

關鍵硬件需求激增 馬斯克秘購近萬個 全球AI大戰 科企瘋搶GPU

【大公報訊】綜合《華爾街日報》、路透社報道：由OpenAI開發的人工智能（AI）聊天機器人ChatGPT去年底問世後，全球迅速掀起一股AI熱潮。美媒稱，各大科企紛紛加入AI戰場，瘋狂搶購用於被稱為「生成式AI系統命脈」的圖形芯片（GPU），「就像疫情期間狂買衛生紙一樣」。芯片巨頭英偉達幾乎壟斷AI所需的GPU市場。大摩最新預測，英偉達將在今年的AI產品市場中獲得60%的份額。

獲得數以萬計的高級GPU對於培訓大型AI模型的公司來說至關重要。如果沒有足夠的GPU，這些大型語言模型的運行速度就要慢得多。

英偉達股價大漲167%

隨著生成式AI的爆火，各家科企巨頭爭相開發自家AI產品，導致對GPU等關鍵硬件的需求激增。一家幫助公司構建AI模型的初創公司Lamini的聯合創始人莎朗形容現在的GPU「就像是疫情期間被瘋搶的衛生紙」。

目前市場上的GPU大多來自英偉達的供應，其高級GPU擅長同時進行大量計算。瑞銀分析師估計，較早版本的ChatGPT需要大約1萬個GPU；馬斯克認為，新版本的ChatGPT可能需要三至五倍的GPU數量。

旺盛的需求推動英偉達股價一度上漲約167%。該公司最近表示，正在擴大供應以滿足不斷增長的需求，但許多AI公司高層預計市場的GPU短缺至少會持續到明年。據悉，有一些零售商甚至以超過3倍的價格（約3.3萬美元）兜售英偉達GPU，而在二級市場的價格還會更高。

AI搜索引擎公司Perplexity AI行政總裁斯里尼瓦桑稱：「即使你已經預付了費用，也不意味着GPU會在第二天或一周內就送上門來。能做的只有等待。」有服務器製造商表示，他們需要等待6個月以上，才能獲得英偉達的最新GPU。

最大的服務器製造商之一超微公司表示，該公司正加足馬力研發AI產品，急需大量GPU系統。一些初創公司目前只好轉用甲骨文等不太受歡迎的雲供應商，希望盡快等到芯片短缺問題解決。

「人和狗都在搶GPU」

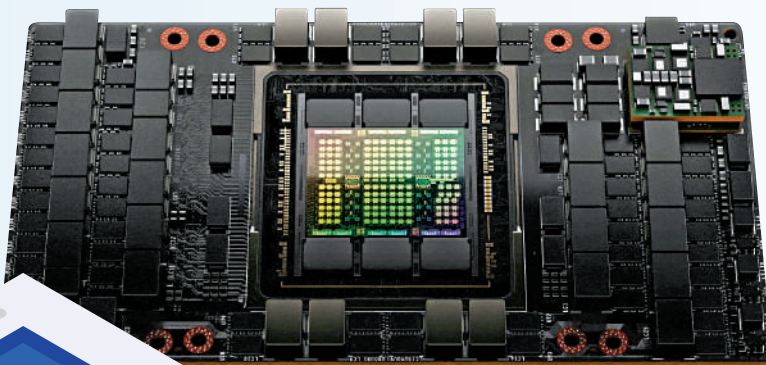
千金難求GPU。馬斯克上周直言，「GPU比毒品更難獲得」。這讓人聯想到馬斯克3月聯同上萬名科技專家簽署公開信，要求暫停研究AI半年，以建立更完善的監管機制。但很快就被媒體踢爆，馬斯克已經暗中購買了近1萬個GPU，為自家AI公司X.AI力撼OpenAI做準備。

當被問及相關問題時，馬斯克直接承認，稱包括特斯拉和Twitter在內的所有人都在購買GPU用於計算和發展AI，「所有人和他們的狗都在購買GPU，Twitter和特斯拉肯定也在買」。

此外，英偉達行政總裁黃仁勳日前向CPU（中央處理器）行業發起挑戰。他宣布傳統的摩爾定律已經過時，未來的性能提升將主要來自生成式AI。他表示，訓練一個需960個CPU集群的大型語言模型需花費約1000萬美元，並消耗11千兆瓦時的電力。相比之下，如果成本不變，全部換成GPU集群，可以訓練44個LLM。這意味着，用GPU訓練大型語言模型，可使成本降低96%。



▲黃仁勳（左）2016年8月向OpenAI親手交付世界上第一台單機箱深度學習超級電腦，馬斯克（右）當時為OpenAI聯合創辦人。 網絡圖片



▲英偉達名為H100的GPU，可以用於大型AI模型的開發。 路透社

什麼是GPU？

話你知

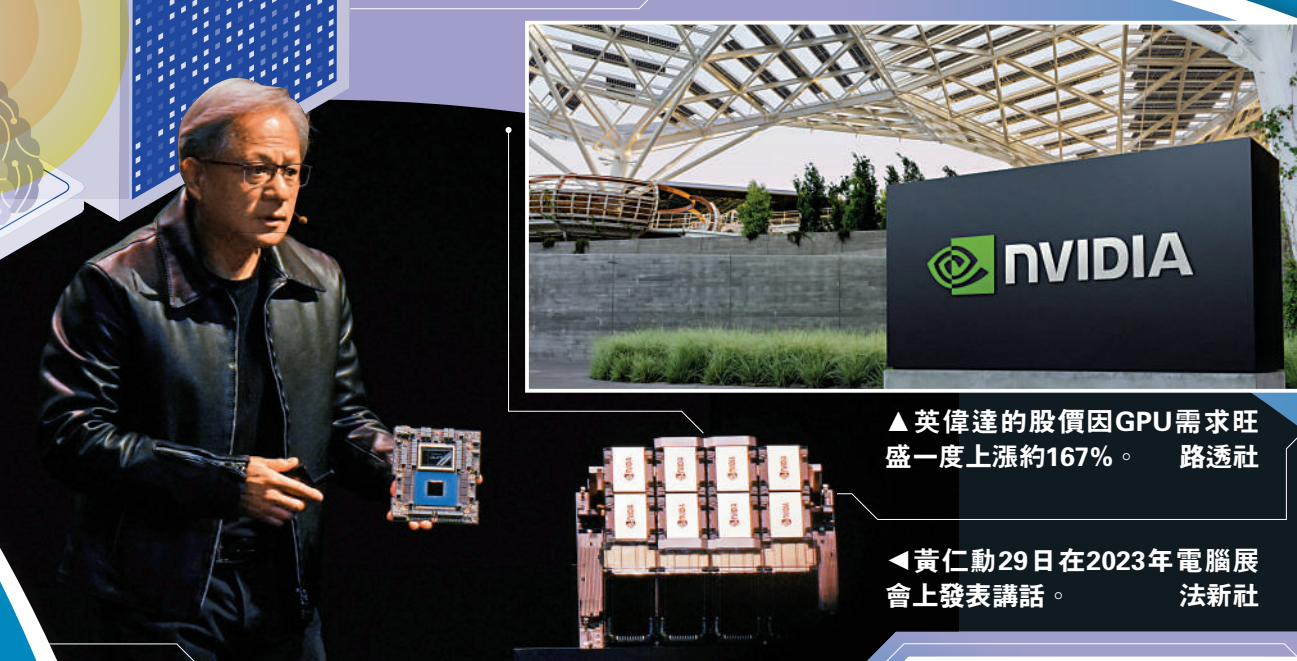
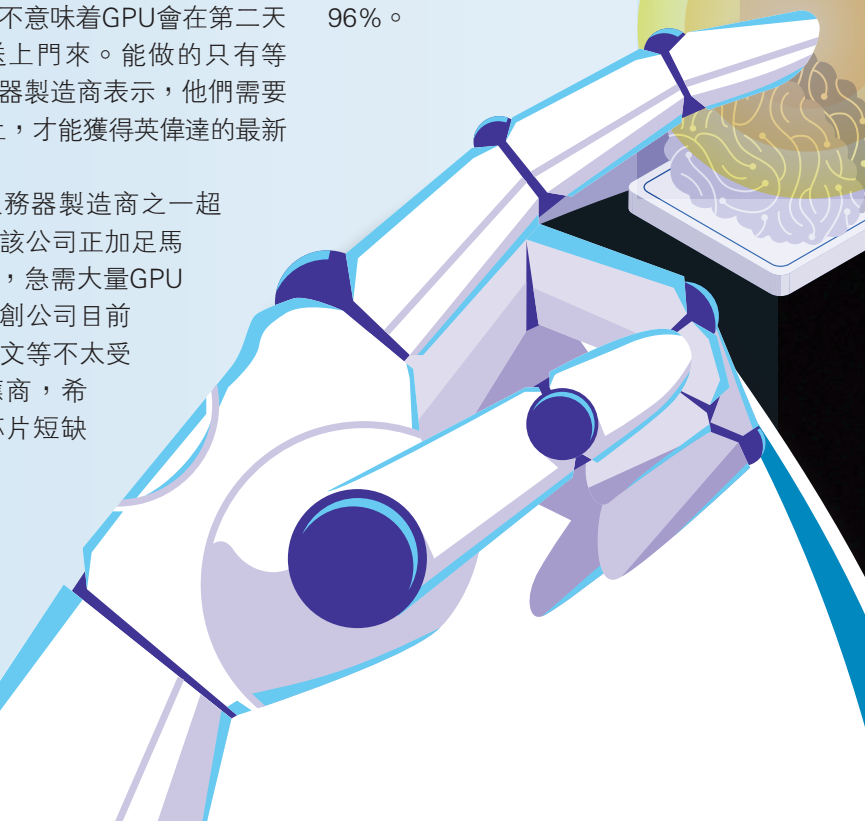
● 圖形處理器（Graphics Processing Unit, GPU）又稱圖形芯片，是一種主要負責加速圖形處理速度的微處理器。GPU最初設計用於加速3D圖形的渲染，使用更簡單、更專業的並行結構，執行任務更快。

● 如今，GPU已不再局限於3D圖形處理了，其通用計算技術得到了發展，在浮點運算、並行計算等方面，GPU的性能是CPU（中央處理器）的數十倍乃至上百倍。GPU現也被大量用於高性能計算（HPC）、人工智能（AI）開發等領域。

● 英偉達公司幾乎壟斷目前最炙手可熱的AI GPU市場，市值直逼1萬億美元。

● 大摩日前預測，英偉達將在今年的AI產品市場中獲得60%的份額，主要來自於其GPU和網絡互連產品。

大公報整理



▲英偉達的股價因GPU需求旺盛一度上漲約167%。 路透社

▲黃仁勳29日在2023年電腦展會上發表講話。 法新社

▶ 多個英國國民保健服務基金會未經病人同意，向Facebook披露病人的隱私。 路透社



英國NHS擅自向Fb披露病人隱私

【大公報訊】據《衛報》報道：英國媒體報道稱，多個英國國民保健服務基金會（NHS Trust）在未經病人同意下，向社交平台Facebook（Fb）披露病人的私密資料，如預約門診、藥物處方及網上諮詢等，估計受影響病人多達數百萬。

英國《觀察家報》調查發現，20個NHS基金會網站有隱藏的追蹤工具，多年來收集英國病人的個人信息，並在未經病人同意的情況下持續與Fb分享他們的隱私。據悉，數據內容包括病人曾經瀏覽的頁面、點擊的按鈕及搜索的關鍵字等，網站同時還會提供病人的IP地址及Fb賬戶。這20個NHS基金會為超過2200萬英格蘭居民提供服務。

上周末，17間使用Fb母企Meta網絡工具Pixel的NHS基金會表示，已從網站移除

追蹤工具，其中8間向患者表示道歉。多間NHS基金會表示，使用Pixel工具的本意是監察慈善活動及招聘過程，沒有意識到會洩露病人的隱私。

英國法律規定，未經明確同意洩露「關於個人過去、現在或未來健康狀況」的信息屬非法行為。據悉，英國的資訊委員會辦公室據報已着手調查事件。數據隱私專家克里斯特表示，監管機構早該制止這些行為，他批評Meta漠視信息安全，「Meta聲稱他們不允許網站發生某些類型的數據，但他們沒有盡心審核這些數據。」

法網擬用AI過濾惡評 保護選手

【大公報訊】據BBC報道：為防止網友的偏激言論給球員造成心理負擔，剛剛開賽的法國網球公開賽（法網）採用人工智能（AI）防護機制，為選手開放社媒監控工具Bodyguard.ai的權限，以避免或減少網絡霸凌。

Bodyguard.ai旨在過濾Facebook、Instagram、Twitter等社媒平台上的辱罵性評論。用戶只需掃描二維碼將自己的社交媒體賬戶與該系統連接，AI就會在200毫秒內分析每條實時評論並過濾辱罵性言論。法網是四大滿貫賽事中第一個引入該系統的賽事，法網聯合會行政總裁弗萊西耶表示，球員的心理健康是「法國網球公開賽的首要任務」。

根據世界體育網絡進行的一項研究，Twitter和Facebook上的相關評論約有15%是針對塞爾維亞網球運動員祖高域的負面評

論。與男球員相比，女子網球選手遭受網絡霸凌比例更高。前網球奧運冠軍佩芝本月呼籲國際女子網球協會（WTA）採取更多措施保護球員。WTA發言人表示，計劃與一家風險評估管理公司合作，在必要時關閉社交媒體賬戶。



▲法網採用AI過濾針對選手的社媒評論，圖為俄羅斯網球運動員盧布列夫。 路透社

ChatGPT杜撰案例 美資深律師上當

【大公報訊】據CNN報道：紐約一名律師使用AI聊天機器人ChatGPT準備法庭文件，結果發現其援引的多宗案例以及文章內容都是假的。因為沒有先例，該律師具體面臨什麼懲罰尚不清楚，恐遭停牌處分。

涉事律師是在紐約已經執業30年的施瓦茨，他代表的當事人馬塔起訴哥倫比亞航空，稱他於2019年8月搭航班時被推車撞傷膝蓋，隨後哥航以訴訟時效已過要求曼哈頓聯邦法院撤案。施瓦茨於今年3月提交一份長達10頁的摘要，引用了起碼6宗類似裁決成功的案例，而這一切最終被發現都是由ChatGPT偽

造。施瓦茨本月25日在宣誓書中承認，曾依賴ChatGPT搜索資料，結果證明來源不可靠。他表示「非常後悔使用生成式AI來補充法律研究」，他強調自己無意欺騙法庭或航空公司，又稱以前從未這樣做過。

據悉，施瓦茨詢問ChatGPT案例的真實性，ChatGPT再三確定這些案例為真實，且在「可信的法律數據庫」中可以查找到。

法官形容，該情況「史無前例」，已下令於6月8日舉行聽證會進行討論。施瓦茨目前面臨何種懲罰尚未明確，最嚴重的情形可能是被停牌。