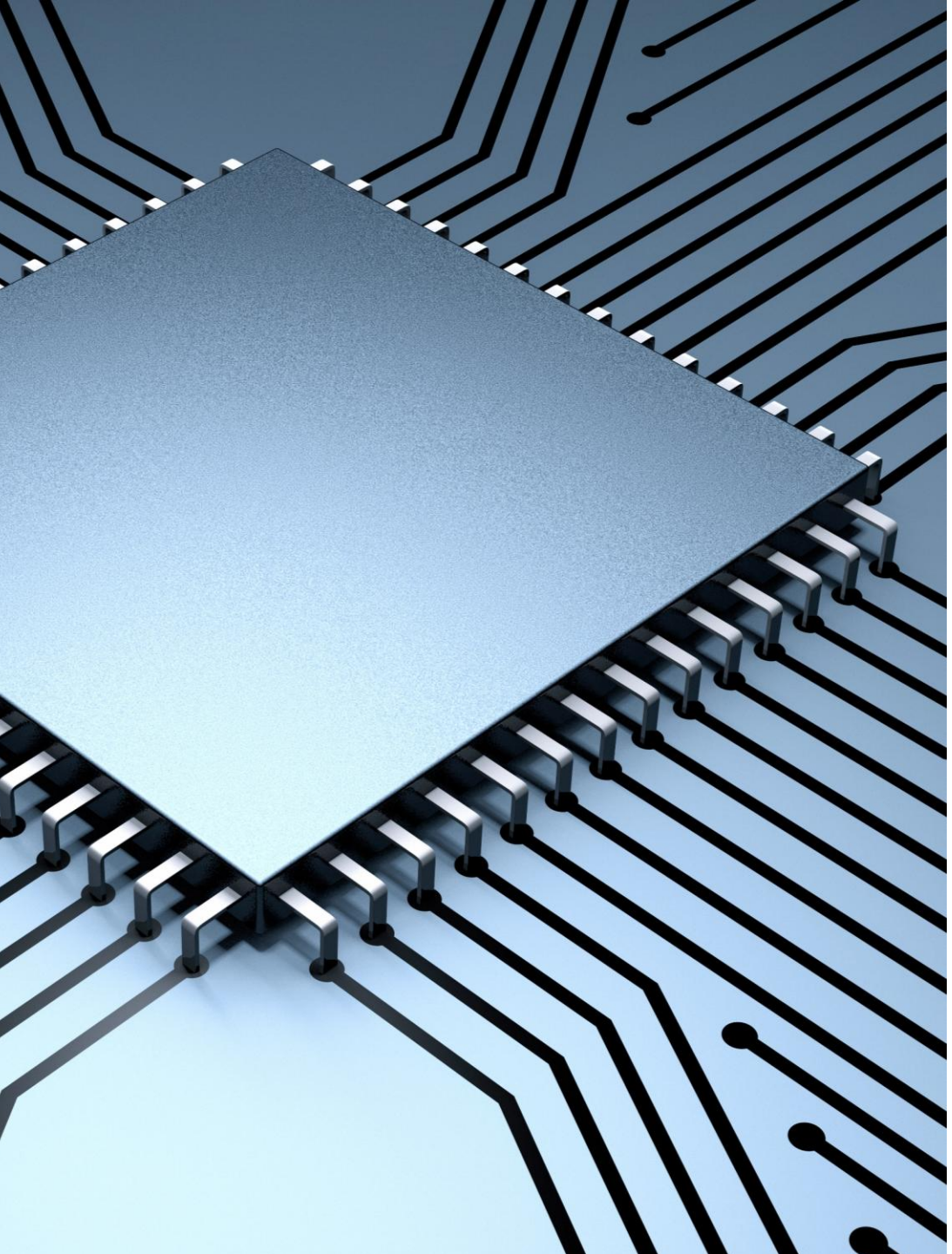


香港個人資料私隱專員公署
Office of the Privacy Commissioner for Personal Data, Hong Kong

「中小企認識AI數據安全及私隱風險」研討會

助理個人資料私隱專員
（合規、環球事務及研究）
王雅媛女士

2025年6月13日

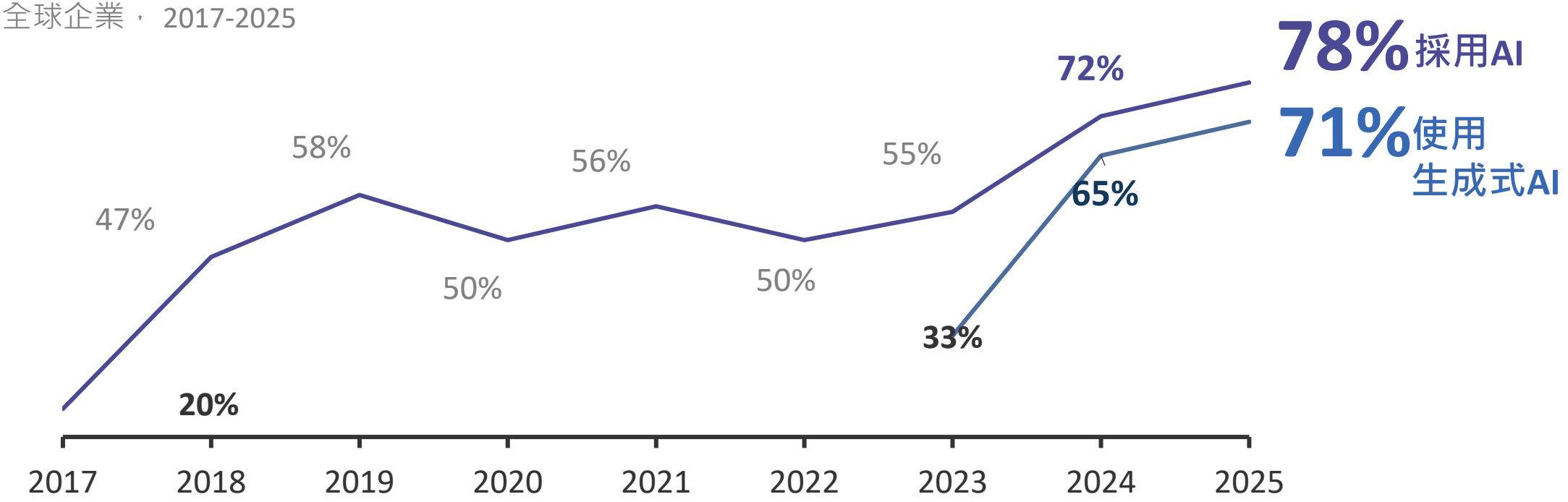
環球趨勢

全球機構正積極採用 AI，但本地中小企使用率仍有上升空間

📈 全球機構AI（包括生成式AI）採用率 於2024-2025年大幅上升

表示至少在一個商業功能上採用AI的受訪機構比例

全球企業， 2017-2025



資料來源: McKinsey

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

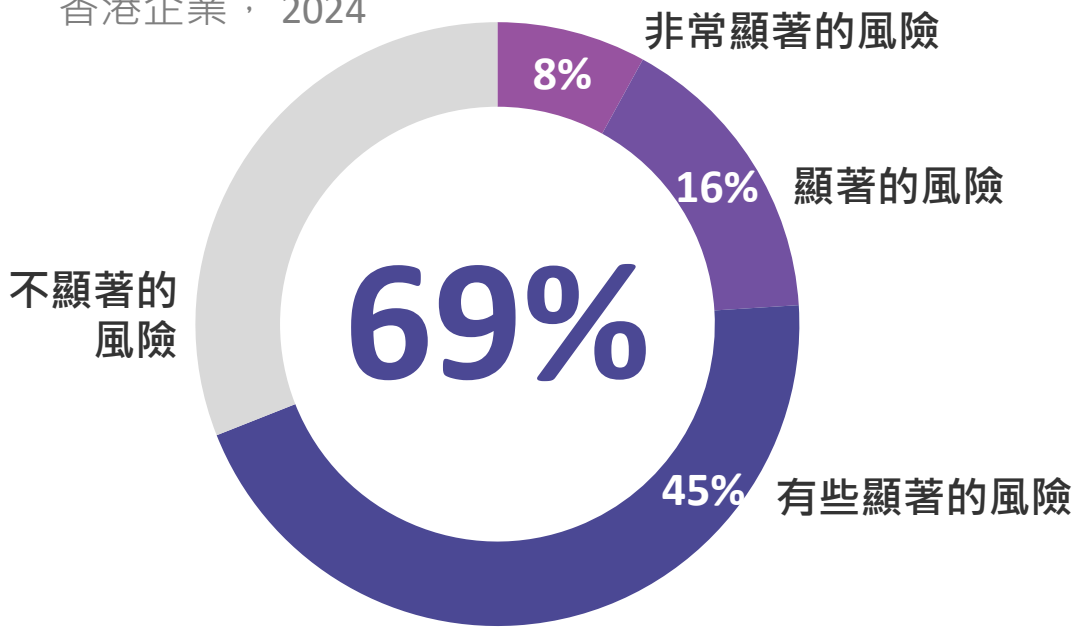
新興技術中的私隱風險
香港企業， 2023

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

認為在營運中使用 AI 會帶來私隱風險的企業百分比
香港企業， 2024



資料來源:私隱專員公署和生產力局

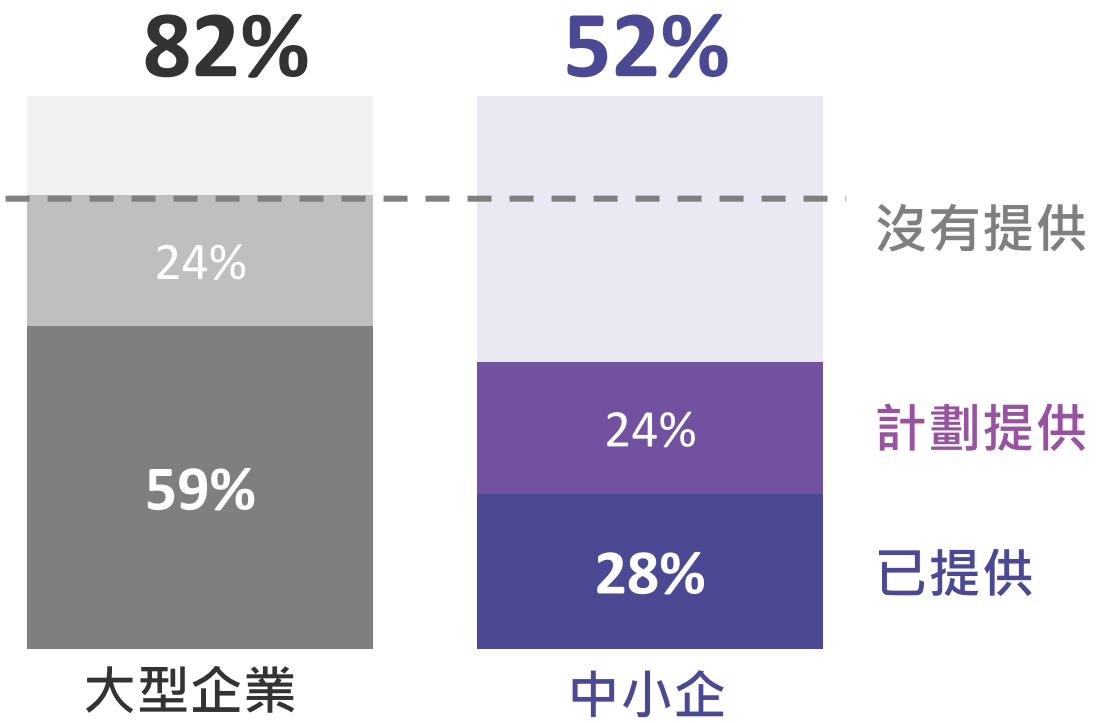
AI安全準備程度

相比於大企，中小企的AI安全準備程度較為不足

較少中小企已經提供或計劃會為員工提供 AI 的培訓

為員工提供有關 AI的培訓情況

於營運中使用AI的香港企業， 2024



資料來源: 私隱專員公署和生產力局

AI相關風險與相應的《個人資料（私隱）條例》資料保障原則

有機會違反資料保障原則的情況

第1原則

收集目的及方式

- 收集過多個人資料
- 在資料當事人不知情的情況下收集其個人資料

第2原則

準確性、儲存及保留

- 不需要保留的／錯誤的資料成為訓練數據的一部分，而且保留時間超過所需

第3原則

資料使用

- 在沒有取得資料當事人的同意下，使用用戶的對話作訓練數據，或用作其他用途

第4原則

資料保安

- 外洩用戶對話數據

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 ）

《人工智能 (AI): 個人資料保障模範框架》

✨ 特點

體現國家的《全球人工智能治理倡議》

人工智能安全是國家安全的重點領域之一

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

✅ 好處

協助機構遵從《私隱條例》的規定

孕育AI在香港的健康發展

促進香港成為創新科技樞紐

推動香港以至大灣區的數字經濟發展

國際標準

《模範框架》反映國際間認受的原則及最佳行事常規

三項數據管理價值

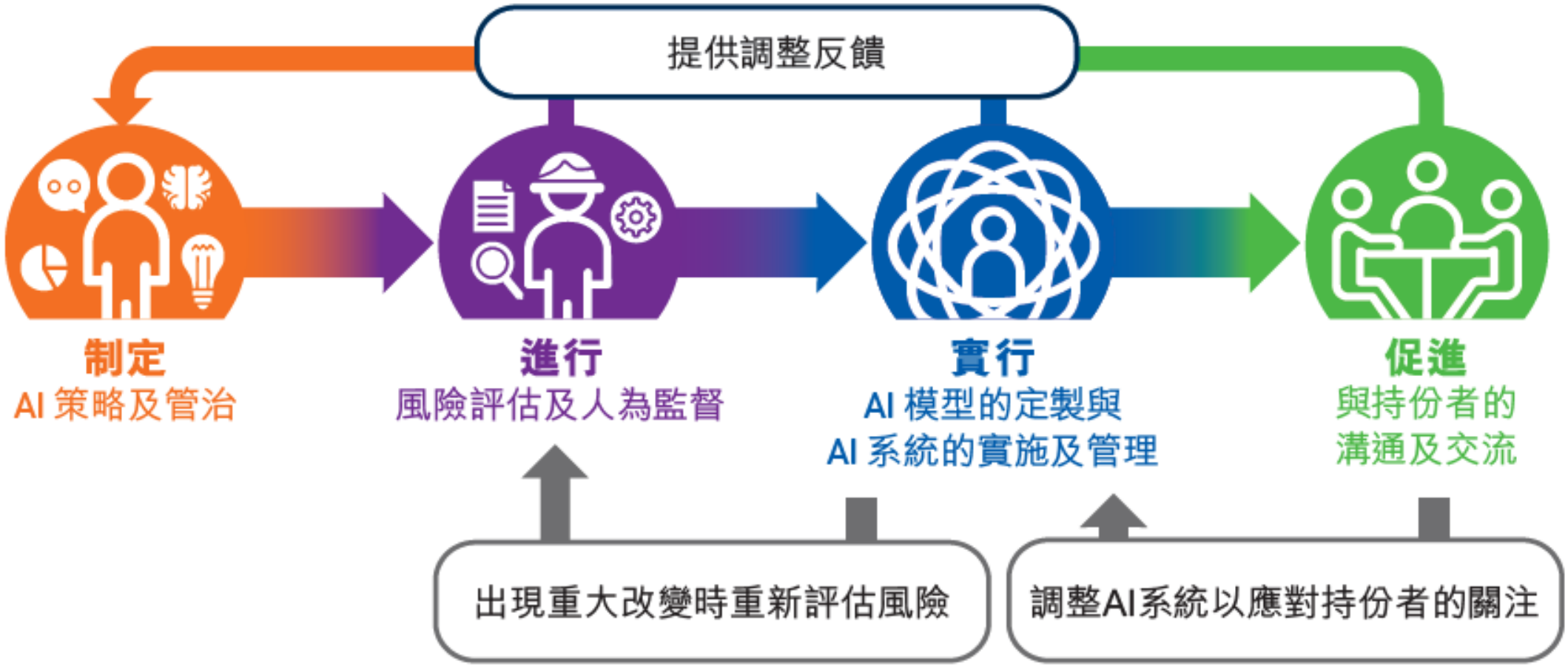
1. 尊重
2. 互惠
3. 公平

七項AI道德原則

1. 問責
2. 人為監督
3. 透明度與可解釋性
4. 數據私隱
5. 公平
6. 有益的AI
7. 可靠、穩健及安全

個人資料保障模範框架

個人資料保障模範框架


提供調整反饋
制定
AI 策略及管治
進行
風險評估及人為監督
實行
AI 模型的定製與
AI 系統的實施及管理
促進
與持份者的
溝通及交流
出現重大改變時重新評估風險
調整AI系統以應對持份者的關注

制定AI 策略及管治

AI 策略包含多項要素，能展示管理層的決心和提供指引

AI 策略

作用

展示高級管理層有決心通過符規、合乎道德標準及負責任的方式採購、實施及使用AI

就採購AI 方案的目的以及如何實施和使用AI 系統提供相關指引

可包含的要素

訂定適用於機構在採購、實施及使用AI 方案方面的**道德原則**

列明AI 系統在機構中**不可接受的用途**

建立**AI 清單**，以幫助機構實施管治措施

就如何符規、合乎道德標準地採購、實施及使用AI 方案制定**具體的內部政策和程序**

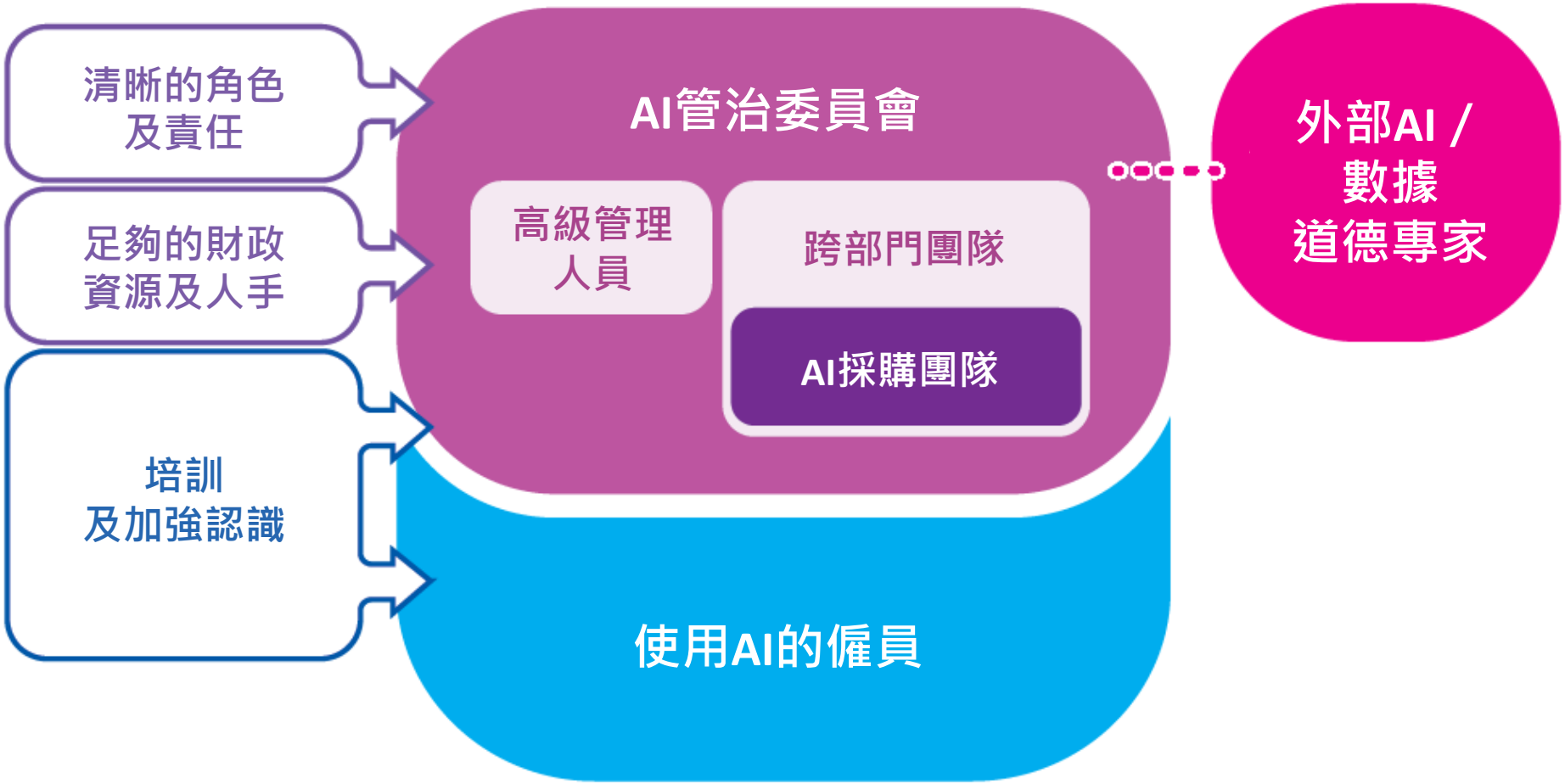
定期與所有相關人士**就AI 策略、政策和程序溝通**

考慮可能將會適用於AI 的採購、實施及使用的**法律和法規**

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

制定AI 策略及管治
管治架構



進行風險評估及人為監督

風險評估的程序

1. 由跨部門團隊在採購過程或對現有的AI系統進行重大更新時**進行風險評估**
2. **識別及評估AI系統的風險**
3. **因應有關風險而採取適當的風險管理措施**

風險類型	考慮因素
私隱風險	• 用來定制所採購AI 方案的資料及／或輸入AI 系統用作決策的資料的准許用途 • 個人資料的數量 • 所涉及資料的敏感程度 • 在使用AI 系統時的個人資料保安
道德風險	• AI 系統對受影響個人、機構及社會大眾的潛在影響 • AI 系統對個人的影響出現的可能性，以及其嚴重程度和持續時間

進行風險評估及人為監督

風險為本的人為監督

如 AI 系統**輸出的結果很可能對個人造成重大影響**，有關系統一般會被視為**高風險**。

較低 ***AI系統的風險程度*** *較高*

人在環外

AI在沒有人為介入下作出決定

人為管控

人類決策者監督AI的運作，在有需要時介入

人在環中

人類決策者在決策過程中保留控制權以防止及／或減低AI出錯

實行AI模型的定制與AI系統的實施及管理

過程	重點建議	例子
1 數據準備	確保遵從私隱法例的規定 收集最少量的個人資料 管理數據質素 妥善記錄處理數據的情況	• 一家**時裝零售平台**正計劃採購**第三方開發的AI聊天機械人**，並將其進行**定制**，以為客戶推薦**時裝建議** • 該公司或會認為需要使用**不同客戶群**過去的**購買記錄**和**瀏覽紀錄**來定制聊天機械人 • 然而，客戶的**姓名**、**聯絡資料**、某些**人口特徵等個人資料並非是需要的**

實行AI模型的定制與AI系統的實施及管理

過程	重點建議	例子
2 AI的定制及實施	☑ 對模型進行嚴格測試及驗證其可靠性、穩健性和公平性 在整合前，根據AI方案所托管的服務器的方式（在機構內部或在第三方的雲端）考慮循規事宜 確保系統安全及數據安全	• 一間**律師事務所**正**定制**第三方開發的**AI聊天機械人**，以協助其員工**草擬法律文件**及進行**文書工作** • 該事務所應提醒員工在使用AI聊天機械人時，**盡量避免輸入個人資料**及／或**客戶的機密資訊**

實行AI模型的定制與AI系統的實施及管理

過程	重點建議	例子
3 AI的管理與持續監察	將記錄妥善地存檔 定期進行審核 制定AI事故應變計劃 隨着風險因素演變而考慮採取檢視機制	• **人為監督**應以**避免及儘量減低AI對個人造成的風險**為目的。 進行人為監督的人員應： • 盡可能**瞭解AI系統的能力和限制**； • **避免過份依賴**AI輸出的**結果**； • **正確地解釋及評估AI輸出**的結果； • 在AI輸出的結果出現**異常時**，作出**標記並在適當情況下不理會、撤銷或推翻結果**；及 • 在AI供應商就AI系統輸出結果提供的資訊協助下，在適當情況**介入及中斷AI系統的運作**。

促進與持份者的溝通及交流

與持份者的溝通

- 披露AI系統的使用
- 提供充足的資訊
- 披露風險

與持份者的交流

- 容許拒絕使用AI及資料查閱和改正
- 按要求提供解釋
- 提供人為介入的選擇

《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

01 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**

02 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**

03 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

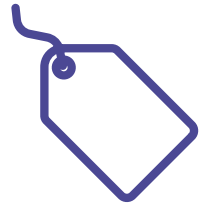
準確度及核實

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**

04 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

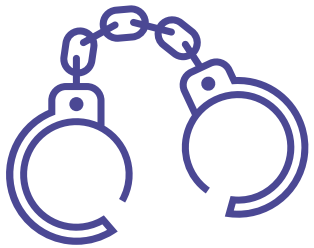
要求僱員根據機構的**AI事故應變計劃報告AI事故**

用戶憑證

要求使用**獨特且高強度的密碼**及**多重認證**

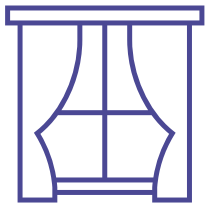
05 違反政策或指引後果和實用貼士

違反後果



- **訂明**僱員違反使用生成式AI政策或指引可引致的**後果**
- 參考公署《**模範框架**》**建議**，以制定生成式AI的**管治架構及措施**

支援僱員使用生成式AI工具的實用貼士



透明度

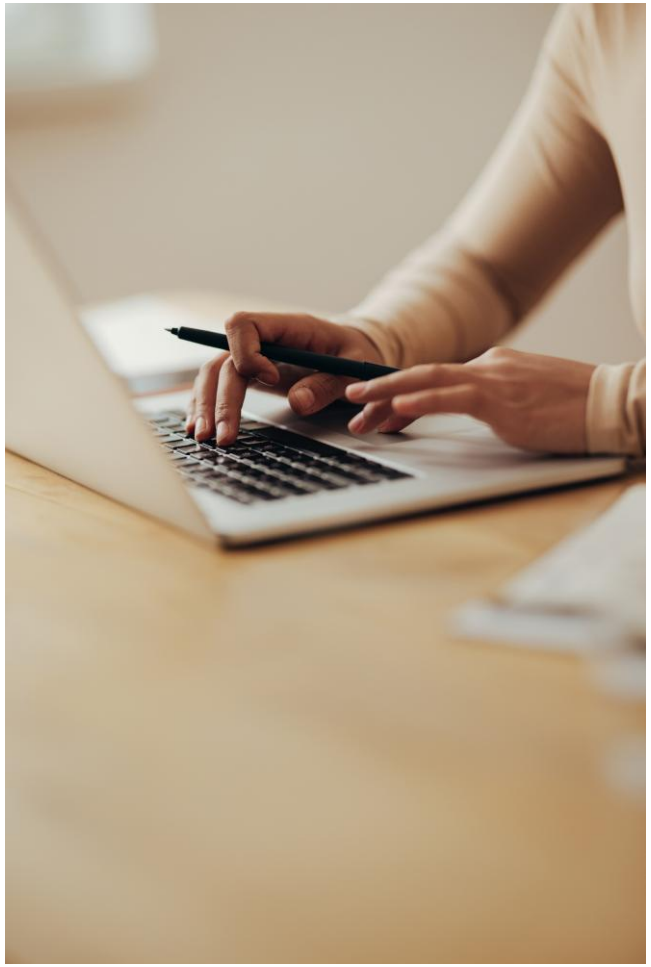
培訓及資源

委派支援隊伍

反饋機制

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
- 參與數據安全和AI安全相關的講座，以及
  申請內部培訓講座

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