



美政府審查加碼 限制 OpenAI 新模型發布

僅供少數客戶試用 白宮逐個批准訪問權限

美媒25日披露，美國政府打着「國家安全」旗號進一步加強對人工智能（AI）模型的審查和管控力度，要求OpenAI公司分階段發布其最新模型GPT-5.6，首階段僅能在少數「得到政府批准」的合作夥伴範圍內發布，而且當局要逐一審批客戶的訪問權限。這是美國政府首次要求AI公司限量發布新模型。此前，美國政府同樣以國家安全為由，要求另一家AI公司Anthropic限制外國公民使用其兩款最新AI模型。美媒稱，此舉標誌白宮對AI的監管立場發生改變。

【大公報訊】美媒25日援引消息人士報道稱，白宮已要求OpenAI分階段發布最新款AI模型GPT-5.6，先在少數得到美國政府批准的合作夥伴範圍內發布，而不是直接面向公眾發布。OpenAI行政總裁阿爾特曼25日在一份備忘錄中告訴員工，GPT-5.6首先以預覽版形式推出，美國政府將「逐一批准客戶的訪問權限」。《金融時報》報道，OpenAI擬先向20多家合作夥伴開放GPT-5.6的使用許可權。報道指出，這是美國政府首次要求AI公司限量發布新模型，表明其對AI行業的管控力度進一步升級。

要求科企「自願」接受審查

關於GPT-5.6的詳細資訊尚未公開。消息人士透露，美國政府認為該模型的性能與早前被當局納入出口管制的Anthropic最新模型不相上下，這些前沿模型「功能強大」，因此白宮希望確保相關公司採取「充分的安全保障措施」。

據悉，OpenAI模型被限制發布，是該公司與隸屬白宮的國家網絡總監辦公室和科學技術政策辦公室溝通後的結果。此前，阿爾特曼24日與美國商務部長盧特尼克會面，後者希望確保GPT-5.6得到美國政府所有相關部門的測試和批准。美媒稱，阿爾特曼要求員工配合，但同時表示這並非公司認可的長期模式，但要為未來模型發布制定可持續方案。阿爾特曼說，希望能夠在幾周後發布GPT-5.6。

根據美國總統特朗普本月初簽署有關AI的行政令，當局要求科企在發布產品前30天，「自願」向政府提交模型以供安全審查，以評估產品可能對金融系統、國安系統和其他關鍵系統所構成的威脅。

6月12日，Anthropic首當其衝，美國政府以「國家安全」為由發布出口管制指令，要求該公司暫停所有外國公民對其最新的Fable 5和Mythos 5兩款模型的訪問權限。Anthropic隨後宣布在全球範圍下架這兩款發布僅數天的產品。Anthropic表示，事件的導火索是Fable 5被發現存在安全漏洞。路透社透露，盧特尼克給出的理由是「擔心」這些模型被中俄等所謂美國「關切國家」的軍事情報人員利用。

Anthropic在聲明中質疑美國政府限制其

模型的做法，稱僅因發現一個「潛在漏洞」就將一款面向數億用戶的商業模型下架，認為這一標準根本「站不住腳」。

監管框架模糊不清

目前美國主要AI技術開發商中，OpenAI、Anthropic、谷歌DeepMind、微軟及xAI等科企均已同意配合美國政府對AI模型的審查措施，只有Meta尚未與當局達成協議。《紐約時報》23日披露，特朗普政府正在對Meta施壓，要求其「自願」提交AI模型接受審查。Meta表示，雖然「細節仍待敲定」，但該公司「希望盡快簽署協議」。

美媒指出，特朗普政府的AI監管仍缺乏透明、統一的框架，現階段科企不清楚應該向政府的哪個部門匯報、何時匯報、距模型發布多長時間前匯報，也不確定應該分享多少細節。此前針對Anthropic的出口管制指令來自美國商務部，而OpenAI接到的命令卻涉及不同機構，這令各大AI公司感到困惑。美國跨黨派超級政治行動委員會「公共優先」負責人卡森表示，特朗普政府目前採用的一種「臨時性的、不透明的甚至可能違法」的做法。

美國強化AI模型審查

頒布行政令

●本月2日，美國總統特朗普簽署關於強化AI模型審查的行政令，計劃建立一個「自願審查」框架，允許美國政府在特定AI模型發布前至多30天獲取其訪問權限。

向Meta施壓

●美媒23日披露，美國政府正在向Meta公司施壓，要求其「自願」提交AI模型接受審查。報道指出，Anthropic、OpenAI、谷歌DeepMind、微軟和xAI等美國科企均已同意接受審查。

對Anthropic新模型設限

●本月12日，美國政府以國家安全為由，要求Anthropic公司限制外國公民使用其AI模型Fable 5和Mythos 5。該公司隨後暫停這兩款模型的所

要求OpenAI分階段發布新模型

●知情人士透露，美國政府要求OpenAI公司分階段發布其最新模型GPT-5.6，第一階段僅向少數經美國政府批准的合作夥伴發布該模型。

大公報整理



▲特朗普（右）和OpenAI行政總裁阿爾特曼17日共同出席一場會議。

美聯社



◀OpenAI旗下的模型面臨美國政府嚴格審查。

法新社



大公報AI製圖



▲Anthropic在其網站上宣布關閉兩款新模型訪問權限。

網絡圖片

OpenAI 被曝考慮延後IPO至明年

【大公報訊】據《紐約時報》報道：美國科企近期面臨特朗普政府加強審查、市場行情起伏不定等挑戰。美媒25日援引3名知情人士披露，美國人工智能（AI）巨頭OpenAI考慮將首次公開招股（IPO）時間推遲至明年。

OpenAI於本月8日宣布，已秘密遞交IPO申請。知情人士稱，OpenAI原計劃最早於今年第三或第四季度上市，但為了實現1萬億美元估值目標，該公司目前傾向於將IPO時間推遲至明年。據悉，為OpenAI提供建議的銀行家警告

說，近期科技股的波動，尤其是太空探索技術公司（SpaceX）創紀錄IPO之後股價劇烈波動，可能削弱投資者對OpenAI的熱情。該公司顧問團向公司高層提出兩種方案，一是等到明年再以1萬億美元估值上市，二是降低估值目標並更早上市。OpenAI行政總裁阿爾特曼對此表示，降低估值目標並不在考慮範圍內。

另據美媒報道，計劃最早今年秋季上市的Anthropic公司也面臨挑戰。特朗普政府本月12日對其兩款新模型發布出口管制指令，迫使其暫時關閉相關模型訪問權限。此前，特朗普政府還曾以國家安全為由將該公司列入黑名單。美媒指出，Anthropic與特朗普政府的矛盾令投資者感到擔憂，勢必對其估值和未來股價造成負面影響。

▲OpenAI近期面臨多項挑戰。圖為該公司位於美國加州的總部。



金建希涉「賣官鬻爵」案 一審判囚7年

【大公報訊】據韓聯社報道：韓國首爾中央地方法院26日就前總統尹錫悅夫人金建希涉嫌「賣官鬻爵」一案作出一審判決，判處其有期徒刑7年。金建希律師團隊隨即表示，將對一審判決結果提起上訴。

法院認定，金建希在尹錫悅2022年3月當選總統後，收受多人贈送的項鍊、手錶、手提包等總價高達約3億韓圓（約153萬港元）的物品，為對方提供職位職務和業務便利，違反《特定犯罪加重處罰法》，據此判處其有期徒刑7年。此外，法院還對多名行賄者分別處以緩刑及罰款等。

今年4月28日，法院已宣布金建希操縱股價和斡旋受賄等罪名成立，二審判處有期徒刑4年並處罰金。她還因牽涉「統一教」信徒集體加入韓國國民力量黨一案被起訴，該案將於8月14日首次開庭。



▲韓前總統尹錫悅夫人金建希（中）26日在首爾出庭。

網絡圖片

2024年12月3日，時任韓國總統尹錫悅發布緊急戒嚴令，韓國國會數小時後通過要求解除戒嚴令的決議案。此後，尹錫悅先後被停職、逮捕、彈劾，他與金建希均面臨多宗訴訟。今年2月，首爾中央地方法院宣布尹錫悅內亂頭目罪名成立，一審被判無期徒刑。4月29日，尹錫悅在其涉嫌妨礙逮捕案二審中被判有期徒刑7年。本月12日，尹錫悅因「製造平壤上空無人機入侵事件」一審被判有期徒刑30年。

英白金漢宮耗資38億翻新 查爾斯不入住

【大公報訊】據《衛報》報道：英國白金漢宮25日首度公開國王查爾斯三世個人稅務資料，證實其2024至2025稅務年度自願繳納了1290萬英鎊（約1.3億港元）的個人所得稅和資本利得稅。與此同時，英國王室宣布，白金漢宮明年完成耗資3.69億英鎊（約38.2億港元）的大規模翻修後，查爾斯和王后卡米拉不會回到這裏居住，以便讓更多公眾有機會參觀這處建築。

25日公開的稅務資料顯示，查爾斯2024至2025稅務年度繳納了1290萬英鎊稅款，自2022年9月登基以來共繳納逾3000萬英鎊（約3.1億港元）稅款；他的長子威廉王子2024至2025稅務年度繳納了776萬英鎊稅款，自父親登基以來共納稅逾2000萬英鎊。英媒指出，這是英國在位君主首次公開個人的納稅細節。白金漢宮稱，這是王室履行提高財務透明度承諾的一部分。不過，查爾斯並未公布其應納



▲遊客26日在白金漢宮外拍照。

路透社

稅收入等資料。

白金漢宮2017年開始翻修，預計明年完工。但英國王室25日宣布，查爾斯和卡米拉決定繼續留在白金漢宮附近的克拉倫斯宮居住。

英國王室司庫查默斯稱，白金漢宮可以舉辦更多活動，給公眾更多機會參觀。白金漢宮始建於1705年，1837年起成為英國君主的寢宮。