

## 校園「AI 冤獄」案件漸多，教師反向偵測軟體誤判引爭議

駐紐約辦事處教育組

自 ChatGPT 於 2022 年問世後，美國各級學校便開始尋求如何因應學生利用 AI 協助完成作業的解決方式，但檢測軟體常有誤判的可能，錯誤的指控也可能對學生心理產生負面影響。現今也有愈來愈多的學生認為自己遭受不當指控使用 AI 作弊。針對此現象，專家分享幾項預防和應對 AI 作弊指控的方法。

普遍使用於偵測 AI 生成內容的系統 Turnitin 在 AI 生成內容低於 20% 時，發生誤判的機率就會變高，該公司表示，他們的系統不該被當作檢測學生作業的唯一依據。ChatGPT 問世後，許多教師使用如 Turnitin 的偵測軟體確認學術原創性，但專家說偵測軟體有將原創內容誤判為 AI 生成的風險。賓州大學 2024 年的一項研究指出，AI 偵測軟體很容易被拼字、標點符號和空格欺騙，建議不要使用於學術領域。史丹福大學 2023 年的一項研究指出，ChatGPT 的偵測系統對母語非英語人士有偏見，ChatGPT 的母公司 OpenAI 更因準確度太低，關閉了他們自己的 AI 偵測平台。波德科羅拉多大學研究科技道德和網路政策的副教授菲斯勒（Casey Fiesler）說，單靠有系統偏見的 AI 偵測軟體來判斷學術原創性是不負責任的，風險太高，面對有瑕疵的演算法時很難為自己辯護。

然而，學校的應對措施跟不上 AI 的快速發展也是一大問題，源於至今尚未有一個通用的標準化規範。非營利教育組織 EDUCAUSE 2024 年針對 AI 的相關調查中，將近一半的教育界高層、員工和教師認為其服務的學校對於學生使用 AI 沒有適當的準則。

在這樣的情況下，學生能以幾個方式保護自己，避免陷入「AI 冤獄」。首先，了解自己修習課程對 AI 工具的使用規範，若不幸被指控使用 AI 作弊，撰寫紀錄可以作為證明，紐約的教育顧問瓦格內羅瓦（Lucie Vágnerová）建議學生使用能儲存草稿的寫作軟體，提早開始撰寫、使用教授的課業時間及校內的寫作資源都能協助證明原創性，甚至可以將寫作過程中在網路上搜尋資料的紀錄畫面截圖作為證明。

為學生辯護的律師阿塞爾塔 (Richard Asselta) 建議學生在回應指控前和信任的長輩或同儕談談，並保持冷靜，因為他看過在相關案件中，被指控者犯的第一個錯誤就是回答前沒有仔細思考，結果說出讓他們被誤解的回應。

撰稿人/譯稿人：駐紐約辦事處教育組

資料來源：2025年2月7日 世界新聞網

[https://www.worldjournal.com/wj/story/121617/8531326?from=wj\\_good\\_index](https://www.worldjournal.com/wj/story/121617/8531326?from=wj_good_index)

