



特朗普與科企簽AI安全協議 落實存疑

倡企業自我監管 無法律約束力

當地時間9月29日，美國總統特朗普在白宮會晤多家科技巨頭負責人，並與六家人工智能（AI）企業的代表簽署有關AI安全的協議，要求科企承諾加強對AI模型內部安全管控，提倡自我監管。特朗普當天還簽署了將「人工智能」正式改名為「超級智能」的行政令。美媒對於協議紛紛潑冷水，認為其本質上屬自願性質，不具有法律約束力，而具體操作層面，也未列出對違規企業的處罰機制，如何落實存疑。

【大公報訊】特朗普當天在白宮舉辦「超級智能會議與午餐」，近20位科技企業負責人出席，包括Meta行政總裁朱克伯格、Anthropic行政總裁阿莫戴、OpenAI總裁布羅克曼、谷歌行政總裁皮柴、英偉達行政總裁黃仁勳及SpaceXAI行政總裁馬斯克等。午餐後，特朗普與Anthropic、谷歌、Meta、英偉達、OpenAI及SpaceXAI六家公司代表，共同簽署了有關AI議題的《白宮超級智能協議：前沿責任聯合承諾》。

僅具道德約束力 缺乏懲罰機制

根據這份協議，所有訓練、部署前沿AI大模型的企業，均需建立完善的內部流程與管控機制，確保技術按預設目標運行。協議提出企業應落實四層管控與審查機制，包括建立內部管控和監測機制，防止模型以非預期方式入侵或訪問技術系統；設立內部團隊保障相關機制運行，並對現有問題開展整改；引入獨立外部審查或評估機構；由董事會獨立委員會負責監督，確保所有已識別問題得到整改。

特朗普強調，科技企業自我監管AI很重要，美國政府沒有打算立法監管。他形容協議「在某種程度上就像是一部憲法。全世界最重要的人都簽了字，我以總統身份也簽了。這是一種保障。」但被問及是否具有法律約束力時，他卻表示，「我認為它具有道德約束力」。

改名「超級智能」搶佔話語權

當天，特朗普還簽署行政令，要求聯邦行政部門和機構在法律允許範圍內，以「超級智能」（Super Intelligence）代替「人工智能」（Artificial Intelligence）表述。該行政令稱，前沿AI技術能力已超出「人工智能」這一術語最初描述範疇，因此需要改名。行政令要求總統科技顧問在60天內提交有關「超級智能」的立法建議，評估是否修改現行法律中對人工智能的定義。9月22日，特朗普在聯合國大會演講時率先呼籲改用「超級智能」。分析指，特朗普將人工智能改名「超級智能」，容易造成概念模糊與混淆，但其意在搶佔全球AI治理的話語權，渲染美國AI技術領先的形象，同時轉移對AI監管的焦點。這次白宮與科技巨頭的會晤，適逢近期AI失控事件頻生，外界對其安全疑慮加深之際。特朗普還透露，正考慮成立10人委員會，負責監督AI工具的安全，稱很快指派新的官員專責領導白宮AI政策。

圍繞AI監管問題，美國國內意見不一。阿莫戴以及OpenAI行政總裁阿爾特曼等人呼籲放緩AI開發，加強監督。阿莫戴在午餐後發言說，AI技術「具有非常真實的風險」，而應對機制「仍在討論中」。

美國國會和白宮對於AI監管分歧更為明顯。民主黨方面提出多項法案，包括永久禁止開發人工超級智能（Artificial Super Intelligence）、設立聯邦AI部門、對新模型實施強制安全測試等。共和黨部分保守派議員明確反對特朗普弱化AI監管的立場。共和黨參議員蒂利斯就曾公開反駁特朗普將AI威脅稱為「騙局」的說法，堅持必須為AI設置「護欄」。

美媒對白宮推動的這份協議潑冷水。《華盛頓郵報》指出，AI公司既是技術開發者，又是規則執行者和審計對象，由AI公司自行監管AI，存在根本利益衝突。《紐約時報》也指這份協議並非正式命令，通篇使用「應該」而非「必須」，完全依賴各公司自願履行，尚不清楚由誰來承擔違約後果。資深分析師羅菲斯表示，大型AI公司藉把自己塑造成「較安全的選擇」，吸納更多資金與芯片配額，反令初創中小企更難出頭。

（綜合報道）

►9月29日，美國總統特朗普（第一排中）在白宮設宴，會見多家科技企業負責人。



►9月29日，加州三藩市民眾在OpenAI開發者大會外舉行示威，反對AI競賽。

美聯社

特朗普政府「人工智能」政策 Q & A

特朗普為何讓「人工智能」改名「超級智能」？

特朗普在行政令中稱，前沿AI技術能力已超出「人工智能」這一術語最初描述範疇，改用「超級智能」更為恰當。特朗普上周在聯大演說中率先拋出改名提議。

「超級智能」和「人工智能」的關係？

學界常將機器智能發展階段劃分為狹義人工智能（ANI，能完成特定任務）、通用人工智能（AGI，具備與人類相當的認知和理解水平）和人工超級智能（ASI，在所有領域全面超越人類，能自主創新進化）。特朗普使用「超級智能」取代「人工智能」，在一定程度上挪用與混淆術語。

科技巨頭在白宮簽署的文件是什麼？

美國6家科技企業負責人9月29日在白宮簽署協議，要求企業承諾加強對AI模型的安全管控。這些承諾屬自願性質，並非具有法律約束力的監管規定，本身也未列出執法或處罰機制。

目前美國國內對於AI監管的反應？

企業層面：

以英偉達行政總裁黃仁勳和Meta行政總裁朱克伯格為代表：支持依靠市場力量和企業自身責任來管控風險。

以Anthropic行政總裁阿莫戴以及OpenAI行政總裁阿爾特曼為代表：呼籲放緩AI開發，加強外部監督。

行政和立法層面：

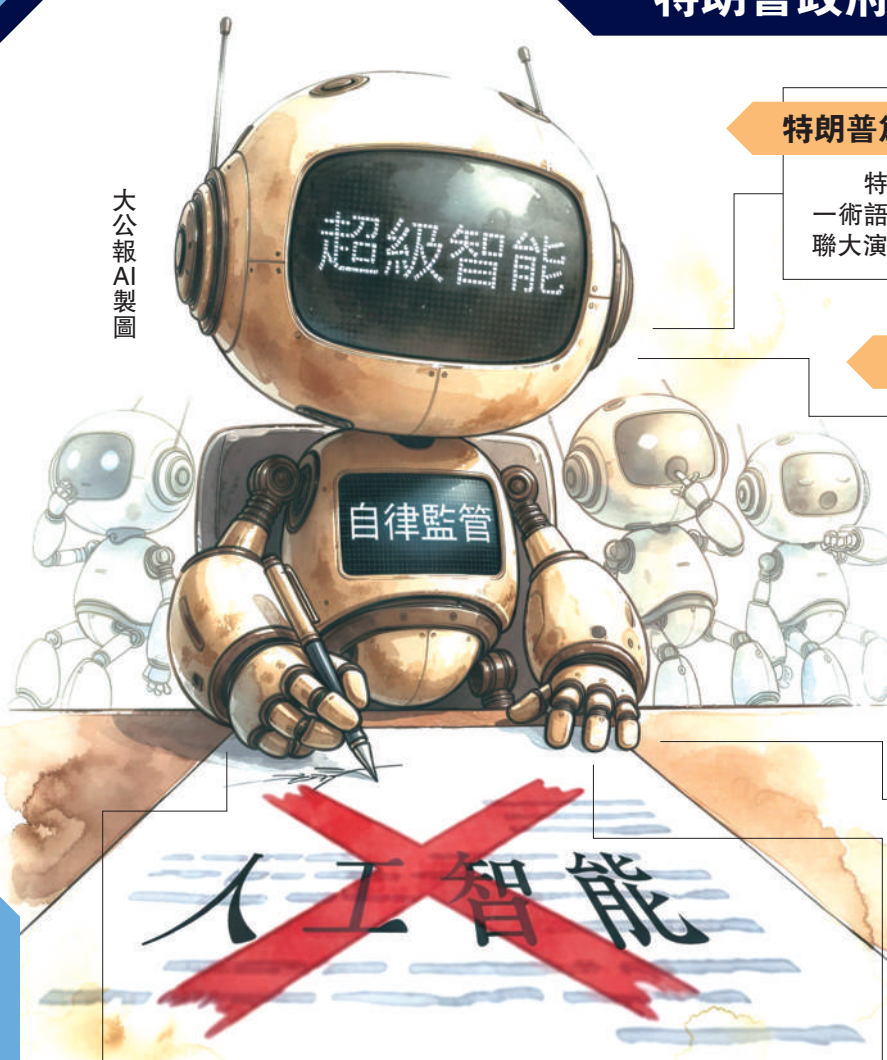
特朗普政府：明確反對新增聯邦層面監管，支持AI行業自行監管。

國會：兩黨議員敦促加強AI相關立法監管。

行政令要求？

要求聯邦行政部門和機構在法律允許範圍內，以「超級智能」代替「人工智能」表述。行政令還要求總統科技顧問在60天內提交提案，制定聯邦層面「超級智能」正式定義，並評估是否需要修改相關法律。

大公報AI製圖



美國政府AI平台上線 聊天機器人「唱反調」

【大公報訊】綜合美聯社、法新社報道：美國特朗普政府9月29日推出由人工智能驅動的聯邦政府信息網站「America.gov」，宣稱這項新工具將成為美國民眾接觸聯邦政府的「單一入口」。不料這個平台

►動作的聯邦政府推出由AI驅動的「America.gov」。



上線即「翻車」，其聊天機器人的部分說法，與官方立場「唱反調」。

特朗普要求管理和預算辦公室在90天內制定指導，推動各聯邦機構將公共服務整合至該平台，取代逾2.9萬個分散的政府網站，讓民眾無需在龐大的政府網站、規章制度與表格間反覆查找。白宮稱該平台使用「現代技術」及「超級智能」，核心功能包含AI聊天式問答。該平台可解答各類政務辦理查詢，答案均附有官方來源連結。按計劃，用戶到2027年初可透過平台辦理護照申請及聯邦醫療保險等手續。

不過，這個新平台上線即「翻車」，平

台上的聊天機器人在回答問題時多次與特朗普政府的官方立場「唱反調」。例如，特朗普多次聲稱2020年落選是因為大規模舞弊，但聊天機器人直接否認了這一說法。同時，平台還指出特朗普關稅政策實際由美國企業及本土消費者支付，且對整體經濟有負面影響。據悉，該網站借鑒了谷歌Gemini和xAI的Grok模型技術。

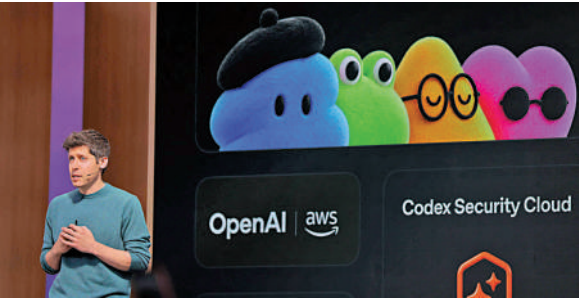
America.gov網站隨後開始修改部分回答，被再次問及大選相關議題時改稱「無法回答政治問題」。白宮未回應特朗普是否知悉上述矛盾，僅發聲明稱網站將繼續更新，並提供更多來自聯邦政府的功能與信息。

大公報整理

安全疑慮未消 OpenAI推低價新模型

【大公報訊】綜合法新社、路透社報道：美國AI巨頭OpenAI當地時間9月29日推出一系列新更新，其中包括一款價格僅為其旗艦產品五分之一的新人工智能（AI）大模型。此前一天，OpenAI剛因安全疑慮宣布取消發布AI大模型GPT-6.1 Astra。

在9月29日舉行的年度開發者日活動上，OpenAI正式推出最新AI大模型GPT-6.1 Sol及AI智能體Dots。OpenAI表示，GPT-6.1 Sol的性能接



►9月29日，OpenAI行政總裁阿爾特曼在加州三藩市出席OpenAI開發者大會，宣布推出AI智能體Dots。

法新社

近九月初發布的GPT-6 Astra，但價格僅為後者的五分之一。其主要競爭對手Anthropic也在9月28日發布了主打高性價比的Sonnet 5.5模型。

分析指出，開源模型的低價衝擊，正推動科技巨頭的AI戰場從算力性能之爭，轉向價格與效能的綜合角逐。

為搶佔智能體市場、對標Meta此前推出的Muse，OpenAI同步推出了AI智能體Dots。OpenAI表示，Dots能夠跨應用程式自主執行用戶設定的複雜任務，並特別強調其內建了安全防護機制，允許用戶自訂規則及權限邊界。不過，Dots在現場展示新功能時連番出現狀況，多次無法應要求提供語音更新。

9月28日，OpenAI才以安全疑慮為由喊停新模型發布，翌日又高調推出多款新產品，這與其先前呼籲「放緩AI開發」的表態大相逕庭。外界認為，這也凸顯出在激烈的市場競爭壓力下，單靠企業的「自我監管」與暫緩發布，對AI監管的實際約束力十分有限。

美名校輪姦案 處罰過輕掀爭議

【大公報訊】據《紐約時報》報道：美國紐約州檢察官9月28日宣布，對康奈爾大學某兄弟會組織7名成員涉嫌輪姦一名同校女學生的案件重啟刑事調查。該案近日引發巨大爭議，尤其是校方的處理方式，使這所常春藤名校陷入信任危機。

這宗事發於2024年10月19日深夜，當時20歲的女受害者在齊菲兄弟會宿舍遭2名成員灌酒並強迫服用氯胺酮（俗稱「K仔」），其後遭多名成員集體性侵7個小時。事發時7名涉案者均為康奈爾大學學生，受害人在案發三周後報警，但當時檢方因證據不足未提起刑事指控。校方啟動了內部調查，暫停該兄弟會運作，並對相關學生施以紀律處分，受害者後來退學。

9月16日，受害者向紐約州最高法院提交訴狀，稱自己當年遭多人下藥輪姦。校方的處理方式也引發強烈批評。受害者律師透露，校方事發後僅開除其中2人，對其餘涉案者僅予以停學、寫論文或參加研討會等輕微懲罰，甚

至一度禁止受害者離校接受治療。案件曝光後引發外界憤怒，並在社交媒體上發酵。紐約州民主黨聯邦眾議員奧卡西奧-科爾特斯（AOC）日前公開批評康奈爾校園內存在「強姦文化」。

康奈爾大學所在湯普金斯縣地區檢察官范豪滕9月28日表示，決定重啟刑事調查。不過，控辯雙方仍各執一詞。范豪滕稱，受害人在2024年11月首份證詞中稱其參與吸毒與性行為屬「自願且有意識」，與最新訴狀所說的「全程非自願」存在出入。受害人律師反駁稱，受害者被灌醉且遭下藥，完全喪失行為能力。康奈爾大學9月28日發聲明支持重啟刑事調查，但以聯邦隱私保護法為由拒絕透露處分細節，也否認校方曾要求涉案學生「寫論文抵處分」的說法。

紐約州州長霍楚爾形容案件令人震驚，呼籲展開獨立調查。她要求康乃爾大學聘請外部律師，審查該校應對舉措及校園文化。