



重新思考 SOC :

從線性通道到自主系統

www.stellarcyber.ai | sales@stellarcyber.ai

執行摘要

安全營運中心 (SOC) 已到了一個轉折點。傳統的線性工作流程——最佳化日誌蒐集、規則式告警與人工分級的作業模式——在現今複雜的資安威脅下已不堪重負。現代 SOC 必須進化為具備遞迴能力與智慧判斷的系統：自主式 SOC (Autonomous SOC)。

本白皮書介紹了一種專為 SOC 轉型量身打造的方法，該方法融合了結果導向 AI、被稱為「判斷訊號檢查 (Verdict Signal Checks, VSC)」的結構化驗證機制（此為 Stellar Cyber 所提出的概念與框架）、確保警報通過輕量級檢查才能獲得優先處理的模組化驗證框架、人機協同設計，以及具成本效益的自動化機制。最終呈現的安全營運架構不僅能夠減少噪音，更重新定義決策方式、界定人員介入的時機，以及讓 AI 代理在偵測與回應生命週期中實現規模化運作。

更重要的是，這場轉型已經在發生。率先導入並測試自主式案件分析與分級處理能力的客戶，已觀察到初階 SOC 分析師角色出現根本性的轉變。他們不再進行繁複且深入的調查，而是逐漸轉為協調並傳遞 AI 產生的案件摘要。這並非未來願景，而是正在發生的改變。隨著技術的成熟，我們預期二、三級分析師的角色也將同步演進，他們的工作重點將轉為調校優化、策略規劃與主動式防禦。

目錄

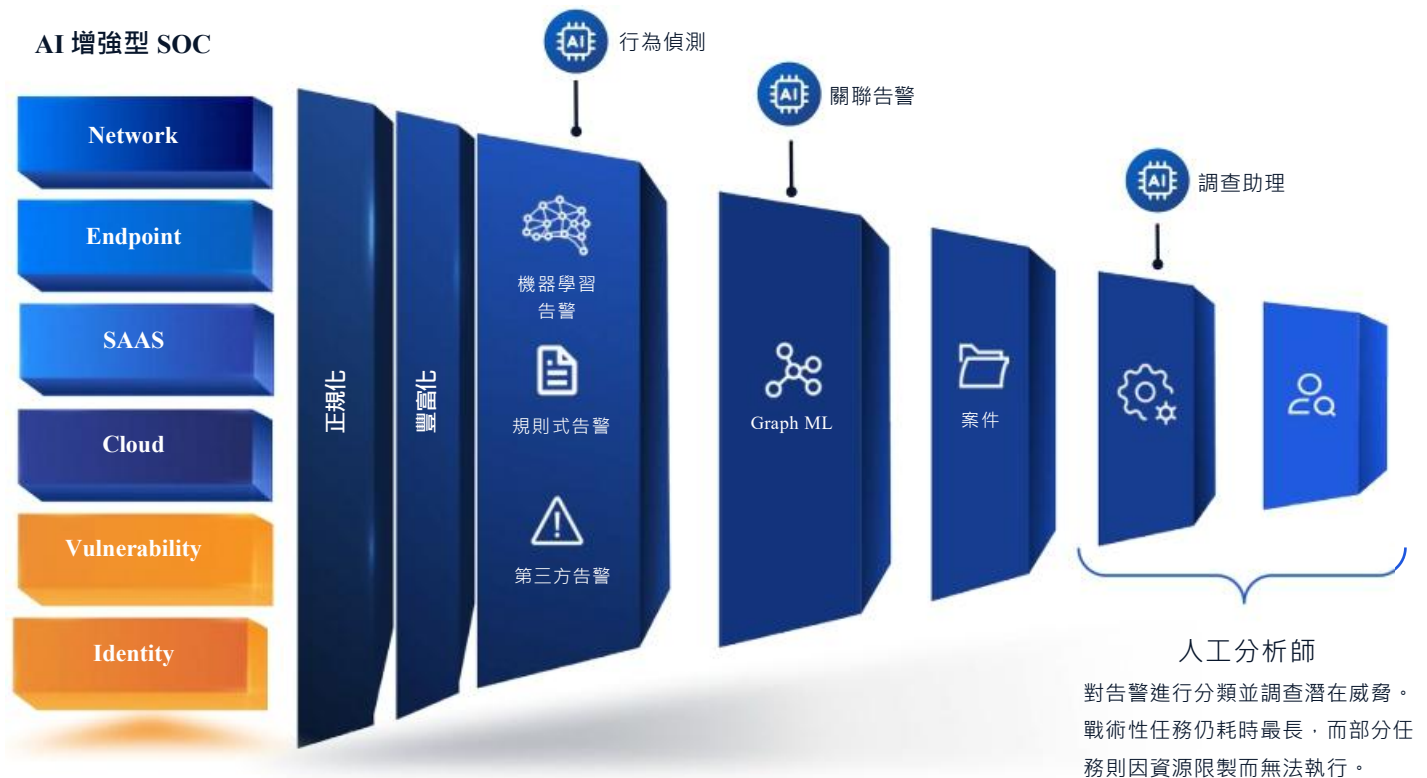
1. 從線性到遞迴：重新思考安全作業流程.....	3
2. 以判斷為核心的 AI：VSC 的角色.....	5
3. 評估樣本集：與 VSC 相匹配.....	7
4. 人機協作的 SOC：影響力金字塔.....	7
5. 資安 AI 經濟學：為什麼光靠 LLM 無法實現規模化.....	9
6. 平台架構：為資安而設計的工程思維.....	11
7. SOC 分析師角色的轉變.....	12

1. 從線性到遞迴：重新思考安全作業流程

二十多年來，安全營運中心 (SOC) 持續遵循著一套可預測的流程：資料收集→規則式告警→人工分類→根據 playbook 進行回應。雖然這種模式能持續運作，但它有三個系統性缺陷：

- **噪音過載**：規則在缺乏情境脈絡下觸發，導致大量誤判，讓分析師被大量誤報淹沒。
- **孤立式調查**：告警往往被各自獨立分級與處理，忽略實體間的關聯性。
- **回應緩慢**：仰賴人工操作，讓調查與遏止時程不斷延長。

自主式 SOC 引入一種遞迴模型，讓每個調查步驟都能產生新的問題、獲取新的情境資訊，為後續行動提供依據。其「遞迴」特性在於，能在調查過程中即時調用原始告警內容外的資料與脈絡——例如，隨著調查的展開，動態引入的補充資訊、威脅情報查詢或實體基準分析。舉例而言，一則「不可能旅行 (impossible travel)」告警，可能會觸發系統取得 ASN 風險資訊、裝置合規狀態或跨雲端服務的登入記錄。這種即時增強機制展示了調查過程如何在初始告警資料的基礎上動態演進。





在傳統的線性工作流程時代，AI 被導入為優化特定步驟的功能：例如，用機器學習進行異常偵測、用 Graph ML 進行跨警報關聯分析，以及用於生成搜尋查詢的 AI 調查助理。這些技術確實降低了作業阻力，但本質上仍屬於漸進式改良。AI 依然只是線性流程中的一項功能，而非典範轉移。

邁向**代理式 AI** 則帶來更具顛覆性的改變。AI 不再只是輔助單一作業步驟，而是成為 **SOC 的核心運作模式**。AI 代理以遞迴方式運作——進行推理、驗證、關聯分析與升級處理——其行為更像分析師，而非靜態工具。調查流程也不再依循線性流程，而是轉為**動態循環**，最終匯聚成可驗證的結果。

人機增強型自主式 SOC

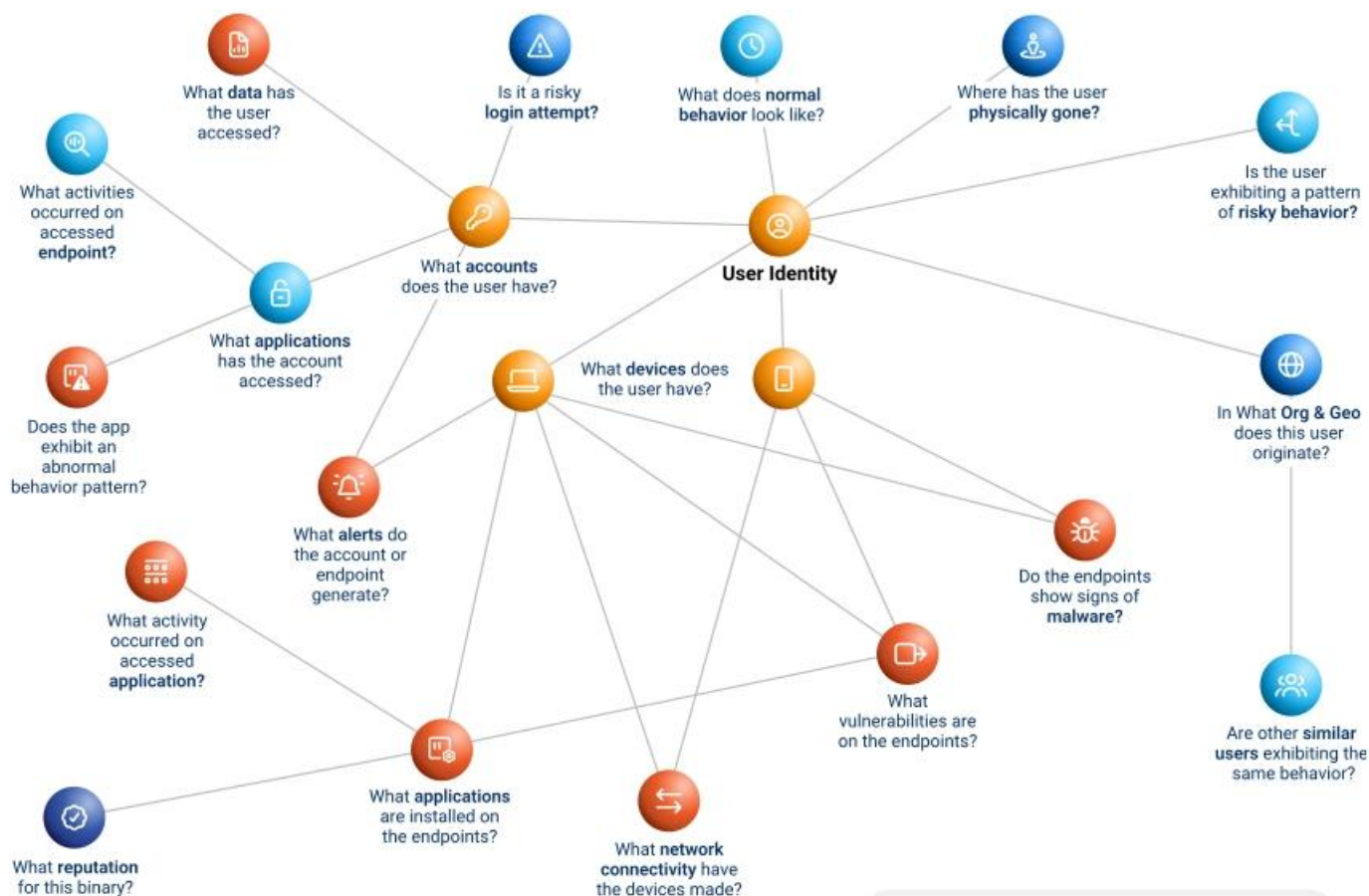


2. 以判斷為核心的 AI : VSC 的角色

傳統偵測系統採取「先觸發、後驗證」的模式：告警由規則或統計模型觸發，再交由人員進行確認。這種方法必然產生大量雜訊，並逐漸削弱分析師對系統的信任。

判斷訊號檢查 (Verdict Signal Check, VSC) 框架則顛覆了這種模式，它不再要求分析師排除雜訊，而是要求偵測結果必須透過結構化且輕量化的驗證流程，逐步提升

可信度。每個 VSC 都是明確的模組化檢核機制，模擬分析師在初步分級處理階段最初幾分鐘的操作。下圖說明 VSC 如何從可觀測指標、事件來源、影響範圍及綜合分析等面向檢查，進而得出最終的判斷。



判斷訊號檢查 (VSC)

- 可觀測指標
- 來源 (發生什麼事?)
- 影響 (代表什麼意義?)
- 後設資訊 (歷史 / 回饋 / 普遍性)
- 威脅等級 (即時威脅情報查詢)

VSC 的類別

此設計框架涵蓋多個面向：



來源檢查：驗證可觀測指標本身是否具有可疑特徵。例如：未知的程序名稱、罕見的父子程序關聯、來自不熟悉 IP 位置的登入行為。



影響檢查：評估惡意事件可能導致的後果。例如：敏感服務存取、異常權限提升、資料暫存活動。



統合分析：與歷史出現頻率及過往判定結果進行比對。例如：「此主機或使用者是否曾觸發類似告警？當時告警是被排除還是確認為真實威脅？」



外部情境補強：從威脅情報 (VirusTotal、AbuseIPDB、IPInfo、MaxMind Pro 等) 取得即時資訊。

實務案例：



端點分級：VSC 會檢查可疑的進程行為。合法的系統二進位檔或受信任目錄中的程序會降低其風險評分；而對於位於暫存目錄、帶有加密參數的可執行檔，則會提高其風險等級。



身分分級：VSC 會驗證是否為不可能旅行 (impossible travel)、高風險 ASN 登入及裝置合規狀態。若登入來自受管理的端點，且該端點先前已與使用者關聯，可能會調低風險；但若是不合規裝置的首次登入，則會提升風險等級。



威脅情報分級：VSC 會評估 IP 或網域風險的可信度。若某個 IP 同時在多個情報來源 (VirusTotal、AbuseIPDB) 被標記，其權重將高於僅在單一來源出現標記的情況。

判斷訊號檢查(VSC)範例

VSC 名稱	主要問題 (VSC #1)	後續問題 (VSC #2)	可能的結果/原因
程序普遍性檢查	在過去 24 小時內，此程序的雜湊值是否曾於超過 20 台不同主機上執行？	該程序雜湊是否經受信任的發行者簽章，或存在於已知安全的目錄清單中？	若為 是 → 良性軟體 (FP)。 若為 否 → 惡意程式散播 (TP)。
密碼異動時間檢查	本次登入失敗前使用者密碼是否在 60 分鐘內更改過？	兩次事件是否均來自相同裝置 / IP 位址？	相同裝置 / IP → 判定為良性誤報。 不同 IP → 可疑事件，列為真陽性 (TP) 候選。
過往分析師判定檢查	過去一週內，該主機上是否曾出現類似警示，且被分析師判定為誤報並排除？	相關情境是否有所改變 (例如：二進位版本、父程序、主機角色)？	情境未改變 → 降低警示風險 (FP)。 情境已改變 → 升級風險 (TP)。

動態訊號強度

每個 VSC 都包含訊號強度（低、中、高）以及執行模式（自動或人工輔助）。例如：

- **Meta-Previous_Override**：先前分析師的判斷，視為高強度自動證據。
- **Meta-Host_Prevalence**：主機反覆觸發相同類型的告警，屬於高強度的輔助訊號。
- **Meta-Previous_Verdict**：先前被判定為誤報的相似告警可能會降低權重（中等強度訊號）。

這種設計確保最終的判斷並非來自單一偵測結果，而是多項檢查累積的推理成果，且每項依據皆可追溯出來源。

為什麼 VSC 至關重要



可解釋性：每個被升級的告警都標示出理由：「由於不可能旅行行為，經地理位置與 ASN 風險驗證後予以升級。」



效率：許多噪音告警在 VSC 階段即被排除，不會增加分析師負擔。



可信：分析師能清楚理解自動化判定之依據，確保其具備推理過程且可被稽核。

VSC 是原始檢測結果和可信結論間的橋樑。它們就像「小型調查流程」，讓 AI 能如一、二級分析師般行動，但又具備機器規模化處理的能力。



3. 評估樣本集：與 VSC 相匹配

為了讓 VSC 發揮效益，AI 代理必須在反映實際運作情況的樣本集上進行訓練和評估。若僅以準確率與召回率作為指標，而忽略對抗性極端狀況或良性但雜訊較大的情境，將難以反映實際效能。

完整的評估框架應：

- 包含乾淨樣本、雜訊樣本及對抗性樣本。
- 涵蓋各種環境（雲端、本地部署、行動裝置、IoT 設備）。
- 同時包含罕見但高影響事件以及常見的良性行為。
- 考量資料時效性問題（日誌延遲、補強資訊缺漏）。

目標不僅是衡量模型的準確性，更要評估在 SOC 情境中的決策實用性。評估機制必須串聯偵測工程、AI 開發與分析師回饋，確保每次風險升級判定在技術上正確，同時在實務運作上具備價值。

4. 人機協作的 SOC：影響力金字塔

SOC 分析師所需的技能正逐漸轉向為 AI 的監督者。在人機協作的自主式 SOC 中，分析師負責驗證、核准、確認或推翻 AI 的判斷與決策。他們的影響力——以及最終績效——取決於回饋的類型與深度。資深分析師理應、也確實會對 AI 產生更大的影響。責任層級可透過人員對 AI 代理所施加的影響程度來衡量。這種現象在許多 AI 驅動的領域中皆可見，但在 SOC 營運中尤為明顯。這是個值得探討的議題，但已超出本白皮書的範圍。

自動化能力固然強大，但分析師的監督不可或缺。為了使兩者形成互補，我們以 **影響力金字塔** 的概念進行延伸，並根據 Stellar Cyber 部落格文章「從痛苦金字塔到影響力金字塔」中描述的反饋影響層級進行了完善。這個強化後的金字塔模型，清楚界定分析師回饋的類型、這些回饋如何影響系統，以及哪些環節中人員介入是不可或缺的。

回饋影響層級

金字塔最底層是簡單的標籤，最高層是策略性變更。每向上一層影響力持續遞增：

1. **被動分類**（影響力最低）：僅標記「真陽性(TP) / 假陽性(FP)」，不包含額外情境說明。這類回饋有助於衡量精確度，但對實質優化的進展有限。
2. **情境註解**：在 TP / FP 回饋中補充說明，具體指出出錯的原因（程序、主機、排程）。這些資訊可直接用於調整偵測規則與告警排除策略。
3. **覆核與排除**：分析師明確指出需停用或調整特定偵測邏輯（例如已知的良性行為、受信任裝置）。此類回饋將影響門檻設定、VSC 權重分配，並從源頭降低雜訊。
4. **偵測創新與邏輯變更**（影響力最高）：分析師提出新的偵測特徵、VSC 分類或模型調整建議。系統邏輯會隨之演進與優化。

人機互動

- **自主層級**：當偵測結果通過多個 VSC 的高可信度驗證，且沒有與過往分析師回饋記錄相互矛盾時，系統無需人工干預即可操作。
- **協作層級**：系統產生告警或判定結果，但將分析師納入決策流程——尤其是在訊號強度接近臨界值，或偵測到異常行為時。
- **判斷層級**：對於罕見、敏感或情境模糊的案件，由人工分析師介入處理，例如內部威脅、策略性決策或政策例外情況。

回饋循環：分析師回饋如何改變系統

1. **收集**：分析師透過操作介面提交回饋：標記、註解或覆核系統輸出結果。
2. **分類與加權**：系統將回饋歸類至對應的影響層級，並記錄其強度、情境脈絡與模式。
3. **執行**：依據不同層級採取相應行動，例如排除後續相似告警、調整模型權重、更新 VSC 設定，或部署新的偵測邏輯。
4. **衡量**：系統會追蹤變更是否降低雜訊、提高精確度或加快回應速度，並將結果回饋至評估機制中。
5. **信任強化**：將成果具體呈現給分析師（例如：「您的排除設定已避免 1,200 則誤判警示」），有助於建立信任並激勵分析師。

「在 Stellar Cyber，我們堅信可視性驅動行動。透過整合來自各種來源的資料，我們的平台改變了資安團隊的工作方式——將數小時的分析縮短為數分鐘的洞察。當組織能清楚掌握全局，便能以自信、快速和智慧的方式採取行動。這正是開放且高度串聯的資安架構所帶來的真正力量。」

— Aimei Wei，技術長暨創辦人

下圖取自 Stellar Cyber 的部落格文章「從痛苦金字塔到影響力金字塔」，可用來視覺化此模型：



5. 資安 AI 經濟學：為什麼光靠 LLM 無法實現規模化

在現今市場中，大型語言模型 (LLM) 常被宣傳為 SOC 的萬用解決方案。許多供應商和新創公司將其定位為可取代現有偵測、關聯和分級流程的引擎。然而，實際情況更為複雜，儘管宣傳中暗示 LLM 可以負責端到端的作業，但經濟和技術現實顯示，它們最適合作為輔助工具，而非核心引擎。

LLM 擅長**加速創意工作**—撰寫正規表示式、建立解析器、設計關聯邏輯，或總結複雜案例。但它們並非為即時、可擴展的偵測而設計。不妨將它想像成「程式碼顧問」，而非整個處理流程的核心引擎。

LLM 擅長的領域

- 產生正規表示式、解析器或關聯邏輯。
- 彙整告警、案件或事件時間軸。
- 用簡單易懂的語言向分析師或高階主管說明威脅情境。

傳統機器學習擅長的領域

- 即時偵測與串流分析。
- 具擴充性的異常偵測（UEBA、DNS 隧道、帳號憑證濫用攻擊）。
- 低延遲、低成本的事件分類。

經濟現實與成本考量

即使 LLM 技術快速成長，它們在大規模資料匯入或即時警示分級處理上的成本仍過高：

- **Token 限制**：情境視窗已擴大 32 倍（自 2023 年約 4000 個 Token 提升至目前約 12.8 萬個）。
- **Token 處理速度**：吞吐量提升約 1500 倍，目前約為每秒 9.4 萬個 Token（約每秒 250 個事件）。
- **成本**：儘管效率有所提高，但若要在 SOC 規模下運行 LLM，每月成本仍可能高達數萬至數十萬美元。



此外，在 VSC 等 AI 代理程式案中，代理程式間可能存在大量的來回呼叫與協作。這種遞迴式推理雖然強大，但也伴隨財務風險：若迴圈未被妥善控制，可能導致 Token 消耗失控。為了確保其永續性，必須建立**成本上限與防護機制**，以避免費用無限制增加。

換言之：LLM 現在已能以極高效率處理深度情境理解與內容摘要，但仍不適合用於串流式偵測、大量日誌解析或大規模告警分類。它們是強化 SOC 的工具，而非取而代之的核心。

實務建議

- 讓生成式 AI 產生解析器，而非讓它處理 2000 萬個日誌。
- 讓它解釋威脅情境，不要期望它監控即時串流資料。
- 將它用於分級處理與摘要整理，而非大規模資料匯入處理。

在合適的階段使用合適的模型至關重要：

- **規則**：快速、可解釋、成本效益佳。
- **傳統機器學習**：專注於特定領域、可擴展、穩定可靠。
- **LLM**：擅長情境推理、摘要整理與創意輔助。

這種混合式架構確保 AI 在經濟上永續、技術上可行且營運上有效益。

成本控制框架

為了控制成本並避免在遞迴式代理工作流程中 Token 過度消耗，應建立完善的成本控管機制：

- **遞迴深度限制**：限制代理程式在升級處理或終止流程前，可互相呼叫的次數。
- **預算上限**：為每個案件、告警或調查設定 Token 或成本上限。
- **備援機制**：當超出預算時，改由確定性規則或傳統機器學習進行處理。
- **監控與稽核**：將 Token 用量、代理互動次數與財務影響納入 SOC 監控範圍。

這些防護措施可確保在充分利用代理推理能力的同時，不會發生成本失控的狀況。

關鍵觀察

為了更直觀地呈現上述權衡與優化效果，總結出以下三點：

- **LLM 功能強大，但並不適用於所有任務**
- **LLM 的力量：成長而非取代（現階段）**
- **大規模應用 LLM 的成本考量**

6. 平台架構：為資安而設計的工程思維

常見的一個問題：這跟將告警內容複製貼上到像 Microsoft Copilot 這類通用工具得出說明，哪裡不同？答案在於架構。我們的流程是專為 SecOps 打造，而非提升生產力。

- **欄位感知處理**：原始 JSON 資料經過最佳化，專注於最關鍵的實體資訊—使用者、IP 位置、主機與時間戳記。
- **擴展性案件處理能力**：系統能處理案件中數十甚至數百則告警，同時維持事件脈絡與敘事一致性。
- **精確時序還原**：透過處理原始 UNIX 時間戳記，確保事件時間軸忠實反映攻擊發生過程。
- **圖形模型降噪**：運用圖形分析技術，在進行摘要前先降低雜訊，解決「大海撈針」的問題。

這正是生產力輔助工具和專為 SOC 打造的處理流程間的差異：前者提供臨時性的說明，後者則在大規模情境下提供可信、可重現的案件分析。

為了實現這點，自主式 SOC 必須遵循一套專為資安場景設計的工程原則，而非沿用一般通用型 AI 平台的思維：

- **佇列感知 AI**：處理案件數量、嚴重性與分析師可用性進行優先排序。
- **成本感知處理**：先執行輕量化調查，將 LLM 運算資源保留給高價值案件。
- **以實體為核心的設計**：將所有相關情境（使用者、主機、檔案、IP）整合成統一的案件視圖。
- **預設可解釋性**：每個決策、評分與行動皆附有可稽核的依據說明。

上述特性讓自主式 SOC 的處理流程不僅具備自動化能力，同時也兼顧透明度、可問責性，並與實際資安營運需求保持一致。



「在 Stellar Cyber，我們的使命是提供能驅動行動的關鍵情報。透過將所有關鍵資料來源整合至單一的開放平台，我們協助 SecOps 團隊在數分鐘內掌握洞察，而非耗費數小時。我們設計的一切都以簡潔、速度及可視性為核心，為每位客戶將複雜轉變為清晰易懂。」

— Subo Guha，產品管理資深副總裁

7. SOC 分析師角色的轉變

以判斷為核心的 AI、VSC 及人機協作模式的導入，從根本上改變了 SOC 分析師的角色。一位客戶在測試這項功能時發現，該系統能在模擬 Web 應用程式滲透測試期間進行深入案件分析。過去需要耗費數小時進行人工日誌檢視與證據拼接的工作，如今可自動彙整為完整的案件敘述，並提供清楚的判定結果。分析師的角色也從逐行調查轉變為**審查、批准和解讀**系統的分析結果。

這代表了一種更廣泛的行業趨勢。下表比較了分析師角色的轉變，並重點說明各層級所需的新技能：

分析師層級	傳統 SOC 角色	自主式 SOC 角色	所需新技能
一級（初級）	手動告警分類、過濾噪音、風險升級	AI 監控、情境補強、流程協調	理解 AI 產出結果、驗證 / 覆核 AI 決策、自動化腳本編寫、溝通與協調能力
二級（中級）	案件調查、關聯告警	引導 AI 回饋循環、優化偵測邏輯、補充與詮釋案件情境	偵測工程能力、查詢語法設計、回饋註解、跨領域知識（雲端、身分管理）
三級（高級）	升級案件、進階調查	策略分析、威脅獵捕、紫隊演練、持續優化偵測與 VSC 設計	威脅獵捕、對抗性模擬、模型 / VSC 設計、指導、策略監督能力

在未來，SOC 的重要性將**不降反升**。各級分析師將從被動的告警分類轉向更高價值的工作：系統調整、監督、策略制定和主動防禦。



結論：前行之路

自主式 SOC 並非要取代分析師，而是以**值得信賴、可遞迴運作、具可解釋性的智慧能力**幫助強化。透過結合 VSC 驅動的驗證機制、完整的評估資料集、優化後的人機協作影響力金字塔，以及混合式 AI 策略，我們可以：

- 消除干擾，讓分析師重新專注於核心業務
- 用較低的成本擴展決策規模
- 為每項行動提供可稽核、易於理解的判斷依據
- 建立一個能隨時間持續調適、學習與優化的 SOC

我們從早期者採用的客戶觀察到的情況證實了這一轉變：**初級 SOC 分析師的角色正在積極變化**。他們不再執行細部調查作業，而是越來越多地協調流程並傳遞 AI 所

產出的綜合摘要。這種轉變並非預測，而是正在發生。早期採用者已經在生產環境中運用自主案件分析與自動分級功能，初階分析師的角色也正逐步轉向更具監督性與高價值的協調職能轉型。

要在這種新模式下取得成功，各級 SOC 分析師都必須培養**新的監督技能**：理解 AI 輸出、掌握自身對模型的影響力類型，以及學會如何有效引導 AI 的決策。這些技能對於最大化人類與 AI 在 SOC 中的效能至關重要。

SecOps 的未來是**循環往復、經過驗證且以人為本的**。它不是線性流程或令人疲憊的儀表板，而是充滿信任和協作的動態循環。

關於代理商 **SPHINXTEC 欣盟科技**
all about data security

欣盟科技 (SphinxTec Co., Ltd.) 以專業的技術服務為基礎，遵循資訊安全的規範，我們以守護客戶的資訊為首要考量，代理銷售不同領域之領導品牌，協助客戶在駭客攻防之中提早預防取得先機，且以欺敵技術降低駭客攻擊突破率，打造堅韌的資安防護網。欣盟科技亦提供資安技術顧問服務可客製化企業需求去分析找出關鍵核心，提供適切的資安策略，配合使用實際的可行性分析，導入適合的企業資安解決方案。

www.sphinxtec.com