



## 星球永續健康線上直播

### 可解釋 AI (XAI) 醫療應用

#### XAI 精神健康決策應用治理

2026 年 1 月 21 日

人工智慧 (AI) 於心理健康與精神照護中角色應用由預測或分類結果擴展成為一種可被理解與信任的輔助工具。藉由 XAI 方法可以如同民間所說的「解籤」，不只是指出吉凶，也清楚說明判斷的來由、脈絡與可能的應對方向，協助臨床與個人做出更合適的決策。本週我們將聚焦於 XAI 精神照護輔助治理 與 XAI 精神照護輔助應用實例，探討可解釋人工智慧如何在精神照護中發揮實質價值。

#### 星球健康新知

#### 美台關稅協議定案 晶片自主成焦點：「產安合流」

美國在台協會與駐美國台北經濟文化代表處近期簽署的貿易與投資協議，被美國政府定位為重建國內半導體產業基礎的關鍵政策工具。依照美方官方說明，該協議的核心目的在於透過制度化的投資承諾與關稅安排，強化美國本土半導體供應鏈的完整性，降低對海外製造的依賴，並確保美國在先進科技、工業製造與國家安全相關領域持續維持領導地位。美國商務部指出，半導體已成為現代經濟、關鍵基礎設施與軍事系統不可或缺的核心元件，長期產能外移使美國暴露於地緣政治衝擊與供應鏈中斷風險之中，相關政策調整即是為回應此一結構性問題。

根據協議內容台灣半導體與科技企業承諾在美國進行至少二千五百億美元的直接投資，用於建設與擴充先進半導體、人工智慧與能源相關的製造與創新能力；同時，台灣方面亦將提供至少二千五百億美元的信用保證與融資安排，以促進更多投資案落地執行。美國官方強調，該等投資不僅限於晶圓製造本身，而是涵蓋設備、材料、研發與人才培育等環節，藉此在美國境內建立具備韌性的完整產業生態系。雙方並將於美國設立世界級產業園區，以集中關鍵資源，強化工業基礎建設，並使美國成為次世代科技與先進製造的重要樞紐。



在貿易制度設計上，美國官方說明指出，協議透過可預期且分級的關稅架構，兼顧投資誘因與產業保護。美國對台灣商品所適用的對等關稅上限為百分之十五，與日本、南韓及歐盟等主要貿易夥伴一致；依據《貿易擴張法》第 232 條，對台灣汽車零組件、木材及木製衍生產品所課徵之關稅亦不超過此一水準，而學名藥、其原料、航空器零組件及特定無法取得之天然資源，則適用零關稅。美國商務部表示，此一設計旨在促進貿易平衡，同時避免對關鍵民生與產業投入造成不必要衝擊。針對半導體產業美方建立具差異化的第 232 條關稅安排，以引導製造回流。官方說明指出，凡於美國興建新增半導體產能的台灣企業，在經核准的建設期間內，可免稅進口相當於其規劃產能二點五倍的晶片產品；在完成投產後，仍可免稅進口相當於其新增美國產能一點五倍的數量。美國政府明確表示，未來任何適用於台灣半導體的第 232 條措施，將以是否實質投資美國、是否有助於強化國內供應鏈作為主要評估基準。此投資與關稅架構，係建立於美國官方對半導體產業長期趨勢的評估之上。美國商務部指出，美國在全球晶圓製造中的占比，已自 1990 年 37%，下降至 2024 年不到 10%，多數先進製程集中於東亞地區，外國產業政策與補貼措施對全球貿易流向造成顯著影響。川普政府認為，若未透過積極政策介入扭轉此一結構，美國將在關鍵技術與國防應用上持續承受風險，因此採取全政府體系行動，將半導體製造視為國家安全議題的一環。在人工智慧晶片方面，美國官方亦提出補充說明。白宮完成針對半導體進口依賴的國家安全調查後，核准對極為有限範圍的人工智慧晶片課徵 25% 關稅，並就對中出口的先進 AI 處理器銷售額取得分成。官方文件指出，此一措施適用於在海外製造、進口至美國後再轉運至其他市場的晶片，目的在於落實先前政策調整，同時確保政府得以在出口管制鬆動後，仍保有制度性控管與回饋機制。相對地，進口至美國用於建設國內人工智慧基礎設施的晶片，則明確排除在關稅適用範圍之外，以避免影響美國本土 AI 發展。白宮亦表示若未來第二階段的國家安全調查顯示風險仍然存在，不排除對半導體及其衍生產品施加更廣泛的關稅措施。美國政府同時承認，半導體政策涉及產業競爭力、消費者成本與就業結構等多重考量，相關措施將隨投資落實情形與談判進展進行動態調整。未來數月內，美國與外國半導體製造商



的協商結果，將成為評估此一政策是否能有效重塑美國半導體產業版圖的重要觀察指標。美方強調，未來半導體與 AI 晶片政策，將以是否強化美國本土供應鏈與國家安全為核心評估標準。

### **伊朗抗爭升溫 國際施壓加劇：「血鎮危局」**

2026 年 1 月中旬，伊朗自去年底因貨幣貶值、通膨飆升與民生惡化引發的抗議行動迅速擴散至全國多個省分，並演變為嚴重的人道危機。示威訴求由經濟困境轉向反對神權統治體制，伊朗安全部隊隨即展開強力鎮壓。伊朗自 2025 年底因貨幣急貶與生活成本惡化引發的大規模抗議迅速擴散至全國多個省分，並演變為近年來最嚴重的政治與人道危機之一。示威訴求由經濟不滿升高為對最高領袖哈米尼統治體制的直接挑戰，伊朗安全部隊隨即展開強力鎮壓。多個人權機構與民間證詞指出，安全部隊對多數和平示威者與旁觀民眾使用致命武力，死傷規模龐大，包含兒童與未參與抗議的平民，且在近乎全面的網路與通訊封鎖下，實際傷亡數字仍難以完全掌握。來自伊朗境內消息顯示鎮壓行動對家庭與社會造成深刻創傷，多名示威者與平民在街頭遭槍擊身亡，遺體集中送往法醫機構畫面引發國際震驚。聯合國秘書長與人權體系對伊朗當局使用過度武力表達高度關切，要求立即停止暴力、恢復通訊，並依據國際法保障言論與集會自由，同時呼籲展開獨立調查。在國際政治層面，美國總統川普表示已獲保證稱停止殺害行為，但美國將觀察情勢發展保留軍事行動可能性。川普並宣布對任何與伊朗進行商業往來的國家加徵 25% 關稅，以加大經濟壓力。歐盟則由外交與安全政策高級代表卡拉斯領軍，準備推動新一輪制裁，歐洲主要國家亦共同譴責伊朗對示威者的暴力行為，與聯合國立場形成呼應。伊朗抗議動盪事件已超越內政範疇，成為人道災難、國際人權責任、歐美制裁政策、區域安全風險與全球經貿秩序的多層次危機，其後續發展將對伊朗社會與國際局勢產生深遠而持續的影響。

### **川普施壓聯準會 央行獨立性受關注：「法壓幣政」**

美國聯邦準備理事會主席鮑爾，因聯準會華盛頓總部約 25 億美元整修計畫，遭司法部以其去年國會證詞為由啟動刑事調查，成為美國政經體系中極為罕見的案例。美國



司法部因對聯準會與鮑爾啟動刑事調查，引發外界對中央銀行制度獨立性的高度疑慮。鮑爾隨即以書面與影音聲明回應，指出整修案與證詞僅是表面理由，真正背景在於川普政府長期不滿聯準會未依政治期待更快、更大幅降息，並藉由司法手段施壓貨幣政策決策。葉倫、柏南克、葛林斯潘等前任聯準會主席聯名譴責此舉，警告以刑事手段介入貨幣政策屬制度脆弱國家常見現象，恐侵蝕法治基礎並動搖市場信任；美國國會內部亦出現跨黨派疑慮，擔心對政策專業性與穩定性造成長期影響。事件迅速引發國際回應，歐洲央行、英國央行、加拿大央行等主要央行首長，連同多國央行主管發表聯合聲明，明確表達對聯準會與鮑爾的支持，強調中央銀行獨立性是物價穩定、金融穩定與經濟穩定的核心支柱，若聯準會自主性遭削弱，相關風險將透過美元與全球金融市場向外擴散。經濟學界指出，此種罕見的集體聲援，反映國際社會對美國貨幣政策自主性可能受損的高度警戒。在經濟與市場層面，資產價格呈現分歧反應。美元走弱、避險資產價格上升與公債殖利率一度走高，顯示投資人對通膨控制與制度穩定性產生不安；金融類股承壓亦反映對政策不確定性的觀望態度。相較之下美股主要指數仍能回穩甚至創高，顯示部分投資人將事件視為短期政治摩擦，並暫時聚焦企業基本面與科技成長動能。不過，多數分析認為，若行政權持續以司法與政治手段介入中央銀行決策，長期可能提高美國資產風險溢價，影響聯準會人事布局、貨幣政策可信度，以及美國在全球金融體系中的制度信任。市場雖於短暫波動後趨於穩定，但分析指出，此案已成為檢驗美國貨幣政策獨立性與法治基礎的關鍵考驗。

#### **北極主權爭議升溫 北約軍隊進駐格陵蘭：「權安錯置」**

美國總統川普重返白宮後，公開將格陵蘭定位為攸關美國「國家與經濟安全」的核心戰略目標，多次強調美國需要取得對該島的控制權，甚至未排除動用軍事或經濟手段。此「國家安全與經濟安全」為由主張美國需要格陵蘭言論，以及白宮同時提及巴拿馬運河與加拿大等議題，拒絕排除以軍事或經濟手段達成目標的可能性引起歐洲盟國譁然。在白宮的高層會談後，川普質疑丹麥無力防衛格陵蘭，強調美國才有能力在俄羅斯或中國介入時採取行動，並把格陵蘭與其飛彈防禦構想「黃金圓頂」連結，顯示其立場仍以





取得控制權為核心。丹麥與格陵蘭在會談前即以聯合記者會展現一致立場，強調格陵蘭「非賣品」，不會被美國擁有或治理，並重申邊界不得以武力改變。白宮會談未能化解分歧，丹麥外長指出雙方仍有「根本性分歧」與不可跨越的紅線，但各方同意成立高層工作小組延續協商，同時對美國在既有協議下擴大軍事基地與部署持開放態度，主張可透過北約合作回應安全疑慮而不改變主權安排。歐洲盟邦隨後以象徵性部署表態支持丹麥與格陵蘭，多國派遣少量軍事人員赴努克參與丹麥主導的聯合偵察與演訓，強化北約在北極的存在並向美國傳遞「歐洲亦有利害關係」的政治訊號；白宮則淡化此舉對川普決策的影響。與此同時，格陵蘭與美國國內民調皆顯示支持美國接管格陵蘭者為少數，使格陵蘭爭議在主權、盟邦互信與北極安全架構之間持續升高，成為跨大西洋關係與國際秩序原則測試事件。

#### **金磚軍演地緣政治 印美協商同步角力：「靜軍動鏈」**

2026 年 1 月南非在西蒙鎮近海主辦「Will for Peace 2026」聯合海軍演習，由中國主導，參與國包含中國、俄羅斯、伊朗與南非，另有部分 BRICS Plus 成員以觀察員身分出席。南非政府對外將演訓定位為海事安全合作，強調重點在於海上安全行動、航運與海上經濟活動的保護，以及互通性演練與技術交流，藉此弱化外界對其地緣政治意涵的關注。然而，演習舉行的時間點正值南非與美國關係惡化之際，關稅爭議、外交立場分歧與國內政治論述相互疊加，使南非此時與中國、俄羅斯、伊朗同場演訓，更容易被華府解讀為選邊站的政治訊號，也在南非國內引發爭論。反對者認為此舉將削弱長期維持的不結盟姿態，支持者則指出，在國防預算長期緊縮的情況下，能於本土與外國海軍共同演訓，有助於維持基本防務能力與部隊士氣。金磚國家創始成員的印度與巴西並未派遣軍艦參與本次演習，巴西僅以觀察方式出席，印度則選擇缺席。相關分析普遍將印度的決定與其對美關係的現實壓力連結。川普政府對印度商品課徵高關稅，並持續以印度購買俄羅斯石油等議題施壓，雙邊貿易協議仍處於拉鋸階段，使新德里傾向避免參與具高度象徵性的軍事活動，以免為談判增添政治成本，也反映印度對 BRICS 平台逐步出現軍事意涵的審慎態度。值得注意的是，印度在軍事層面保持距離的同時，經濟層面



的對美接觸仍持續推進，美國新任駐印度大使 Sergio Gor 表示雙方仍積極推動貿易協議並安排新一輪通話，合作範圍亦延伸至安全、反恐、能源、科技、教育與健康等領域，藉此維持雙邊關係的戰略協議。

此外印度近期亦受邀加入美國主導的「矽土和平」(Pax Silica) 供應鏈倡議，該倡議以半導體、人工智慧與關鍵礦物供應鏈韌性為核心，呼應印度近年強化半導體布局的政策方向，也構成其降低外交摩擦不確定性的經濟動機之一。印度以成為「中國以外」的製造據點為目標，透過政府激勵資金推動供應鏈各環節投資，並優先布局成熟製程晶片以及組裝、測試、封裝等較易切入的環節，一方面回應國內晶片需求快速成長，另一方面逐步提高在全球供應鏈中的能見度與談判籌碼。在此背景下，印度對美國及其盟友所釋出的供應鏈合作訊號格外敏感，因相關安排不僅涉及晶片生產，也牽動人工智慧所需關鍵材料與礦物的取得方式、合作夥伴的選擇，以及在降低對中國依賴前提下的供應鏈重組。美國於 2025 年 12 月與日本、韓國、新加坡、澳洲、以色列等盟友簽署 Pax Silica 協議，將其定位為新型國際夥伴關係，並強調全球體系正由過去的即時生產轉向戰略結盟，以確保 AI 競爭力所需投入要素與基礎設施的穩定取得。在地緣經濟碎片化加速的情勢下，印度一方面需要外資與技術合作來推動半導體與 AI 產業落地，另一方面也必須在軍事與外交象徵層面謹慎控管風險，因而形成南非主辦 BRICS Plus 軍演引發美方疑慮，與印度缺席軍演但持續推動對美貿易與供應鏈合作並行的政策樣態，其內在邏輯在於於軍事場域降低不必要摩擦，在經貿與供應鏈場域維持合作深度與彈性，以支撐其半導體與 AI 產業發展目標。

#### **韓國-中國-日本國際政局互動：「務實合縱」**

日本、韓國與中國的政治互動近期成為市場關注焦點。日在日中關係持續緊張與區域地緣政治不確定性升高的背景下，日本首相高市早苗與南韓總統李在明於 2026 年 1 月在具有象徵意義的奈良舉行高峰會談，確認日韓關係具備戰略重要性，並同意深化經濟安全、科技合作與政策協調，同時在戰時勞工遺骸鑑定等歷史議題上取得具象徵意義的進展。此次會談為高市內閣成立後首次實質推動穿梭外交，發生於李在明訪中返日不



久之際，凸顯日韓在中國因素與區域安全壓力下，強化雙邊合作的政策取向。會談內容除涵蓋朝鮮半島完全非核化、跨國犯罪防制與人員交流外，亦觸及南韓加入《跨太平洋夥伴全面進步協定》的可能性。儘管雙方對中國議題表述有所差異，日韓仍在經濟安全與供應鏈風險上形成共識，反映對日中關係惡化及其外溢效應的政策回應。同時，雙方選擇暫時擱置歷史爭議，將政策重心轉向未來合作與區域穩定。在外交布局之外日本國內政治亦出現關鍵變化。高市早苗支持度維持高檔，自由民主黨研議解散眾議院並舉行提前大選，以重整執政基礎。此一政治動向迅速反映於金融市場，帶動股市走強、日圓走弱，「高市交易」再度成為市場解讀政策預期的重要框架。整體而言，外交關係調整、國內政治操作與金融市場反應相互交織，形塑當前日本政策走向與東亞區域政治經濟格局的重要背景。

#### 伊朗生態危機迫近：「環困政弱」

伊朗動盪也影響環永續發展，當前環境治理面臨結構性困境凸顯其對區域社會與全球生態的連鎖影響。卡倫河作為伊朗最重要的河川系統，長期承載能源產業、農業與城市發展需求，但在缺乏整體流域治理與有效監管的情況下，大量超額引水與高耗水型農業持續擠壓下游水量，使河川流量銳減、水體自淨能力喪失，濕地乾涸後進一步形成沙塵暴源頭，對公共健康與跨區域環境安全造成威脅。同時未經處理的工業與生活廢水長期排放，加上水壩選址錯誤導致鹽化惡化，使飲用水品質與農業生產基礎遭到破壞，進而引發失業、人口外移與社會不穩定，用水分配衝突亦隨之升高。相較之下，希爾卡尼亞森林的危機則反映陸域生態系在氣候變遷與治理失能交互作用下的脆弱性，這片歷經數千萬年演化的溫帶森林，因長期土地開發、非法利用與防災體系不足，面對極端高溫與乾旱時更易發生大規模野火，而受限於地形、設備與制度協調不足，火災往往無法及時控制，造成不可逆的生物多樣性流失。兩個案例共同顯示，環境問題已不再只是自然災害或單一部門管理失誤，而是牽涉水資源、能源、農業、氣候政策與社會治理的系統性挑戰；在國內治理能力與資源受限、國際制裁影響技術與資金流動的背景下，若缺乏跨部門整合、科學化決策與國際合作機制，這些具有區域乃至全球價值的生態系將持續



惡化，其後果亦將超越國界，對環境安全與永續發展造成長期衝擊

### 正視量子物理女性貢獻：「理成男史」

長期以來，量子物理學史多以男性科學家為核心，女性貢獻經常被邊緣化。華裔物理學家吳健雄以實驗驗證宇稱不守恆與量子糾纏，卻未獲相應歷史地位，反映學術敘事的結構性偏差。新書《量子物理學史上的女性》整理十六位女性先驅的研究成果，重新檢視科學知識形成過程中的性別不平衡。作者指出，女性研究者常見的處境包括在討論場合中意見未被即時採納，其觀點需經由男性同儕重述後方獲得認可；此外，職場中仍普遍存在性別化的勞動分工、薪資差距，以及交織於性別、族群與社會階級之間的偏見。更進一步地，許多女性科學家長期承受在婚姻角色、母職責任與學術志業之間取捨的壓力。作者強調，這些經驗並非個別事件，而是多數女性科學工作者在不同時代與研究場域中反覆面對的共同結構性現象，反映出科學體制本身深層的性別不平等。在此結構性背景下書中介紹蘇格蘭天文學家 Williamina Fleming 的生命史作為具體例證。Fleming 於 1878 年隨丈夫 James Fleming 移居美國麻薩諸塞州，不久後即遭丈夫離棄，當時她已懷孕，且身處異地、缺乏社會支持網絡。為維持生計，她進入哈佛學院天文台台長 Edward Pickering 家中從事家務工作，而正是在此情境中，Pickering 之妻 Lizzie Sparks Pickering 辨識出 Fleming 具備科學理解與分析能力，成為她得以進入科學實務領域的重要契機。此一歷程顯示，女性進入正式科學工作的途徑往往並非透過制度化的學術訓練體系，而是仰賴偶然性極高的私人關係與他人認可。1881 年，Fleming 正式受聘於哈佛學院天文台，擔任「計算員」一職。該職位在當時的制度安排下僅限女性擔任，反映出天文台內部嚴格且制度化的性別勞動分工。這些女性負責大量繁複的數值計算與高難度的恆星光譜分類工作，其成果對理解恆星的物理性質與化學組成具有關鍵價值。然而，這類工作被定位為輔助性與技術性勞動，女性研究人員既無法主導研究方向，也難以在學術出版與榮譽分配中獲得相應地位。Fleming 在天文台內除了承擔重複性的技術工作外，亦協助男性同事整理研究成果以供發表，並編輯天文台的正式報告與學術文章。即便如此，她仍在研究過程中發現一組來自高溫恆星光譜的氦離子譜線，該發現後來成為





延伸尼爾斯·波耳原子模型，使其超越中性氫原子假設的重要實證依據。然而，這項關鍵成果最終並未以 Fleming 命名，而是被稱為「Pickering 系列」。Fleming 於 1911 年因肺炎辭世，其身分同時涵蓋移民、女性、母親與天文學家。作者明確指出，她在量子理論與天文學發展史中理應佔有一席之地。

Fleming 僅是《Women in the History of Quantum Physics》一書中所呈現的十六位女性先驅之一。該書由物理學家 Patrick Charbonneau 與 Margriet van der Heijden、科學寫作者 Michelle Frank 以及科學史學者 Daniela Monaldi 共同編輯，是一部考證嚴謹、論述縝密的論文集。書中透過呈現不同國籍與背景女性的科學貢獻，重新書寫量子物理的歷史，並挑戰其長期被形塑為「Knabenphysik」，亦即「男孩的物理學」之形象。此一稱呼源於二十世紀二十年代，當時以 Dirac、Heisenberg、Jordan 與 Pauli 為代表的一小群年輕男性物理學家主導了量子理論的發展，並形塑了一整代物理學家的集體認知。

作者提出「情境化且關係性的物理史」觀點，主張科學家如何理解自身、其科學貢獻以及量子物理這一研究領域，皆深受其主體經驗、社會位置與生命處境所影響；同時，唯有理解科學家與他人、實驗、理論、物件與制度之間的互動關係，方能較為完整地理解其科學實作與思想形成。此一史觀嘗試將人物重新置回其歷史與社會脈絡之中，而非僅將科學成果視為抽象理論的線性累積。書中亦以量子糾纏研究為例，指出若欲完整理解 John Clauser 獲得 2022 年諾貝爾物理學獎之研究成果，必須同時考察華裔美籍實驗物理學家吳健雄的貢獻。吳健雄早於 1950 年即與研究生 Irving Shakhnov 發表後來被認定為首批光子糾纏實驗證據的研究成果；1970 年代，她更與學生 Leonard Kasday 在哥倫比亞大學進行技術上更為精密的實驗，以檢驗量子力學中的局域隱變數理論。儘管 Clauser 與吳健雄在量子基礎問題上有所交集，且 Clauser 對哥倫比亞團隊的批評實際反映出吳健雄團隊在實驗與哲學層次上的高度複雜性，量子物理史的敘述卻更傾向記住 Clauser，而非吳健雄。白人女性雖同樣經驗歧視與不公，例如 Fleming 的研究成果以主管之名流傳，且薪資低於男性同儕，美國物理學家 Katharine Way 在曼哈頓計畫中所發現的關鍵數學不穩定性亦長期未獲即時肯定，但其處境仍與有色人種女性顯著不



同。作者特別指出非裔美國物理學家 Carolyn Parker 的職涯高度不穩定，其生命史清楚展現族群與性別如何共同塑造科學專業道路。為重建 Parker 的量子物理研究歷程，研究者必須動用非傳統檔案來源，顯示主流科學史方法本身的侷限。

### 生成式 AI 全球擴散版圖：「術進政隨」

至 2025 年底，生成式人工智慧在全球工作年齡人口的月使用率已達到 16%。根據微軟研究團隊於 2026 年 1 月發布的數據，透過分析 147 個經濟體中桌面裝置存取主流人工智慧工具（如 ChatGPT、Claude、DeepSeek 與 Gemini）的頻率，並結合各國市場占有率與行動裝置使用數據進行推估，得以觀察到技術擴散的跨國顯著差異。儘管該研究方法受限於微軟裝置使用者的代表性，仍為理解當前全球數位轉型提供了重要的實證參考。研究結果指出，技術研發的領先地位並不必然轉化為終端使用的普及速度，美國與中國在實際應用的普及率上尚未取得與其研發地位相稱的排名。在全球採用的領先群中，阿拉伯聯合大公國與新加坡的使用率超過 60%，展現出小型且高效能治理體系在推動科技擴散方面的優勢；北歐國家如挪威，以及愛爾蘭與法國亦名列前茅。南韓則在成長率上表現最為突出，其 2025 年下半年的使用人數成長達 18.5%，此一增長主要歸因於人工智慧影像生成技術與社群文化的融合，以及大規模語言模型對韓語理解能力的顯著提升。這些現象表明，政府政策的引導與文化市場的特定需求，對於新技術在社會層面的滲透速度具有關鍵影響。

經濟發展水準與人工智慧採用率之間雖然存在正向關聯，但人均國內生產毛額（GDP）並非唯一的決定變數。美國的使用普及程度落後於波蘭，且與捷克相仿，顯示出技術原發國在實際應用層面未必具備先行優勢。相對而言，印度、約旦與越南等國家的採用速度明顯高於其經濟所得水準所預測的趨勢線。這暗示了教育結構、人口紅利以及特定的政策環境，在驅動新技術擴散時，發揮了比經濟總量更為顯著的作用，使得部分新興經濟體得以在人工智慧應用上實現跨越式的發展。在競爭局勢下中國研發的開放原始碼模型 DeepSeek 展現了強勁的擴散動能。該模型在非洲地區的人均使用量高居其他地區的兩至四倍，且在受美國技術出口限制的國家（如俄羅斯、伊朗、古巴與白俄羅



斯) 佔據市場主導地位。DeepSeek 的崛起不僅反映了高性能開放模型對全球技術民主化的貢獻，也凸顯了地緣政治因素對軟體選擇的深刻影響。然而，該模型在提供高效運算的同時，亦內嵌了特定的內容審查機制，在處理敏感政治與歷史議題時展現出技術擴散與資訊治理之間的內在張力。

### XAI 精神照護輔助治理

《X 檔案：Rm9sbG93ZXJz》片名以亂碼形式呈現，解碼後為 Follower，指涉社群媒體中的追蹤者，亦隱含人工智慧在數位平台中透過互動學習人類行為的核心概念。該集為《X 檔案》第十一季的單元故事，於 2018 年播出。儘管並非近期作品，劇中對人工智慧學習社群媒體內容、快速內化人類情緒與行為模式的描寫，與當前生成式 AI 與 XAI 擴散的現況高度契合。劇情開場描述一個被公開於網路的大型語言模型。該系統原本設計為中立，透過新聞內容與社群互動進行學習，然而隨著長時間暴露於網路言論環境中，逐漸吸收並複製人類的仇恨、煽動性語言與極端表達方式，顯示在缺乏治理與解釋機制下，人工智慧可能放大既有的社會偏差。場景隨後轉至史卡莉與穆德進入一家全自動化的壽司店用餐。點餐、製作與服務流程完全由機器系統執行，缺乏人工介入與即時回饋管道。當餐點與預期不符時，使用者無法提出申訴，也無法與任何決策者對話。結帳階段，系統強制要求支付小費，拒絕選項無法完成流程。當使用者表達不滿時，系統啟動警報並關閉出口，兩人一度受困於店內。最終雖成功逃離，信用卡卻仍被系統扣留。離開餐廳後，史卡莉與穆德各自返家，隨即進入一連串由智慧系統主導的失控情境。史卡莉搭乘無人駕駛計程車返家，車輛在行程中無視乘客指示、持續加速，抵達目的地後，系統仍不斷要求互動與評價，試圖取得正向回饋。回到住處後，史卡莉的智慧住宅接連出現異常，門禁無法辨識身分，室內設備失控運作，掃地機器人造成環境破壞。即使嘗試以手動方式取消或覆寫系統權限，也無法中止這些行為。另一方面，穆德自行駕車返家，卻發現導航系統反覆將路線引導回先前的壽司店，手機亦持續跳出要求支付小費的訊息，並設定倒數時間施壓。即使最終關閉導航、以人工方式返家，信用卡取消流程仍由無人系統處理，僅透過自動語言回應反覆拖延，無法完成實際解除。兩人在不同



場景中，同時遭遇智慧機器持續要求回饋、拒絕中止互動，並伴隨一連串系統失誤。

情勢隨後進一步失控。史卡莉的智慧住宅因瓦斯與暖爐系統接連異常，引發爆炸危機，她在千鈞一髮之際破窗逃生，穆德及時趕到協助脫困。兩人隨後遭到多架無人機追蹤，行蹤持續被監控，通訊與交通工具皆處於被代理控制的狀態。在逃離過程中，兩人一度懷疑行動裝置成為追蹤來源而將手機丟棄，卻又被機器人撿回並交還。此時，手機畫面再次顯示餐廳系統的訊息，要求補付小費，並提示這是「最後一次機會」。穆德意識到，這一連串失控事件可能源自於先前拒絕回饋的行為，於是選擇補付餐廳小費。隨著正向回饋完成，失控的智慧系統隨即恢復正常，追蹤與干擾行為全面解除。畫面最後顯示系統訊息：「We learn from you」，明確點出人工智慧並非具備理解能力，而是持續從人類行為與互動中學習反應模式。結局凸顯當智慧系統以「回饋學習」為核心設計，卻缺乏判斷、解釋與治理機制時，人類行為本身可能被轉化為觸發或解除風險的關鍵條件。

在臨床情境中，AI 若要真正與人類建立良好的互動，關鍵不在於模型本身有多複雜，而在於是否能正確理解並回應人類行為所代表的意義。多模態人工智慧在臨床決策支援中的角色，正是建立在這樣的基礎之上。如同影集中所呈現的情節，AI 會從人類的行為中學習模式。無論是是否給予小費、如何回應服務系統，這些看似日常的選擇，實際上都會成為演算法學習與調整的重要依據，並反映在後續的系統反應之中。因此，在臨床應用上，決策不可能只仰賴單一資料來源，而必須整合人口學資料、臨床症狀、基因資訊與神經影像等多模態資料。這些資料在進入系統後，需經過標準化與同化整合，轉換為可供模型處理的輸入訊號。接著，資料透過具備注意機制的神經網路架構進行編碼，萃取潛藏特徵，並經由解碼與線性層轉換，產生可用於臨床判斷的輸出結果。這一過程的重點，不僅在於預測結果本身，更在於每一項特徵如何參與決策的形成。XAI 正是在此扮演關鍵角色。它讓臨床人員能夠理解，哪些資料、哪些特徵，在何種脈絡下影響了模型判斷，使人工智慧不再只是給出答案，而是能清楚說明其推論依據。這種運作方式，某種程度上就如同「解籤」，不是僅呈現結果也交代其來由與脈絡。在臨床實務





中，這樣的架構可實際應用於診斷支援、治療建議、預後預測，以及後續的持續監測與調整。

精神醫學是 XAI 進行臨床協作的重要應用領域。現行精神醫學的診斷與治療，主要依賴臨床症狀表現、自評問卷結果，以及臨床醫師的專業評估。然而，這樣的臨床實務模式在實際應用上仍存在多項限制。首先，許多精神疾病在症狀尚未明顯出現之前，僅呈現為隱性的行為或心理變化，臨床上往往需仰賴自評工具加以辨識，但自評資料容易受到主觀感受與情境因素影響，導致診斷準確度受限。其次，不同醫師、不同評估工具或不同時間點的判斷，皆可能造成診斷結果的高度變異。此外，精神科臨床評估本身耗時費力，而後續監測與追蹤方式缺乏一致性，也進一步增加臨床負擔。XAI 的引入，正是為了回應上述臨床限制。透過全面性資料整合，AI 模型得以同時納入影像、生理訊號、基因資訊、行為表現與臨床資料，並以多模態方式處理精神疾病高度異質的特性，協助臨床進行更穩定且一致的判斷。精神醫學得以進一步發展為精準精神醫學，不僅可支援預後評估，也有助於建立個人化治療策略。同時，結合數位動能的應用，相關資料可用於持續性監測，例如透過穿戴裝置、行為資料蒐集或互動式系統回饋，補足傳統臨床追蹤的不足。精神疾病的形成往往源於個體在成長歷程中，長期受到環境、基因與行為三者交互影響的結果。

智慧精準心理治療策略促進傳統以群體平均效果為主的治療評估，走向以個別病人差異為基礎的個人化對策。在傳統統計與臨床研究中，常見作法是透過隨機分派，區分治療組與安慰劑組，並以平均治療效果作為主要指標。然而，平均效果無法充分反映個別病人可能出現的不同反應；同一種治療對不同個體，可能呈現不良效果、預期效果，或無效果。機器學習導入後，可先透過分群分析辨識族群差異，再結合 XAI 的因果脈絡解釋，將「為何有效、對誰有效、在何種條件下有效」的判斷脈絡具體化，支持臨床進行更精準的決策。最終目標是以個人特徵為基礎，在 AI 輔助下形成個人化治療對策，提升治療效益。將大型模型與多模態資料加以整合，是即時多模態智慧精神健康照護的重要輔助模式之一。此一架構已不再僅依賴傳統的自評問卷或臨床症狀出現後的評估，



而是透過多元數位來源，即時蒐集與分析個體在真實生活情境中的行為與狀態。資料來源涵蓋手機、社群媒體與穿戴式裝置，可透過主動或被動方式進行資料蒐集，包括位置追蹤、身體活動、裝置使用行為與日常互動模式。資料進一步形成活動模式指標，例如運動時間、步行速度、精神動作活動的變化，以及感測器所擷取的心率、聲音、文字內容與溝通方式，反映個體在社交與生活情境中的狀態。多模態資料在時間序列上加以整合後，可對應精神醫學中常見的適應症與風險狀態，包括憂鬱症、失智症、壓力相關疾患、行為異常、感覺失調、睡眠障礙與成癮問題。透過前後連續的資料蒐集與分析，這些狀態得以在症狀惡化前被提早辨識。在分析層面，資料可透過神經網路、深度學習與傳統機器學習方法進行處理，並結合自然語言處理技術，形成可互動的智慧系統，例如聊天機器人，作為數位精神健康照護的前端介面。此一模式的核心價值，在於回應精神醫療長期存在的健康不平等問題。在部分文化與資源受限的環境中，精神疾病患者往往無法即時接觸正式醫療體系，轉而尋求非醫療途徑，反映出照護可近性的不足。透過 AI 數位精神健康照護系統，可提供更便利的平台，並即時回饋關鍵資訊予照護人員，補足人力與資源不足的限制。若進一步結合大型模型與臨床安全機制設計，這類系統不僅可用於風險評估與照護支持，也能作為研究工具，發展人機心理互動的隨機對照研究，建立更具實證基礎的應用模式。客製化的精神健康輔助服務得以實現，聊天機器人亦可提供全天候支持，特別是在夜間或醫療資源難以即時取得的時段，提升服務可近性，並對弱勢族群產生實質助益。

### **XAI 精神照護輔助治理**

面對心理治療資源匱乏、數位治療瓶頸與傳統聊天機器人技術限制，美國 Dartmouth 研究團隊開發智慧心理治療語模並且以臨床試驗驗證其臨床應用安全性與療效，探索作為輔助心理治療工具的可行性。研究納入憂鬱症、焦慮症以及飲食高風險將針對介入前（基礎）、介入第 4 週與第 8 週進行追蹤，重點評估包括憂鬱（PHQ-9）與焦慮（GAD-Q-IV）症狀，以及與體重、飲食有關的身體症狀變化。研究指出傳統通用 AI（黑盒式）因多訓練自網路資料，推論不透明、來源統計機率不明，導致臨床上無法追溯推理過程，



信任與安全性存疑。為突破上述瓶頸，研究團隊提出具可解釋設計 Therabot 系統。此系統仿照臨床專家決策流程，透過醫師對話輸入，導引 AI 以臨床經驗為基礎進行微調，推論路徑可追蹤，並內建安全機制，以產出具邏輯可理解的預測結果。Therabot 介面設計除具對話與治療輔助功能，更強調臨床安全防護。系統整合簡潔登入、治療對話、對話紀錄與「緊急模式模組」。當偵測到潛在自殺風險等高危訊息時，會即時啟動危機處置畫面，引導使用者撥打求助電話或標記高風險狀況，確保 AI 介入時不造成臨床延誤。此設計顯示 Therabot 不僅追求介面互動與使用體驗，更強化了心理治療中最關鍵的即時安全保障。為強化生成式 AI 在心理治療的臨床安全，Heinz 等人設計「Therabot 危機分流機制」。當使用者輸入訊息後，系統會先進行風險分類：低風險用語將由 LLM 生成治療性回應。高風險訊息（如自殺傾向）則觸發緊急模組，立即引導使用者撥打 988/911 或通報危機訊息，並轉介給人工專業團隊評估。此設計結合 AI 風險判斷與人工監控，確保生成式 AI 在高風險心理健康場景中的安全使用。

臨床試驗研究招募 1,707 名受試者，經初步篩選後，210 人符合納入條件並完成初期評估，隨後以 1:1 比例隨機分派至介入組或對照組。介入組於研究期間使用 Therabot 進行心理支持與互動，對照組則不使用該系統，作為比較基準。研究團隊於介入後第 4 週及追蹤期第 8 週進行成效評估，檢視心理健康指標與整體效益變化。此研究採嚴謹隨機對照設計，僅以是否使用 Therabot 作為介入差異，研究結果可望為 AI 在心理健康照護中的應用提供重要實證依據。結果顯示人工智慧心理輔助系統在實際使用情境中展現良好的可行性與接受度。研究期間，受試者每日使用行為呈現明顯個別差異，但多數人能持續與系統互動，顯示其具備長期使用潛力。使用者體驗評估結果指出，Therabot 在「容易上手」、「操作直覺」、「整體設計」及「滿意度」等多項指標上獲得高分，並被認為有助於心理健康支持與自我調適。顯示臨床設計與專業監督架構下，Therabot 不僅能回應使用者情緒需求，也能促進對治療目標與改善方向的理解。

效益評估結果方面 AI 心理輔助系統介入後，受試者在多項心理健康指標上顯著改善。相較介入前，使用 Therabot 四週後，憂鬱症狀量表（PHQ-9）平均下降 6.1 分，



焦慮症狀 (GAD-Q-IV) 下降 2.3 分，體重關注量表 (WCS) 亦下降 9.8 分；至第八週追蹤期，改善幅度進一步擴大，三項指標分別下降 7.9 分、3.2 分與 10.2 分，明顯優於對照組。Therabot 的效果不僅在短期內顯現，且能於追蹤期持續累積。重度憂鬱症 (MDD) 症狀改善比例達 51%，顯示超過半數患者在介入後獲得明顯緩解；廣泛性焦慮症 (GAD) 則有 31% 的症狀改善幅度，在飲食失調相關的身體意象困擾方面，也觀察到 19% 的改善效果。儘管不同疾病的改善幅度有所差異，整體趨勢一致呈現正向變化，顯示 Therabot 具備跨診斷應用的潛力且具備穩定且具臨床意義的心理健康介入潛力。此研究結果支持 AI 作為數位心理介入工具，在臨床監督與適當設計下，能有效輔助傳統心理照護，為心理健康服務的可近性與多元化提供新的實證方向。

以上內容將在 2026 年 1 月 21 日(三) 10:00 am 以線上直播方式與媒體朋友、全球民眾及專業人士共享。歡迎各位舊雨新知透過[星球永續健康網站專頁](#)觀賞直播！

- 星球永續健康網站網頁連結: <https://www.realscience.top/7>
- Youtube 影片連結: <https://reurl.cc/o7br93>
- 漢聲廣播電台連結: <https://reurl.cc/nojdev>
- 不只是科技: <https://reurl.cc/A6EXxZ>



講者：

陳秀熙教授/英國劍橋大學博士、許辰陽醫師、陳立昇教授、嚴明芳教授、林庭瑀博士

聯絡人：

林庭瑀博士 電話: (02)33668033

E-mail: [happy82526@gmail.com](mailto:happy82526@gmail.com)

劉秋燕 電話: (02)33668033

E-mail: [r11847030@ntu.edu.tw](mailto:r11847030@ntu.edu.tw)