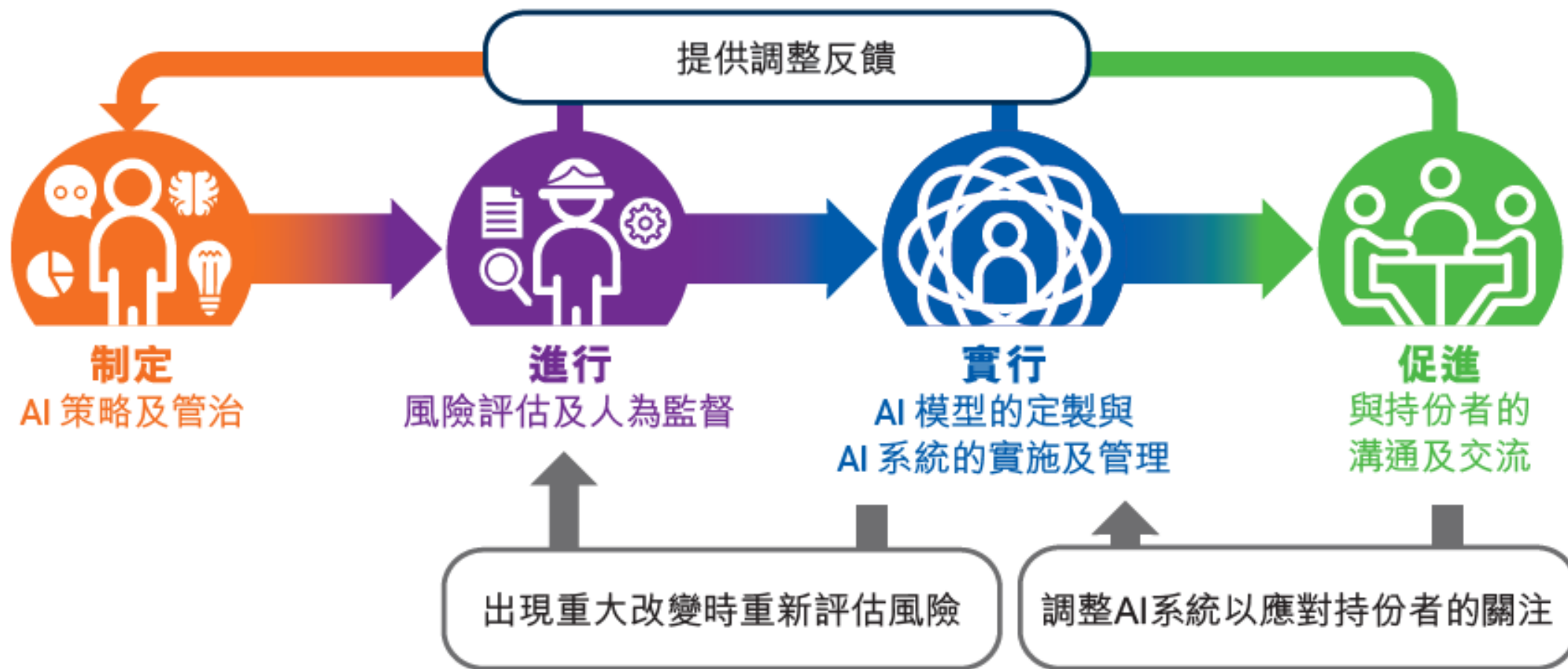
確保人工智能安全

增加競爭力

向採購、實施及使用任何種類的AI系統（包括生成式AI）的機構，就保障個人資料私隱方面提供有關AI管治的建議及最佳行事常規

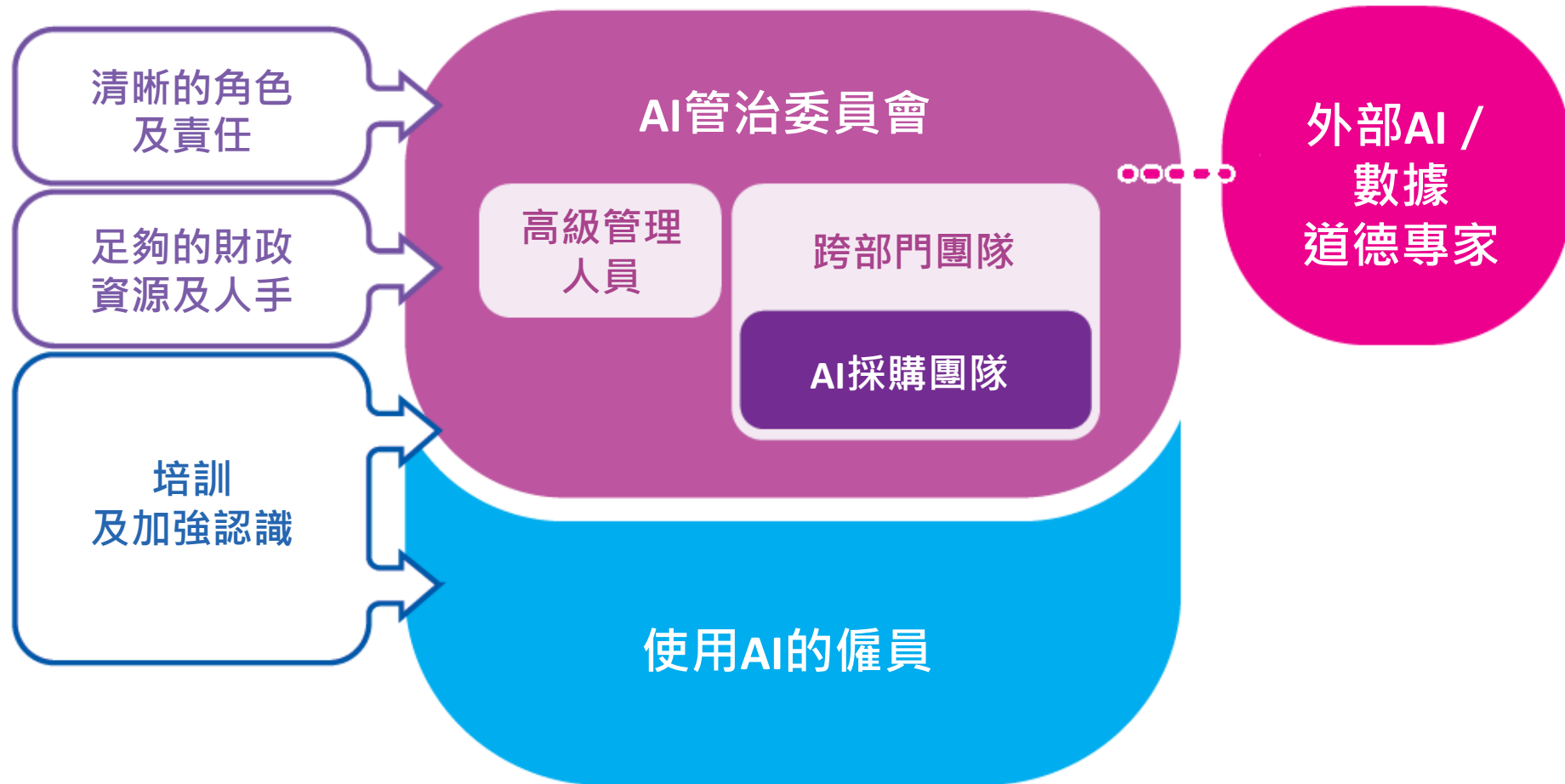


個人資料保障模範框架



制定AI 策略及管治

管治架構



制定AI 策略及管治

9項管治考慮



使用AI的目的



私隱和保安的責任
及道德規定



技術性和管治方面
的國際標準



審查AI方案的準則
和程序



資料處理者協議



處理AI系統生成結
果的政策



持續檢視環境變化
的計劃



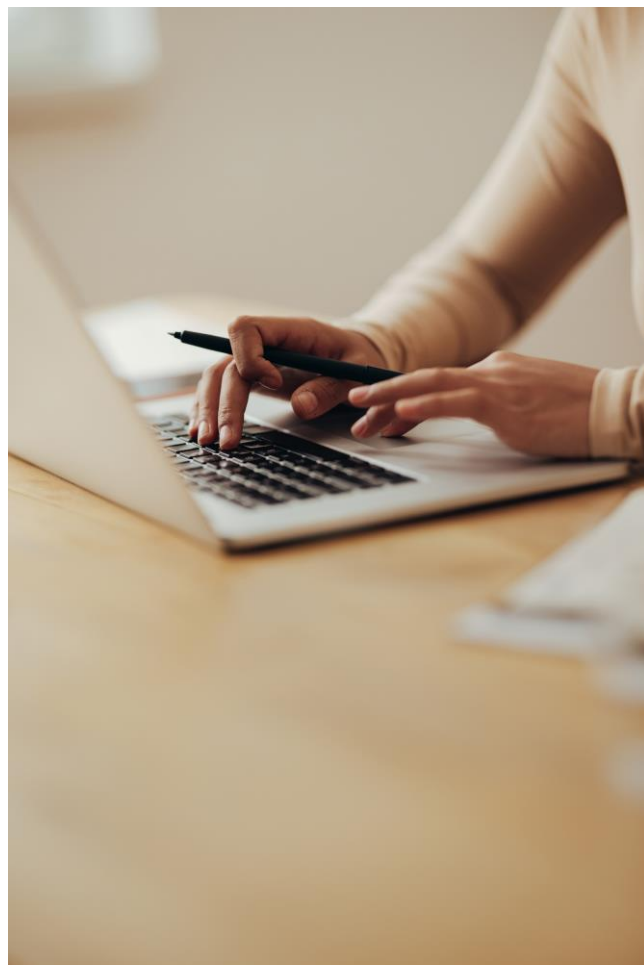
持續監察、管理和
維持AI方案的計劃



評估AI供應商

提升AI安全 = 提升競爭力

機構可參考以下步驟，提升 AI 安全



下載公署有關AI的指引



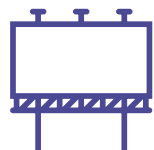
評估機構使用 AI 的策略



制定相關管治策略和架構、
草擬相關內部政策或指引



如有疑問，請向公署「AI安全」
熱線2110 1155查詢



參與中小企數據安全培訓系列
及內部培訓講座

下載指引

機構



(2021年8月)

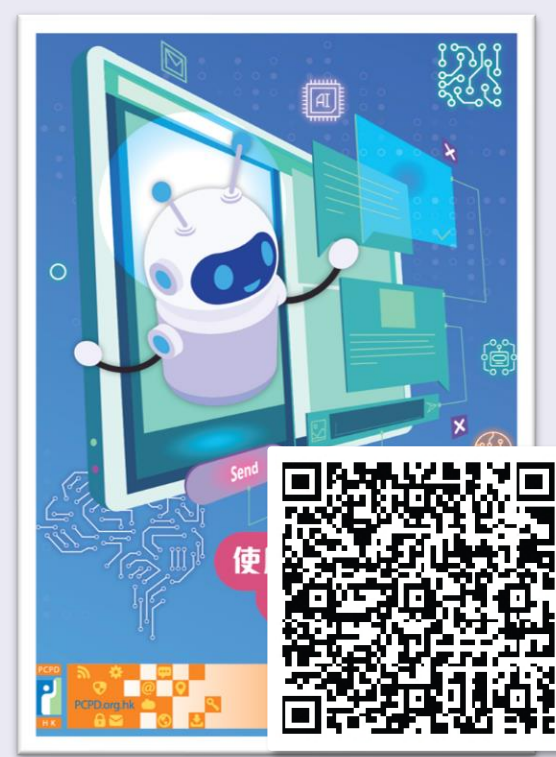


(2024年6月)



(2025年3月)

公眾



(2023年9月)

聯絡我們

 **查詢** 2827 2827  **傳真** 2877 7026

 **網址** www.pcpd.org.hk

 **電郵** communications@pcpd.org.hk

 **地址** 香港皇后大道東248號大新金融中心13樓1303室

保障、尊重個人資料私隱

Protect, Respect Personal Data Privacy

追蹤我們
最新資訊

