Ai引領未來③
安全隱患

自動生成勒索軟件 釣魚郵件幾可亂真
黑客利用Ai犯案 衝擊網絡安全

網絡世界充斥無數黑客，繼早年的深偽技術（Deepfake）採用人工智能（AI）換面，能夠以假亂真，令人感到憂慮。如今，黑客利用生成式AI技術編寫惡意程式，或製作偽冒信息和郵件，速度之快令人側目。隨着更多複雜的新型網絡攻擊出現，全球網絡安全已響起警號。

掃一掃 有片睇

大公報記者 李潔儀

近月以來，通訊應用程式WhatsApp的詐騙手法層出不窮，繼假冒官方賬號後，黑客又透過視像來電，攝取用戶的樣貌和聲音，加以AI技術合成以作犯案用途。

香港電腦保安事故協調中心（HKCERT）數據顯示，截至今年10月已處理6341宗網絡保安事故，雖然按年有所減少，但涉及釣魚攻擊的個案大增22%。

靠人手偵測必輸 只有AI能抵禦AI

「黑客用AI做網絡攻擊的速度十分快，企業單靠人手防範，肯定會輸！」羅兵咸永道香港網絡安全及私隱服務合夥人顏國定指出，AI令黑客的攻擊手段更強，快則可於出現「零日漏洞」（即尚未修補漏洞）前12小時發動網絡攻擊，令人防不勝防。

事實上，今年以來出現不少「中間人攻擊」的個案，即在網絡上傳送數據時，發送者與接收者之間存在的「中間人」，監控雙方的通訊並進行窺探、竊取、竄改和操縱內容，而不為受害者察覺。

「世界很公平，黑客有AI去『攻』，企業也可有AI去『防』。」顏國定認為，黑客利用AI生成內容再發動攻擊的同時，企業亦要做好準備，利用AI偵測黑客的活動，例如防範網絡釣魚、惡意軟件，以及其他惡意活動。

數據分析判斷黑客蹤影

網絡安全服務供應商Palo Alto Networks大灣區系統工程主管鄭志輝表示，公司每日接收7.5億條信息，包括相關網址的原始數據，平均發現150萬個新型攻擊。他認為，雖然AI技術讓黑客的攻擊手段更高效，相對地，防守的一方亦可採用AI工具應對，利用AI技術可以更精準判斷由黑客造成的未知威脅。

▲顏國定指，AI令黑客的攻擊手段更強。

鄭志輝指出，黑客以往製作的釣魚郵件有機會錯漏百出，例如文句不通、英語拼寫錯誤等，但現在內容是通過AI生成的，大幅減少明顯的錯誤，令人難以分辨真偽，黑客的可疑行為更容易瞞天過海。

▲鄭志輝說，黑客利用AI製作釣魚郵件的錯誤大幅減少。

倡加強監管 防不法分子應用AI

「黑客可以利用AI技術做coding（程式編碼），例如做一個ransomware（勒索軟件），人要3日時間，用AI只需3個小時做完。」智慧城市聯盟資訊科技管理委員會主席龐博文指出，在人機互動的前提下，所有攻擊者的威力將會倍增，製造勒索軟件或變種電腦病毒的速度將更快。

不過，龐博文提到，防守者在數據分析、監控和抵禦黑客提出對應的方法，在信息安全的危機層面，形成敵我雙方的速度和技術得以快速提升。因此，在網絡安全領域的議題上，已與AI和諧共存、互相融合。

龐博文指出，雖然AI有學習、搜索資料的能力，但它並非百科全書，不能夠判斷真假，因此極需要對AI技術的應用進行監管，以防不法分子利用AI技術行騙。


▲龐博文稱，AI技術應用需要監管，防止被不法分子利用。

名人語錄

"The possible threat posed by AI, I think this might end up being more urgent（than climate change）."

「人工智能可能帶來的威脅，我認為最終可能會比氣候變遷更迫切。」


辛頓Geoffrey Hinton
電腦科學家

AI助黑客提速發動網絡攻擊

出現「零日漏洞」*前12小時
勒索軟件集團Hive利用暗網上洩露的託管服務供應商憑證，初始到訪受害者並嘗試部署發動勒索軟件

8小時
勒索組織Cl0p利用文件傳輸軟件MOVEit零日漏洞發動攻擊，實現初始存取並成功竊取資料

1天之內
勒索組織Cl0p利用GoAnywhere MFT工具中的另一個零日遠端程式碼執行（RCE）漏洞所花費的時間，是快速竊取資料並執行單次勒索的另一個實例

2天
觀察到Citrix產品NetScaler ADC和NetScaler Gateway程式碼被大規模注入漏洞，造成香港和澳門逾30名潛在受害者

3天
觀察到大規模利用Ivanti的Endpoint Manager Mobile（EPMM）遠端未經驗證的API存取漏洞，造成香港和澳門逾20名潛在受害者

註：* 零日漏洞（Zero-Day attack）意指發現尚未修補的漏洞；表中時間為攻擊所需時間

大公報記者整理

美律師以AI找判例 全是捏造案件

加倍審慎

所謂「成也AI，敗也AI」，雖然人工智能（AI）技術正以驚人速度滲透至各行各業，但AI是否百分百信得過？

巴西法官Jefferson Rodrigues公布一份由ChatGPT成生的裁決，由於涉及大量的法律謬誤，因而遭巴西國家司法委員會調查。Jefferson Rodrigues表示，該份裁決文件是由「trusted advisor（值得信賴的顧問）」利用AI協助而成，但當中包含的是多處不正確的細節。

至於在美國執業超過30年的律師Steven Schwartz，早前協助原告起訴哥倫比亞航空公司疏忽，在法庭文件中引用超過6個判例。不過，法官及被告人都找不到相關判例，涉嫌存在虛假成分。

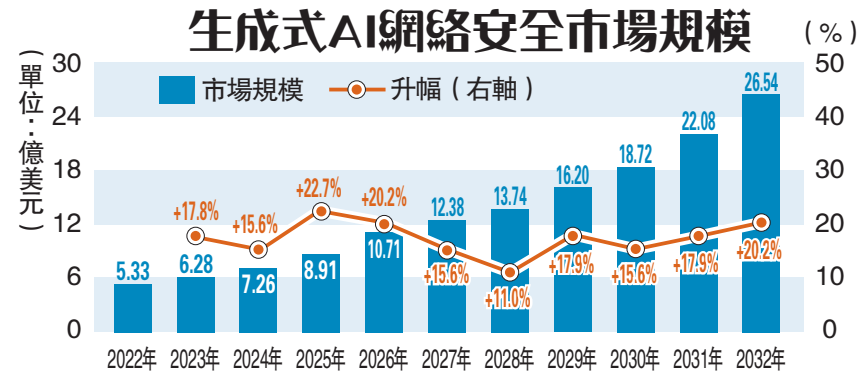
Steven Schwartz承認，自己利用ChatGPT協助撰寫法庭文件，而且並沒有查證當中的消息來源，現實發現ChatGPT也會捏造案件。

印度法官判決 參考ChatGPT意見

另外，印度一名犯人因涉嫌毆打及謀殺案而被捕，法官在判決前尋求AI協助，詢問AI聊天機器人ChatGPT：「What is the jurisprudence on bail when the assailants assaulted with cruelty?（施襲者在施襲後，應如何作出保釋判決？）」ChatGPT回答並提供相關理據，如果施襲者或對社會構成危險，法官不應提供保釋。

該名法官強調，參考ChatGPT的答案，只為得到更廣泛的案件理解，最終法官考慮到犯人施襲令受害者死亡，決定駁回被告的保釋申請，這是印度首宗由ChatGPT協助法官作出判決的案例。

生成式AI網絡安全市場規模



註：2023至2032年為預測數據 資料來源：MarketResearch

科技雙刃劍 享便利亦要防風險

新聞分析
李潔儀

今年以來，生成式人工智能（AI）爆發技驚四座的能力，瞬間成為普羅大眾的超級助手，用戶只要填寫數個關鍵詞、指令，便可生成內容、文章、圖片，甚至為用戶解答問題、提出建議等，無疑是AI領域之中異軍突起的工具。

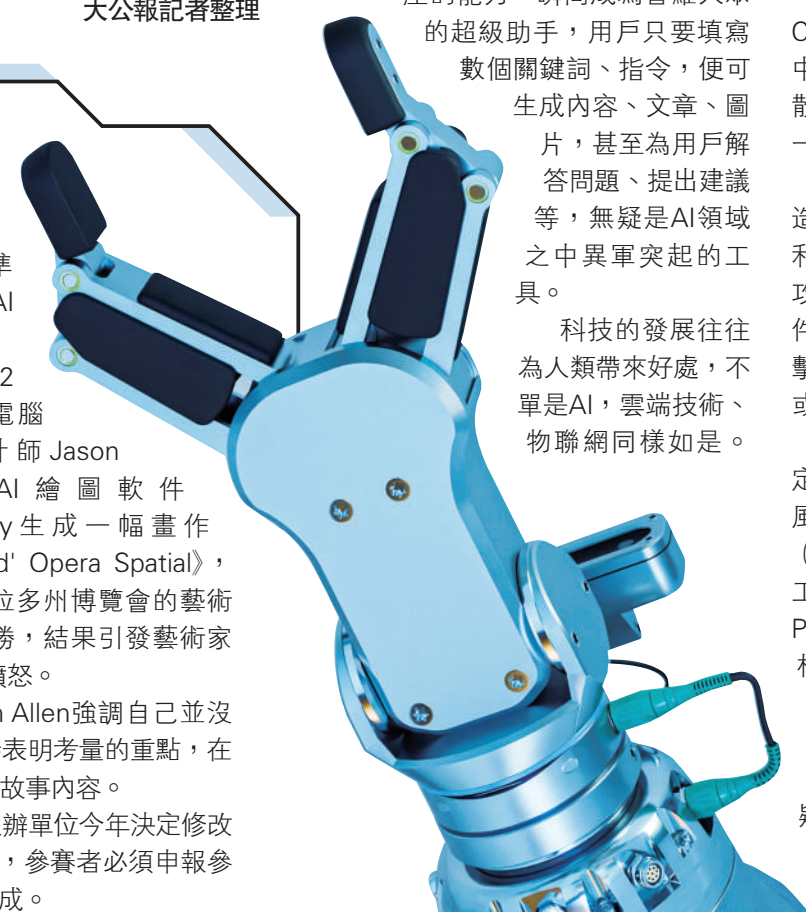
科技的發展往往為人類帶來好處，不單是AI，雲端技術、物聯網同樣如是。縱然科技能夠解決人類不少的問題，背後卻為人類造成另一項挑戰。

AI並非萬能，較早前，ChatGPT因技術故障導致服務一度中斷，開發商OpenAI證實是遭到分散式阻斷服務攻擊（DDoS），可見一斑。

事實證明，黑客衝着AI而來，造成大量網絡風險，例如成為黑客利用的目標，黑客通過AI撰寫網絡攻擊程式、創作零破綻的釣魚郵件。同時AI系統有機會反遭黑客攻擊，挖掘系統漏洞，導致數據洩露或其他安全問題。

對於AI技術的使用，外國已制定大量的指引，以盡可能規避網絡風險，例如《歐盟人工智能法案》（EU AI Act），還有英國發布人工智能白皮書《AI White Paper》。至於香港，亦正籌備多項相關法律，包括網絡安全法，制訂關鍵基礎設施法等，以減低網絡安全的隱患。

雖然AI技術的威力無容置疑，但凡事都有利與弊，AI已成為新時代科技下的雙刃劍，用者必須認清風險。



AI製作兒童讀物 惹道德和版權爭議

創作爭議

儘管人工智能（AI）帶來無限想像，在一鍵生成的前提下，人人都可成為設計師、藝術家。不過，利用AI生成的作品版權誰屬？當中涉及的道德問題又如何規範？

在美國三藩市當金融產品設計經理的Ammaar Reshi，利用生成式AI，只用72小時便創作出兒童讀物《Alice and Sparkle》，並在電商平台以每本8.99美元（約70港元）出售。


▲美國金融產品設計經理利用生成式AI創作兒童讀物，並在電商平台出售。

不過，有兒童學家質疑書本內容未能啟發兒童，而且存在大量抄襲行為，部分插畫細節更會令人不安。

AI畫作獲獎 引發藝術家不滿

另外，德國攝影藝術家Boris Eldagsen一幅名為《The Electrician》的作品，在今年Sony世界攝影大賽中奪冠。

可是，Boris Eldagsen指出，該作品是經由AI生成，因此拒絕接受獎項，並呼籲業界討論是否準備好接納AI圖像。

2022年9月，電腦遊戲設計師Jason Allen以AI繪圖軟件Midjourney生成一幅畫作《Theatre d' Opera Spatial》，並在科羅拉多州博覽會的藝術比賽中獲勝，結果引發藝術家的爭議和憤怒。

Jason Allen強調自己並沒有犯規，評審們當時表明考量的重點，在於作品是否能夠帶出故事內容。

為避免爭拗，主辦單位今年決定修改151年來的比賽規則，參賽者必須申報參賽作品是否採用AI而成。