大型語言模型進入「平權時代」？從Gemini 2.5到DeepSeek V3的觀察

Alice and Bob｜Mike Wong

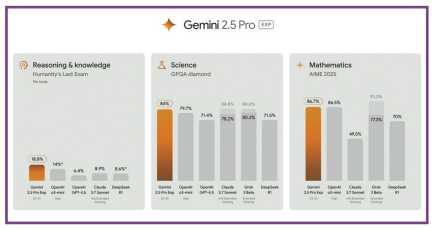
Mike Wong，香港搵樓平臺House730的首席科技官(CTO)。本科為電腦科學，後在香港中文大學修讀哲學完成自己的哲學夢，工商管理碩士預備生。熱愛科技，認為科技可以和人文主義一起帶領人類前進。文章內容多從人文學科的角度去看科技發展對我們的影響和反思；而作為管理人員，也會分享科技與管理的發展。

2024年初，Google正式發表最新的大型語言模型Gemini 2.5 Pro，聲稱為旗下最強AI，具備前所未有的理解與推理能力。與此同時，OpenAI、Anthropic、微軟及來自中國的DeepSeek亦先後發布新一代模型，令AI領域再度進入一波「百模爭鋒」的局面。

然而，今次與過往不同的是：這些旗艦級模型的實力，開始趨向一致。

性能收斂 差異縮小

從目前多項基準測試(benchmark)結果來看，Gemini 2.5在表格解讀、視覺推理(Vista Benchmark)上表現突出，接近人類水平，更首次實現長達1百萬tokens的上下文理解能力，遠超GPT-4 Turbo的128K tokens。然而，在常識推理(SimpleBench)上，Gemini 2.5的表現與Claude 3.7不相伯仲，各得5/10，顯示不同模型在不同領域有優勢，但整體差距正在收窄。



OpenAI的GPT-4 Turbo雖然未見重大升級，但已在價格、速度及可用性方面佔據領先優勢。Anthropic的Claude 3.7對語境的掌握較佳，尤其在對話連貫性方面表現穩定。這場競爭，已不再是「誰遙遙領先」，而是多方拉鋸。

算力驅動成主流 模型商品化加速

微軟CEO Satya Nadella最近提出一個值得深思的觀點：AI模型已進入商品化階段(commoditization)。他指出，當前各家模型背後使用的架構與訓練方法已趨相近，真正的差異愈來愈來自於資源——特別是算力與資料量。

過往模型之間的突破多建基於架構創新，例如Transformer、RLHF、Mixture-of-Experts等技術。但進入2024年後，效能的提升更多來自規模的擴展與調整。例如Gemini 2.5雖未公開具體參數，但從各項能力判斷，其應為極大型、具多模態整合的模型。

這也說明了：未來AI模型的競爭核心，很可能不再是「技術秘密」，而是誰擁有更多資源、更強運營能力。

中國模型的突圍與全球競局

值得關注的是，來自中國的DeepSeek V3表現令人驚艷，尤其在數學與編程能力上，逼近甚至超越GPT-4 Turbo。這意味著，在compute與資料基礎條件接近的情況下，西方實驗室並不再具壓倒性優勢。隨着開源社群活躍、中國本地模型訓練架構成熟，全球AI發展格局正快速多極化。

小結：進入AI的「平權時代」？

若說過去是GPT-4的一枝獨秀，現在則是「多模鼎立」。每個模型皆有長處，但表現愈見一致。對用戶而言，重點已不在於誰是最強模型，而在於如何部署、整合及落地應用。AI模型的門檻降低，意味企業可更靈活選擇工具，建立專屬的知識引擎與流程自動化。

當AI模型愈來愈像，人類的選擇，不再是信仰，而是實用。

文章同時發表於House730.com

作者電郵：mikewong@house730.com

am 今期流行

「塑造品格．成就專才—生命教育X職業專才計劃」正式啟動

明愛馬鞍山中學於2022-2023年度規劃「醫護賢才發展藍圖」，並於2023-2024年度創建於香港中學界罕有的「健康護理實習專室」，在本年度(2024-2025年度)正式推展醫護相關的職專才教育項目及應用學習課程，讓學生及早規劃未來發展，建立健康的生活方式，貢獻社會。「健康護理實習專室」設有護士站、病房、浴室、工作室及狹窄居住環境五個模擬區域，加強同學在技能上的訓練，為香港培訓醫護賢才。由香港教育大學主辦，明愛馬鞍山中學協辦，將於3月29日舉辦「塑造品格．成就專才—生命教育X職業專才計劃」啟動禮，將帶大家參觀校園及由退役巴士改裝的「有價巴士」生命體驗館。

理大學樂齡與家庭研究中心「賽馬會齡活共創計劃」啟動禮

「賽馬會齡活共創計劃」啟動禮暨最新研究結果分享會已於3月24日順利舉行。當日透過一系列研究成果發佈及深入討論，研究成果顯示顯示，實施長者及年齡友善措施可提升50+人士消費意欲達146%。其中員工友善態度、產品服務改良、營銷信息優化以及建築環境改善，是提升長者消費意欲的四大關鍵措施。但目前僅25%長者滿意現有商戶措施，反映市場仍有龐大改善空間。為進一步推動長者及年齡友善文化，「賽馬會齡活城市『全城．長者友善』計劃24/25」現正接受報名，截止日期為4月11日。網站：https://jccitypartnership.hk/

報料電郵：amjetso@am730.com.hk

面對「漏缺」的自己

萬紫千虹｜曹虹(盲言)

在我們的生活中，總有一些人在工作和人際關係中遊刃有餘，每一件事都能做到圓滿無缺。他們的能力和魅力使他們成為眾人仰望的對象，甚至讓人不禁想要模仿。這種「無漏缺」的形象，究竟是現實還是幻影？值得注意的是，所謂的「無漏缺」其實是一種社會化的產物。在許多文化中，成功和完美被視為理想的標準。這導致許多人在社交場合中努力塑造一個無懈可擊的形象，讓他們看起來總是能夠應對自如。然而，這種形象往往是表象，背後卻可能隱藏著許多未被揭示的困惑和挫折。

在職場中，有些人擁有出色的專業技能和良好的人際關係，這使得他們在團隊中備受推崇。但事實上，這些人也許在背後經歷了無數次的挑戰和失敗，甚至在個人生活中面對著不為人知的困擾。這種情況下，「無漏缺」的形象只是他們應對困難的一種策略，而非真實的自我。漏缺本身並不是一件壞事。每個人都有自己的強項和弱點，這是人性的一部分。承認自己的不足，反而能讓我們更真實地面對自己和他人。漏缺所帶來的反思和學習，往往能成為我們成長的催化劑。在這個過程中，我們不僅能夠提升自己的能力，還能更好地理解他人的困難，增強同理心。

當我們把「無漏缺」視為成功的唯一標準時，許多人會因此感到焦慮和挫敗。每當面對挑戰時，他們可能會選擇隱藏自己的脆弱，而不是勇敢地展現真實的自我。這樣的心態不僅會影響個人的心理健康，還可能在團隊中造成隔閡，降低合作的效果。

因此，我們需要重新思考「漏缺」的意義。與其追求一個無懈可擊的形象，不如學會接受自己的不完美。這樣，我們能夠在失敗中吸取教訓，從中成長，並在生活的各個方面都能保持真誠和透明。在這種情況下，漏缺不再是弱點，而是一種力量。

只有當我們接受並擁抱自己的不完美，才能在不斷變化的世界中找到真正的自我，建立起更為真實和深厚的人際關係。正是在這些漏缺中，我們看到了做人的智慧和生活的真諦。

香港恒生大學傳播學院院長