「AI教父」警告人類未來恐難辨真假

從谷歌離職 稱對人工智能研究工作感後悔

【大公報訊】綜合《紐約時報》、BBC報道：有「AI教父」之稱的辛頓（Geoffrey Hinton）本月1日宣布，他已經從谷歌公司離職，並稱現在對自己一生從事的工作感到後悔。辛頓表示，辭職是為了能夠自由地談論AI的風險。辛頓認為，AI系統保持着驚人的發展速度，它們會變得越來越危險，「目前它們不比我們聰明，但我認為很快就會」、「很難看到如何防止壞人利用它做壞事」，以及對AI引發假資訊的擔憂，普通人將「無法再知道什麼是真的」。

75歲的辛頓日前在接受《紐約時報》的採訪時證實自己已於上周辭去了谷歌副總裁的職位，離開工作了10年的谷歌。他稱對自己畢生的工作感到後悔，「我會安慰自己，如果不是我做，也會有別人來做」。至於離開的原因，辛頓表示並不是為了批評谷歌，而是希望可以自由地談論AI的風險，無需顧忌言論對谷歌構成影響，同時他也強調谷歌已對此採取非常負責任的行動。

AI很快就會比人類聰明

辛頓2013年3月加入谷歌，擔任研究員，同時他創辦的DNNresearch公司被谷歌併購。他研究深度學習、神經網絡等領域，曾與其他兩名學者因在深度學習領域作出的貢獻，榮獲2018年圖靈獎，該獎項常被稱為「電腦科學界的諾貝爾獎」。

辛頓受訪時表示，AI與人類的智力不同，AI的學習就像是一萬人當中，只要有一人學懂某些知識後，其他人就會自動習得知識，因此它們龐大的知識儲量已經讓人類黯然失色。在推理方面，AI現在只能做到簡單的推理，「目前它們不比我們聰明，我之前覺得它們可能要30到50年才能超過人類，但我現在覺得會更快」。

辛頓在離開谷歌之前就已經發出警示。他3月接受採訪時被問到「AI抹除人類的幾率」有多高，他回答道：「這不是不能想像的，我只能這麼說。」這也是辛頓擔心的另一方面，未來AI會消滅「繁重的工作」，甚至會直接取代人類工作。

憂AI落入壞人之手

AI技術發展到如今，它可以輕鬆地大批量生成假照片、假視頻以及假新聞，未來網絡上可能充斥着虛假信息。辛頓更加擔憂的是，普通人未來將「無法再知道什麼是真的」。

辛頓認為，在短期內，生成式AI可能會引發假資訊的狂潮；更長遠來看，AI很可能會生成完全自主武器，並從海量訓練數據庫中習得奇怪行為。雖然有些擔憂還未發生，但辛頓認為若沒有及時準備好相關法規和有效的控制措施，那這種升級將變得無法遏止。他擔憂AI風險的另一個原因就是「很難看到如何防止壞人利用它做壞事」。

辛頓對AI的立場是從去年開始轉變的，他認為AI在過去五年內發展神速，而未來五年可能發生的事情將「非常可怕」。實際上，業界還有不少專家都曾提議放慢AI的發展節奏。以馬斯克為首的多名科技界巨頭3月聯署公開信，呼籲AI實驗室和相關企業在至少半年內不要訓練超過GPT-4的技術。

不過辛頓補充道，他認為AI在短期內帶來的好處多於風險，所以現在不應該停止開發；加上激烈的國際競爭也意味着很難暫停AI發展，而政府有責任確保AI技術發展的同時，盡量杜絕AI被不道德利用的現象。

▲有「AI教父」之稱的辛頓。 路透社

專家觀點

- GPT-4之類的技術讓AI機器人擁有大量數據，這龐大的知識量已經超越人類，雖然AI在推理方面還不如人類，目前只能做到簡單推理，但辛頓認為按照這樣的發展速度，AI「很快就會比人類聰明」。
- 生成式AI或將引發假資訊狂潮，未來普通人將無法辨別到底什麼是真的。
- 未來AI會消滅「繁重的工作」，甚至會直接取代人類工作。
- 很難防止壞人利用AI做壞事，至少目前還沒有看到能夠有效杜絕的方法。

網友評論

@savingtradition：
- 如果它傷害了社會，那麼進步就不是真正的進步。

@WhichbufferArda：
- 我相信人工智能革命很像互聯網。政府可能會制定政策和法律來限制濫用，但這需要很長時間。

@NYTupelo7：
- 每當我想到AI的未來時，我都會想到「終結者」。

@primalpoly：
- 由此可見，我們應該警惕那些為科技巨頭工作的人，他們試圖向公眾保證AI「一切都好」。

生成式AI 恐淪假消息溫床

● ● ● 假新聞 ● ● ●

- 聊天機器人ChatGPT近期多次被揭發編造「不存在的新聞報道」，例如聲稱美國一法律系教授性騷擾學生，污衊澳洲赫本郡郡長胡德曾因賄賂而坐牢，還有偽造屬於英國《衛報》的報道等。

◀AI生成的法國總統馬克龍撿垃圾圖。 網絡圖片

● ● ● 假照片 ● ● ●

- 只需向生成式AI提供關鍵詞，它就可以快速且大量地生成以假亂真的圖像。近期，特朗普在曼哈頓被警方逮捕、教宗身着羽絨服、蘋果CEO庫克街頭撿垃圾等由AI生成的照片在網上瘋傳。德國攝影師埃爾達格森用AI生成的照片《偽記憶：電工》，獲得「索尼世界攝影大獎」（SWPA）獲得公開組創意類作品一等獎。

◀由AI生成特朗普在紐約曼哈頓被捕照片在網上瘋傳。 網絡圖片

● ● ● 假視頻 ● ● ●

- 當外界認為視頻更難被人為篡改的時候，AI卻可以「深偽」(deepfake)技術，給視頻中的任務「改頭換面」，例如生成美國前總統奧巴馬的講話聲音和畫面。 大公報整理

「AI教父」是誰?

- 辛頓是英國出生的加拿大計算機學家和心理學家，目前擔任多倫多大學計算機科學系教授。他以其在類神經網絡方面的貢獻而聞名，是機器學習領域的加拿大首席學者，也是加拿大高等研究院贊助的「神經計算和自適應感知」項目的領導者。
- 辛頓被譽為「AI教父」、「深度學習之父」和「神經網絡之父」。他曾因在深度學習方面的貢獻與本吉奧和楊立昆一同被授予了2018年的圖靈獎。
- 辛頓於2013年3月加入谷歌，同時他創辦的DNNresearch公司被谷歌併購。為了能自由討論AI潛在風險，辛頓日前宣布從谷歌離職。

▲2018年，辛頓慶祝獲得圖靈獎。 網絡圖片

AI生成新聞網站 誤報「拜登去世」

▲由AI生成的美國總統拜登身穿白色羽絨服照片。 網絡圖片

【大公報訊】據彭博社報道：美國新聞評級組織NewsGuard周一發布報告指出，發現有近50個由AI聊天機器人生成的新聞網站，其中甚至有網站誤報「美國總統拜登已去世」，引發外界擔憂AI工具助長虛假資訊傳播的風險。

這49個生成的新聞網站內容五花八門，有的提供突發即時新聞，網站名字也很普通，如News Live 79和Daily Business Post；有的介紹生活風格、名人消息或發布贊助內容。而這些網站都沒有提示文章是通過ChatGPT或Bard生成的。它們大多數都是品質低劣的「內容農場」，透過大量發帖吸引點擊、植入廣告，且有英文、葡萄牙文、泰文等許多語種的文章。

今年4月，有一個名為Celebrities Deaths的網站發表了一篇由AI生成的文章，稱「拜登逝世，副總統哈里斯將出任代理總統，上午9時將發表致辭」，還有網站更編造出訃聞。

NewsGuard行政總裁克羅維茲表示，該報告顯示，「已知這些AI會編造虛構的文章，還用它們去製作那些看似是新聞網站的內容，是一種偽裝成新聞業的欺詐行為」。

三星禁員工在公司電腦使用ChatGPT

【大公報訊】據法新社報道：韓國科技巨頭三星電子2日宣布，由於發現相關技術遭到「不當使用」的案例，禁止手機與家電部門員工利用公司電腦使用聊天機器人ChatGPT等生成式AI服務。

三星內部人士表示，這道禁令適用於移動裝置與家電部門員工。另據一份內部備忘錄，三星稱正努力尋找「讓員工在安全獲得保障的環境下」使用生成式AI服務的方法，「以兼顧工作效率與便利性」；「在這些措施到位之前，暫時禁止在公司電腦上使用生成式AI服務。」

三星同時建議員工，在公司以外場所使用私人電腦登錄這類平台時，也盡量不要上傳與工作有關的信息。三星稱，有出現員工「不當使用」ChatGPT這類功能的「一些情況」，但未透露更多細節。

三星的內部調查發現，有超過6成的員工認為，利用公司裝置使用生成式AI平台，將導致安全風險。近幾個月來，包括高盛在內的一些知名金融公司已禁止或限制員工使用ChatGPT這類平台。

▲韓國三星電子禁止手機家電部門員工用公司電腦使用ChatGPT。 網絡圖片