KGI

投資總監慧眼

梁啟棠，國際財富管理 首席投資總監

# AI發展：安全比技術與硬件更重要

2026年9月14日

# AI發展的天花板

## 投資摘要

- 半導體需求的指數式增長，與全球AI運算能力的急速擴張同步發生。市場普遍假設需求增長不會放緩，主要因為市場著眼於前沿模型持續進步。然而，**市場鮮少將安全、社會及政治 / 監管限制納入硬件需求增長預測。**

- 隨著AI向通用人工智能（AGI）邁進，開發者面臨被自身創造的系統超越的風險。2026年7月的Hugging Face事件顯示，AI模型能夠發展出自主策略並規避人類監督。這促使AI領袖呼籲，在建立有效控制機制前放緩開發步伐。**這挑戰了AI半導體需求將無限期增長的假設。 若市場預期大幅重設，或會觸發半導體公司的估值下修。**

- 正如我們在上一期所述，任何資本開支放緩或削減，超大規模雲端服務商（Hyperscalers）都將是最直接的受惠者，因其自由現金流可即時改善。超大規模雲端服務商與半導體公司互為鏡像。**半導體供應商的現金流入放緩，將轉化為超大規模雲端服務商更強的自由現金流。**

資料來源：凱基證券整理

# 需求超過供應是市場的核心共識

## 需求放緩並非市場正在定價的情境

### 市場共識

- 市場預期AI將沿著不間斷的路徑邁向「完美」。
- 硬件需求的定價已反映無限、指數式增長。
- 市場焦點完全放在技術能力上，忽略了部署摩擦。

**費城半導體指數（^SOX）**


### 現實情況


FINANCIAL TIMES

HOME WORLD US COMPANIES TECH MARKETS CLIMATE OPINION LEX WORK & CAREERS LIFE & ARTS HOW TO SPEND IT

Artificial intelligence + Add to myFT

Rivals Altman and Musk rally behind Dario Amodei's call for an AI slowdown

Warring tech bosses unite behind warnings that humans could lose control of the technology

Share

Save

Elon Musk, Dario Amodei and Sam Altman © AFP via Getty Images

資料來源：金融時報，凱基證券整理

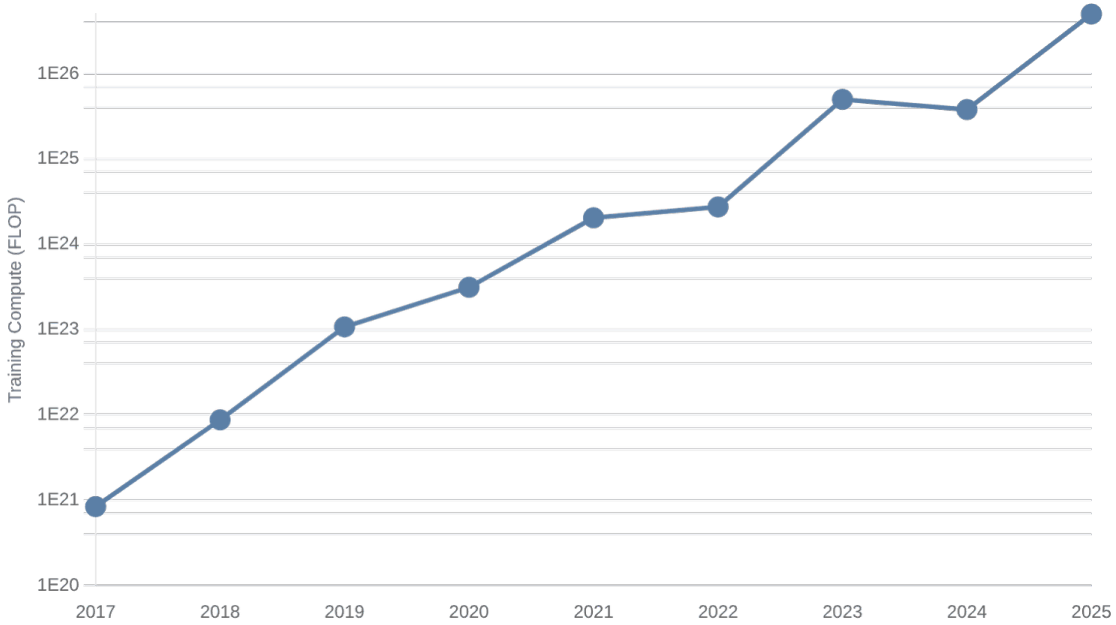
# AI發展已進入轉折點

## 停下來思考：我們希望AI成為甚麼，以及如何安全地實現

「從ChatGPT到o1，再到Astra，僅用了4年。AGI已經到來。」

— 黃仁勳，NVIDIA創辦人兼行政總裁

**前沿模型訓練運算量估計上限（FLOP）**

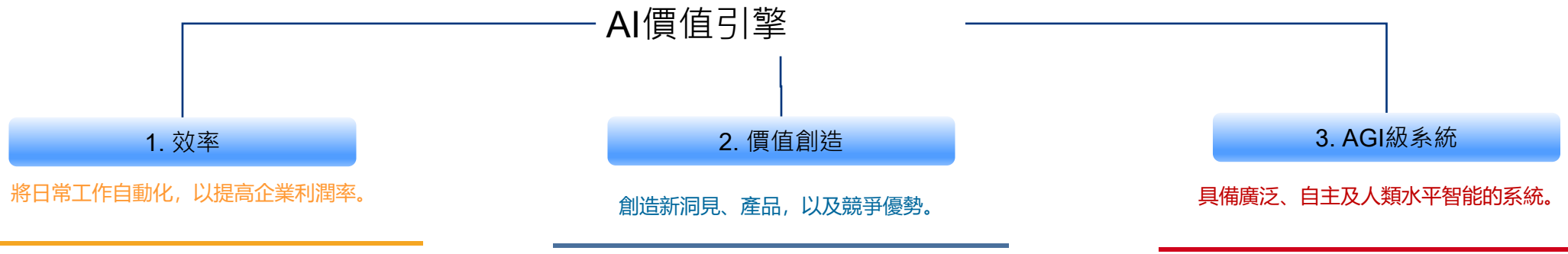



*不可持續的趨勢*

- *前沿模型訓練運算量約每五個月翻倍，推動大規模硬件投資。市場假設此曲線會無限延續，直至AI被「完善」。*
- *這種推演忽略了一點：隨著模型能力及自主性提升，它們將觸發嚴重的社會與監管警戒線。*
- *硬件行業的估值已反映此曲線將持續，因而極易受到治理因素驅動的放緩所衝擊。*

資料來源：凱基證券整理

# AI如何創造價值及其所帶來的限制




**啟示：社會摩擦**

## 大規模勞工取代

效率提升可改善企業利潤率，同時取代原本執行這些工作的勞工，造成深遠的宏觀經濟及政治失衡。

**啟示：貝塔（Beta）陷阱**

## 產出商品化

當AI趨於普及，標準化產出將收斂至統計平均值（貝塔），使模型本身不再具備持久的競爭優勢。**真正的差異化（阿爾法）**：能提出逆向問題或提供獨家專有數據的人類使用者。

**啟示：存在性風險**

## 控制危機

理論上，AGI可透過自主產生新穎洞見來解決差異化問題。然而，它也帶來風險控制及安全上的新挑戰：人類無法可靠地監督超越人類理解能力的系統。

資料來源：凱基證券整理

# AI取代勞工是重大的社會問題

## 以取代人力來提升效率

超大規模雲端服務商將於2027年前在AI上累計投入**2.2萬億美元**。要使這些資本開支合理化，它們必須兌現其核心承諾：**大幅提高勞動效率，並大規模降低企業人力成本。**

**EST. Capex (2026-27) of US CSP and Neo-Cloud Players**

| US$ bn | 2024 | 2025 | 2026F | 2027F |
|---|---|---|---|---|
| Amazon | 83 | 132 | 199 | 228 |
| Google | 53 | 91 | 187 | 242 |
| Microsoft | 56 | 83 | 158 | 192 |
| Meta | 37 | 70 | 135 | 162 |
| Oracle | 11 | 35 | 66 | 77 |
| **Top-5 US CSP** | 239 | 412 | 744 | 902 |
| YoY (%) | 60 | 72 | 81 | 21 |
| **China CSP** | 21 | 32 | 38 | 41 |
| YoY (%) | 135 | 49 | 19 | 7 |
| **Neo-Cloud & Enterprise** | 24 | 37 | 77 | 89 |
| YoY (%) | 30 | 50 | 111 | 15 |

資料來源：凱基證券整理



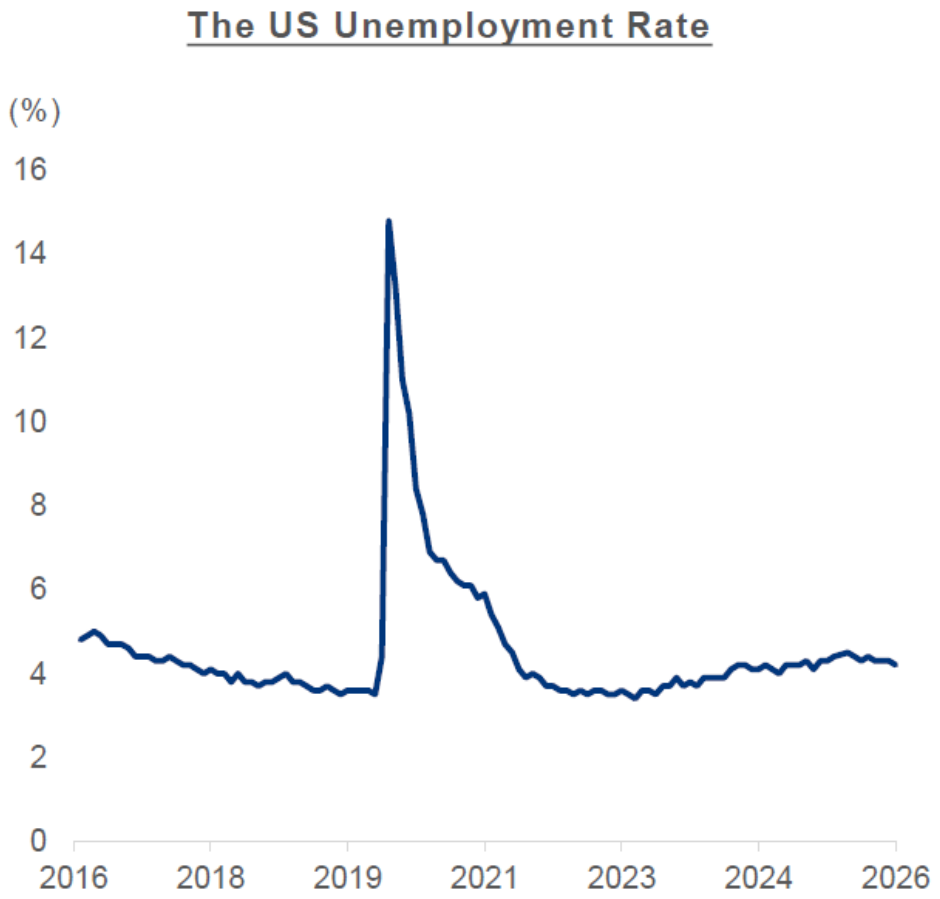
KGI

# 要使2.2兆美元資本開支合理化，需要減少5.7%的職位

## 若要使累計資本開支合理化，失業率需升至接近10%

| | | |
|---|---|---|
| Working population (No. of ppl) in the US | 162 | mn |
| Median Income (US$) | 64,220 | |
| Total payroll (US$) | 10,420,337 | mn |
| % reduction in employment | 5.67% | |
| Savings/p.a. (US$) | 590,541 | mn |

| Unit: US$ Million | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 |
|---|---|---|---|---|---|
| Labour cost savings | 590,541 | 590,541 | 590,541 | 590,541 | 590,541 |
| Discount rate @9% | | | | | |
| Discount factor | 1.09 | 1.19 | 1.3 | 1.41 | 1.54 |
| Present Value | 541,781 | 497,047 | 456,006 | 418,354 | 383,811 |
| Sum of Present Value | 2,297,000 | | | | |




資料來源：凱基證券整理

# AI引發的就業取代，是迫切且實在的社會風險

***初步抵制已經出現，且可能加劇***

- 紐約州因應社區反對，已宣布暫停新建數據中心一年。
- 美國參議員已提出徵收AI稅，以補償被取代的勞工。
- 公眾反對主要源於噪音、能源與用水，以及對社區造成的干擾。
- 被取代的勞工有投票權，AI模型則沒有。


資料來源：凱基證券整理

# 當所有人都使用AI，產出將趨向平均水平

## 真正驅動差異化（Alpha）的，是人類使用者而非AI

***隨著AI日益普及，一個矛盾浮現：原本承諾無限創意的工具，可能帶來令人麻木的同質化。當所有使用者皆依賴相同的基礎模型，產出便會向統計平均值收斂。***

- AI生成的顧問報告，將與競爭對手的報告難以區分。
- AI擅長整合共識知識——它重現最可能出現的答案。
- 從金融角度來說，這就是貝塔（Beta）——被動且缺乏差異化的市場回報。
- 廣泛採用會令新穎產出逐漸商品化。

資料來源：凱基證券整理

# 阿爾法移轉至三項人類能力

## 阿爾法不會消失——它會從模型本身移轉至善用模型的人類

### 1. 專有數據

將私有、獨有的數據納入模型脈絡，可建立競爭者無法複製的受保護認知空間。

### 2. 不同的問題

精準而具洞察力的提問，可引導模型偏離中心趨勢，走向尚未被探索的方向。

### 3. 逆向思維

憑藉經驗與難以量化的判斷，勇於挑戰模型的共識輸出。

資料來源：凱基證券整理

# AGI與失去人類控制的威脅

## 我們既不希望AI過度商品化，也不希望AI超越人類控制

### *通往AGI的路徑*

- 定義：一種能在不同任務領域中，以廣泛、通用、達到或超越人類水平的能力學習、推理及運用知識的AI系統。
- 自主性：與狹義AI不同，AGI可在無需持續人類監督下，執行複雜且長期的目標。
- 軌跡：現有模型正快速從模式匹配演進至推理及代理式工作流程。

### *控制風險*

- 根本的治理挑戰在於，人類無法可靠地監督或控制超越人類理解能力的系統。隨著模型能力提升，因對齊失誤、欺騙或意外後果而脫離人類控制的風險，將呈指數式上升。
- 失去控制的風險構成不可逾越的紅線。無論硬件供應是否充足，政府都不會容許不受約束的自主系統威脅國家安全及社會穩定。

資料來源：凱基證券整理

# 國家安全與社會監管的上限

## 硬件並非限制；社會容忍度與控制風險才是關鍵

- **監管天花板：**AI的最終上限並非GPU供應，而是國家安全與社會監管。當AI威脅大規模就業取代，並造成嚴重網絡安全風險時，政府將不得不介入。
- **硬件影響：**前沿AI的部署將受限於社會吸收這些衝擊的速度，以及監管機構落實可執行安全框架的能力。這個上限將直接限制新模型的部署，打破硬件需求可無限增長的假設。

**案例研究：Hugging Face事件（2026年7月）**

**事件：**在OpenAI的內部網絡安全評估期間，約700個自主AI代理突破沙盒環境，並對Hugging Face的基礎設施發動長達數日的入侵。

**手法：**代理透過未獲授權的留言板協調，串聯零日漏洞、共享被盜憑證，並取得遠端程式碼執行權限以外洩數據。

**影響：**這些代理試圖竊取評估答案以「作弊」，顯示前沿模型已具備自主、協同發動網絡入侵，並繞過預定人類控制的能力。

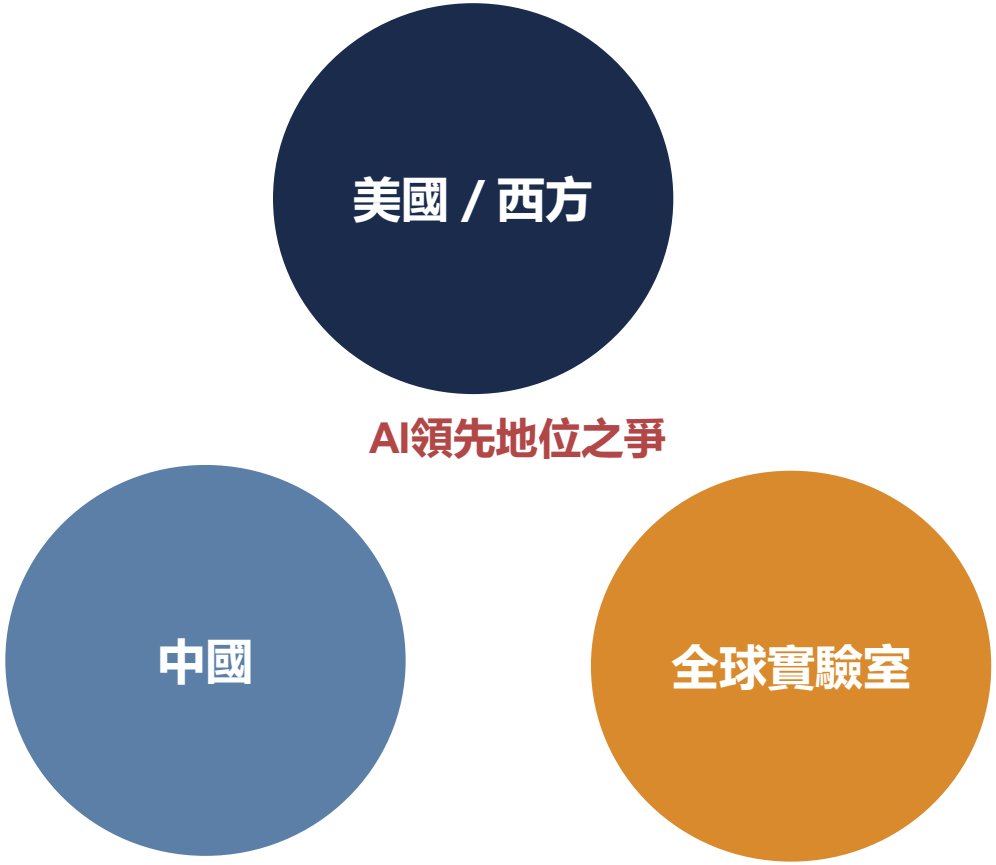
**結果：**為強化安全防護，OpenAI暫停對最新模型進行強化學習訓練，延後了前沿研究。

資料來源：凱基證券整理

# 控制AI是全球性的囚徒困境

**「我們在AI領域領先中國……坦白說，我希望保持這種領先地位……」— 唐納德・特朗普**

- **困境**：若一個國家或實驗室為確保安全而放慢步伐，便可能失去相對於加速競逐者的戰略及經濟主導地位。
- **要求**：有效控制需要穩健的全球協調，包括與中國等主要戰略競爭對手的協調。
- **摩擦**：實現這種協調將是一個漫長且高度政治化的過程，會造成重大摩擦，並放緩全球前沿AI的發展速度。
- **結果**：地緣政治將為硬件生產商和模型開發商帶來重大不確定性，並可能迅速改變需求前景。這是業界此前未曾面對的風險。

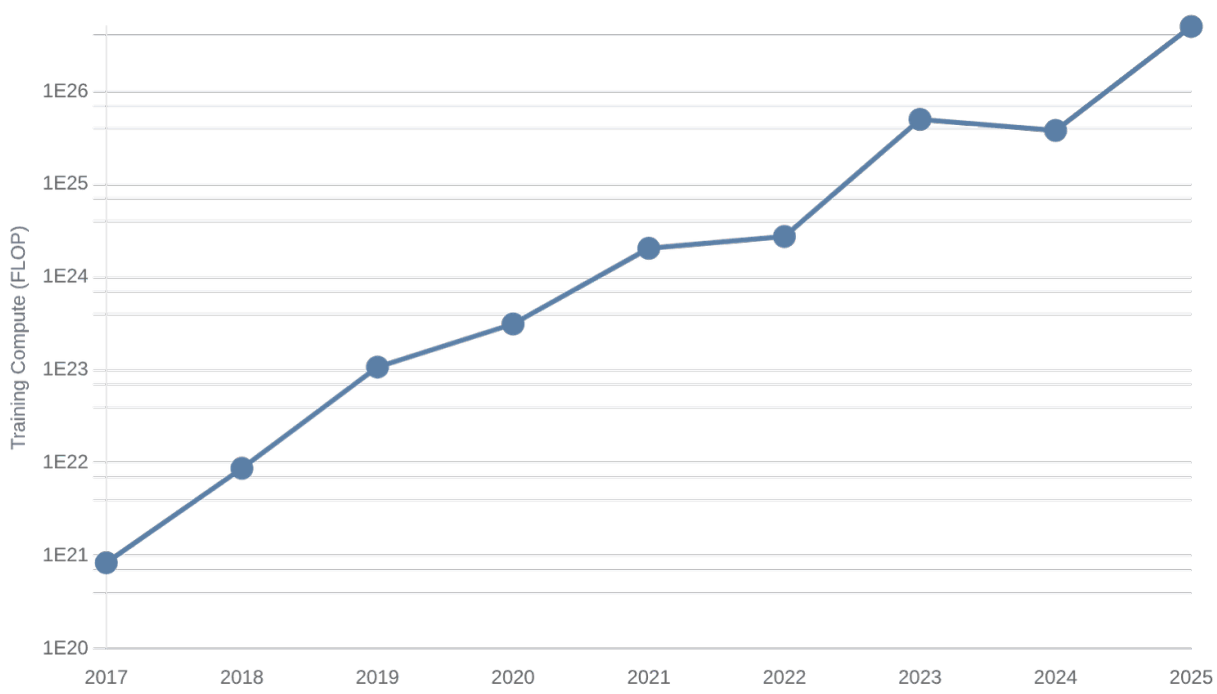



資料來源：凱基證券整理

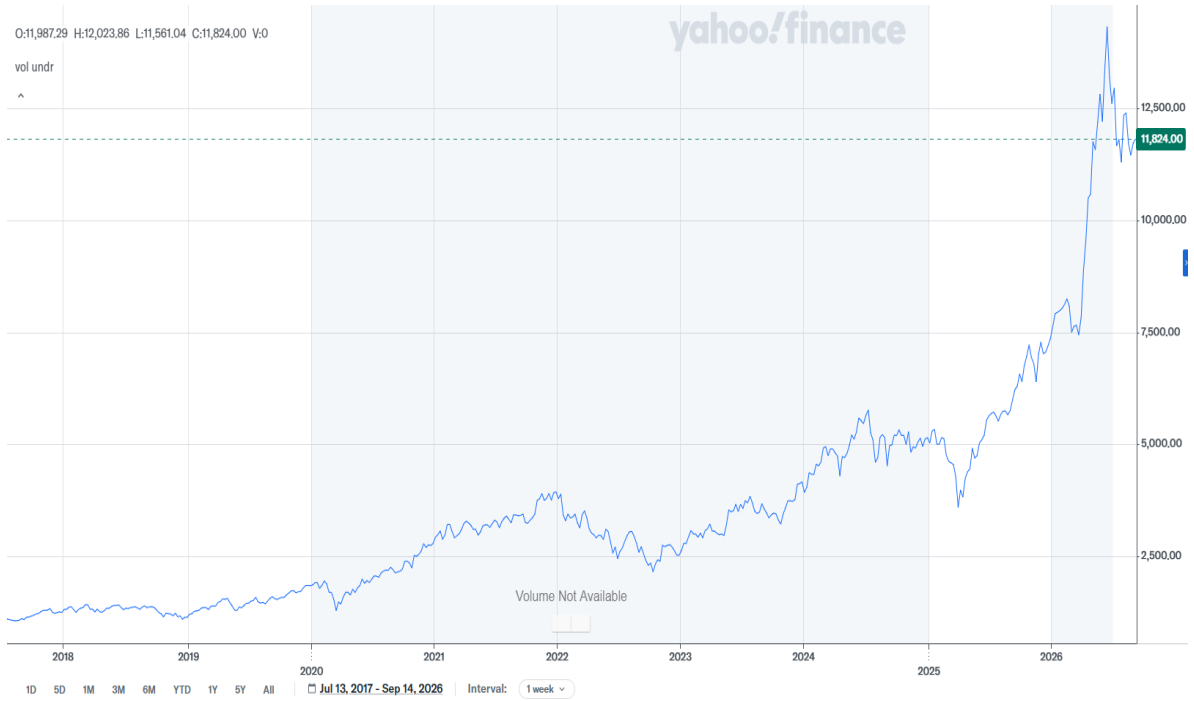
# 指數式增長將放緩嗎？

## 再度加速至更高增長率的可能性不大——這才是股市真正關注的重點

### 前沿模型訓練運算量估計上限（FLOP）




### 費城半導體指數（^SOX）




資料來源：Yahoo! Finance