

AI工具開支驚人 企業收巨額賬單 硅谷巨頭濫用人工智能 成本貴過真人

以硅谷巨頭為代表的全球企業正競相引入人工智能（AI）工具提升生產力，甚至以AI更節省開支為由裁員。然而，隨著企業大量消耗代表AI算力的Token（詞元），巨額的賬單也讓不少企業大喊吃不消，他們發現，AI在工作流程中消耗的算力成本往往遠超人力成本。分析認為，企業一邊大肆裁員、一邊濫用AI，不僅不能節省成本，反而導致開支飆升，有部分企業已經開始對內部AI使用「踩煞車」，以重新計算如何平衡AI成本與人力開支。

大公報記者 陳明

全球企業紛紛引入AI工具，AI智能體（AI代理）的使用成為潮流。與此同時，Meta、微軟、甲骨文等大型企業近期都以提升效率、加大AI投入為由，宣布大幅裁員。職場追蹤網站Layoffs.fyi數據顯示，今年第一季度有超過9萬人遭解僱。然而，不少科技行業高管認為，現階段AI未必能夠幫助企業省下人力開支，因為算力成本仍然遠遠超過人力成本。

Uber 4個月燒光全年AI預算

AI算力成本最直觀的衡量單位便是Token。在AI技術語境中，Token是語言模型處理文字的最小計算單位，Token使用量愈大、費用愈高。隨著算力爭奪戰日趨白熱化，以硅谷巨頭為代表的科企們幾乎都在強調Token的使用量。

谷歌CEO皮查伊19日透露，谷歌目前每月處理的Token數量已超過320萬億，是一年前的7倍，顯示了企業正在瘋狂地使用AI。儘管如此，這些企業卻沒獲得想像中的成本大幅縮減。

網約車平台Uber首席技術官內帕利此前透露，公司2026年的AI編碼工具預算，在短短4個月內便已耗盡。Uber所用的正是Anthropic的編碼工具Claude Code。Uber每月使用AI工具的工程師比例已達95%，其中約11%的後端更新完全由AI自動生成。

隨著AI普及應用，伴隨而來是成本的急劇增加。內帕利表示：「我現在只能重新從頭規劃，原來的預算早就嚴重超支了。」研究機構Gartner指出，即便Token單價預計在2030年前下降近9成，企業AI總成本仍可能持續攀升，因AI智能體執行每項任務所需的Token量，遠超傳統模型，使用量攀升速度或無法抵銷降價帶來的紅利。

「不要為了用AI而用AI！」

大規模消耗Token更是讓不少科技主管與工程師開始質疑：燒掉更多AI成本，真能帶來更多產出嗎？Uber首席運營官麥克唐納日前直言，尚未觀察到Token用量與生產力提升之間存在明確關聯。

亞馬遜近日關閉了一個內部AI使用量排行榜，該排行榜原本是為了鼓勵員工使用AI，但結果適得其反：員工刻意誇大Token消耗量，導致亞馬遜額外支出巨額算力成本。亞馬遜高級副總裁特雷韋爾為此告誡員工：「不要為了用AI而用AI！」

分析認為，AI取代人力的論調可能沒有想像中那麼誇張。英偉達負責應用型深度學習研究的副總裁卡坦扎羅近日在受訪時表示：「就我所在的團隊來說，算力成本其實遠高於人力成本。」麻省理工學院此前的一項研究表明，在以視覺作業為核心的崗位中，僅有23%適合通過AI自動化實現降本增效，其餘77%的崗位，繼續聘用人類員工反倒成本更低。

耶魯預算實驗室稱，儘管目前尚無確鑿證據證明AI能提升生產效率，也沒有充分數據支撐「AI取代就業崗位」的說法，但科技巨頭仍在大舉加碼AI布局。摩根士丹利數據顯示，今年科企在AI領域的資本開支已達7400億美元，較2025年大增69%。瑞士人工智能研究院旗下戈登商學院AI與金融專業教授基思·李表示，明明僱用人工更省錢，企業卻一邊大把砸錢投入AI，一邊大規模裁員，暴露出AI投資在經濟上明顯存在不合邏輯之處。他認為，AI未來的經濟效益如何，還取決於這項技術能否真正兌現自身價值。其必須證明自身的可靠性，減少幻覺、降低對人工審核的依賴，並順暢融入企業現有體系架構。

Anthropic 估值達 7.6 萬億 超越 OpenAI

美國AI初創公司Anthropic於28日表示，在最新一輪融資中籌得650億美元，最新估值達到9650億美元（約7.6萬億港元），首次超過長期競爭對手OpenAI。

據悉，最新一輪融資由紅杉資本、高度計資本、Dragoneer及Greenoaks領投。谷歌亦投入數十億美元，這是此前宣布的對Anthropic未來投資總額高達400億美元承諾的一部分；亞馬遜的投資額為50億美元。另外，韓國三星電子及SK海力士亦有份參與，但未披露具體金額。

Anthropic公司稱，自2月完成G輪融資後，全球企業客戶的採用率持續增長，本月稍早公司的年化營收運行率已突破470億美元。

評論指出，Anthropic並未盲目追求聊天機器人的用戶規模，而是將目光鎖定在企業客戶身上，尤其編程自動化領域。AI智能體在今天的爆火大大推動了Anthropic的發展，以其旗下AI大模型Claude為核心的一系列編程工具產品在行業引發巨大反響。作為Anthropic的核心產品，編程工具Claude Code自今年年初以來，年化收入翻了一倍以上，在AI編碼的份額達到54%。

Anthropic的增長速度比預期快了8倍多，這也使其陷入了算力瓶頸。一些用戶反映，算力短缺導致了服務中斷或性能下降。

Anthropic積極尋求私募資金之際，也正為首次公開募股（IPO）做準備。與此同時，OpenAI也被曝很快將提交上市申請，這使得兩家公司的競爭更加白熱化。

狂燒Token VS 大肆裁員

Uber 4個月燒光全年AI預算

● 網約車平台Uber今年的AI編碼工具預算，在短短4個月內便已耗盡。平均每位工程師每月在AI工具上的花費介於150美元至250美元，重度使用者可達2000美元。

Meta 下架競賽榜單

● Meta曾於3月在內部上線了一個名為「Claude經濟學」的榜單，員工消耗的Token越多，排名就越高。榜單上線一個月後，第一名的員工將Token消耗刷到200萬美元。Meta現已下架這一榜單。

亞馬遜關閉使用量排行榜

● 亞馬遜近日關閉了一個內部AI使用量排行榜，分數完全基於其AI活動量。員工為衝排名、刻意誇大Token消耗量，導致公司算力成本激增。

OpenClaw 月燒千萬

● OpenClaw創始人施泰因貝格爾此前公布其賬單，顯示其30天內的平台調用開銷高達130萬美元，總計消耗了近6030億個Token，日均消耗規模堪比中小型AI初創公司的月費。



今年科企宣布裁員人數

AI盤點工具錯漏百出 星巴克北美棄用

適得其反

在使用一款AI自動盤點庫存工具僅9個月後，星巴克北美日前宣布棄用，因為該工具「經常出錯」。

星巴克總裁尼科爾表示，公司去年9月在北美引入自動盤點牛奶、糖漿等飲品庫存的AI工具。該工具的設計初衷是讓門店能夠實時掌握庫存狀況、加快補貨速度，同時讓員工更專注於為顧客服務。預想中，員工只需將平板電腦舉到貨架前，對準糖漿、牛奶和其他飲品，該工具就會利用激光雷達和攝像頭掃描並盤點庫存。

然而，實際操作中問題不斷。路透社2月曾報道，該AI工具經常出現計算和標籤錯誤的情況，例如分不清燕麥奶、牛奶等相似的奶類，或點漏部分款式的牛奶。星巴克上傳的視頻顯示，該工具在計算相鄰樽裝貨品時，未能辨識出貨架上的一樽薄荷糖漿。

星巴克本月18日發出公司內部通訊，宣布即日起停止使用該AI工具，會沿用點算其他商品庫存的方式，來點算飲品庫存。星巴克21日發聲明證實消息，稱公司決定停用該AI工具，旨在統一所有庫存盤點方式，將繼續專注於規模化運營的一致性和執行力。星巴克還表示正加強門市補貨頻密度，並持續改善供應鏈，希望做到只要餐牌上列出的商品，顧客都能購買。



什麼是Token？

1 企業員工大量使用AI工具

輸入信息轉化為AI運算詞條越長，消耗越多



2 AI進行運算，推理時消耗Token



Token的消耗類似電量

3 AI大模型供應方以Token進行定價、計費



用量越大，模型越強，費用越高

有多貴？

Token的使用費用是由單價乘以消耗量決定的，意味着Token消耗得越多，其費用就越高。以Anthropic的Claude Opus 4.6模型為例，其輸出價格高達每百萬Token約25至37.5美元。

為何變貴？

今年，幾乎所有的模型廠商和雲廠商的Token集體漲價。AI智能體的火爆推動了Token需求端的增長，其消耗的Token是傳統對話模式（例如聊天機器人）的數十倍甚至數百倍；供給端方面，內存和GPU又供給不足，帶動價格上漲。



「Token消耗量至上」美科企員工畸形內捲

本末倒置

今年，美國硅谷掀起一場關於「Token極大化（Tokenmaxing）」的激烈爭論。愈來愈多的企業要求員工大量使用AI工具、提高Token消耗量，希望藉此提升效率，甚至把Token用量視為工作積極度和成果的單一衡量標準，引發外界質疑。

科技媒體The Information上月披露，Meta一名員工自建名為「Claude經濟學」的榜單，追蹤同事的Token消耗量，優勝者可獲得「Token Legend」等榮譽稱號，導致全公司30天內的Token總消耗量從6.02萬億，急升11倍至73.7萬億，出現大規模濫用Token的亂象。

排名最高的用戶在30天內平均消耗2810億至3285億個Token，費用接近200萬美元（約1567萬港元）。報道刊發後，該排行榜下線，員工猜測是因為「畸形競爭下消耗的Token成本已遠超Meta預期」。

報道更援引多名Meta員工稱，部分工程師為提升排名採取不當手段，有人指示AI智能體生成大量、細碎的代碼變更，這些變更對功能毫無改善，卻能大幅拉高Token消耗的統計量。

值得注意的是，不少科企效仿Meta，推出類似排行榜。迪士尼在內部推出名為「AI應用儀表盤」的系統，追蹤員工使用AI工具的情況。《華爾街日報》指出，當大多數員工為了績效而用AI時，反而製造更多「AI垃圾」。



▲ Anthropic行政總裁阿莫戴在Claude開發者大會上講話。

競爭激烈