

# AI有「缺陷」 使用莫大意

## 專家：人工智能當助理 人類做最終決策

去年11月，聊天機器人ChatGPT的發布，掀起人工智能（AI）熱潮。隨着生成式人工智能飛速發展，令AI更好地成為了人們日常工作的好幫手，不僅可以提高效率、節省時間，在一定程度上降低人類可能犯錯的風險。專家提醒，當下AI系統仍有缺陷，尤其是在特殊的領域或情況下，AI工具一旦發生錯誤，將有可能導致嚴重的後果，因此人為干預不可少。

大公報記者 楊金宇



AI利弊

香港城市大學計算機科學系助理教授宋林琦接受採訪時指出，現今的AI技術在辦公中多扮演「寫作助手」的角色，協助撰寫新聞稿件、廣告語、演講詞、程序代碼等，且已在金融、交通、醫療、教育、零售、自媒體平台等多個領域中展開了實際應用。

以金融領域為例，美國的AIEQ股票即為使用AI驅動的ETF，完全借助AI的分析進行炒股操作。雖然該股票在短期和長期的狀況仍比不上標準普爾500指數，但在某種程度上已經開啟了先河，證明了AI能夠在金融市場上進行分析判斷並進行投資操作。

### AI優勢：降低系統性風險

在ChatGPT之前，AI在金融領域中的使用仍處於比較弱勢的狀態，但是，以ChatGPT背後使用的大型語言模型（LLM）為首的AI技術，則直接將金融研究帶入了新的階段，並迅速投入了實地使用。宋林琦指出，金融領域的特點是很多信息均以文本和數據的形式呈現，該特點正好極大地發揮了LLM強大的語言文字處理能力。AI可以協助人們處理各種文件報表，並迅速提取出要點及風險點，大量節省了人工成本。

宋林琦指

### 天氣預報員

谷歌DeepMind的天氣預報AI模型GraphCast在準確預測未來3到10天的天氣方面首次超越傳統方法，在預測1380個天氣指標中，90%結果優於歐盟傳統預報系統。

### 醫療助手

谷歌12月發布了基於Med-PaLM 2構建醫療保健專用AI模型MedLM，該模型在對醫學問題作出回答時準確率高達92.6%，僅比人類專業人士的準確率低0.3%。

### 會計師

Intuit公司利用自己的生成式AI作業系統GenOS平台實現專門研究稅務、會計、現金流量等的金融大型語言模型，有助於減少工作人員的重複性任務，並簡化和減少資料。

### 人力「專員」

亞馬遜在招聘過程中利用AI來篩選簡歷、將候選人與最合適的角色相匹配，進行初步簡歷評估，從而加快招聘過程和招聘時間。AI還可以用於進行初始視頻採訪。

### AI技術落地多領域



Tesla創辦人 馬斯克（Elon Musk）：

「我認為人工智能是最大的威脅之一。我不清楚我們是否能真正控制這種東西，但我認為可以引導它朝着對人類有益的方向發展。」

微軟聯合創始人 蓋茨（Bill Gates）：

「未來在人工智能能夠完成大部分工作的情况下，人類每周只需要工作3天。」

人工智能之父 辛頓（Geoffrey Hinton）：

「人工智能對世界的威脅，可能比氣候變化更為緊迫。」

美國股神 巴菲特（Warren Buffett）：

「人工智能可以改變世界上的一切，除了人類的思維和行為方式。」



### 造假換臉

今年4月，由AI圖片軟件Midjourney生成的前美國總統特朗普被捕照瘋傳，而可換臉的深偽技術（deepfake）等AI，更成為荷里活演員大罷工的導火索之一。

### 抄襲爭議

今年年初，從美國紐約到法國巴黎，部分學校禁止學生使用ChatGPT，以防止其成為寫作業和作弊的「槍手」。到了今年秋季開學，絕大部分學校已經出台了有關AI工具的使用指南。

### 詐騙勒索

隨着AI技術發展，使用門檻降低，不少犯罪分子利用AI技術詐騙，包括生成不雅裸照勒索，偽造親友聲音詐騙等等。

### 「AI幻覺」

今年最熱門的AI相關話題之一，AI因為出現「幻覺」造假說謊的例子層出不窮。

### 機器傷人

今年11月，韓國一名工人被工業機器的機械爪壓死，警方稱疑為機器的AI傳感器誤將工人當成紙箱。

### 版權爭議

OpenAI與其他AI開發商利用網上可得的新聞報道、詩詞、劇本等文本訓練聊天機器人，不少藝術家和《紐約時報》已正式起訴OpenAI和微軟。

### AI監管？

隨着AI進入各行各業，監管的權力歸屬和尺度備受關注，如何平衡技術發展、經濟利益與安全風險是其中的關鍵。目前，歐盟在監管問題上已經先行一步，已達成了全球首份AI監管草案。

### 政府

英國首相蘇納克：「到目前為止，測試新人工智能模型安全性的人員都來自於開發這些模型的公司。這種情況必須改變。」

### 企業

OpenAI行政總裁阿爾特曼：「無論政府做什麼，（AI）公司都要承擔自己的責任，這一點很重要。」

### AI「創作權」？



AI使藝術影視行業內容更多元，製作成本更低，但給版權和就業帶來了挑戰。AI是否擁有作品「創作權」？訓練AI的數據是否需要支付版權費？

### AI入課堂？

由ChatGPT引領的風潮，震蕩了教育界和學術界。AI會成為學生的私人導師還是作弊利器？AI會給學生帶來更深入的學習還是帶走學生的自我創造力？AI應當被禁止還是被引入課堂？

### 學生

丹麥學生 HORSENS GYMNASIUM 中學學生諾爾斯科夫：「我覺得我們都會認為使用它是作弊，我們都把它當作課堂上的一個東西來使用。哦，看看這個新奇的東西，它能在幾秒鐘內給出你想要的所有答案。」

### 老師

英國密德薩斯大學科技學院院長巴爾比爾：「作為學術工作者，我們有責任教育學生，讓他們了解如何以批判的方式有效地使用這些工具，並認識到它們的局限性。」

### 支持

AI視頻初創公司Runway行政總裁瓦倫蘇埃拉：「AI是一種開創性的新工具，它使電影製作更快、更容易、更便宜，它為更多人打開了製作電影的大門。」

### 反對

荷里活影帝湯漢斯：「任何人都可以通過人工智能或深度造假技術在任何年齡重現自己，明天他可能被汽車撞倒，但可以一直繼續表演。這不僅是一個藝術挑戰，也是一項法律挑戰。」

來源：BBC