DELL Technologies

NVIDIA

生成式 AI (GenAI) 的 5 大安全考量

透過 Dell AI Factory with NVIDIA，
加速採用安全且可擴充的基礎結構

GenAI 的轉型潛力

GenAI 具備改變遊戲規則的潛力，
超越有遠見的業界人士當前的想像。

76%

的 IT 和業務主管認為，
GenAI 可為組織創造
改變現狀的價值[1]

AI

運用先進的分析和邏輯技巧來解釋事件，並支援及自動化執行決策和動作。

生成式 AI

利用大量資料，從自然語言提示或其他非程式碼和非傳統輸入產生新內容的科技與技術。

模擬

- 數位分身
- 合成資料
- 設計框架
- 預測

內容探索

- 自然語言搜尋
- 大型資料集分析
- 知識管理
- 個人化教育與訓練

內容生成

- 編碼
- 數學
- 寫作/口說
- 影像/影片
- 音效

使用者經驗

- 超過 70 種語言的即時翻譯
- 使用自然面部表情和肢體語言進行個人化互動

1《Dell Technologies Innovation Catalyst Study》(Dell Technologies 創新催化劑研究)，2024 年 2 月

潛力增加，風險增加

對企業領導者而言，能迅速採取行動，繞過涉及資料、法規遵循、治理和其他風險的影響，是相當誘人的。但在安全性方面，GenAI 是一把雙刃劍。

優點

- 改善威脅偵測能力
- 提升營運效率
- 個人化安全意識訓練

缺點

- 攻擊手法日益複雜
- 增進社交工程技術
- 影子 AI

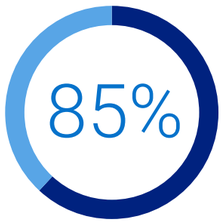
33%
的受訪者將網路安全性列為其組織正在努力減輕的首要 GenAI 風險。[2]

2 《McKinsey Global Survey on AI: The state of AI in early》(McKinsey 全球 AI 調查：AI 的早期狀態)，2024 年 5 月

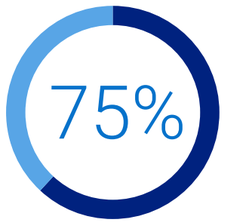
考量 1

全新的威脅態勢

伴隨 GenAI 的承諾而來的是一個嚴肅的現實：攻擊者正在創造更複雜的全新攻擊手法，有能力繞過傳統防禦，使網路安全團隊難以跟上。

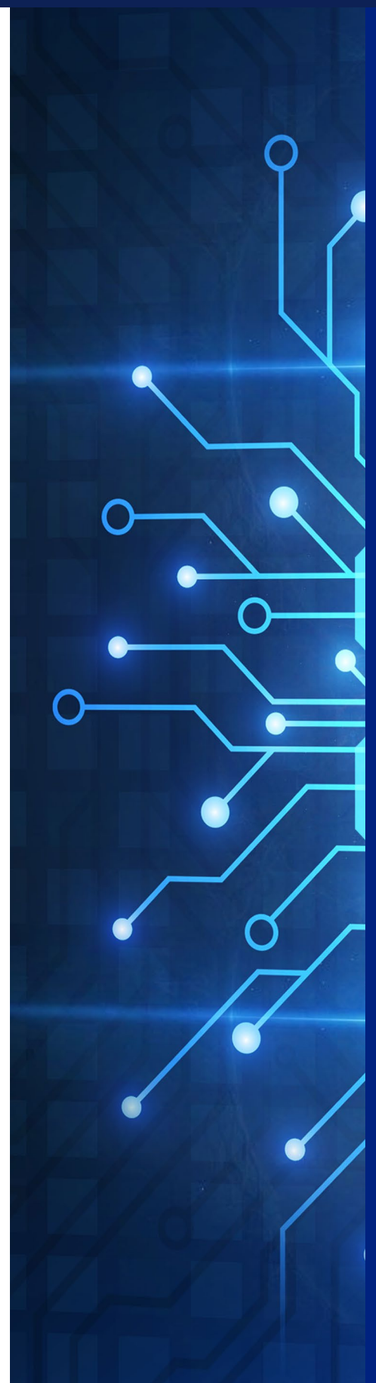


的受訪者認為，AI 讓網路安全攻擊變得更加複雜。[3]



的安全專業人士發現，過去 12 個月的攻擊事件增加。[4]

為了防範這些新出現的威脅，公司必須專注於透過滲透測試、監控和稽核等方式，將攻擊面降至最低。



新興攻擊手法

進階惡意軟體
日益複雜的惡意軟體使用 GenAI 來「自我進化」，不斷變更其程式碼，以避免被現有的安全性措施 (例如以簽名為基礎的偵測) 偵測到。

高度個人化的網路釣魚電子郵件和活動
缺乏常見詐騙跡象、看似真實的惡意電子郵件出現的頻率增加。

令人信服的深偽資料
模仿人類行為 (例如寫作、口說、影像或影片) 的能力使得身分竊盜、金融詐欺和錯誤資訊變得更加容易。

自動偵查
收集資訊，識別潛在目標網路或系統中的漏洞和弱點，以促成更具針對性的攻擊。

3 《2024 Human Risk in Cybersecurity Survey》(2024 年網路安全中的人為風險調查)，EY，2024 年 5 月
4 Voice of SecOps Report，《Generative AI and Cybersecurity: Bright Future or Business Battleground?》(生成式 AI 和網路安全性：光明的未來還是商業戰場？) 2023

考量 2

部署和實作風險

想要利用 GenAI 潛在優勢的組織需要大量的高品質資料，亦即模型可用於產生最佳成果的輸入內容。但資料與風險密切相關。在利用任何資訊之前，公司必須仔細評估並考慮其獨特的條件、投入和風險。

大型語言模型 (LLM) 漏洞

GenAI 服務容易受到提示插入攻擊，攻擊者會操縱輸出以繞過安全護欄，或未經授權存取可能用於調整模型的檔案。

資料中毒

攻擊者可以在訓練階段，故意將竄改過的資料提供給 LLM。這可能導致模型容易因資料中嵌入的後門而遭受攻擊。一個真實的例子是透過垃圾郵件訓練，來攻擊和利用垃圾郵件篩選器。

法規複雜性

世界各地的監管機構都在競相瞭解、控制和保證 GenAI 的安全性。雖然 GenAI 模型受當前資料主權規則的約束，資料的儲存、處理和使用方式皆有其規範，但主管機關仍在定義該如何監管智慧財產權和受版權保護的資訊。遵守法規可能成本高昂，但不遵守既定法規和新興法規可能會導致罰款和其他處罰。

考量 3

影子 AI

如今許多員工已經在使用公開文字、影像和影片生成器 (如 ChatGPT) 來增強日常工作流程。然而，如果在沒有適當控管的情況下使用這些工具，會對試圖保護公司智慧財產權和資料的組織構成嚴重威脅。這種未經授權使用 GenAI 的行為稱為影子 AI。

智慧財產權損失

目前公司已經開始需要處理員工在公有 GenAI 工具中分享敏感資訊所造成的智慧財產權損失。

原始碼資料洩露

開發人員試圖使用 ChatGPT 來最佳化原始碼，進而造成資料洩露。

為了應對影子 AI 的挑戰，公司應該成立一個統籌全公司的理事會或委員會，有權做出涉及安全 AI 治理的決策。

您的資料儲存在哪裡？工作負載應放置在何處？

無論資料存放於何處，AI 搭配資料使用時可發揮最佳效能。透過對基礎結構和 LLM 的完全控制，沒有智慧財產權損失或原始碼資料洩露的風險。

成本

運用內部部署可在 3 年內使總體擁有成本降低多達 75%。[5]

安全性與隱私

透過內部部署工作流程和作業，在整個組織內建立安全的 AI / GenAI 環境。嚴格控制資料安全性並遵守法規，特別是處理敏感資料的產業。

5 根據 2024 年 4 月 Dell 委託撰寫的企業策略集團研究，其中比較了內部部署 Dell 基礎結構與原生公有雲基礎結構即服務。分析的模型顯示，對於擁有 5 千名使用者的組織，運用 RAG 的 7B 參數 LLM 的成本效益高出 38%，而對於擁有 5 萬名使用者的組織，運用 RAG 的 70B 參數 LLM 的成本效益高出 75%。實際結果可能有所差異。經濟摘要

考量 4

評估準則

在過去的一年裡，AI 社群越來越關注三個關鍵議題：負責任的開發和部署、評估影響和降低風險。公司在評估 GenAI 模型時必須考慮幾個重要的注意事項：

報告要求不一致

引領業界的開發人員在測試模型時採用的負責任的 AI 基準各不相同。由於報告嚴重缺乏標準化，因此很難有條理地比較頂級 AI 模型的風險和限制。

輸出中包含受版權保護的內容

熱門 LLM 的輸出可能包含受版權保護的內容，因此可能會違反法律，並導致使用這些素材的公司面臨受到處罰的風險。

漏洞日益複雜

研究人員發現一些較不明顯的策略會導致 LLM 表現出有害行為，例如要求模型無限重複隨機字詞。

開發人員缺乏透明度

在許多情況下， AI 開發人員不會主動提供他們的訓練資料和方法。這妨礙我們進一步瞭解 AI 系統的健全性和安全性。

考量 5

安全性優勢

GenAI 除了安全性風險外，亦有其潛在的安全性優勢。GenAI 正逐漸成為網路安全的重要後盾，開闢出一條新的保護途徑。

現在，您可以開始組建可擴充的安全作業，更快取得更豐富的深入見解和自動威脅偵測，從而提高效率並支援人手不足的安全團隊。

威脅偵測與回應

透過分析歷史資料及識別模式和異常情況，GenAI 能即時識別不斷演變的新威脅。它能持續監控網路流量、系統記錄和使用者行為，並及時識別可能表示安全威脅的異常活動。

其結果就是極具適應性的威脅偵測功能，能夠快速回應不斷變化的攻擊手法，並提供針對新興網路威脅的主動防禦機制。

威脅模擬和訓練

借助 GenAI，公司可以在受控環境中，模擬各種網路安全威脅和攻擊情境。因此，在分秒必爭時，團隊可以有更充分的準備，識別、回應和緩解網路威脅。

深入分析和摘要

GenAI 使團隊能夠調查來自不同來源或模組的資料，進而更快、更準確地執行傳統上耗時又繁瑣的資料分析工作。團隊還可以建立事件和威脅評估的自然語言摘要，從而提高效率和團隊輸出。

個人化安全意識訓練

透過將對話式 AI 包裝在 GenAI 之上，並將 AI 虛擬人偶整合到使用者介面中，組織可以使用自然的面部表情和肢體語言，提供個人化的互動 (24 小時全年無休且大規模)。這可用於安全性訓練和教育，提供更自然、自訂式和互動式的學習體驗及自動化評估等。

Dell AI Factory with NVIDIA

透過業界首創的全方位一站式 AI 解決方案，加速您的 AI 旅程，並安全地將您的資料轉化為深入見解。Dell AI Factory with NVIDIA 可滿足企業在運用 AI 和生成式 AI 方面的複雜需求。透過領先業界的基礎結構和服務，再加上 NVIDIA AI 軟體，您可以簡化開發和部署作業，使專案能更快將時間轉換為價值。

- 運用具備內建安全性的基礎結構，包括信任根和其他重要功能，降低遭到入侵的風險。
- 透過您控制的內部部署 AI 解決方案，保護您的資料免遭可能導致損失智慧財產權的外洩情形。
- 透過安全存取功能將 AI 技術導入資料中，以滿足嚴格的法規遵循和資料主權要求。
- 透過控制有權存取資料的位置和人員，保護利益關係人的隱私。

Dell AI Factory with NVIDIA

業界首款端對端企業 AI 解決方案

資料驅動 AI Factory 和您的使用案例

您最重要的資料位於內部部署和邊緣裝置。Dell Technologies 協助您運用 AI 處理重要的資料，也是儲存、保護及管理這些資料的業界領先廠商。

從使用案例到成果

AI Factory 利用您優先順序最高的使用案例來產生業務成果。Dell Technologies 透過經驗證的解決方案和量身打造的服務，協助您簡單部署最重要的 AI 使用案例。

別讓安全風險成為創新的阻礙

讓我們協助您自信穿梭 AI 與 GenAI 的世界，收穫豐碩的果實。

策略規劃

免費且適用於 GenAI 的 Accelerator Workshop

- 展開制定制勝策略的過程
- 解決挑戰和差距，確定目標的優先順序並找出機會
- 取得準備程度評估，深入瞭解基礎結構需求、AI 模型、營運整合等

技術準備

立即可用的行動實驗室

快速展開您的成功之旅。包括搭載 NVIDIA GPU 的 Dell Mobile Precision Workstation 5690 / 7780，以及兩天的諮詢服務，協助您開始使用。

- 用於 GenAI 測試和示範的可攜式沙箱環境
- 透過 NVIDIA AI Workbench 平台預先驗證，可立即供開發人員使用
- 以您的資料實作的初始聊天機器人使用案例
- 以具成本效益且低風險的方法，進行實驗和培訓 GenAI 技能

搭載 NVIDIA GPU 的 DELL MOBILE PRECISION WORKSTATION 5690 / 7780

立即開始

DELL Technologies AI Factory WITH NVIDIA