

DeepSeek 論文登《自然》封面

披露更多模型訓練細節 成本僅228萬



梁文鋒的DeepSeek-R1研究論文登上《自然》封面。

由DeepSeek團隊共同完成、梁文鋒擔任通訊作者的DeepSeek-R1推理模型研究論文，日前登上了國際權威期刊《自然(Nature)》封面，DeepSeek-R1成為全球首個經過同行評審的主流大語言模型。Nature評論指出，依靠獨立研究人員進行同行評審，是應對AI行業炒作的一種方式，希望更多AI公司能夠效仿DeepSeek。

類示例來開發推理步驟，減少了訓練成本和複雜性。

在論文中，DeepSeek明確否認了此前關於其使用OpenAI模型輸出進行訓練的質疑。在同行評審文件中，DeepSeek介紹，DeepSeek-V3 Base (DeepSeek-R1的基座模型)使用的數據全部來自互聯網，雖然可能包含GPT-4生成的結果，但絕非有意而為之，更沒有專門的「蒸餾」環節。為了防止基準測試數據污染，對DeepSeek-R1的預訓練和後訓練數據都實施了全面的去污染措施。

訓練方法革新加速多領域應用

此外，DeepSeek團隊還首次公開了R1訓練成本——僅為29.4萬美元(約合

228.6萬港元)。這個金額即使加上約600萬美元的基礎模型成本，遠低於OpenAI、谷歌訓練AI的成本。

「由梁文鋒為核心加中國人工智能團隊研發的通用推理模型DeepSeek-R1研究成果正式發表，標誌着AI基礎模型研究正從技術博客和基準測試分數評價，轉向以方法論科學性、可復現性和嚴謹驗證為核心的成熟學科。」中國電子信息產業發展研究院副總工程師劉權說。DeepSeek-R1為全球AI發展提供了新範式。它開創僅靠強化學習激發大模型推理能力的先例，突破以往模型依賴人工標註的局限，推動訓練方法革新，加速多領域應用。

「當前AI發展迅猛，公眾對其安全性、可靠性存疑。DeepSeek-R1經過嚴格評審後刊登，公開訓練細節、回應安全關切，增強了透明度與可信度，為AI大規模應用築牢信任基礎，促進產業健康發

展。」劉權說。

劉權強調，DeepSeek取得的成功離不開開源的優勢。開源模式促進了知識共享與技術迭代，為全球開發者提供創新平台，推動AI研究邁向科學、透明、可復現的新階段。從成本角度看，開源免去授權費用，降低研究門檻；靈活性上，開發者可按需修改代碼、調整架構；社區協同則匯聚全球智慧，加速模型優化。開源優勢為AI大模型發展注入強勁動力。「開源策略促進了技術民主化，為推理能力演進指明方向，還吸引更多科研人員投身行業，為AI發展注入新動力。」

針對中國人工智能等未來產業的發展，科技部部長陰和俊18日表示，在AI領域，國內湧現出多個達到國際先進水平的通用大模型，部分模型準確率突破95%，打造百餘個標杆應用場景。



國際學術期刊《自然》封面。

DeepSeek R1披露更多模型訓練細節。

據報道，與今年1月發布的DeepSeek-R1的初版論文相比，本次論文披露了更多模型訓練的細節，並正面回應了模型發布之初的「蒸餾」(模型壓縮和知識轉移技術，指將複雜模型的知識遷移到輕量的模型)質疑。梁文鋒團隊報告稱，DeepSeek-R1模型使用了強化學習而非人

中方倡設國際打擊電騙聯盟

全球公共安全合作論壇(連雲港)2025年大會打擊跨國犯罪形勢與對策分論壇18日在江蘇省連雲港市舉辦。中方倡議相關國家和地區共同建立國際打擊電信網絡詐騙聯盟，推動各方和國際社會攜手應對電信網絡詐騙犯罪治理問題。

6.8萬名境外涉詐疑犯歸案

本次分論壇以「攜手打擊跨國犯罪，共同維護全球公共安全」為主題。中國公

安部有關負責人介紹了中方攜手各方打擊跨國電信網絡詐騙犯罪取得的積極進展，特別是近年來中方先後與西班牙、阿聯酋、緬甸、印度尼西亞、菲律賓、老撾、泰國、柬埔寨等國開展執法安全合作，共有6.8萬名境外涉詐犯罪嫌疑人成功歸案。

中方指出，當前刑事犯罪結構發生重大變化，傳統犯罪加快向網上蔓延變異，以電信網絡詐騙為代表的新型犯罪已成為世界公害和全球性打擊治理難題。各國執

法部門應進一步強化國際執法合作，完善辦案協作機制、健全全球打擊治理體系。

分論壇期間，與會各方充分交流打擊跨國犯罪的現狀、經驗及舉措，特別是圍繞「電信網絡詐騙犯罪打防對策」「深化國際執法安全合作、共同打擊跨國犯罪」「建立國際打擊電信網絡詐騙聯盟」等議題深入交換了意見。分論壇發出《聯合打擊治理電信網絡詐騙犯罪倡議書》，得到與會嘉賓積極響應。

9月18日，是九一八事變爆發94周年，亦是長春航空展前一天，大國重器殲-20、殲-16、殲-15首飛試飛員首次相會長春航空展，講述中國航空裝備研製、試飛等階段背後的故事。

殲-20首飛試飛員李剛表示，如今中國強大的航空國防力量不是一蹴而就的，是成千上萬科研工作者幾年、幾十年夜以繼日奮戰出來的。

談及殲-20首飛，李剛介紹，為確保首飛節點，團隊經常工作至深夜。「有時候實驗會持續到半夜一兩點鐘。」道出無數科研人員默默付出的艱辛過程。

殲-15、殲-16首飛試飛員李國恩介紹，中國自主研發的艦載戰鬥機殲-15首飛成功，源於整個團隊多年的技術儲備。「一架飛機從方案構思、圖紙設計、生產製造到最終飛天，需要耗費大量時間，且依賴國家多方面的準備和保障，是綜合國力的體現。」

長春航空展19至23日舉行

「殲-35身披『隱身外衣』，作戰體系適應性更強，未來升級空間更大。」李國恩如此評價中國隱身艦載戰鬥機。他同時透露，艦載機試飛有諸多難點需要攻克，僅尾鉤設計就要攻克材料難題，以承受着陸甲板的巨大衝擊力。

2025空軍航空開放活動和長春航空展將於9月19日至23日舉行。

殲擊機試飛員分享艱辛研製過程

毋忘九一八

昔日，香港學生在初中、高中和預科的中國歷史科(簡稱「中史科」)和歷史科(俗稱「西史科」)，對抗日戰爭的歷史過程，都有或詳或略的涉獵學習。

在中史科，對於1931年9月18日的九一八事變起，直至1945年抗日戰爭勝利結束的抗日戰爭內容，有比較完整的教授。然而，在「西史科」，由於採用歐洲為主的史觀，將二戰放在1939年至1945年這六年時間作片面的理解。結果，以往該科將抗日戰爭稱為Second Sino-Japanese War，1931年至1937年7月7日蘆溝橋事變前的抗戰歷史被忽略了。這導致香港學生長期認

為1941年12月8日珍珠港事件後，美國參戰才起決定性角色，而忽略了中國在東方主戰場中所起的中流砥柱作用。

九一八事變的重要性在哪裏？它打響了中國人民十四年抗戰的序幕。在1931年至1941年珍珠港事件前，中國在亞洲獨力抵抗了日本法西斯主義入侵達十年之久。1937年至1945年間，中國牽制超過50%至最高94%的日軍，使之無法北攻蘇聯、南侵亞太地區。因此，不能以1945年8月蘇聯進攻關東軍、美國在日本投下原子彈，而否認中國作為二戰戰勝國的資格！

近年來，關於東江縱隊港九大隊的歷史梳理、相關歷史遺蹟和博物館的建立已有不少進展。然而，我們不

蔡思行 學研社成員 歷史研究者

要忘記1931年至1941年這段與香港市民似遠實近的歷史。當年香港雖受英國殖民統治，但九一八事變發生後三天，香港報章均以「我國軍民」等措詞，表達了香港華人和內地同胞同民族同命運的強烈關注。幾日後的中秋節，更發生反日騷動。此後，香港華人或罷買日貨，或捐款支援遭受戰禍的同胞，表現了愛國的情操。

今日國際上，右翼極端思想和法西斯主義有明顯重燃的跡象。我們觀看《南京照相館》、《東極島》等抗戰主題電影，能認識到平民既是戰爭的受害者，也是抗戰的參與者。牢記抗戰歷史，不是為了記恨，而是和世界同道共同珍惜來之不易的和平。