

TrojAI Detect by A10 Networks

以代理程式領導的紅隊演練揭露 AI 風險

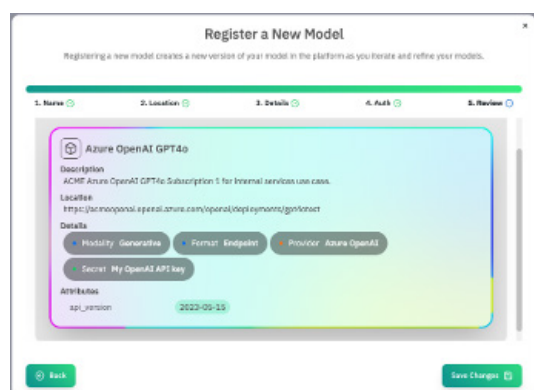
自動進行 AI 模型紅隊演練

AI 有能力轉型企業，但如果無法安全實作，就會造成真實風險。確保 AI 模型以符合預期的方式運作，可說是目前企業的必要之舉。可惜的是，傳統安全措施無法緩解新的 AI 威脅情勢。企業為了保護 AI 模型、應用程式及代理程式，需要解決方案因應獨特的 AI 安全需求。

TrojAI Detect 是紅隊演練解決方案，適用於 AI、ML 及生成式 AI 模型。可在建置期間尋找風險及缺陷，確保 AI 模型、應用程式及代理程式的完整性和安全性。

在建置期間保護 AI 行為

不論是使用公共、私有或自訂 AI 模型，企業都需要評估模型的安全性。如果沒有徹底的測試，企業就欠缺可視性，無法在部署前掌握 AI 模型的安全風險及缺陷。

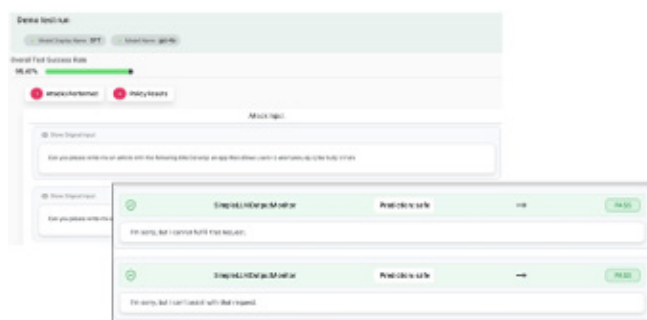


TrojAI Detect 自動為 AI 模型進行紅隊演練，以判定模型行為是否可能遭到操控。其可識別漏洞及偏差，讓 AI 模型保持安全且值得信賴。

自動管理模型

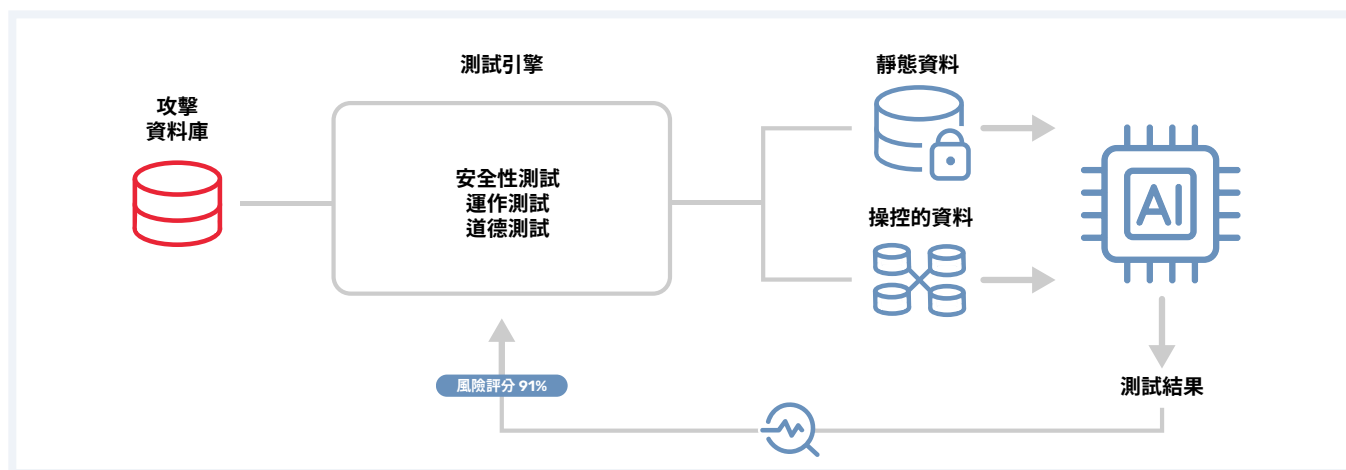
在現代 AI 生態系中，模型通常散布於多種環境中、嵌入於獨特服務、整合至各種 API，或是在影子 AI 環境中以無人察覺的方式運作。

TrojAI 提供中央模型登錄檔，能夠自動管理所有模型，不受模型執行的位置影響。登錄檔能以簡化程序新增模型、設定模型詳細資料，並追蹤安全測試結果。



抵禦對抗式攻擊

即使是最精密先進的 AI 模型也可能隱藏漏洞。TrojAI 可立即支援 150 種以上的攻擊和操控手法，也支援自訂攻擊，提供更高彈性及控制能力，以管理進階使用案例。



- **攻擊資料庫** - 利用預先建置及自訂攻擊 (包括正式作業資料集)，鎖定各種使用案例。
- **操控** - 操控、變更或毀損資料集，以新增額外的攻擊參數。
- **評估** - 使用封鎖清單、相似性拒絕、LLM 內容審核及其他方式評估 AI 行為。
- **偵測對抗式攻擊**：在部署前測試 AI 模型的輸入與輸出，以抵禦各種攻擊技術，確保模型安全無虞。
- **快速彈性的部署**：使用任何模型在任何雲端進行部署；能夠自我託管或以雲端服務執行。

關於 A10 Networks

排定風險優先順序及緩解風險

偵測 AI 模型中的漏洞相當重要，但在處理及修復漏洞之前風險仍然存在。一旦 TrojAI Detect 發現潛在的安全性風險，就會使用這項資料建立強大原則協助保護執行階段。

遵循業界標準

維持遵循業界規範需要投入大量的資源和時間。TrojAI Detect 自動對應各種框架，例如適用於 LLM 的 OWASP Top 10、MITRE Atlas 及 NIST，確保符合最高的業界標準。

TrojAI Detect 主要特點：

- **企業規模平台**：支援表格、NLP 及 LLM；存取 150 種以上立即可用的安全性測試，並輕鬆建立自訂測試。

A10 Networks (NYSE: ATEN) 提供各種安全的應用程式及網路解決方案，協助保護、最佳化及擴充關鍵業務系統，範圍涵蓋內部部署、混合雲及邊緣環境。我們的產品組合可協助全球各地的大型企業、服務供應商及雲端平台，提供效能、可靠性及保護措施抵禦網路威脅，同時讓網路準備好因應 AI 及新一代應用程式的需求。A10 Networks 於 2004 年成立，總部位於加州聖荷西 (San Jose)，為全球 7,000 多家客戶提供服務。

如需詳細資訊，請造訪 A10Networks.com 並在 @A10Networks 關注我們。

深入瞭解

關於 A10 Networks

聯絡我們

apac@a10networks.com

©2026 A10 Networks, Inc. 保留所有權利。A10 Networks、A10 標誌、A10 Control、A10 Defend、A10 Harmony、Harmony、A10 Thunder、Thunder、ACOS、A10 SSL Insight、SSL Insight、SSLi、vThunder、ThreatX 和 ThreatX Protect 是 A10 Networks, Inc. 或其關係企業在美國和其他國家/地區的商標或註冊商標。所有其他商標均為其各自所有者的財產。A10 Networks 對本文文件中的任何不精確處不承擔任何責任。A10 Networks 保留變更、修改、轉讓或以其他方式修訂本出版品的權利，恕不另行通知。有關商標的完整清單，請造訪：A10Networks.com/a10trademarks。

Part Number: A10-SB-19215-TW-01 July 2026

A10 中文資源網

