《香港家書》 2023 年 4 月 29 日（香港電台第一台）

個人資料私隱專員鍾麗玲 ：

人工智能系統有「機」亦有「危」　科企確保數據安全責無旁貸

王教授：

近期人工智能（AI），特別是生成式 AI 驅動的聊天機械人的崛起，例如 ‘ChatGPT’、 ‘Bard’、「文心一言」和「通義千問」，已在不同司法管轄區和各行各業引起了震盪。不知道你有否使用過這些聊天機械人，去幫助你教學或者研究的工作呢？

與早期主要透過分析大數據作出決策的人工智能不同，聊天機械人的「神奇」之處在於它們可以秒速回應用戶的請求和問題，並根據所給予的提示，像人類般提供看來理性、完整及令人信服的回應，甚至可用來創作詩詞歌賦，這使它成為人工智能發展史上的一個里程碑。

最近有報道指，生成式 AI 的邏輯理解和分析能力已相當接近人類，例如用生成式 AI 去解答 DSE 中史科試卷，可以 1 分鐘內完成試卷並得到 5 級的佳績。最近更有世界攝影比賽的冠軍作品是由人工智能生成。究竟我們面對的是真實的人的作品還是人工智能的作品，真是可謂「人機難辨」！

然而，有「機」亦有「危」。雖然很多人視這科技突破為人類的福音，但人工智能的使用所帶來的私隱及道德問題同時引起各界關注。

很多人可能不知道，生成式 AI 的運作建基於「深度」學習技術，當中涉及使用大量的原始資料或數據作為訓練數據，而這些資料或數據可以包括敏感的個人資料，可以是由用戶所提供的，或是由互聯網收集及複製的。

試想一想，若你的朋友在你完全不知情、遑論同意的情況下，將你的姓名、職位、住址、身份證號碼，甚至信用卡號碼提供予聊天機械人作為對話的一部分，視乎有關程式的演算法，這些個人資料便可能會被儲存到機械人的資料庫中，並可能成為它回答下一個用戶的材料！

從以上的例子可以看到，由於用戶的對話可能被用作訓練人工智能模型的新訓練數據，若是用戶不經意地向人工智能系統提供了敏感的資料，那麼他們的個人資料便可能會被濫用，超出用戶提供他們個人資料的原本目的。同時，若是相關的資料收集過程透明度不足，又是否會出現在用戶不知情及不公平的情況下收集個人資料，從而構成私隱風險呢？除卻資料的性質及收集過程外，還有其他同樣複

雜的問題有待考慮及處理，包括用戶可如何行使他們查閱及改正資料的權利、如何保護兒童私隱、在開發及使用生成式 AI 的過程中如何減少不準確的資料、具歧視性或偏見的內容，以及資料安全的風險。

最近更有報道指，有些員工因為使用聊天機械人而將公司機密資料，例如公司系統原始碼、甚至一些醫療紀錄放入聊天機械人的系統裏，引來企業機密資料和私隱外洩的憂慮。

最近，全球超過二萬七千人，當中包括科技專家，聯署了一封公開信，呼籲所有人工智能實驗室六個月內暫停訓練更強大的生成式 AI 系統，並在這期間共同制定和實施一套安全守則。

究竟現時是否將人工智能的發展好好地規劃及管理、甚至規管的契機？

我認為科技公司，尤其是人工智能開發科企，有責任確保人工智能系統的數據安全，相關科企應檢視及嚴謹地評估他們的人工智能系統在運作上對數據私隱及道德的影響，並確保遵從相關法例或適用的指引。

在這方面，私隱專員公署於 2021 年 8 月發出《開發及使用人工智能道德標準指引》，協助機構以保障私隱、符合道德及負責任的方式開發及使用人工智能系統。《指引》臚列了國際認可的人工智能道德原則，涵蓋問責、人為監督、透明度與可解釋性、公平、數據私隱、有益的 AI，以及數據可靠、穩健及安全等方面的標準。機構在開發及使用生成式 AI 時，除了須要符合《個人資料（私隱）條例》的規定外，亦應該遵守這些原則，以減輕私隱及道德風險。

展望未來，生成式 AI 的出現已經是一個不可逆轉的事實。我贊成對這新興科技採取一個謹慎開放的態度，在不窒礙新科技長遠發展的同時，亦應考慮以法律法規、指引、行業標準，或甚至國際標準去作出規範，令生成式 AI 在一個健康、合法及符合道德的空間中發展。王教授，你從事科技鑽研多年，在這方面我也想聽一下你的意見。

麗玲

2023 年 4 月 29 日

（1,515 字）