

機率的海洋， 與那個始終在說話的東西

成大交管系\黃奎臻



重讀馬克·科拿柏格（Mark Coeckelbergh）與大衛·岡克·歐克爾（David J. Gunkel）的《當AI取得話語權，人類還剩下什麼？》（Communicative AI）

從「消失的密室」到無所不在的對話機器，我們在跟誰說話？

在正式翻開《當AI取得話語權，人類還剩下什麼？》這本書之前，我對大型語言模型（LLM）的認知，多半停留在「這是一個極度方便的生產力工具」或「一個似乎無所不知的虛擬助理」。在115年當前社會環境中，似乎不難洞見它的存在。潤飾報告、生成大綱，甚至在思緒卡關時尋求靈感，成為我使用它的幾個時機點。然而，本書導讀文章中引用的一個文學隱喻，給了我一記極為沉重的當頭棒喝：導讀者林宛瑩以《哈利波特：消失的密室》裡金妮與瑞斗日記的互動為例，生動地描繪了當人類向一個「未知的實體」傾訴，並在不知不覺中產生

信任與依賴時，會面臨何等巨大的危險。或許我們太容易被流暢的文字表象所說服，卻忘了追問最根本的哲學問題：「在對話的另一端，究竟是誰或者什麼在說話？」

這本書由馬克·科拿柏格（Mark Coeckelbergh）與岡克·歐克爾（David J. Gunkel）所著，其核心價值跳脫我們對於一般技術型教學書籍，沒有把篇幅著重於教導我們如何寫出更好的提示詞（Prompt），或如何更有效率地使用AI。相反地，它提供了一套結合語言學、倫理學、傳播學與存在哲學的「後設視角」，迫使我們停下腳步，重新審視那些我們以為理所當然且可能不曾質疑過的概念。從第一章對技術底層邏輯（機率預測）的拆解，一路推進到第七章對「書寫未來」的終極叩問，這本書的內容如同一面被擦拭乾淨的鏡子，清楚照出機器的本質，與人類自身對「智慧」、「意識」、「真理」、「作者」乃至於「語言」本身未經檢視的傲慢與偏

見。帶著對自身使用 AI 習慣的反思與質疑，我走進了這本書的學術與哲學深淵，在機器的喧囂不間斷地當代，嘗試找回人類主體性的錨點。

猜字的機器，與預測車流的模型

大型語言模型的本質，不過是建立在Transformer架構與注意力網路（Attention Network）上的機率預測機器。它的核心任務極其單純：預測序列中下一個最有可能出現的字詞。它捨棄了早期符號推理法（Symbolic Approach）那種依賴人工定義語法規則的笨拙，轉而將詞彙化為高維度的「詞向量（Word Embeddings）」，從極大規模的語料庫中提取統計規律。它不具備對真實世界的理解，卻能輸出極度擬真的語言，重點是它還不會感到自我懷疑，充滿自信地給予使用者提出要求之回應。

讀到這裡，我想起了自己的本科系專業：過去進行交通工程相關的計劃，於微觀車流模擬軟體中，若要針對號誌控制策略進行設置，會參考車流量相關的歷史時序資料，並從中適當提取統計規律，對特定時刻的狀態嘗試做出推估，兩者的底層邏輯似曾相似，只不過預測的對象從「車輛行為」換成了「人類語言的符號序列」。書中用莎士比亞名句中「question」與「toaster」的機率差異，冷酷地揭示了真相：大型語言模型只不過是在進行極度高維度且精密的「猜字遊戲」。

這個跨領域的類比讓我深刻意識到，把LLM擬人化是多麼危險的直覺。我們不會認為交通預測模型「理解」了駕駛者的心情；那我們憑什麼認為，一個依賴統計機率的語言模型「理解」了我們的喜怒哀樂？這種技術上的解魅，是我建立批判性思考的第一道裂縫，也是全書論述層層遞進的起點。

胡扯的結構性必然，以及我曾是幫兇

閱讀第一章「大型語言模型101」後，讓我解除了對LLM的神祕感，後續的第二章與第六章合起來，我對整個當代資訊生態感到前所未有的焦慮。

LLM消耗了龐大的環境資源，用於訓練模型等同於美國三萬三千戶家庭的用電，ChatGPT訓練期間的碳排放約等於一千六百萬次飛機旅行。更嚴重的是，它在無意間複製並放大了人類語料庫中的偏見與刻板印象。在版權爭議上，作者引用了柯比費格茲的混成公式，指出AI的生成本質是「複製→轉化→結合」，這讓現有的著作權法律框架面臨「西部荒野（Wild West）」般的治理真空。

而「胡扯（bullshit）」的哲學定義，則是精準得令人發毛。哲學家法蘭克福指出：說謊者知道真相而刻意隱瞞，但胡扯者對真假「完全漠不關心」，它的唯一目標是產出聽起來合理、流暢的語法序列。書中更進一步精準區分了錯誤資訊、假訊息與「幻覺（Hallucination）」，幻覺不是謊言，因為LLM根本無所謂「說謊的意圖」，它只是在機率的指引下，填出了一個聽起來最合適的詞。

當我把這個定義套回自己的使用習慣，我看見了一個令我不安的畫面，那些因為ChatGPT語氣無比自信且通順，便直接貼進報告的段落，以及那些省略查證步驟、以「它說得很有道理」替代批判思考的時刻。我其實在無意間，一次又一次地成為散布「完美胡扯」的幫兇，我就是那個兇手。而這種胡扯機制的氾濫，正嚴重威脅著以「知識民主（epistocracy）」為基礎的現代社會，大衛·艾斯倫所擔憂的，是一個不在乎真假的系統一旦被大規模應用，將以前所未有的速度稀釋整個社會的信任基礎。面對仍在

「西部荒野」階段的法規現實，批判性地驗證資訊，已成為當代公民不可推卸的道德責任。

中文房間的震撼，以及他心深淵的對稱性焦慮

第三章「智慧、意識與他心問題」探討的哲學深水區，是我認為全書最燒腦、也最令人著迷的部分。

面對「機器是否有智慧」的提問，作者以「移動目標問題」指出AI智慧定義的無盡迴圈，意即每當機器達到一個智慧門檻，人類就立刻提高標準，將其排除在外，始終達不到我們設立的變動式標的。而舍爾的「中文房間（Chinese Room）」思想實驗，則完美擊破了圖靈測試「只要表現得像人就有智慧」的行為主義迷思。操作語法（syntax）的符號操作，永遠不等於語義（semantics）的真實理解，這句話讓我重新審視「理解」這兩個字本身的重量。

更讓我陷入認識論焦慮的，是書中關於「他心問題（Problem of Other Minds）」的論述，我們永遠無法直接感知他人的主觀體驗，我們相信他人有意識，純粹是基於「行為相似性」的類比推論。作者無情地指出，我們對AI意識的推論，與對他人意識的推論，在結構上是完全對稱的。既然我們無法證明AI有意識，我們其實也無法百分之百確定別人有意識，這正是哲學殭屍概念所揭示的困境。克勞德·夏農（Claude Shannon）警告我們「外觀並不是現實」，AI表現出情緒不代表擁有情緒。這種對稱性焦慮提醒我，面對AI時必須保持一種「疏離的清醒」，身為使用者的我們，應該拒絕在技術的流暢感與欺騙性設計中，交出盲目的信任與情感依附。

語言是存在之家，還是符號的遊戲場

第四章「語言、意義與溝通」的語言學轉向，徹底改變了我看待語言的方式。

在結構主義與海德格「語言是存在之家（the house of Being）」的觀點下，語言不再只是傳遞資訊的

工具，而是我們界定世界、棲居於世的根本框架。相較之下，LLM因為沒有肉身、沒有生活經驗、沒有在世存有的掙扎，它生成的只是「被石化的文字」，它充其量是一隻只會玩弄符號差異的「隨機鸚鵡（stochastic parrot）」，缺乏任何真正的在世存有經驗。

書中對Shannon and Weaver「傳播觀點（Transmission View）」與凱瑞「儀式觀點（Ritual View）」的區分，給了我極大的啟發。一切都可以從用途出發，若溝通只是為了傳遞訊息，那現階段的LLM確實已經蠻符合大多數需求了，但如果溝通是為了「維繫社群、分享共同信念、確認我們屬於同一個文化」，我想LLM就徹底失效了。

這讓我聯想到，當我用AI幫我撰寫一封給教授的道歉信，或者給朋友的生日祝福，在文中雖然精準地傳達了資訊，但我這個角色卻悄悄退出了人類社會的「意義共創儀式」。語言的形式被保留了，但語言的重量與誠意卻直接遭到掏空。那封信最終送達，但送達的究竟是什麼？

作者已死，書寫的私生子何去何從

將第五章「作者身分與權威」對作者身分的解構，與第七章「書寫有沒有未來」對書寫未來的探討結合起來閱讀，並嘗試從這部分的視角切入，激起我對於「創作」這件事產生前所未有的懷疑，或許這是現今面臨的一大課題。

「作者」並非自古皆然，而是印刷術與啟蒙運動之後被建構出的「家長主義（Paternalism）」產物，是一種理體中心的權威，現代版權法將作者固化為財產與責任的持有者。但LLM沒有意圖、不具備身世，它的文字缺乏指涉真實世界的能力，這讓傳統版權法作為「主體化」的基礎徹底動搖。由於它生成的內容缺乏一個可被究責與賦予所有權的真實人類主體，導致版權法賴以運作的核心前提面臨崩解與解構。

回溯到柏拉圖《費德羅篇》中蘇格拉底對書寫的質疑，這深刻提醒著我們，書寫從一開始就是脫離作者掌控的「私生子」。LLM只是將這種技術化推向了極致，它瓦解了西方傳統的「理體中心主義」（logocentrism），展示了語言可以在沒有作者意圖的情況下自行運作。這並不代表書寫沒有未來，而是意味著我們正在進入一個「後作者書寫（post-authorial writing）」的時代。在未來，寫作或許不再是從無到有的神聖創造，而是人類與演算法之間、意圖與機率之間的無盡協商與編輯，並經過多次修正而產出得結果。

寫下這份心得的此刻，我認真思索了一番，如果讓AI幫我生成段落，並且請它針對全書的所有內容，加以統整並模仿人的撰寫方式將反饋完整寫出，那這份心得的「作者」還是我嗎？我對於此問題無法給出肯定的答案，但也許，意識到這個問題本身，就已經是屬於人類的那一份主體性了。

機器的喧囂裡，守住語言的重量

《當AI取得話語權，人類還剩下什麼？》這本書沒有給我一個「擁抱AI」或「抵制AI」的懶人包結論。它透過跨學科的廣闊視野，做了一件更為困難且意義深遠的事，它拆解了語言、意識、真理與權威的底層邏輯，清楚洞見隱藏在流暢對話背後的哲學深淵與社會危機。

回到導讀中那個令人不安的問題：「是誰在說話，重要嗎？」在讀完全書之後，我認為極度重要。

語言可以最為符號的交換，但它也是意義的生成、責任的承擔，亦是人類社群賴以維繫的儀式。大型語言模型確實在技術上創造了一場奇蹟，它能夠在機率的汪洋中拼湊出幾乎無懈可擊的文法與邏輯，顛覆過往人們在這方面的想像。但正如舍爾的中文房間與海德格的存在論所揭示的，沒有真實生命經

驗的符號操作，永遠無法孕育出真正的語義理解。它技術所及範圍內能給出答案，但它無法針對答案「負責」，這很可怕，你可能從未思考過它能模擬同理心，但它不會為人類的痛苦流下一滴眼淚。

書寫有沒有未來？人類的主體性會不會被淹沒？這些問題的解答，不在於科技巨頭的演算法能迭代到多麼強大的版本，而在於我們這些身處歷史轉捩點的使用者，選擇以何種姿態面對這項技術。面對這個充斥著「胡扯」、「幻覺」與「版權爭議」的生成式對話AI時代，我們不應該僅停留在實用主義的層次，沾沾自喜於生產力的提升。我們必須培養一種深刻的「哲學病識感」，時刻警惕技術的欺騙性外觀，拒絕盲目的信任與情感投射。

我認為，唯有深刻理解這本書中所揭示的哲學與倫理挑戰，努力成為一個具備高度反思能力的「批判性使用者（critical user）」，我們才能在現有機器的喧囂與胡扯中，穩住人類語言的重量，捍衛屬於人類的存在之家，這也是人存在的價值，機器或許可以成為一個協作者，但人才是詮釋一切的主體。

