

把 ChatGPT 當傾訴對象？荷蘭萊頓大學警告潛在風險

駐歐盟兼駐比利時代表處教育組

荷蘭萊頓大學研究人員對越來越多人使用 AI 作為傾聽者的攀升趨勢展開調查。他們指出 AI 可為心理健康問題提供快速的心理支持，但也指出圍繞在隱私和對話品質方面的風險。在荷蘭和歐洲，越來越多的人正在使用 Wysa、Woebot 和 Replika 等聊天機器人進行關於心理健康的對話。這在歐盟《人工智慧法案》(AI Act) 和《一般資料保護規則》(General Data Protection Regulation, 簡稱 GDPR) 的框架下，對醫療機構和政府提出了新的課題。

AI 傾聽服務正嶄露頭角

AI 傾聽服務是一種透過文字或語音進行對話並嘗試提供支持的 AI 聊天機器人服務。例如包括科技公司 Touchkin 的 Wysa、AI 心理健康平台 Woebot Health 的 Woebot，以及科技公司 OpenAI 更廣為使用的 ChatGPT。這類工具提供 24 小時不打烊且低門檻的服務，對於不想尋求專業人士協助的人來說也是如此。因此，從學生福祉到企業員工關懷，其使用量正在增長。

其應用範圍從簡單的日常關懷到認知行為治療的練習。有些聊天機器人使用固定的問卷運作，有些則使用能自行生成句子。生成式模型（通常是大語言模型）會根據海量的文本預測下一個單字。這能讓對話變得流暢，但同時也增加錯誤建議的風險。（AI 聊天機器人是一種自動化的對話系統，能透過文字或語音對使用者的問題或訊號做出回應。）

教育界、市政府及「精神衛生與成癮醫療體系」（Geestelijke Gezondheidszorg, GGZ）機構正在探索將這些工具作為接觸專業人士之外的補充。他們希望能藉此緩解排隊等待的時間，並做到更早期的預警。萊頓大學的研究人員強調，AI 是一項輔助工具，而非治療師。因此，選擇使用此類工具需要有明確的目標和界限。

低門檻、隨時可用

其最大的前景在於便利性。匿名性降低了那些感到羞怯或猶豫不

決之人的求助門檻。這些工具及系統隨時開放，即使在深夜或週末也是如此，可處理輕度問題來減輕求助專線和家庭醫生的負擔。

萊頓大學對其效果的研究顯示出了好壞參半但令人振奮的結果，尤其是在自我幫助和心理健康衛教方面。聊天機器人可以協助進行情緒追蹤、寫日記和呼吸練習，與此同時，它並不能取代專業人員的治療。對於較嚴重的症狀，仍應以人際接觸為核心。

安全的設計至關重要，系統中必須始終有一個顯眼的緊急求助按鈕，以便能直接轉接到 112（荷蘭緊急求難電話）或 113（荷蘭自殺預防專線）。提供荷蘭語及弗里西語（Frisian）的語言及文化支持以增強其可用性。此外，清晰易懂的語言品質和無障礙網頁規範（WCAG）也是基本的前提條件。

隱私與品質風險

與 AI 討論感受和健康會涉及敏感數據，在歐盟《一般資料保護規則》GDPR 框架下將其視為「特殊類別個人資料」，需要取得用戶明確且知情的同意。資料最小化和端點加密應該成為標準配置。在歐盟內部進行數據存儲，以及對向第 3 國傳輸數據進行仔細評估，依然至關重要。

這種生成式 AI 模型（聊天機器人）可能會產生「幻覺」，它們會自信地給出不正確或不安全的答案。關於藥物治療或自我傷害的無意建議會帶來風險。在這種情況下，僅靠免責聲明是不夠的。明確的界限、過濾機制和人工介入機制必須深植於系統設計之中。

此外還存在社會風險。擬人化的語言和虛擬分身可能會激發出虛擬的親密感與依賴性，就像之前在聊天機器人 Replika 所發生過移情效應案例。兒童和青少年尤為脆弱，因此需要年齡驗證和家長溝通機制。對於「在另一端」的究竟是人還是機器，必須始終保持透明。

《一般資料保護規則》與《人工智慧法案》

歐盟《人工智慧法案》（AI Act）將應用程序劃分為不同的風險等級。一個實際提供心理健康照護的聊天機器人，可能會被歸類為醫療器材並受《醫療器材法規》（MDR）管轄，從而被視為「高風險」。此時將適用嚴格的要求：風險管理、數據治理、日誌記錄、人工監督及

帶有通過歐盟的安全法規審查並取得歐盟認證(CE 標籤)的評估程序。即使不具備該身份，透明度義務和基本安全要求在任何情況下依然有效。

在《人工智慧法案》下，生成式 AI 模型的供應商在文件記錄和測試方面承擔額外義務。向公民提供聊天機器人的主體必須明確表明該系統為 AI 系統。禁止使用具操縱性的設計選擇(暗黑模式*故意利用人類的心理弱點或視覺習慣，來誤導或強迫使用者做出「對廠商有利，但對使用者不利」的決定)。對於政府和醫療提供者，在採購和使用時則適用應盡的責任與審查標準。

《一般資料保護規則》GDPR 依然是所有數據處理的主導標準。組織必須進行資料保護影響評估、引入個人資料保護官並簽署數據處理協議。只能處理必要的數據，且保存時間應儘可能縮短。用戶必須能夠輕鬆行使其權利，例如查詢權和刪除權。

監督及實踐

在荷蘭，健康與青年監察局(IGJ)負責監督醫療軟體和醫療效用說明，具有醫療照護功能的 AI 輔助工具應具備 CE 標誌。在採購招標中，健康應用程式的品質標準 EN-ISO 82304-2 越來越頻繁地被要求，以上做法有助於安全性、可用性和臨床實證上的評估。

地方政府、應用科學大學及其他大學院校正在探索利用 AI 聊天機器人來為學生和居民提供支援。相關採購條件中越來越要求資料本地化、稽核日誌及易用性。此外，清晰的荷蘭語和 B1 語言程度的說明對於可信任的互動至關重要。針對荷蘭語的數據及口音培訓也有助於提高服務品質。

其次，與 113 自殺防治熱線、家庭醫生、夜間門診服務及精神醫療機構的合作對於妥善轉介也至關重要。一旦發生意外，必須迅速通報並展開調查，獨立審查及用戶評估應成為改進工作的重點工作。荷蘭致力於人工智慧 ELSA 倡議，可在這方面提供方法與知識。(註 *ELSA 為倫理 Ethical、法律 Legal、社會 Social 及方法 Approach 4 個英文字的縮寫，主張發展科技時必須將上述 4 因素影響納入考量。)

使用建議

-針對醫療服務提供者的使用建議：

發佈清晰的安全說明書，規範限制、操作規範和禁忌事項。針對危機訊號出現、涉及未成年人及複雜個案時引入人工監督。限制資料收集，針對所有個資內容進行加密，並參考荷蘭的實際情況進行測試。制定事件回應和強制通報計畫。

• 針對醫療機構的使用建議：

首先進行試點計畫及資料保護影響評估。邀請法律專家、臨床醫生、倫理學家和經驗豐富的專家共同進行評量。針對宣稱有醫療效果工具要求 CE 標誌。明確約定關於模型更新、偏差測試及日誌存取權限的協議。申請醫療保健索賠的 CE 標誌。

-針對使用者的使用建議：

將聊天機器人視為自我輔助工具，而非診斷手段。儘可能少地分享個人敏感資料。在出現嚴重念頭或面臨即時危險時：請撥打荷蘭防治專線 112 或 113。使用者應確認該服務是否明確標示其為 AI，以及了解個人資料是如何被處理的。

撰稿人/譯稿人：駐歐盟兼駐比利時代表處教育組

資料來源：2026 年 7 月 8 日，AI Insider，

<https://aiinsider.nl/nieuws/chatgpt-als-luisterend-oor-universiteit-leiden-waarschuwt-voor-risicos/>