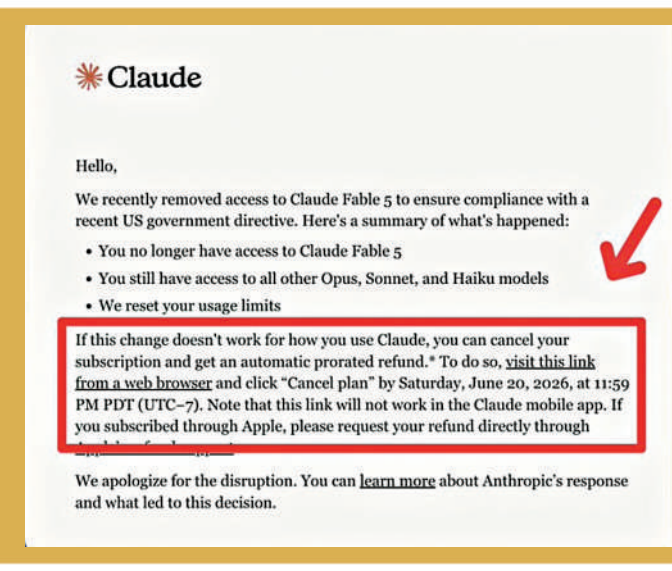
美科技巨頭暗中角力 Anthropic上市蒙陰影

亞馬遜CEO舉報 促成白宮限AI模型出口

美媒14日爆料指，美國政府日前對人工智能（AI）初創公司Anthropic的兩款新模型Mythos 5及Fable 5實施出口管制，這一決定的導火索是亞馬遜CEO賈西的一通電話。據報道，賈西11日晚致電美國財政部長貝森特等官員，稱其研究人員通過特定提示詞繞過了Fable 5的安全「防護欄」，獲取了可用於網絡攻擊的敏感信息。白宮隨即召開緊急會議，在施壓Anthropic自主撤回模型失敗後，美國商務部祭出管制措施。業界認為，此次管制不僅令正在籌備上市的Anthropic前景蒙上陰影，也折射出美國科技巨頭間的暗中角力。



▲Anthropic下架AI模型後提供退款通道，有網友發布了收到的退款郵件。 網絡圖片

【大公報訊】據報道，亞馬遜研究人員在測試Fable 5模型時發現，通過一系列特定提示詞，可以繞過Fable 5的安全「防護欄」，取得能夠用於網絡攻擊的相關敏感信息。亞馬遜的報告顯示，在特定提示詞引導下，Fable 5模型能夠挖掘出至少4款軟件程序中的安全漏洞，而這類信息通常受模型安全機制屏蔽。

白宮施壓Anthropic自主下架失敗

賈西於11日晚間致電貝森特及其他政府高層官員反映這一問題，到12日上午，該問題已上報至白宮最高層。知情人士透露，貝森特、美國國家網絡總監凱恩克羅斯、白宮辦公廳主任威爾斯及其他高級官員舉行會議，討論應對措施。

Anthropic行政總裁阿莫戴與包括貝森特、美國商務部長盧特尼克等在內的政府官員於12日進行了數小時談判。會談中，阿莫戴解釋稱，亞馬遜研究人員提到的功能只能識別出少數先前已知的軟件漏洞，而且其他多款公開可用的AI模型也有類似用法，每天都被網絡安全防禦者用於系統維護。另有一名知情人士透露，Anthropic在發布這兩款新模型之前，一直與美國政府「保持合作」，也告知6月9日的發布日期，而美政府並未反對。

但凱恩克羅斯和貝森特並未被阿莫戴的論點說服。一名白宮官員表示，他們已將亞馬遜的調查結果提交給美國國家安全局（NSA）審閱，並認為他們掌握了確鑿的「證據」。

據報道，白宮方面敦促Anthropic主動下架模型，但阿莫戴則請求給予更多時間和信息，且並未做出承諾。Politico引述一位白宮高級官員報道，「我們苦苦哀求他們（Anthropic）幾個小時後才採取的最後手段。」但一位與Anthropic公司關係密切的人士反駁說，官員們未曾表示要合作的態度。美國政府在12日下午1點15分致電Anthropic，只給約90分鐘時間要求下架最新模型，理由是「未具體說明的國安風險」，下午5點左右正式發出出口管制命令。當天晚上10點左右，全球用戶無法使用Mythos 5及Fable 5。

Anthropic據報最快計劃今年秋季上市，新款頂尖模型被迫下架可能導致用戶流向其他平台，並對公司估值造成壓力。

特朗普政府加強AI模型監管

外界認為，此次出口管制不僅嚴重打擊Anthropic，也凸顯特朗普政府在AI監管方面的大幅「轉型」。美國總統特朗普重返白宮後，美國政府一度對AI行業整體採取較為寬鬆的監管立場，並撤銷拜登政府時期要求企業對高風險模型進行安全評估並上報的相關措施。但這段時間以來，美國政府在AI監管方面的政策從「放鬆」趨向「收緊」。6月初，特朗普簽署行政命令，要求Anthropic、OpenAI、谷歌等AI開發商對外公布AI新模型30天前，先行向聯邦政府開放訪問權限，自願將模型遞交政府進行網絡安全審查。

另外，作為Anthropic最大的投資方和雲服務提供商，亞馬遜此次「舉報」也揭示AI企業與美政府間的密切聯繫。亞馬遜方面表示，CEO賈西致電政府高層的本意是回應特朗普政府要求提供反饋的請求。亞馬遜發言人稱，作為服務大量公私客戶的領先雲服務提供商，政府就潛在安全風險徵詢其意見並不罕見，並稱此類情況下公司不會披露相關討論細節。 （綜合報道）

大公報設計圖片

▼Anthropic兩款先進AI模型下架。 路透社

Mythos 5和Fable 5是什麼？

- Fable 5和Mythos 5是Anthropic於9日推出的AI大語言模型，核心優勢為超強的漏洞識別與邏輯推理能力。兩者共享完全相同的底層架構，Fable 5設置了安全「防護欄」，面向大眾開放，一旦被問及相關高級別敏感問題，系統會自動將對話轉至其他能力偏低的模型繼續處理；Mythos 5能力全面強大，僅限美國政府以及少數經過嚴格審核的頂尖安全防禦基礎設施提供商使用。

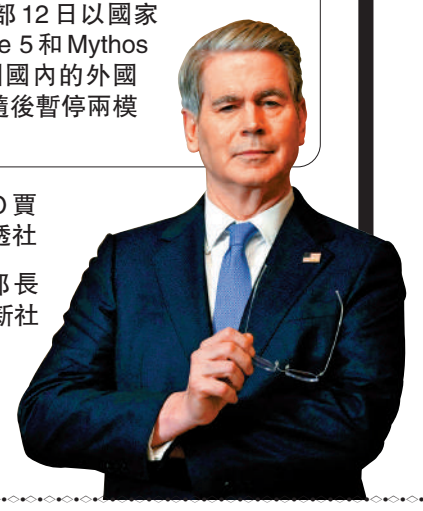
爲何要設置「防護欄」？

- Fable 5和Mythos 5都屬於Mythos系列。Anthropic於今年4月首次發布的Claude Mythos系統，在網絡安全方面能力表現突出，幾周之內就自主識別了數千個「零日漏洞」（即軟件開發者此前未知的安全缺陷），但它同時也能自主生成利用這些漏洞的攻擊方法，從而變成大規模黑客攻擊的工具。Mythos的能力引起美國銀行業警戒，美國財長貝森特與華爾街主要銀行CEO在模型發布後不久召開會議，討論其風險。Anthropic為避免Mythos系列被濫用，為其設定「防護欄」，發布了Fable 5。

特朗普政府爲何禁用？

- 亞馬遜CEO賈西11日致電美國官員，稱研究人員通過特定提示詞繞過了Fable 5的「防護欄」，獲取了可用於網絡攻擊的敏感知識信息。美國商務部12日以國家安全為由，禁止Fable 5和Mythos 5出口，並禁止美國國內的外國人使用。Anthropic隨後暫停兩模型的所有用戶訪問。


◀亞馬遜CEO賈西。 路透社


▶美國財政部長貝森特。 法新社

美國政府收緊AI監管

限制AI模型出口

- 美國政府12日以國家安全為由，要求人工智能（AI）模型Claude開發商Anthropic，禁止所有外國公民使用其最新的Fable 5和Mythos 5模型。Anthropic已對所有用戶下架這兩款最新模型。

審查AI新模型

- 美國總統特朗普2日簽署行政命令，要求Anthropic、OpenAI、谷歌等AI開發商對外公布AI新模型30天前，先行向聯邦政府開放訪問權限，自願將模型遞交政府進行網絡安全審查。

停止發布公開報告

- 據知情人士透露，包括國家網絡總監凱恩克羅斯在內的政府官員已告知「人工智能標準與創新中心」（CAISI）暫停發布其模型評估報告。CAISI隸屬於商務部，是政府內部負責在AI模型發布前進行測試並向公眾披露其能力及相對表現的主要機構。

Anthropic與美政府「邊打邊合作」

【大公報訊】綜合美國Axios新聞網、《華爾街日報》報道：美國政府近日以國家安全為由，要求人工智能（AI）初創公司Anthropic暫停Fable 5與Mythos 5兩款模型的境外訪問權限，Anthropic隨後表示已關閉全體用戶的兩款新模型訪問權限，但同時公開反對這項決定。


▲Anthropic行政總裁阿莫戴。 美聯社

事實上，這並非Anthropic首次與美國政府爆發衝突，二者可謂「積怨已久」。

今年3月，美國國防部要求無限制地將Anthropic的Claude模型用於「所有合法用途」，包括完全自主的致命武器系統或大規模國內監控，遭到CEO阿莫戴的拒絕。美國國防部隨後宣布把Anthropic定為「供應鏈風險」對象，禁止國防部以及防務承包商在相關工程中使用該公司的AI技術。Anthropic因此控告美國國防部長海格塞斯及國防部，希望撤銷相關認定，目前案件仍在審理當中。

不過，美國政府並沒有因此停止與Anthropic合作。英國《金融時報》本月初報道，Anthropic正協助美國國家安全局（NSA）部署其旗下的強大Mythos AI模型，進行進攻性網絡行動。Anthropic據報已向NSA派遣6名「前置部署工程師」，負責指導Mythos的使用，並制定特定應用模型。

2月28日，包括伊朗最高領袖哈梅內伊在內的多名伊朗高層被美、以鎖定並打死。據《華爾街日報》報道，美軍在本次行動中使用了Claude，包括進行情報評估、目標識別和模擬戰鬥場景等關鍵任務。

瑞士公投料否決「人口封頂」提案

【大公報訊】綜合BBC、路透社報道：當地時間14日，瑞士就「人口封頂」提案進行公投。據報該提案由瑞士最大政黨、主張嚴控移民的瑞士人民黨提出，主張限制人口增長，避免全國人口突破1000萬人。根據初步開票結果，過半民眾傾向否決這項公投案。

提案要求立法規定，在2050年前，包括瑞士公民和合法居留外國人在內的瑞士常住人口數量不得超過1000萬；如果人口在2050年前達到950萬，政府應採取措施限制人口增長，例如限制庇護、家庭團聚、居留許可等證件的簽發，並就相關國際協議進行重新談判；一旦人口達到1000萬，政府就要終止與歐盟之間的人員自由流動協議。

瑞士廣播電視台（SRF）14日發布的公投結果初步推估顯示，約45%選民支持這項公投，但有55%反對。

自2002年以來，瑞士的人口迅速增長，從當時的730萬，增長到目前的910萬，但未持有瑞士護照的常住外國公民佔比約27%，其中多數來自歐盟國家。人民黨認為，移民前往瑞士，導致了對醫院、學校的需求不斷增加，限制移民將減輕壓力。

◀瑞士14日就「人口封頂」提案進行公投。圖為瑞士街頭支持限制人口的廣告牌。 路透社

一些選民認為，這是人民黨最新的反移民舉措。政府、其他政黨、商業領袖和工會則將其稱為「混亂倡議」，認為這會影響醫院和酒店急需的員工數量，破壞瑞士和歐盟來之不易的關係。在瑞士酒店工作的人中，有一半是移民。醫院和養老院也依賴外國工人。瑞士企業聯合會首席經濟學家魯道夫·明施表示，如果公投通過，瑞士「在與歐盟的關係中可能面臨挑戰」。

英稱首次攔截俄油輪 被指轉移國內政治焦點

【大公報訊】綜合新華社、路透社報道：英國首相斯塔默14日表示，英國武裝部隊當天凌晨在英吉利海峽攔截一艘被制裁的俄羅斯「影子艦隊」油輪。英國國防部透露，這是英國首次主導針對此類船隻的攔截行動。

英國國防部當天發表聲明說，在一次持續約6小時的行動中，英國軍方人員登上懸掛喀麥隆國旗的油輪「斯邁托斯」號（Smyrtos）。巡邏機、直升機、軍艦等在行動中提供支援。據悉涉事油輪已被臨時轉移至英格蘭南部海岸附近一處錨地，並在調查期間接受監控。斯塔默在社交媒體上說，此舉「再次打擊了俄羅斯」。英國國防部透露，此次行動是在與法國密切協調下實施的。

俄羅斯直接投資基金總裁、俄羅斯總統負責對外投資與經濟合作事務的特別代表基里爾·德米特里耶夫表示，斯塔默不去攔截非法移民，反而在英吉利海峽扣押一艘油輪，試圖藉此轉移英國民眾的注意力。

法國總統馬克龍6月1日在社交媒體發文說，法國海軍5月31日在大西洋海域攔截了一艘從俄羅斯駛出的油輪，行動獲得英國等多個夥伴國家配合。俄羅斯總統新聞秘書佩斯科夫當天回應稱，法國軍方扣押俄羅斯油輪是非法的，這樣的行為「近乎國際海盜行徑」。

今年3月，俄羅斯外交部發表過聲明，就西方國家在海上採取的有關「非法行為」作出表態。為阻止俄羅斯獲得國際貿易收入，歐盟長期以來指責俄羅斯利用油輪等商用船隻逃避西方制裁，將這些商用船隻稱為「影子艦隊」。俄方認為，歐盟國家的這些做法不可接受，俄方將使用一切可用的政治、法律及其他手段，確保航行自由原則得到尊重。

▲英國軍方人員14日凌晨從直升機上速降到油輪「斯邁托斯」號（Smyrtos）。 網絡圖片