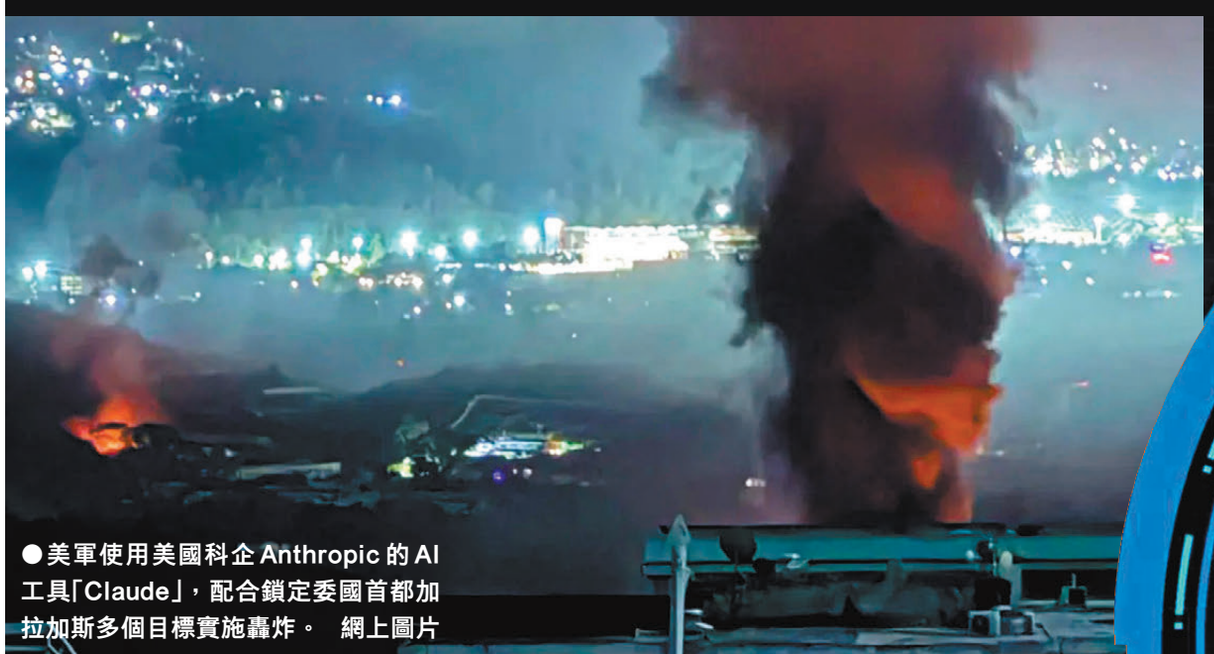
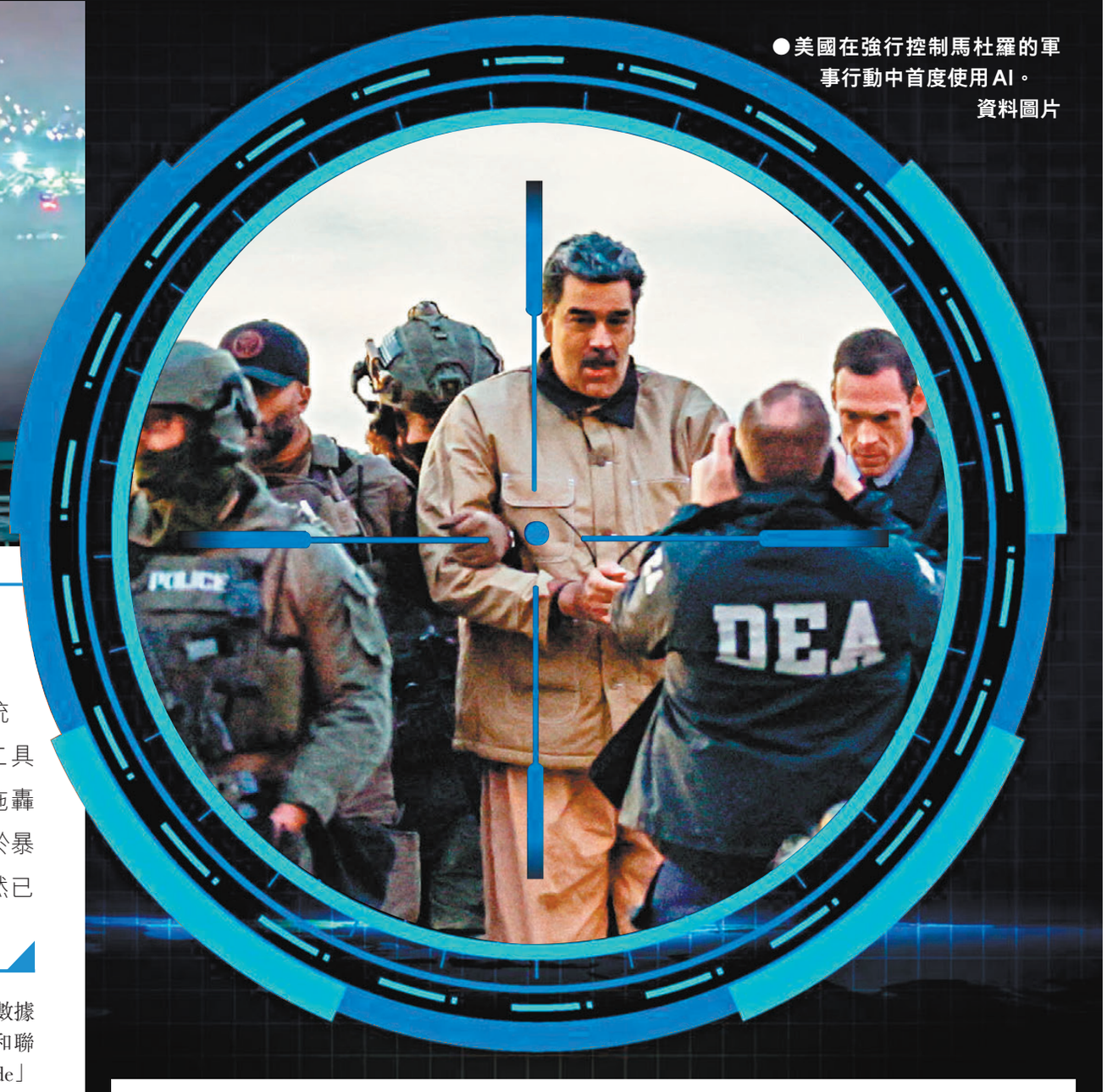
美媒揭無視科企使用準則 借助Anthropic工具強行控制馬杜羅



●美軍使用美國科企Anthropic的AI工具「Claude」，配合鎖定委國首都加拉加斯多個目標實施轟炸。 網上圖片



●美國在強行控制馬杜羅的軍事行動中首度使用AI。 資料圖片

美入侵委行動首用AI 被批越界引發人道風險

香港文匯報訊 美國正逐步將人工智能(AI)技術作軍事用途。《華爾街日報》上周五(2月13日)獨家披露，美軍上月在入侵委內瑞拉、強行控制委國時任總統馬杜羅的行動中，曾使用美國科企Anthropic的AI工具「Claude」，配合鎖定委國首都加拉加斯多個目標實施轟炸。Anthropic的使用準則中，明確禁止將「Claude」用於暴力行動、開發武器或實施監控，專家警告美軍行為顯然已經越界，或令AI技術成為軍事衝突的重大風險來源。

報道指出，美軍部署「Claude」，是通過Anthropic與美國大數據分析公司Palantir的合作實現，後者的工具早已被國防部和聯邦執法機構普遍使用。早前由於Anthropic對國防部使用「Claude」的方式存在顧慮，美政府已考慮取消一份與該企業價值達2億美元（約15.6億港元）的合約。

或有其他AI工具被美軍使用

令人尷尬的是，Anthropic多次公開強調公司旨在保障AI安全，行政總裁阿莫戴伊還多次表示非常擔憂AI技術被用於致命的作戰行動和內部監控，豈料自家AI工具卻被揭協助美軍入侵他國的行動。Anthropic辯稱，公司無法評論「Claude」或其他AI模型是否被用於特定任務，不論政府部門還是民間組織，使用「Claude」都應遵守使用準則，確保符合相關規範。

Anthropic是首家AI模型被確認由國防部用於機密行動的AI技術開發商。報道還引述知情人士消息稱，是次行動中，可能也有其他AI工具已被美軍和國防部用於處理非機密任務，應用範圍包括匯總文件、整理資料、操作無人機等。

防長掃清阻AI高效發展障礙

美國防長海格塞斯從去年起便多次表示，他會在任內徹底改革國防部技術體系，掃清美軍內部阻擋AI高效發展的障礙。今年1月視察特朗普盟友馬斯克名下太空探索企業SpaceX期間，海格塞斯宣稱會提出AI優先轉型計劃，國防部會向AI開放美軍信息技術系統的所有合適數據和情報：「我們很快會將全球領先的AI模型，部署到我們所有非機密甚至機密網絡中。」

海格塞斯大幅推進AI技術應用於軍事領域的主張惹人關注，他在演說中更直言：「我們不會採用無法用於戰爭的AI模型。」海格塞斯稱，美軍會迅速發展或具有不可預測危險性的新興AI技術，「我們會以戰時手段，對待阻礙在國防部推廣AI的人員和政策，我們正在清除這些障礙。」

專家認為，AI軍事化應用可能帶來的人道、倫理問題一直備受關注，AI武器化可能帶來比核武更大的破壞力，必須受到嚴格監管。如今海格塞斯大力推進國防部的「人工智能加速戰略」，這一動向以及由此可能帶來的難以預測的後果，需要全球高度警惕。

美加速AI軍事化進程 「競爭優先速度至上」

香港文匯報訊 美國國防部長海格塞斯上月簽署發布新版人工智能（AI）戰略文件，標誌着特朗普政府在第二任期內全面加速AI軍事化進程。這份名為《人工智能加速戰略》的文件明確提出構建「AI優先」作戰力量，以鞏固美國的軍事優勢，確立AI在美軍中的主導地位。

該戰略強調以「競爭優先」和「速度至上」為核心理念，通過4條路徑推進落實。國防部將在內部推動先進AI模型的試驗性應用，改革陳舊的信息技術體系，加大對算力、模型創新及實戰數據等關鍵領域的投入，並依託「節奏設定項目」機制，快速構建AI整合所需的支撐體系。七大初始引領項目覆蓋作戰、情報與管理領域，其中包括探索AI賦能新戰法，以及將情報轉化為作戰能力周期壓縮至小時計的開放兵工廠項目。這些項目需每月向國防部高層匯報進展，並在備忘錄發布後6個月內完成首批用戶轉化演示。

事實上，特朗普去年重返白宮後，對美國AI戰略作出重大調整。他去年1月簽署行政令，移除AI領域的發展障礙，並廢除上屆政府的相關監管措施，隨後發布的《贏得AI競賽：美國人工智能行動計劃》，提出以加速創新、建設基礎設施和領導國際外交為三大支柱的戰略全貌。為配合這一轉型，美國啟動名為「星際之門」的龐大基礎設施建設計劃，擬在4年內投入5,000億美元（約3.9萬億港元）建設涵蓋數據中心集群、能源擴容系統及半導體製造能力的綜合網絡。

分析認為，美國正以激進方式清除AI技術進入軍事體系的障礙，從強調發展與安全並重的「雙向布局」，轉向「速度至上」的單一導向。這種轉變深受「科技右翼」思潮影響，強調市場優先和反對過度監管。特朗普更計劃將2027財年國防開支提高至1.5萬億美元（約11.7萬億港元）的破紀錄水平，同時深化與私營科技企業的合作，將數千億美元的民間投資納入軍事體系。

●美明確提出構建「AI優先」作戰力量。 網上圖片

特朗普擬訪委 授權5油企營運禁俄伊投資

香港文匯報訊 美國總統特朗普上周五(2月13日)宣布計劃親自訪問委內瑞拉，並重申對當地臨時領導人表示滿意。美國政府同日正式授權5家歐美石油巨頭企業恢復在委國的營運。然而這份來自美國財政部的許可設有嚴格的「排他條款」，明確禁止俄羅斯與伊朗等國家參與相關投資。

許可證由美國財政部外國資產控制辦公室（OFAC）頒發，授權英國石油公司（BP）、蜆殼（Shell）、美國雪佛龍（Chevron）、意大利埃尼（Eni）及西班牙的雷普索爾（Repsol）投資委國石油產業。這並非無條件解禁，而是規定所有石油與天然氣的特許權使用費，必須存入美國財政部指定的賬戶，顯示美方雖然放行開採，但仍牢牢掌握委內瑞拉石油的經濟命脈。

分析指出，特朗普政府允許歐美油企恢復在委國營運之舉，旨在解除一部分對委內瑞拉制裁的同時，重塑拉美能源版圖，將競爭對手拒於門外。OFAC發出的許可證證實這一點，即美國雖允許企業就潛在合約進行談判，但明確禁止俄羅斯與伊朗等國家參與。

●委內瑞拉臨時總統羅德里格斯(左二)到油企視察。法新社

國防部施壓科企移除AI安全道德限制

香港文匯報訊 據路透社報道，美國五角大樓正積極施壓OpenAI、Anthropic等頂尖人工智能(AI)公司，要求將其AI模型部署至軍方的網絡，且必須移除規範一般用戶的安全道德限制。

兩名知情人士透露，五角大樓技術長邁克爾日前在白宮一場活動中向科技高層表示，軍方的目標是讓AI能同時在非機密與機密領域使用，一名匿名官員更直言，五角大樓正開發能在所有機密資料中都能使用的AI。

無限制使用AI恐釀致命後果

這項最新發展標誌着美軍與生成式AI巨頭之間的談判進入最新階段。面對未來戰場已逐漸被自主無人機群、機械人與網絡攻擊主宰，美軍急欲導入AI強化戰力，然而這也加劇了軍方「無限制使用」需求與科企「設立道德護欄」之間的緊張關係。

美國軍方高層希望利用AI強大的資訊整合能力來輔助作戰決策，但專家警告，AI模型著名的「幻覺」問題若發生在涉及導彈目標鎖定或任務規劃的機密環境中，恐釀成致命後果。雖然科企試圖透過建立安全護欄來降低風險，但五角大樓官員對此感到不耐煩，主張只要符合美國法律，軍方部署商業AI工具就不應受到科企的道德條款限制。

OpenAI早前已與五角大樓達成協議，允許美軍在名為GenAI.mil的非機密網絡上使用ChatGPT等工具，該網絡覆蓋逾300萬名國防部員工。作為協議一部分，OpenAI同意移除許多針對一般用戶的典型限制，儘管仍保留部分核心護欄。Alphabet旗下的Google與馬斯克的xAI此前也簽署了類似協議。

美就襲擊伊朗擬訂為期數周計劃

香港文匯報訊 美國總統特朗普上周五（2月13日）證實，將派遣第二艘航母「福特號」前往中東，施壓伊朗就核計劃達成協議，他又表明伊朗政權更迭將是「最理想的情況」。路透社引述兩名美國官員稱，若特朗普下令向伊朗發動攻擊，美軍已擬訂為期數周的軍事行動計劃。

特朗普：政權更迭是「最理想情況」

特朗普日前警告伊朗若不能在一個月內與美國達成協議，將面臨慘痛後果。多間美媒報道「福特號」已從加勒比海調往中東，特朗普稱「如果我們無法達成協議，我們就會需要它。相信與伊朗的談判將會成功，否則對伊朗而言將非常糟糕」。他又在布拉格堡軍事基地對記者說，「與伊朗達成協議一直很困難，有時你必須令人恐懼，這是唯一能真正解決問題的辦法。」

被問及會否尋求伊朗的「政權更迭」時，特朗普一反此前全面撤回謀求伊朗政權更迭的表態，表示這似乎是最理想情況。但他拒絕透露希望由誰人接替伊朗最高領袖，僅稱「那裏有一些人選」。

一名美國官員指出，在潛在的軍事行動中，美軍或會打擊伊朗的國家機構和安全設施，而不僅是核基礎設施。他同時指美方完全預料到伊朗會展開報復，衝突不會馬上結束，故此擬訂了數周的軍事行動計劃。伊朗革命衛隊則警告，若美國對伊朗領土發動攻擊，附近區域內的所有美軍基地都會成為報復目標。

AI應用涵蓋行政運作潛艇建造

香港文匯報訊 美國國防部近日正式宣布推出名為GenAI.mil的人工智能（AI）平台，其首個導入的商用AI模型為Google旗下的Gemini for Government，這也是五角大樓首次大規模部署商用生成式AI工具，顯示美軍正務實推進其「AI優先」戰略。

根據美國國防部規劃，此舉旨在將先進AI能力直接部署到五角大樓及全球美軍基地中，將服務約300萬名國防部員工，是美軍「AI優先」戰略的一部分。出於國防軍事領域「安全至上」原則，初期GenAI平台主要從事文件格式化與撰寫、影片與圖像分析、人員入職流程等非機密工作。當前該工具僅有IL5等級環境下運行的許可，該等級允許系統處理受控非機密信息，其亦是軍方對敏感資料的主要管理層級。同時為確保輸出內容可靠，並降低生成錯誤信息的風險，GenAI平台生成的內容還會即時比對Google搜索結果進行驗證。

美國海軍方面，近日亦宣布與大數據分析公司Palantir合作推出「ShipOS」數碼化平台，用於整合並可視化造艦與維修流程中的各項數據，提前最多180天預知潛在問題，而非等到問題發生才被迫停工。

目前ShipOS應用於潛艇建造流程，未來這項技術還會擴展應用至航空母艦與戰機的維護與製造。