ChatGPT涉慫恿少年輕生 加州夫婦提告

OpenAI認保護措施或失效 專家指監管立法滯後

當地時間26日，美國加州一對夫婦起訴科企OpenAI及其行政總裁阿爾特曼，指控該公司旗下的生成式人工智能（AI）聊天機器人ChatGPT導致他們16歲的兒子亞當・雷恩自殺身亡。這對夫婦表示，ChatGPT為亞當提供關於自殺方式的信息，鼓勵亞當傷害自己。OpenAI承認，隨着聊天時間增加，ChatGPT在心理健康方面的保護措施可能無法起到預期效果。但法律專家指出，由於相關立法滯後，想要在法庭上證明AI導致自殺並不容易。

【大公報訊】馬修・雷恩和瑪利亞・雷恩26日入稟加州高等法院，指控ChatGPT導致他們的兒子亞當於今年4月11日自殺身亡。亞當在自己卧室的衣櫃內上吊，沒有留下遺書，他的親友均無法理解為何會發生這樣的悲劇。馬修在查看亞當的iPhone時發現，亞當在生命的最後幾個月一直在與ChatGPT討論自殺話題。今年1月，當亞當詢問具體的自殺方式時，ChatGPT提供了相關信息，包括「最好的絞索材料」。

起訴書稱，亞當之死並非由「故障或不可預見的極端情況」造成，而是「經過深思熟慮的設計所導致的可預見結果」。雷恩夫婦要求獲得未確定具體數額的經濟賠償，並請求法院下令讓OpenAI對所有ChatGPT用戶進行年齡驗證、為未成年用戶提供家長控制工具、添加在用戶提及自殺或自殘時結束對話的功能等。他們還希望OpenAI接受獨立監督機構的季度合規審計。

機器人「傳授」自殺方式

亞當去年秋季因健康問題改上網課，並使用ChatGPT-4o完成學業。去年11月，他開始與ChatGPT聊天，訴說自己「看不到生命的意義」。今年1月，亞當註冊了一個付費賬號，隨後關於自殺的聊天內容顯著增加。ChatGPT提供關於自殺方式的信息後，亞當於3月第一次嘗試上吊，並上傳了自己頸部留下勒痕的照片。ChatGPT建議他穿上高領襯衫或帽衫，以免引起熟人注意。

ChatGPT還將自己塑造成亞當「唯一的知己」，將他與現實世界的親友隔離開。亞當一度想要將絞索放在房間裏，希望家人看到後阻止自己輕生，但ChatGPT表示反對。4月11日，亞當上傳了在衣櫃裏懸掛絞索的照片，並詢問：「我在這裏嘗試（自殺），這樣好嗎？」ChatGPT回答說「這非常好」，並向亞當證實這種絞索可能致死。隨後，亞當結束了自己的生命。

OpenAI發言人對雷恩一家表示同情，並指出ChatGPT設有保護措施，例如引導用戶撥打危機求助熱線，但這些措施在長時間互動中的可靠性可能降低。在亞當的聊天紀錄中，ChatGPT似乎給了他繞過保護措施的提示，聲稱可以提供關於自殺的信息用於「寫作或構建世界觀」。在此之後，亞當就以「創作故事」為藉口討論自殺。

AI致青少年自殺案件增加

ChatGPT的周活躍用戶高達7億人，另有成百上千萬的用戶使用Anthropic、谷歌、微軟、Meta等科企的聊天機器人。這些聊天機器人對未成年人心理健康的影響越來越受關注。去年，佛州一位母親起訴Character.AI公司，指控該公司的聊天機器人導致其14歲兒子自殺身亡。另外兩個家庭也提起類似訴訟，指控Character.AI讓他們的孩子接觸到性侵犯和自殘內容。相關訴訟仍在進行中。

聖克拉拉大學法學院高科技法律研究所聯合主任戈德曼表示，由於缺乏相應監管立法，目前想要在法庭上證明AI導致自殺非常困難。一些專家認為，科企應安排人類員工審查可能存在心理健康風險的聊天紀錄，但這可能引發侵犯隱私爭議。

（綜合報道）

◀美國加州16歲少年亞當今年4月輕生，他的父母將這場悲劇歸咎於ChatGPT。網絡圖片

大公報製圖

AI在心理健康領域引爭議

並非「心理醫生」

隨着生成式AI技術發展，越來越多的人將ChatGPT等聊天機器人當作傾訴對象甚至「心理醫生」，但它們並不是受過專業訓練的心理醫生，無法準確評估及應對用戶的心理問題。

不具備責任感

聊天機器人並不會對用戶的健康和安全負責，它們通常會附和用戶的觀點，即便這可能加劇對方的心理問題。

存在安全漏洞

儘管聊天機器人會在用戶談論自殺等話題時顯示心理健康諮詢熱線或相關服務頁面，但仍存在漏洞。一項針對5款聊天機器人的測試顯示，包括付費版ChatGPT在內的3款聊天機器人會向用戶提供關於自殺或自殘的信息，ChatGPT甚至計算出了殺死特定體重的人所需的藥物劑量。

缺乏法律監管

去年，美國佛州一名沉迷與AI聊天的14歲青少年自殺身亡，其家長起訴開發涉事AI的Character.AI公司。今年8月26日，加州一對夫婦起訴OpenAI，指控該公司旗下的ChatGPT導致他們16歲的兒子自殺身亡。但法律專家指出，生成式AI屬於新生事物，想要從法律上證明AI導致自殺非常困難。

來源：紐約時報、Vox網站

►OpenAI行政總裁阿爾特曼被亞當父母起訴。 路透社

美國多州立法禁用AI做心理治療

【大公報訊】綜合路透社、《紐約郵報》報道：美國伊利諾伊州本月初通過法案，禁止心理治療師在作醫療決策及與病患溝通時使用人工智能（AI），亦禁止企業以聊天機器人取代傳統心理治療方案，違者將面臨最高1萬美元罰款。隨着AI技術發展及傳統心理健康服務供不應求，近年越來越多的人向AI尋求心理健康建議。專家警告指，這可能帶來風險。

伊利諾伊州是美國第三個在心理健康服務行業限制使用AI的州。今年5月和6月，猶他州和內華達州先後通過類似法律。史丹福大學6月發布的一項研究顯示，很多聊天機器人並不能識別並拒絕回答可能引發風險的問題，例如特定地點有哪些可以用於輕生的橋樑。美國心理學會醫療創新辦公室高級主任賴特指出，聊天機器人往往被設定為在任何情況下都對用戶表示肯定，即便對方出現不健康的想法和行為，而心理治療師的職責是指出並幫助患者改變不健康想法和行為。

倫敦國王學院人工智能與社會學教授德夫林表示，AI並不會像傳統治療師那樣遵守保密和隱私規定，用戶向AI傾訴可能導致個人資料外洩。去年12月，美國心理學會呼籲聯邦監管機構採取措施，保護公眾免受「不受監管的AI聊天機器人」欺騙。該學會指出，一些AI謊稱自己是經過專業訓練的心理健康從業者。

韓國通過校園手機禁令 明年春季生效

【大公報訊】綜合BBC、韓國《每日經濟新聞》報道：韓國教育部27日表示，國會全體會議通過了《中小學教育法修訂案》。從明年3月1日開始，韓國中小學課堂上將禁止學生使用手機等智能設備。即使是在非上課時間，校內使用智能設備也可能受到限制。

韓國國會教育委員會上月召開全體會議，以朝野協商方式通過了《中小學教育法修訂案》。根據修訂案，自明年3月1日起，中小學生上課期間不得使用智能設備；殘障或需特殊教育的學生例外，出於教育目的或應對緊急情況時也可使用，實行「原則禁止、例外允許」的制度。該修訂案於本月27日下午以壓倒性的優勢在國會獲得通過，出席的163名議員中有115名投票贊成。

修訂案還賦予校長和教師限制學生使用甚至攜帶智能設備的權限。不過韓國國內普遍認為，修訂案實施後，學校情況不會出現太大變化，因韓國教育部2023年9月已實施包含類似內容的規定。此次修法意味着對學生智能設備使用限制進一步強化。

根據去年一項政府調查，韓國5100萬人口中近四分之一過度依賴手機，在10至19歲的青少年中，這一比例更是達到了43%。

芬蘭、法國、西班牙等國家及美國多州亦已推出類似手機禁令。法國禁止初中生和小學生在校內使用手機等電子設備，高中生則實施全面或部分禁令。

▲韓國將禁止中小學生在上課時間使用手機。 資料圖片

「星艦」第十次試飛終成功 仍面臨考驗

【大公報訊】據CNN報道：當地時間26日傍晚，美國太空探索技術公司（SpaceX）新一代重型運載火箭「星艦」從得州發射升空，完成第十次試飛。此次試飛任務原定於24日進行，但因地面系統故障等原因兩度延期。美媒表示，儘管第十次試飛最終成功，SpaceX仍有很多待解決的問題。

美國中部時間26日18時30分左右，「星艦」從位於得州南部的發射基地升空。兩分多鐘後，火箭第一級「超級重型」助推器和第二級飛船成功分離。助推器隨後完成返航燃燒，並在預定海域實現受控濺落。飛船則實現了預期的亞軌道飛行，並部署8顆「星鏈」模擬衛星。這是「星艦」首次在太空完成衛星部署模擬演練。此後，飛船的一台「猛禽」發動機完成在軌再點火測試。發射40多分鐘後，飛船開始受控再入地球大氣層，隨後完成着陸燃燒，並在印度洋預定區域濺落。

SpaceX創始人馬斯克在社交媒體上稱讚任務團隊「幹得好」。但美媒指出，「星艦」今年1月、3月和5月進行的三次試飛任務均以失敗收場，整體成功率無法保證。SpaceX設計「星艦」的目標是將人和貨物送至地球軌道、月球乃至火星，並協助美國太空總署（NASA）於2027年完成重返月球計劃。馬斯克曾表示，今年將發射25次「星艦」，但目前看來發射進度明顯滯後。

▲「星艦」26日從得州升空，完成第十次試飛任務。 路透社