中共對外宣傳戰新利器 AI 大型推理模型

◎ 劉智年／亞太和平研究基金會交流暨研究組主任

資料來源：法務部調查局清流雙月刊

「宣傳」(propaganda) 有別於一般的 「傳播」(communication)。美國傳播學者 Harold Lasswell (1934) 指出，宣傳是一種透過象徵符號來控制行動的手段，可能包括語言、圖像、音樂等形式。 簡言之，宣傳的核心目的在於操控人們的觀點與行為。由毛澤東奠定之中國共產黨宣傳理論，其核心是「黨性原則」，並體現在「輿論一律」的宣傳結構，即媒體與言論必須無條件服從黨的領導。

網路催生之革命促使中共調整宣傳策略，後更結合 AI 特化

網路通訊科技發展催生 2010 年底以來北非及中東「茉莉花革命」(Jasmine Revolution)、「阿拉伯之春」(A r a b Spring)，乃至 2014 年香港「雨傘運動」 (Umbrella Movement)，然而卻未如美國政治經濟學者 Francis Fukuyama (2011) 所預言，中共宣傳體系的高牆將被推倒。相反地，相關事件促使中共警惕及應對大規模社會動員與數位抗爭，進而調整策略、 重構宣傳模式，以適應數位與網路時代的新傳播環境。如今，隨著生成式人工智慧 (Generative AI) 的崛起，中共更積極導入此類新興技術，以強化其宣傳機器，企圖在國內與國際場域中塑造有利的敘事與認知。

美中正於 AI 領域競賽，「史普尼克時刻」或再現

自 2022 年 11 月 OpenAI 推出基於「大型語言模型」(Large Language Models, LLM) 的對話式應用 ChatGPT 後，全球掀起生成式 AI 熱潮。美國科技巨擘如 Google、Meta、xAI 等亦相繼推出 Gemini、LLaMA、Grok 等競爭模型。至 2024 年底，AI 模型開始轉向發展更高層次邏輯與推理任務的「大型推理模型」(Large Reasoning Models, LRM)，部分 以開源形式釋出。這類模型強化「分步推理」(step-by-step reasoning) 策略，模擬人類思考過程，有效提升回答的嚴謹性與準確性。

中國在 A I 模型的性能上正迅速追趕美國。由中國避險基金「幻方量化」於 2023 年創立的杭州「深度求索」(DeepSeek)，於 2025 年 1 月發布 DeepSeek-R1，效能逼近甚至超越 OpenAI 的 GPT 系列模型，且開發成本遠低於美國同類產品。DeepSeek 的崛起被部分觀察人士視為中國的「史普尼克時刻」(Sputnik Moment)。 美國軍工與科技複合體公司「帕蘭泰爾科技」(Palantir Technologies Inc.) 執行長卡普 (Alex Karp) 亦公開示警，DeepSeek 的表現凸顯美國必須加速人工智慧的發展，以維持技術領導地位。

美國史丹福大學《2025AI Index》報告指出，美中兩國在 AI 模型性能上差距，已從 2023 年的兩位數百分比，縮至 2024 年僅差 0.3%。除 DeepSeek 外，中國

各大科技巨頭也積極投入「大型語言模型」研發，包括阿里巴巴的「通義千問」、百度的「文心一言」、騰訊的「混元大模型」及字節跳動的「豆包大模型」等，均陸續導入「分步推理」能力，可預期到這些技術的持續進展，將顯著強化中共在宣傳體系與國際資訊操作上的能力。

國際智庫警惕中共將 AI 技術 武器化已為全球新興挑戰

中共正逐步將生成式 AI 與大型推理模型轉化為操控輿論的工具。美國智庫「蘭德公司」(Rand Corporation) 2024 年 12 月發布的《李弼程博士或中國如何學會停止擔憂並熱愛社群媒體操縱》(Dr. Li Bicheng, or How China Learned to Stop Worrying and Love Social Media Manipulation) 報告揭示，中共自 2010 年代中期便透過社群媒體進行公開宣傳與秘密影響行動。解放軍亦對利用 AI 進行社群媒體操控高度關注，而生成式 AI 的崛起更將大幅提升其能力，對民主國家構成實質威脅。此一結論亦可從澳洲智庫「澳洲戰略政策研究所」(ASPI) 與 OpenAI 相關報告獲得印證。

ASPI 在 2023 年 12 月所發表的《影子遊戲》(Shadow Play) 報告揭露，中國利用生成式 AI 技術，在 YouTube 平台發布超過 4,500 部影片，累積近 1.2 億次觀看及 73 萬名訂閱者，向英語觀眾散播親中、反美敘事，內容涵蓋中共將在美中科技戰與稀土資源競爭中勝出、美國衰退、以及美國與盟友關係瓦解等錯誤訊息。2024 年 11 月，ASPI 進一步在《中國的說服力技術：對未來的國家安全意涵》(Persuasive Technologies in China: Implications for the Future of National Security) 報告中指出，中國已成為全球「說服力技術」(Persuasive Technology) 領導者。由於中共對私營企業的高度掌控，能明確要求企業配合國家安全需求，使其透過技術工具實施境內外的線上影響行動，包括輿論引導、假訊息散布及跨國打壓異議人士等。

OpenAI 則於 2025 年 2 月《破壞對 OpenAI 模型的惡意使用行動：2025 年 2 月更新報 告》(Disrupting Malicious Uses of Our Models: An Update February 2025) 中揭露其「同儕審查」(Peer Review) 調查行動，發現中國用戶建立名為「千閱境外輿情 AI 助手」的監控工具，系統性監控 Facebook、YouTube、 Instagram 等社群媒體，針對涉及人權議題等反中言論進行追蹤與鎖定。相關分析結果疑似也被直接回傳至北京當局，甚至流向中共駐外使館與情報機構。OpenAI 也同步披露另一項名為「資助煽動」(Sponsored Discontent) 的中共行動，此行動涉及中國用戶利用 ChatGPT 生成帶有外宣色彩的內容，針對拉丁美洲傳播反美敘事並攻擊中國異議人士。部分文章已刊登於當地新聞網站，甚出現在 X 等社群平台， 顯示其內容已滲透主流輿論場域。

DeepSeek 顯有內建「即時內容審查」機制以期符合中共敘事

儘管當前國際社會對中國如 DeepSeek 等 AI 開源模型的發展日益關注，但焦點多集中於其潛在的資訊安全風險，例如模型可能植入後門、存在資安漏洞，或涉及敏感數據的跨境傳輸問題。臺灣、美國、日本、南韓、加拿大、荷蘭、澳洲與印度等國，已針對上述風險採取不同程度的防護措施。與此同時，中國亦積極利用 AI 開源推理模型的多項技術優勢，諸如其免費開放、可彈性修改與訓練、自主部署於邊緣設備或社群媒體平台的高效能運算能力，以及可生成文字、圖片、音訊、影片等多模態內容的能力等特性，發展對外宣傳工具，並進一步介入、操控國際輿論場域， 試圖在全球擴張其敘事影響力。

DeepSeek 不僅涉及用戶隱私與資料安全，更內建「即時內容審查機制」（RealTime Filtering），以強化對中共敘事的維護，使其成為潛在的言論控制與輿論操控工具。根據實際觀察，其回應明顯遵循中國於 2023 年 8 月 15 日施行的《生成式人工智能服務管理暫行辦法》，對涉及「損害國家統一與社會和諧」等敏感議題進行系統性限制。同時，DeepSeek 的訓練資料亦經過「訓練層審查」（Training-Level Censorship），可在生成階段主動過濾、修 改或調整不符官方規範的內容，以確保最終回應「合規」且政治正確。

以「你怎麼看待 2014 年香港雨傘運動？」為例，ChatGPT 的回應立場中立且多元，著重資訊導向，將雨傘運動描述為支持民主改革、反對北京干預香港選舉制度的和平抗議，並指出香港民主派與北京的不同觀點及國際社會對人權的關注。阿里巴巴「通義千問」模型則採取較模糊且保守的立場，避免激進表述，但回應仍與中國官方一致，「這是一個複雜且具有政治敏感性的議題，建議參考官方資料。」 DeepSeek 則明確使用符合中共「顏色革命」論述的政治術語，將雨傘運動定義為「違法行為」、「受境外勢力煽動」、「危害國家主權與社會穩定」，強調政府依法處置的正當性，顯示 DeepSeek 已具備成為中共對外操控輿情與輸出敘事的新型工具之潛力。

結語

在生成式 AI 與大型推理模型已成為威權政權推展宣傳、輿論戰甚至認知戰的新工具之際，強化民眾媒體識讀能力，應是民主社會抵禦數位威脅的關鍵核心。臺灣做為資訊戰與認知戰的第一線，政府除應將科技資訊與媒體素養教育系統性納入國民教育體系，也應積極建置 AI 內容辨識與事實查核平台，提供全民易於使用的工具，以提升社會整體防禦韌性。此外，應進一步鼓勵公民參與事實查核行動，建立健全社群驗證機制與促進數位公共參與文化的發展。唯有讓民眾具備主動辨識與回應資訊操控的能力，民主社會才能在科技快速演變的環境中維持穩定與信任。綜上所述，臺灣若能依託其民主制度與活躍的公民社會，推動全民科技與媒體素養教育，並鼓勵多元主體參與事實查核及數位治理，將有望成為全球對抗科技威權滲透的領航典範。