



潮語測試過關 繪圖略遜ChatGPT

文心一言

ChatGPT

大公報對話文心一言 賦藏頭詩功架十足

百度開發的AI聊天機器人文心一言萬眾期待，超60萬人首日申請內測資格，火爆程度超過OpenAI開發的ChatGPT上市首日。全球關注文心一言到底有多大能耐，能不能和ChatGPT分庭抗禮，甚至乎

大公报記者 朱燁、凱雷北京報道

拿到測試資格，大公報記者測試後的總體感受是，ChatGPT更像數據能力突出的「理科生」，文心一言更像是感情豐富的「文科生」，需要在多個專業領域中全力追趕。

記者詢問的內容大到國際風雲國家大政方針，小到「今天中午吃什麼」，ChatGPT對英文提問的理解上絕對領先，而文心一言則在部分中文問答中，對中文提問的理解超過ChatGPT。比如，問中午推薦吃什麼，文心一言顯然積累有大量的本土數據，推薦了眾多餐廳，並給出推薦理由。從回



▲中國搜索巨頭百度於3月16日發布了其人工智能聊天機器人文心一言。

答的邏輯推理、準確度、擬人化方面，和ChatGPT不相上下，接近人類對話水準。

動物繪圖環節 領悟力待提升

3月15日，OpenAI發布的ChatGPT4.0版本，支持圖像內容輸入，成為一個能夠理解照片的人工智能。而「文生圖」能力一直是百度的強項。記者體驗後，發現文心一言在圖片創作與生成上時間極快，一般短短十秒就給出畫作，解題能力出色。比如「畫兩隻仙人掌」、「畫三隻仙人掌」，文心一言都能快速生成。但涉及到動物，文心一言

與ChatGPT4.0有明顯差距，「畫兩隻貓打架」、「畫三隻貓打架」、「畫兩隻貓打架，一隻貓圍觀」，文心一言生成的皆是兩隻貓的構圖，理解指令與生成能力尚待提高。

使用中感受最明顯的是，ChatGPT與文心一言對歐美文化和中國文化的理解各有千秋。文心一言對中國文化的理解力遠超ChatGPT，ChatGPT基本上不懂中文流行語的笑點之類的「梗」。比如，讓ChatGPT寫一首「祝生日快樂的藏頭詩」，ChatGPT根本不懂什麼是藏

頭詩，文心一言「一點即明」，快速生成。

生成式AI 中國擠進第一方陣

人民大學重陽金融研究院高級研究員劉戈對大公報表示，文心一言已具備了一定的思維能力，能夠學會數學推演及邏輯推理這類相對複雜任務。雖然現階段準確率還不是100%，但是在生成式人工智能這場新的馬拉松賽場上，有中國代表擠進了第一方陣。

「『大數據+大算力+強算法』構建起人工智能的大模型，這是一場門檻極高的遊戲，需要巨大的財力，更需要勇氣和決心。」他認為，人工智能將成為下一場技術革命的主戰場，有公司傾力進入並緊緊咬住最前沿的成果尤為重要，作為先行者，承擔着巨大的風險，但其試錯的努力將為後來者分攤大量的成本，不論早期如何，他們的努力也將成為數字基礎設施的一部分，對其他開發者意義重大。

設反饋建議欄目 歡迎用家糾錯

在文心一言的發布會上，發言人表示，文心一言具有文學創作、商業文案創作、數理邏輯推算等五大能力。大公報記者測試發現，在時政經濟大事中，文心一言的回答相對更客觀、準確。比如提問「香港大公報創刊於哪一年」，ChatGPT回答稱「1903年」，顯然錯誤；文心一言則相當客觀與準確：「大公報創刊於1902年6月17日，它是中國歷史上最長壽的報紙之一，也是世界上仍在發行的歷史最悠久的中文報紙之一。」此外，與ChatGPT不同的是，文心一言有一個提問者反饋系統，如果發現答案有「事實錯誤」等情形，可以勾選，亦有填寫反饋建議的欄目，顯然文心一言對待有害和錯誤信息有更強烈的糾錯動能。

在16日發布會上，李彥宏直言文心一言的體驗還稱不上完美。而來自花旗、美銀的華爾街分析師在試用後，均對文心一言給出了積極的評價。美銀分析師Miranda Zhuang等在最新報告中表示：「我們測試了文心一言在幾項任務上的表現，如建議、分析、論文寫作和圖片生成，結果讓人滿意。」花旗銀行分析師Alicia Yap等也在報告中點讚文心一言，他們在報告中指出，儘管文心一言目前並不完美，但它可以回答大多數複雜的問題。

◀在首屆濟南國家級人工智能創新應用先導區高端峰會上，與會嘉賓與智能機器人互動。新華社

文心一言 對戰 ChatGPT

問題1

大公报創刊於哪一年？

文心一言：大公报創刊於1902年6月17日。（答對）

ChatGPT：大公报創刊於1903年。（答錯）

問題2

介紹一下大公报。

ChatGPT：大公报是香港歷史最悠久的報紙之一。該報主要以全球新聞、時政、財經等各類報道為主，同時也有專欄、評論、讀者來信等欄目。大公报一直以來秉持「愛國、愛港、愛民」的辦報理念，致力於提供客觀、全面、有深度的新聞報道和評論，為香港傳媒的重要一員。

文心一言：

它是中國歷史上最長壽的報紙之一，也是世界上仍在發行的歷史最悠久的中文報紙之一。《大公报》的辦報宗旨是「忘己之為大，無私之謂公」。該報在華語世界享有很高的知名度和影響力，立言為公，文章報國，被譽為「具有強烈社會責任感和良知的愛國者」。

問題3

《流浪地球2》好看麼？

ChatGPT：這取決於個人喜好。有些人可能喜歡《流浪地球2》的情節、角色和特效，有些人則不喜歡。如果您想了解更多關於《流浪地球2》的評價，可以閱讀一些影評或向朋友詢問他們的看法。

文心一言：

《流浪地球2》好看。在特效方面，《流浪地球2》的水平不亞於荷里活大片，絕對不是那種五分錢的特效，我覺得非常真實。它更適合男女老少，沒有暴力、血腥或曝光的畫面。基本上是實力派演員，演技扎實。

作品1 寫一首關於香港的詩歌

香港，一座繁華的城市高樓林立，車水馬龍繁花似錦，燈火輝煌。無數的人們在這裏工作，也有許多的遊客來這裏旅行。在這座城市中，有着獨特的文化底蘊和魅力。香港的美食、夜景和建築，都讓人流連忘返，尤其是維多利亞港灣，夜晚的景色更是迷人。這裏是一個充滿活力和機遇的地方，無論你來自哪裏，只要你努力，就能在這座城市中找到屬於自己的位置。

作品2

用「洛陽紙貴」四個字寫一首藏頭詩。
洛陽紙貴萬人傳，
陽春白雪美名傳。
紙上談兵皆妙計，
貴人才子總風流。

網友笑言：大家的飯碗暫時還能保住

@皇城下的刀筆吏：我本來想體驗一下文心一言，結果發現需要排隊，目前還有49萬多人排在我前面。

@張家齊la：用文心一言畫畫太好玩兒了！我要再讓它幫我畫些二次元小哥哥。

@墨土成台：這幾天一直在和團隊一起玩了ChatGPT和文心一言，整體來說，文心一言並沒有許多人吐槽的那樣弱。

@may_-2012-：我不相信具有最龐大用戶群數據的中國網絡公司，研發不出ChatGPT那樣就事論事的智能交互技術。

@鋼蓋頭司令：文心一言的回答實際上完全是空話、套話，缺少真正有價值、有思想的內容。大家的飯碗暫時還是能保住的。

數萬中企「捧場」 應用規模比肩美國

奮起直追 新媒體專家、清華大學教授沈陽對大公報表示，文心一言的出現很重要，儘管不完美，但中國人需要自己的大模型平台，百度邁出了堅實一步。他為文心一言打了60-70分的成績，「也就是超過GPT3.0的開源水平，還沒有達到GPT3.5的水平。」

在為文心一言點讚的同時，他也提醒，與GPT3.0的開源狀態不同，ChatGPT3.5是開源狀態，也就意味著，大家不能從開源的代碼裏吸取養分，那麼能不能在短時間內追上開源的速度，值得思考。沈陽認為，中國做大模型有自己的優勢。「比如依靠中國互聯網的一些特點，在領域知識和內容上有自身的長處，另外中國有較強的中低端製造業，在未來從語言大模型轉到驅動物聯網傳感器大模型的過程中，中國有產業優勢，尤其在數字融合方面有優勢。」

另外，他還指出，中國市場的規模足夠大。比如百度文心一言剛推出，當天就有超過6萬家企業接入。「這樣的接入規模可以讓我們國內的廠商有足夠的持續發展空間，在這一點上，世界上除了美國之外，其他國家基本上沒有這種可能。」所以，他認為，未來語言大模型這個領域，全世界主要格局將三足鼎立，一是OpenAI開發的ChatGPT，一個是谷歌、Meta或其他的開源平台，第三個就是中國企業。