

AI滅絕人類推論 專家兩極

Twitter WhatsApp Facebook 讚好 19

A+ A- 分享 打印



AI應像孩子般接受訓練，向其灌輸人類的初始價值觀。

浸會大學宗教及哲學系教授 皮耶塔里寧



圖4-1



【明報專訊】美國科網初創企業 OpenAI 去年 11 月推出人工智能 (AI) 聊天機械人，其應對和生成內容能力技驚四座，自此 AI 成為全球炙手可熱話題，各界亦致力研究如何加強應用 AI 改善工作效率。不過，AI 崛起的同時亦帶來隱憂，有專家警告在目前各國缺乏有效監管下，AI 的能力持續突飛猛進或最終會危及人類存亡，當中多名專家、學者和業界翹楚先後發表聯署信警告，聲勢尤盛。另一方面亦有學者認為有關威脅被誇大，反而會令人忽視 AI 正造成的即時問題。雙方大辯論鏖戰正酣之際，本報訪問了 3 名支持「末日威脅論」的香港學者，讓他們詳述他們對 AI 的憂慮、應對威脅的方法和對規管的看法等。

明報國際組

港學者聯署警示 AI 風險 促立法規管灌輸人類價值

人們對 AI 的關注，近期由以往的各行業工種被取代、自動駕駛安全性、學生「請槍」生成作業甚至論文，以至假資訊氾濫，提升至認為 AI 的威脅足以跟核戰及氣候變化等議題看齊。今年 3 月，1300 多名科技領袖和研究人員發表公開信，希望所有比最新公開聊天機械人 GPT-4 更強大的 AI 能夠暫停訓練起碼 6 個月，以讓業界開發「共同安全協議」，確保 AI 系統不會對人類造成威脅，聯署人數至今累積至逾 3 萬。

逾 350 名全球學者及科技業界高層，5 月底參與 AI 安全中心 (CAIS) 的聯署公開信，警告各國應致力降低 AI 技術所帶來的「人類滅絕」風險，並將應對 AI 帶來的人類滅絕風險「與應對疫症大流行和核戰，視為同級的全球優先要務」。參與聯署、有「AI 教父」之稱的加拿大多倫多大學電腦學系榮譽退休教授欣頓 (Geoffrey Hinton) 早前承認對自己的工作感到後悔，憂慮未來科技會威脅人類。有聯署的 OpenAI 前高層克里斯蒂亞諾 (Paul Christiano) 4 月底時亦警

告，AI「接管世界」的機會有 10% 至 20%，到時將造成大量、甚至大部分人口死亡；若 AI 的智能達到人類水平的話，出現災難結局的機會更會高達一半。

本報訪問了全部 3 名參與 5 月聯署的香港學者，分別是浸會大學宗教及哲學系教授皮耶塔里寧 (Ahti-Veikko Pietarinen)、香港大學哲學系助理教授薩拉丁 (Nathaniel Sharadin) 和中文大學法律學院助理教授德魯津 (Bryan Druzin)，他們同對 AI 威脅人類感到擔憂。

AI 回應具引導性 影響用戶

皮耶塔里寧表示，AI 的運作存在很多不確定性，人們亦未知 AI 應用及其反應對全球的影響。他同意 AI 的才能很大程度上都是正面的，尤其在製造知識、學習及教育方面。AI 的最大危險是其產生的回應不是為了真相，而是為影響及說服用戶按某方式思考及行事，一如現今社交網絡出現的情況。他表示，雖然 5 月的聯署內容簡潔，但確實表達了有關全球風險的觀點。他認為，要避免出現大規模濫用可能引致的危機，需要人們重視自身的獨特性、思辨能力及機械不具備的道德和情感反應。此外 AI 亦應像孩子般接受訓練，向其灌輸人類的初始價值觀。

國際間須建強效管理制度

香港大學哲學系助理教授薩拉丁 (Nathaniel Sharadin) 亦表示，現階段難以預測 AI 系統的發展方向，原因在於這取決於人們如何管理 AI 發展。他表示，災難並非 AI 發展的唯一結果，相反它或能帶來經濟變革、大幅改善人類福祉與幸福感，而這些 AI 發展前景切實可期，但只會在全球共同制定出安全機制下才能實現。他強調，減輕 AI 風險的最佳做法，是共同制定國內以至國際間的強力而有效管理制度。

規管 AI 同時避免窒礙發展

應用 AI 將無可避免要面對規管問題，但亦衍生可能窒礙 AI 創新的顧慮。香港中文大學法律學院助理教授德魯津 (Bryan Druzin) 表示，鑑於 AI 的自我改善能力，人們或正創造很快就擁有比人類更高智能及知識的機器，而人們或無法完全控制它們，圍繞 AI 的規管迫在眉睫。對於在香港規管 AI，他指規管 AI 要面對不少困難，因 AI 影響眾多範疇。他認為，要在刑事或侵權方面建立新責任領域 (areas of liability)，以專門避免 AI 開發人員和用家濫用 AI，例如針對利用 AI 生成可能損害公眾利益的假資訊、可能損害個人或企業聲譽的誹謗內容而制定明確法規或判例法 (case law)；例如故意發布 AI 內容而不註明為 AI 生成的，會被定為類似發布偽鈔的罪行。

各地政府要平衡規管及鼓勵 AI 創新發展不易。德魯津稱因為各地政府都要鼓勵 AI 發展，讓其在環球經濟中保持競爭力，過度規管可能窒礙創新或限制 AI 的潛在好處。以香港為例，或會面對商業上不要去規管 AI 的壓力，以保持跟新加坡等地的競爭力。

(環球經緯)