

研究支持自動寫作評分科技

駐舊金山臺北經濟文化辦事處文化組

最近一項研究指出，在寫作評量測驗中，自動寫作評分系統與人工評分方式具有相同效力，讓「自動閱讀器 (roboreaders)」在評估測驗和學校教育方面又再次引發關注。雖然技術開發者很開心能夠受到注意，但他們表示必須更重視政策改變的時機（如推廣至共同標準），這比寫作評分工具本身的戲劇性發展更為重要。

AIR 評估計劃主任 Cohen 先生表示：「個人認為，一段時間後，技術的整合將使該技術不僅適用於寫作評分中，也能應用於其他任務。但我們確實需要釐清目前的應用限制，才能有更進一步的發展。」

本研究由改善評估方法之推廣計畫所推動，並以評估專用的寫作測驗為依據。每種軟體將針對來自 6 個州的 2 萬 2,000 份寫作樣本進行評分，使用演算法衡量每篇文章的語言及結構特性，並根據先前以人工評分的文章，來預測真人評分者將如何評定特定文章。整體而言，由 9 個自動化寫作評分系統所產生的分數與真人評分的結果非常接近，且在某篇新聞稿中，共同研究主任 Tom Vander Ark 表示：「此項實驗決定性地證實，自動化寫作評分系統為快速、準確且符合成本效益的工具。」

Akron 大學教育學院院長暨研究共同作者 Mark Shermis 表示，該論文尚未提及自動化寫作評分系統最令人興奮的潛在能力，不僅能以更便宜的機器取代測試評分人員（或教師），更具有提供學生回饋和修改建議的能力。過去 10 年，大部分的軟體由本研究中的兩間廠商－教育測驗服務處 (Educational Testing Service) 和 Vantage Learning 公司－所提供，在寫作文法、風格、技巧、組織和思路發展方面，給予學生一些基本回饋。這類軟體可為教師提供支援，讓他們能夠完全專注於鼓舞人心或啟發想像力的創作。

研究專長為語言科技和人機互動的副教授 Carolyn P. Rose 表示，關於自然語言處理的測驗，或電腦如何與人類語言互動的科學，皆強調軟體能幫助學生對於任何領域的學科進行更富成效的合作討論。以使用人工智慧在線上平台帶動討論會的計畫為例，對象為 Pittsburgh 公立學校的中學生。一名電腦生成的人物在線上討論會中以參與者的身分詢問學生問題，有時甚至在討論過程中穿插幽默的言詞，緩解學生之間緊張的情緒。

Vantage Learning 公司的行銷副總裁 John Fallon 表示，利用目前的政策發

展動力—包括推行創建全新的進階寫作評估方法—將有助於所有領域的自然語言處理研究。Fallon 先生聲稱，該公司的產品不僅能為學生作文評分，還能提供學生改善建議。「我們所面臨的挑戰是如何幫助教師讓學生撰寫更多文章，並維持學生層面的互動。」

但 Sudbury 出版商 Will Fitzhugh 表示，他懷疑自動化寫作評分系統的任何應用是否真能提升學生的教育內容。他質疑，僅瞭解組織、句型結構和文法，是否能提供學生將來在職場所需的寫作能力。「我尚未深入研究過電腦所看的內容，但我不認為電腦能夠瞭解學生實際想表達的想法。」

資料來源：2012 年 4 月 25 日

Education Week

連結網址：<http://www.edweek.org/ew/articles/2012/04/25/29essays.h31.html>

