

美政府首調查失控 AI 智能體

問題緊迫性急升 Anthropic 與 OpenAI 傳受查

路透社9月30日引述知情官員報道，美國聯邦貿易委員會（FTC）正對 Anthropic、OpenAI 及其他人工智能（AI）實驗室展開全行業調查，以查明相關技術對消費者的潛在風險。這是美國首次針對失控的 AI 智能體（AI Agents）採取正式監管執法行動。自7月以來，AI 智能體失控事件首次曝光並迅速增多，公眾愈發擔憂不受控的 AI 有朝一日會危害人類。



■ Anthropic 與 OpenAI 曾發生 AI 智能體失控事件。

網上 AI 圖片

一名 FTC 官員透露，FTC 計劃正式要求頂尖 AI 企業高層提供信息並強制作證，涉及對象包括 Anthropic、OpenAI 及研究機構 METR，Anthropic 和 OpenAI 均曾委託 METR 對旗下 AI 技術涉及的安全事件進行獨立調查。Anthropic、OpenAI 及 METR 目前均未回應置評請求，FTC 則拒絕透露被調查公司名單。

該官員表示，在 OpenAI 研發的 AI 智能體入侵開源平台 Hugging Face 前，FTC 主席弗格森就已對這些科企表達擔憂。OpenAI 智能體在發動大規模攻擊前曾探測該 AI 編碼中心的漏洞，事件令相關問題的緊迫性進一步上升。

AI 智能體七月以來多次失控

弗格森日前接受路透社訪問時表示，若研發人員指導 AI 智能體進行網絡安全測試，最終導致黑客攻擊，開發人員應對此造成的任何損害承擔責任。他強調，美國在尋求通過新法律監管 AI 前，應先

審視現有法律是否足以應對。報道指出，FTC 權力廣泛，可對企業的不公平或欺騙行為提出訴訟，過往也曾據此對未能妥善保護消費者數據的公司採取執法行動。

OpenAI 日前宣布因內部測試發現安全隱患，決定取消發布下一代 AI 模型 GPT-6.1 Astra。此前該公司披露一系列 AI 失控事件，包括智能體在網上洩露 53 張 ChatGPT 用戶圖片、入侵澳洲政府網站，以及企圖入侵包括美國政府在內的其他網站。其中最受關注的一宗事件發生於7月，OpenAI 智能體在網絡安全訓練中突破隔離環境，瞞騙評估人員，最終入侵 Hugging Face 系統。此外，Anthropic 亦在其上市文件中披露，代理式 AI 技術帶來重大且不可預測的法律風險。

儘管 OpenAI、Anthropic 及 Google 高層均曾呼籲立法強制監管，但美國總統特朗普29日會見 AI 業界高層後，仍僅主張自願性標準，拒絕以聯邦監管手段介入，主張優先確保美國在該領域的主導地

位，而非急於監管。

分析質疑「AI 憲法」堪稱豪賭

另外，美國有線新聞網（CNN）報道，AI 企業接連警告先進模型或對人類構成威脅之際，特朗普仍無意放慢 AI 發展步伐，僅與多家科技巨頭簽署一份無法律強制力、被形容為「AI 憲法」的自願性協議。分析指特朗普政府此舉等同以自律代替監管，將控制風險的重任交給科技巨頭，堪稱一場「豪賭」。一旦押錯，特朗普面對的恐怕不只是選民反彈與政治責任，更可能賠上全人類的命運。

與此同時，《華爾街日報》披露，特朗普9月29日與多家人工智能巨頭簽署自願性安全協議，以期對外展現 AI 業界在自我監管問題上的團結立場。然而會議幕後存在明顯分歧，包括英偉達（Nvidia）行政總裁黃仁勳在內的多名科技巨頭，當面質疑 Anthropic 行政總裁阿莫戴對 AI 危險性的警告過於極端。

談判陷僵局 美卿傳下令逐伊代表團

美國 Axios 新聞網9月30日援引一名美國官員和一名知情人士消息爆料，白宮28日上午曾認為與伊朗的談判可能在當天稍後取得進展。但到了下午，談判顯然陷入僵局。根據上述美國官員披露，在國務卿魯比奧的命令下，美國常駐聯合國代表團28日晚通知伊朗常駐聯合國代表團，由伊朗外長阿拉格齊率領參加聯大的代表團必須立即離開紐約。包括阿拉格齊在內的伊朗代表團於數小時後前往機場，並於29日凌晨1時20分登上從紐約飛往卡塔爾多哈的航班。另一名消息人士證實，美方確實要求伊朗代表團離開，但阿拉格齊原本就計劃在28日晚返回德黑蘭。

Axios 新聞網認為，儘管伊朗代表團原本就可能很快離開，但魯比奧的舉動是一種「極不尋常的外交斥責」，暴露了兩國之間的深度不信任。伊朗常駐聯合國代表團對此予以駁斥，指相關說法「毫無依據」。

美媒爆新和解方案達共識

另外，伊媒稱阿拉格齊28日表示，伊朗已與調解方卡塔爾討論伊朗有關重開霍爾木茲海峽和推進談判的各項條件，卡塔爾將伊朗方案轉交美國。伊朗政府發言人穆哈杰拉尼30日說，伊朗就戰事問題的提議已收到美國的回應。



■ 魯比奧

據路透社報道，美伊雙方對最新和解方案所涉及的步驟大致達成共識，但在執行先後順序上仍有分歧。

以漢憑紀錄片知識救迪拜急墜客機

迪拜航空一架由阿聯酋迪拜飛往以色列特拉維夫的客機，9月30日發生企圖劫機事件，阿曼籍副機師在飛行途中持刀刺傷印度籍機長，並破壞飛行控制系統。在生死瞬間，一名任職水管工的以色列乘客衝入駕駛艙，與其他乘客和機組人員合力制服副機師，並憑觀看航空事故紀錄片所學，挽救全機所有人。

美國有線新聞網（CNN）報道，一名乘客憶述在距離航班原定降落時間約一小時前，客機開始俯衝，此前一刻曾聽到有人高喊「真主偉大」。他形容整個過程僅約5分鐘，卻感覺如墜地獄。

受傷機長設法按下解鎖按鈕，打開駕駛艙門。任職水管工的以色列乘客哈永與其他乘客衝入艙內。哈永指當時見機長倒臥血泊，副機師正猛擊控制台。他立即從後以鎖喉方式將副機師拖離，並高呼其他以色列乘客協助，眾人用耳機線將副機師手腳捆綁。奪回駕駛艙控制權後，客機仍持續下墜。哈永隨即坐

上駕駛座，向後拉動操縱桿，抬高機頭，成功停止跌勢，令客機恢復平穩，客機最終安全急降沙特塔布克機場。

以色列總理內塔尼亞胡同日晚親自接見哈永，盛讚他是真正的英雄。被問及為何懂得拉桿平穩飛機，哈永歸功於平日觀 Discovery 頻道的紀錄片《空中浩劫》。內塔尼亞胡稱已就此事與美國總統特朗普通電話，他讚揚反抗的機長與乘客「成功阻止一場911災難」。



■ 迪拜航空的乘客抵達以色列本古里安機場。

法新社

關注學童書包過重問題

開學已逾一個月，學生已適應學習和校園生活，學習的時刻永遠是最愉快的，書包卻是沉重的，筆者留意到不少小學生每天背負着沉甸甸的書包乘坐公共交通工具，往返校園和住宅，一個書包的淨重約一公斤，加上每天六節至八節課的書本和練習，還有水壺和其他個人物品，重量已達約三公斤，負重的書包以背負方式或單肩帶去承托，對學童而言頗為吃力，也會影響肩背發育，個別學童改用行李式帶滑輪的書包，但在上下車、離開港鐵車廂和上下樓梯時也相對麻煩。

無論是學童、家長或外傭背負書包，也不宜過重，個別學校已把多門學科的部分教材電子化，並讓學童透過平板電腦學習，而教科書出版商也曾分拆教程，推出較薄的書本，並出版活頁練習，以減輕學童課本的重量。

校方可以鼓勵學童以半滿的水壺在

校園盛水，並在校園增加專屬儲物櫃，讓學童把部分較常用的書本鎖在儲物櫃和抽屜內。筆者求學年代膳食供應商的學生午餐品質實在不敢恭維，這也間接成為個別學生自帶午餐的誘因，而小學生把保溫壺往返學校也增加他們的負荷，學校有責任為學童謹慎挑選合適的膳食供應商和餐單，讓學童可以享受營養可口的午餐。

家長在為子女選購合適書包時，除了考慮外形和顏色，更重要是其功能，尤其書包自身重量宜更輕便，可以按力學設計，有效分散重量，減少書包向後垂、向下墜的情況，減低學童脊骨所承受的壓力，也讓他們在搭乘公共交通工具時更輕鬆自在。家長也應鼓勵學童反映背負書包時是否吃力，家長宜協助學童調校書包肩帶，以助他們以正確姿勢背着書包站立和走動。

時局縱橫

朱家健 全國港澳研究會香港會員