

編者按

隨着人類進入人工智能（AI）時代元年，AI正迅速改變很多人的生活方式和思維習慣。AI就像一面鏡子照出不同人生，而鏡子也有自己的陰暗面。香港文匯報近期推出「AI之暗面」系列，與讀者一起走進AI使用日常，從讓渡決定權到收購臉權，在肯定AI提升效率的同時，警惕AI失控的風險，提醒人們須緊扣AI韁繩，防止被AI異化，反噬人類主體性。

加密貨幣富豪孫宇晨近期在《我的女友景甜》一文中提到，面對一筆5,000萬美元的款項要不要支付，他沒有自己拿主意，而是問了AI，回答是「不要把這五千萬美元給她」。

這不是抖音短劇裏的狗血劇情，而是真實發生在我們身邊的一幕幕情景：從孩子取名到要不要買房、該不該換工作，再到看什麼病、投什麼項目，越來越多人習慣做決定前先問一下AI。當答案越來越容易獲得，人類在不經意間，已把一部分「做決定」的權利交了出去，奪回AI決定權已刻不容緩。

●香港文匯報記者 曾萍 廣西報道

AI進入日常生活時間雖短，但已令人類習慣出現重大改變。QuestMobile數據顯示，豆包、DeepSeek、文心一言三家合計月活超6.79億，每天有30億次提問被敲下。這些提問裏，相當一部分不是「查資料」，而是「求決定」。美國布朗大學對OpenAI平台25億條對話的分析顯示，49%的用戶提問集中在「尋求建議與資訊」——「要不要辭職」「這段關係要不要繼續」。過去人們拋硬幣、問朋友，現在敲進對話框，等AI給出一串有理有據的理由。

AI輔助決策滲透高風險領域

數字背後，是真實發生的「認知拉鋸」。北京大學第三醫院內科診室裏，60歲老人拿着AI開的注射劑處方反問醫生：「AI說瑞百安降脂快」，堅持換掉醫生開的口服他汀，這場爭論持續十分鐘才落幕；2023年，美國紐約南區法院開出全球首張「AI幻覺罰單」：一名律師提交ChatGPT生成的辯護狀，引用的6個判例全屬虛構，律師本人被罰5,000美元……從買榴蓮讓豆包「看圖挑果」，到問AI「該不該給5,000萬」，決策的顆粒度正從日常瑣事向醫療、投資、法律這些高風險領域悄悄蔓延。

香港文匯報記者在內地社交平台發起一項關於AI決定權的投票，參與的網友中逾54%會經常參考AI建議再做決定，僅有9%的人幾乎不借助AI做判斷，AI輔助決策已經成為多數人的日常習慣。也有不少人開始讓AI從日常瑣事延伸到更重要的人生選擇，只是依賴程度各有不同。

曉禾（化名）在主副業兩頭忙時糾結是否還要花時間考取職業證書，她把自身處境、職業規劃告知AI，借助算法拆解問題、梳理核心訴求。對她而言AI更像一面「思路鏡子」，最終決定權始終握在自己手裏。職場人小檸（化名）對AI的依賴度更高，去年面對兩份工作offer難以抉擇時，她將前路交給AI分析，她坦言那次人生決策裏，AI的決定權超過一半。

但把決策過多託付給AI，也容易踩坑。「它從不告訴你它有盲區。」網友阿哲（化名）曾把一件難抉擇的事交給豆包。AI的方案邏輯清晰、建議篤定，他被說服照做了，兩周後卻發現並不合適。

提升決策效率 卻暗藏算法偏見

AI參與決策的優勢首先在於效率。「AI降低了許多壁壘，降低了普通人的信息門檻，並且對許多信息進行了平權化處理。」廣西社會科學院社會學研究所所長姚華指出，面對海量信息，人很難在短時間內完成全面檢索、歸納和比較，而AI可以迅速梳理信息、提供多個選項。不過，姚華強調，需要劃清的是「輔助決策」與「最終拍板」的邊界。

「AI可以向我們提供建議或者觀點，然而是否採納決策還是需要人來做出，而不應該把決策的權利推給AI。」姚華指出，AI幻覺、訓練數據被人為污染的問題始終存在，更值得警惕的是算法自帶的偏見，它不會憑空產生立場，卻會把歷史數據裏的不公固化下來，復刻甚至放大現實裏的階層差異。他特別提到小語種的困境：使用人口越少的語言，AI訓練數據越匱乏，最終只會在算法時代陷入越來越弱勢的循環，「固化原有的不公」。

這種風險落在社會層面，就是集體思辨能力的退化。姚華坦言自己不主張研究所裏的年輕人過度依賴AI，當所有人都習慣接收AI輸出的、邏輯自洽的現成結論，公共討論會漸漸變得模板化，失去了不同觀點碰撞、權衡、辯論的過程。在姚華看來，AI浪潮就像當年汽車取代馬車，風險必然存在，但歷史的前進不可逆。人類會在長期的動態博弈裏，慢慢找到約束技術、制衡資本的規則。而對每個人來說，最核心的底線始終清晰：「AI應該是輔助工具，不要讓它代替我們進行思考。」

「它不會承擔選擇帶來的後果」

孫宇晨當初如果沒有讓AI幫其決定，結局是否會有所不同？無從得知。但把決定的責任推卸給AI，不僅是對自己人生的不負責，更是把自己的未來活成了別人的數據庫，它不會替你承擔選擇帶來的後果，最後的得失、喜悲，都還是落在自己身上。

AI之暗面系列

上篇

數據來源：埃森哲《中國消費者洞察》、麥肯錫消費行業AI研究、中國互聯網絡信息中心《生成式人工智能應用發展報告（2026）》

AI不同領域參與度

美妝消費

AI決策參與度：超80%
AI角色定位：消費決策輔助
風險等級：低



消費電子/旅行行業

AI決策參與度：40%-55%
AI角色定位：購買決策參考
風險等級：中低



醫療健康諮詢

AI決策參與度：使用率較高（問診前參考普遍）
AI角色定位：初步信息參考
風險等級：中高



情感與個人抉擇

AI決策參與度：持續上升
AI角色定位：梳理情緒與處境，幫人提煉真實訴求
風險等級：中



學習教育場景

AI決策參與度：86%（全球學生），24%每日使用
AI角色定位：信息整合與作業輔助
風險等級：中



AI時代由誰做主？ 莫讓算法接管人生

專家警示人類須守住自主判斷底線



香港文匯報AI製圖

AI決策削弱判斷力 人機邊界亟需重新釐清

「榴蓮挑哪顆？」「男朋友這話什麼意思？」「要不要辭職？」當人類習慣把選擇權逐步交付給算法，當AI取代人類做出大部分決定，個體心理與社會運行將會發生怎樣的深層變化？

「首先是『決策肌肉』正在悄悄變生疏。」華東師範大學應用心理學碩士、心理諮詢師余夏茜指出，人做選擇需要經歷界定目標、搜集證據、權衡方案、預估結果的完整過程，「當人用AI做決定，不再反覆練習決策過程，獨立判斷的能力可能變得生疏。」

「機器能給出概率 不能產生權威」

今夏「AI挑榴蓮」刷屏，可實測AI推薦款平均出肉率僅32.1%，還不如水果店老闆憑手感挑的38.7%——AI看不見氣味、捏不到彈性，會把「死包」誇成「報恩果」。余夏茜說，AI能說服人「往往不是更正確，而是語言流暢、結構完整、語氣確定」，讓人把「像個好答案」誤當成「真的是好答案」，本質上就是權威依賴，「權威從人換成了算法」。



●今夏內地民眾流行用AI鑒定榴蓮好壞，AI還會給出具體原因。

比判斷力下滑更危險的，是自我能动性被稀釋。上海一位廣告從業者的十年感情，終結在丈夫用AI給她貼的「邊緣性人格障礙」標籤上：一次忘倒垃圾的爭執，被算法升級為人格判決。余夏茜直言，當AI從輔助變成事實上的決策者，「人仍然可以點『同意』，卻不再感到選擇真正出自自己」，AI算得出收入、風險、穩定性，卻算不出意義、關係和「我願意接受的遺憾」。

從技術邏輯上看，AI本身的特性，決定了它無法承擔人類決策的重量。AI領域博主席梁賽直言：「機器學習能給出概率，但不能產生權威。」梁賽劃出技術硬邊界：「錯誤可檢測、後果可撤銷、責任可歸屬、決策可解釋並接受質證——四條同時滿足才適合交由AI定論；缺一條，就應限定為輔助。」

他提醒，大模型「說服人的能力」是訓練出來的：「預訓練學的是連貫，不是真實；對齊階段獎勵的是好評，不是校準。」那些讓你覺得「好有道理」的話，是被設計成「讓你覺得好有道理」的。一段被算法勸分的感情、一個漏診、一次誤判，都回不了頭。

算力壟斷資本逐利 警惕技術私化

特稿

當人們習慣把購物比價、職業規劃甚至人生困惑都交給AI梳理，主動將決策的參考權一步步讓渡給算法，很少有人退後一步審視：輸出這些結論的核心技術，究竟掌握在誰手中？我們在依賴AI提升效率的同時，也在默許一場靜默的權力轉移：全球頂尖的大模型技術與核心算力正快速向少數私人科技企業集中，而各國公共部門普遍存在AI專業人才缺口，缺少制度防火牆的技術私化，正在成為算法時代最值得警惕的深層隱憂。

公共部門須掌規則制定權

從技術監管的維度看，公私之間的能力差正在讓監管流於形式。AI領域博主席梁賽直言，這種失衡首先帶來「審查空心化」：「政府只能核對模型卡與評測報告，無法獨立復現結論。這比沒有監管更危險，因為它生產了合法性的外衣。」梁賽認為，公共部門的AI人才儲備普遍落後產業一到兩代，神經網絡的內部邏輯連開發者都難以完全釐清，風險標準、整改方案大多由被監管企業自行提供，信息不對稱直接演變為監管的被動，最終形成「企業自己定規則、自己做評估」的偽監管局面。

從社會治理的層面看，資本逐利的屬性會讓算法偏離公共利益。廣西社會科學院社會學研究所所長姚華點出了本質：「資本是逐利的，AI有可能無意或者有意地誤導公眾，做出對資本有利但是對全社會大眾不利的決策。」更隱蔽的是黑箱，當商業偏好悄悄嵌入算法，滲透進信貸審核、崗位篩選、內容分發甚至公共決策場景，普通人往往渾然不覺自己正在被評估、篩選，公共資源與社會機會會在無聲中向資本傾斜。而一旦算法決策出錯，責任會在多方之間推諉，代價最終由公眾承擔。

奪回AI決定權，本質是守住技術的公共屬性。給算法裝上制度防火牆，把規則制定權、最終監督權、邊界劃定權牢牢握在公共手中，技術才能真正服務於人。這才是「奪回AI決定權」的深意：不是拒斥技術，而是守住「最後拍板」的那道閘門。