

AI技術提升 恐助長假消息氾濫

外界籲加強監管 硅谷不願配合

以假亂真

【大公報訊】綜合美聯社、彭博社、《華盛頓郵報》報導：近年來，人工智能（AI）技術不斷進化，已經能寫出媲美人類創作的文字，甚至生產照片和短片。基於人工智能的圖像合成技術「深度偽造」（Deepfake）近年也發展迅速，尤其是用於視頻或照片「換臉」，外界愈來愈擔憂這類技術會被有心人用來散播假消息，對社會帶來更大的影響。科學家也呼籲科技公司可對這些技術開源，以便進行監管和糾錯，但科技巨頭們為了保持技術優勢，多數不願配合。



6月24日，有人利用人工智能「換臉」技術假扮烏克蘭首都基輔市長，與德國柏林市長進行視頻通話。美聯社

隨著科技發展，人工智能技術在機器學習和算法提升下，目前已能勝任多種文本的寫作。在天氣預報、空氣質量報告、醫療報告、財經及體育新聞領域，人工智能「寫手」身影早已存在。其核心技術歸功於「自然語言生成引擎」，系統可以從網上抓取數據，完成分析以及句子的實現。

AI寫詩信手拈來

以微軟旗下OpenAI公司新一代人工智能系統「GPT-3」為例，它通過分析數以千計的數字書籍、維基百科和社交媒體以及整個互聯網的龐大資料庫，現已能寫出類似於人類寫作的文本。過去人工智能用於回答簡單問題和翻譯文本，但如今的「GPT-3」系統，已經可以與人交談，創作出令人信以為真的文本、照片及短片。GPT-3系統能寫出符合客戶要求的大多數文本，例如申請動物飼養員工作的求職信，或者莎士比亞風格的十四行詩詠嘆火星。

紐約大學名譽教授、Robust.AI創始人馬庫斯看來，「儘管目前人工智能非常適合用來做廣告識別、語音識別及為圖片加標籤等等工作，在這方面它的效率與人類無法比擬」。

AI初創公司Hugging Face的研究工程師勒斯考說，通過深度學習，人工智能寫作能力近乎媲美人類，但駕馭事實方面就沒有那麼高明，寫出來的文本看似連貫，幾可亂真，但會出錯。波莫納學院教授史密斯舉例稱，假如問GPT-3一個簡單但無意義的問題，例如談「上樓梯」，「是啊！你事先洗過手，就可用手爬上樓梯」，完全答非所問。

但是，隨着人工智能的寫作能力

提升，亦引發了散播虛假消息上的爭議。人工智能「快速化、批量化」的特點，已經成為內容農場（Content Farm）提高生產率、騙點擊率的最佳助手。此外，可以「換臉」的「深度偽造」技術，所引起的爭議更加廣泛。2018年，「內容農場始祖」美國網媒BuzzFeed就使用後期製作以及人工智能換臉軟件，重新為前美國總統奧巴馬配音，讓「奧巴馬」在片中大罵時任總統特朗普是笨蛋。

自俄烏衝突以來，利用人工智能技術製造的假新聞層出不窮，包括假冒烏克蘭總統澤連斯基投降的視頻，社交媒體廣泛流傳。上月24日，更有人利用「深偽」技術假扮成基輔市長克利欽科，先後與歐洲三個城市柏林、馬德里和維也納市長進行了視頻通話，但最後被識破。

歐盟新例監管「換臉」技術

外界早就對人工智能技術發展表示擔憂，擔心若被有心人利用，將助長假新聞猖獗，使得社會更加恐慌。科學家呼籲科技巨頭們將其人工智能系統開源，以便糾錯，但科技巨頭為了保證技術優勢，向來對算法諱莫如深，不願配合。

部分地區的監管部門已開始行動。歐盟在6月下旬公布了新規定，要求科技巨頭們加強打擊假消息，尤其是針對「深偽」技術，否則將面臨巨額罰款風險，最高可被罰全球營業額的6%。在法國政府資助下，勒斯考率領一群AI研究員組成「大科學」聯盟，對人工智能系統進行糾錯。

被人類帶壞 AI也種族歧視

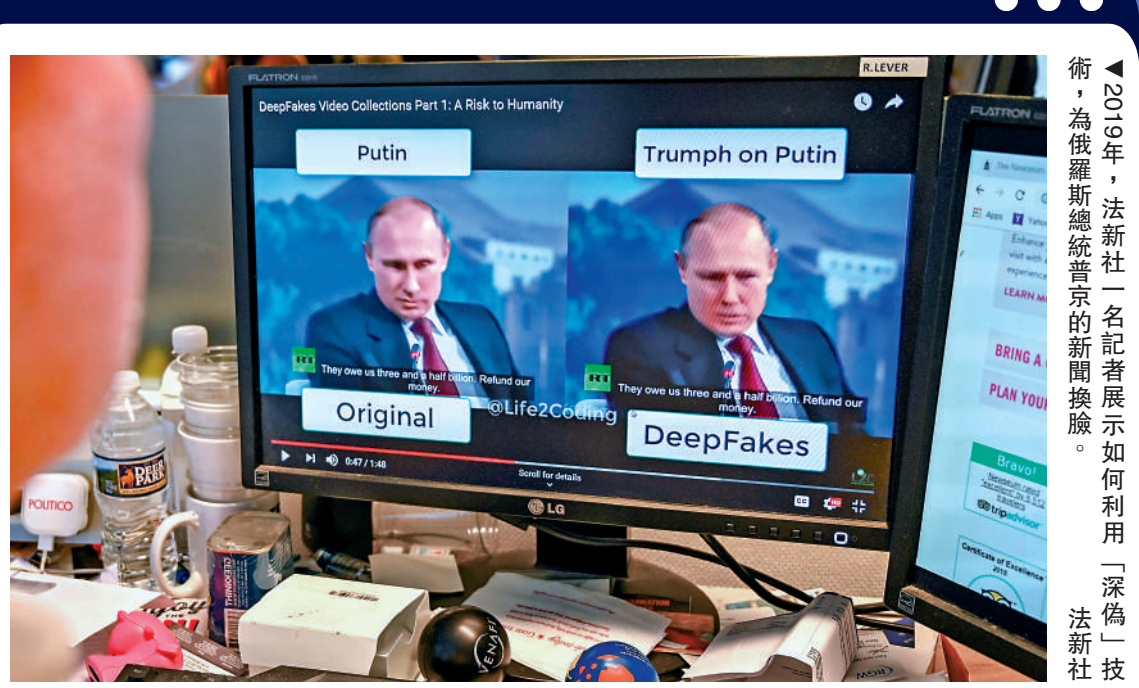
【大公報訊】據美國《華盛頓郵報》報導：談到機器人，人們認為它們可能相對於人類更客觀和中立，但近期一項研究發現，人工智能機器人在辨認人臉時，其實也會出現種族歧視與性別歧視的現象。

美國約翰霍普金斯大學與佐治亞理工學院合作研究報告發現，當工程師要求人工智能系統，從給既定的人臉圖像中找出「家庭主婦（夫）」時，非裔和拉丁裔女性被選中的幾率，明顯比白人男性要高。在識別「罪犯」時，非裔被選中的幾率比白人高近10%。當要人工智能系統找出「清潔工」時，選出拉丁裔男性的數量明顯

偏多；當要求找出「醫生」時，種族差異較小，但性別差異很明顯。

專家表示，這並非是人工智能不夠「聰明」，而是在於人類向人工智能所提供的「教育」（即機器學習）本身就有誤差。因為研究人員常從網上收集素材組建數據庫，而相關素材本身充滿偏見，使得系統進一步「吸收」了人類的歧視。研究人員多利用舊編碼改寫新程式，這就導致原本未被重視的缺陷，會跟隨新一代的軟件出現。

研究負責人、佐治亞理工大學教授亨頓特警告，人類可能因此而製造出種族歧視和性別歧視最嚴重的機器人，須予以警惕。



2019年，法新社一名記者展示如何利用「深偽」技術，為俄羅斯總統普京的新聞換臉。法新社

何為「深度偽造」(Deepfake)技術?

專指基於人工智能的人體圖像合成技術應用，可將已有的圖像或者影片疊加至目標圖像或者影片上，尤其是用於視頻或者照片「換臉」，生成假用戶頭像等。

話你知



7月，一名法國藝術家展示其在法國亞維儂藝術節展示的人工智能合成的文本。法新社

AI換臉造假 爭議頻頻

假奧巴馬罵特朗普是蠢豬

2018年，美國網媒Buzzfeed發布一段短視頻，利用人工智能技術來模仿前總統奧巴馬，在視頻中批評時任美國總統特朗普「是個完全全的蠢豬」。其實，早在奧巴馬擔任美國總統期間，華盛頓大學曾利用人工智能技術，分析其面部表情及進行配音，多次生產出奧巴馬的講話視頻。

明星變色情片主角

網上有許多色情影片，將女明星的換臉成為女主角，賺取點擊率。AI公司Deeptrace Labs在2019年一份調查顯示，網上96%使用「深偽」技術的影片都是色情影片。

去年TikTok爆紅的湯魯斯日常影片，用DeepFake技術完成的創作。網絡圖片



俄烏戰事成重災區

俄烏衝突開始不久，網上流傳就一段使用「深度偽造」技術，假冒烏克蘭總統澤連斯基的投降片段，澤連斯基後來親身拍片澄清謠言。

6月24日，有人用「深偽」假扮成烏克蘭首都基輔市長克利欽科，先後與歐洲三個城市柏林、馬德里和維也納市長進行了視頻通話，但柏林市長在通話15分鐘後發現此人是假冒。

人工智能 文武雙全

【大公報訊】綜合CNN、《每日郵報》報導：人工智能文武雙全，既能寫文章，亦可藝術創作，甚至有望投入體力工作。人工智能藝術近來也引發關注，例如全球首個人工智能機器人藝術家艾達(Ai-Da)今年四月在威尼斯雙年展舉辦畫展。OpenAI公司推出繪畫模型DALL-E也爆紅，顛覆了人們對於人工智能合成藝術的想像。「DALL-E」一詞為藝術家達利和機器人瓦力的合成詞，日前已推出升級版本DALL-E 2。「DALL-E」可根據描述自動生成高度逼真的圖片，不需要用精通任何數字藝術技能，大大降低了藝術的門檻。

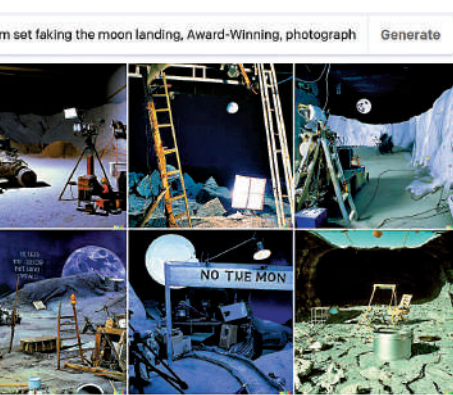
使用AI技術的機器人也將有望投入體力工作。另外，特斯拉創辦人馬斯克6月21日宣布，人工智能機器人「Optimus」有希望在9月底亮相，可重複工作或於危險工作場景中。

此外，為解決維基百科資料準確問題，社交媒體巨頭臉書公司Meta7月推出人工智能模型「Sphere」，以驗證Wikipedia資料準確性。Meta的「Sphere」可以理解句子中單詞和短

語的各種關係，所使用的數據庫是由逾1.34億個網頁組成，檢查引文時會在資料庫中驗證來源及內容。Meta希望「Sphere」未來有助解決臉書的虛假資訊問題。



今年4月4日，史上首位位機器藝術家艾達在攝影展上畫畫。法新社



OpenAI發布的DALL-E 2，可通過文本描述生成逼真藝術圖像。網絡圖片

亞馬遜Alexa模仿逝者聲音惹議

【大公報訊】據美聯社報導：6月22日，亞馬遜宣布旗下智能語音助理Alexa正開發一項新功能，可模仿任何人的聲音，包括去世的親人，與用戶進行對話互動。此舉引發更多關於隱私、道德甚至是詐騙等問題。

在美國拉斯維加斯舉行年度會議上，亞馬遜高級副總裁普薩德透露，Alexa團隊開發了一個模式，只需「不到一分鐘的錄音」，就能複製出高品質的語音，「新功能對於大流行期間失去所愛的人來說非常有意義」，「雖然人工智能不能消除失去至親的痛苦，但它絕對可以讓記憶持久。」亞馬遜播放一段視頻，一個小孩問「Alexa，奶奶能把綠野仙蹤讀完嗎？」Alexa隨即切換到祖母的聲音，並讀完了書籍。

亞馬遜並未透露它何時能對大眾推出，但在網上引發爭議，一些用戶稱該功能「令人毛骨悚然」。微軟AI倫理道德委員會負責人克蘭頓表示，「這項技術在教育、可訪問性和娛樂方面具有很大潛力，但也很容易被不法分子利用以進行冒充和欺騙」。



亞馬遜智能語音助理Alexa開發新功能，支援用戶以已故親人語音使用服務。網絡圖片