DELL Technologies

5

安全地充分運用 GenAI 的建議

1	2	3	4	5
保護 GenAI 系統的各層	**啟用零信任原則**	**維持治理和人員監督**	**善用可用的 GenAI 安全性工具**	**放心創新**
基礎結構	絕不信任，一律驗證	讓關鍵利益關係人參與	內容	以網路安全性為目標，促進而非阻礙任務
作業系統和 Kubernetes	最低權限存取	制定道德和法規遵循、資料管理原則	風險預測	讓網路安全成熟度建立起組織創新的信心
GenAI 應用程式	系統強化	監控和執行問責制	知識與自動化	
資料	身分識別管理	訓練和教育		
	分段			
	記錄、監控和稽核			

生成式 AI 技術承諾帶來轉型能力，但也帶來了獨特的安全性挑戰。

生成式 AI 正以前所未有的方式革新業務，推動創新並提供無與倫比的優勢，從而帶來競爭優勢。雖然這項技術具有轉型潛力，但也帶來了一系列安全挑戰。

服務產品經理 Steve Brodson、網路安全顧問 Eitan Lederman，以及 APEX 和 AI 行銷團隊的 Chris Cicotte 這幾位 Dell 主題專家，共同探討這些問題，並討論如何安全地發揮 GenAI 最大效能。請繼續閱讀對談摘要和有關此主題的其他見解，並在 dell.com/cybersecuritymonth 觀看完整的討論。

> 「重點在於訓練人員。人員需要知道如何使用 GenAI 系統，例如該做什麼，以及不該做什麼。」
>
> ***Eitan Lederman***
> *Dell 網路安全顧問*

保護 GenAI 系統的各層

雖然 GenAI 是一項相對較新的技術，但大多數安全通訊協定，與用於保護其他工作負載的既定網路安全技術相同。

基礎結構 - 著重於將攻擊面降至最低：

- 漏洞和滲透測試
- 修補
- 強化
- 身分識別管理，包括強式密碼、多因素驗證 (MFA)
- 監控和稽核
- 確保第三方供應鏈安全無虞

作業系統和 Kubernetes - 也是減少攻擊面的重點，包括：

- 漏洞掃描
- 定期修補
- 更新 Kubernetes 元件
- 根據身分識別管理、角色型存取 (RBAC) 和最低權限存取，限制存取控制
- 保護控制平面，包括 API 伺服器、密碼、kubelet 和其他元件
- 使用命名空間

GenAI 應用程式 - 針對 GenAI 建立的新攻擊面實施安全動作：

- 身分識別管理，以解決提示插入、機密資訊披露、模型盜用、訓練資料中毒等問題
- 資料來源驗證，以防止訓練資料中毒、模型偏差
- 監控和稽核，以識別和防止模型 DOS、模型盜用、機密資訊披露，用於進行異常偵測、鑑識

資料 - 結合強大的資料保護措施，以保護語言模型和應用程式中的資料：

- 實體隔離網路存放庫
- 加密
- 事件應變計畫
- 監控和稽核訓練資料與輸出

確保將資料保護原則應用於所有資料，包括訓練輸入、模型輸出，以及檢索增強生成 (RAG) 中涉及的任何資料 (如有使用)。此外，確保持續遵守所有適用的資料保護法規。

啟用零信任原則

前面已經提過身分識別管理、最低權限存取、系統強化和修補等數種零信任原則的作用，這些都指出了零信任原則在保護 GenAI 工作負載方面的價值。零信任架構也需要持續記錄、監控及稽核網路活動，以防範 GenAI 特定風險，例如結果操縱和資料中毒。

DELL Technologies

此外，零信任還鼓勵微切分，從而減少系統入侵的影響。零信任也需要傳輸中和靜態資料加密，這是整體資料保護策略的重要組成部分。

雖然這些只是零信任可保護 GenAI 工作負載的部分方式，但採用零信任原則應視為最佳實務做法。

維持治理和人員監督

GenAI 大部分的價值在於自動執行人類通常會執行的任務，但人類治理對於確保應用程式的安全性和正常運行至關重要。治理模型通常涉及整個組織的關鍵利益關係人，他們為道德和法規遵循、資料管理原則和程序制定指導方針和要求，並最終執行問責制。

適當的治理和監督有助於解決模型過度依賴、偏差、結果操縱、機密資訊披露和資料中毒等問題。

Lederman 還指出了訓練的重要性：「重點在於訓練人員。人員需要知道如何使用 GenAI 系統，例如該做什麼，以及不該做什麼。」

除了組織的 GenAI 應用程式帶來的風險外，採用 GenAI 技術的網路攻擊也隨之大增，這通常需要人為介入。例子包括使用深度偽造來煽動人類行為的惡意分子，以及透過更準確地模仿人類的寫作或說話風格，而讓網路釣魚攻擊變得更加有效。持續的訓練和教育是應對這些風險最有效的方法之一，這也再次加強了人的因素。

善用可用的 GenAI 安全性工具

雖然大部分焦點都集中在風險上，但 GenAI 也有潛力加強安全工作。雖然這些功能還處於起步階段，但將在三個關鍵領域提供優勢：

- **內容：**產生安全原則、個人化訓練、資料分類和報告
- **預測：**風險和攻擊活動，建議補救措施
- **知識：**查詢環境 (與系統對話)、鑑識、自動化

GenAI 對安全性工具的貢獻，有助於最大限度地提高安全團隊的能力、降低成本並強化防禦。隨著這些解決方案的成長和成熟，加以善用。

放心創新

最重要的是，不要讓安全性風險阻礙您運用潛在的革命性技術。效率、自動化、降低成本、解決問題和推動創造力只是 GenAI 實現業務轉型的部分方式。

雖然 GenAI 需要強大的、有時甚至是新的網路安全措施，但目標應該是促進組織的使命，而不是成為阻礙。制定正確的網路安全策略，應該能讓組織放心成長和創新。

造訪 **dell.com/cybersecuritymonth**，瞭解如何解決當今一些主要的網路安全挑戰

DELL Technologies