

AI「產生幻覺」致假消息氾濫

舊詞獲新解 入選年度熱話

【大公報訊】綜合《衛報》、美聯社、彭博社報道：人工智能（AI）全球爆紅，英國劍橋詞典15日將「Hallucinate」，即「產生幻覺」一詞評選為本年年度詞彙。劍橋詞典稱，隨着AI的發展，該詞於今年獲得了新的含義：AI在處理信息時亦可能會「產生幻覺」，從而「編造」出一系列看似具有說服力的錯誤信息，恐導致假消息氾濫。

「Hallucinate」的原始詞義是「看到、聽到、感覺到或聞到一些不存在的東西」的體驗，通常是由於健康狀況或服用藥物所產生。該詞今年獲得了全新的含義，即指代ChatGPT等AI系統產生「幻覺」，生成虛假信息的情況。

劍橋詞典在其官網稱，之所以選擇「產生幻覺」一詞作為年度詞彙，是因為新含義「觸及了人們談論AI的核心」。ChatGPT等生成式AI，是一種「強大」的工具，但「遠不完美」，人們仍在學習如何安全有效地與AI互動，而這意味着必須要先意識到AI的潛在優勢和其當前的弱點。

另外，劍橋詞典亦正式收錄了一系列與AI相關的詞彙條目進入詞典，包括大型語言模型（LLM）、生成式AI（GenAI）和GPT（生成式預先訓練轉換器的縮寫）等。

ChatGPT現「幻覺」幾率約3%

ChatGPT等聊天機器人爆紅，但不斷被外界發現，其實AI系統也會「一本正經」胡說八道，可能會對原文增刪、竄改，甚至杜撰內容，這一過程通常被稱為AI「幻覺」。例如，今年2月，谷歌發布AI聊天機器人Bard之際，其宣傳視頻中言之鑿鑿地聲稱，韋布望遠鏡拍攝了首張太陽系外

行星的照片。然而該說法實際是錯誤的，第一批系外行星的照片實際是由歐洲南方天文台的甚大望遠鏡（VLT）所拍攝。這個杜撰錯誤，一度讓谷歌市值蒸發1000億美元。

據統計，目前流行的AI模型，都有可能產生「幻覺」，其中ChatGPT產生「幻覺」的幾率約為3%，而谷歌的用於社交平台AI的Palm-Chat模型則擁有高達27.2%的發生AI「幻覺」幾率，為主要AI模型中最差。

需要批判性思維

目前，AI產生「幻覺」的原因目前仍未清楚，因為其背後的大型語言模型擁有海量數據，訓練過程極其複雜，難以確定具體的出錯環節。有研究人員認為，最主要的問題出在訓練AI算法的數據質量上。過時、低質量的訓練數據在訓練過程中，可能誤導AI模型。

隨着越來越多的領域都開始在工作中使用AI，這可能會促進假信息的傳播。專家指出，當人們利用AI技術尋找可能影響其健康、投票意向和其他潛在敏感話題的答案時，AI「幻覺」更會帶來風險。

有人已經成為AI「幻覺」的受害者。今年4月，美國喬治華盛頓大學法學院教授特利出現在性騷擾學生的教授名單中，是ChatGPT給出杜撰的答案造成。澳洲赫本郡郡長胡德被ChatGPT誣賴他曾因賄賂而坐牢，他一度提出要控告OpenAI誹謗。

另外，AI工具越來越強大，專業人士在疏忽下，也會對AI生成的假信息難辨真假。5月，美國紐約一名執業30年的律師施瓦茨被發現法庭文件中援引的多宗案例和文章內容均為虛假。施瓦茨承認自己曾使用ChatGPT準備材料，並稱曾向AI多次確認信息的真實性；本月，巴西的一名法官更因在使用AI撰寫裁決時，在其中引用了具有錯誤細節的法庭案例，而被當局調查。

劍橋詞典表示，AI「幻覺」的存在提醒着人們，人類仍然需要將批判性思維，運用到AI工具的使用中，人類自身的專業知識「比以往任何時候都更加重要」。

AI預測天氣 90%結果勝過傳統方法

【大公報訊】據《金融時報》報道：谷歌旗下的人工智能公司DeepMind取得新突破，其AI天氣預報AI模型GraphCast在準確預測未來3到10天的天氣方面，已經首次超越傳統方法。

根據DeepMind開發人員14日在《科學》雜誌上研究文章，GraphCast在預測包括不同大氣層的溫度、壓力、風速、風向以及濕度等1380個指標中，當中90%的結果優於歐洲中期天氣預報中心（ECMWF）運行的傳統預報系統。GraphCast根據當前時間和六小時前的全球大氣狀態，通過雲端電腦計算，能夠在一分鐘內產生未來10天的天氣預測。而歐洲中期天氣



▲人工智能天氣預報系統GraphCast提前9天預測到了颶風「李」登陸加拿大。網絡圖片

預報中心和多國氣象局目前使用的傳統方法，則是使用超級電腦根據大氣物理學進行計算，完成相關預測則需要幾個小時。

GraphCast使用一種圖像神經網絡（GNN）的機器學習架構，以ECMWF天氣預報系統過去40多年的歷史資料作為訓練模型。ECMWF機器學習協調員錢特里表示，經過訓練的GraphCast運行成本非常低，能源消耗可能降低一千倍。

DeepMind研究人員表示，「標誌着天氣預報領域的一個轉折點」，提到了9月北大西洋的颶風「李」，作為成功預測的範例。GraphCast能夠在颶風登陸加拿大新斯科舍省前9天作出正確預測，而傳統方法僅能提前6天。不過，研究人員強調，暫時AI系統還不能完全取代傳統方法，但可以作為輔佐使用。

TikTok在澳洲成求職利器

【大公報訊】據《衛報》報道：越來越多澳洲的年輕人，改為使用社交媒體平台求職。連鎖酒店希爾頓澳洲在社交媒體上發布視頻，鼓勵求職者在申請該酒店職缺時放棄傳統的履歷，轉而發布TikTok作為申請材料。

希爾頓澳洲在其TikTok賬戶@hiremehilton上發布了一段34秒的視頻，要求有興趣申請職位的候選人，製作一個1分鐘或更短的TikTok視頻，來展示自己將如何為希爾頓的客人留下難忘的住宿體驗。申請人需要從他們的TikTok賬戶公開發布內容，並添加主題標籤#hiremehilton。酒店招募團隊將在TikTok上對影片進行評



▲希爾頓澳洲要求求職者上傳TikTok影片，替代傳統簡歷。網絡圖片

估，然後與入圍的候選人建立聯繫並邀請其參加面試。

希爾頓澳洲地區人力資源總監霍

格表示，該公司選擇TikTok試點是為了吸引從小接觸互聯網的Z世代（1990年代末出生）員工。她補充稱，當酒店需要具有良好人際交往能力、能夠處理賓客關係，以及其他方面事物的人時，僅從紙面上「不知道他們是否有能力」。

希爾頓委託Randstad公司進行的一項調查顯示，大多數澳洲Z世代求職者正在轉向TikTok、Instagram和Snapchat等平台。該研究發現，每四名35歲以下的調查對象中，就有一人通過社交媒體平台成功找到了工作。對於不使用TikTok的申請者，希爾頓表示仍將接受傳統的書面簡歷。



塑膠包裝污染環境 紐約州起訴百事

【大公報訊】綜合路透社及《紐約時報》報道：美國紐約州檢方周三對飲料巨頭百事公司提出起訴，指控百事製造的塑膠垃圾危害環境，並抨擊百事就環境問題對公眾發表「誤導性」聲明。

紐約州檢察長詹姆斯（Letitia James）在紐約州最高法院提起民事訴訟，尋求法院就百事公司對紐約水牛河造成「公害」一案做出裁決。詹姆斯請求法院對百事進行罰款並要求其賠償損失，要求百事停止銷售未標明對環境有害警告的一次性塑料產品，還要求該公司為生產產品在布法羅地區造成的污染提供補救措施，減少百事公司塑料產品進入河流的數量。

起訴書指出，檢方認為，百事生產

的塑料包裝產品是水牛河上最大的塑料污染源，其污染規模是第2名麥當勞的3倍。詹姆斯稱，百事不負責任的包裝及銷售塑料產品，對公眾健康、環境以及水牛河的供水造成了廣泛傷害，而在人類和魚類的體內都發現存在塑膠微粒。起訴書還表示，百事的塑料包裝恐產生健康相關的問題。研究人員認為，人類接觸塑料添加劑和塑膠微粒，可能面臨生殖器官功能改變以及更高的致癌風險。檢方還提到，百事宣稱減少包裝中非再生塑料含量，實際上增加了非再生塑料的使用量，涉嫌欺騙消費者。百事發言人表示，非常看重減塑以及有效回收，並會努力與當地社區進行合作。



▲美國紐約州水牛河上漂浮大量的塑膠廢物。網絡圖片