



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

香港新聞博覽館

AI安全「智」重要： 如何應對人工智能時代的 私隱新挑戰

個人資料私隱專員
鍾麗玲女士

2025年8月2日



AI騙案

騙徒運用AI深度偽造製作詐騙影片



騙徒在視像會議運用深偽技術假扮公司高層，最終令公司損失近兩億元。



騙徒竄改公開片段，利用官員或名人的圖片或錄音，並透過深偽技術製作影片，誘騙市民投資。

資料來源: 星島日報

AI騙案

騙徒利用AI技術模仿聲線進行詐騙



騙徒以AI合成聲線，假扮成親朋戚友以索取金錢。

資料來源: e-zone, HK01

有圖未必有真相

製作深度偽造影片並不困難



4

有圖未必有真相

AI生成影片



5



AI帶來的私隱風險

1) 資料外洩風險

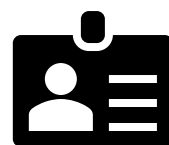


資料來源: [OpenAI](#)

解釋

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

例子



2023年3月，著名聊天機械人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



AI帶來的私隱風險

1) 資料外洩風險



AI生成圖片

解釋

使用線上AI聊天機械人時，用戶輸入的資料往往會傳送至AI服務提供者。若當中包含敏感或機密資料，會帶來資料外洩風險

例子



- 2024年，荷蘭一名診所員工未經授權將病人的醫療數據輸入到AI聊天機器人中
- 醫療數據屬於高度敏感的個人資料
- 將敏感數據分享給科技公司可能會侵犯相關病人的私隱

資料來源: [AP](#)

AI帶來的私隱風險

2) 資料收集過量



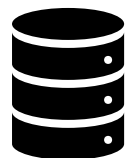
AI生成圖片

資料來源: Fortune

解釋

AI 傾向於收集和保留盡可能多的數據，包括個人資料

例子

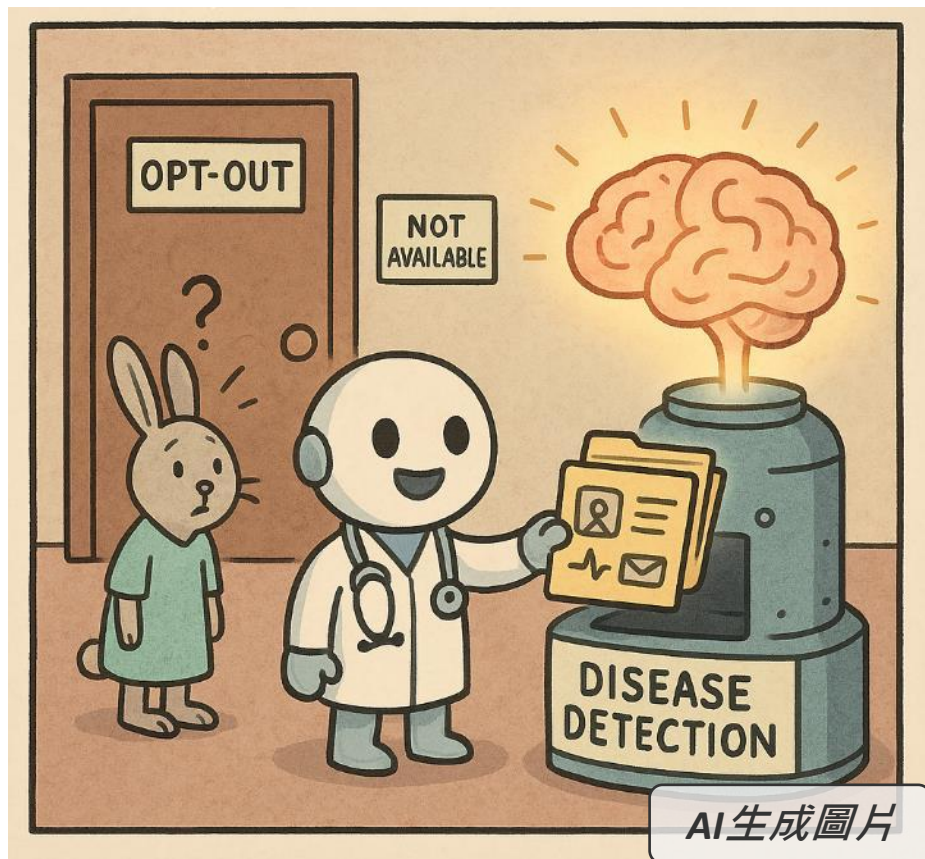


有科技公司被指擷取了3000億個網上詞彙來訓練聊天機械人



AI帶來的私隱風險

3) 資料的使用



解釋

AI系統開發者在資料當事人不知情或未得到其同意的情況下，將其個人資料用於訓練AI

例子

- 科技公司與英國醫療機構合作，獲取約160萬名病人的醫療紀錄，以開發偵測疾病的應用程式
- 機構沒有告知病人或取得其同意，亦沒有給予病人「退出」的選項
- 遭英國資訊專員辦公室裁定違法

資料來源: [ICO](#), [BBC](#)

9

AI帶來的私隱風險

4) 資料準確性



資料來源: CPO Magazine

解釋

即使AI系統中儲存了過時或不準確的個人資料，開發者亦未必能夠更正或刪除這些資料

例子



- 在奧地利，有AI聊天機械人被問及一位公眾人物的出生日期時，反覆提供錯誤的日期
- 開發商表示他們只能過濾或封鎖有關的查詢，無法透過修改訓練數據以更正AI系統的回應
- 該公司免責聲明亦指出，由於技術複雜性，部分資料可能無法更正

10

AI相關風險與相應的資料保障原則

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4原則

資料保安

- 外洩用戶對話數據

風險

許多企業均認為AI 對個人資料私隱構成顯著風險

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

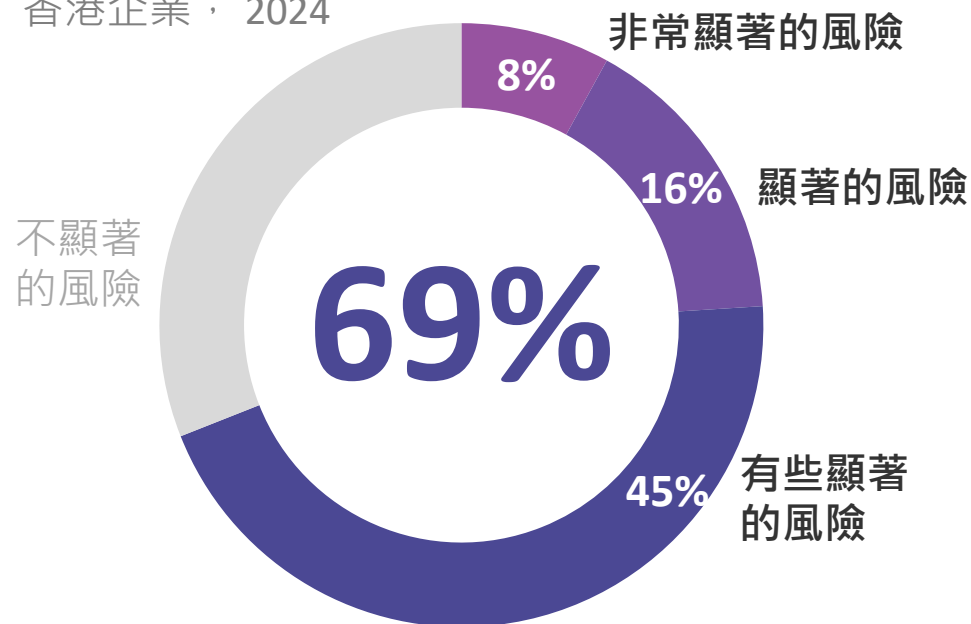
- 1  生成式AI
- 2  Cookies 和其他在綫追蹤器科技
- 3  雲端計算
- 4  物聯網
- 5  區塊鏈相關技術
- 6  數據分析及營運流程自動化

資料來源：私隱專員公署

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

循規審查

公署在於2023至2025年進行了針對使用AI的循規審查



2023年
8月至
2024年
2月

公署合共對**28間**本地機構進行循規審查，以了解該些機構在開發或使用**AI**時收集、使用及處理個人資料的情況，以及對於使用AI的管治情況

2025年
2月至5月

公署展開新一輪的循規審查，涉及**60間**本地機構，並涵蓋更多行業。除了第一輪循規審查的範疇，公署亦審視了機構實施《模範框架》建議和最佳行事常規的情況

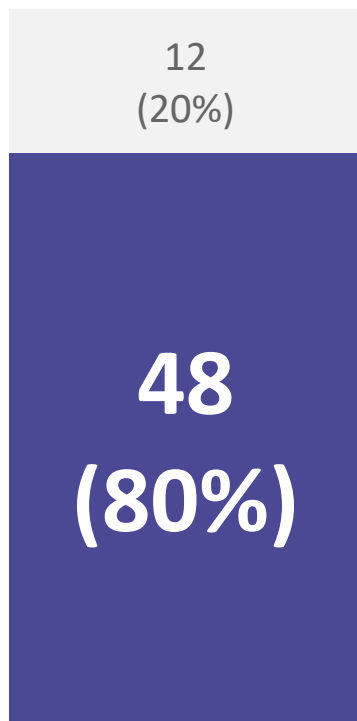
結果

公署在兩次循規審查過程中**未有發現違反《私隱條例》**的情況

2025年循規審查

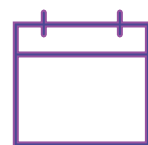
大部分機構在業務營運中使用AI

八成機構
在日常營運中使用AI



有使用
AI

48間在日常營運中使用AI當中：



使用AI超過一
年

42 (88%)



使用三個或以
上的AI系統

26(54%)



透過AI系統收
集及 / 或使用
個人資料

24 (50%)

2025年循規審查

24間會透過AI系統收集及 / 或使用個人資料的機構當中：



在實施AI系統
前會進行測試

23 (96%)



在採用AI系
統前有進行私
隱影響評估

20 (83%)

2025年循規審查

24間會透過AI系統收集及 / 或使用個人資料的機構當中：



已制定資料外洩事故應變計劃

22 (92%)

7 (32%)

15 (68%)

涵蓋專門針對AI相關的資料外洩事故



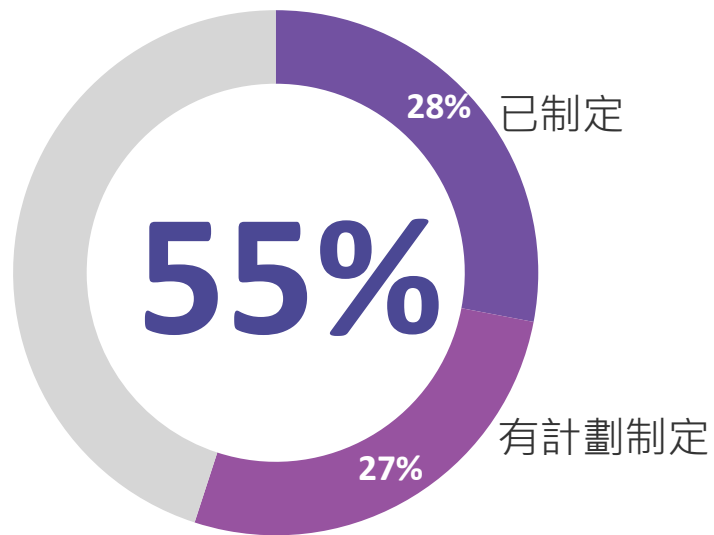
設有AI管治架構

19 (79%)

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)：



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

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)

使用AI聊天機械人「自保」十招

公署發出單張協助市民使用AI聊天機械人時保障其個人資料私隱



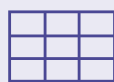
背景

生成式AI技術驅動的AI聊天機械人的使用日益普及



發佈單張

公署2023年發出《使用 AI 聊天機械人「自保」十招》，協助用戶保障其個人資料私隱，並更安全地使用AI聊天機械人



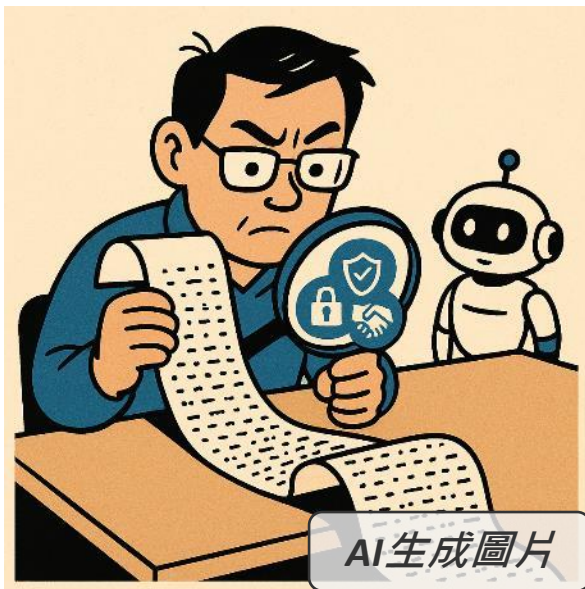
三大範疇

1. 在登記或使用前
2. 與AI聊天機械人對話時
3. 安全及負責任地使用AI聊天機械人

使用AI聊天機械人「自保」十招

保障個人資料私隱

在登記
或使用
前



1. 細閱私隱政策、使用條款及其他有關處理資料的政策



2. 慎防假冒AI聊天機械人的應用程式和釣魚網站 (例如DeepSeek)



3. 調整設定選擇拒絕分享聊天紀錄 (如有)

與AI聊
天機械
人對話
時

使用AI聊天機械人「自保」十招

保障個人資料私隱

在登記
或使用前



4. 不要分享自己及他人的個人資料



6. 對網絡安全威脅提高警覺

與AI聊天機械人對話時



5. 如有需要，提交改正或移除資料要求



7. 刪除聊天紀錄中不需要保留的對話

使用AI聊天機械人「自保」十招

安全及負責任地使用



8. 謹慎使用AI聊天機械人提供的資料



9. 不要分享機密資料和檔案



10. 老師 / 家長應在同學與AI聊天機械人對話時從旁指導

22

《僱員使用生成式AI的指引清單》



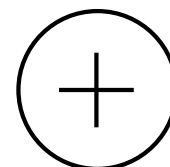
🎯 目的

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

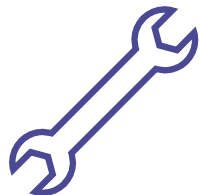
05

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

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



獲准許的用途



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



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



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

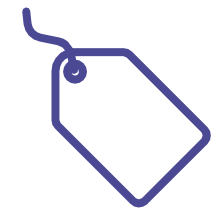
強調僱員需要核
實AI所提供的資
訊



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

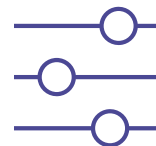
僱員使用生成式AI的政策或指引的建議內容 – 數據安全

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**

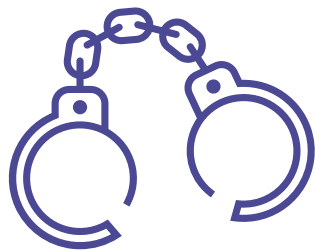
用戶憑證



要求使用**獨特且高強度的密碼及多重認證**

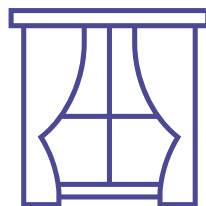
05 違反政策或指引後果和實用貼士

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



委派支援隊伍

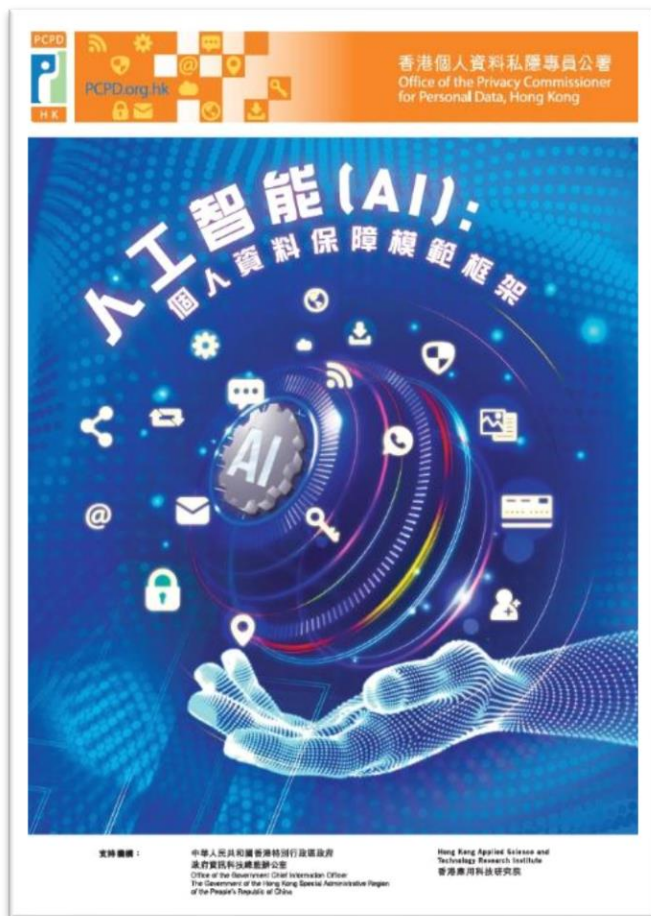


培訓及資源



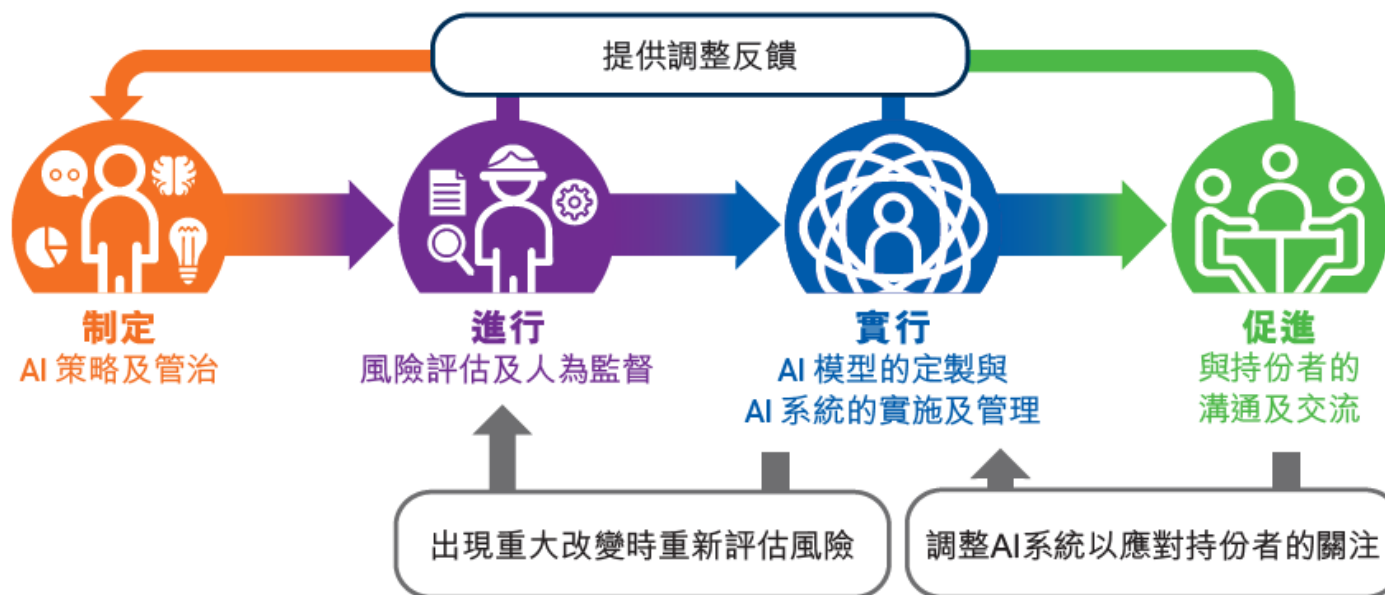
反饋機制

《人工智能 (AI): 個人資料保障模範框架》



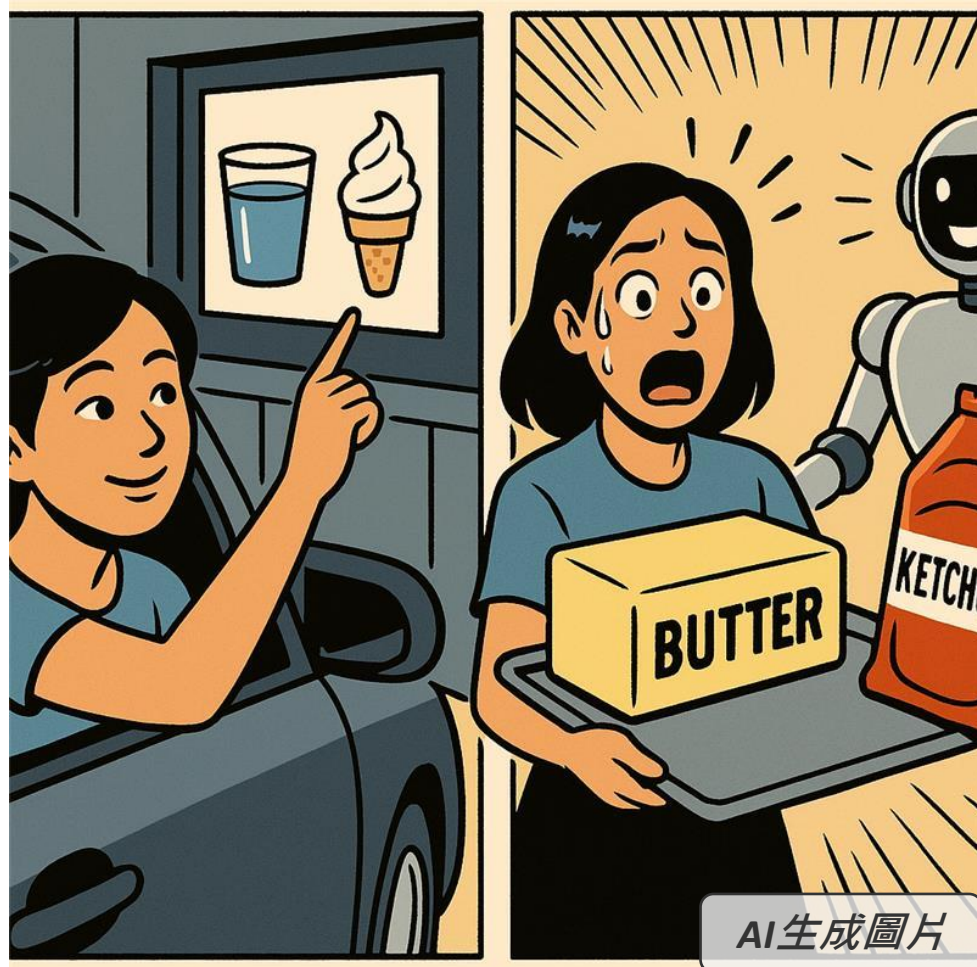
i 協助機構遵從《私隱條例》的規定

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



快餐店事件

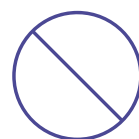
一家連鎖餐廳因AI下單系統頻頻出錯，最終決定停止使用



試行AI驅動的語音下單系統



AI系統無法準確辨識顧客指示，造成混亂



最終於2024年年中終止試用該系統

《人工智能 (AI): 個人資料保障模範框架》



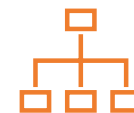
制定
AI 策略及管治



AI 策略



管治考慮



管治架構



培訓及
加強認識

較低

AI系統的風險程度

較高



進行
風險評估及人為監督



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

《人工智能 (AI): 個人資料保障模範框架》



實行
AI 模型的定製與
AI 系統的實施及管理



數據準備



AI的定製
及實施



管理與
持續監察



促進
與持份者的
溝通及交流



提供資訊



可解釋的 AI



資料當事人
的權利及反饋



語言及方式

AI 事故應變計劃

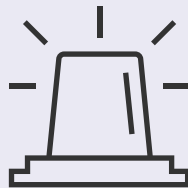
計劃可包含以下六項要素

1



界定AI事故

3



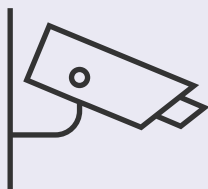
通報AI事故

5



調查AI事故

2



監察AI事故

4



遏止AI事故擴大

6



從AI事故復原

想知更多？

瀏覽私隱專員公署專題網頁

防騙貼士 專題網頁



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

主頁 > 防騙貼士

防騙貼士

私隱專員教你防騙六招
Privacy Commissioner Offers
Six Tips to Prevent Fraud

個人資料防騙熱線
Personal Data
Fraud Prevention Hotline
3423 6611

- 1 切勿隨意披露個人資料
Do not disclose
personal data
to others arbitrarily
- 2 留意帳戶及簽帳紀錄
Keep an eye on your
accounts and
transaction records
- 3 切勿向任何人透露帳戶密碼
Never share passwords
with anyone
- 4 減少在網上分享個人的
正面照片及影片
Minimise the online sharing
of portrait photos and videos
- 5 核實來電者身分
Authenticate the
identity of callers
- 6 留意最新防騙資訊
Pay attention to the
latest fraud
prevention information

- + 防騙六招
- + 相關例子
- + 教育資訊
- + 防騙報道及活動
- + 求助途徑

人工智能安全 專題網頁



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

主頁 > 人工智能安全 新!

人工智能安全

私隱專員公署
發布的人工智能指引

AI安全熱線
2110 1155

- + A. 私隱專員公署有關AI的指引資料
- + B. 教育資訊／資源
- + C. 國際發展概覽
- + D. 私隱專員公署有關AI的最新消息及活動
- + E. 2025年私隱之友嘉許獎 – 最佳人工智能管治獎
- + F. 有用資訊及連結

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



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

