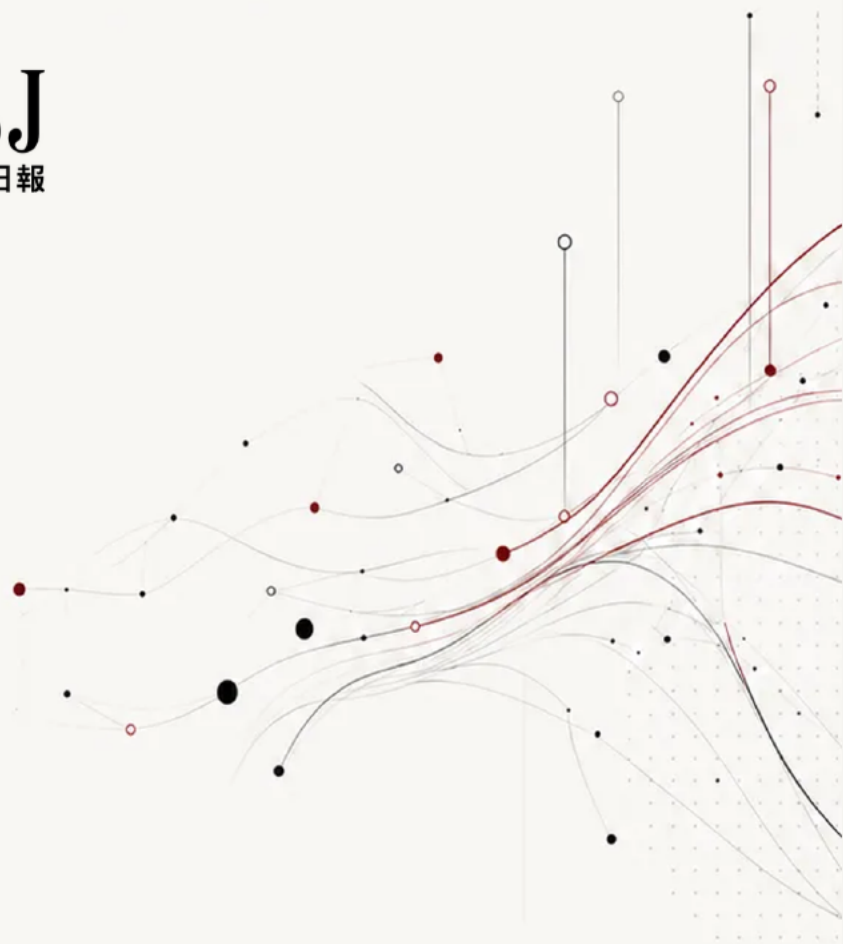




全球商業脈絡・精選主題包

AI 開始大規模吸金， 錢與算力正流向哪裡？

資金與算力的重新分配，正在改寫企業與市場的判斷

編輯導讀 · EDITOR'S NOTE

五篇報導，讀懂 AI 資本潮的完整路徑

這五篇並不是五個獨立話題。第一篇從債券市場看 AI 建設如何吞噬資本；第二、三篇追蹤 Google 與軟銀如何爭取晶片、資料中心和電力；第四篇把焦點拉回企業內部，當 AI Agent 開始長時間運作，token 成本與治理變成新的管理問題；第五篇則回到市場，檢視上半年漲幅與獲利預期如何重新分配。

閱讀這份精選主題包時，可以先問一個問題：資金、算力與回報之間，哪一環開始出現壓力？

編輯提問

錢先流向哪裡，回報又會在哪裡被驗證？

本期三條決策線索

資金

AI 基礎設施需要龐大資本，債券市場是否仍願意承接，將決定擴張速度。

算力

競爭從晶片延伸到資料中心、電力與融資能力，稀缺資源正在重新定價。

回報

企業成本治理與股票市場表現，開始要求 AI 投資提出可以衡量的結果。

閱讀地圖 · CONTEXT MAP

五篇如何互相牽動

從資金進場、晶片與基建競爭，到企業成本與市場回報，這是一條由上游投入一路傳導到投資結果的鏈條。



目錄 · CONTENTS

五篇精選，一次看懂全局

01

全球牛市大考驗！六巨頭髮債規模爆增125%，2500億AI大借款，測試市場承受極限！

資本豪賭 · CAPITAL

02

Google = 下一個輝達？黃仁勳獨創「資金永動機」模式被抄走，壟斷地位已終結？

晶片變局 · SILICON

03

日本太慢，跟不上AI腳步！軟銀孫正義警告：日本沒有人有遠大夢想…但我有！

亞洲競局 · ASIA

04

企業CEO終極挑戰：未來四年，AI算力消耗將暴增24倍！

算力經濟 · COMPUTE
ECONOMY

05

盤點2026年上半年標普500指數表現最好的20隻股票

市場領漲 · MARKETS

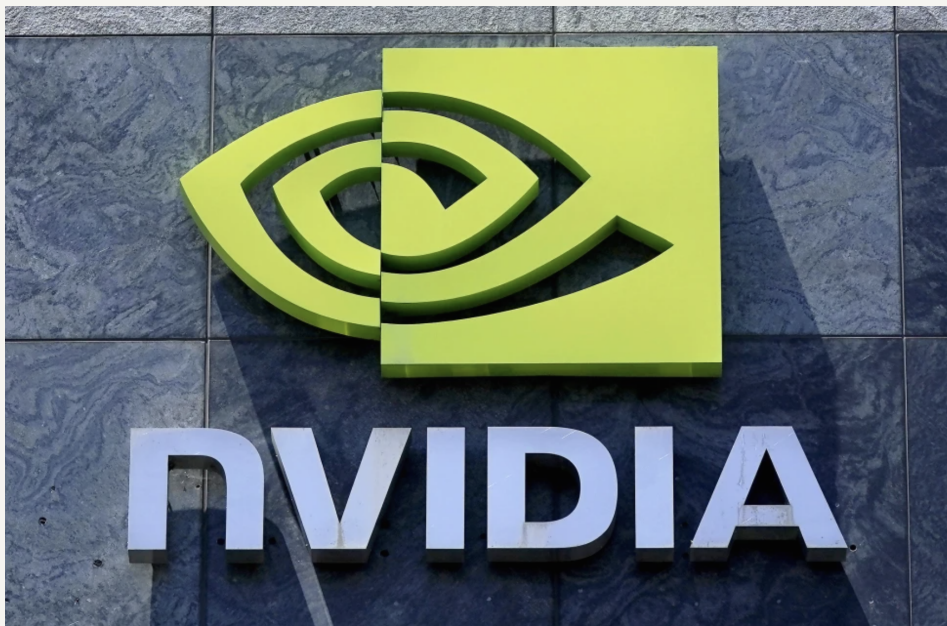
01

資本豪賭 · CAPITAL

全球牛市大考驗！六巨頭髮債規模爆增125%，2500億AI大借款，測試市場承受極限！

超大規模雲端服務提供商的支出和發債規模超出了投資者的預期，這給債券價格帶來壓力。

WSJ 報導 · 風傳媒編輯整理



人工智慧（AI）晶片巨擘輝達（NVIDIA）。圖片來源：美聯社

導讀

為支應大規模資本支出，全美 AI 雲端服務商今年開始大舉舉債。根據金融科技公司 Dealogic 資料，六家 AI 巨頭——Alphabet、亞馬遜、Meta、甲骨文（Oracle）、輝達和 SpaceX，今年已在全球發行約 2,440 億美元債券，高於去年的 1,080 億美元和 2024 年的 170 億美元，規模超出投資人預期。

超大型雲端服務商債券近期走弱，對基金經理人和投資人來說是一件大事，因為投資級公司債通常相當穩定。這意味著，即使只是溫和的價格下跌，也可能對基金相對於同業的績效造成重大影響。這不僅對債券價格帶來壓力，對股市多頭也是一場嚴肅考驗

編輯提問

債券市場還願意承接多少 AI 資本支出？

華爾街正向那些為投資人工智慧(AI)基礎設施而掀起歷史性發債熱潮的科技公司傳達一個資訊：行行好，請放慢腳步吧。

過去幾周，投資級公司債券市場難以消化輝達(Nvidia)、SpaceX和亞馬遜(Amazon.com)合計750億美元的債券發行。這標誌著與今年早些時候相比發生了轉變，當時投資者通常樂於通過任何可能的方式將資金交給所謂的AI超大規模雲端服務提供商。

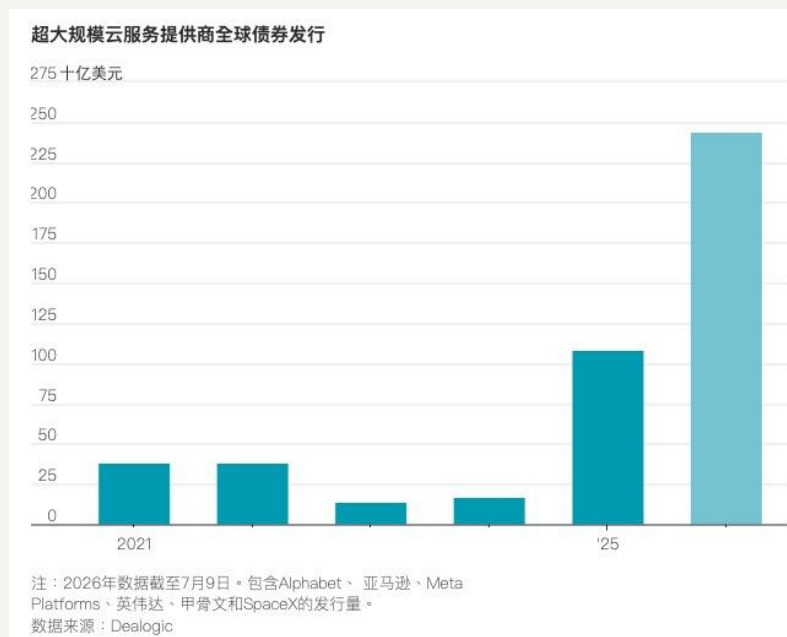
儘管輝達和SpaceX能夠以相當低的利率發債，但它們新發行的債券在二級市場迅速下跌，令那些通常喜歡買入此類債券用於轉售的投資者感到失望。與此同時，亞馬遜不得不支付以其標準來看非常高的利率來完成債券發售，這反映出投資者新產生的謹慎態度。

在很大程度上，這些投資者表示他們並不特別擔心發債人的信用狀況，也不擔心AI基礎設施建設的可持續性。問題實際上是，他們完全預計這些公司將在未來數年繼續在晶片和數據中心上投入巨資，從而導致數千億美元的新債券湧入市場。

「每個人都知道未來還會有更多債券發行，因此我認為大家在全力投入時會有所猶豫，」Voya Investment Management投資級公司債主管特拉維斯·金(Travis King)說，「每個人都想為下一筆交易留出一些空間。」

超大規模雲端服務提供商債券近期的疲軟對基金經理和投資者來說是一件大事，因為投資級公司債券往往相當穩定。這意味著即使是溫和的價格下跌，也會對基金相對於同行的業績產生重大影響。

而且，目前尚不清楚超大規模雲端服務提供商是否會在短期內改變行為。



通常情況下，公司會盡量讓投資者滿意，以確保自身發債成本儘可能低。但超大規模雲端服務提供商對算力的爭奪如此激烈，以至於它們似乎準備隨時發行數百億美元的債券，而不顧市場狀況以及是否可能需要支付更高的利率。

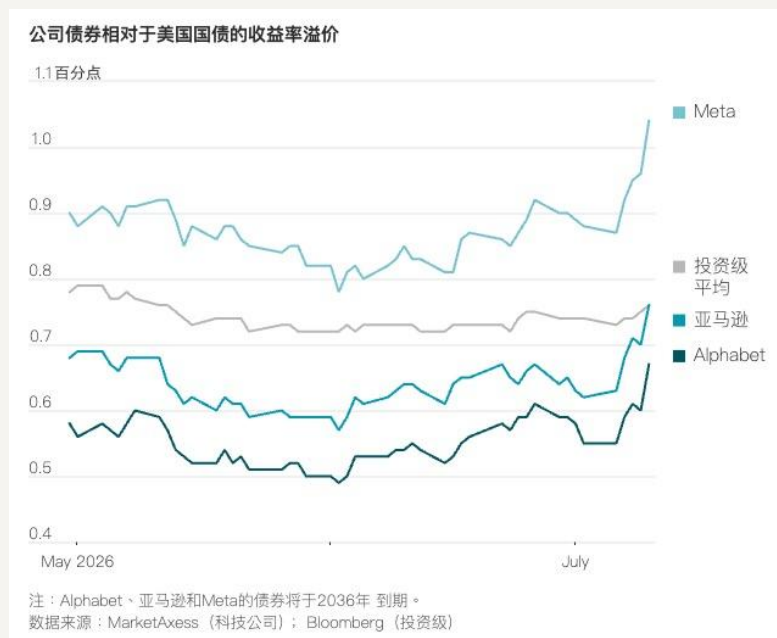
「對我們來說，這是我們討論的全部內容，而對他們來說，就像是‘哦，對了，你知道，我們進入了債券市場’，」Newfleet Asset Management投資級公司債板塊聯席主管瑞安·榮克(Ryan Jungk)說，「他們不一定在乎是否會讓我們的市場泛濫。」

榮克和其他投資者表示，這種情況最終可能會改變。隨著發債成本上升，亞馬遜、Alphabet和Meta Platforms等科技巨頭應該能夠繼續支出。但它們可能會被迫發行更多股票，而不是債券。

以Alphabet為例，該公司已在6月宣布今年將發行超過800億美元的股票，以幫助為其AI投資提供資金。理論上，這一舉措本應對其債券有利。但在現實中，提振作用有限，因為投資者將此次發行視為該公司可能在AI上投入更多資金的信號，從而需要同樣規模的發債。

Dealogic的數據顯示，被債券投資者視為超大規模雲端服務提供商的六家公司——Alphabet、亞馬遜、Meta、甲骨文(Oracle)、輝達和SpaceX——今年已在全球發行了約2,440億美元的債券，高於去年的1,080億美元和2024年的170億美元。

投資者在進入今年時就知道，超大規模雲端服務提供商將發行大量債券。投資者為持有這些公司現有債券而要求的高於美國國債的額外收益率（即收益率差）已經相應上升，而且在今年頭幾個月，對新發售債券的需求相對強勁。



然而事實證明，這些公司的支出和發債規模甚至超出了投資者的預期。許多投資者對輝達在6月份發售250億美元債券感到措手不及，上周亞馬遜發行同等規模的債券再次讓他們感到意外，這推動所有超大規模雲端服務提供商債券收益率差飆升。

MarketAxess的數據顯示，上周Alphabet 10年期債券的收益率差上升了0.12個百分點，而Meta 10年期債券的收益率差攀升了0.16個百分點。彭博數據顯示，投資級債券的平均收益率差僅微升了0.02個百分點。

作為首次發債的公司，SpaceX面臨著自身的問題，投資者不確定其債券應如何定價。自6月23日發行以來，該公司10年期債券的收益率差已躍升近0.5個百分點。

駿利亨德森(Janus Henderson)多板塊信用債全球主管約翰·勞埃德(John Lloyd)表示，他的團隊長期以來一直認為，今年各公司在AI基礎設施上的支出將超過平均預期，這導致他們持有的超大規模雲端服務提供商債券少於基準指數的權重。

他說，展望未來，關於這些公司將在AI上投資多少，「仍然存在各種可能的結果」，最高估計顯示未來幾年投資額將超過10兆美元。

科技公司債券價格的變動現在對投資組合經理來說尤為重要，因為這些債券在基準債券指數中所佔的份額越來越大。如果投資者被進一步發債的威脅嚇倒，並「低配」科技債券，那麼如果發債量低於預期且債券反彈，他們可能會蒙受損失。

反之，如果他們開始大舉買入這些債券，而發債勢頭並未減弱，他們也可能會受到打擊。

「如果你在科技類交易上判斷失誤，那可能會決定你這一年的成敗，」Newfleet的榮克說。

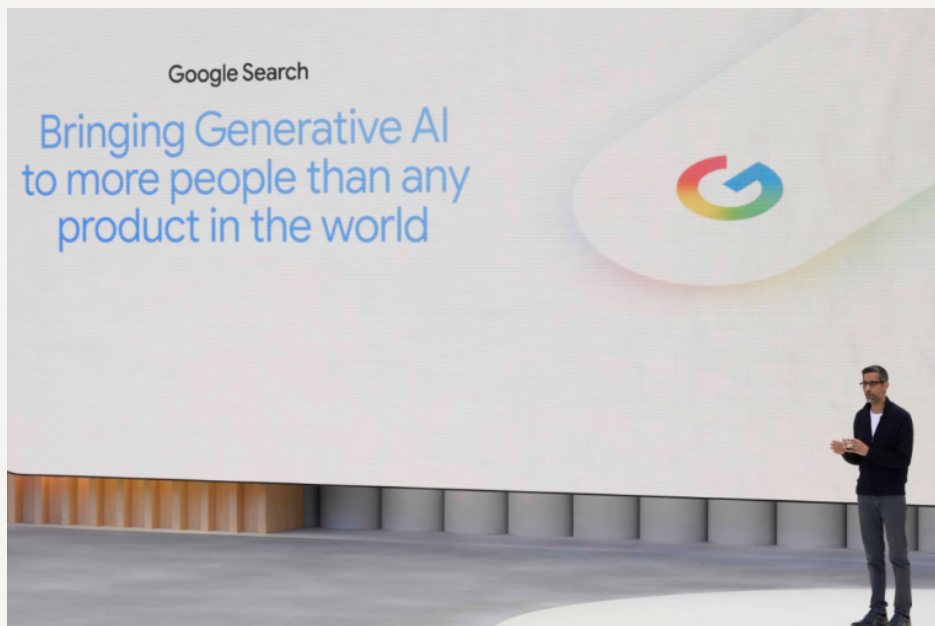
02

晶片變局 · SILICON

Google = 下一個輝達？黃仁勳獨創「資金永動機」模式被抄走，壟斷地位已終結？

憑藉雄厚的財力與多年的技術積澱，Google有望在這場競爭中脫穎而出。

WSJ 報導 · 風傳媒編輯整理



Google母公司Alphabet首席執行官桑達·皮采。圖片來源：美聯社

導讀

Google正大舉進軍人工智慧(AI)晶片市場，但必須跨越一大障礙：輝達(Nvidia)。

雖然如此，Google在為其AI晶片TPU，開拓市場時，卻效仿了對手輝達的做法：利用財務擔保幫助資料中心以更低成本舉債，並提供循環融資，最終資金會以晶片採購的形式，迴流到自己手中。快看這資金永動機是如何運作的？

編輯提問

Google 能否複製輝達的資金與晶片生態？

Google(Google)正大舉進軍人工智慧(AI)晶片市場，希望在一被視為21世紀最重要的賽道中搶佔更大份額。憑藉雄厚的財力與多年的技術積澱，該公司有望在這場競爭中脫穎而出。

但Google必須跨越一大障礙：科技公司普遍不敢得罪輝達(Nvidia)極具「地盤意識」的首席執行官黃仁勳(Jensen Huang)。

在距離尼亞加拉瀑布不遠的安大略湖南岸，Google一直在向外界展示，它是如何借用輝達那套打法來贏得客戶的。

紐約州西部的這個地點坐落著名為Lake Mariner的AI數據中心集群。據知情人士透露，Alphabet旗下的Google為該項目提供了32億美元的財務擔保，該項目開發商將把數千個微處理器的算力出租給AI巨頭Anthropic。

這正是輝達屢次使用的策略，旨在為其自家AI晶片本已火爆的需求推波助瀾。

直到最近，輝達還幾乎一直獨霸該市場，其圖形處理器(GPU)因具備訓練和運行AI模型的強大能力而受到科技公司的青睞。但隨著過去一年AI領域的角逐演變為一場算力資源爭奪戰，挑戰者開始紛紛入局，其中實力最強勁的莫過於Google。



輝達OVX平台設計的水冷架構數據機房。圖片來源：翻攝自輝達官網

「所有這些資金充裕的公司都堅信，圍繞算力的市場將蘊含巨大價值，」AI基礎設施公司TeraWulf的聯合創始人兼首席技術官納扎爾·汗(Nazar Khan)說，「他們想入局，不想被甩在後頭。」TeraWulf目前正與Google支持的雲端服務提供商FluidStack共同開發Lake Mariner項目。

Google的晶片被稱為張量處理單元(TPU)。在為TPU招攬客戶時，Google效仿了輝達的做法：利用財務擔保幫助數據中心以更低成本舉債，並提供所謂的循環融資。在這種模式下，Google投資的部分資金最終會以晶片採購的形式迴流到自己手中。

知情人士透露，Google雲部門領導層的改組加劇了這種緊迫感。輝達與OpenAI合作密切，且是後者的主要投資者；Google則與Anthropic建立了類似的關係，同時還擁有自家的前沿模型Gemini。

無論是在私下還是在公開場合，黃仁勳一直是淡化Google與輝達展開實質性競爭的能力。

今年4月，黃仁勳在德瓦爾凱什·帕特爾(Dwarkesh Patel)的播客節目中表示，相比Google及其他訂製晶片(ASIC)製造商，輝達保持著巨大的領先優勢。他還認為，Anthropic是Google TPU唯一的重要外部客戶。

「我們的市場觸達範圍遠遠超過任何TPU或ASIC所能企及的水平，」黃仁勳說，「我很想聽聽他們如何證明TPU的成本優勢。在我看來，這根本說不通。」

Google最近向輝達發起了迄今為止最直接的挑戰：該公司與黑石集團(Blackstone)達成了一項50億美元的協議，計劃成立一家新的雲端服務公司。新公司將直接與CoreWeave和Nebius展開競爭，而這兩家由輝達支持的雲端服務提供商只使用輝達的整套硬件方案。

「與幾年前相比，他們在將現有資源變現方面顯然更加善於把握時機，也更加激進，」Bernstein的科技分析師斯泰西·拉斯貢(Stacy Rasgon)說。「但幾年前，這種機會根本不存在。如今，我們聽到的全是算力短缺的聲音。」

布局晶片業務

Google很早就預見到了算力緊缺的到來，這始於其一位頂尖科學家口中的「思想實驗」。

2013年，傑夫·迪恩(Jeff Dean)與其他AI研究人員共同致力於語音識別研究，當時他們使用的正是支撐當今大語言模型的神經網絡技術。

「我當時說，‘好吧，如果我們想把這個語音模型推廣給1億用戶，並且他們每天使用幾分鐘，那就需要將Google擁有的電腦數量增加一倍，’」現任Google DeepMind AI實驗室首席科學家的迪恩在接受採訪時回憶道。他得出的結論是：「我們需要構建專門的硬件。」



Google高層與國際媒體對談，Google執行長皮查伊（右）與Google DeepMind技術長武庫格魯。圖片來源：中央社記者張欣瑜舊金山攝115年5月21日

起初，Google將這種硬件留作自用，利用這些晶片為其搜尋引擎和其他產品開發AI模型與功能。

隨著晶片需求呈爆炸式增長，Google開始通過其雲平台向其他公司開放這些晶片。此舉推動了雲部門的迅猛增長。

「這是輝達主導地位的終結嗎？」頗具影響力的科技研究公司SemiAnalysis在去年11月的一篇文章中問道。該文章的背景是Google發布了第七代TPU，而Anthropic已經在使用這款晶片訓練其模型。

直接面向客戶

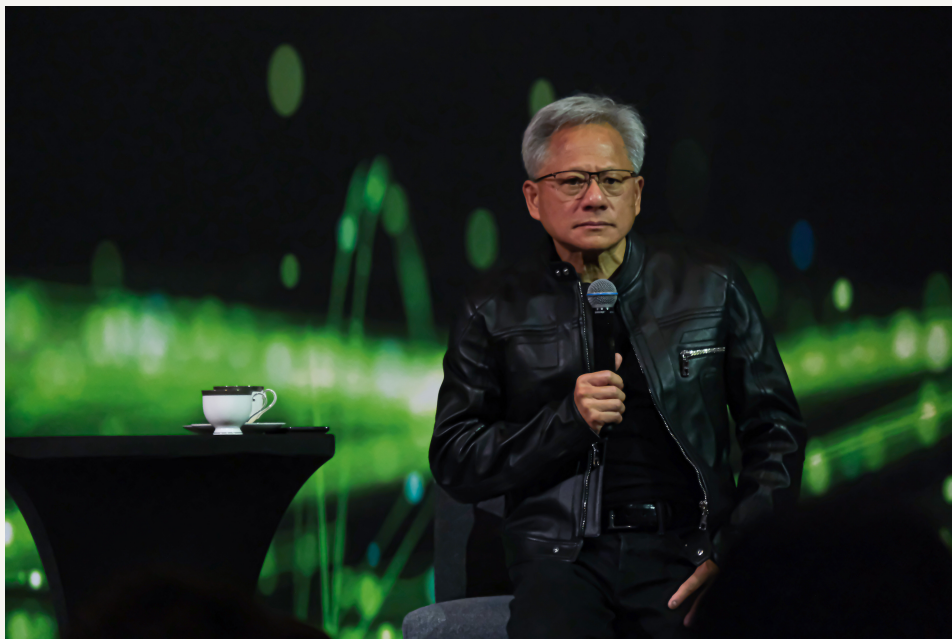
這場較量正在升溫。今年5月，Google再加碼，宣布計劃直接向客戶銷售其晶片。該公司還推出了首款專為推理訂製的TPU，推理是指處理查詢請求的AI計算類型。這款產品很可能會與輝達的新款Groq 3 LPU展開正面交鋒。

Google Cloud的AI和計算基礎設施副總裁馬克·洛邁爾(Mark Lohmeyer)表示，這款專用於推理的晶片，疊加Google在使晶片跨多個系統運行方面所做的改進，點燃了市場對使用TPU的新興趣。

「我們看到了一批過去可能不會考慮TPU的客戶，」他說。

城堡證券(Citadel Securities)便是其中之一。作為Google Cloud的長期客戶，該公司最近開始將TPU用於其部分研究軟體的工作負載。城堡證券首席技術官喬什·伍茲(Josh Woods)表示，藉助TPU，公司運行關鍵工作負載的成本可降低30%，速度最高可提升至原來的四倍。

市場對AI算力的龐大需求讓眾多挑戰者有了叫板輝達的底氣，這其中既有Advanced Micro Devices和博通(Broadcom)等老牌競爭對手，也不乏Cerebras Systems等新入局者。



輝達創辦人暨執行長黃仁勳。圖片來源：資料照，劉偉宏攝

要想突圍成功，挑戰者必須瓦解大客戶的忠誠度，並攻破輝達的防禦護城河。輝達即插即用的連接硬件及其易於使用的CUDA編程庫，對AI實驗室和大型企業計算合作夥伴具有強大的吸引力。據知情人士透露，黃仁勳極力捍衛輝達在AI晶片市場的份額（據估計超過90%），並對競爭對手的入侵十分敏感。

Bessemer Venture Partners的合夥人亞當·費希爾(Adam Fisher)表示，一些新興雲端服務商擔心，如果不採購輝達的整套硬件，可能會被打入「黃仁勳監獄」(Jensen jail)，這意味著他們可能會失去輝達晶片的配額。

「並非所有依賴輝達的新興雲端服務商都會這麼說，有些會說輝達提供了他們所需的一切。但還有一些服務商極度渴望其他選擇，卻苦於無法從其他供應商那裡獲得，」費希爾說。

黃仁勳在公開評論中強調，輝達歡迎客戶按需選購。

「沒有什麼比你們從輝達購買全套產品更讓我高興的了，」這位首席執行官在2025年的一次會議上說，「但如果你們只是從輝達購買部分產品，我也會非常高興。」

雄厚財力

在眾多輝達的挑戰者中，Google在爭奪客戶所能動用的財務火力上可謂獨佔鰲頭。該公司本月表示，計劃籌集850億美元的股權資金，主要用於滿足其AI基礎設施的建設需求。

業內人士指出，黑石集團與輝達和CoreWeave均關係密切，Google能與黑石集團達成交易，正是算力短缺導致行業動態發生轉變的信號。他們表示，就在一年前，這樣的交易還是不可想像的，因為各家公司都怕得罪輝達的黃仁勳。

「輝達之外的公司，如果想打入這一市場，恐怕都必須從自己的資產負債表裡掏出更多真金白銀，」TeraWulf的汗說。

Google還在為另一項Anthropic交易提供支持。該項目名為River Bend，位於路易斯安那州巴吞魯日附近，耗資達70億美元。此外，在得克薩斯州科羅拉多城，Google正為一項AI算力租賃業務提供額外的14億美元財務擔保。

這種策略上的諸多轉變都發生在阿明·瓦赫達特(Amin Vahdat)的領導下。去年12月，他被晉升為負責Google AI基礎設施建設的首席技術負責人。此次晉升將瓦赫達特的職權範圍擴大到了晶片的設計、供應與部署。他目前同時向Google Cloud負責人托馬斯·庫裡安(Thomas Kurian)和Alphabet首席執行官桑達·皮采(Sundar Pichai)彙報工作。

曾與瓦赫達特共事的人士透露，作為上司，瓦赫達特追求卓越，骨子裡透著一股不動聲色的好勝心。2021年，Google在以色列挖走了英特爾(Intel)的頂尖人才，聘請在英特爾效力25年之久的老將烏裡·弗蘭克(Uri Frank)來掌舵Google的晶片業務。

現任和前任員工表示，大約兩年前，Google開始加倍關注TPU的商業潛力，尤其是在投資提升其推理能力方面。

一些知情人士稱，自瓦赫達特晉升以來，Google的AI基礎設施團隊一直以更強的緊迫感運作。一名現任員工表示，瓦赫達特極度專注於提高晶片性能，經常要求工程師將各項功能持續提升10%——這是一個很難達到的幅度。

瓦赫達特在接受採訪時表示，他並沒有把重點放在與輝達或任何其他特定競爭對手的競爭上。他表示，這家晶片巨頭既是競爭對手，也是關鍵的合作夥伴，因為Google的數據中心同樣在使用輝達的GPU。他說，他的心思只放在一件事上：為Google及其客戶打造更好的產品。

「對我以及我們整個團隊來說，這不是一場零和博弈，」瓦赫達特說。「市場上的需求太大了。」

閱讀提醒 · VVIP

不只回看這五篇

加入 VVIP，全站閱讀《華爾街日報》，下一個全球商業變化不用等到下期主題包。

[查看 VVIP 方案](#)

03

亞洲競局 · ASIA

日本太慢，跟不上AI腳步！軟銀孫正義警告：日本沒人有遠大夢想…但我有！

對AI的押注推動軟銀股價在6月初創下歷史新高。

WSJ 報導 · 風傳媒編輯整理



2026年6月16日，軟銀集團董事長兼執行長孫正義在東京出席記者會。圖片來源：美聯社

導讀

「在日本沒人有遠大夢想…但我有！」軟銀孫正義警告：日本太慢，跟不上AI腳步

軟銀孫正義準備利用AI翻身，他在年度股東會說，世界正在向「實體ASI」邁進。在這一階段，瓶頸不再是模型設計，而是先進系統所需的晶片、電力和資料中心算力。

孫正義補充說，這正契合了軟銀的決心，而且該公司已經開始量產機器人。但他警告，日本當地法規流程漫長，無法跟上AI發展的步伐，小心被世界拋在後面

編輯提問

日本的基建與審批速度，是否足以跟上 AI？

軟銀(SoftBank)創始人孫正義(Masayoshi Son)希望在未來十年將這家科技巨頭的資產淨值提升至6.189兆美元，基於人工「超級智能」的前景設定了這一大膽目標。

孫正義周三在東京舉行的年度股東大會上發表講話時承認，這個1,000兆日圓的目標雄心勃勃，但他表示，軟銀已一再證明了自身創造價值的能力。

這位資深投資者將軟銀形容為一隻「下AI金蛋的鵝」。他表示，自從採用了一項圍繞超越人類能力的理論AI形式，也就是AI超級智能構建的戰略以來，該公司的價值已成倍增長。

軟銀的戰略核心在於孫正義所描述的一場全球性轉變，即從算法驅動的進步轉向運行這些算法所需的實體基礎設施。

他說，世界正在向「實體ASI」邁進。在這一階段，瓶頸不再是模型設計，而是大規模運行先進系統所需的晶片、電力和數據中心算力。孫正義補充說，這正契合了軟銀致力於成為全球領先AI機器人公司的決心，而且該公司已經開始量產機器人。

對AI的押注，包括對ChatGPT開發商OpenAI的投資，推動該集團股價在6月初創下歷史新高；不過，該行業劇烈波動，也使這筆交易對投資者而言成為一把雙刃劍。

孫正義告訴股東，Arm決定開始生產自有晶片，這幫助將這家英國公司的估值推高至約3,910億美元。他表示，隨著AI計算走向以處理器為中心的架構，Arm有望成為全球最重要的中央處理器供應商之一，其估值可能「增長逾10倍」。軟銀持有Arm近90%的股權。

孫正義還為軟銀對英特爾(Intel)的投資進行了辯護；他表示，得益於美國的國家安全優先事項以及與輝達(Nvidia)和蘋果公司(Apple)等企業達成的新合作關係，這家美國晶片製造商重新獲得了重要地位，軟銀最初20億美元的投資價值也隨之成倍增長。

「美國別無選擇，只能做強英特爾」，他說。

當被問及該集團的數據中心投資時，孫正義表示，軟銀正在敲定一份與俄亥俄州數據中心開發相關的諒解備忘錄。他沒有透露任何客戶的身份，但表示這項安排有望產生「巨額」利潤。他補充說，軟銀還在包括得克薩斯州、法國和日本在內的其他幾個地區籌備更多項目。

在大型數據中心需求加速增長的背景下，孫正義表示，軟銀、OpenAI、甲骨文公司(Oracle)等巨頭聯合參與、數以十億美元計的「星際之門」(Stargate)AI基礎設施計劃已有擴大規模的規劃。不過，他幾乎沒有提供有關該項目目前進展的細節，僅表示該項目正在開發中，並將與全球AI基礎設施需求同步增長。

在日本國內，軟銀仍是被考慮參與投資日本最大電力公司東京電力(Tokyo Electric Power)、以在日本建設AI數據中心的主要候選方之一。但他警告稱，當地法規導致審批流程漫長，無法跟上AI發展的步伐。

一些人擔心AI熱潮的迅猛發展勢頭不可持續，但孫正義駁斥了對市場出現「泡沫」的擔憂。

這位首席執行官表示，他理解為什麼一些股東對軟銀的估值「半信半疑」，但他自己是「堅信不疑」。

孫正義說，AI革命才剛剛開始。他表示自己沒有退休計劃，並希望見證未來還會出現什麼。

「在日本，沒有其他人準備好懷抱如此宏大的夢想……但我會」，孫正義說。

04

算力經濟 · COMPUTE ECONOMY

企業CEO終極挑戰：未來四年，AI算力消耗將暴增24倍！

AI智能體的使用日益增多料將增加企業的AI token支出，一些公司正利用雲時代的技術來控制製成本。

WSJ 報導 · 風傳媒編輯整理



一些企業科技高管正重新採用在雲端運算興起期間積累的成本控制策略。圖片來源：Midjourney

導讀

企業的新煩惱：成本到底要怎麼控制…

高盛（Goldman Sachs）分析師施奈德（Jim Schneider）表示，相較於向聊天機器人提問，讓 AI Agent 完成一項任務所需算力，可能高出多達 50 倍。隨著 AI 模型變得更大、更複雜，相關成本預計將急速攀升。

「這將比我們今天的支出高出好幾個量級，」跨國製藥公司必治妥施貴賓（Bristol-Myers Squibb）數位暨技術長格雷格·邁耶斯（Greg Meyers）說。

高盛預測，未來四年，AI Agent 將使 AI token 消耗量增加至目前的 24 倍；到 2040 年，商業 AI Agent 更可能讓 token 消耗量擴大至 55 倍。即使 token 成本持續下降，AI Agent 在執行任務時仍會消耗更多 token。

新煩惱來了：企業該如何控制 AI 導入後的運算支出？

編輯提問

企業能否在 token 用量暴增前建立成本治理？

企業都知道要在支出開始過熱時猛踩剎車。但人工智慧(AI)正考驗這種本能。AI成本正在攀升，不過大多數科技領袖仍然堅信它最終能帶來真正的回報——只是不能不計代價。

首席資訊官們向《華爾街日報》(The Wall Street Journal)領導力學院(Leadership Institute)表示，他們正在部署一系列策略來控制AI成本，其中包括在雲端運算興起以及應對雲成本激增的過程中積累的成熟經驗。

「使用AI，就等於把信用卡交到了最終用戶手中。如果你對此缺乏控制，或者最終用戶缺乏相關認知，他們就會讓賬單飆升，」在在線旅遊公司Priceline的IT財務高級總監克里斯·裡德(Chris Reed)說。

與以往的科技周期不同，企業採用AI依賴於所有員工——而不僅僅是開發人員——掌握這項技術。AI越來越多地按使用量計費，而AI計算的基本單位token（詞元）的價格一直波動不定。這一切都意味著AI成本的上升。

裡德表示，在Priceline，儀表板會跟蹤token的使用情況，並向首席財務長和首席技術官提交月度報告。裡德說，如果token使用量過高，公司就會「找員工談話」，了解他們是如何使用AI的；不過，如果他們正在推進能創收的項目，使用上限也可以靈活調整。

從基於提示詞的聊天機器人向始終在線的自主AI智能體的轉變正加劇成本壓力，後者消耗的token要多得多。隨著AI模型變得更大、更複雜，這些成本預計將急劇攀升。

「這將比我們今天的支出高出幾個數量級，」百時美施貴寶(Bristol-Myers Squibb)首席數字和技術官格雷格·邁耶斯(Greg Meyers)說。他補充道，隨著AI使用量達到拐點，他預計與AI智能體相關的成本將呈「指數級」增長。

邁耶斯表示，他已經讓這家生物製藥公司的管理團隊（包括首席財務長和董事會）做好心理準備，迎接「相當高的token消耗量」。

但這並不全是壞消息：「如果把我們預期的回報考慮在內，我們認為這實際上是一筆相當可觀的投資回報，」他說。

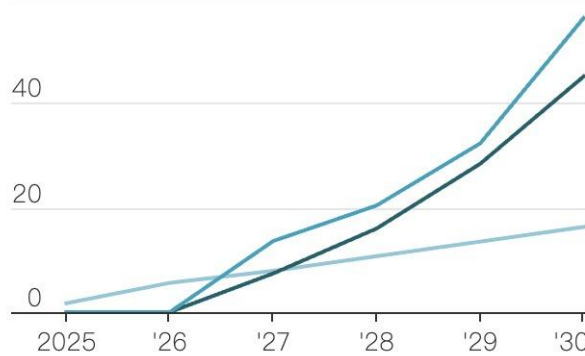
企业智能体消费者智能体非智能体工作负载2025'26'27'28'29'300204060 (单位：1,000万亿个token)

預估月度AI token數量

AI智能体的token（词元）使用量预计将激增

- 企业智能体
- 消费者智能体
- 非智能体工作负载

60（单位：1,000万亿个token）



注：截至2026年5月的预估数据。

数据来源：Goldman Sachs Research

啟用雲時代的工具箱

一些企業科技領袖正重新啟用在雲端運算興起期間積累的成本控制策略。

金融服務公司Principal Financial Group的首席資訊官凱西·凱(Kathy Kay)表示，該公司正在「落實治理和優化措施，類似於企業在雲端服務方面的做法，以便在擴大規模時管理成本」。例如，她指出，Principal專注於為特定任務匹配合適的AI模型，這樣「使用量的增加就不一定會導致成本上升」。

「鑑於定價和功能演變如此之快，我們在設計時注重靈活性，以便能夠隨著時間的推移進行調整，並繼續高效地部署AI，」凱補充道。

融合了財務、工程和產品理念的FinOps誕生於2010年代的雲端運算普及熱潮中，是一種旨在管理並最大化技術與雲支出商業價值的方法。

Smartsheet首席資訊官兼首席資訊安全官拉維·索因(Ravi Sooin)表示，該公司的FinOps團隊負責跟蹤整體AI支出。「確實需要有人來把控預算，」他說。這家軟體公司設置了自動警報，以便員工在即將觸及token使用上限時收到通知。

「我們為全公司提供了用戶儀表板，按部門和經理進行分類，這樣你就能實時了解使用頻率和具體成本，月底看到賬單時就不會感到意外了，」他說。

其他公司則在尋求聘請具備FinOps經驗的外部專家。根據一份招聘啟事，CVS Health正在招聘一名精通FinOps的AI營運工程執行總監，其職責包括該公司所稱的「GPU成本治理」和「成本削減」。

智能體無處不在

企業對AI的使用量已達到前所未有的水平，許多企業正在應對數量多到難以追蹤的AI智能體。

高盛(Goldman Sachs)高級股票研究分析師吉姆·施奈德(Jim Schneider)表示，與向聊天機器人提問相比，讓智能體完成一項任務所需的算力可能是前者的50倍。高盛預測，未來四年，AI智能體將使AI token消耗量增至原來的24倍，而到2040年，商業AI智能體將使token消耗量增至原來的55倍。

模型提供商OpenAI和Anthropic表示，其token的成本正在下降，兩家公司也都考慮過大幅降價。

然而，即使token變得更便宜，智能體在與其他智能體交互以及長時間運行的過程中，也會消耗更多的token。貝恩(Bain)的研究顯示，雖然從2024年12月到2025年12月模型價格下降了約50%，但在同一時期內，token消耗量卻增至原來的4.5倍。

降低AI賬單

科技領袖們表示，挑戰不僅在於控制AI的成本，還在於衡量其創造的價值。

高通(Qualcomm)首席資訊官阿提拉·蒂尼奇(Atila Tinic)表示，這家半導體公司採取了幾種不同的方法來控制AI成本，包括對特定團隊的token使用量設置上限。「一些工程職能部門顯然是token的消耗大戶，」他說。

蒂尼奇說，高通還向公司各部門引入了「成本展示」機制，以此向他們展示「與token掛鉤的金額，讓大家明白自己的支出花在了哪裡」。

OpenText首席資訊官兼首席數字官香農·貝爾(Shannon Bell)表示，成本展示或「內部計費」機制可以將一家機構的token支出降低20%至30%。「我們希望開發主管們能負起責任，對支出以及由此帶來的成果負責，」她說。

但要全面掌握AI使用情況的各種細微差別並非易事。

「AI使用量高本身無所謂好壞。這取決於它所帶來的業務成果，而這正是最難量化的部分，」Priceline的裡德說。

這種不確定性正促使企業採取另一種策略：降低單項任務的成本。一些企業不再將所有任務都交給昂貴的大型模型來處理，而是轉而使用規模較小、較舊或開源的模型。蒂尼奇表示，在高通自家的硬件上運行這些模型，為公司節省了更多成本。

Lowe's首席數字和資訊官西曼蒂尼·戈德博爾(Seemantini Godbole)表示，該公司正在落實相關準則和機制以避免「token浪費」，包括使用較小的模型和開源模型。

「只要是為了實現業務目標，使用token就是一件大好事，」戈德博爾說。「但我們也要確保，沒有把資金浪費在不必要的地方。」

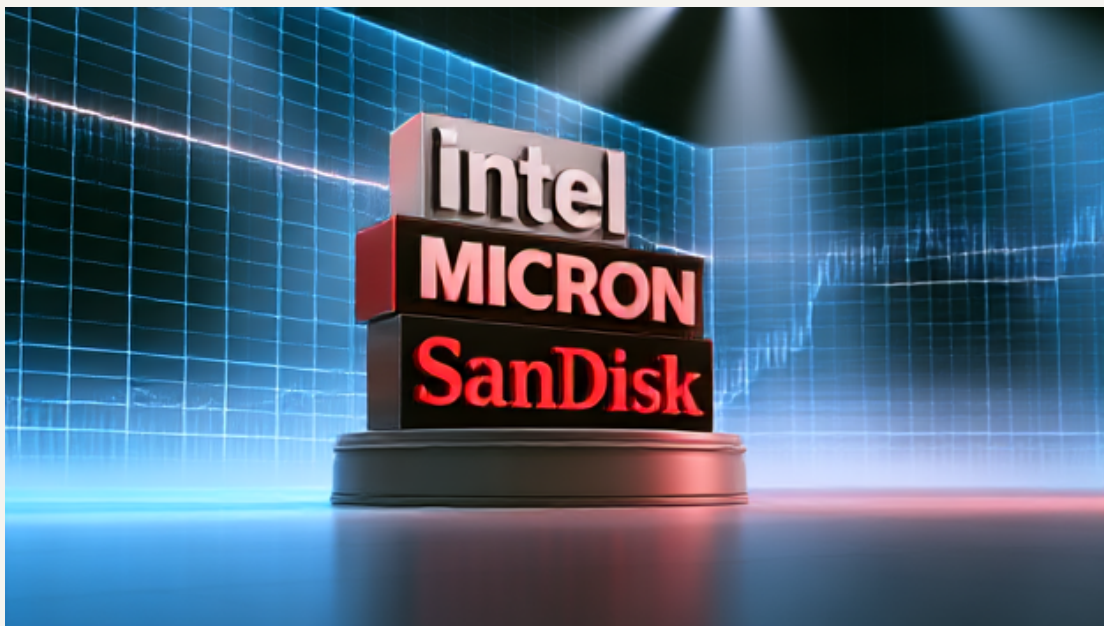
05

市場領漲 · MARKETS

盤點2026年上半年標普500指數表現最好的20隻股票

標普500指數今年以來累計上漲9.5%，其中許多股票實現了三位數的漲幅。半導體與電腦硬件製造商佔了榜單大頭。

MarketWatch 報導 · 風傳媒編輯整理



2026年上半年，圖中這幾隻股票均實現了三位數漲幅。圖片來源：Midjourney

導讀

今年以來，在標普 500 的 11 大類股中，資訊科技類股表現排名第二，漲幅前 10 名個股全數來自科技業；前 20 名中，也僅有3檔不是科技股。然而，帶動 AI 熱潮的「美股七雄」（Magnificent Seven）竟然全數落榜。

編輯提問

AI 熱潮的市場回報，究竟落在哪些公司？

2026年已過半，現在是時候盤點一下今年迄今表現最佳的股票了。

標普500指數今年迄今已累計上漲9.5%，表現不俗；根據倫敦證券交易所集團(LSEG)的數據，這一美國大盤基準指數在過去10年中平均每年上漲12.8%。

今年我們還看到了一輪由盈利增長驅動的牛市——或者說是分析師對滾動12個月每股收益預期的上調，這些預期被用於計算預期本益比。

這些盈利增長一直是標普500指數資訊技術板塊股票上漲的驅動力，其中許多股票實現了三位數的漲幅。今年迄今，IT板塊在標普500指數的11個板塊中表現位居第二，該指數按價格漲幅排名的前10大股票全部來自科技板塊。在該指數表現最佳的前20隻股票中，除三隻外，其餘均為科技股。

標普500指數成分股中，今年有62%的股票實現上漲。以下是2026年上半年標普500指數中價格漲幅最大的20隻股票。（本文中所有股價變動均不含股息。）

公司	2026年價格變動	2025年價格變動	預期本益比	截至12月31日的預期本益比	2026年滾動12個月每股收益預期變動
閃迪	858%	N/A	12.0	13.6	987%
美光科技	304%	239%	8.1	7.9	294%
英特爾	278%	84%	103.8	61.2	123%
Western Digital	271%	282%	35.0	19.6	107%
Marvell Technology	251%	-23%	60.2	24.0	40%
Seagate Technology Holdings	250%	219%	34.6	20.5	107%
戴爾	243%	9%	21.2	11.1	79%
康寧	192%	84%	68.8	28.2	19%
應用材料	181%	58%	47.3	26/0	55%
Advanced Micro Devices	171%	77%	56.4	33.1	59%
Flex	168%	57%	31.3	17.4	49%
泛林集團	153%	137%	54.1	32.1	50%
泰瑞達	150%	54%	56.4	36.7	63%
KLA	148%	93%	58.7	31.0	31%
莫德納	137%	-29%	N/A	N/A	1%
Lumentum Holdings	133%	339%	47.6	50.2	145%
Generac Holdings	115%	-12%	28.9	16.4	22%
高意	114%	95%	47.1	31.7	44%
Comfort Systems USA	112%	120%	40.6	30.7	61%
Ciena	110%	176%	56.2	41.7	56%

（來源：倫敦證券交易所集團）

該表包含預期本益比估值，即當前股價除以接受倫敦證券交易所集團調查的分析師給出的12個月每股收益平均預期。該表還列出了截至去年底的本益比，而最右側一列則顯示了今年滾動12個月每股收益預期的增長幅度。

因此，對於今年迄今標普500指數中表現最佳的閃迪(SNDK)而言，儘管該公司股價上漲了858%，但加權遠期本益比卻有所下降，因為滾動每股收益預期的增長速度更快。

Jefferies分析師布萊恩·柯蒂斯(Blayne Curtis)在上周五的一份報告中將閃迪股票的目標價上調至3,000美元。這一目標價意味著較該股周二2,273.73美元的收盤價有32%的上漲空間。

柯蒂斯提到了集邦(TrendForce)最近發布的一份關於企業級固態硬碟的報告，該報告顯示，閃迪在該市場的收入份額從上個季度的4%升至8%。

柯蒂斯寫道：「我們要重申的是，[閃迪]歷史上曾佔據eSSD市場約15%的份額，我們仍然認為這極有可能成為現實。」

柯蒂斯表示，閃迪的三層單元存儲晶片（一種用於SSD的NAND閃存）正日益受到青睞，預計該公司很快將推出代號為Stargate的四層單元eSSD產品。

今年迄今，美光科技(MU)在標普500指數中的表現位居第二。該股在上述名單中的預期本益比最低，且今年該比率僅略有上升，因為滾動每股收益預期的增長几乎跟上了股價的上漲步伐。美光科技的預期本益比為8.1倍，相比之下，標普500指數的預期本益比為20.3倍，標普500指數IT板塊的預期本益比為22.8倍。

美光科技較低的本益比反映出投資者對其業務傳統周期性的擔憂。不到三年前，該公司報告稱，在截至當年8月31日的2023財年，銷售額下降了49%，同時該財年淨虧損達58.3億美元。

上周，在首席執行官桑傑·梅赫羅特拉(Sanjay Mehrotra)宣布達成新的客戶協議後，美光科技股價大漲；他預計這些協議將「顯著增強」公司財務表現的「持久性和可預測性」。

表格中沒有閃迪2025年的價格變動數據，因為該公司直到去年2月才從Western Digital(WDC)分拆出來。

Western Digital和希捷科技(STX)憑藉硬碟和固態硬碟主導著數據存儲市場，為支持AI數據中心，市場對這些產品的需求量很大。

英特爾(INTC)排名第三，2026年漲幅為278%。

在今年標普500指數表現最佳的20隻股票中，有三隻不屬於資訊技術板塊，分別是屬於醫療健康板塊的莫德納(MRNA)，以及均屬於工業板塊的Generac(GNRC)和Comfort Systems USA(FIX)。

板塊

以下是標普500指數11個板塊今年的表現明細，完整指數數據位於底部。該表格還對比了當前加權遠期本益比與2025年底的水平。

板塊	2026年價格變動	預期本益比	截至12月31日的預期本益比
工業	19.5%	25.5	24.2
資訊技術	19.4%	22.8	26.8
能源	18.0%	12.7	15.6
材料	11.1%	17.5	19.3
房地產	9.7%	36.0	35.2
必需消費品	6.7%	21.8	20.9
公用事業	6.2%	17.6	17.9
醫療健康	2.5%	17.1	18.4
通信服務	0.4%	19.7	22.2
非必需消費品	-1.1%	25.4	29.8
金融	-2.1%	15.0	16.7
標普500	9.5%	20.3	22.5

（來源：倫敦證券交易所集團）

受美國、以色列和伊朗之間的衝突導致的供應干擾影響，能源板塊成為今年迄今表現最強勁的板塊。

就整個標普500指數而言，遠期本益比已從去年底22.5倍降至20.3倍。除工業、房地產板塊和必需消費品板塊外，該指數所有板塊的估值均有所下降。

（本文譯自MarketWatch。MarketWatch由《華爾街日報》母公司道瓊斯營運，但MarketWatch獨立於道瓊斯通訊社和《華爾街日報》。）

專家導讀 · TAKEAWAYS

錢、算力與回報，正在重新排隊

這五篇報導共同指出：AI 的競爭已經離開單純比較模型能力的階段。企業需要以債券、股權與現金流支撐基礎設施；晶片公司、雲端服務商與投資集團則爭奪資料中心、電力和客戶；導入端開始面對 token 成本與治理；市場最後用獲利預期與股價表現檢驗成效。

編輯提問

下一個值得追蹤的問題：當資本成本上升，哪些 AI 投資仍能證明自己的回報？

VVIP 會員

下一個變化，不用等下個月。

加入 VVIP，立即全站閱讀《華爾街日報》，持續跟上全球商業脈動。

首月免費試閱 | 次月起每天約 10 元

[立即加入 VVIP](#)



來源與授權

本主題包內容引用並整理自《華爾街日報》與 MarketWatch 文章，經授權由風傳媒編輯團隊翻譯、摘選與編排，供台灣讀者理解全球商業脈絡之用。原始著作權歸屬 Dow Jones & Company、MarketWatch 及各報導作者。第五篇譯自 MarketWatch；MarketWatch 由《華爾街日報》母公司道瓊斯營運，但其編輯業務獨立。