用ChatGPT找法院判決？從Roberto Mata v. Avianca, Inc.案淺析生成式AI之侷限

用ChatGPT找法院判決？從Roberto Mata v. Avianca, Inc.案淺析生成式AI之侷限

資訊工業策進會科技法律研究所

2023年09月08日

生成式AI是透過研究過去資料，以創造新內容和想法的AI技術，其應用領域包括文字、圖像及影音。以ChatGPT為例，OpenAI自2022年11月30日發布ChatGPT後，短短二個月內，全球月均用戶數即達到1億人，無疑成為民眾日常生活中最容易近用的AI科技。

惟，生成式AI大量使用後，其中的問題也逐漸浮現。例如，ChatGPT提供的回答僅是從所學習的資料中統整歸納，無法保證資料的正確性。Roberto Mata v. Avianca, Inc.案即是因律師利用ChatGPT撰寫訴狀，卻未重新審視其所提供判決之正確性，以致後續引發訴狀中所描述的判決不存在爭議。

壹、事件摘要

Roberto Mata v. Avianca, Inc.案[1]中，原告Roberto Mata於2019年8月搭乘哥倫比亞航空從薩爾瓦多飛往紐約，飛行過程中膝蓋遭空服員的推車撞傷，並於2022年2月向法院提起訴訟，要求哥倫比亞航空為空服員的疏失作出賠償；哥倫比亞航空則主張已超過《蒙特婁公約》（Montreal Convention）第35條所訂之航空器抵達日起兩年內向法院提出損害賠償之請求時效。

然而，法院審理過程中發現原告訴狀內引用之六個判決無法從判決系統中查詢，進而質疑判決之真實性。原告律師Steven A. Schwartz因而坦承訴狀中引用的六個判決是ChatGPT所提供，並宣稱針對ChatGPT所提供的判決，曾多次向ChatGPT確認該判決之正確性[2]。

貳、生成式AI應用之潛在風險

雖然運用生成式AI技術並結合自身專業知識執行特定任務，可能有助於提升效率，惟，從前述Roberto Mata v. Avianca, Inc.案亦可看出，依目前生成式AI技術之發展，仍可能產生資訊正確性疑慮。以下彙整生成式AI應用之8大潛在風險[3]：

一、能源使用及對環境危害

相較於傳統機器學習，生成式AI模型訓練將耗費更多運算資源與能源。根據波士頓大學電腦科學系Kate Saenko副教授表示，OpenAI的GPT-3模型擁有1,750億個參數，約會消耗1,287兆瓦/時的電力，並排放552噸二氧化碳。亦即，每當向生成式AI下一個指令，其所消耗的能源量相較於一般搜尋引擎將可能高出4至5倍[4]。

二、能力超出預期（Capability Overhang）

運算系統的黑盒子可能發展出超乎開發人員或使用者想像的隱藏功能，此發展將會對人類帶來新的助力還是成為危險的阻力，則會隨著使用者之間的相互作用而定。

三、輸出結果有偏見

生成式AI通常是利用公開資料進行訓練，若輸入資料在訓練時未受監督，而帶有真實世界既存的刻板印象（如語言、種族、性別、性取向、能力、文化等），據此建立之AI模型輸出結果可能帶有偏見。

四、智慧財產權疑慮

生成式AI進行模型訓練時，需仰賴大量網路資料或從其他大型資料庫蒐集訓練資料。然而，若原始資料來源不明確，可能引發取得資料未經同意或違反授權條款之疑慮，導致生成的內容存在侵權風險。

五、缺乏驗證事實功能

生成式AI時常提供看似正確卻與實際情形不符的回覆，若使用者誤信該答案即可能帶來風險。另外，生成式AI屬於持續動態發展的資訊生態系統，當產出結果有偏誤時，若沒有大規模的人為干預恐難以有效解決此問題。

六、數位犯罪增加與資安攻擊

過去由人工產製的釣魚郵件或網站可能受限於技術限制而容易被識破，然而，生成式AI能夠快速建立具高度說服力的各種擬真資料，降低詐騙的進入門檻。又，駭客亦有可能在不熟悉技術的情況下，利用AI進一步找出資安弱點或攻擊方法，增加防禦難度。

七、敏感資料外洩

使用雲端服務提供商所建立的生成式AI時，由於輸入的資料存儲於外部伺服器，若要追蹤或刪除有一定難度，若遭有心人士利用而導致濫用、攻擊或竄改，將可能產生資料外洩的風險。

八、影子AI（Shadow AI）

影子AI係指開發者未知或無法控制之AI使用情境。隨著AI模型複雜性增加，若開發人員與使用者未進行充分溝通，或使用者在未經充分指導下使用 AI 工具，將可能產生無法預期之風險。

參、事件評析

在Roberto Mata v. Avianca, Inc.案中，法院關注的焦點在於律師的行為，而非對AI技術使用的批判。法院認為，隨著技術的進步，利用可信賴的AI工具作為協助用途並無不當，惟，律師應踐行其專業素養，確保所提交文件之正確性[5]。

當AI科技發展逐漸朝向自主與獨立的方向前進，仍需注意生成式AI使用上之侷限。當個人在使用生成式AI時，需具備獨立思考判斷的能力，並驗證產出結果之正確性，不宜全盤接受生成式AI提供之回答。針對企業或具高度專業領域人士使用生成式AI時，除確認結果正確性外，更需注意資料保護及治理議題，例如建立AI工具合理使用情境及加強員工使用相關工具之教育訓練。在成本能負擔的情況下，可選擇透過企業內部的基礎設施訓練AI模型，或是在訓練模型前確保敏感資料已經加密或匿名。並應注意自身行業領域相關法規之更新或頒布，以適時調整資料使用之方式。

雖目前生成式AI仍有其使用之侷限，仍應抱持開放的態度，在技術使用與風險預防之間取得平衡，以能夠在技術發展的同時，更好地學習新興科技工具之使用。

[1] Mata v. Avianca, Inc., 1:22-cv-01461, (S.D.N.Y.).

[2] Benjamin Weiser, *Here's What Happens When Your Lawyer Uses ChatGPT*, The New York Times, May 27, 2023, https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html (last visited Aug. 4, 2023).

[3] Boston Consulting Group [BCG], *The CEO's Roadmap on Generative AI* (Mar. 2023), https://media-publications.bcg.com/BCG-Executive-Perspectives-CEOs-Roadmap-on-Generative-AI.pdf (last visited Aug. 29, 2023).

[4] Kate Saenko, Is generative AI bad for the environment? A computer scientist explains the carbon footprint of ChatGPT and its cousins, The Conversation (May 23, 2023.), https://theconversation.com/is-generative-ai-bad-for-the-environment-a-computer-scientist-explains-the-carbon-footprint-of-chatgpt-and-its-cousins-204096 (last visited Sep. 7, 2023).

[5] Robert Lufrano, *ChatGPT and the Limits of AI in Legal Research*, National Law Review, Volume XIII, Number 195 (Mar. 2023), https://www.natlawreview.com/article/chatgpt-and-limits-ai-legal-research (last visited Aug. 29, 2023).

許嘉芳
法律研究員 編譯整理

上稿時間：2023年09月

文章標籤

人工智慧

推薦文章