AI 風險：別跟 ChatGPT 聊太多，個資可能被偷記？

OpenAI 創辦人之一兼執行長奧特曼（Sam Altman）坦言，AI 可能對網路資訊安全的危害程度之高，讓他覺得不堪設想。

Google 和 OpenAI 都多次以「需要處理好潛在危險」為理由，延遲新功能的 AI 產品公開時程。

但近期微軟大動作推廣 ChatGPT，甚至採取訂閱制，並積極與其搜尋引擎 Bing 進行深度整合；Google 也被動推出了 Bard 聊天機器人來應戰，究竟這是「出奇招」、還是「下險棋」？

ChatGPT 在公開版本中，宣稱自己的安全與道德圍牆，舉凡犯罪、恐怖主義、種族歧視、仇恨言論、性騷擾等不道德的問題，或是涉及個人隱私、資安攻擊的問題，都會被拒絕回答。但在實際實驗中，初期版本就曾被國外資安專家「設局」，透過問題的誘導與情境設定，寫出毀滅人類的計劃書，詳細描述入侵各國網路、控制武器、破壞基礎建設等 SOP，還提供了對應的 Python 程式碼，說明其安全圍牆還是可能被繞過。

ChatGPT 目前是全球最潮的技術，如果對其優勢過度樂觀，反而可能減少其學習改進的機會，若能集眾群力找出其缺失並加以修正，從錯誤中學習，相信 ChatGPT 在發展之路上將走得更穩健，並對普羅大眾做出更大的貢獻！

文章出處:遠見 https://www.gvm.com.tw/article/100138