



如何因應 AI 時代帶來的 資安挑戰

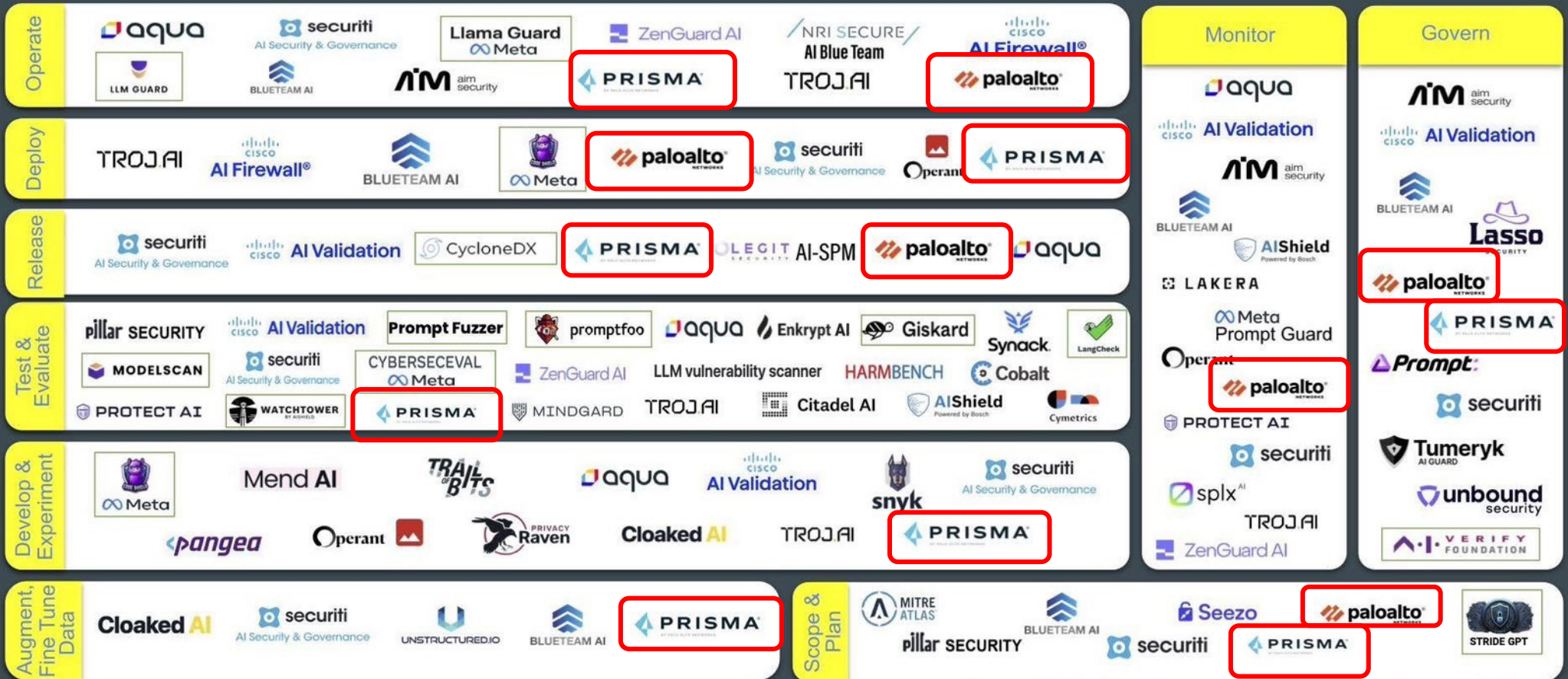
如何安全地應用 GenAI 的服務

CHEAT SHEET

LLM & GenAI Security Landscape – 2025, Q1

<https://genai.owasp.org/ai-security-solutions-landscape/>

Open Source



HOW PEOPLE USE GENERATIVE AI in 2025

This analysis by Marc Zao-Sanders for **Harvard Business Review** groups insights from thousands of forum posts in the last year.



2024		2025	Change in rank from 2024
Generate Ideas	1	Therapy & Companionship	+1
Therapy & Companionship	2	Organize Life	New
Specific Search	3	Find Purpose	New
Edit Text	4	Enhance Learning	+4
Explore Interests	5	Generate Code	+42
Fun & Nonsense	6	Generate Ideas	-5
Troubleshoot	7	Fun & Nonsense	-1
Enhance Learning	8	Improve Code	+11
Personalize Learning	9	Creativity	+18
General Advice	10	Healthy Living	+65
Draft Emails	11	Interview Preparation	+24

Explainers	12	Generate Images	+53
Write & Edit Résumé	13	Specific Search	-10
Excel Formulas	14	Explainers	-2
Email Writing	15	Cooking Guidance	+10
Evaluate Copy	16	Troubleshoot	-9
Improve Decisions	17	Personalize Learning	-8
Translation	18	Boost Confidence	New
Improve Code	19	Email Writing	-4
Draft Document	20	Explain Legalese	+21
Navigate Personal Disputes	21	Child Entertainment	New
Summarize Content	22	Corporate LLM	New
Make a Complaint	23	Student Essays	New
Recommendations	24	Travel Itinerary	New
Cooking Guidance	25	Childcare Help	+8
Generate Appraisals	26	Medical Advice	+2
Creativity	27	Navigate Personal Disputes	-6
Medical Advice	28	Generate Legal Document	+1
Generate Legal Document	29	Conversations	New
Fix Code	30	Anti-trolling	New

VISUAL CAPITALIST

voronoi
BY VISUAL CAPITALIST

Where Data Tells the Story

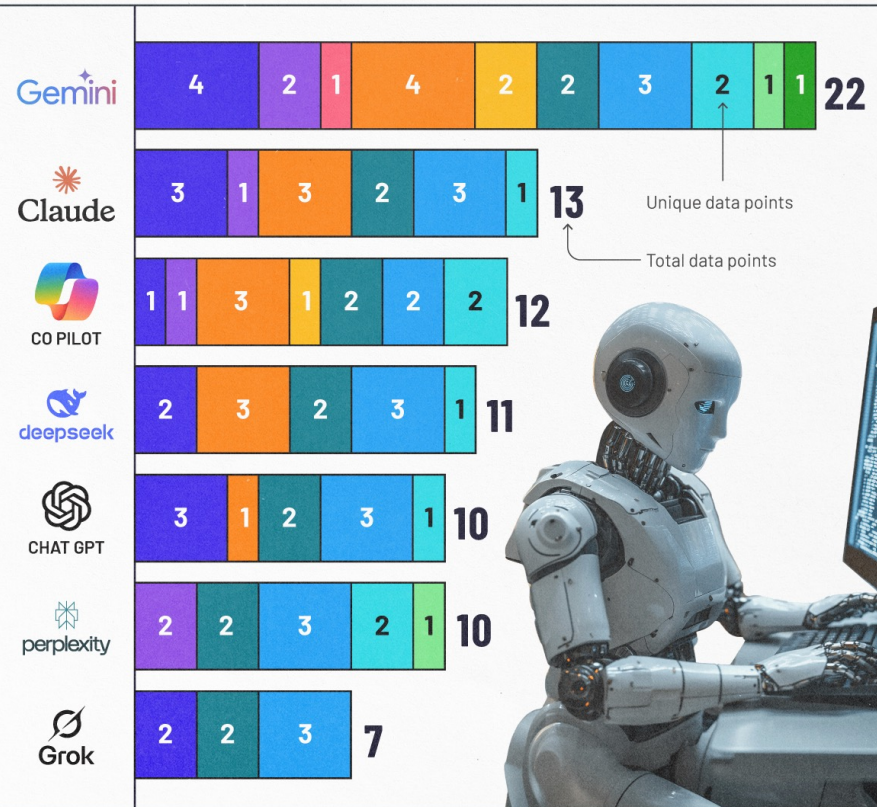
Labels lightly edited from source for readability
Source: Marc Zao-Sanders, Harvard Business Review



WHAT KIND OF DATA

DO AI CHATBOTS COLLECT?

Contact Info Location Contacts User Content History
 Identifiers Diagnostics Usage Data Purchases Other Data

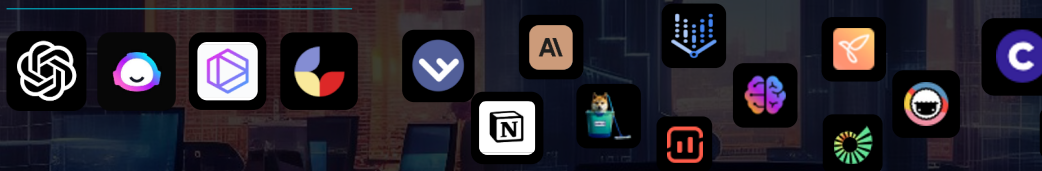


As of Feb 2025. Source: Surfshark

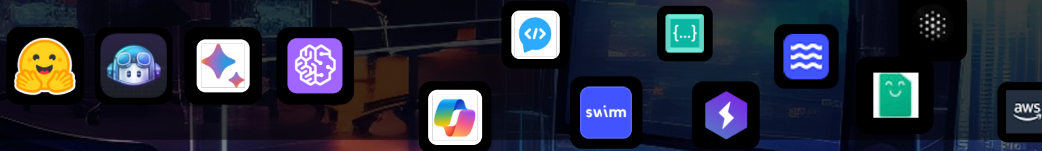
VISUAL CAPITALIST

AI 正在改變員工生產力

Conversational Chats



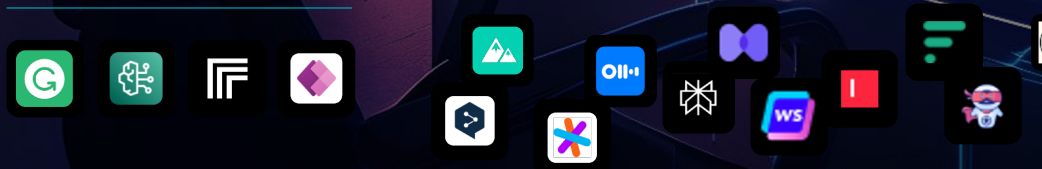
Code Assistants & Generators



Video & Image Generators



Writing Assistants



And more...

原生 AI 應用程式正在爆炸式增長

12,000+

預計到 2030 年使用的 AI 應用程式。

Source: Pitchbook's Generative AI Emerging Space, Artificial Intelligence & Machine Learning Analyst Curated Vertical and SaaS Vertical

隨著 GenAI 應用程式使用量的增加， 安全風險也隨之增加

55%

的員工在工作中使用
過未經批准的 GenAI
工具。



影子 AI 應用程式
產生安全盲點



敏感數據洩漏
源於未經檢查的 GenAI 輸入
和回應



惡意內容
來自 GenAI 回應
對用戶構成風險



IOTW: Samsung employees allegedly leak proprietary information via ChatGPT

Three separate employees have allegedly leaked information to the AI chatbot



US House forbids staff members from using AI chatbot Microsoft Copilot

House Office of Cybersecurity has deemed Microsoft Copilot a risk to users because of the threat of leaking House data to non-House approved cloud services.



Microsoft accidentally exposes 38TB of internal data via GitHub repository

38 terabytes of internal data, including passwords, publicly accessible through a GitHub repository that Microsoft's artificial intelligence research group uses

Source: Salesforce 2023



GenAI 應用程式需要嚴格審查數據上傳

員工在使用 GenAI 應用程式時可能會在不知不覺中使用機密數據。

上傳的數據可用於訓練 AI 模型，可能會洩露專有資訊或導致意外數據洩露。



未檢測到的 GenAI 外掛程式
可能會導致未經授權的數據
訪問

具有 AI 驅動的第三方整合的互連
SaaS 生態系統會產生安全漏洞，使檢
測和控制複雜化。

這會導致安全漏洞和不遵守法規要求
的風險增加。

安全確保 GenAI 應用程式採用需要什麼

挑戰

AI 應用程式創建的自由化導致了不安全應用程式的蔓延。

AI 應用程式存儲、學習和迭代數據。

AI 應用程式攝取非結構化輸入並生成多樣化的輸出。

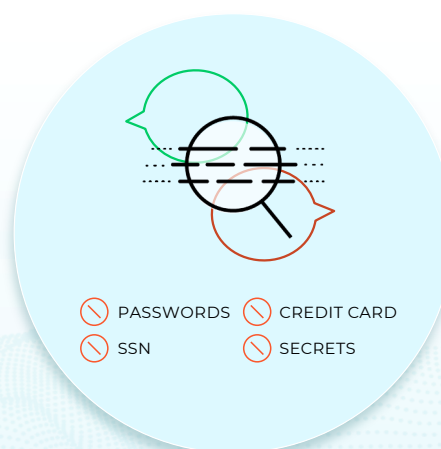
快速發展、功能強大的應用程式可以輕鬆存取。



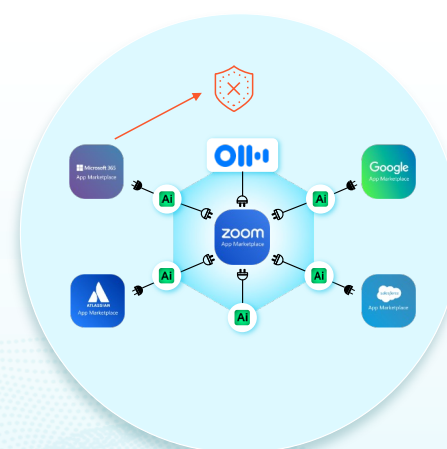
全面的 GenAI 應用程式目錄，保持同步。



了解哪些應用程式正在對您的資料進行訓練



使用基於上下文感知 ML 的檢測器進行精細數據控制。
能夠在 GenAI 回應中防範威脅。



第三方 AI 外掛程式的可視性和控制。
檢測許可權過高的外掛程式。

解決方案

Palo Alto 的做法：使用 AI Access Security 以實現 AI 的安全使用



AI 使用方式的即時可視性

查看使用了哪些 AI 應用程式以及由誰使用。

強大的存取控制

阻止未經批准的應用程式，實施安全策略，並抵禦威脅。

精細的數據保護

掃描共享的數據、金鑰和IP(專利權)。

步驟 1：AI 應用的可視性



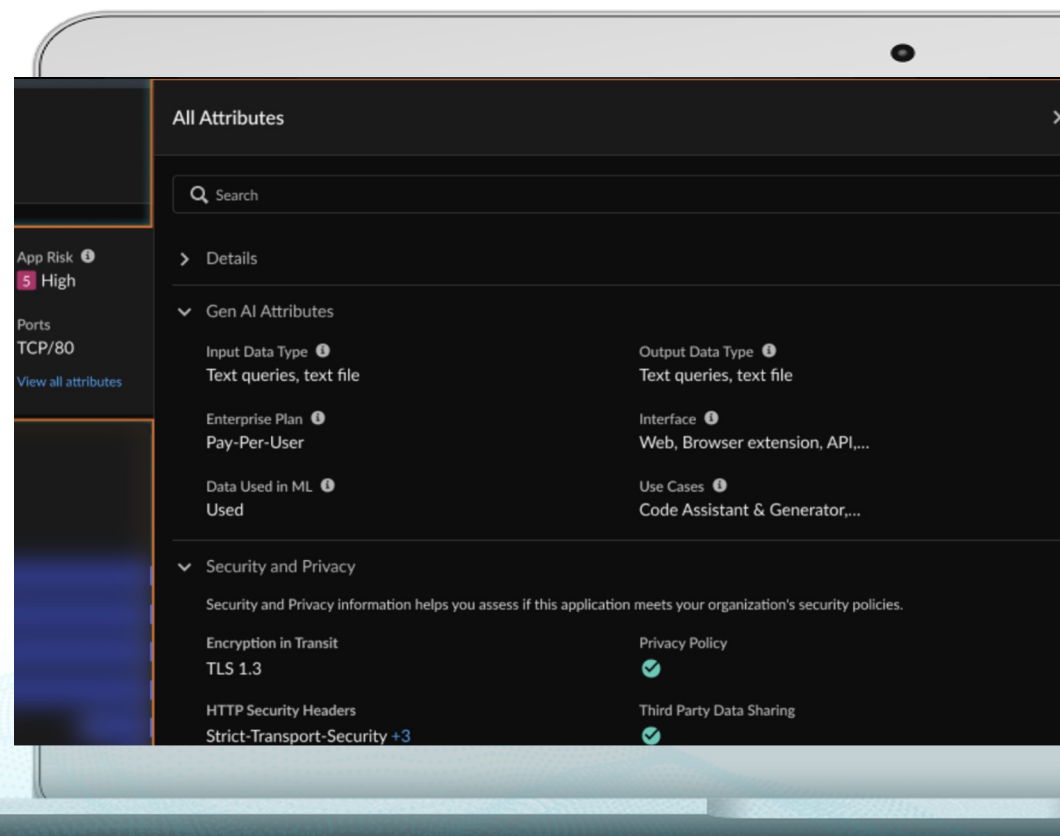
瞭解不同使用案例中的 **GenAI** 應用程式使用方式。



2000+ GenAI 應用程式的詳細目錄。



深入瞭解 **60+** 應用程式屬性。



步驟 2：分類和應用程式存取控制



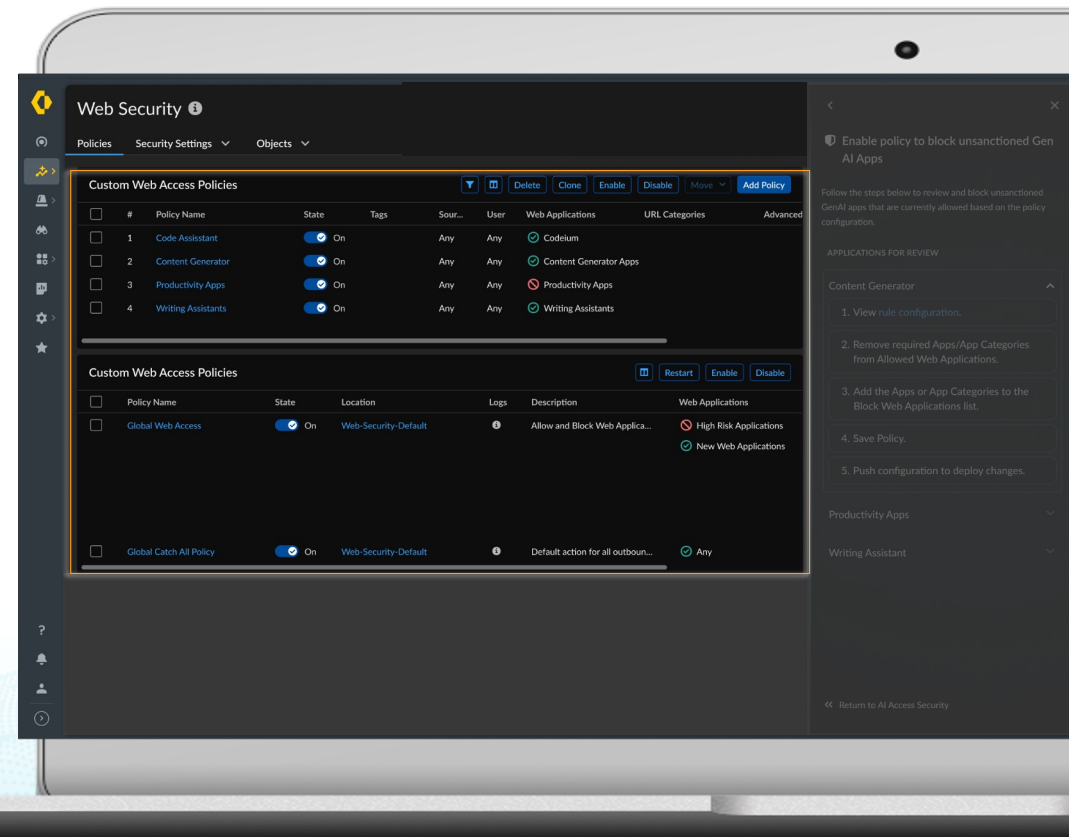
將應用程式分為 Sanctioned、Tolerated 和 Unsanctioned。



為每個使用案例獲得量身定製的可見性和控制措施。



使用 **OOTB (開箱即用)** 最佳實踐策略設置強大的應用程式存取控制。



步驟 3：數據存取控制



利用 LLM 支援的上下文感知 ML 模型，對 **300+** 種類別的數據進行分類。



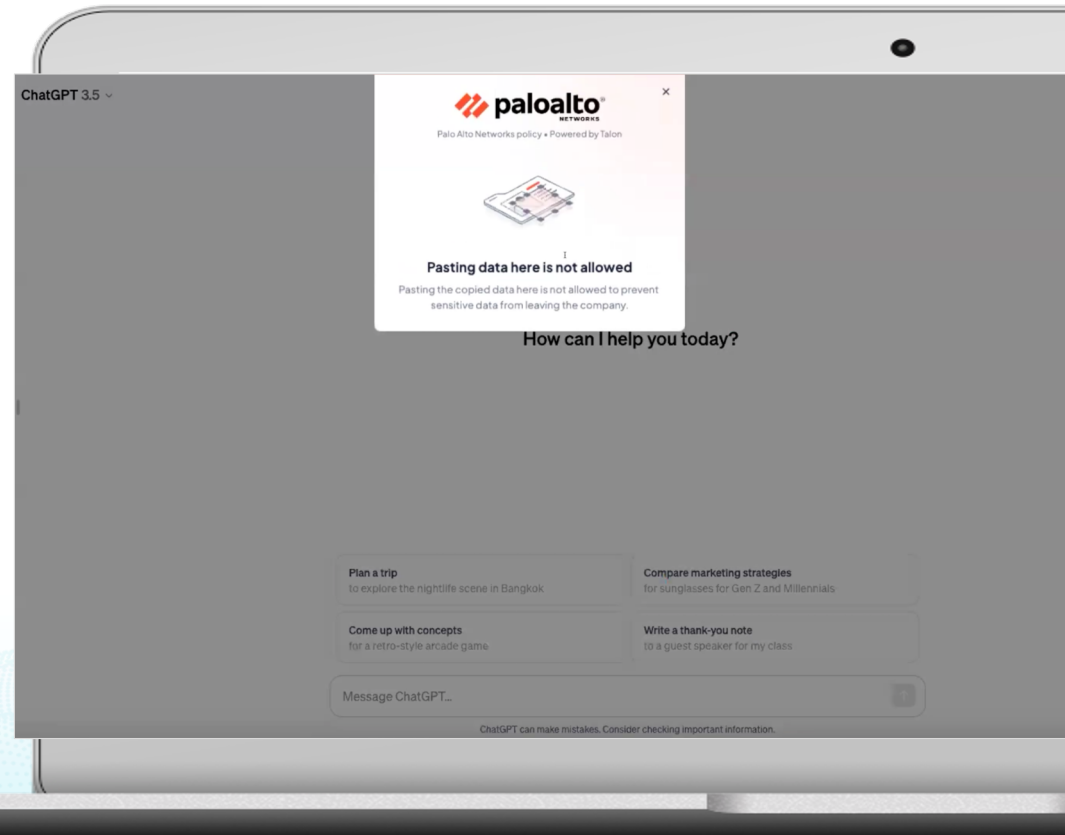
設置上下文內聯策略，以防止敏感數據洩露到 GenAI 應用程式。



通過 Prisma Access 代理(Agent)和瀏覽器整合進行終端用戶教育。



直接通過 **Prisma Access Browser** 查看加密流量。



步驟 4：安全控制



在 SaaS 服務中發現互連的 **GenAI** 應用程式。



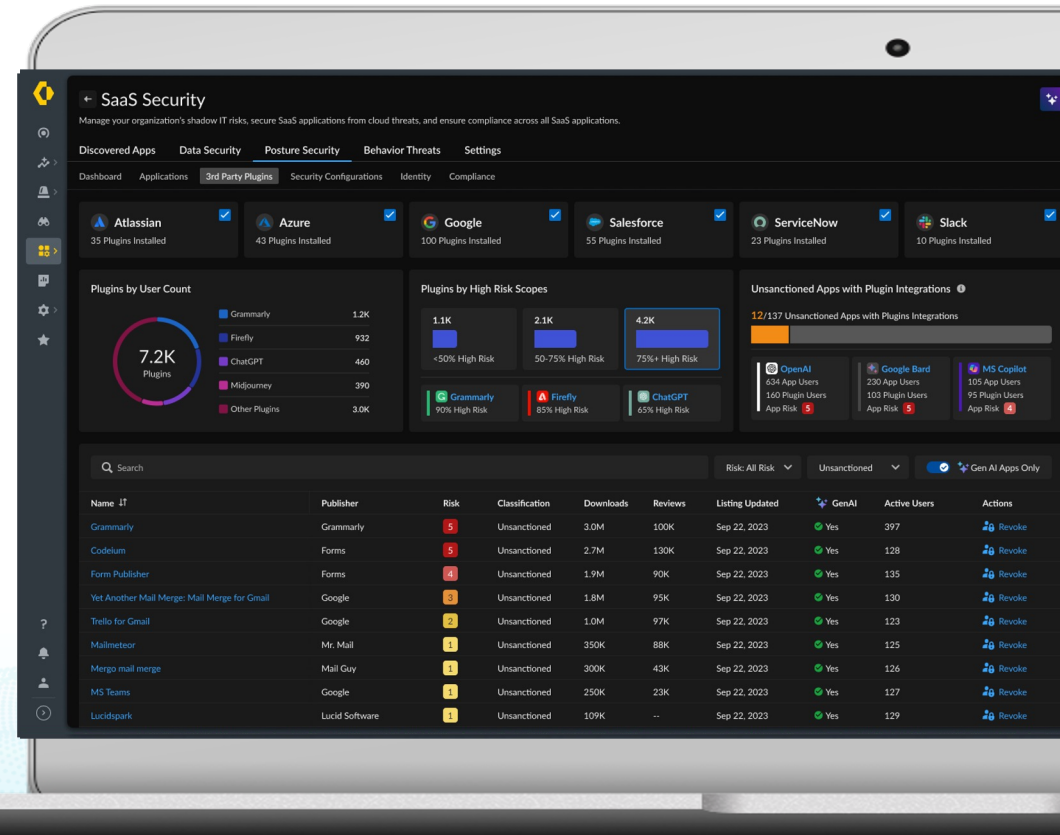
識別、監控和修復未經授權的 **AI** 機器人。



檢測 GenAI 應用程式回應中的威脅（惡意 URL、檔案）。



監控和維護 **GenAI** 應用程式狀況以確保合規性。



Visibility

App Access Controls

Data Access Controls

Security Controls

Continuous Monitoring

步驟 5：持續風險監控



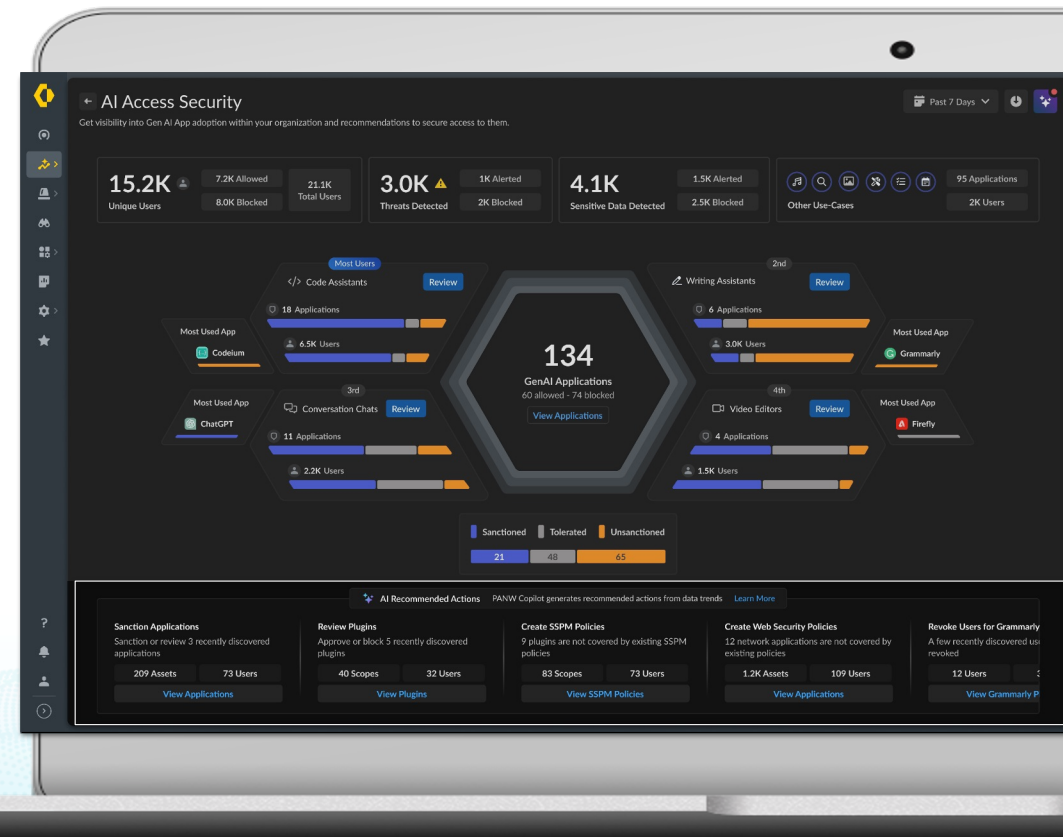
查看各種 **GenAI 應用類別**的採用和使用狀況。



提供 GenAI 應用程式使用方式、風險、安全性和合規性的**綜合報告**。



用於加強 GenAI 應用程式安全控制的**上下文建議**。



參考案例

Palo Alto Networks IT

團隊 試圖...

- 發現網路上的所有 GenAI 應用程式。
- 對現有和新興的 GenAI 應用程式進行分類。
- 防止敏感數據丟失。
- 指導最終使用者安全使用 GenAI。

主要成效

- 發現了 124 個 GenAI 應用程式，每周有 3 個新應用程式。
- 將應用程式分類為「已批准」（11 個）、「容忍」（18 個）和「未批准」（95 個），並能夠自動阻止新應用程式。
- 實施了 12 個新的 DLP 策略，以保護敏感的原始碼、金鑰、訪問 token 和憑證等資訊。



“雖然 AI 應用程式提高了我們員工的工作效率，但它們也可能帶來安全問題。AI Access Security 使我的安全團隊不僅能夠全面瞭解 AI 應用程式，還可以全面做到存取控制和持續監控。”

米拉·拉賈維爾
Palo Alto Networks 首席資訊官

隨時隨地即時防範威脅

統一管理與營運

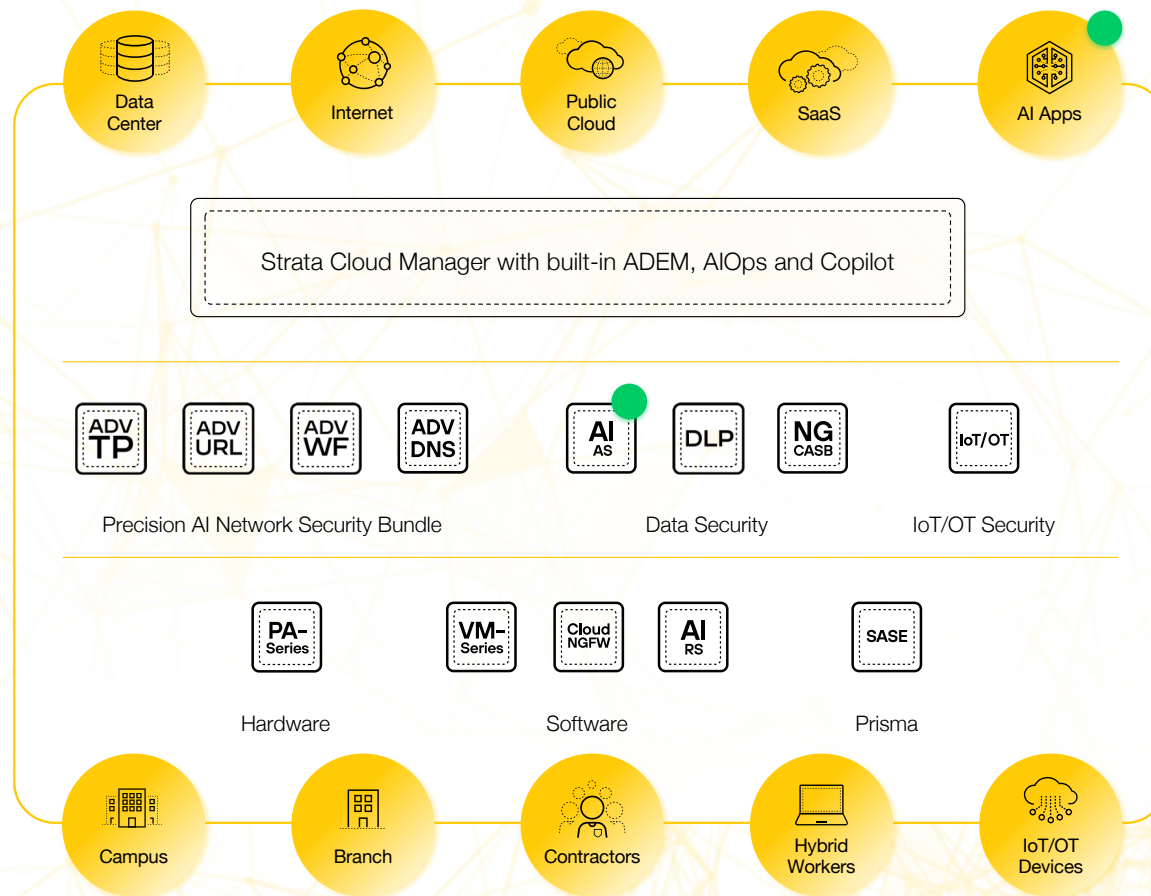
政策制定一次，隨處執行。使用生成式人工智慧主動加強安全性並防止中斷。

AI 驅動的即時安全

使用機器學習和深度學習應用於來自 70,000 多個客戶的豐富資料，即時預防威脅。

單一感測器，多種外形種類

透過一致的作業系統簡化安全性。透過合適的建置保護每個位置。



Palo Alto Networks AI Access Security 保護員工對 Gen AI 的使用



1.Estimated based on data from Tracxn.com

Thank You