

chatgpt不是萬能，AI對話式機器人正確性存疑

刊出日期：2023/11/07 | 文字：林聖恩 | 責任編輯：錢品臻

全文共2503字，閱讀大約需要4分鐘

隨著AI技術的日新月異，其正確與準確性逐漸成為人們關注的焦點，能否全然相信AI生成的內容與如何查證將是人們在未來使用AI工具需要注重的事，也將是本篇探討的重點。

AI生成內容的崛起：挑戰與機會

2023年4月，中國甘肅省警方偵破了一起利用ChatGPT編造假新聞的犯罪事件，該新聞聲稱有列火車撞到修路工人，導致9人死亡。經警方調查發現，相同的內容被21個帳號發布，且瀏覽量超過1.5萬次，而這些帳號全都屬於深圳的一家自媒體公司。被拘捕的公司法人代表承認，為了賺取流量，他使用ChatGPT修改熱門時事新聞元素，製造不同版本的假新聞，並透過批量發布以躲過百度內容審查。(資料來源：[數位時代](#))

此類ChatGPT編造假新聞的事件並不少見，內容生成式AI的發展使人們能夠快速且大量地生成一篇又一篇的文章，但這些發布者不經思考、接收者難以追溯源頭的資訊卻造成隱憂，引起社會對於AI生成內容可信度的廣泛質疑，在這個充滿倫理考驗的時代，如何保持對資訊的掌控與判斷也變得至關重要。

▲ ChatGPT 登入畫面。(圖片來源：Unsplash 攝：Jonathan Kemper)

自2019年微軟投入數十億美金研發聊天機器人開始，人工智慧領域掀起了一場革命，而其中最具有代表性的工具就是OpenAI開發的ChatGPT。從作家在創作小說中使用ChatGPT獲得靈感，到學生在論文和課堂作業中尋求協助，ChatGPT開始成為各行各業執行各種任務時的得力助手。這個聊天機

器人在2022年11月問世，在僅僅兩個月內就迅速吸引了一億活躍用戶，這樣的成就超越了其他熱門應用程式，例如TikTok和Instagram，刷新了用戶增長的歷史紀錄。

然而，ChatGPT的應用也引起了一些問題，例如，儘管其資料庫龐大且擅長模仿人類寫作，但它卻缺乏情感和合理性，因此，當這些生成的內容被廣泛地運用在新聞中，無事實根據的虛假資訊便透過新聞被快速傳播，對社會造成困擾。

AI聊天機器人的崛起也在技術上引起了一些爭議。2023年3月，為了與OpenAI開發的ChatGPT抗衡，Google推出了一款AI聊天機器人程式——Bard，然而，Bard在回答問題時卻出現了準確性問題。它指出韋伯太空望遠鏡拍攝了太陽系外行星的第一張照片，事實卻是由歐洲南方天文台（ESO）的甚大望遠鏡（Very Large Telescope）在2004年所拍攝。這樣的失誤，導致其母公司Alphabet股票狂瀉8%，股價也連帶下跌。於此同時，ChatGPT也受到了一些質疑。Buzzfeed與其合作後指出，OpenAI難以辨認文章來源，甚至連聖經和美國獨立宣言的段落都被誤認為是由AI生成的。這一系列事件引起大眾對AI技術的擔憂，人們擔心它可能被用於生成假帳號，進而騙取個人資訊，甚至可能使選舉受到假資訊的影響、社交媒體的疏忽導致種族滅絕危機等，而AI聊天機器人的出現可能進一步加劇這些問題，特別是當生成的文本用於新聞傳播時，假資訊四處傳播更是一發不可收拾。

除了假資訊外，AI聊天機器人還可能生成不存在的科學研究。Men's Journal使用AI生成以專家口吻編寫的文章，它聲稱引用學術文獻，但卻被醫學主任查出諸多錯誤，誤導讀者。此外，AI還存在抄襲問題，CNET發現至少77篇AI生成的文章涉及剽竊。

在當今資訊大爆炸的環境中，AI生成內容的傳播速度和影響力日益增大，我們需要更主動地辨別資訊，以避免受到虛假資訊的影響。AI生成內容如何進入主流媒體？它在短時間內如何傳播？這是訊息盲信還是用戶對AI生成內容的無知所致？這些也都成為AI世代下值得深入思考的問題。（資料來源：[卓越新聞電子報](#)）

面對 AI 生成的虛假消息

AI技術的進步使得自動生成文本變得更加容易，但這也帶來了虛假訊息的風險。當我們與ChatGPT等AI聊天機器人對話時，我們經常被其一本正經的陳述所迷惑，很難快速察覺其中的錯誤，這讓我們不禁思考，如何確保AI生成的內容是正確的，並且應該對AI的話語保持質疑的態度。

然而，我們不需完全否定AI生成內容的可信度，隨著技術的進步，AI生成內容的正確性也逐漸提高。而我們是否能夠通過更先進的工具和方法來協助判斷AI生成內容的真實性？現今市面上的許多工具已具備事實查核的能力。

如AI Fact Checker 中的 Debunkd，它是一款 Chrome 擴充程式，可以幫助使用者判斷文字敘述的正確性。它的介面簡潔直覺，擁有文字糾錯以及 AI 圖片協助判斷功能。不過，這些工具也面臨著一些困難，因為它們多半依賴於大型數據庫的結果，而這樣的結果本身也可能是由 AI 生成的。這種無窮循環使得我們難以確定糾錯工具提供的訊息是否真的正確。因此，雖然這些工具能作為一種參考，但我們不能完全依賴它們。

有些小工具可以提供幫助

儘管 AI Fact Checker 工具存在侷限性，但還是有一些小工具可以用來簡單糾錯文字型 AI 生成的內容。例如，一些大型科技公司推出的 AI 聊天機器人，如 Bing Chat 和 Google Bard，可以作為一種交叉驗證的手段。通過查看這些不同 AI 機器人的回答，我們或許能夠發現其中的一些差異，從而更好地判斷 AI 生成內容的正確性。

然而，這些小工具和機器人也並非百分之百可靠，它們仍然是基於目前的技術水平而存在一定的缺陷。因此，在使用這些工具時，我們需要謹慎思考，不能完全依賴它們的判斷。（資料來源：[遠傳心生活](#)）

數位時代的真假

在這個不斷發展的領域中，AI生成內容的進步和挑戰同時存在。

隨著AI的不斷發展，人機界線變得模糊，人工智慧雖然有提高工作效率的潛力，但同時也帶來了一系列的問題，包括真實可信的問題以及對人際關係的考驗，傳統的圖靈測試似乎已經不再適用，人們難以根據對話判斷對方是人還是機器，這增加了假消息和帶風向的風險。加強技術監管、提高用戶的訊息識別能力，或許可以在這場抗衡假消息的戰爭中取得更多進展。保持對訊息的警覺，成為每一個用戶在資訊時代下的重要使命。

綜上所述，儘管AI聊天機器人的技術發展在工作上、生活上都帶來便利，但其應用仍存在許多問題和挑戰，除了期待AI的發展必須在監管和倫理框架的指導下進行，我們也應該更謹慎地使用這些技術，以確保生成的內容是真實、準確且不帶偏見的。

© 2023 All rights Reserved

[AI上工·編劇罷工？](#)