

# 低成本AI模型帶來重新評估GenAI策略的時刻

## 低成本模型推動AI巨輪（2）

新一代生成式 AI 模型的震撼登場，可能會加劇市場競爭、推動創新、降低成本，並使生成式 AI 計畫更加經濟實惠。或許是時候讓 CIO 重新評估他們的 AI 策略了。

編譯／酷魯

中國 AI 新創公司推出號稱效能優於 OpenAI o1 GPT 的全新開放原始碼推理模型 DeepSeek-R1，為開年不久的2025年擲下一枚驚動全球 AI 界的震撼彈。伴隨而來的，還有從紐約到東京等市場的科技股拋售潮，專門生產資料中心 AI GPU 的領導廠商 Nvidia，在 1/27 當天市值蒸發了近 6,000 億美元。

這家中國新創之所以迅速竄紅，主要歸功於 R1 模型的高品質與成本效益。該公司聲稱，其 AI 模型的表現可媲美 OpenAI 的 o1 推理模型，甚至在某些情況下超越對手，同時使用更少的 GPU，使得整體成本大幅降低。

DeepSeek 的火熱引來正反兩面的聲浪，一方面該新創的 AI 聊天機器人App 可以免費存取 R1，並已登上多國應用商店排行榜的首位。一方面，開始質疑其 AI 模型的開發過程，並質疑會有隱私、安全與倫理方面的疑慮。

Gartner 副總裁暨分析師 Chirag Dekate 認為，市場對於 DeepSeek 的訓練技術細節及其創新來源的反應有些過度。「這種情況加深了我們與某個未明對象之間在認知上的對立，甚至助長了沙文主義或民族主義的論述，」他表示。這樣的論述的確吸引眾人的目光並造成熱烈的討論，畢竟聳動的新聞標題比起技術報告更吸引人們的興趣和想像力。

### 顛覆以算力為依歸的傳統模式，AI 訓練不再是超大規模業者的禁嚮

雖然該模型在業界引發正反兩面的激烈論戰，但不意味著該模型的創新可以被忽視。Gartner 在一份研究報告中指出，DeepSeek 挑戰了主流生成式 AI（Gen AI）的成本架構與方法論，突顯了當前市場領先供應商的定價模式效率低下，進而導致大規模部署高價值使用案例的投資報酬率出現負值的狀況。

由此可見，新 AI 模型的創新似乎顯著降低了對先進硬體和計算資源的依賴，同時還能以更低成本提供與其他市場主流大型語言模型（LLM）相當的效能。

Gartner 在 1/29 發布的一份研究報告認為，DeepSeek-R1 代表了一場關鍵轉變，表明生成式 AI 的未來應該由創新且具成本效益的方法來驅動，而不是仰賴純粹的算力進行擴充的傳統模式。

IDC 軟體市場研究與諮詢業務全球 AI、自動化、資料與分析研究部門集團副總裁兼總經理 Ritu Jyoti 表示，DeepSeek 的橫空出世可能會帶來更易獲取且更經濟的 AI 解決方案，但仍需就策略、競爭、品質和安全等因素進行審慎考量。

IDC 高效能運算與全球基礎設施研究副總裁 Peter Rutten 表示，從 DeepSeek 成果中得到的最

重大收獲是，當前 AI 的訓練方法是不合理的，這方法是基於 AI 只能透過更大、更快、更強的架構來提升的理論。

「透過創新 AI 開發演算法、框架及軟體，我們可以在只需使用現有基礎設施成本的一小部分且相同準確性的情況下，實現與最新版 ChatGPT 相當，甚至更優異的成果表現，」他進一步表示。「這意味著，AI 訓練不再是那些能夠投入數十億美元建置大型基礎設施之超大規模業者（hyperscaler）的禁臠。」

### 掌握各國中國新型 AI 模型禁令狀況，納入企業 AI 開發策略

除此之外，隨著 DeepSeek 的崛起，各國基於倫理、安全性、隱私性、透明度及公平性等考量，開始紛紛祭出禁令。所以企業在擬定今後 AI 開發或部署策略時，也必須將這點納入考量，否則會有違反禁令的不必要風險。

2025 年 1 月 31 日，美國德州州長 Greg Abbott 成為首位禁止 DeepSeek 的美國官員，理由是該模型會讓美國基礎設施面臨來自境外安全威脅的可能。緊接其後的是美國紐約州州長 Kathy Hochul，他因為中國新創的言論審查，以及使用者資訊會有被中國政府監控的可能風險，而禁止在政府裝置上安裝該模型。

美國眾議院行政主管部門管理長（CAO）已發出通知，禁止所有國會職員不得在官方手機、平板或電腦上安裝該模型。此外，美國國防部、海軍、太空總署（NASA）也陸續發布禁令。

除了美國之外，義大利是最早禁用該模型的國家之一，隨後，包括台灣、南韓（東亞首國）、澳洲、加拿大、印度（該國連 ChatGPT 也禁用）、丹麥 3/5 區域，以及多家日本企業也陸續實施相關禁令。

### 立刻展開行動！避免錯失良機、落於人後風險

由 DeepSeek 所開發的方法顯示，從成本與資源需求的角度來看，今後大型 AI 開發不論是在成

本或資源需求方面，對企業而言已變得更容易實現。這也讓包括模型客製化/微調，以及在模型上進行推理的中小型 AI 計畫變得更加負擔得起。隨著 DeepSeek 這類技術的發展，AI 最終將變得更加經濟實惠，甚至隨著時間推移，會像其他工作負載一樣負擔得起。

創新工場（Sinovation Ventures）董事長兼執行長李開復曾做過預測，中國 AI 新創公司將專注於提升 AI 的運行效率。Dekate 也認同地表示，DeepSeek 展示了一個全新趨勢：生成式 AI 的成本結構將變得更有效率、更具可行性。

Dekate 認為，DeepSeek 的推出息再次提醒企業，AI 創新正在以驚人的速度發展，對於尚未開始投入生成式 AI 的 CIO 們來說，現在正是全面擁抱的時候，否則將面臨被市場淘汰的風險。

即使市場可能對 DeepSeek 的任何正面擁抱或負面疑慮的態度都有反應過度之嫌，但 Dekate 仍認為這股風潮帶了一個極其明確的訊號，那就是 CIO 不能再等下去了，否則恐有坐以待斃的風險。

他認為，CIO 現在有兩條路可選，一條是立刻行動，展開實驗，制定並執行生成式 AI 策略，以全面推動部署。另一條是，CIO 也可基於安全或倫理疑慮而選擇靜觀其變，而不對既有 AI 政策做任何改變，企業恐將因此錯失良機，最終面臨落於人後的地步。