

人工智慧於公務運作中的利與弊（上）

駐休士頓辦事處教育組

人工智慧是一種透過計算機系統模擬人類推理思維的技術，比如能夠透過使用者提示生成獨特答案的 ChatGPT 即是生成式人工智慧的一種；此外，人工智慧也能夠分析大型數據，並利用數據自動執行過去許多須由人工進行的工作。然而隨著人工智慧的應用愈趨廣泛，人們也開始為其功能及使用倫理而感到擔憂。

2024 年初，美國司法部開始調查賓州一項有關於人工智慧使用不當的案例。該人工智慧模型旨在用以協助改善兒童福利，然而卻因系統錯誤判定而導致身障父母最終失去了對其子女的監護權。德州大學奧斯汀分校助理教授 Samantha Shorey 說：「若賦予人工智慧政策決策權，將有可能複製並加劇社會的不平等。人們應對引入人工智慧參與決策過程的行為更加謹慎，因為這有關乎誰能擁有獲取社會福利的資格、或者犯罪者假釋時間的判定。」

布朗大學技術責任中心主任 Suresh Venkatasubramanian 也說：「人工智慧將有可能因為其模型訓練的數據，對特定性別、宗教、種族產生歧視。此外，大量蒐集數據也可能引起隱私外洩問題，隨著人工智慧技術愈趨複雜，人們將更難理解這些系統是如何運作的，以及它們如何做出決策。」

華盛頓特區 AI 政策中心執行董事 Jason Green-Lowe 則認為：「如果技術加速發展卻又缺乏監督，未來我們將有可能活在一個由人工智慧控制、並以其利益而非人類利益為主的世界。」目前，AI 政策中心持續推動並遊說華盛頓特區為人工智慧安全設立更嚴格的標準。

不過，支持者卻對人工智慧在政府機構中逐漸佔據更多主導地位抱持樂觀態度，他們認為政府將因人工智慧帶來的自動化運作而得益，比如加速管理社會服務案件、針對冗長的政策分析做出及時摘要，或簡化雇員的招聘培訓流程等。德州大學奧斯汀分校電機工程和計算機科學教授 Alex Dimakis 表示：「人們不能躲在人工智慧的背後，人工智慧系統應是透明且須接受獨立評估的，即所謂的『紅隊演練』（red teaming，註 1）。在這個過程中，該技術的基礎數據和決策過程將被

多名專家審查，以確保是否需要更強大的安全措施。此外，當技術觸法時，州政府應強制執行現行法規，不應被人工智慧所迷惑。」

註 1：紅隊演練（red teaming）：指在軍事或網路安全演習中，由「紅隊」與「藍隊」分別扮演敵方角色和己方角色，扮演紅隊者須以敵方立場進行思考作戰，以驗證藍隊之偵測防守能力。

撰稿人/譯稿人：賴晏如

資料來源：<https://www.texastribune.org/2024/01/02/texas-government-artificial-intelligence/>

