微軟出錯，將機密電郵暴露喺 AI 工具 Copilot

2 日前

分享 保存

利夫·麥克馬洪
科技記者

盧克圖片

微軟承認咗一個錯誤，令到佢嘅 AI 工作助理錯誤噉存取同總結咗部分用戶嘅機密電郵。

呢間科技巨頭推動咗微軟365副機師聊天，作為職場同佢哋嘅員工使用佢嘅生成式人工智能聊天機械人嘅安全方法。

但係佢話最近嘅一個問題令到呢個工具向部分企業用戶展示咗儲存喺佢哋草稿同傳送嘅電郵資料夾入面嘅訊息嘅資料 —— 包括標記為機密嘅訊息。

微軟話佢哋已經推出咗一個更新嚟解決呢個問題，而且佢哋「冇畀任何人存取佢哋未獲授權睇到嘅資料」。

不過，有專家警告，公司競爭加入新 AI 功能嘅速度，意味著呢類錯誤係必然嘅。

副機師即時通訊可以喺 Microsoft 程式入面用，例如 Outlook 同 Teams ，用嚟發送電郵同即時通訊功能，嚟獲取問題嘅答案或者總結訊息。

微軟發言人同 BBC 新聞講：「我哋發現同解決咗一個問題，即係 Microsoft 365副機師聊天可能會喺由用戶撰寫並儲存喺 Outlook 桌面嘅草稿同已發送項目入面嘅標籤為機密嘅電郵入面返回內容。」

佢哋補充：「雖然我哋嘅存取控制同資料保護政策保持完整，但呢個行為並唔符合我哋預期嘅 Copilot 體驗，而呢個體驗係為咗將受保護嘅內容排除喺 Copilot 存取之外。」

「已經為企業客戶喺全球部署咗一個配置更新。」

呢個錯誤首先由科技新聞媒體**Bleeping Computer**報導，佢哋話佢哋見到一個服務警示確認咗呢個問題。

佢引用咗微軟嘅通知，話「微軟 365 Copilot 即時通訊處理錯誤嘅係用戶嘅電郵訊息，而呢啲訊息係套用咗機密標籤」。

通知補充， Copilot Chat 入面嘅工作標籤已經總結咗儲存喺用戶草稿同傳送資料夾入面嘅電郵訊息，就算佢哋有敏感標籤同埋設定咗防止未經授權嘅資料共享嘅防止資料遺失政策。

報導指出，微軟喺一月先發現呢個錯誤。

佢對呢個錯誤嘅通知亦都喺英國 NHS 工作人員嘅支援資訊主頁上面分享 - 根本原因係歸因於「程式碼問題」。

NHS IT 支援網站上面嘅通知有部分暗示佢受到影響。

但係佢同 BBC 新聞講，任何由 Copilot Chat 處理嘅草稿或者發送嘅電郵嘅內容都會保留喺佢哋嘅創作者手上，而病人資料亦都冇被公開。

「數據洩露會發生」

企業 AI 工具，例如 Microsoft 365副駕駛聊天 - 訂閱咗 Microsoft 365嘅組織可以使用 - 通常會有更嚴格嘅控制同安全保護，以防止共享敏感嘅企業資料。

但對於一啲專家嚟講，呢個問題仍然突顯咗喺某啲工作環境下採用生成式人工智能工具嘅風險。

Gartner 嘅數據保護同人工智能治理分析師 Nader Henein 話，鑑於「新同新穎嘅人工智能功能」發佈嘅頻率，「呢種失誤係避免唔到嘅」。

佢同 BBC 新聞講，用呢啲人工智能產品嘅機構通常缺乏保護自己同管理每個新功能所需嘅工具。

Henein 話：「喺正常情況下，組織只會閂咗呢個功能，等治理追上嚟。」

「可惜，由未經證實嘅人工智能炒作所引起嘅壓力，令到呢個幾乎係唔可能嘅，」佢補充道。

薩里大學嘅網絡安全專家 Alan Woodward 教授話，呢個表明咗將呢類工具預設為私有同埋只可以選擇加入嘅重要性。

「呢啲工具難免會有錯誤，尤其係因為佢哋以極快嘅速度前進，所以就算資料漏洞未必係有意嘅，都會發生。」佢同 BBC News 講。

亞馬遜股價下跌，因為佢加入咗 **Big Tech** 人工智能支出狂潮
Xbox 製作人叫員工用 **AI** 去緩解失業嘅痛苦
微軟推出人工智能截圖工具被稱為「私隱噩夢」
