ChatGPT帶來人工智慧新衝擊 資策會科法所：應關注其風險層級

由OpenAI開發的人工智慧聊天機器人程式ChatGPT，自2022年11月正式推出以來，引起全世界的使用熱潮，上線兩個月內使用者數目達到1億人以上，成為史上增長最快的消費應用程式。ChatGPT的主要功能除能與自然人使用者流暢互動，難以被辨認為電腦軟體的對話方式外，亦可勝任部分複雜的資訊處理工作，包括且不限於依使用者所提的條件自動產生文字，包含詩詞、介紹敘述甚至論文內容；依使用者要求自行編寫，及審視程式碼內容是否包含錯誤等等，已經大幅超出過去大眾對聊天機器人的功能認知。

然而，ChatGPT學習能力也可能讓其生成的內容產生隱私、道德與歧視的風險。OpenAI聯合創始人，現已離開董事會的馬斯克（Elon Musk）及OpenAI執行長奧特曼（Sam Altman）皆曾表示， ChatGPT的回應偶有出現偏見、歧視並具有道德風險。美國加利福尼亞州民主黨眾議員劉雲平（Ted Lieu）也認為ChatGPT應受到限制與監管，並使用ChatGPT「生成」了一份決議草案，呼籲國會應關注人工智慧並「確保人工智慧的開發和部署，應以安全、道德、尊重所有美國人的權利和隱私的方式進行」。

資策會科法所表示，對於人工智慧的監管法制，現行歐盟發展腳步相對較快，以歐盟「人工智慧法」（Artificial Intelligence Act）草案為例，該法係以「風險」為判斷基礎（risk-based approach），將人工智慧管制體系依對於歐盟基本價值、公共安全、人權等領域的風險、可能危害程度作成金字塔結構分級，主要有：不可接受風險（unacceptable risk）、高風險（high risk）、低度風險（low risk）與風險極小（minimal risk）四個級別。分級的意義在於區別各類AI應用所應受到的管制程度，以令各成員國在立法上不致因採取相同管制手段，而導致個別廠商受到過度限制

資策會科法所進一步解釋，歐盟人工智慧法草案所定義的「不可接受風險」層級，是指已顯然侵害歐盟基本權等重大權益的應用，原則上予以禁止，僅在極少數受嚴格限制的狀況下例外准許；「高風險」是指對歐盟基本權或相關法規所保護的公益，造成重大風險的應用，一般需符合歐盟法規的要求，在上市前提出申請，且經過第三方機構的評估驗證後方受核准上市；「低度風險」是指特定有欺瞞風險的AI應用；「風險極小」則是指目前該應用所產生的風險能受當前法規管制，因此原則上不課予額外的義務及規範。

一般而言，聊天機器人多歸類於「可能有欺瞞風險」的「低度風險」級別，雖然，部分專業人士目前也認為ChatGPT僅僅是個擁有多樣功能的便利工具，無須過度反應而進行特別規制。不過，ChatGPT不同於一般大眾所認知的聊天機器人，當其經過持續改良及大量機器學後，轉化應用在不同領域，加上未予以管控的狀況下，可能對於基本權等等重大權益保障產生威脅。ChatGPT能多樣轉化應用於不同領域潛力，也可能放大其威脅範圍，偏見、歧視對「平等權」的衝擊應屬首當其衝。在針對此類AI系統所生爭端，目前尚無法有效歸責的情況下，或許應審慎評估其風險級別並思考相對應的管理規範或手段，以符合科技發展實際情狀。

中央社

上稿時間： 2023年03月06日

新聞來源： https://www.cna.com.tw/postwrite/chi/336122

文章標籤