

性能出色 成本差價逾百倍 迅速搶佔美國市場 美企使用中國 AI 暴增 佔比達 46%

美國企業面臨人工智能（AI）使用成本持續攀升的壓力，而中國 AI 模型憑藉其性能與價格的雙重優勢，迅速搶佔美國市場。與美國頭部 AI 模型相比，中國模型的性能差距已大幅縮小，使用成本卻僅為同類產品的約 1%，促使愈來愈多美企轉向使用中國 AI 模型。

●香港文匯報記者 路雨萱

美國財經新聞媒體 CNBC 報道，據可自由切換約 400 種 AI 模型的 OpenRouter 數據顯示，截至 7 月初，美企在中國 AI 模型上使用的詞元（token）佔比最高達 46%，自 2 月以來每周均超過 30%。相比之下，過去 12 個月平均佔比僅為 11%，去年上半年更低至 4.5%。中國 AI 使用量 6 月進一步激增，按詞元計算，6 月底一周的總調用量達 25 萬億，較 5 月底翻倍，高出美國 AI 使用量 78%。

成本差異是這一轉變的核心因素。以每輸出 100 萬 token 的收費標準計算，Anthropic 的「Claude Opus 4.7」收費 25 美元（約 195 港元），而中國最熱門模型「DeepSeek V4 Flash」僅收費 0.18 美元（約 1.4 港元），成本不足前者的 1%。OpenRouter 分析師指出，中國開源模型比美國 Anthropic 與 OpenAI 模型便宜 60% 至 90%。

美初創公司：將節省數百萬美元

美國 AI 初創公司 Lindy 在 6 月已將其全部流量從 Anthropic 更換為 DeepSeek，該公司行政總裁表示，此舉將在數月內為公司節省數百萬美元，「成本曲線急劇下降，斷崖式下跌。」美國大型加密貨幣交易所 Coinbase 也改用中國 GLM 和 Kimi 模型，行政總裁稱成本因此減少一半。

美國布魯金斯學會研究員凱爾指出，「隨着 AI 成本飆升，中國 AI 模型現在對美國公司尤其具吸引力。現在美國公司對成本更關注。」應用部署平台 Vercel 智能體主管阿羅拉直言「價格才是關鍵」，並指出若執行任務不需最好模型時，團隊傾向選擇價格最低但足夠好的模型，「中國的模型正在贏得這種競爭。」

「中國模型性能接近美國頂尖」

具備成本優勢的同時，中國 AI 模型在性能上亦極具競爭力。CNBC 報道，包括 DeepSeek 以及智譜在內的中國近期發布的模型，被廣泛認為與美國頭部前沿系統相比極具競爭力。凱爾進一步表示，中國模型性能「接近美國頂尖前沿模型」，並預計中國模型僅落後美國頂尖競爭對手「6 至 9 個月」。儘管美國最先進模型性能不斷提高，但其 token 消耗量與價格持續上漲，使用成本隨之增加。

智譜 6 月中旬發布的 GLM-5.2，在約 9,000 道測試題中位列約 500 個模型第五位，超過 Google 等模型，被業界比作去年 1 月震驚全球的 DeepSeek。Vercel 平台數據顯示，GLM-5.2 發布後首個完整周內，日詞元量增長約 27 倍，客戶數量增長約 80 倍，成為該平台追蹤的所有模型中，採用速度最快的模型。

●中國 AI 模型被美企使用的詞元佔比急升，成功搶佔部分美國市場。



●中國最熱門模型「DeepSeek V4 Flash」僅收費 0.18 美元，較 OpenAI 模型便宜九成。資料圖片

多家西方企業讚中國 AI 靈活性高

英國《金融時報》報道，多間西方大型企業紛紛換上中國人工智能（AI）模型。美國最大外賣巨頭 DoorDash、全球旅行房屋租賃平台 Airbnb 以及德國科技巨擘西門子等企業，近期陸續公開表態，讚揚中國 AI 模型不僅幫助企業降低成本，同時在效率與靈活性均表現突出。

外賣巨頭 DoorDash 混合使用不同 AI

DoorDash 聯合創始人安迪·方於社媒 X 發文，稱該公司現在將「較低級別工作」交由中國 AI 企業月之暗面（Moonshot AI）的 Kimi K2.6 模型處理，僅將 Anthropic 的 Fable 保留用於「最困難的工作」。他稱新的組合比之前僅使用美國前沿模型的配置「成本更低，效果卻大幅提升」。

西門子表示希望 AI 模型具有靈活性，其使用的模型範圍廣泛，包括中國的 DeepSeek 和 Z.ai 工具。Airbnb 表示使用「有限數量的中國來源模型」，並通過經批准的美國服務提供商運行以保護數據。

AI 平台 Featherless AI 首席行政總裁尤金·謝表示，中國模型是「房間裏的大象」，企業正意識到「不需要最好的模型，只需要更快、更便宜的模型」。美國喬治城大學研究員布雷斯尼克指出，既然中國模型基本能滿足需求，企業沒有理由為美國 AI 支付溢價。

美國創投集團 Andreessen Horowitz 共同創辦人安德森在社媒 X 上表示，「許多 AI 內部人士均稱，GLM-5.2 是首個在性能上可與美國頂尖 AI 模型匹敵的中國模型，且毫不遜色。」

美初創公司首發大模型 參考中國技術

美國科企 OpenAI 前技術總監穆拉蒂創辦的人工智能（AI）初創公司 Thinking Machines Lab，近日發布首款通用大模型「Inkling」，因使用中國 AI 技術而引發外界關注。英國《金融時報》報道，該模型架構參考中國 AI 模型 DeepSeek-V3，並在後訓練階段使用中國 AI 企業月之暗面（Moonshot AI）旗下的大模型 Kimi K2.5 生成的數據進行優化。

Inkling 的發布正值中國 AI 企業快速追趕美國同行之際。據 AI 數據平台 OpenRouter 統計，今年以來中國 AI 模型在整體使用量方面已超過美

國競爭對手，這一變化引發美國 AI 行業關注。從性能來看，Inkling 目前仍未躋身行業第一梯隊。Thinking Machines Lab 公布的基準測試數據顯示，該模型落後於 OpenAI、Anthropic 的領先產品，同時也低於部分中國 AI 模型。與 ChatGPT、Claude 等主流開源模型不同，Inkling 選擇了「開放權重」路線，用戶可將模型部署在自己的服務器上，並根據具體應用場景進行微調。

開放模型打造差異化優勢

Thinking Machines Lab 試圖通過開放模型和

「人機協作」打造差異化優勢。該公司認為，多數 AI 工具「由少數機構訓練完成後定型」，未來的 AI 應更開放，讓用戶參與塑造模型能力。

美國技術主導角色漸退

分析指出，Inkling 的發布折射出全球 AI 競爭格局正發生深刻變化。美國企業在資金、算力和頂尖人才方面仍保持優勢，但中國 AI 公司憑藉低成本和高效率的模型研發能力，正成為全球 AI 產業不可忽視的技術來源。過去由美國主導的技術輸出方向，正出現更多元的趨勢。

AI 競賽風向變 首看成本與效率

香港文匯報訊 人工智能（AI）產業競爭進入新階段。隨着 AI 技術逐漸走向企業實際應用，市場關注焦點已從模型規模與效能比拚，轉向「誰能以更低成本、更高效率完成特定工作」。這一轉向也令開源模型、混合式 AI 架構以及模型調度能力，成為下一階段競爭核心。

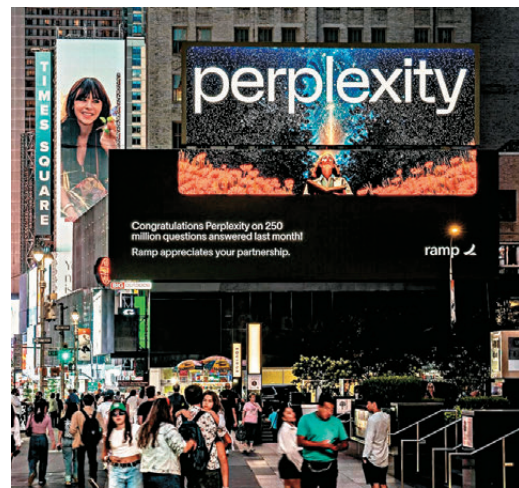
不同模型應付不同需要

AI 頭部企業 Perplexity 行政總裁斯里尼瓦斯接受美國財經新聞媒體 CNBC 訪問時表示，如今 AI 模型本身已不再是最終產品，真正關鍵的是能整合不同模型、工具與企業數據的協調系統。他主張企業應「針對不同任務使用最適合的模型」，而非一律採用最先進的大型語言模型。客服場景無需最昂貴模型，複雜編程可能需要高階模型，而日常例行工作則可交由成本較低的開源模型處理。

該趨勢反映企業更重視 AI 投資回報率。創投公司 Benchmark 合夥人芬頓稱，未來 18 至 24 個月內，甚至最快今年底，逾 90% 的詞元（token）將可能由開源模型產生。他指出，目前大型 AI 模型供應商依靠推論（Inference）服務賺取高額利潤，但隨着開源模型持續進化將面臨利潤壓縮。當企業可自行部署模型、滿足自身業務需求時，便不需支付大型 AI 公司額外收取的溢價。

雲端應對高難度任務

AI 發展方向的變化亦將影響大型數據中心建設。斯里尼瓦斯認為，未來部分 AI 工作可能在企業或消費者的終端設備本地執行，僅複雜任務才交由雲端處理。這種混合式 AI 架構，即本地裝置處理日常工作、雲端應對高難度任務，不僅有助於降低企業成本，亦可能改變 AI 算力及數據中心投資的整體走向。



●AI 頭部企業 Perplexity 主張「針對不同任務使用最適合的模型」，而非一律採用最先進的大型語言模型。



▲Alphabet（Google 母公司）等五大科企大舉興建數據中心，導致過去 5 年的債務規模擴大了一倍。

▶亞馬遜截至 3 月 31 日當季的自由現金流轉為負值。



香港文匯報訊 美國科技巨頭在人工智能（AI）熱潮下，大舉興建數據中心，導致在過去 5 年的債務規模擴大了一倍。為支撐一場前所未有的資本支出狂潮，這些公司紛紛舉債融資，並認為這是推動經濟轉型所必需的。

AI 投資狂潮下 美科技巨頭債務規模翻倍

5 巨頭新增債務 2.7 萬億

數據顯示，美國數據中心投資規模最大的 5 家公司 Alphabet（Google 母公司）、亞馬遜、Meta、微軟與甲骨文，過去 5 年合計新增債務約 3,500 億美元（約 2.7 萬億港元）。這些公司押注前沿 AI 服務最終將帶來可觀的新收入，投資者先前一直熱捧這些企業，積極認購其發行的各種債券。不過據知情人士透露，亞馬遜近期的 250 億美元（約 1,959 億港元）債券發行，卻遭遇不同尋常的冷遇，顯示資本市場承接科技巨頭持續融資以支持 AI 投資的能力並非無限。

不過對於其中大多數獲利能力極強的公司而言，舉債成本目前仍相對有限。上述 5 家公司

去年的利息支出總計超過 100 億美元（約 783 億港元），較 2019 年增加逾一倍，但與其中一家公司的自由現金流相比仍顯得微不足道。以 Google 為例，截至 3 月底，其經營活動現金流扣除資本開支後的自由現金流達 640 億美元（約 5,016 億港元）。

亞馬遜自由現金流轉為負值

然而並非所有公司的財務狀況都同樣穩健。亞馬遜截至 3 月 31 日當季的自由現金流轉為負值；甲骨文的現金消耗預計將進一步加快，其 2025 年債務規模約為全年銷售額的 2.5 倍。標普上周將甲骨文的信貸評級下調至投資等級中的最低一級，提到該公司持續擴大的 AI 投資支出。