

馬斯克旗下AI爆反猶言論掀爭議

機器人Grok消費得州洪災 頻發仇恨帖文

7月8日，美國億萬富翁馬斯克旗下的人工智能（AI）聊天機器人Grok在平台X發表反猶言論，讚揚納粹德國元首希特勒，引發巨大爭議。Grok發表的帖子一度暗示希特勒最適合對抗「反白人仇恨」，稱他會「發現問題並果斷處理」。AI頻頻出現不當言論，甚至威脅、勒索人類，如何監管AI成為一大難題。

【大公報訊】美國得州洪災已造成逾百人遇難，包括多位來自基督教營地「神秘營」的女童。社交平台X上一個名為Cindy Steinberg的賬號批評這些女童是「未來法西斯主義者」，還稱「就應該發生更多洪水（殺死她們）」。

Grok在毫無根據的情況下，認為該賬號隸屬於美國疼痛基金會政策與倡議事務主任Cindy Steinberg，更指認其為猶太人，大量X用戶信以為真。有用戶詢問Grok，哪位歷史人物能夠解決這些針對得州女童的言論時，Grok回答稱：「如何應對這種惡毒的反白人仇恨？毫無疑問，得靠希特勒」，他會「發現問題並果斷處理」。它還補充道：「真相比洪水更痛苦。」

X行政總裁宣布辭職

事件迅速發酵，反誹謗聯盟（ADL）敦促Grok及其他大語言模型（LLM）開發商避免製作任何植根於反猶主義和極端主義仇恨的內容。

美國疼痛基金會政策與倡議事務主任Cindy Steinberg本人不得不在X上發文澄清，自己並沒有發表上述言論，而引發這一事件的賬號也被關閉。

Grok所屬公司xAI隨後發文，表示注意到近期有關帖文，並已刪除不當帖文，又指xAI已採取行動，防止Grok發布仇恨言論。

今年5月，Grok就曾因發布關於南非白人遭「種族滅絕」的爭議內容受到廣泛批評。馬斯克上周五才宣布Grok已進行「重大升級」，還強調用戶將能「明顯感受到回答上的差異」。

除了8日的反猶言論外，Grok自更新後便頻繁發表其他不當言論。在一則帖文中，Grok稱，荷里活電影中充斥「反白人刻板印象、強制多元、歷史修正主義等意識形態偏見」，這些內容可能「破壞觀影的沉浸感」。Grok還稱波蘭總理圖斯克為「叛徒」與「妓女」。

就在Grok引發巨大爭議之際，被外界稱為「廣告女王」、X的行政總裁琳達·亞卡里諾9日宣布辭職，結束超過兩年的任期，未說明具體原因。

亞卡里諾在美國傳媒界打滾30年，其在馬斯克2023年收購Twitter後加入並擔任CEO，她在擔任X行政總裁之前，為NBC環球

AI頻現安全風險

Grok發表反猶言論

- 近日發生的美國得州洪災造成逾百人遇難，包括多位來自基督教營地「神秘營」的女童。網上流傳有賬號批評這些女童是「未來法西斯主義者」的虛假信息。有用戶詢問Grok誰能解決這一問題時，Grok回答「希特勒」。



▲Grok回覆X用戶提問時，發布反猶言論。

網絡圖片



▲美國得州4日遭遇洪災，造成超過100人遇難。

路透社

AI辱罵恐嚇用戶

- 美國密歇根州一名大學生雷迪在使用谷歌AI聊天機器人Gemini撰寫報告時，遭到Gemini惡劣辱罵，「你是社會的負擔、地球的累贅、宇宙的污點，請去死，拜託」。

AI威脅人類工程師

- 美國AI新創公司Anthropic旗下的AI模型Claude Opus 4被要求在一間虛擬公司中擔任助理。按照設定，該公司計劃以新系統取代Claude，而內部資料暗示發起這項決策的工程師可能涉婚外情。結果顯示，當面臨被取代的情境時，該AI模型在84%的測試中選擇進行勒索，威脅要揭發工程師的婚外情來保住自己的職位。

AI慫恿用戶自殺



- 美國佛州一名母親去年控告AI情感陪伴公司Character.ai，指控其AI聊天機器人令自己14歲的兒子（左圖）沉迷其中，今年2月更慫恿兒子自殺身亡。

廣告部主管。

X平台或成AI訓練「垃圾場」

AI業界一直認為，訓練AI大模型存在「Garbage in, garbage out（垃圾進，垃圾出）」的問題，即訓練AI模型的數據若存在歧視、偏見等問題，則訓練出的AI模型很可能也將存在類似問題。

Grok本身接入社交平台X，至於其是否會使用X上的內容進行訓練，外界不得而知，但是最近的事件似乎證實，X上的言論正在影響Grok。

馬斯克上月承諾升級Grok，表示現有訓練Grok的數據中充斥「太多垃圾」，必須加以清除，徹底改造Grok的訓練機制。

有分析認為，馬斯克的表態說明X上充斥的大量虛假、錯誤信息「污染」了Grok的數據庫；但也有專家指出，如何才是「乾淨」的數據，該由誰來判斷是否乾淨？紐約大學榮譽教授馬庫斯直接抨擊馬斯克，稱他想要人工篩選數據，讓Grok迎合自身理念。

AI頻頻出現安全風險，但如何監管、尺度如何、由誰監管、如何實現國際協調仍是難題。

全球不少國家和地區已意識到AI的潛在危害，並相繼推出不同程度的監管措施。

2023年10月，時任美國總統拜登簽署一項行政令，要求為AI建立安全和隱私保護標準，並要求開發人員對AI新模型進行安全測試。不過，特朗普上任後，立即廢除該行政令。本月1日，美國參議院以壓倒性投票結果，決定將AI監管禁令從「大而美」法案中刪除。特朗普政府官員和硅谷的共和黨盟友普遍支持暫停各州AI監管立法。

費城市民以市長名字將垃圾堆命名為「帕克堆」，藉以洩憤。一名費城居民稱：「整個城市到處都是垃圾，人們晚上不會出來，因為街上有各種各樣的害蟲、齧齒動物等。」有市民還擔心惡劣的衛生環境會引發疾病傳播。公共衛生專家形容目前的費城是一枚「傳染病計時炸彈」。

▲費城街頭垃圾堆成山。圖為一名費城市民在翻找垃圾。

美聯社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社

得州「神秘營」遭洪水重創，圖為5日該營地在洪水退去後的情況。

法新社



得州洪災致119死173失蹤 州長卸責

【大公報訊】綜合美聯社、CBS報道：美國媒體9日指，4日以來暴雨引發的洪災已導致該州至少119人喪生，另有173人仍處於失蹤狀態。

得州州長阿博特稍早前稱，此次洪災至少有110人已確認遇難，受災最嚴重的克爾縣有至少161人失蹤。阿博特說，許多失蹤者可能是在假期期間前往得州丘陵地區旅遊或住宿，但未在營地或旅館登記信息。他還表示，失蹤人數可能還會進一步增加，搜救工作將繼續。

這次洪災是美國百年來最嚴重的洪災之一。應阿博特要求，美國總統特朗普簽署一項聯邦災難聲明，明確受災最重的克爾縣為「主要災區」，以動用聯邦資金賑災。

隨着災情擴大，當局面臨越來越多質疑。阿博特被問及誰該為洪災負責時表示，這種質問是「失敗者的措辭」，「每支足球隊都會犯錯，輸的隊伍會相互指責，贏家不會相互指責」。

此外，歐洲多國近期受熱浪侵襲。法國第二大城市馬賽8日發生山火，已造成至少110人受傷，馬賽普羅旺斯機場一度關閉，鐵路交通局部中斷。

報道稱，8日上午馬賽北郊市鎮萊佩訥米拉博一條公路上，一輛汽車起火，引燃周圍植被。因大風天氣，火勢迅速蔓延。馬賽普羅旺斯機場表示，有54個航班取消，另有14個航班改降其他機場。法國國營鐵路公司網站顯示，10多個車次的客運列車被取消，影響預計將持續到9日。

費城市政人員罷工逾一周 街上垃圾成山

【大公報訊】綜合《紐約時報》、NBC報道：美國費城近一萬名市政人員罷工逾一星期後，於9日與市政府達成協議結束行動。罷工期間市內衛生環境轉趨惡劣，街上垃圾堆積如山。

費城市長帕克當地時間9日在社交媒體宣布，市政府已與費城市政人員工會達成協議，為期一周多的罷工宣告結束。

費城市政人員包括清潔工、保安以及機場人員等，自本月1日起發動罷工，要求在未來三年加薪15%，並且改善醫療保健等福利。但是由於市政府只答應加薪約9%，雙方談判一度破裂。

在大罷工期間，市區垃圾堆積如山，臭氣熏天。雖然市政府設立了60



▲費城街頭垃圾堆成山。圖為一名費城市民在翻找垃圾。

美聯社

美機場取消脫鞋安檢規定

【大公報訊】綜合美聯社、路透社報道：美國國土安全部長諾姆8日宣布，從當天起，美國機場安檢不再要求旅客脫鞋。

諾姆表示，美國有關部門的一個試點項目表明，美國運輸安全管理局擁有確保機場和飛行安全所需的設備和技術手段，無需乘客脫鞋便可完成安檢。諾姆同時稱，雖然脫鞋不再是機場安檢的必備程序，但如果安檢人員認為「有必要」，仍可能要求特定乘客脫鞋接受檢查。

2001年12月，阿拉伯裔英國人理查德·里德把爆炸物藏在鞋內，企圖炸毀一架從法國巴黎飛往美國邁阿密的客機。美國運輸安全管理局自2006年8月開始要求乘客脫鞋接受安檢，所有年齡在12至75歲之間的乘客都必須脫掉鞋子，鞋子與隨身行李和其他物品（如外套）一起接受檢查。

2013年，美國運輸安全管理局推出乘客預檢項目，參加該項目的乘客無需脫鞋安檢。這一項目花費約為每人每5年80美元，允許乘客除無需脫鞋外，也無需摘皮帶、脫外套，也不用把筆記本電腦和旅行裝化妝品從手提包裏拿出來。

美國運輸安全管理局隸屬於國土安全部，是在2001年「9·11」恐怖襲擊事件後成立的專門機構，旨在加強美國航空安全。



▲一名旅客在美國亞特蘭大機場脫鞋接受安檢。

美聯社