生成式AI技術發展對國家安全的影響與挑戰

◆ 國防大學中共軍事事務研究所副教授 — 董慧明

「生成式 AI」（Generative Artificial Intelligence）技術的發展提高了人類工作與生活的效率，然亦同樣引發各種安全、法律和道德面的爭議，如何在應用此技術的同時亦合乎情、理、法準則，已成為大眾必須關注的焦點。

生成式 AI 技術的誕生

2022 年 11 月，美國人工智慧研究公司（OpenAI）推出 AI 聊天機器人 ChatGPT，掀起了「生成式 AI」技術應用熱潮。透過機器學習模型和神經網路技術運算大量數據，生成式 AI 創作出包括文書、對話、故事、圖像、影片和音樂等類成品；因具有高效提升處理、編輯、編碼等日常與專業任務能力，讓各種 AI 聊天機器人成為備受矚目的應用工具。以 ChatGPT 為例，根據以色列一家軟體和資料公司 SimilarWeb 統計數據顯示，自問世到今年 5 月的網路流量達 18 億高峰，目前已躋身全球前 25 大網站之列。

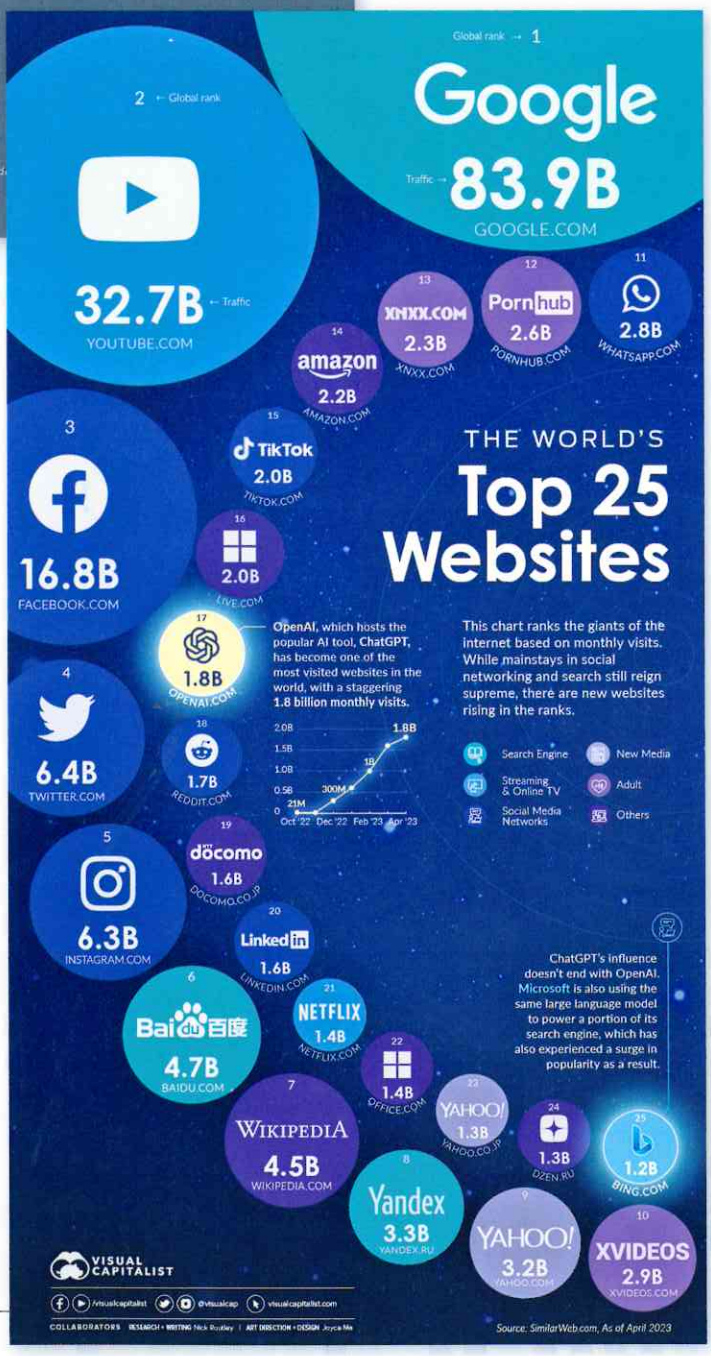
生成式 AI 技術因具有提高工作效率和多領域的創新潛力而受到各界青睞，然而

透過機器學習模型和神經網路技術運算大量數據，生成式 AI 根據指示創作出包括文書、對話、故事、圖像、影片和音樂等類成品。（Photo Credit: Visual Capitalist, https://www.visualcapitalist.com/generative-ai-explained-by-ai）

當眾人引頸期盼這項技術還能夠為我們的生活和工作帶來更多便利之際，使用上的安全顧慮、法律和道德方面的爭議等話題，也產生愈來愈多的討論，特別是政府公務、個人資料等攸關國家安全問題均值得深思。

生成式 AI 技術的特性與運用範疇

相較於近年來同樣受到關注的大數據分析、文字探勘、機器學習等新穎資訊科技，生成式 AI 因具有更大的發展潛力和多重應用影響力而迅速成為引領下波科技革命的創新力量。它所具有的「創造性」、「逼真性」，以及「可擴展性」三大特性，更是可能全面性地影響國家安全。



根據統計數據顯示，ChatGPT 自問世到今年 5 月的網路流量達 18 億高峰，而原先為微軟搜尋引擎的 Bing 也因推出 Bing Chat 流量躍升至 12 億，兩者皆躋身全球前 25 大網站之列。（Photo Credit: Visual Capitalist, https://www.visualcapitalist.com/ranked-the-worlds-top-25-websites-in-2023）

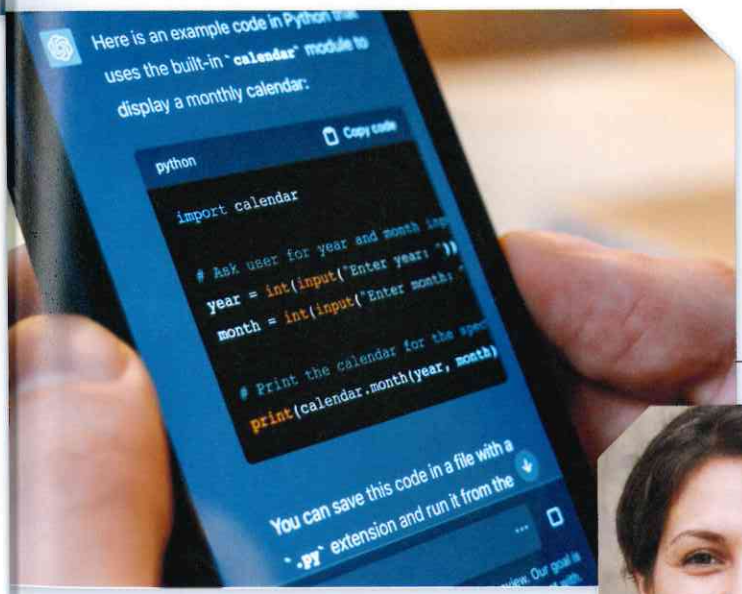
生成式 AI 技術讓不是藝術家的使用者，只要輸入適當的提示詞或是點選繪圖指示，便能利用演算法分析多組網路圖片，按照其畫風與美感創作出新的藝術作品；只是隨著使用率愈高，所衍生的藝術界定、著作版權及作品擁有權等爭論也應運而起。

首先，生成式 AI 技術能夠分析現有數據，創建全新內容，其創造性成為藝術創作、產品設計和研究開發等領域所用，例如 Midjourney、NightCafe、Stable Diffusion、Deep Dream Generator 皆屬時下熟悉的AI圖像生成器，簡單易用且免費。即使不是藝術家的使用者，只要懂得輸入適當的提示詞或是點選繪圖指示，生成器便能利用演算法分析多組網路圖片，並且按照其畫風與美感創作出新的藝術作品。只是隨著使用率愈來愈高，所衍生的藝術界定、著作版權及作品擁有權等爭論，也應運而起。

其次，生成式 AI 技術的逼真性也是各界躍躍欲試的主因。娛樂、教育和廣告工作者等皆因這項新科技能生成逼真的內容，使他們能夠創造更引人入勝的遊戲、教材和廣告，進而吸引更多的觀眾和學習者。例如：OpenAI 的 GPT-3.5、GPT-4 語言模型可以生成逼真的詩歌、程式碼、腳本、音樂作品、電子郵件、信件等文字內容。

最後，在求快、求變的網路科技時代，能夠快速有效地蒐集、整理與活用資訊，已是決策制定和競爭優勢的成功關鍵。生成式 AI 技術的可擴展特性，可應用於各種數據查整與分析，並隨著資料量的增加而提高生成效果。例如：由「輝達」公司（NVIDIA）開發的 ProGAN、StyleGAN 生成式對抗網路模型，能夠生成高解析度的人臉和風景圖像，而模擬仿真的應用，更提供了創新方法來解決現實世界中的複雜問題。

由此可知，生成式 AI 技術的運用幾乎涵蓋政府、企業乃至個人創作等各領域，潛力無窮。「臺灣人工智慧學校校務長」蔡明順受訪時曾提到，當前生成式 AI 在企業應用中的五個層次分別是：一、學習溝通方式，找出合適的人機協作模式；二、串接公司服務；三、對語言模型進行「微調」（Fine-Tune）；四、使用自己的「數據集」（Dataset）來訓練語言模型；五、開發多模態處理的模型，建立企業本身的生態系。綜上，生成式 AI 正在成為工作、生活不可或缺的利器，在社會各領域扮演關鍵角色，為人們帶來更多的便捷和創新。



OpenAI 的 GPT-3.5、GPT-4 語言模型可以生成逼真的詩歌、程式碼、腳本、音樂作品、電子郵件、信件等文字內容。

對國家安全的影響

就國家安全面向，生成式 AI 豐富的功能有助於提高情報蒐集分析的效率與品質。以公開來源情報（Open Source Intelligence, OSINT）工作為例，因以蒐集、研析公開可得、合法獲得、可供大眾使用的訊息為特色，且必須從包羅萬象的情資細節中分析出具有重要價值的訊息，導入生成式 AI 技術能快速分析大量數據、自動識別並提取關鍵資訊。有別於過去需耗費大量時間和人力進行蒐集、分析，生成式 AI 讓情報單位能更快速、精確地識別潛在威脅和線索機會，並藉由分析多種文本、圖片、影像和聲音等數據來源，提供相關事件、行為和趨勢的輔助判斷。目前以色列國安單位已採用生成式 AI 作為情報工具，打擊潛在國家安全威脅，甚至協助政府執

由輝達公司開發的 ProGAN、StyleGAN 生成式對抗網路模型，能夠生成高解析度、多層次的人臉和風景圖像。（Photo Credit: Nvidia Corporation, https://github.com/NVlabs/stylegan）

法單位打擊犯罪。美國國防部也發現生成式 AI 具有強化情報工作、作戰計畫與行政程序效率的潛力，因而成立「利馬工作小組」（Task Force Lima）負責五角大廈生成式 AI 能力的評估、同步與應用等工作，以確保國家安全、最大程度降低風險，並且整合技術發展。

我國對生成式 AI 應用於各公部門工作的情形也相當重視。今（112）年 8 月 31 日通過的「行政院及所屬機關（構）使用生成式 AI 參考指引（草案）」即指出：生成式 AI 快速發展，功能極為多元，利用此技術協助執行業務或提供服務，有助於行政效率之提升。然而，生成式 AI 亦因大量蒐集、學習與產出之資料，可能涉及智慧財產權、人權或業務機密之侵害，且其生成結果，有可能存在真偽難辨或創造不存在的資訊，須客觀且專業評估其產出資訊與風險。

因此在前開草案的十點指引內容中，特別提醒「為避免其可能帶來之國家安全、資訊安全、人權、隱私、倫理及法律等風險，各機關人員使用生成式 AI 時，應秉持負責任及可信賴之態度，掌握自主權與控制權，並秉持安全性、隱私性與資料治理、問責等原則，不得恣意揭露未經公開之公務資訊、不得分享個人隱私資訊及不可完全信任生成資訊」。生成式 AI 在國家安全領域的應用層面廣泛，對於政府提高行政效能也有助益，但確保此技術在國家安全領域的合法與負責任的使用也十分重要，必須謹慎確保公務機密、個人隱私和安全不受威脅。

安全挑戰的因應

當各界聚焦在生成式 AI 的應用層面，以及思索未來有哪些行業或將遭到 AI 取代之際，也應反思其衍生的使用安全問題及因應作法，包括法律層面的「資料治理」（Data Governance）議題，以及 AI 產製內容所造成的倫理和道德問題。無論是哪種功能的生成式 AI 產品，「數據集」與「資

養成對生成式AI的正確觀念

- 掌握自主權與控制權
- 客觀且專業評估生成式AI產出之資訊與風險

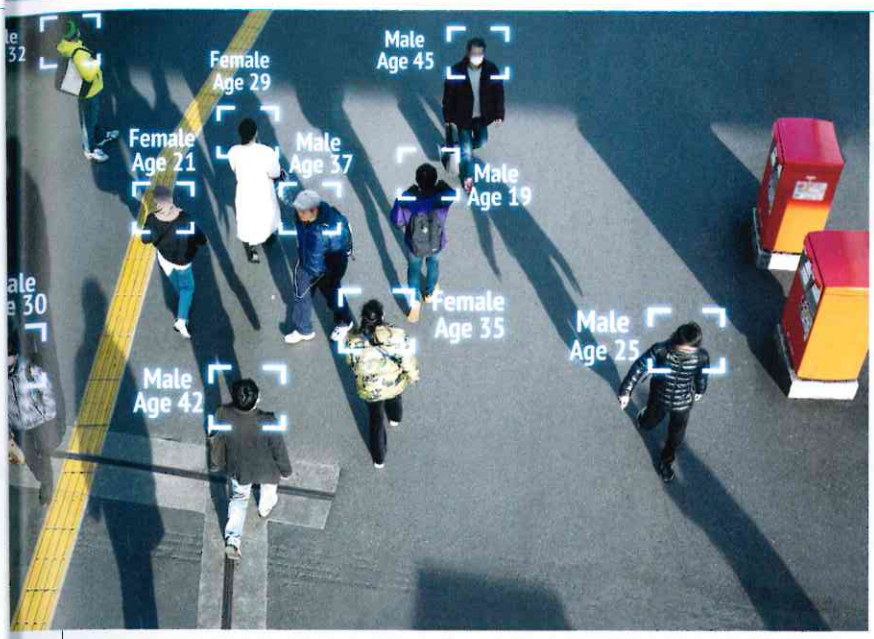
界定技術/工具運用的責任

- 保持公務之機密性及專業性
- 注意著作權及人格權等

建立必要的安全與內控機制

- 秉持負責任及可信賴之態度使用
- 得視需求訂定內控管理措施

因應 AI 發展日新月異，行政院國科會將持續觀察全球相關趨勢與作法，目前採滾動式調整「行政院及所屬機關（構）使用生成式 AI 參考指引」，使規範保留彈性，力求於使用生成式 AI 之風險管理與創新發展之間取得平衡。（資料來源：行政院，https://www.ey.gov.tw/Page/9277F759E41CCD91/e63572a7-fb79-4c02-9ea3-b731e7c06a56）



歐洲議會審議《人工智慧法》將著重禁止人工智慧系統進行社會評分、生物識別分類和情感識別。

料處理能力」是兩大核心要素，缺一不可，因此從資料治理所延伸出的智慧財產權歸屬，以及在訓練模型時所涉及的個人資料保護、數據安全爭議，成為相應制度與法規制定的重點。資料治理不僅關注數據的品質和安全性，還涵蓋了所有資料來源、清理、更新、儲存、分析、傳輸、備份和刪除等方面的生命週期，必須透過規範和程序來管理生成式 AI 所使用的數據，確保資料的存管運用皆能得到妥善管理。

生成式 AI 使用在非法用途所涉及的道德層面問題，如不法人士將訓練資料提供給相對應功能的各種語言模型，再用以作為學術倫理、假訊息詐騙、洗錢、釣魚郵件、詐騙電話、深度偽造（Deepfake）、網路攻擊、假新聞宣傳、認知作戰，甚至攻擊國家關鍵基礎設施、竊取機敏數據等態樣，已成為全民難以迴避且應有效反制的新形態威脅，對經濟、社會與國家安全危害更鉅；制定相關法律法規，規範生成式 AI 的研發、使用和應用，以防範負面衝擊，保障公民權益，恐有其必要。

鑒此，透過立法監管生成式 AI 已成為各國政府和法律機構的當務之急，英國、澳洲政府也曾提出公務員使用生成式 AI 指南，歐盟則是在今年 6 月通過《人工智慧法》（Artificial Intelligence Act）草案，且有望於年底生效，成為全球首部管理人工智慧的法律。美國聯邦貿易委員會（Federal Trade Commission）也去函 OpenAI、微軟（Microsoft）公司瞭解其解決虛假、誤導資訊風險的作法，以及新「必應（Bing）」搜索引擎是否因使用 OpenAI 技術侵害民眾權益。面對數位智能時代的來臨，相信惟有制定適切的法律規範，方能確保生成式 AI 技術在未來的發展中既能推動創新，又能保護個人和國家的權益。

結語

生成式 AI 展現出強大的創造力、逼真性和可擴展性，在各個領域都具有廣泛的應用前景，它能提高工作效率、促進創新，並為國家安全情報工作提供助力；然而，在 AI 技術快速發展的同時，其涉及個人隱私、機密資訊、智慧財產權等問題，也帶來了一系列法律和倫理道德挑戰。除在引入應用時，必須確保資料品質與使用生命週期的完善，均衡技術應用與確保安全更是重點，各界推動創新應用的同時，必須保持警覺，建立制衡機制，並持續周延監管法規，讓生成式 AI 技術的發展，為人類社會帶來正面效益，維護公共利益。