《香港家書》 2024 年 6 月 29 日（香港電台第一台）

個人資料私隱專員鍾麗玲 ：

採用 AI 模範框架 減私隱風險

欣欣：

最近工作忙碌嗎？近年生成式人工智能（AI）工具越來越普及，你又有沒有在工作中運用這些工具呢？上星期與你姊姊吃飯時，我們也有提起 AI 這個時下熱門話題。她大讚生成式 AI 令她的工作量減輕了不少，她說只需向聊天機械人輸入一兩句指示，就可得到一篇大方得體的文件初稿，甚至幫助她分析財務報表。也難怪現時許多機構都打算採購各種 AI 系統用於營運之中，從而提升生產力。

不過，值得留意的是，聊天機械人提供的資訊，很多時並非百分百準確。聊天機械人有些時候甚至在它的回應中，以出乎意料的方式向你洩漏個人資料。舉例來說，若你向提供予公司內部員工使用的聊天機械人輸入自己的個人資料，聊天機械人或會將你輸入的資料儲存至它的資料庫，並且在回答其他員工的問題時披露這些資料，構成私隱風險。此外，去年有研究人員發現，當他們要求市場上某聊天機械人不斷重複一些單字時，它雖然會先按指示重複該單字，但不久之後，它就會開始輸出訓練資料，當中包括姓名、電話、電郵地址等個人資料。欣欣，我相信你也知道，訓練 AI 模型需要海量數據，隨

着 AI 模型不斷「升呢」，它們對數據的「胃口」越來越大。 有報道更指出，一個知名品牌的聊天機械人模型的開發者從互聯網上收集了多達 3000 億字來訓練模型，而這些數據很大機會包含個人資料。

此外，AI 輸出的資訊可能不準確，具歧視性或帶有偏見。美國有研究顯示，許多人臉識別技術系統錯誤識別有色人種面孔的可能性比白人高達 10 至 100 倍，原因是訓練數據中屬有色人種的數據不足。若 AI 系統是用於診斷病人，但分析錯誤，導致醫生「斷錯症」、延誤治療，後果就不堪設想。事實上，AI 系統「擺烏龍」的情況時有發生。據報道，美國有一名血癌病人被 AI 判斷為患有敗血症，因此需要進行抽血等進一步醫療程序，結果卻發現 AI 誤診，白白地增加了病人的感染風險和醫藥費。

另一方面，AI 的數據安全風險同樣不容忽視。AI 系統與其他資訊科技系統一樣，有資料外洩風險。我們身處數碼年代，無論是香港或其他地區，資料外洩事故都呈現上升的趨勢。同樣地，AI 系統亦可能因為黑客入侵或系統設定出錯等不同原因而導致資料外洩。去年正正有一宗牽涉 AI 聊天機械人的嚴重資料外洩事故，除了洩露了部分用戶過往對話的標題，更洩露了用戶的姓名、電郵地址和信用卡號碼的部份數字。

針對香港機構使用 AI 的情況，私隱專員公署於今年初已就 28 間機構使用 AI 的情況完成循規審查，並未有發現任何違規情況。為應對 AI 所帶來的風險，公署早前出版了《人工智能（AI）：個人資料保障模範框架》（《模範框

架》），協助機構在採購、實施及使用 AI，包括生成式 AI 處理個人資料時，遵從《私隱條例》的相關規定。這套《模範框架》建基於一般業務流程，涵蓋四個範疇，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，包括建議機構應制定 AI 策略及管治架構、評估相關風險、採取「風險為本」的方式以決定所需的人為監督程度、實行 AI 模型的定製與 AI 系統的實施及管理，以及促進與持份者的溝通及交流。業界對公署推出《模範框架》，反應相當正面，許多機構都認同《模範框架》有助它們安全、合規地使用 AI。《模範框架》以淺白的文字，具體的個案、例子、圖表說明怎樣在採購、實施及使用 AI 模型時保障個人資料，並引入「紅隊演練」、AI 事故應變計劃等建議，提升系統安全和數據安全。我希望《模範框架》可幫助中小企對症下藥，更安心地採購、實施及使用各式各樣的 AI 系統。

為了協助公眾和機構理解《模範框架》的建議，我們亦已推出懶人包，將《模範框架》濃縮至兩頁簡單又易明的單張，令到大家可以透過該單張輕易掌握框架的精髓。

我相信在採取適當的保障措施下，AI 的使用將可以為大家帶來更大的益處。機構若對《模範框架》的內容有甚麼疑問，可致電私隱專員公署的中小企熱線查詢。我們也會主動為業界提供協助，幫助機構將《模範框架》的各項建議納入至他們的業務流程之中，提升 AI 管治水平。公署將會舉辦講座及工作坊，為業界詳細介紹框架內容。

欣欣，我深信《模範框架》會協助孕育 AI 在香港的健康及安全發展，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。

最後姨姨想提提你，機構固然有責任在使用 AI 之餘，保障用戶的個人資料；而你我作為不同 AI 工具的使用者，也要學懂保障自己及他人的個人資料私隱。你使用聊天機械人時，要切記不可隨便透露任何個人資料呀。私隱專員公署早前都出版了《使用 AI 聊天機械人「自保」十招》，閒時記得看看這單張，學多幾招傍身、保障自己的私隱！

姨姨

2024 年 6 月 29 日