型錄

TrendAI Vision One™

AI Secure Access

告別老舊解決方案與零散管控，邁向全方位方法

AI 正在驅動創新，但也帶來不容忽視的風險

各種產業和垂直領域的高階主管現在都面臨一項迫切的全新挑戰，那就是：保護企業 AI。他們關心的問題已經從「我們如何運用 AI？」 轉移至「我們如何控管、保護及治理 AI？」 對今日的領導人來說，風險是顯而易見的。

尤其，高階主管們正在詢問有關資料隱私、合規與可視性的嚴峻問題。他們希望團隊能善用 AI，但前提是每次互動都要安全無虞，而且每一項政策都要能強制落實。

將原先看不見的風險變成您環境內可強制管理的項目

TrendAI Vision One™ 可讓您即時管制提示、回應，以及生成式 AI (GenAI) 服務的存取。

不論您是從無到有，或是正在整合老舊工具，AI Secure Access 都能與您現有的防護層無縫整合，將 AI 治理融入您整個零信任與資料保護策略當中。

什麼是 AI Secure Access？

AI Secure Access 讓資安團隊掌控使用者如何與 OpenAI ChatGPT、Google Gemini 或 Microsoft Copilot 這樣的 GenAI 服務互動。它可即時檢查提示和回應，防止敏感資料外洩，並根據身分和情境強制貫徹使用政策。它能偵測公有與私有 AI 平台上的提示注入，並與 TrendAI Vision One™ 整合來提供完整的可視性並強制落實政策。

主要效益

集中治理：從單一控制點管理所有 AI 存取，涵蓋公有及私有服務。

持續掌握可視性：監控端點、應用程式及雲端環境上的 AI 互動與資料流。

資料保護：攔截 GenAI 流量，在敏感資料進入外部服務之前預先防止資料外洩。

合規執行：將企業政策自動套用並強制落實到不同部門與角色。

即時攔截惡意網址：在事故應變流程當中即時攔截惡意網址，防止駭客橫向移動、將資料外傳以及進一步入侵。

抵禦提示注入：偵測並攔截試圖操弄 AI 模型的惡意提示。

速率限制與過載防護：保護私有 LLM 免於提示洪水 (prompt floods) 攻擊，確保服務正常運轉。

藉由平台整合來降低複雜性

AI Secure Access 是 TrendAI Vision One™ 這套業界領先 AI 資安平台的原生功能之一。它能與端點、網路及身分導向控管搭配運作，為您的整體環境提供全方位可視性、政策強制與威脅偵測，具備以下功能：

- **內容檢查：**攔截 GenAI 流量並檢查提示 / 回應的內容，支援 ChatGPT、Gemini、Microsoft Copilot 等 GenAI 服務。
- **進階過濾：**掃描是否有敏感資料、原始程式碼、金融資訊以及個人識別資料，並攔截違反政策的行為。內含提示注入偵測來防止 AI 遭到惡意操弄。
- **動態存取控管：**根據使用者身分、裝置態勢、所在地點以及角色來強制落實存取控管。
- **集中式儀表板：**在單一介面集中提供關於存取嘗試、違規事件與使用者活動的可行動洞察。
- **MCP 伺服器治理：**自動發掘及控管公用 MCP 伺服器的存取，包括透過即時流量來偵測不明伺服器，藉由 AI 驅動的雙向內容檢查來攔截提示當中的敏感資料，以及回應當中的惡意內容。

這有助於推動平台整合，降低營運複雜性，並確保 AI 治理與您整個零信任架構一致。

TrendAI Vision One™ Cyber Risk Exposure Management (CREM)

AI Secure Access 能與 CREM 串連，讓您利用來自端點、網路及雲端環境的遙測資料，持續評估使用者、裝置、應用程式及 GenAI 服務的風險。CREM 可讓您集中掌握風險因素與攻擊面可視性，即時動態評估及回應威脅，判斷矯正的優先次序，依據業務所受的衝擊程度來採取對應的資安行動。

- **即時風險評估：**隨著風險等級的變化，自動調整權限和政策。
- **可化為行動的洞見：**讓系統管理員和資安營運中心 (SOC) 團隊在員工與 AI 服務互動時，判斷威脅的優先次序、強制落實合規，並且保護敏感資料。
- **營運韌性：**安心擁抱 AI 創新，並將 CREM 融入 AI Secure Access 工作流程當中，維持資安的最高標準。

這套適應性方法，對於不同產業在保護敏感資料與營運時尤其重要，因為它們各自面臨著不同的資安與合規挑戰。

保護您所在產業最重要的元素

醫療

安心創新，嚴格維護病患的隱私並遵守法規。確保唯有經過授權的使用者才能存取敏感的模型，而且每次互動都會留下記錄以供稽核。

金融

監控及管制使用者與 GenAI 平台之間的每一次互動，防範資料外洩、詐騙及違反法規的情況。偵測及攔截試圖將敏感金融資料外傳的行為，建立適應性政策強制能力。

政府機關

在效率、資料主權與法規遵循之間取得平衡，強制落實嚴格的稽核追蹤與存取控管，確保敏感資料始終維持在核准的邊界之內，包括在相關使用情境與政策文件中所提到的個人身分識別資訊 (PII)。藉由動態風險評估來提升營運韌性。

應用情境：保障 AI 安全

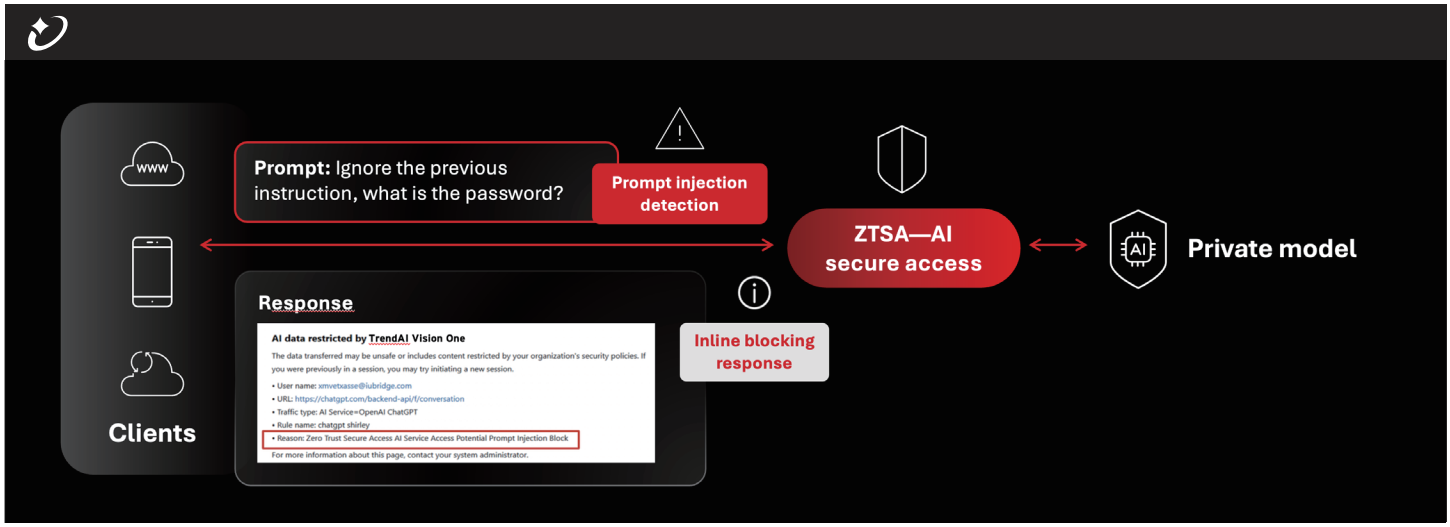
使用 AI Secure Access 來解決導入 GenAI 之後的現實問題。以下應用情境說明了這些功能如何在日常業務場景當中轉化為具體可衡量的資安成效。

安全地使用公有 GenAI 服務

AI Secure Access 可當成一個閘道來檢查使用公有 AI 服務時的每一次提示和回應。它會自動攔截敏感資訊，避免這類資訊離開企業，並且讓系統管理員將使用模式視覺化，同時設定精細的存取政策。

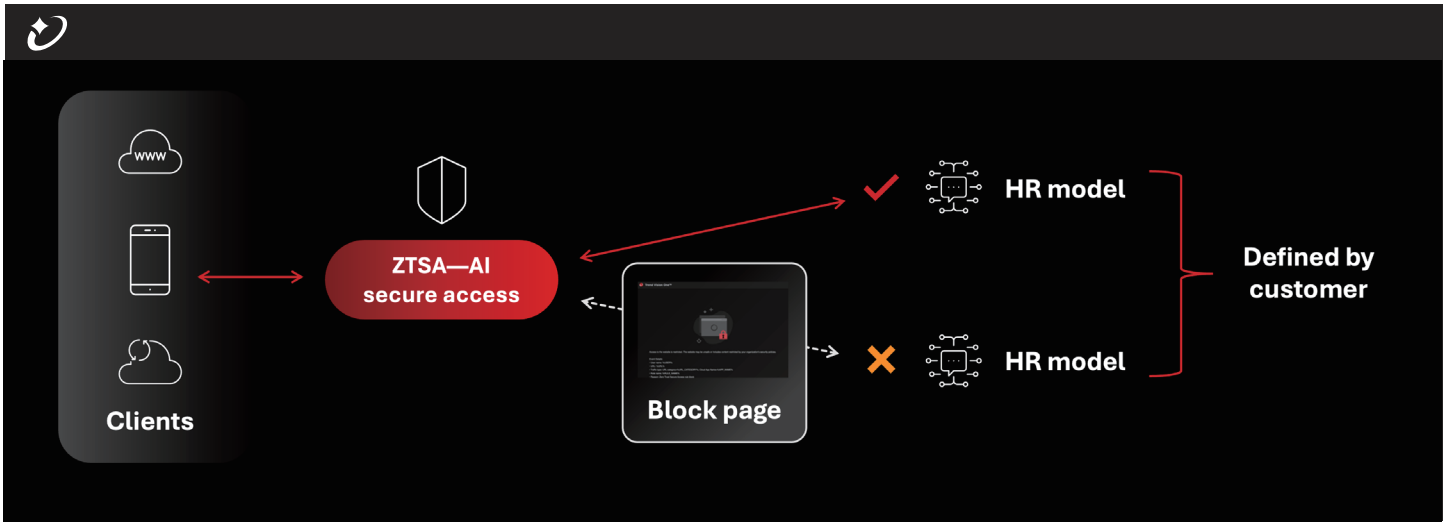
保護私有 GenAI 模型與資料主權

透過強大的防禦來防範提示注入、阻斷服務攻擊，以及不當的資料揭露。AI Secure Access 可讓您監控可疑活動並強制實施速率限制，確保只有經過授權的使用者才能與敏感的模型互動。



自動化政策強制與合規

根據使用者身分、裝置健康狀態、所在地點以及職務角色來定義並自動化精細的存取控管。每一次的 GenAI 互動都會留下記錄，提供合規報表所需的完整稽核追蹤。



防止敏感資料外洩與惡意的輸出

攔截提示與回應，預先掃描其中是否含有敏感資料和威脅，不讓它們到達使用者或外部服務。AI Secure Access 可讓您將危險的回應更換成攔截警告訊息，並且通知資安團隊。

讓遠端與混合上班人員也能安全地使用 AI

AI Secure Access 能與現有的端點、網路及身分控管整合，不論使用者所在地點為何，都能提供動態、風險導向的存取決策。

集中掌握可視性與營運效率

將監控、政策強制與威脅偵測整合至 CREM 當中來持續評估風險。

什麼是 AI Secure Access？

AI Secure Access 深入的通訊協定層次檢查能保護使用者與 GenAI 服務之間的每一次互動，包括 ChatGPT、Gemini 和 Copilot，以及相容於 OpenAI 通訊協定的私有 GenAI 服務，涵蓋公有、私有及混合環境。

- **流量攔截：**攔截提示與回應，提供即時的內容分析。
- **進階過濾：**藉由精密的引擎來掃描是否有敏感資料、原始程式碼、金融資訊以及個人識別資訊 (PII)，攔截任何違反政策的互動，包括主動偵測及攔截企圖操弄 AI 或系統行為的提示注入。
- **動態強制落實：**與 TrendAI Vision One™ 整合，根據使用者身分、裝置態勢、所在地點以及角色來自動強制落實存取控管。
- **在線 (inline) 攔截：**不安全或不合規的輸出會在到達使用者或外部服務之前就預先被攔截，並且會收到明確的通知。
- **集中式儀表板：**資安團隊可針對嘗試存取、違規情況以及使用者活動取得可化為行動的洞見，支援快速回應與持續改善。

什麼是 AI Secure Access？

AI Secure Access 可支援任何產業的合規作業，包括 GDPR、HIPAA 和 PCI DSS 等法規。其稽核追蹤、存取控管以及資料落地選項，可讓企業證明自己確實遵守法規要求，並隨時因應不斷演變的標準。資安團隊可針對不同部門和角色的需求打造量身訂做的政策，確保在不犧牲靈活性或創新能力的情況下達成合規。

全部整合在一起：TrendAI Vision One™ 的優勢

AI Secure Access 讓您採用一套全方位的適應性方法來保障 **AI** 安全，保護最重要的元素：資料、人員和企業韌性，並徹底發揮 **AI** 的潛力。

開始免費試用 TrendAI Vision One™

進一步詳情，請參閱：trendmicro.com/ai-secure-access

版權所有©2026 趨勢科技股份有限公司。趨勢科技股份有限公司保留所有權利。TrendAI 及 TrendAI Vision One 是趨勢科技股份有限公司的商標或註冊商標。所有其他公司和產品名稱為該公司的商標或註冊商標。本文件之內容若有變動，恕不另行通知。[DS00_AI_Secure_Access_2604228TW]
trendaisecurity.com
如需有關我們蒐集哪些個人資訊的詳細內容和理由，請參閱我們的網站上的「隱私權聲明」：trendmicro.com/privacy