生成式 AI 安全性控制

Information Security for Workload Factory

NetApp
September 16, 2026

目錄

生成式 AI 安全性控制

了解 NetApp Workload Factory 中的 AI 控制

NetApp Workload Factory 包含生成式 AI 功能，可加快診斷速度，同時對模型存取、資料範圍和執行時間行為實施安全性控制。

Workload Factory 預設啟用 AI 診斷。您可以隨時透過 Workload Factory **Administration** 設定停用生成式 AI 功能。"了解更多關於管理 GenAI 功能的資訊。"

安全性範圍 AI 功能總覽

Workload Factory 目前將生成式 AI 套用至以下功能：

- Ask Me 助手（RAG，包含精選 NetApp 文件和有來源支援的回應）
- 延遲分析
- EMS 事件分析
- 錯誤日誌分析

適用於 AI 診斷的 AWS Bedrock 總覽

Workload Factory 使用 Amazon Bedrock 進行 AI 推論。推論在您的 AWS 帳戶中執行，使用您設定的憑證和推論設定檔。

相關資訊

- "Amazon Bedrock 選擇加入 NetApp Workload Factory AI 診斷"
- "NetApp Workload Factory 的人工智慧工具邊界"
- "NetApp Workload Factory 的 AI 模型參數與計費"

Amazon Bedrock 選擇加入 NetApp Workload Factory AI 診斷

NetApp Workload Factory AI 診斷使用 Amazon Bedrock 分析事件，此功能預設為啟用。在 Workload Factory 呼叫選定的 Bedrock 模型並收集診斷資料之前，您必須設定 IAM 權限、推論設定檔以及具有 SSH 功能的 Lambda 連結。如果缺少任何元件或 Bedrock 組態已移除，AI 診斷將無法使用。

AI 診斷功能預設為啟用。

開始之前

- 請確保您的帳戶包含 AWS Bedrock 推論設定檔。
- 擁有具備 SSH 功能的已連線 Lambda 連結，以進行診斷資料收集。

步驟

1. 在您的 IAM 角色中授予 Bedrock 權限：
   - `bedrock:InvokeModel`
   - `bedrock:InvokeModelWithResponseStream`
   - `bedrock:ListInferenceProfiles`
   - `bedrock:GetInferenceProfile`
2. 在 Workload Factory 設定中設定推論設定檔 ARN。

如果缺少任何先決條件，系統會報告特定缺少的元件，且 AI 功能將保持非作用中狀態。

停用 **AI** 診斷

若要停用 AI 診斷：

- 從您的 IAM 角色中移除 Bedrock 權限，或者
- 在 Workload Factory 設定中刪除 Bedrock 組態。

AI 診斷立即無法使用，且不會進行任何剩餘資料處理。

區域和跨區域推論設定檔

Bedrock 推理會在您推理設定檔指定的 AWS 區域中執行。如果您使用跨區域推理設定檔，根據 AWS Bedrock 的標準行為，AWS 可能會將請求路由到設定檔組態中的任何區域。資料不會離開 AWS 基礎架構。

NetApp Workload Factory 的人工智慧工具邊界

NetApp Workload Factory 對其 AI 功能與您的 AWS 和 ONTAP 環境互動的方式實施嚴格的界線，因此您可以信賴 AI 輔助診斷，而不會有意外變更的風險。Amazon Bedrock 使用唯讀的 AWS CLI 和 ONTAP CLI 診斷工具，這些工具會區塊修改命令並依大小和執行限制輸出，且所使用的任何資料僅為暫時性，不會由 NetApp Workload Factory 或 AWS 保留。

工具安全控制

工具	功能	安全控制
AWS CLI（唯讀）	執行唯讀的 AWS CLI 指令（describe、list、get）	阻止所有修改動詞（建立、刪除、修改、更新、放置、移除）。每步最多 10 個命令。輸出大小截斷至 20 KB。
ONTAP CLI（唯讀）	使用 SSH 執行唯讀 ONTAP CLI 指令	封鎖所有變動動詞（刪除、毀損、建立、修改、移動）。輸出已截斷。

代理程式無法修改、建立或刪除資源。它只能觀察和診斷。

資料保留和模型訓練邊界

根據 AWS 原則，Amazon Bedrock 不會儲存或使用您的輸入和輸出來訓練或改進基礎模型。對於隨需推論（由 Workload Factory 使用），AWS 會暫時處理您的資料，並在回傳回應後不予保留。

NetApp Workload Factory 的 AI 模型參數與計費

NetApp Workload Factory 定義了與 AI 治理相關的模型組態、使用情況和計費邊界。設定的模型使用確定性參數並傳回結構化 JSON，以支援一致的分析結果。AWS 會將 Amazon Bedrock 推理費用計入您的帳戶，而 Workload Factory 則會追蹤 AI 使用情況並提供每個帳戶的每月成本可見度。這些小節說明了模型參數以及適用於 AI 消耗的限制。

模型和參數

參數	價值
模型	Anthropic Claude（使用跨區域推論設定檔）
溫度	0（確定性，無隨機性）
最大輸出 Token	2,048
最大推理步驟	每次分析 15
輸出格式	結構化 JSON（總結、嚴重程度、根目錄原因、糾正措施）

計費和使用限制

- AWS 會將 Bedrock 推論費用計入您的帳戶（您的推論設定檔、您的憑證）。
- Workload Factory 會在內部追蹤 AI 使用情況。
- Workload Factory 提供每個帳戶的每月成本可見度。

  "在 NetApp Workload Factory 中追蹤儲存設備成本"

Ask Me AI 助理

NetApp Workload Factory 的 Ask Me 安全性架構

Ask Me 是內建於 NetApp Workload Factory 的生成式 AI 助手。它為 Amazon FSx for NetApp ONTAP 和 Workload Factory 主題提供即時的對話式支援。回應均基於經過驗證的 NetApp 文件，且 Ask Me 無法存取公共網際網路或使用客戶資料來訓練模型。多層安全性機制有助於區塊惡意查詢，將回應限制在支援的網域內，並防止使用者輸入覆寫系統指令。當 Ask Me 使用工具時，授權邊界、唯讀查詢強制執行、臨時沙箱和稽核記錄有助於約束和監控執行。

您可以隨時透過 Workload Factory 管理 設定選擇退出 Ask Me，這將停用您使用者帳戶的助理。"了解更多關於管理 GenAI 功能的資訊。"

Ask Me 中的 **RAG** 架構

Ask Me 使用擷取增強生成（RAG）。

- 精心整理的知識庫：回覆基於經過驗證、已發佈的 NetApp 文件所組成的代管語料庫。
- 擷取評分與篩選：模型上下文中僅包含足夠相關的內容。
- 沒有網路存取：此模型無法取得、瀏覽或抓取外部內容。
- 來源標註：回覆包含所用來源文件的引用。

多層提示安全性

- 第 1 層：安全性分類器（LLM 作為評判者）封鎖惡意查詢（提示注入、權限提報、資料外洩、社交工程、原則違規）。

  系統會記錄被封鎖的查詢，生成式模型永遠不會看到它們。

- 第 2 層：系統提示強化可防止使用者輸入覆寫系統指令。
- 第 3 層：網域範圍限制對 FSx ONTAP 和 Workload Factory 主題的回應。
- 第 4 層：工具使用治理在加固的沙箱中驗證參數並執行查詢；此模型不能執行任意操作。

模型資料隔離和隱私

- 沒有使用客戶資料進行模型訓練
- 要求層級隔離（無跨租戶資料外洩）
- 工作階段範圍的對話上下文，具有有限的歷史記錄
- 無持久模型記憶體容量
- 代管推論基礎架構

工具使用和代理安全性

- 授權範圍工具（強制執行使用者權限邊界）
- 嚴格的工具執行逾時
- 用於擷取資料的臨時性、工作階段範圍沙箱
- 使用允許清單強制執行唯讀查詢
- 工具呼叫稽核及敏感參數清理

NetApp Workload Factory 的 Ask Me 存取控制

在 NetApp Workload Factory 中，Ask Me 存取安全性的核心在於經過驗證的工作階段、使用者層級的責任歸屬，以及在執行階段作業期間受保護的機密處理。驗證控制會驗證每個工作階段，並將互動與特定使用者建立關聯。敏感憑證會在執行階段從代管的機密保存庫中擷取，並從記錄中刪除。該服務也套用基礎架構控制，包括非根目錄 Container 執行和受限的 CORS 允許清單。

驗證

- 使用加密驗證進行簽署 JWT 驗證（包括對象、發行者和演算法檢查，例如 RS256）
- 服務邊界處的中介軟體強制驗證
- 使用者身分範圍限定，以便將互動歸因於特定使用者

憑證保護和機密處理

- 安全保存庫儲存設備：服務使用的敏感憑證儲存在代管的機密保存庫中，並在執行階段擷取。應用程式程式碼、組態檔案或 Container 映像檔中均不儲存任何機密。
- 日誌清理：在寫入日誌之前，會將敏感值（認證、權杖、密碼、存取金鑰、憑證）從日誌中遮蔽。

基礎架構安全性

- 非根目錄 Container 執行
- CORS 僅限於受信任 NetApp 網域的明確允許清單

NetApp Workload Factory 的 Ask Me 執行模式

NetApp Workload Factory 為 Ask Me 提供多種執行模式，每種模式都有不同的安全性和隱私特性。當 Ask Me 使用工具整合時（例如查詢客戶本身的 Workload Factory 資源），工具的執行會透過分層防護來控制。啟用工具的執行會在經過驗證的使用者權限範圍內運作，並對每次操作的持續時間和範圍套用限制。擷取到的資料會保留在短暫的工作階段沙箱中，該沙箱會封鎖外部存取，並在工作階段結束時毀損。獨立的提示與程式碼控制會強制執行唯讀查詢，而稽核記錄會擷取工具活動，並清除敏感值。

工具執行安全措施

- 授權範圍工具在已驗證使用者的權限範圍內運作。
- 工具執行逾時限制了長時間運行的操作。
- 臨時資料沙箱將工具檢索的資料儲存在工作階段範圍內，阻止外部存取、檔案 I/O、網路呼叫和引擎層級的擴充載入，並在工作階段結束時毀損。
- 唯讀查詢強制執行透過嚴格的允許清單來驗證產生的查詢。
- 提示強制執行與程式碼強制執行的安全性會獨立套用。
- 工具呼叫稽核會記錄工具名稱、參數、時間戳記和結果狀態，並進行敏感值清理。

輸出控制

- Token 限制執行
- 用於分類/安全性關鍵操作的確定性輸出（溫度 0）
- 具有提前終止與清理的串流

NetApp Workload Factory 的 Ask Me 隱私權與可用度

在 NetApp Workload Factory 中，Ask Me 隱私控制可限制資料暴露，讓工作階段處理保持隔離的，並維護使用者互動的隱私邊界。您的輸入由代管 AI 服務處理，該服務不會使用這些資料來訓練或改進基礎模型。擷取到的操作資料在工作階段期間保留在短暫的記憶體儲存設備中，不會持久化或是共享的。跨多個雲端區域的多模型後援鏈可在主要模型受限或無法使用時支援可用度，同時相同的安全性控制適用於所有後援模型。

資料處理和隱私

- 提示中的使用者資料：由代管 AI 服務處理，該服務不會使用您的輸入來訓練或改進基礎模型。
- 知識庫內容：僅包含精選、已發佈的 NetApp 文件；不包含您的資料。
- 操作資料：僅當使用者查詢自己的資源時才會擷取；在工作階段期間保存在短暫的記憶體容量儲存設備中，不會持續保留或共享的。
- 稽核記錄：已記錄查詢中繼資料（匿名使用者 ID、時間戳記、持續時間），並清理敏感欄位。
- 回饋資料：單獨儲存並與查詢 ID 關聯，不與匿名使用者 ID 以外的個人識別資訊關聯。

可用性和隔離控制

Ask Me 採用跨多個雲端區域的多模型回退鏈。

如果主模型受到限制或無法使用，系統可以容錯移轉至替代模型。

所有模型均受到相同的安全性控制、系統提示強化與隔離保證所規範。