

人工智慧時代的十字路口*

蘇碩斌**

* 本文整理自作者在 2023 年 7 月 7 日「政治大學 AI 讀書會夏日論壇：AIGC 對大學教育的影響與衝擊」的發言稿。AI 讀書會係由政治大學傳播學院名譽教授鍾蔚文與政治大學創新與創造力研究中心合作於 2022 年 9 月開始舉辦，並於寒暑假擴大為論壇，邀請不同領域專家針對特定主題交流。

** 蘇碩斌為國立臺灣大學臺灣文學研究所教授、國科會人文與社會科學處處長，e-mail: sbsu@ntu.edu.tw。

本文引述格式：

蘇碩斌 (2024)。〈人工智慧時代的十字路口〉，《新聞學研究》，159，183-189。https://doi.org/10.30386/MCR.202404.0011

這個論壇的主題是 AI 與高等教育，我就用自己方式、包括剛剛做的筆記，以及之前的想法跟大家分享。我先自我介紹，本來是社會學背景，後來在文學所任教，現在借調到國科會的人文處工作。

從印刷術到 AI：媒介革命與知識的傳播

我的研究領域之一是印刷術。五百七十年前，印刷術其實是世界的超級新科技，或許人們以為它就是一個工具、就是幫我們把思想印出來而已。這是需要修正的，傳播學院非常熟悉的媒介思想家 McLuhan，已經告訴我們，印刷術改變了很多，包括宗教改革、民族主義，因為印刷術用大量的方言印製聖經，造成人們用同一種語言來讀聖經，方言於是變成國家語言，所以民族主義是因為印刷術這個媒介新革命而出現的。

AI，其實就是我們的媒介新革命，從工業革命算起是第一次技術大革命，到電腦出現的資訊革命，到現在，又遠比資訊革命更跨一步。雖然 AI 自從 Alan Turing 在 1950 年代提出至今已存在很久，但我想先來界定我們在這裡這麼驚恐，絕對都是因為 2022 年 11 月底發佈的 ChatGPT 所引起。我們要談的 AI，並不是弱 AI (Weak AI) 像是專用型 AI，而是強 AI (Strong AI)，或是所謂的通用型人工智慧 (Artificial General Intelligence)。

國科會支持的 AI 研究，其實過去幾十年都是由學者各自做研究，但 ChatGPT 一出來，國科會馬上緊急組成國家隊，也就是日前國科會發布新聞的「可信任人工智慧對話引擎」(Trustworthy AI Dialogue Engine，簡稱 TAIDE)。當然很多人問說，為什麼國家要做這個？怎麼可能拼得過 OpenAI 等大型企業？人家是用上萬片 GPU 運作、臺灣可能就只有兩千多片，怎麼拼？但其中原因應該是因為內容的問題。

通用型 AI 的挑戰與機遇

所以我要講的第一點，是通用型 AI 馬上會碰到一個難題，要如何訓練 AI？要不要讓通用型 AI 連上網路。我們知道 GPT 3.5 並不會連上網路更新資料，所以每次問到奇怪問題，GPT 3.5 就會說「對不起，我的資料只到 2021 年 9 月」，還要開外掛才能夠連網路。為什麼不讓 AI 上網，就是因為擔心資料污染，所以 TAIDE 強調可信任（Trustworthy）概念，也就是臺灣的資料是可信任的，不被污染的。所以資料污染這件事情，是人類跟 AI 相處非常重要的一條界線。這條界線什麼時候會被突破？接下來 Google 的 Bard 或者是 Microsoft 的 Bing，這二個 AI 就一定要上網，之後資料會混雜到什麼程度，我覺得是人類前所未有的挑戰。

如果還堅持要求通用型 AI 的資料一定是正確無誤，事實上不可能。這也導向一個微妙的預測，通用模型會再釋放、容許各個單位思考。人類會因為 AI 而活在真理或者說科學愈來愈穩固的時代，或是相反？我在思維上比較偏好後現代主義，滿樂見這種真理不穩固的世界，誇張一點就是一人一把號、各吹各的論調。我認為這未必會導向世界一片混亂，而是 AI 可能會培養出一群群各自堅信自己知識的社群，而未必是廣泛、普遍、被自然科學或啟蒙理性所引導為唯一真理的世界。

AI 與高等教育的未來

通用型 AI 因為思維更接近人類，甚至引起「取代人類」之憂，讓人類面臨很多不知答案的十字路口，以下提出幾個可能。

第一個十字路口是大學知識的問題。剛剛幾位講者都提到，尤其鍾蔚文老師講到從 *biversity* 到 *university*、以及真正的 *university* 會不會存在的問題？東京大學吉見俊哉教授月前曾到政大演講，他近年從東大副校長職位退下之後一直在省思大學前途問題，已有五本討論大學的書、質問大學應該是什麼？

當然大學會變，中世紀的大學最重要的學問是修辭學 *Rhetoric*，一個人即使讀非常多的書，知識的意義也都在最後剎那爆發出來的表現。有了印刷術以後，知識的表現就不在人的剎那表現，而變成看書、比對、寫字的緩慢累積。數位化、AI 到來之後呢？又推動查詢資料、累積資料變得無比容易，可以想見，當世界有極大量的資料，人類、大學造成什麼質變？知識，恐將不再以沉澱、思辯為主了。

其實我昨天問 *ChatGPT*，想知道 *ChatGPT* 對大學造成什麼正面、負面的影響。很厲害，真的回答不錯的意見。它說短期影響有好有壞，好的是可以協助自動化，但是壞處就是人類會習慣依賴。

這裡就看到很微妙的十字路口：AI 可能讓大學趨向單一思維、也可能造成多元思維。多元思維的理由是，每個學生都可以取得自己的個人助理、可以發展獨有的思考；但 *ChatGPT* 也告訴我們，AI 是大數據訓練出來的模型，生成的知識是過去資料的機率堆疊，雖然不會每個問題都得到一樣的答案，但是總體結果，答案其實差不多，最後其實是知識單一化。所以，大學因為 AI 而造成知識單一化或多元化，這是我們走到的一個十字路口。

第二個十字路口是社會治理的問題。我有一本很喜歡的書叫《透明社會》（*Transparent Society*），David Brin 所寫。一開始就講監視器社會，未來每個人可能都會隨身帶著監視器，不只汽車、摩托車，早晚個人也有行動監視器——監視器拍的資料最後誰可以看？兩種可能，一種

是所有人都可以相互看到監視資料，第二種是只有政府授權的官員可以看。

第一種是可怕的相互監視社會，第二種是接近我們現行體制，最極端的就是中共模式。Brin 說，我們無從預測未來社會一定走向哪一方向。我們也將無法知卜 AI 的運用會走向哪裡？如果每個人都可以透過 AI 去查知對方生平、點閱喜好、預測行為，其實我們幾乎已經是這樣被廣告攻擊、過著透明社會。這是我們喜歡的社會嗎？但如果說 AI 的一切不准公眾使用、只有政府或權力者可用，變成公眾是透明的、完全攤在陽光下任由政府監控？未來是何種意義的透明社會治理，這是第二個十字路口。

第三個十字路口是人類作者（authorship）的問題。聯合國的高等教育指引提到，雖然多數期刊還不同意 AI 可列名共同作者，但其實有期刊已經接受。國科會申請計畫時，可能只能要求自我揭露是否使用 AI，短期不能接受 AI 掛名共同作者，但其實已可預見未來的倫理問題將層出不窮。論文或計畫書有用到 AI 協助？是否超越作者的自主範圍？有沒有 AI 能夠反偵查、像 Turnitin 偵測出抄襲比例來偵測使用 AI 的比例？其實科技業者已開發出 Originality.ai 這樣的系統在判斷。然而不論如何，AI 是否視同論文作者的問題其實十分困難，因為這其實涉及到，到底 AI 是不是「人類」這個最可怕的問題。

人類、AI 與生命的本質

上面就是我聯想到的人類十字路口。鍾老師開了很多書單，其中一本我最近讀過，就是小說《Machines Like Me》，中譯《機器如我，人類如你》。書的故事時間發生在 1980 年，小說採用了一個反事實設定

——Alan Turing 沒死，甚至為英國開發出世界尖端的機器人。小說場景就是管家機器人開始商業販售之後。男主角是機器人 Adam，最特殊的特質不是聰明全能，而是他能進行性愛動作，甚至愛上一個人。略去小說情節，故事最令人驚訝的是，有些機器人經過學習，他們開始憂鬱、厭世。為什麼會？一個理由，永遠無法活得像人類那樣有意義！人的意義？其實就在問人類最重要的特質。或可聯想到 1940 年代的艾西莫夫（Isaac Asimov）的小說《正子人》，或改編成電影的《變人》，百年不死、愛上人類的男主角機器人，一路追求著接近人類，最後的人生期望是什麼？他想得到死亡的能力。只有死，機器人才能成為人類。

所以面對 AI，我大膽預測人類最珍貴的能力，就是人類的死亡。就像海德格（Martin Heidegger）所講，人類因為死亡，才有時間感，而 AI 沒有辦法有時間感。所以從《正子人》和《Machines Like Me》這些科幻小說，談到機器人和人類的差異，最後答案是死亡。如果死亡是人類最需要正視的資產、是人類遺忘的特異能力，那我們如何面對步步追求長壽不老的我們？或許應該告訴 AI 同時告訴人類自己，不要害怕死亡，應該要珍惜死去、也就是珍惜生命。但結果對人類的危害是變大或是變小？因為 AI 雖然不會永垂不朽、但卻更像人類了。

我們其實沒有辦法預測，人類將開發出什麼樣的 AI。你說，能不能讓它有情感？能不能讓它做判斷？如果設計的時候就要求 AI，只要預測癌症的百分比達到一定比率，就直接開刀治療，即使 AI 不知道為什麼，AI 也會照做。所以也能訓練 AI，只要和使用者應答二十句時發現使用者沒有弄懂，AI 就要生氣、就要罵人、就講髒話。當然這都真的可能發生。人類設計的 AI，有情緒，有暴怒，有愛，都可能設計出來。我相信，我們沒有辦法預測 AI 的發展，除了前面的三個十字路口，更麻煩的是科幻小說預告的死亡和愛情的難題。如果按照黃仁勳的預測——

兩年算力還會增加一倍，這個 AI 社會應該會把人類嚇死。

2022 年底出現的 ChatGPT 嚇到我們，就是因為那麼快，甚至問題還未問完，答案已經出來。這樣的發展速度，我們能想到的事情，AI 應該很快都做得到。從 1980 年代人類使用電腦科技至今幾十年，資訊、網路的發展，讓我們人生的節奏還勉強應付得來，但是生成式 AI 的出現，我覺得已經嚴重影響到、威脅到人類未來的生存狀態。真的茲事體大，我們今天還沒有辦法找到人類未來世界應該要走的路，能做的，大概只能哀悼人性的不再、稍微有點自信的，就只能堅信人類不會被取代、警告資訊科技開發者必須擁有人文素養、知道自己在開發什麼。如果資訊科技沒有人文素養，那麼人文社會科學將只能永遠落在科技後面哀悼。因為不可能、也不應該告訴 AI 開發者不准研究、不准進步。

所以我只能講說，我們應該多討論人類對於世界 to be or not to be 的兩難，設法緩一下、想一下，因為一個程式調整，可能就把人類帶到另外一個不同的境地。

參考文獻

- 國家科學及技術委員會（2023 年 6 月 14 日）。〈TAIDE 對話引擎初步成果發表，產官學齊聚見證〉【新聞稿】。https://www.nstc.gov.tw/folksonomy/detail/f094b57a-204e-4114-99d3-a9412ae42d7a?l=ch
- 葉李華譯（2000）。《正子人》。天下文化。（原書 Asimov, I. & Silverberg, R. [1995]. *The positronic man*. Bantam.）
- 趙丕慧譯（2023）。《機器如我人類如你》。啟明出版。（原書 McEwan, I. [2019]. *Machines like me*. Nan A. Talese.）
- Brin, D. (1999). *The transparent society: Will technology force us to choose between privacy and freedom?*. Addison-Wesley.

• 新聞學研究 • 第一五九期 2024 年 4 月