NetApp Data Classification文檔

NetApp Data Classification

NetApp
February 06, 2026

目錄

NetApp Data Classification文檔

發行說明

NetApp Data Classification分類的新功能

了解NetApp Data Classification的新功能。

2026年1月14日

版本 **1.50**

本次資料分類版本更新包含錯誤修復和以下更新：

自訂分類改進

資料分類功能現在支援建立自訂資料類別。您可以上傳檔案來微調資料分類所使用的 AI 模型，以便將類別標記套用至資料。所有自訂分類的介面都得到了改進。

有關詳細信息，請參閱 "建立自訂分類"。

自訂過期資料定義

資料分類功能現在可讓您自訂過期資料的定義，以滿足您組織的需求。此前，過時數據被定義為三年前最後一次修改的數據。現在，可以根據上次訪問時間或上次修改時間來識別過時的數據；時間段可以從 6 個月前到 10 年前不等。

有關詳細信息，請參閱 "自訂過期資料定義"。

性能提升

資料分類、資料對應報告和調查頁面上的篩選器等所有頁面的載入時間都已縮短。

調查報告預計完成時間

下載調查報告時，「資料分類」現在會顯示下載完成的預計時間。

2025年12月8日

版本 **1.49**

本次資料分類版本更新包含錯誤修復和以下更新：

在健康監測儀表板中監控指標和效能

資料分類現在提供了一個健康監控儀表板，可以即時監控您的資源，並提供有關記憶體使用情況、磁碟使用情況、磁碟使用率等方面的見解。借助健康監控儀表板提供的信息，您可以查看部署的基礎架構，並獲得優化儲存和效能的見解。

有關詳細信息，請參閱 "監控資料分類的運作狀況"。

提升了裝載性能

資料分類中所有頁面的載入效能均已提升，從而創造了更有效率的使用者體驗。

2025年11月10日

版本 **1.48**

本次資料分類版本更新包括錯誤修復、安全性改進和效能提升。

增強掃描進度清晰度

掃描配置現在包含對掃描完成情況的更深入洞察。以前，進度條只在掃描進行時顯示。現在，掃描完成後進度條仍然可見，以確認掃描已成功完成。您也可以查看已對應和已掃描的檔案數量。

有關掃描設定的更多信息，請參閱 "更改儲存庫的NetApp Data Classification掃描設置"。

2025年10月6日

版本 **1.47**

BlueXP classification現為NetApp Data Classification

BlueXP classification已重新命名為NetApp Data Classification。除了重命名之外，使用者介面也得到了增強。

BlueXP現在是NetApp Console

BlueXP已重新命名並重新設計，以更好地反映其在管理資料基礎架構中的作用。

NetApp Console提供企業級跨本地和雲端環境的儲存和資料服務的集中管理，提供即時洞察、更快的工作流程和簡化的管理。

有關更改的詳細信息，請參閱 "NetApp Console發行說明"。

增強調查體驗

使用新的可搜尋篩選器、每個值的結果計數、總結關鍵發現的即時見解以及具有可自訂列和滑出詳細資訊窗格的刷新結果表，更快地查找和理解您的資料。

有關更多信息，請參閱"調查數據" 。

新的治理與合規儀表板

透過直覺的小部件、更清晰的視覺效果和改進的加載性能更快地獲得關鍵見解。有關詳細信息，請參閱"審查有關您的數據的治理信息"和"查看有關您的數據的合規性信息"。

已儲存查詢的策略（預覽）

資料分類現在使您能夠透過條件操作實現治理自動化。您可以建立保留規則，設定自動刪除和定期電子郵件通知，所有這些都可以透過更新的已儲存查詢頁面進行管理。

有關更多信息，請參閱"創建策略" 。

操作（預覽）

從調查頁面直接控制 - 單獨或批次刪除、移動、複製或標記文件，以實現高效的資料管理和補救。

有關更多信息，請參閱"調查數據" 。

支援Google Cloud NetApp Volumes

資料分類現在支援在Google Cloud NetApp Volumes上進行掃描。從NetApp Console輕鬆新增Google Cloud

NetApp Volumes，實現無縫資料掃描與分類。有關詳細信息，請參閱 "掃描Google Cloud NetApp Volumes"。

2025年8月11日

版本 **1.46**

此資料分類版本包括錯誤修復和以下更新：

審計頁面中增強的掃描事件洞察

審計頁面現在支援對BlueXP classification的掃描事件的增強洞察。審計頁面現在顯示系統掃描的開始時間、系統狀態以及任何問題。共享和系統的狀態僅適用於映射掃描。

有關審計頁面的更多信息，請參閱"監控NetApp Console操作"。

支援 **RHEL 9.6**

此版本增加了對 Red Hat Enterprise Linux v9.6 的支持，用於手動本地安裝BlueXP classification，包括暗站部署。

以下作業系統需要使用 Podman 容器引擎，並且需要BlueXP classification版本 1.30 或更高版本：Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4 和 9.5。

2025年7月14日

版本 **1.45**

此BlueXP classification版本包括優化資源利用率的程式碼變更以及：

改進了新增文件共用進行掃描的工作流程

將文件共享新增至文件共享組的工作流程已經簡化。該流程現在還根據身份驗證類型（Kerberos 或 NTLM）區分 CIFS 協定支援。

有關更多信息，請參閱"掃描檔案共享" 。

增強文件所有者信息

現在您可以在「調查」標籤中查看有關捕獲的文件的文件所有者的更多資訊。在「調查」標籤中查看文件的元資料時，找到文件擁有者，然後選擇「查看詳細資料」以查看使用者名稱、電子郵件和 SAM 帳戶名稱。您也可以查看該用戶擁有的其他物品。此功能僅適用於具有 Active Directory 的工作環境。

有關更多信息，請參閱"調查組織中儲存的數據" 。

2025年6月10日

版本 **1.44**

此次BlueXP classification發布包括：

改進了治理儀表板的更新時間

治理儀表板各個組件的更新時間已改善。下表顯示了每個組件的更新頻率。

成分	更新時間
數據時代	24小時
類別	24小時
數據概覽	5分鐘
重複文件	2小時
文件類型	24小時
非業務數據	2小時
開放權限	24小時
已儲存的搜尋	2小時
敏感資料和廣泛權限	24小時
數據大小	24小時
陳舊數據	2小時
按敏感度等級劃分的頂層資料儲存庫	2小時

您可以查看上次更新的時間，並按敏感度等級手動更新重複文件、非業務資料、已儲存的搜尋、陳舊資料和頂級資料儲存庫元件。有關治理儀表板的更多信息，請參閱"查看有關組織中存儲的數據的治理詳細信息"。

性能和安全性改進

已做出改進以提高BlueXP分類的效能、記憶體消耗和安全性。

錯誤修復

Redis 已升級，以提高BlueXP classification的可靠性。BlueXP classification現在使用 Elasticsearch 來提高掃描期間文件計數報告的準確性。

2025年5月12日

版本 **1.43**

本次BlueXP分類版本包含：

優先進行分類掃描

資料分類除了支援僅映射掃描之外，還支援對映射和分類掃描進行優先排序的功能，可讓您選擇先完成哪些掃描。在掃描開始期間和開始之前，支援對地圖和分類掃描進行優先排序。如果您選擇在掃描過程中確定掃描優先級，則映射掃描和分類掃描都會優先處理。

有關更多信息，請參閱"優先掃描" 。

支援加拿大個人識別資訊 **(PII)** 資料類別

資料分類掃描識別加拿大 PII 資料類別。這些類別包括加拿大所有省份和地區的銀行資訊、護照號碼、社會保險號碼、駕駛執照號碼和健康卡號碼。

有關更多信息，請參閱"個人資料類別" 。

自訂分類（預覽）

資料分類支援地圖和分類掃描的自訂分類。透過自訂分類，您可以自訂資料分類掃描，以使用正規表示式擷取特定於您的組織的資料。此功能目前處於預覽狀態。

有關更多信息，請參閱"新增自訂分類" 。

已儲存的搜尋標籤

政策 選項卡已重新命名"已儲存的搜尋"。功能沒有改變。

將掃描事件傳送至審核頁面

資料分類支援發送分類事件（掃描啟動時和掃描結束時）到"NetApp Console 稽核頁面"。

安全性更新

- Keras 套件已更新，緩解了漏洞（BDSA-2025-0107 和 BDSA-2025-1984）。
- Docker 容器配置已更新。容器不再有權利存取主機的網路介面來製作原始網路封包。透過減少不必要的訪問，此更新可減輕潛在的安全風險。

效能增強

已經實施了程式碼增強，以減少 RAM 使用率並提高資料分類的整體效能。

錯誤修復

導致StorageGRID掃描失敗、調查頁面過濾選項無法載入以及無法下載大容量評估的資料發現評估的錯誤已修復。

2025年4月14日

版本 **1.42**

此次BlueXP classification發布包括：

工作環境批次掃描

BlueXP classification支援工作環境的批次操作。您可以選擇啟用對應掃描、啟用對應和分類掃描、停用掃描或在工作環境中跨磁碟區建立自訂設定。如果您對單一磁碟區進行選擇，它將覆寫批次選擇。若要執行批次操作，請導覽至配置頁面並進行選擇。

本地下載調查報告

BlueXP classification支援將資料調查報告下載到本地以便在瀏覽器中查看。如果選擇本機選項，資料調查僅以 CSV 格式提供，並且僅顯示前 10,000 行資料。

有關更多信息，請參閱"使用BlueXP classification調查組織中儲存的數據" 。

2025年3月10日

版本 **1.41**

此BlueXP classification版本包括一般改進和錯誤修復。它還包括：

掃描狀態

BlueXP classification追蹤磁碟區上的初始映射和分類掃描的即時進度。單獨的進度條追蹤映射和分類掃描，顯示掃描文件總數的百分比。您也可以將滑鼠停留在進度條上以查看已掃描的檔案數和檔案總數。追蹤掃描狀態可

以更深入地了解掃描進度，使您能夠更好地規劃掃描並了解資源分配。

若要查看掃描狀態，請導覽至BlueXP classification中的配置，然後選擇工作環境配置。每卷的進度均以行顯示。

2025年2月19日

版本 **1.40**

此BlueXP classification版本包括以下更新。

支援 **RHEL 9.5**

此版本除了支援先前支援的版本外，還提供對 Red Hat Enterprise Linux v9.5 的支援。這適用於任何手動本機安裝的BlueXP classification，包括暗站部署。

以下作業系統需要使用 Podman 容器引擎，並且需要BlueXP classification版本 1.30 或更高版本：Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4 和 9.5。

優先進行僅映射掃描

當進行僅映射掃描時，您可以優先考慮最重要的掃描。當您擁有多個工作環境並希望確保首先完成高優先掃描時，此功能會有所幫助。

預設情況下，掃描會按照啟動的順序排隊。透過設定掃描優先權，您可以將掃描移至佇列的最前面。可以對多個掃描進行優先排序。優先權會依照先進先出的順序指定，這表示您優先考慮的第一個掃描將移至佇列的最前面；您優先考慮的第二個掃描將成為佇列中的第二個掃描，依此類推。

優先權是一次性授予的。映射資料的自動重新掃描按照預設順序進行。

優先權僅限於"僅映射掃描"；它不適用於地圖和分類掃描。

有關更多信息，請參閱"優先掃描" 。

重試所有掃描

BlueXP classification支援批次重試所有失敗掃描的功能。

您可以使用全部重試功能以批次操作的方式重新嘗試掃描。如果分類掃描因網路中斷等臨時問題而失敗，您可以使用一個按鈕同時重試所有掃描，而不必單獨重試。可根據需要重試掃描多次。

若要重試所有掃描：

1. 從BlueXP classification選單中，選擇 配置。
2. 若要重試所有失敗的掃描，請選擇*重試所有掃描*。

提高分類模型的準確性

機器學習模型的準確率"預定義類別"提高了11%。

2025年1月22日

版本 **1.39**

此BlueXP classification版本更新了資料調查報告的匯出流程。此匯出更新對於對您的資料執行額外分析、對資料建立額外視覺化或與他人分享資料調查結果很有用。

以前，數據調查報告匯出限制為 10,000 行。在此版本中，限制已被取消，以便您可以匯出所有資料。此變更使您能夠從數據調查報告中匯出更多數據，從而為您的數據分析提供更大的靈活性。

您可以選擇工作環境、磁碟區、目標資料夾以及 JSON 或 CSV 格式。匯出的檔案名稱包含時間戳，以協助您識別資料的匯出時間。

支援的工作環境包括：

- Cloud Volumes ONTAP
- 適用於ONTAP的 FSx
- ONTAP
- 共享群組

從數據調查報告匯出數據有以下限制：

- 每種類型（檔案、目錄和表格）最多可下載 5 億筆記錄
- 預計匯出一百萬筆記錄大約需要 35 分鐘。

有關數據調查和報告的詳細信息，請參閱 "調查組織中儲存的數據"。

2024年12月16日

版本 **1.38**

此BlueXP classification版本包括一般改進和錯誤修復。

2024年11月4日

版本 **1.37**

此BlueXP classification版本包括以下更新。

支援 **RHEL 8.10**

此版本除了支援先前支援的版本外，還提供對 Red Hat Enterprise Linux v8.10 的支援。這適用於任何手動本機安裝的BlueXP classification，包括暗站部署。

以下作業系統需要使用 Podman 容器引擎，並且需要BlueXP classification版本 1.30 或更高版本：Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3 和 9.4。

詳細了解 "BlueXP classification"。

支持 **NFS v4.1**

此版本除了支援先前支援的版本外，還提供對 NFS v4.1 的支援。

詳細了解 "BlueXP classification"。

2024年10月10日

版本 **1.36**

支援 **RHEL 9.4**

此版本除了支援先前支援的版本外，還提供對 Red Hat Enterprise Linux v9.4 的支援。這適用於任何手動本機安裝的BlueXP classification，包括暗站部署。

以下作業系統需要使用 Podman 容器引擎，並且需要BlueXP classification版本 1.30 或更高版本：Red Hat Enterprise Linux 版本 8.8、9.0、9.1、9.2、9.3 和 9.4。

詳細了解 "BlueXP classification部署概述"。

改進的掃描性能

此版本提供了改進的掃描效能。

2024年9月2日

版本 **1.35**

掃描**StorageGRID**數據

BlueXP classification支援掃描StorageGRID中的資料。

有關詳細信息，請參閱"掃描StorageGRID數據"。

2024年8月5日

版本 **1.34**

此BlueXP classification版本包括以下更新。

從 **CentOS** 改為 **Ubuntu**

BlueXP classification已將其針對 Microsoft Azure 和 Google Cloud Platform (GCP) 的 Linux 作業系統從 CentOS 7.9 更新為 Ubuntu 22.04。

有關部署詳細信息，請參閱 "在具有網際網路存取權限的Linux主機上安裝並準備Linux主機系統"。

2024年7月1日

版本 **1.33**

支援 **Ubuntu**

此版本支援 Ubuntu 24.04 Linux 平台。

地圖掃描收集元數據

在映射掃描期間從文件中提取以下元數據，並將其顯示在治理、合規性和調查儀表板上：

- 工作環境
- 工作環境類型
- 儲存庫
- 文件類型
- 已用容量
- 文件數量
- 文件大小
- 文件創建
- 文件上次訪問
- 文件上次修改時間
- 文件發現時間
- 權限擷取

儀表板中的附加數據

此版本更新了映射掃描期間治理、合規和調查儀表板中顯示的資料。

有關詳細信息，請參閱"映射和分類掃描之間有什麼區別" 。

2024年6月5日

版本 **1.32**

配置頁面中的新映射狀態列

此版本現在在設定頁面中顯示一個新的對應狀態列。新列可協助您識別映射是否正在運行、排隊、暫停或更多。

有關狀態的解釋，請參閱 "更改掃描設定"。

2024年5月15日

版本 **1.31**

分類是**BlueXP**中的一項核心服務

BlueXP classification現在作為BlueXP中的一項核心功能提供，每個連接器最多可免費掃描 500 TiB 的資料。無需分類許可或付費訂閱。由於我們將BlueXP classification功能的重點放在新版本掃描NetApp儲存系統上，因此某些舊功能將僅對先前已支付授權費用的客戶可用。當付費合約到期時，這些舊功能的使用將失效。

資料分類不會對其可以掃描的資料量施加限制。每個控制台代理程式支援掃描和顯示 500 TiB 的資料。要掃描超過 500 TiB 的數據，"安裝另一個控制台代理"然後"部署另一個資料分類實例"。 + 控制台 UI 顯示來自單一連接器的資料。有關查看來自多個控制台代理的資料的提示，請參閱"使用多個控制台代理"。

2024年4月1日

版本 **1.30**

增加了對 **RHEL v8.8** 和 **v9.3 BlueXP classification**的支持

此版本除了先前支援的 9.x 之外，還支援 Red Hat Enterprise Linux v8.8 和 v9.3，它需要 Podman，而不是 Docker 引擎。這適用於BlueXP classification的任何手動本機安裝。

以下作業系統需要使用 Podman 容器引擎，並且需要BlueXP classification版本 1.30 或更高版本：Red Hat Enterprise Linux 版本 8.8、9.0、9.1、9.2 和 9.3。

詳細了解 "BlueXP classification部署概述"。

如果您在本機的 RHEL 8 或 9 主機上安裝連接器，則支援BlueXP classification。如果 RHEL 8 或 9 主機位於 AWS、Azure 或 Google Cloud 中，則不受支援。

刪除了啟動審計日誌收集的選項

啟動審計日誌收集的選項已停用。

掃描速度提高

輔助掃描節點的掃描性能得到了改善。如果您需要額外的掃描處理能力，您可以新增更多掃描器節點。有關詳細信息，請參閱 "在可以存取網際網路的主機上安裝BlueXP classification"。

自動升級

如果您在具有網路存取權限的系統上部署了BlueXP classification，則系統會自動升級。以前，升級發生在自上次用戶活動以來經過特定時間之後。在此版本中，如果當地時間在凌晨 1:00 至凌晨 5:00 之間， BlueXP classification將自動升級。如果當地時間不在這些時間範圍內，則升級將在使用者上次活動後經過特定時間後進行。有關詳細信息，請參閱 "在可以存取網際網路的 Linux 主機上安裝"。

如果您在沒有網路存取的情況下部署了BlueXP classification，則需要手動升級。有關詳細信息，請參閱 "在沒有網路存取的 Linux 主機上安裝BlueXP classification"。

2024年3月4日

版本 **1.29**

現在您可以排除駐留在特定資料來源目錄中的掃描數據

如果您希望BlueXP classification排除駐留在特定資料來源目錄中的掃描數據，則可以將這些目錄名稱新增至BlueXP classification的設定檔。此功能可讓您避免掃描不必要的目錄，或避免傳回錯誤的個人資料結果。

"了解更多" 。

超大型實例支援現已合格

如果您需要BlueXP classification來掃描超過 2.5 億個文件，您可以在雲端部署或本地安裝中使用超大實例。這種系統最多可以掃描 5 億個檔案。

"了解更多" 。

2024年1月10日

版本 **1.27**

調查頁面結果顯示總大小以及項目總數

調查頁面中的過濾結果除了顯示文件總數外，還顯示項目的總大小。這在移動檔案、刪除檔案等操作時很有幫助。

將其他群組 **ID** 配置為“向組織開放”

現在，如果群組最初沒有設定該權限，您可以直接從BlueXP classification將 NFS 中的群組 ID 配置為「向組織開放」。任何附加了這些群組 ID 的文件和資料夾都將在調查詳情頁面中顯示為「向組織開放」。了解如何"添加其他群組 ID 作為“對組織開放”"。

2023年12月14日

版本 **1.26.6**

此版本包含一些小的改進。

該版本還刪除了以下選項：

- 啟動審計日誌收集的選項已停用。
- 在目錄調查期間，無法使用目錄計算個人識別資訊 (PII) 資料數量的選項。請參閱"調查組織中儲存的數據"。
- 使用 Azure 資訊保護 (AIP) 標籤整合資料的選項已停用。

2023年11月6日

版本 **1.26.3**

此版本已修復以下問題

- 修正了儀表板中顯示系統掃描的檔案數量不一致的問題。
- 透過處理和報告名稱和元資料中帶有特殊字元的檔案和目錄來改善掃描行為。

2023年10月4日

版本 **1.26**

支援在 **RHEL** 版本 **9** 上本機安裝**BlueXP classification**

Red Hat Enterprise Linux 8 和 9 版本不支援 Docker 引擎；而BlueXP classification安裝則需要引擎。我們現在支援在 RHEL 9.0、9.1 和 9.2 上使用 Podman 版本 4 或更高版本作為容器基礎架構進行BlueXP classification安裝。如果您的環境需要使用最新版本的 RHEL，現在您可以在使用 Podman 時安裝BlueXP classification（版本 1.26 或更高版本）。

目前，在使用 RHEL 9.x 時，我們不支援暗站安裝或分散式掃描環境（使用主節點和遠端掃描器節點）。

2023年9月5日

版本 **1.25**

中小型部署暫時無法使用

當您在 AWS 中部署BlueXP classification實例時，此時無法選擇 部署 **>** 配置 並選擇小型或中型實例。您仍然可以透過選擇*部署>部署*來使用大實例大小部署實例。

在調查結果頁面中為最多 **100,000** 個項目新增標籤

過去，您一次只能在調查結果頁面中將標籤套用至單一頁面（20 個項目）。現在您可以在調查結果頁面中選擇*所有*項目並將標籤應用於所有項目 - 一次最多 100,000 個項目。

識別最小檔案大小為 **1 MB** 的重複文件

BlueXP classification僅用於在檔案大小為 50 MB 或更大時識別重複檔案。現在可以識別以 1 MB 開頭的重複檔案。您可以使用調查頁面過濾器「檔案大小」和「重複」來查看您的環境中哪些特定大小的檔案是重複的。

2023年7月17日

版本 **1.24**

BlueXP classification識別出兩種新的德國個人數據

BlueXP classification可以識別和分類包含以下類型資料的檔案：

- 德國身分證 (Personalausweisnummer)
- 德國社會安全號 (Sozialversicherungsnummer)

"查看BlueXP classification可以在您的資料中識別的所有類型的個人數據" 。

BlueXP classification在限制模式和私人模式下完全受支持

BlueXP classification現在完全支援沒有網路存取（私人模式）和有限的外部網路存取（受限模式）的網站。"了解有關連接器的BlueXP部署模式的更多信息" 。

升級**BlueXP classification**的私人模式安裝時可以跳過版本

現在，即使 BlueXP 分類不是連續的，您也可以升級到較新版本的BlueXP classification。這意味著不再需要目前一次升級BlueXP classification的一個版本的限制。此功能從 1.24 版本開始適用。

BlueXP classificationAPI 現已可用

BlueXP classificationAPI 可讓您執行操作、建立查詢以及匯出有關您正在掃描的資料的資訊。互動式文件可透過 Swagger 取得。該文件分為多個類別，包括調查、合規、治理和配置。每個類別都是對BlueXP classificationUI 中的選項卡的引用。

"了解有關BlueXP classificationAPI 的更多信息" 。

2023年6月6日

版本 **1.23**

搜尋資料主體名稱時現在支援日語

現在，在回應資料主體存取請求 (DSAR) 時搜尋主體名稱時可以輸入日文名稱。您可以生成"資料主體存取請求報告"以及由此產生的資訊。您還可以在"資料調查頁面中的「資料主體」過濾器"識別包含主題名稱的文件。

Ubuntu 現在是受支援的 **Linux** 發行版，您可以在其上安裝**BlueXP classification**

Ubuntu 22.04 已被認定為BlueXP classification的支援作業系統。您可以在網路中的 Ubuntu Linux 主機上安裝BlueXP classification，或使用安裝程式 1.23 版本在雲端中的 Linux 主機上安裝。 "看看如何在安裝了 Ubuntu 的主機上安裝BlueXP classification" 。

新的**BlueXP classification**安裝不再支援 **Red Hat Enterprise Linux 8.6** 和 **8.7**

這些版本不支援新的部署，因為 Red Hat 不再支援 Docker，而 Docker 是先決條件。如果您有在 RHEL 8.6 或 8.7 上執行的現有BlueXP classification機器， NetApp將繼續支援您的設定。

BlueXP classification可以配置為 **FPolicy** 收集器，以從**ONTAP**系統接收 **FPolicy** 事件

您可以啟用檔案存取稽核日誌功能，在BlueXP classification系統上收集在工作環境中的磁碟區上偵測到的檔案存取事件。 BlueXP classification可以擷取以下類型的 FPolicy 事件以及對您的檔案執行操作的使用者：建立、讀取、寫入、刪除、重新命名、變更擁有者/權限以及變更 SACL/DACL。

暗網現已支援 **Data Sense BYOL** 許可證

現在，您可以將 Data Sense BYOL 授權上傳到暗站中的BlueXP digital wallet中，以便在許可證不足時收到通知。

2023年4月3日

版本 **1.22**

新數據發現評估報告

數據發現評估報告對掃描環境進行了高級分析，以突出顯示系統的發現並顯示關注區域和潛在的補救步驟。本報告的目標是提高人們對資料治理問題、資料安全漏洞以及資料集的資料合規性差距的認識。 "了解如何產生和使用數據發現評估報告" 。

能夠在雲端中的較小實例上部署**BlueXP classification**

在 AWS 環境中從BlueXP連接器部署BlueXP classification時，現在您可以從兩個比預設執行個體更小的執行個體類型中進行選擇。如果您正在掃描小型環境，這可以幫助您節省雲端成本。但是，使用較小的實例時存在一些限制。 "查看可用的實例類型和限制" 。

現在可以使用獨立腳本在**BlueXP classification**安裝之前驗證您的 **Linux** 系統

如果您想獨立於執行BlueXP classification安裝來驗證您的 Linux 系統是否符合所有先決條件，您可以下載一個單獨的腳本，該腳本僅測試先決條件。 "了解如何檢查您的 Linux 主機是否已準備好安裝BlueXP classification" 。

2023年3月7日

版本 **1.21**

從**BlueXP classificationUI** 新增您自己的自訂類別的新功能

BlueXP classification現在可讓您新增自己的自訂類別，以便BlueXP classification能夠識別適合這些類別的檔案。 BlueXP classification有很多 "預定義類別"，因此此功能可讓您新增自訂類別，以識別在資料中找到組織獨有的資訊的位置。

現在您可以從**BlueXP classificationUI** 新增自訂關鍵字

BlueXP classification已經能夠添加自訂關鍵字， BlueXP classification將在未來的掃描中識別這些關鍵字。但是，您需要登入BlueXP classificationLinux 主機並使用命令列介面新增關鍵字。在此版本中，新增自訂關鍵字的功能位於BlueXP classificationUI 中，這使得新增和編輯這些關鍵字變得非常容易。

當「上次訪問時間」發生變化時， **BlueXP classification**不會掃描文件

預設情況下，如果BlueXP classification沒有足夠的「寫入」權限，系統將不會掃描磁碟區中的文件，因為BlueXP classification無法將「上次存取時間」恢復為原始時間戳記。但是，如果您不介意將上次訪問時間重置為文件中的原始時間，則可以在配置頁面中覆蓋此行為，以便BlueXP classification可以掃描卷，而不管權限如何。

與此功能結合，新增了名為「掃描分析事件」的新篩選器，以便您可以查看未分類的文件，因為BlueXP classification無法恢復上次存取時間，或即使BlueXP classification無法恢復上次存取時間也已分類的文件。

"詳細了解「上次造訪時間戳記」以及BlueXP classification所需的權限" 。

BlueXP classification可識別三種新的個人資料類型

BlueXP classification可以識別和分類包含以下類型資料的檔案：

- 波札那身分證（奧芒）號碼
- 波札那護照號碼
- 新加坡國民登記身分證（NRIC）

"查看BlueXP classification可以在您的資料中識別的所有類型的個人數據" 。

更新了目錄的功能

- 資料調查報告的「精簡版 CSV 報告」選項現在包含來自目錄的資訊。
- 「上次造訪」時間過濾器現在顯示檔案和目錄的上次存取時間。

安裝增強功能

- 對於沒有網路存取的網站（暗站）， BlueXP classification安裝程式現在會執行預檢查，以確保您的系統和網路要求符合成功安裝的要求。
- 安裝審計日誌檔案現在已儲存；它們被寫入 `/ops/netapp/install_logs`。

2023年2月5日

版本 **1.20**

能夠向任何電子郵件地址發送基於策略的通知電子郵件

在BlueXP classification的早期版本中，當某些關鍵策略傳回結果時，您可以向您帳戶中的BlueXP使用者發送電子郵件警報。此功能使您能夠在不在線時收到通知以保護您的資料。現在，您也可以從策略向不在您的BlueXP帳戶中的任何其他使用者（最多 20 個電子郵件地址）發送電子郵件警報。

"詳細了解如何根據策略結果發送電子郵件提醒" 。

現在您可以從**BlueXP classificationUI** 新增個人模式

BlueXP classification已經能夠添加自訂“個人資料”， BlueXP classification將在未來的掃描中識別這些資料。但是，您需要登入BlueXP classificationLinux 主機並使用命令列新增自訂模式。在此版本中，使用正規表示式新增個人模式的功能位於BlueXP classificationUI 中，因此可以非常輕鬆地新增和編輯這些自訂模式。

使用**BlueXP classification**可以移動 **1500** 萬個文件

過去，您可以透過BlueXP classification將最多 100,000 個來源檔案移至任何 NFS 共用。現在您一次最多可以移動 1500 萬個檔案。

能夠查看有權存取 **SharePoint Online** 檔案的使用者數量

過濾器「具有存取權限的使用者數量」現在支援儲存在 SharePoint Online 儲存庫中的檔案。過去僅支援 CIFS 共享上的檔案。請注意，此時不基於活動目錄的 SharePoint 群組將不會計入此篩選器。

操作狀態面板中新增了新的「部分成功」狀態

新的「部分成功」狀態表示BlueXP classification作業已完成，有些專案失敗，有些專案成功，例如，當您移動或刪除 100 個檔案時。此外，「完成」狀態已重新命名為「成功」。過去，「完成」狀態可能會列出成功和失敗的操作。現在「成功」狀態意味著所有項目上的所有操作都成功。 "了解如何查看操作狀態面板" 。

2023年1月9日

版本 **1.19**

能夠查看包含敏感資料和過於寬鬆的文件圖表

治理儀表板新增了一個新的「敏感資料和廣泛權限」區域，該區域提供了包含敏感資料（包括敏感資料和敏感個人資料）且過於寬鬆的文件的熱圖。這可以幫助您了解敏感資料可能的風險。 "了解更多" 。

資料調查頁面新增三個過濾器

新的過濾器可用於優化資料調查頁面中顯示的結果：

- 「具有存取權限的使用者數」過濾器顯示哪些檔案和資料夾對一定數量的使用者開放。您可以選擇一個數字範圍來優化結果 - 例如，查看 51-100 個使用者可以存取哪些檔案。
- 現在，「建立時間」、「發現時間」、「上次修改時間」和「上次存取時間」篩選器可讓您建立自訂日期範圍，而不僅僅是選擇預先定義的日期範圍。例如，您可以尋找「建立時間」超過 6 個月的文件，或「上次修改時間」在「最近 10 天」內的文件。
- 現在，「檔案路徑」篩選器可讓您指定要從篩選查詢結果中排除的路徑。如果您輸入包含和排除某些資料的路徑， BlueXP classification會先在包含的路徑中找到所有文件，然後從排除的路徑中刪除文件，然後顯示結果。

"查看可用於調查資料的所有過濾器的列表" 。

BlueXP classification可以辨識日本個人編號

BlueXP classification可以識別和分類包含日本個人編號（也稱為 My Number）的檔案。這包括個人和企業我的號碼。 "查看BlueXP classification可以在您的資料中識別的所有類型的個人數據" 。

NetApp Data Classification的已知限制

已知限制標識了此版本中不受支援或無法正確互通的功能。仔細審查這些限制。

NetApp Data Classification停用選項

2023 年 12 月（版本 1.26.6）版本刪除了以下選項：

- 啟動審計日誌收集的選項已停用。
- 在目錄調查期間，無法使用目錄計算個人識別資訊 (PII) 資料數量的選項。
- 使用 Azure 資訊保護 (AIP) 標籤整合資料的選項已停用。

資料分類掃描

資料分類掃描存在以下限制。

資料分類僅掃描卷下的一個共享

如果單一磁碟區下有多個檔案共用，資料分類將掃描具有最高層次的共用。例如，如果您有以下共享：

- /一個
- /A/B
- /C
- /D/E

在此配置中，僅掃描 /A 中的資料。 /C 和 /D 中的資料未被掃描。

解決方法

有一種解決方法可以確保您掃描卷中所有共享的資料。請依照以下步驟操作：

1. 在系統中，新增要掃描的磁碟區。
2. 資料分類完成掃描磁碟區後，請前往「資料調查」頁面並建立篩選器以查看正在掃描的共用：

   透過「系統名稱」和「目錄類型 = 共享」過濾數據，以查看正在掃描哪個共享。

3. 取得磁碟區中存在的共用的完整列表，以便您可以看到哪些共用未被掃描。
4. "將剩餘的共享新增至共享群組" 。

   單獨新增所有共享，例如：

   ```
   /C
   /D
   ```

5. 對系統中具有多個共享的每個磁碟區執行這些步驟。

上次訪問的時間戳

當資料分類對目錄進行掃描時，掃描會影響目錄的上次存取欄位。當您查看上次造訪欄位時，該元資料反映掃描的日期和時間或使用者上次存取目錄的時間。

開始

了解NetApp Data Classification

NetApp Data Classification是NetApp Console的資料治理服務，它可以掃描您的企業內部和雲端資料來源以對應和分類資料並識別私人資訊。這可以幫助降低您的安全和合規風險，降低儲存成本，並協助您的資料遷移專案。

從 1.31 版開始，資料分類作為NetApp Console中的一項核心功能提供。無需額外付費。無需分類許可或訂閱。 + 如果您一直在使用舊版本 1.30 或更早版本，則該版本在您的訂閱到期之前可用。

NetApp Console

可以透過NetApp Console存取資料分類。

NetApp Console提供企業級跨本機和雲端環境的NetApp儲存和資料服務的集中管理。需要控制台才能存取和使用NetApp資料服務。作為管理介面，它使您能夠從一個介面管理許多儲存資源。控制台管理員可以控制企業內所有系統的儲存和服務的存取。

您不需要許可證或訂閱即可開始使用NetApp Console，並且只有當您需要在雲端部署控制台代理程式以確保與儲存系統或NetApp資料服務的連線時才需要付費。但是，一些可從控制台存取的NetApp資料服務是需要授權或基於訂閱的。

詳細了解"NetApp Console"。

特徵

資料分類使用人工智慧 (AI)、自然語言處理 (NLP) 和機器學習 (ML) 來理解其掃描的內容，以便提取實體並對內容進行相應的分類。這使得資料分類能夠提供以下功能領域。

"了解資料分類的用例" 。

保持合規

資料分類提供了多種工具，可以幫助您實現合規性。您可以使用資料分類來：

- 識別個人識別資訊 (PII)。
- 根據 GDPR、CCPA、PCI 和 HIPAA 隱私法規的要求識別廣泛的敏感個人資訊。
- 根據姓名或電子郵件地址回應資料主體存取請求 (DSAR)。

加強安全

資料分類可以識別可能被犯罪分子存取的資料。您可以使用資料分類來：

- 識別向整個組織或公眾公開的所有具有開放權限的文件和目錄（共用和資料夾）。
- 識別位於初始專用位置之外的敏感資料。
- 遵守資料保留政策。

- 使用*Policies*自動偵測新的安全性問題，以便安全人員可以立即採取行動。

優化儲存使用情況

資料分類提供可協助您降低儲存總擁有成本 (TCO) 的工具。您可以使用資料分類來：

- 透過識別重複或與業務無關的資料來提高儲存效率。
- 透過識別可以分層到較便宜的物件儲存的非活動資料來節省儲存成本。 "了解有關Cloud Volumes ONTAP系統分層的更多信息" 。 "了解有關本地ONTAP系統分層的更多信息" 。

支援的系統和資料來源

資料分類可以掃描和分析來自以下類型的系統和資料來源的結構化和非結構化資料：

系統

- Amazon FSx for NetApp ONTAP管理
- Azure NetApp Files
- Cloud Volumes ONTAP （部署在 AWS、Azure 或 GCP 中）
- 本地ONTAP集群
- StorageGRID
- Google Cloud NetApp Volumes

資料來源

- NetApp檔案分享
- 資料庫:
  - 亞馬遜關係型資料庫服務 (Amazon RDS)
  - MongoDB
  - MySQL
  - 甲骨文
  - PostgreSQL
  - SAP HANA
  - SQL 伺服器 (MSSQL)

資料分類支援 NFS 版本 3.x、4.0 和 4.1，以及 CIFS 版本 1.x、2.0、2.1 和 3.0。

成本

資料分類可以免費使用。無需分類許可或付費訂閱。

基礎設施成本

- 在雲端安裝資料分類需要部署雲端實例，這會導致部署雲端的雲端供應商收取費用。看為每個雲端提供者部署的執行個體類型 。如果您在本機系統上安裝資料分類，則無需付費。

- 資料分類要求您部署控制台代理程式。在許多情況下，由於您在控制台中使用其他儲存和服務，因此您已經擁有控制台代理程式。控制台代理執行個體會導致其部署所在的雲端提供者收取費用。查看 "為每個雲端提供者部署的執行個體類型"。如果您在本機系統上安裝控制台代理，則無需付費。

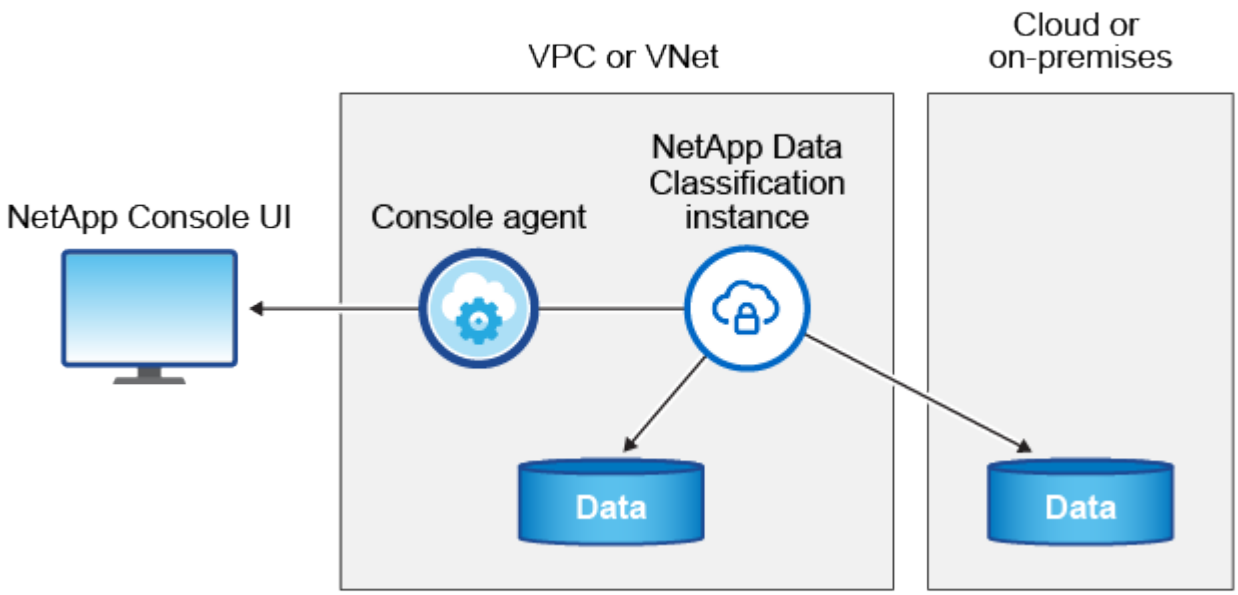
資料傳輸成本

資料傳輸成本取決於您的設定。如果資料分類實例和資料來源位於相同可用區和區域，則沒有資料傳輸成本。但是，如果資料來源（例如Cloud Volumes ONTAP系統）位於不同的可用區或區域，那麼您的雲端供應商將向您收取資料傳輸費用。請參閱以下連結以了解更多詳細資訊：

- "AWS：Amazon Elastic Compute Cloud (Amazon EC2) 定價"
- "Microsoft Azure：頻寬定價詳情"
- "Google Cloud：儲存傳輸服務定價"

資料分類實例

當您在雲端部署資料分類時，控制台會將執行個體部署在與控制台代理相同的子網路中。 "了解有關控制台代理的更多資訊。"



請注意預設實例的以下幾點：

- 在 AWS 中，資料分類在 "m6i.4xlarge 實例"帶有 500 GiB GP2 磁碟。作業系統映像是 Amazon Linux 2。在 AWS 中部署時，如果您要掃描少量數據，則可以選擇較小的執行個體大小。
- 在 Azure 中，資料分類在"Standard_D16s_v3 VM"帶有 500 GiB 磁碟。作業系統映像是 Ubuntu 22.04。
- 在 GCP 中，資料分類在"n2-standard-16 虛擬機"配備 500 GiB 標準持久性磁碟。作業系統映像是 Ubuntu 22.04。
- 在預設實例不可用的區域中，資料分類在備用實例上運行。"查看替代實例類型" 。
- 此實例名為 *CloudCompliance*，並帶有與之連接的產生的雜湊值（UUID）。例如：*CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- 每個控制台代理程式僅部署一個資料分類實例。

您也可以在您的場所內的 Linux 主機上或您首選的雲端提供者的主機上部署資料分類。無論您選擇哪一種安裝方法，軟體的功能都完全相同。只要實例可以存取互聯網，資料分類軟體的升級就會自動進行。

實例應始終保持運行，因為資料分類會持續掃描資料。

在不同的實例類型上部署

查看實例類型的以下規範：

系統大小	規格	限制
特大號	32 個 CPU、128 GB RAM、1 TiB SSD	最多可掃描 5 億個文件。
大（預設）	16 個 CPU、64 GB RAM、500 GiB SSD	最多可掃描 2.5 億個文件。

在 Azure 或 GCP 中部署資料分類時，如果您想使用較小的實例類型，請發送電子郵件至 ng-contact-data-sense@netapp.com 尋求協助。

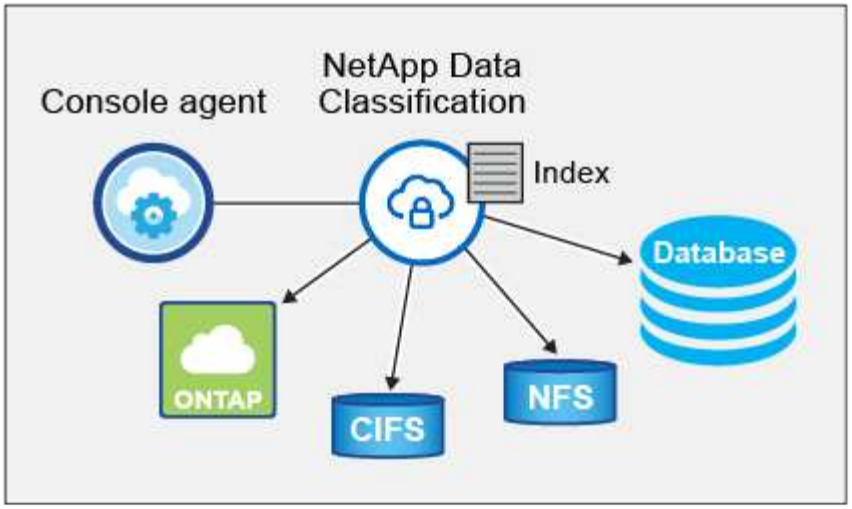
資料分類掃描的工作原理

從高層次來看，資料分類掃描的工作原理如下：

1. 您在控制台中部署資料分類實例。
2. 您可以在一個或多個資料來源上啟用進階映射（稱為「僅映射」掃描）或深層掃描（稱為「映射和分類」掃描）。
3. 資料分類使用人工智慧學習過程掃描資料。
4. 您可以使用提供的儀表板和報告工具來幫助您實現合規性和治理工作。

啟用資料分類並選擇要掃描的儲存庫（這些是磁碟區、資料庫模式或其他使用者資料）後，它會立即開始掃描資料以識別個人和敏感資料。在大多數情況下，您應該專注於掃描即時生產數據，而不是備份、鏡像或 DR 網站。然後，資料分類會對應您的組織數據，對每個檔案進行分類，並識別和提取資料中的實體和預定義模式。掃描結果是個人資訊、敏感個人資訊、資料類別和文件類型的索引。

資料分類透過安裝 NFS 和 CIFS 磁碟區像任何其他客戶端一樣連接到資料。 NFS 磁碟區會自動以唯讀方式訪問，而您需要提供 Active Directory 憑證來掃描 CIFS 磁碟區。



初步掃描後，資料分類將以循環方式持續掃描您的資料以偵測增量變化。這就是為什麼保持實例運行很重要。

您可以在磁碟區層級或資料庫模式層級啟用和停用掃描。

> 資料分類不會對其可以掃描的資料量施加限制。每個控制台代理程式支援掃描和顯示 500 TiB 的資料。要掃描超過 500 TiB 的數據，"安裝另一個控制台代理"然後"部署另一個資料分類實例"。 + 控制台 UI 顯示來自單一連接器的資料。有關查看來自多個控制台代理的資料的提示，請參閱"使用多個控制台代理"。

映射掃描和分類掃描之間有什麼區別

您可以在資料分類中進行兩種類型的掃描：

- 僅映射掃描僅提供資料的進階概覽，並在選定的資料來源上執行。僅映射掃描比映射和分類掃描花費的時間更少，因為它們不存取文件來查看其中的資料。您可能希望首先執行此操作來確定研究領域，然後對這些領域執行地圖和分類掃描。
- 地圖和分類掃描 為您的資料提供深層掃描。

有關映射掃描和分類掃描之間的差異的詳細信息，請參閱"映射和分類掃描之間有什麼區別？"。

資料分類所分類的信息

資料分類收集、索引並分配以下資料的類別：

- 關於文件的*標準元資料*：文件類型、大小、建立和修改日期等等。
- 個人資料：個人識別資訊 (PII)，例如電子郵件地址、身分證號碼或信用卡號，資料分類使用檔案中的特定單字、字串和模式進行識別。"了解有關個人資料的更多信息" 。
- 敏感個人資訊：《一般資料保護規範》（GDPR）和其他隱私法規定義的特殊類型的敏感個人資訊（SPII），例如健康資料、種族血統或政治觀點。"了解有關敏感個人資料的更多信息" 。
- 類別：資料分類將掃描的資料分為不同類型的類別。類別是基於 AI 對每個文件的內容和元資料的分析的主題。"了解有關類別的更多信息"。
- 名稱實體識別：資料分類使用人工智慧從文件中提取人們的自然名稱。"了解如何回應資料主體存取請求" 。

網路概述

資料分類可以在您選擇的任何地方部署單一伺服器或叢集：在雲端或本地端。伺服器透過標準協定連接到資料來源，並在 Elasticsearch 叢集中對結果進行索引，該叢集也部署在同一台伺服器上。這使得能夠支援多雲、跨雲端、私有雲和本地環境。

控制台使用安全性群組部署資料分類實例，該安全性群組啟用來自控制台代理程式的入站 HTTP 連線。

當您在 SaaS 模式下使用控制台時，與控制台的連線透過 HTTPS 提供，並且您的瀏覽器和資料分類實例之間發送的私人資料使用 TLS 1.2 進行端對端加密保護，這表示NetApp和第三方無法讀取它。

出站規則完全開放。需要網路存取來安裝和升級資料分類軟體以及發送使用情況指標。

如果您有嚴格的網路要求，"了解資料分類聯繫的端點" 。

存取NetApp Data Classification

您可以透過NetApp Console存取NetApp Data Classification。

若要登入控制台，您可以使用您的NetApp支援網站憑證，也可以使用您的電子郵件和密碼註冊NetApp Console登入。"了解有關登入控制台的更多信息" 。

特定任務需要特定的控制台使用者角色。"了解所有服務的控制台存取角色" 。

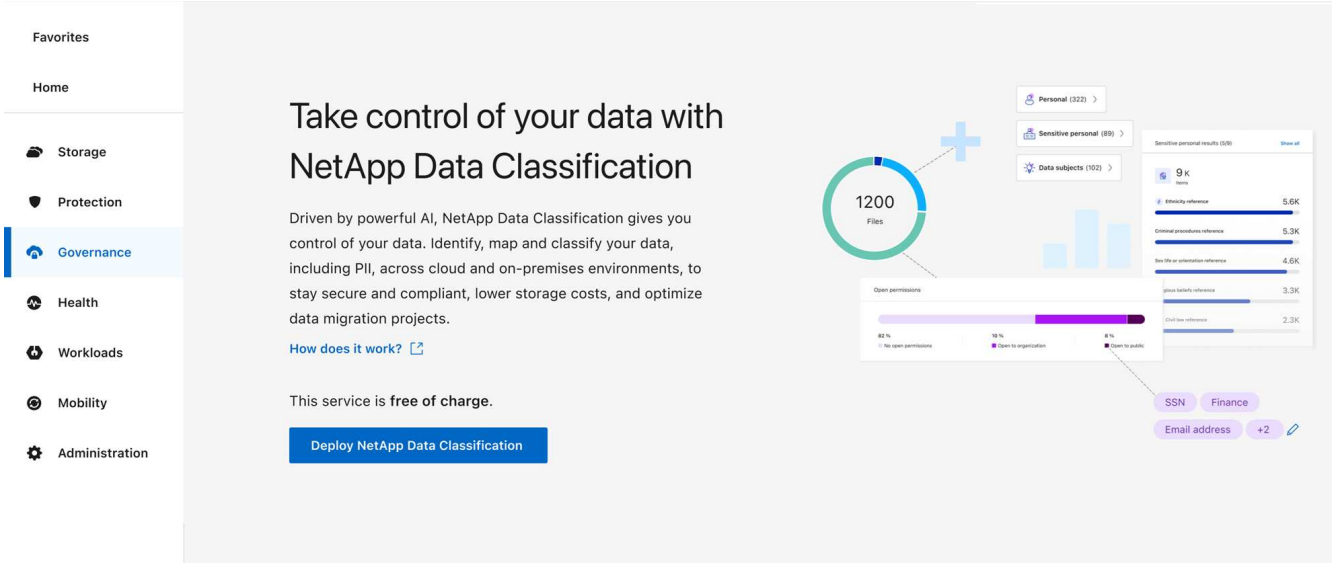
開始之前

- "您應該新增一個控制台代理程式。"
- "了解哪種資料分類部署樣式適合您的工作負載。"

步驟

1. 在 Web 瀏覽器中，導覽至"安慰"。
2. 登入控制台。
3. 從NetApp Console主頁中，選擇「治理」>「資料分類」 。
4. 如果這是您第一次造訪資料分類，則會出現登入頁面。

   選擇*在本機或雲端部署分類*以開始部署您的分類實例。有關詳細信息，請參閱"您應該使用哪種資料分類部署？"



否則，將出現資料分類儀表板。

部署資料分類

您應該使用哪種**NetApp Data Classification**部署？

您可以透過不同的方式部署NetApp Data Classification。了解哪種方法可以滿足您的需求。

資料分類可以透過以下方式部署：

- "使用控制台在雲端部署" 。控制台將資料分類執行個體部署在與控制台代理相同的雲端提供者網路中。
- "在可以存取網際網路的 Linux 主機上安裝" 。在您的網路中的 Linux 主機上或雲端中的 Linux 主機上安裝資料分類，前提是該主機可以存取網際網路。如果您希望使用同樣位於本機的資料分類實例來掃描本機ONTAP系統，則這種類型的安裝可能是一個不錯的選擇，儘管這不是必需的。
- "在沒有網際網路存取的本機站點的 Linux 主機上安裝"，也稱為_私人模式_。這種安裝類型使用安裝腳本，與控制台 SaaS 層沒有連接。

> BlueXP私有模式（傳統BlueXP介面）通常用於沒有網路連線的本機環境和安全雲端區域，其中包括 AWS Secret Cloud、AWS Top Secret Cloud 和 Azure IL6。NetApp繼續透過傳統的BlueXP介面支援這些環境。有關舊版BlueXP介面中的私有模式文檔，請參閱"BlueXP私人模式的 PDF 文檔"。

無論是在有網路存取的 Linux 主機上安裝，或是在沒有網路存取的 Linux 主機上進行本機安裝，都會使用安裝腳本。腳本首先檢查系統和環境是否符合先決條件。如果滿足先決條件，則安裝開始。如果您想獨立於執行資料分類安裝來驗證先決條件，您可以下載一個單獨的軟體包，該軟體包僅測試先決條件。

請參閱"檢查您的 Linux 主機是否已準備好安裝資料分類" 。

使用**NetApp Console**在雲端部署**NetApp Data Classification**

您可以使用NetApp Console在雲端部署NetApp Data Classification。控制台將資料分類執

行個體部署在與控制台代理理相同的雲端提供者網路中。

請注意，您還可以"在可以存取網際網路的 Linux 主機上安裝資料分類"。如果您希望使用同樣位於本機的資料分類實例來掃描本機ONTAP系統，則這種類型的安裝可能是一個不錯的選擇 - 但這不是必需的。無論您選擇哪一種安裝方法，軟體的功能都完全相同。

快速啟動

按照以下步驟快速開始，或向下捲動到其餘部分以獲取完整詳細資訊。


建立控制台代理

如果您還沒有控制台代理，請建立一個。看"在 AWS 中建立控制台代理"，"在 Azure 中建立控制台代理" ， 或者"在 GCP 中建立控制台代理"。

你也可以"在本機安裝控制台代理"在您網路中的 Linux 主機上，或是在雲端的 Linux 主機上。


先決條件

請確保您的環境能夠符合先決條件。這包括該實例的對外網際網路存取、Console agent 與 Data Classification 之間透過 443 埠的連線，以及其他要求。查看完整列表。


部署資料分類

啟動安裝精靈以在雲端部署資料分類實例。

建立控制台代理

如果您還沒有控制台代理，請在您的雲端提供者中建立控制台代理。看 "在 AWS 中建立控制台代理"或者 "在 Azure 中建立控制台代理"， 或者 "在 GCP 中建立控制台代理"。大多數情況下，您可能需要在嘗試啟動資料分類之前設定控制台代理，因為大多數 "控制台功能需要控制台代理"但有些情況下，你需要現在就設定一個。

在某些情況下，您必須使用部署在特定雲端提供者中的控制台代理：

- 在 AWS 中的Cloud Volumes ONTAP或Amazon FSx for ONTAP儲存桶中掃描資料時，您可以使用 AWS 中的控制台代理程式。
- 在 Azure 中的Cloud Volumes ONTAP或Azure NetApp Files中掃描資料時，您可以使用 Azure 中的控制台代理程式。
  - 對於Azure NetApp Files，它必須部署在與您要掃描的磁碟區相同的區域中。
- 在 GCP 中的Cloud Volumes ONTAP中掃描資料時，您可以使用 GCP 中的控制台代理程式。

使用任何這些雲端控制台代理時，都可以掃描本機ONTAP系統、 NetApp檔案共用和資料庫。

請注意，您也可以 "在本機安裝控制台代理"在網路或雲端的 Linux 主機上。一些計劃在本機安裝資料分類的使用者可能也會選擇在本機安裝控制台代理程式。

可能有些情況下你需要使用 "多個控制台代理"。

資料分類不會對其可以掃描的資料量施加限制。每個控制台代理程式支援掃描和顯示 500 TiB 的資料。要掃描超過 500 TiB 的數據，"安裝另一個控制台代理"然後"部署另一個資料分類實例"。 + 控制台 UI 顯示來自單一連接器的資料。有關查看來自多個控制台代理的資料的提示，請參閱"使用多個控制台代理"。

政府區域支持

當控制台代理部署在政府區域（AWS GovCloud、Azure Gov 或 Azure DoD）時，支援資料分類。以這種方式部署時，資料分類有以下限制：

"了解如何在政府區域部署控制台代理"。

先決條件

在雲端部署資料分類之前，請查看以下先決條件，以確保您具有受支援的配置。當您在雲端部署資料分類時，它與控制台代理程式位於同一子網路中。

啟用資料分類的出站互聯網訪問

資料分類需要出站網路存取。如果您的虛擬或實體網路使用代理伺服器進行網際網路訪問，請確保資料分類執行個體具有出站網際網路存取權限以聯絡下列端點。代理必須是非透明的。目前不支援透明代理。

根據您是在 AWS、Azure 還是 GCP 中部署資料分類，請查看下面的對應表格。

AWS 所需的終端節點

端點	目的
\ https://api.console.netapp.com	與控制台服務（包括NetApp帳戶）的通訊。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	與控制台網站通信，實現集中用戶身份驗證。
\ https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com \ https://hub.docker.com \ https://auth.docker.io https://production.cloudflare.docker.com/ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/	提供對軟體映像、清單和範本的存取。
\ https://kinesis.us-east-1.amazonaws.com	使NetApp能夠從稽核記錄中串流資料。
\ https://cognito-idp.us-east-1.amazonaws.com \ https://cognito-identity.us-east-1.amazonaws.com \ https://user-feedback-store-prod.s3.us-west-2.amazonaws.com \ https://customer-data-production.s3.us-west-2.amazonaws.com	啟用資料分類來存取和下載清單和模板，以及發送日誌和指標。

Azure 所需的終點

端點	目的
\ https://api.console.netapp.com	與控制台服務（包括NetApp帳戶）的通訊。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	與控制台網站通信，實現集中用戶身份驗證。
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https https://dseasb33srnrn.cloudfront.net/ // https://production.cloudflare.docker.com/	提供對軟體映像、清單、範本的存取以及發送日誌和指標。
\ https://support.compliance.api.console.netapp.com/	使NetApp能夠從稽核記錄中串流資料。

GCP 所需的端點

端點	目的
\ https://api.console.netapp.com	與控制台服務（包括NetApp帳戶）的通訊。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	與控制台網站通信，實現集中用戶身份驗證。

端點	目的
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https https://dseasb33srnrn.cloudfront.net/ // https://production.cloudflare.docker.com/	提供對軟體映像、清單、範本的存取以及發送日誌和指標。
\ https://support.compliance.api.console.netapp.com/	使NetApp能夠從稽核記錄中串流資料。

確保資料分類具有所需的權限

確保資料分類具有部署資源和為資料分類實例建立安全性群組的權限。

- "Google Cloud 權限"
- "AWS 權限"
- "Azure 權限"

確保控制台代理可以存取資料分類

確保控制台代理程式和資料分類實例之間的連線。控制台代理程式的安全性群組必須允許透過連接埠 443 進出資料分類實例的入站和出站流量。此連線支援部署資料分類實例，並允許您查看「合規性和治理」標籤中的資訊。 AWS 和 Azure 的政府區域支援資料分類。

AWS 和 AWS GovCloud 部署需要額外的入站和出站安全群組規則。看 "AWS 中的控制台代理程式規則"了解詳情。

Azure 和 Azure 政府部署需要額外的入站和出站安全群組規則。看 "Azure 中的控制台代理程式規則"了解詳情。

確保資料分類能夠持續運行

資料分類實例需要保持開啟狀態以持續掃描您的資料。

確保 **Web** 瀏覽器連接到資料分類

啟用資料分類後，請確保使用者從與資料分類實例有連接的主機存取控制台介面。

資料分類實例使用私人 IP 位址來確保索引資料無法被網際網路存取。因此，您用來存取控制台的 Web 瀏覽器必須連線到該私人 IP 位址。此連線可以來自與雲端提供者的直接連線（例如 VPN），也可以來自與資料分類執行個體位於同一網路內的主機。

檢查您的 **vCPU** 限制

確保您的雲端提供者的 vCPU 限制允許部署具有必要數量的核心的執行個體。您需要驗證控制台運作區域中相關執行個體系列的 vCPU 限制。"查看所需的實例類型" 。

有關 vCPU 限制的更多詳細信息，請參閱以下連結：

- "AWS 文件：Amazon EC2 服務配額"
- "Azure 文件：虛擬機器 vCPU 配額"

- "Google Cloud 文件：資源配額"

在雲端部署資料分類

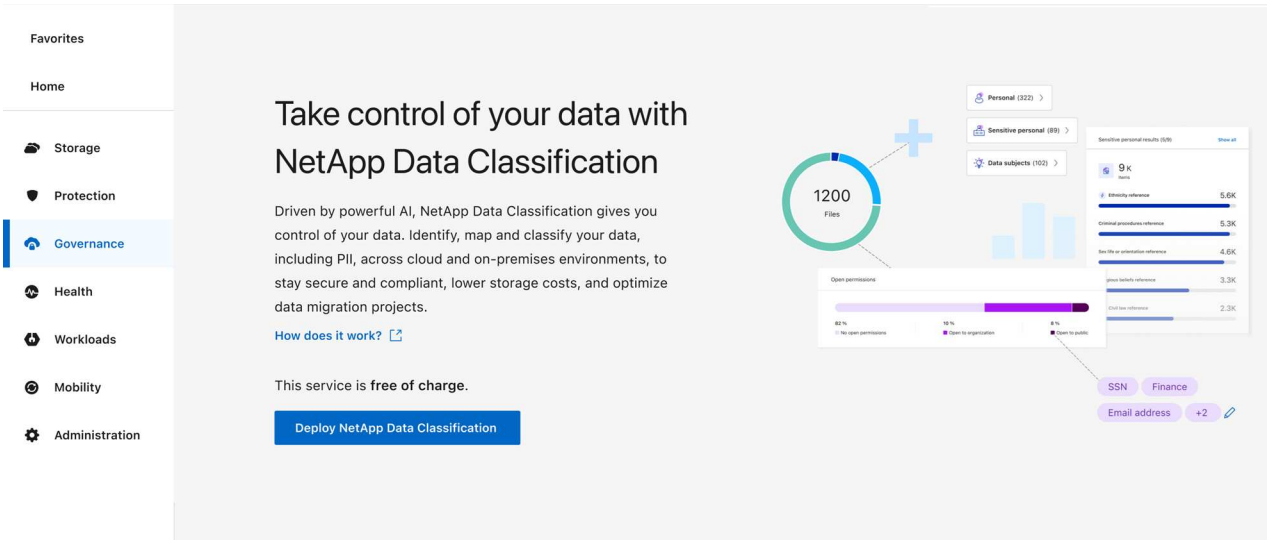
請依照下列步驟在雲端部署資料分類實例。控制台代理程式將在雲端部署實例，然後在該實例上安裝資料分類軟體。

在預設實例類型不可用的區域中，資料分類在"備用實例類型"。

在 **AWS** 中部署

步驟

1. 從資料分類主頁中，選擇*在本地或雲端部署分類*。

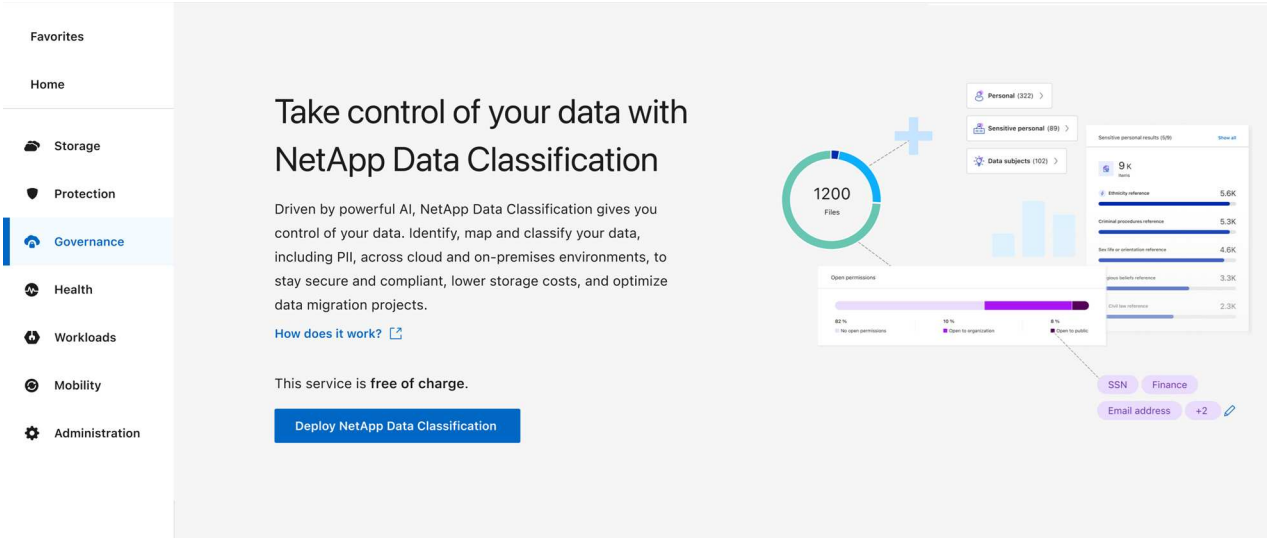


2. 從「安裝」頁面，選擇「部署」>「部署」以使用「大型」實例大小並啟動雲端部署精靈。
3. 精靈在執行部署步驟時會顯示進度。當需要輸入或遇到問題時，系統會提示您。
4. 當實例部署完畢並安裝資料分類後，選擇「繼續設定」進入「設定」頁面。

在 **Azure** 中部署

步驟

1. 從資料分類主頁中，選擇*在本地或雲端部署分類*。

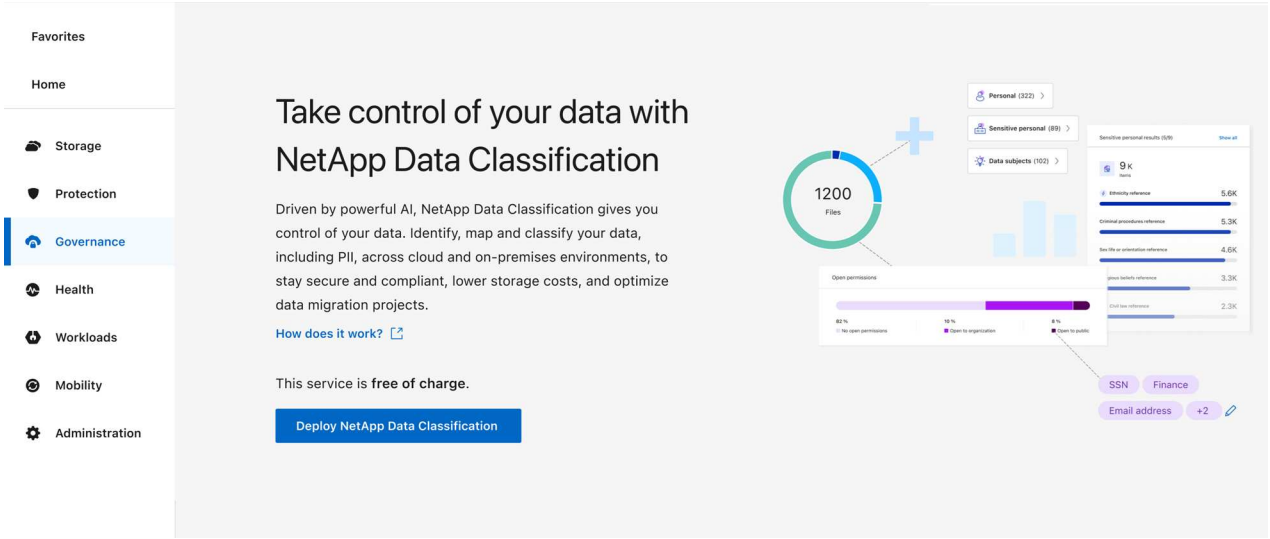


2. 選擇*部署*以啟動雲端部署精靈。
3. 精靈在執行部署步驟時會顯示進度。如果遇到任何問題，它將停止並提示輸入。
4. 當實例部署完畢並安裝資料分類後，選擇「繼續設定」進入「設定」頁面。

在 **Google Cloud** 部署

步驟

1. 從資料分類主頁中，選擇*治理>分類*。
2. 選擇*在本機或雲端部署分類*。



3. 選擇*部署*以啟動雲端部署精靈。
4. 精靈在執行部署步驟時會顯示進度。如果遇到任何問題，它將停止並提示輸入。
5. 當實例部署完畢並安裝資料分類後，選擇「繼續設定」進入「設定」頁面。

結果

控制台在您的雲端提供者中部署資料分類執行個體。

只要實例具有互聯網連接，控制台代理和資料分類軟體的升級就會自動進行。

下一步

您可以從設定頁面選擇要掃描的資料來源。

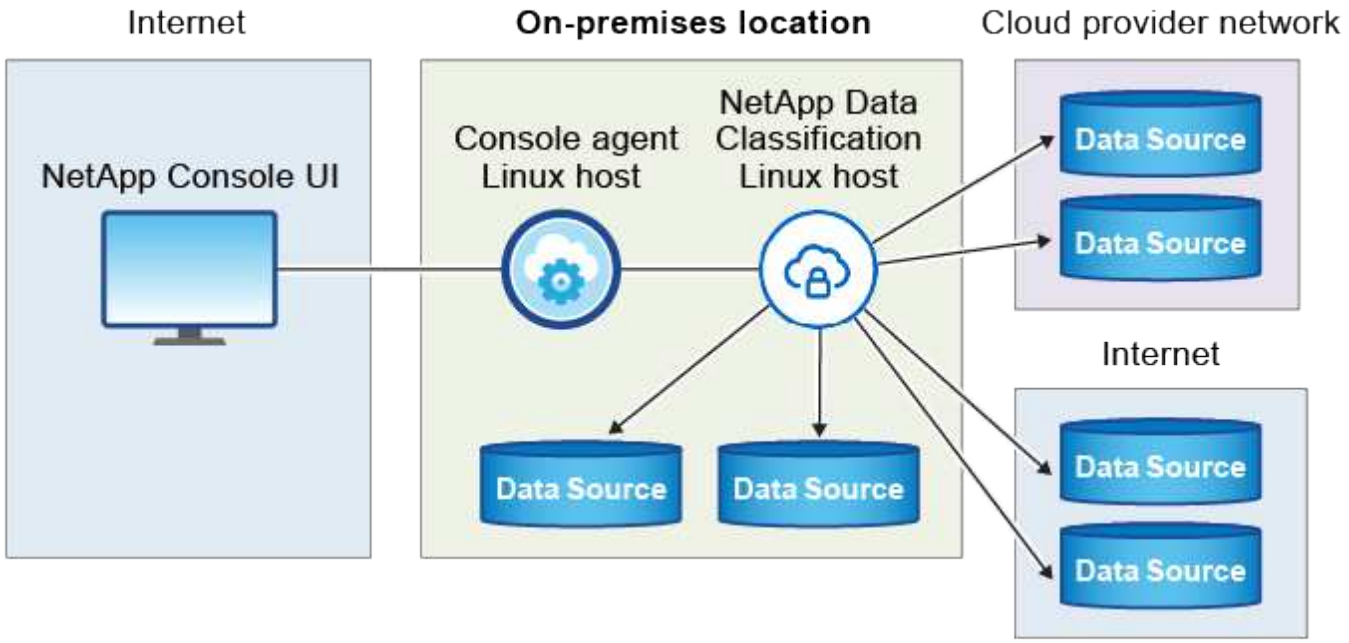
在可存取網際網路的主機上安裝**NetApp Data Classification**

要在您的網路中的 Linux 主機或具有 Internet 存取權限的雲端中的 Linux 主機上部署NetApp Data Classification，您需要在您的網路或雲端中手動部署 Linux 主機。

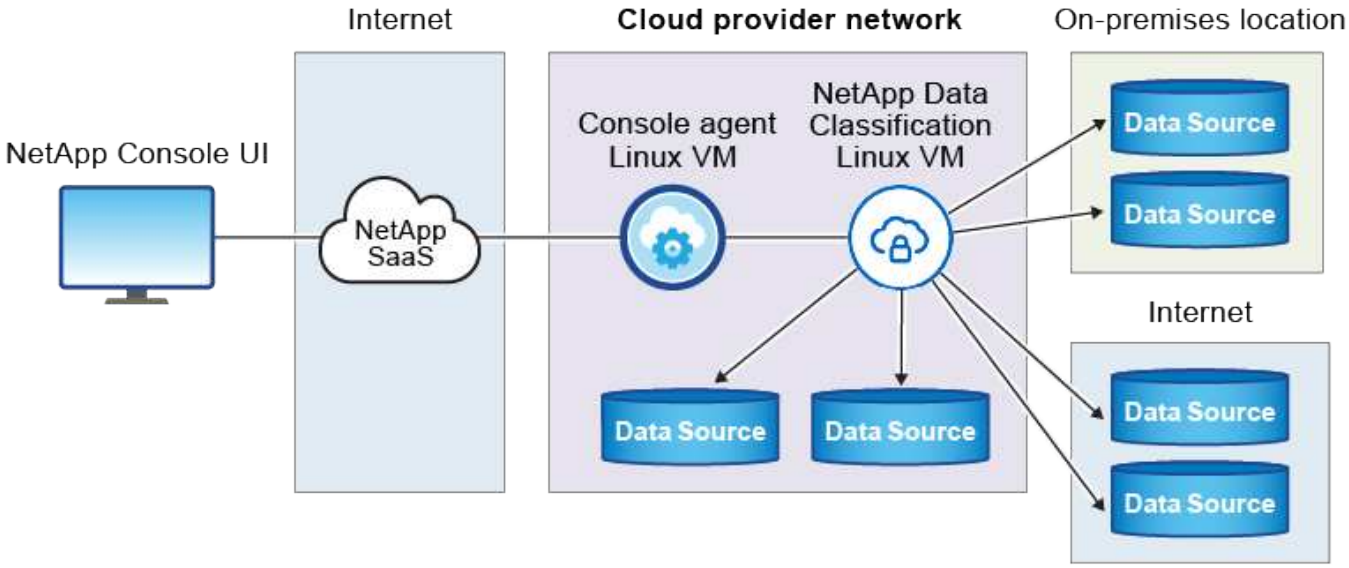
如果您希望使用同樣位於本機的資料分類實例來掃描本機ONTAP系統，則本機安裝是個不錯的選擇。這不是必需的。無論選擇哪一種安裝方法，軟體的功能都是相同的。

資料分類安裝腳本首先檢查系統和環境是否滿足所需的先決條件。如果所有先決條件都滿足，則安裝開始。如果您想獨立於執行資料分類安裝來驗證先決條件，您可以下載一個單獨的軟體包，該軟體包僅測試先決條件。"了解如何檢查您的 Linux 主機是否已準備好安裝資料分類" 。

您所在場所的 Linux 主機上的典型安裝具有以下元件和連接。



雲端 Linux 主機上的典型安裝具有以下元件和連接。



快速啟動

按照以下步驟快速開始，或向下捲動到其餘部分以獲取完整詳細資訊。

1 建立控制台代理

如果您還沒有控制台代理， "在本機部署控制台代理"在您的網路中的 Linux 主機上，或在雲端的 Linux 主機上。

您也可以與您的雲端提供者一起建立控制台代理程式。看 "在 AWS 中建立控制台代理"， "在 Azure 中建立控制台代理" ， 或者 "在 GCP 中建立控制台代理"。

2 審查先決條件

確保您的環境能夠滿足先決條件。這包括實例的出站互聯網存取、控制台代理和資料分類之間透過連接埠 443 的連接等等。查看完整列表 。

您還需要一個符合以下條件的 Linux 系統遵循要求。

3 下載並部署資料分類

從NetApp支援網站下載雲端資料分類軟體，並將安裝程式檔案複製到您打算使用的 Linux 主機。然後啟動安裝精靈並依照指示部署資料分類實例。

建立控制台代理

您需要先安裝控制台代理，然後才能安裝和使用資料分類。在大多數情況下，您可能會在嘗試啟動資料分類之前設定控制台代理，因為大多數 "控制台功能需要控制台代理"，但有些情況下您需要立即設定一個。

若要在您的雲端提供者環境中建立一個，請參閱 "在 AWS 中建立控制台代理"， "在 Azure 中建立控制台代理" ， 或者 "在 GCP 中建立控制台代理"。

在某些情況下，您必須使用部署在特定雲端提供者中的控制台代理：

- 在 AWS 或Amazon FSx for ONTAP中的Cloud Volumes ONTAP中掃描資料時，您可以使用 AWS 中的控制台代理程式。
- 在 Azure 中的Cloud Volumes ONTAP或Azure NetApp Files中掃描資料時，您可以使用 Azure 中的控制台代理程式。

  對於Azure NetApp Files，它必須部署在與您要掃描的磁碟區相同的區域中。

- 在 GCP 中的Cloud Volumes ONTAP中掃描資料時，您可以使用 GCP 中的控制台代理程式。

可以使用任何這些雲端控制台代理程式來掃描本機ONTAP系統、 NetApp檔案共用和資料庫帳戶。

請注意，您還可以 "在本機部署控制台代理"在您的網路中的 Linux 主機上或雲端中的 Linux 主機上。一些計劃在本機安裝資料分類的使用者可能也會選擇在本機安裝控制台代理程式。

安裝資料分類時，您將需要控制台代理系統的 IP 位址或主機名稱。如果您在您的場所安裝了控制台代理，您將獲得此資訊。如果控制台代理程式部署在雲端中，您可以從控制台中找到此資訊：選擇幫助圖標，然後選擇*支援*，然後選擇控制台代理。

準備 **Linux** 主機系統

資料分類軟體必須在滿足特定作業系統需求、RAM 需求、軟體需求等的主機上運作。 Linux 主機可以在您的網路中，也可以在雲端。

確保您可以保持資料分類運行。資料分類機器需要保持開啟以持續掃描您的資料。

- 資料分類必須運行在專用主機上。主機不能與其他應用程式或第三方軟體（例如防毒軟體）共用。
- 選擇與您打算使用資料分類掃描的資料集相符的大小。

系統大小	中央處理器	**RAM**（必須停用交換記憶體）	磁碟
超大	32 個 CPU	128 GB 內存	• / 上 1 TiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 上可用 895 GiB • /tmp 上 5 GiB • 對於 **Podman**，**/var/tmp** 上有 **30 GB**
大的	16 個 CPU	64 GB 內存	• / 上 500 GiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 或 Podman /var/lib/containers 上可用 400 GiB • /tmp 上 5 GiB • 對於 **Podman**，**/var/tmp** 上有 **30 GB**

- 在雲端為資料分類安裝部署運算執行個體時，建議您使用符合上述「大型」系統需求的系統：
  - **Amazon Elastic Compute Cloud (Amazon EC2)** 執行個體類型：「m6i.4xlarge」。"查看其他 AWS 執行個體類型"。
  - **Azure VM** 大小：「Standard_D16s_v3」。"查看其他 Azure 執行個體類型"。
  - **GCP** 機器類型：「n2-standard-16」。"查看其他 GCP 實例類型"。
- **UNIX** 資料夾權限：需要以下最低 UNIX 權限：

資料夾	最低權限
/tmp	`rwxrwxrwt`
/選擇	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/系統	`rwxr-xr-x`

- 作業系統:
  - 以下作業系統需要使用 Docker 容器引擎：
    - Red Hat Enterprise Linux 版本 7.8 與 7.9
    - Ubuntu 22.04（需要資料分類版本 1.23 或更高版本）
    - Ubuntu 24.04（需要資料分類版本 1.23 或更高版本）
  - 以下作業系統需要使用 Podman 容器引擎，並且需要資料分類版本 1.30 或更高版本：
    - Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5 和 9.6。
  - 必須在主機系統上啟用進階向量擴充 (AVX2)。

- **Red Hat** 訂閱管理：主機必須在 Red Hat 訂閱管理中註冊。如果未註冊，系統將無法存取儲存庫來在安裝期間更新所需的第三方軟體。
- 附加軟體：安裝資料分類前，必須在主機上安裝以下軟體：
  - 根據您使用的作業系統，您需要安裝其中一個容器引擎：
    - Docker Engine 版本 19.3.1 或更高版本。 "查看安裝說明" 。
    - Podman 版本 4 或更高版本。若要安裝 Podman，請輸入(`sudo yum install podman netavark -y`) 。
- Python 版本 3.6 或更高版本。 "查看安裝說明" 。
  - **NTP** 注意事項： NetApp建議設定資料分類系統以使用網路時間協定 (NTP) 服務。資料分類系統和控制台代理系統之間的時間必須同步。
- **Firewalld** 注意事項：如果您打算使用 `firewalld`，我們建議您在安裝資料分類之前啟用它。運行以下命令進行配置 `firewalld`以便與資料分類相容：

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

如果您打算使用其他資料分類主機作為掃描器節點，請在此時將這些規則新增至您的主系統：

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

請注意，每次啟用或更新時都必須重新啟動 Docker 或 Podman `firewalld`設定.

安裝後，資料分類主機系統的 IP 位址無法變更。

啟用資料分類的出站互聯網訪問

資料分類需要出站網路存取。如果您的虛擬或實體網路使用代理伺服器進行網際網路訪問，請確保資料分類執行個體具有出站網際網路存取權限以聯絡下列端點。

端點	目的
\ https://api.console.netapp.com	與控制台的通信，其中包括NetApp帳戶。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	與控制台網站通信，實現集中用戶身份驗證。

端點	目的
\ https://support.compliance.api.bluexp.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ https://production.cloudflare.docker.com/.	提供對軟體映像、清單、範本的存取以及發送日誌和指標。
https://support.compliance.api.bluexp.netapp.com/	使NetApp能夠從稽核記錄中串流資料。
https://github.com/docker https://download.docker.com	提供docker安裝的必備包。
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	提供 Ubuntu 安裝的必備軟體包。

驗證所有必要的連接埠均已啟用

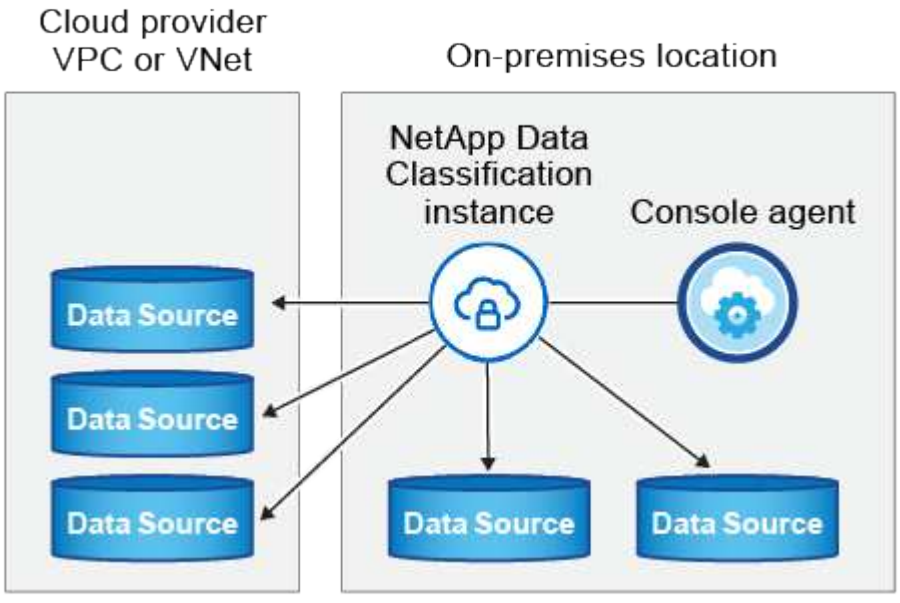
您必須確保所有必要的連接埠都已打開，以便在控制台代理程式、資料分類、Active Directory 和資料來源之間進行通訊。

連接類型	連接埠	描述
控制台代理<>資料分類	8080 (TCP)、443 (TCP) 和 80。9000	控制台代理程式的防火牆或路由規則必須允許透過連接埠 443 進出資料分類實例的入站和出站流量。確保連接埠 8080 已打開，以便您可以在控制台中看到安裝進度。如果 Linux 主機上使用防火牆，則 Ubuntu 伺服器內的內部進程需要連接埠 9000。
控制台代理<> ONTAP叢集 (NAS)	443（TCP）	控制台使用 HTTPS 發現ONTAP叢集。如果您使用自訂防火牆策略，則它們必須符合以下要求： • 控制台代理主機必須允許透過連接埠 443 進行出站 HTTPS 存取。如果控制台代理程式位於雲端中，則預先定義的防火牆或路由規則允許所有出站通訊。 • ONTAP叢集必須允許透過連接埠 443 進行入站 HTTPS 存取。預設的「mgmt」防火牆策略允許來自所有 IP 位址的入站 HTTPS 存取。如果您修改了此預設策略，或建立了自己的防火牆策略，則必須將 HTTPS 協定與該原則關聯並啟用從控制台代理主機的存取。

連接類型	連接埠	描述
資料分類 <> ONTAP集群	• 對於 NFS - 111 (TCP\UDP) 和 2049 (TCP\UDP) • 對於 CIFS - 139 (TCP\UDP) 和 445 (TCP\UDP)	資料分類需要與每個Cloud Volumes ONTAP子網路或本地ONTAP系統建立網路連線。 Cloud Volumes ONTAP的防火牆或路由規則必須允許來自資料分類實例的入站連線。 確保這些連接埠對資料分類實例開放： • 對於 NFS - 111 和 2049 • 對於 CIFS - 139 和 445 NFS 磁碟區匯出策略必須允許從資料分類實例進行存取。
資料分類<> Active Directory	389（TCP 和 UDP ）、636（TCP）、3268（TCP）和 3269（TCP）	您必須已經為公司使用者設定了 Active Directory。此外，資料分類需要 Active Directory 憑證來掃描 CIFS 磁碟區。 您必須具有 Active Directory 的資訊： • DNS 伺服器 IP 位址，或多個 IP 位址 • 伺服器的使用者名稱和密碼 • 網域名稱（Active Directory 名稱） • 您是否使用安全 LDAP (LDAPS) • LDAP 伺服器連接埠（LDAP 通常為 389，安全 LDAP 通常為 636）

在 **Linux** 主機上安裝資料分類

對於典型配置，您將在單一主機系統上安裝該軟體。請參閱此處的步驟 。



看準備 Linux 主機系統和審查先決條件了解部署資料分類之前的完整要求清單。

只要實例具有互聯網連接，資料分類軟體的升級就會自動進行。

當軟體安裝在本機時，資料分類目前無法掃描 S3 儲存桶、 Azure NetApp Files或 FSx for ONTAP 。在這些情況下，您需要在雲端中部署單獨的控制台代理程式和資料分類實例，並且 "在連接器之間切換"適用於不同的資料來源。

典型配置的單主機安裝

在單一本機上安裝資料分類軟體時，請查看要求並遵循下列步驟。

"觀看此視頻"了解如何安裝資料分類。

請注意，安裝資料分類時會記錄所有安裝活動。如果您在安裝過程中遇到任何問題，您可以查看安裝審計日誌的內容。它被寫給 `/opt/netapp/install_logs/`。

開始之前

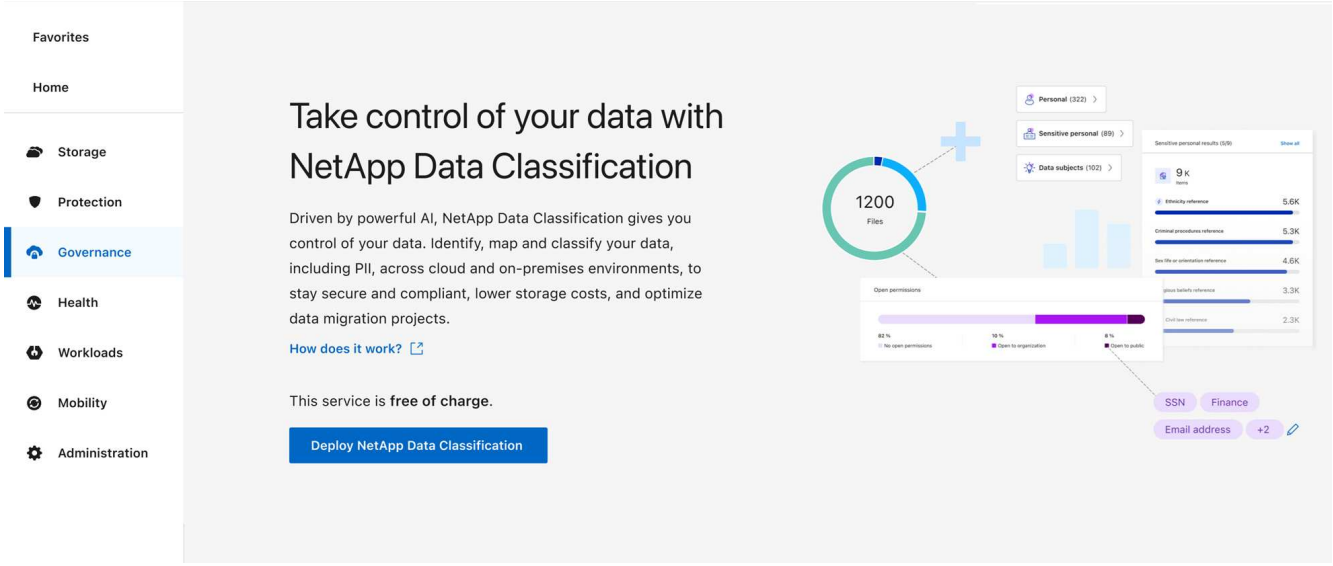
- 驗證您的 Linux 系統是否滿足主機需求。
- 驗證系統是否安裝了兩個必備軟體套件（Docker Engine 或 Podman 和 Python 3）。
- 確保您在 Linux 系統上擁有 root 權限。
- 如果您使用代理程式存取互聯網：
  - 您將需要代理伺服器資訊（IP 位址或主機名稱、連接連接埠、連接方案：https 或 http、使用者名稱和密碼）。
  - 如果代理程式正在執行 TLS 攔截，您需要知道資料分類 Linux 系統上儲存 TLS CA 憑證的路徑。
  - 代理必須是非透明的。資料分類目前不支援透明代理。
  - 該用戶必須是本機用戶。不支援網域用戶。
- 驗證您的離線環境是否符合要求權限和連線性。

步驟

1. 從下載資料分類軟體 "NetApp支援站點"。您應該選擇的檔案名稱為 **DATASENSE-INSTALLER-<version>.tar.gz**。
2. 將安裝程式檔案複製到您打算使用的 Linux 主機（使用 `scp`或其他方法）。
3. 在主機上解壓縮安裝程式文件，例如：

```
tar -xzf DATASENSE-INSTALLER-V1.25.0.tar.gz
```

4. 在控制台中，選擇*治理>分類*。
5. 選擇*在本機或雲端部署分類*。



6. 根據您是在雲端中準備的實例上還是在本機準備的實例上安裝資料分類，選擇適當的*部署*選項來啟動資料分類安裝。
7. 將顯示「在本機部署資料分類」對話方塊。複製提供的命令（例如： `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq` ）並將其貼到文字檔案中，以便稍後使用。然後選擇*關閉*以關閉對話框。
8. 在主機上，輸入您複製的命令，然後按照一系列提示進行操作，或者您可以提供包含所有必需參數的完整命令作為命令列參數。

   請注意，安裝程式會執行預檢查以確保您的系統和網路要求滿足，以便成功安裝。 "觀看此視頻"了解預檢資訊和意義。

根據提示輸入參數：	輸入完整命令：
a. 貼上從步驟 7 複製的命令： `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` 如果您在雲端實例上安裝（而不是在您的本地），請新增 `--manual-cloud-install <cloud_provider>`。 b. 輸入資料分類主機的 IP 位址或主機名，以便控制台代理系統可以存取它。 c. 輸入控制台代理主機的 IP 位址或主機名，以便資料分類系統可以存取它。 d. 根據提示輸入代理詳細資料。如果您的控制台代理已經使用代理，則無需在此處再次輸入此信息，因為資料分類將自動使用控制台代理所使用的代理。	或者，您可以提前建立整個命令，提供必要的主機和代理參數： `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token> --host <ds_host> --manager-host <cm_host> --manual-cloud-install <cloud_provider> --proxy-host <proxy_host> --proxy-port <proxy_port> --proxy-scheme <proxy_scheme> --proxy-user <proxy_user> --proxy-password <proxy_password> --cacert-folder-path <ca_cert_dir>`

變數值：

- *account_id* = NetApp帳號 ID

- *client_id* = 控制台代理客戶端 ID（如果客戶端 ID 中沒有後綴“clients”，則新增後綴）
- *user_token* = JWT 使用者存取權令牌
- *ds_host* = 資料分類 Linux 系統的 IP 位址或主機名稱。
- *cm_host* = 控制台代理系統的 IP 位址或主機名稱。
- *cloud_provider* = 在雲端實例上安裝時，根據雲端提供者輸入「AWS」、「Azure」或「Gcp」。
- *proxy_host* = 如果主機位於代理伺服器後面，則為代理伺服器的 IP 或主機名稱。
- *proxy_port* = 連接到代理伺服器的連接埠（預設為 80）。
- *proxy_scheme* = 連接方案：https 或 http（預設 http）。
- *proxy_user* = 如果需要基本驗證，則經過驗證的使用者連接到代理伺服器。使用者必須是本機使用者 - 不支援網域使用者。
- *proxy_password* = 您指定的使用者名稱的密碼。
- *ca_cert_dir* = 資料分類 Linux 系統上包含附加 TLS CA 憑證包的路徑。僅當代理執行 TLS 攔截時才需要。

結果

資料分類安裝程式安裝套件、註冊安裝並安裝資料分類。安裝可能需要 10 到 20 分鐘。

如果主機和控制台代理執行個體之間透過連接埠 8080 建立連接，您將在控制台的「資料分類」標籤中看到安裝進度。

下一步

您可以從設定頁面選擇要掃描的資料來源。

在沒有網路存取的 **Linux** 主機上安裝**NetApp Data Classification**

在沒有網路存取權限的本機站點的 Linux 主機上安裝NetApp Data Classification稱為_私有模式_。這種安裝類型使用安裝腳本，與NetApp Console SaaS 層沒有連線。

BlueXP私有模式（傳統BlueXP介面）通常用於沒有網路連線的本機環境和安全雲端區域，其中包括 AWS Secret Cloud、AWS Top Secret Cloud 和 Azure IL6。NetApp繼續透過傳統的BlueXP介面支援這些環境。有關舊版BlueXP介面中的私有模式文檔，請參閱"BlueXP私人模式的 PDF 文檔"。

檢查您的 **Linux** 主機是否已準備好安裝**NetApp Data Classification**

在 Linux 主機上手動安裝NetApp Data Classification之前，可以選擇在主機上執行腳本以驗證安裝資料分類的所有先決條件是否都已具備。您可以在網路中的 Linux 主機上或雲端中的 Linux 主機上執行此腳本。主機可以連接到互聯網，或者主機可以駐留在沒有互聯網訪問的站點（暗站）。

資料分類安裝腳本包含一個測試腳本，以確保您的環境符合要求。您可以單獨執行此腳本來驗證 Linux 主機是否已準備就緒，然後再執行安裝腳本。

入門

您將執行以下任務。

- 如果您尚未安裝控制台代理，則可以選擇安裝。您可以在未安裝控制台代理的情況下執行測試腳本，但腳本會檢查控制台代理程式和資料分類主機之間的連線 - 因此建議您安裝控制台代理程式。
- 準備主機並驗證其是否滿足所有要求。
- 啟用資料分類主機的出站網際網路存取。
- 驗證所有系統上的所有必要連接埠是否都已啟用。
- 下載並執行先決條件測試腳本。

建立控制台代理

您需要先安裝控制台代理，然後才能安裝和使用資料分類。但是，您可以在沒有控制台代理的情況下執行先決條件腳本。

你可以 "在本機安裝控制台代理" 在您網路中的 Linux 主機上，或是在雲端的 Linux 主機上。如果控制台代理程式安裝在本機，您也可以在本機安裝資料分類。

若要在您的雲端提供者環境中建立控制台代理，請參閱：

- "在 AWS 中建立控制台代理"
- "在 Azure 中建立控制台代理"
- "在 GCP 中建立控制台代理"

執行先決條件腳本時，需要控制台代理系統的 IP 位址或主機名稱。如果您在本機安裝了控制台代理，則可以取得此資訊。如果控制台代理程式部署在雲端，您可以從控制台中找到此資訊：選擇說明圖標，然後選擇「支援」；在「代理程式和審核」部分，選擇「前往代理程式」。

驗證主機需求

資料分類軟體必須運行在滿足特定作業系統要求、記憶體需求和軟體要求的主機上。

- 資料分類必須運行在專用主機上。主機不能與其他應用程式或第三方軟體（例如防毒軟體）共用。
- 選擇與您打算使用資料分類掃描的資料集相符的大小。

系統大小	中央處理器	**RAM**（必須停用交換記憶體）	磁碟
超大	32 個 CPU	128 GB 內存	• / 上 1 TiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 上可用 895 GiB • /tmp 上 5 GiB • 對於 **Podman**，**/var/tmp** 上有 **30 GB**

系統大小	中央處理器	**RAM**（必須停用交換記憶體）	磁碟
大的	16 個 CPU	64 GB 內存	• / 上 500 GiB SSD，或 /opt 上 100 GiB 可用 • /var/lib/docker 或 Podman /var/lib/containers 上可用 400 GiB • /tmp 上 5 GiB • 對於 **Podman**，**/var/tmp** 上有 **30 GB**

- 在雲端為資料分類安裝部署運算執行個體時，建議您使用符合上述「大型」系統需求的系統：
  - **Amazon Elastic Compute Cloud (Amazon EC2)** 執行個體類型：「m6i.4xlarge」。"查看其他 AWS 執行個體類型"。
  - **Azure VM** 大小：「Standard_D16s_v3」。"查看其他 Azure 執行個體類型"。
  - **GCP** 機器類型：「n2-standard-16」。"查看其他 GCP 實例類型"。
- **UNIX** 資料夾權限：需要以下最低 UNIX 權限：

資料夾	最低權限
/tmp	`rwxrwxrwt`
/選擇	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/系統	`rwxr-xr-x`

- 作業系統:
  - 以下作業系統需要使用 Docker 容器引擎：
    - Red Hat Enterprise Linux 版本 7.8 與 7.9
    - Ubuntu 22.04（需要資料分類版本 1.23 或更高版本）
    - Ubuntu 24.04（需要資料分類版本 1.23 或更高版本）
  - 以下作業系統需要使用 Podman 容器引擎，並且需要資料分類版本 1.30 或更高版本：
    - Red Hat Enterprise Linux 版本 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5 和 9.6。
  - 必須在主機系統上啟用進階向量擴充 (AVX2)。
- **Red Hat** 訂閱管理：主機必須在 Red Hat 訂閱管理中註冊。如果未註冊，系統將無法存取儲存庫來在安裝期間更新所需的第三方軟體。
- 附加軟體：安裝資料分類前，必須在主機上安裝以下軟體：
  - 根據您使用的作業系統，您需要安裝其中一個容器引擎：
    - Docker Engine 版本 19.3.1 或更高版本。"查看安裝說明"。
    - Podman 版本 4 或更高版本。若要安裝 Podman，請輸入(`sudo yum install podman netavark -y`)。

- Python 版本 3.6 或更高版本。 "查看安裝說明" 。
  - **NTP** 注意事項： NetApp建議設定資料分類系統以使用網路時間協定 (NTP) 服務。資料分類系統和控制台代理系統之間的時間必須同步。
- **Firewalld** 注意事項：如果您打算使用 `firewalld`，我們建議您在安裝資料分類之前啟用它。運行以下命令進行配置 `firewalld`以便與資料分類相容：

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

如果您打算使用其他資料分類主機作為掃描器節點（在分散式模型中），請在此時將這些規則新增至您的主系統：

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

請注意，每次啟用或更新時都必須重新啟動 Docker 或 Podman `firewalld`設定.

啟用資料分類的出站互聯網訪問

資料分類需要出站網路存取。如果您的虛擬或實體網路使用代理伺服器進行網際網路訪問，請確保資料分類執行個體具有出站網際網路存取權限以聯絡下列端點。

對於安裝在沒有網路連線的站點的主機系統，則不需要本節。

端點	目的
\ https://api.console.netapp.com	與控制台服務（包括NetApp帳戶）的通訊。
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	與控制台網站通信，實現集中用戶身份驗證。
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https https://dseasb33srnrn.cloudfront.net/ // https://production.cloudflare.docker.com/	提供對軟體映像、清單、範本的存取以及發送日誌和指標。
\ https://support.compliance.api.console.netapp.com/	使NetApp能夠從稽核記錄中串流資料。

端點	目的
https://github.com/docker https://download.docker.com	提供docker安裝的必備包。
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	提供 Ubuntu 安裝的必備軟體包。

驗證所有必要的連接埠均已啟用

您必須確保所有必要的連接埠都已打開，以便在控制台代理程式、資料分類、Active Directory 和資料來源之間進行通訊。

連接類型	連接埠	描述
控制台代理<>資料分類	8080 (TCP)、443 (TCP)和 80。9000	控制台代理程式的防火牆或路由規則必須允許透過連接埠 443 進出資料分類實例的入站和出站流量。確保連接埠 8080 已打開，以便您可以在控制台中看到安裝進度。如果 Linux 主機上使用防火牆，則 Ubuntu 伺服器內的內部進程需要連接埠 9000。
控制台代理<> ONTAP叢集 (NAS)	443（TCP）	控制台使用 HTTPS 發現ONTAP叢集。如果您使用自訂防火牆策略，控制台代理主機必須允許透過連接埠 443 進行出站 HTTPS 存取。如果控制台代理程式位於雲端中，則預先定義的防火牆或路由規則允許所有出站通訊。

執行資料分類先決條件腳本

請依照以下步驟執行資料分類先決條件腳本。

"觀看此視頻"了解如何執行先決條件腳本並解釋結果。

開始之前

- 驗證您的 Linux 系統是否滿足主機需求。
- 驗證系統是否安裝了兩個必備軟體套件（Docker Engine 或 Podman 和 Python 3）。
- 確保您在 Linux 系統上擁有 root 權限。

步驟

1. 從下載資料分類先決條件腳本 "NetApp支援站點"。您應該選擇的檔案名稱為 **standalone-pre-requisite-tester-<version>**。
2. 將檔案複製到您計劃使用的 Linux 主機（使用 `scp`或其他方法）。
3. 分配運行腳本的權限。

```
chmod +x standalone-pre-requisite-tester-v1.25.0
```

4. 使用以下命令運行腳本。

```
./standalone-pre-requisite-tester-v1.25.0 <--darksite>
```

僅當您在沒有網際網路存取的主機上執行腳本時才新增選項「--darksite」。當主機未連接到網際網路時，某些先決條件測試將被跳過。

5. 腳本會提示您輸入資料分類主機的 IP 位址。
   - 輸入 IP 位址或主機名稱。
6. 該腳本會提示您是否安裝了控制台代理程式。
   - 如果您沒有安裝控制台代理，請輸入 **N**。
   - 如果您確實安裝了控制台代理，請輸入 **Y**。然後輸入控制台代理的 IP 位址或主機名，以便測試腳本可以測試此連線。
7. 該腳本在系統上執行各種測試，並在運行過程中顯示結果。完成後，它會將會話日誌寫入名為 `prerequisites-test-<timestamp>.log`在目錄中 `/opt/netapp/install_logs`。

結果

如果所有先決條件測試均成功運行，您可以在準備就緒後在主機上安裝資料分類。

如果發現任何問題，則將其歸類為“建議”或“需要”修復。建議的問題通常是會使資料分類掃描和分類任務運作速度變慢的項目。這些項目不需要糾正 - 但您可能需要解決它們。

如果您有任何「必要」問題，您應該修復這些問題並再次執行先決條件測試腳本。

啟動資料來源掃描

使用**NetApp Data Classification**掃描資料來源

NetApp Data Classification會掃描您選擇的儲存庫（磁碟區、資料庫模式或其他使用者資料）中的數據，以識別個人和敏感資料。然後，資料分類會對應您的組織資料、對每個文件進行分類並識別資料中的預定義模式。掃描結果是個人資訊、敏感個人資訊、資料類別和文件類型的索引。

初步掃描後，資料分類將以循環方式持續掃描您的資料以偵測增量變化。這就是為什麼保持實例運行很重要。

您可以在磁碟區層級或資料庫模式層級啟用和停用掃描。

映射掃描和分類掃描之間有什麼區別

您可以在資料分類中進行兩種類型的掃描：

- 僅映射掃描僅提供資料的進階概覽，並在選定的資料來源上執行。僅映射掃描比映射和分類掃描花費的時間更少，因為它們不存取文件來查看其中的資料。您可能希望首先執行此操作來確定研究領域，然後對這些領域執行地圖和分類掃描。
- 地圖和分類掃描 為您的資料提供深層掃描。

下表顯示了一些差異：

特徵	映射和分類掃描	僅映射掃描
掃描速度	慢的	快速地
定價	自由的	自由的
容量	限制為 500 TiB*	限制為 500 TiB*
文件類型和已使用容量列表	是的	是的
文件數量和已用容量	是的	是的
文件的年齡和大小	是的	是的
能夠運行"數據映射報告"	是的	是的
數據調查頁面查看文件詳細信息	是的	不
在文件中搜尋名稱	是的	不
創造"已儲存的查詢"提供自訂搜尋結果	是的	不
能夠運行其他報告	是的	不
能夠查看文件中的元資料**	不	是的

* 資料分類不會對其可掃描的資料量施加限制。每個控制台代理程式支援掃描和顯示 500 TiB 的資料。要掃描超過 500 TiB 的數據，"安裝另一個控制台代理"然後"部署另一個資料分類實例"。 + 控制台 UI 顯示來自單一連接器的資料。有關查看來自多個控制台代理的資料的提示，請參閱"使用多個控制台代理"。

** 在映射掃描期間從檔案中提取以下元資料：

- 系統
- 系統類型
- 儲存庫
- 文件類型
- 已用容量
- 文件數量
- 文件大小
- 文件創建
- 文件上次訪問
- 文件上次修改時間
- 文件發現時間
- 權限擷取

治理儀表板差異：

特徵	地圖和分類	地圖
過時的數據	是的	是的
非業務數據	是的	是的
重複文件	是的	是的
預定義已儲存的查詢	是的	不
預設儲存的查詢	是的	是的
DDA 報告	是的	是的
地圖報告	是的	是的
靈敏度等級檢測	是的	不
具有廣泛權限的敏感數據	是的	不
開放權限	是的	是的
數據時代	是的	是的
數據大小	是的	是的
類別	是的	不
文件類型	是的	是的

合規性儀表板差異：

特徵	地圖和分類	地圖
個人資訊	是的	不
敏感個人資訊	是的	不
隱私風險評估報告	是的	不
HIPAA 報告	是的	不
PCI DSS 報告	是的	不

調查過濾器差異：

特徵	地圖和分類	地圖
已儲存的查詢	是的	是的
系統類型	是的	是的
系統	是的	是的
儲存庫	是的	是的
文件類型	是的	是的
文件大小	是的	是的
創建時間	是的	是的
發現時間	是的	是的
上次修改時間	是的	是的
上次訪問	是的	是的
開放權限	是的	是的
檔案目錄路徑	是的	是的
類別	是的	不
敏感度等級	是的	不
識別符數量	是的	不
個人資料	是的	不
敏感個人數據	是的	不
資料主體	是的	不
重複項	是的	是的
分類狀態	是的	狀態始終為“見解有限”
掃描分析事件	是的	是的
文件哈希	是的	是的
有存取權限的使用者數	是的	是的
使用者/群組權限	是的	是的
文件所有者	是的	是的
目錄類型	是的	是的

使用NetApp Data Classification掃描Amazon FSx for ONTAP卷

完成幾個步驟即可使用NetApp Data Classification掃描Amazon FSx for ONTAP磁碟區。

開始之前

- 您需要 AWS 中一個活動的控制台代理來部署和管理資料分類。
- 建立系統時選擇的安全群組必須允許來自資料分類實例的流量。您可以使用連接到 FSx for ONTAP檔案系統的 ENI 找到關聯的安全性群組，並使用 AWS 管理主控台對其進行編輯。

  "Linux 執行個體的 AWS 安全群組"

  "Windows 執行個體的 AWS 安全性群組"

  "AWS 彈性網路介面 (ENI)"

- 確保以下連接埠對資料分類實例開放：
  - 對於 NFS – 連接埠 111 和 2049。
  - 對於 CIFS – 連接埠 139 和 445。

部署資料分類實例

"部署資料分類"如果尚未部署實例。

您應該在與 AWS 主控台代理程式和要掃描的 FSx 磁碟區相同的 AWS 網路中部署資料分類。

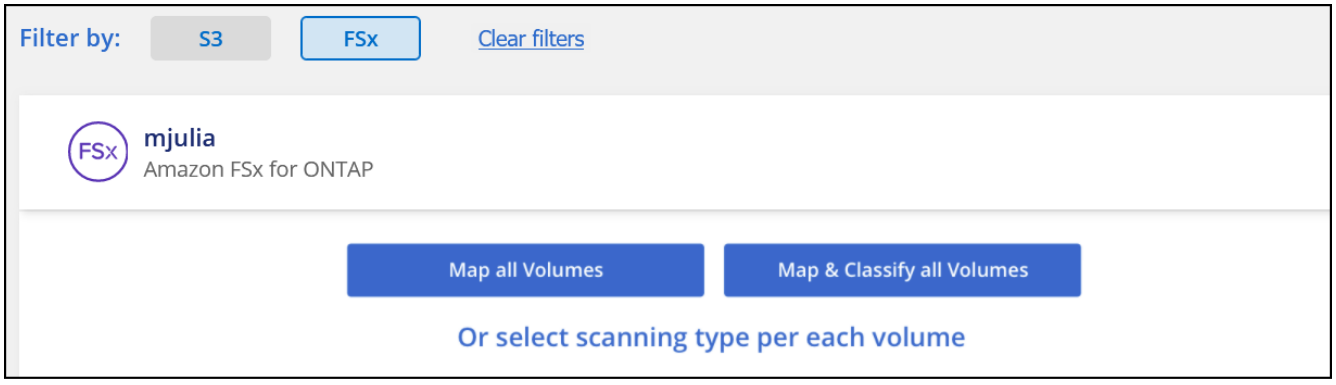
*注意：*掃描 FSx 磁碟區時，目前不支援在本機位置部署資料分類。

只要實例具有互聯網連接，資料分類軟體的升級就會自動進行。

在您的系統中啟用資料分類

您可以為 FSx for ONTAP磁碟區啟用資料分類。

1. 從NetApp Console，治理 > 分類。
2. 從資料分類選單中，選擇*配置*。



3. 選擇如何掃描每個系統中的磁碟區。"了解映射和分類掃描"：
   - 若要對應所有捲，請選擇*映射所有磁碟區*。
   - 若要對應和分類所有捲，請選擇*映射和分類所有捲*。
   - 若要自訂每個磁碟區的掃描，請選擇*或選擇每個磁碟區的掃描類型*，然後選擇要對應和/或分類的磁碟

區。

4. 在確認對話方塊中，選擇「核准」以使資料分類開始掃描您的磁碟區。

結果

資料分類開始掃描您在系統中選擇的磁碟區。一旦資料分類完成初始掃描，結果將在合規性儀表板中提供。所需時間取決於資料量——可能是幾分鐘或幾小時。您可以透過導航到配置選單然後選擇系統配置來追蹤初始掃描的進度。在進度條中追蹤每次掃描的進度；您可以將滑鼠懸停在進度條上，以查看相對於磁碟區中總檔案數的掃描檔案數。

- 預設情況下，如果資料分類在 CIFS 中沒有寫入屬性權限，或者在 NFS 中沒有寫入權限，系統將不會掃描磁碟區中的文件，因為資料分類無法將「上次存取時間」還原為原始時間戳記。如果您不介意上次造訪時間是否重設，請選擇*或為每個磁碟區選擇掃描類型*。結果頁面有一個您可以啟用的設置，以便資料分類可以掃描卷，而不管權限如何。
- 資料分類僅掃描卷下的一個檔案共用。如果您的磁碟區中有多個共享，則需要將這些其他共享作為共享群組單獨掃描。"查看有關此數據分類限制的更多詳細信息" 。

驗證資料分類是否有權存取卷

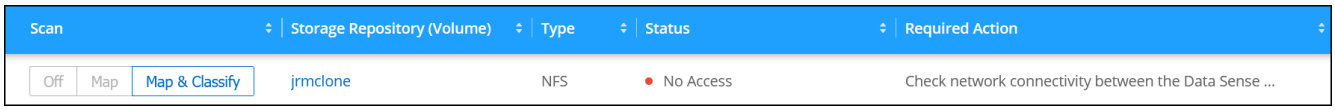
透過檢查網路、安全群組和匯出策略，確保資料分類可以存取磁碟區。

您需要向資料分類提供 CIFS 憑證，以便它可以存取 CIFS 磁碟區。

步驟

1. 從資料分類選單中，選擇*配置*。
2. 在配置頁面上，選擇*查看詳細資訊*以查看狀態並修正任何錯誤。

   例如，下圖顯示由於資料分類實例和磁碟區之間的網路連線問題，資料分類無法掃描磁碟區。



3. 確保資料分類實例與包含 FSx for ONTAP磁碟區的每個網路之間存在網路連線。

   對於 FSx for ONTAP，資料分類只能掃描與控制台位於相同區域的磁碟區。

4. 確保 NFS 磁碟區匯出策略包含資料分類執行個體的 IP 位址，以便它可以存取每個磁碟區上的資料。
5. 如果您使用 CIFS，請向資料分類提供 Active Directory 憑證，以便它可以掃描 CIFS 磁碟區。

   a. 從資料分類選單中，選擇*配置*。

   b. 對於每個系統，選擇*編輯 CIFS 憑證*並輸入資料分類存取系統上的 CIFS 磁碟區所需的使用者名稱和密碼。

      憑證可以是唯讀的，但提供管理員憑證可確保資料分類可以讀取任何需要提升權限的資料。憑證儲存在資料分類實例上。

      如果您想要確保檔案的「上次存取時間」不會因資料分類掃描而改變，建議使用者在 CIFS 中具有寫入屬性權限或在 NFS 中具有寫入權限。如果可能，請將 Active Directory 使用者設定為組織中具有所有檔案權限的父群組的一部分。

輸入憑證後，您應該會看到一條訊息，表示所有 CIFS 磁碟區均已成功驗證。

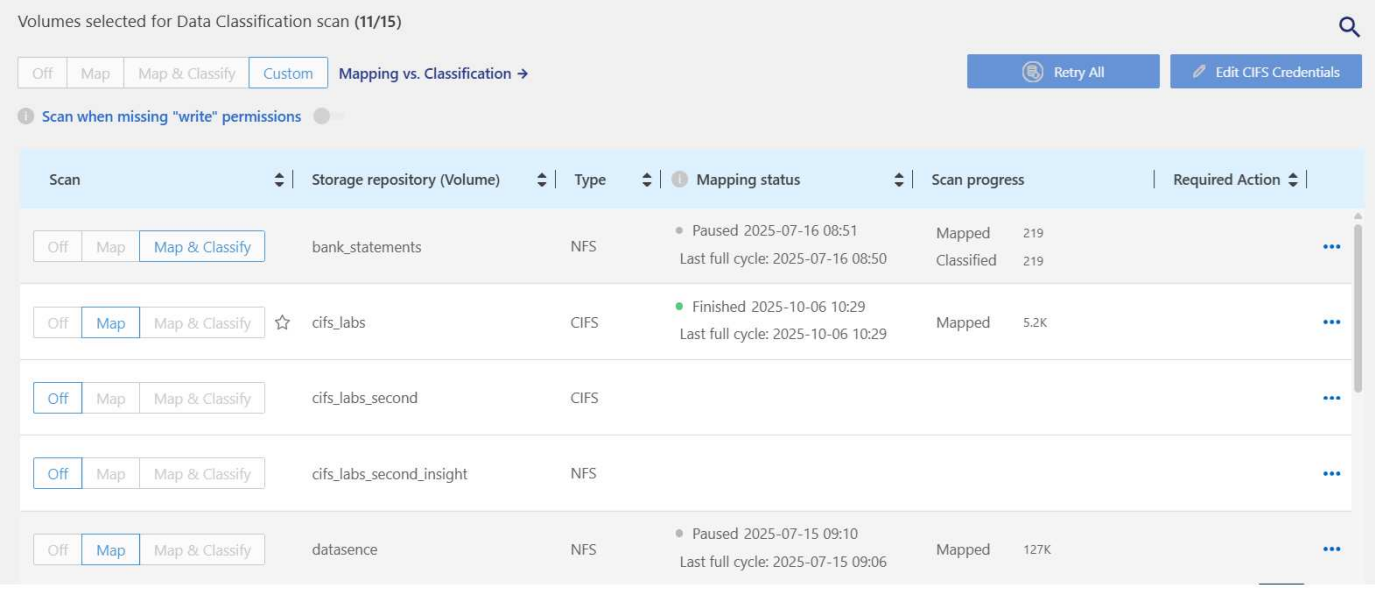
啟用和停用磁碟區上的掃描

您可以隨時從設定頁面啟動或停止任何系統上的掃描。您也可以將掃描從僅映射掃描切換到映射和分類掃描，反之亦然。建議您掃描系統中的所有磁碟區。

只有當您在標題區域中選擇了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增至系統的新磁碟區。當在標題區域設定為*自訂*或*關閉*時，您需要在系統中新增的每個新磁碟區上啟動對應和/或完整掃描。

頁面頂部的「缺少「寫入」權限時掃描」開關預設為停用狀態。這意味著，如果資料分類在 CIFS 中沒有寫入屬性權限或在 NFS 中沒有寫入權限，系統將不會掃描文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。如果您不介意是否重設上次造訪時間，請開啟開關，無論權限為何，都會掃描所有檔案。"了解更多"。

只有當您在標題區域中設定了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增到系統的新磁碟區。當所有磁碟區的設定都是「自訂」或「關閉」時，您需要為新增的每個新磁碟區手動啟動掃描。



步驟

1. 從資料分類選單中，選擇*配置*。
2. 選擇一個系統，然後選擇*配置*。
3. 若要啟用或停用所有磁碟區的掃描，請在所有磁碟區上方的標題中選擇對應、對應和分類或關閉。

   若要啟用或停用對單一卷的掃描，請在清單中找到該卷，然後選擇卷名稱旁的映射、映射和分類或關閉。

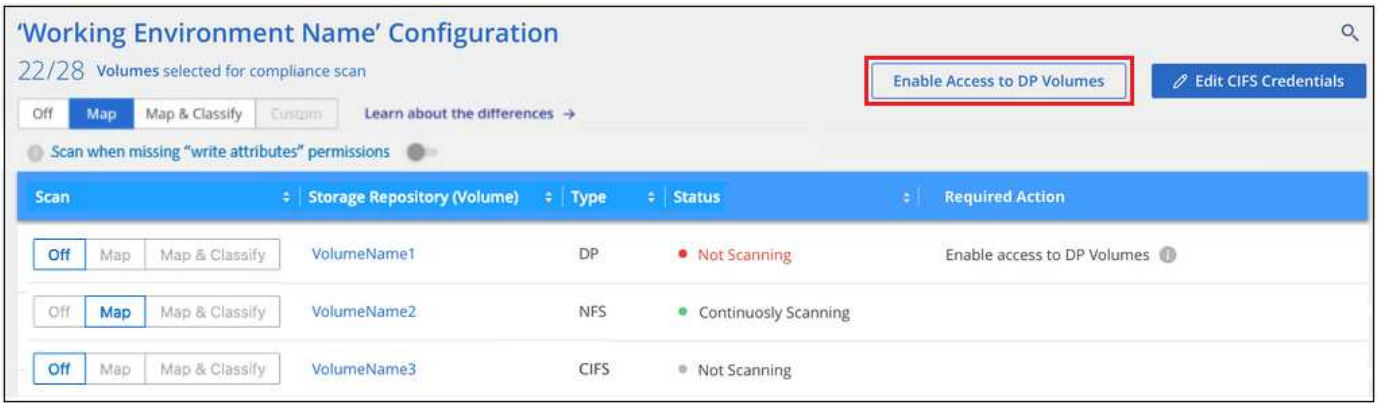
結果

當您啟用掃描時，資料分類將開始掃描您在系統中選擇的磁碟區。一旦資料分類開始掃描，結果就會開始出現在合規性儀表板中。掃描完成時間取決於資料量，從幾分鐘到幾小時不等。

掃描資料保護卷

預設情況下，不會掃描資料保護 (DP) 卷，因為它們未暴露在外部，且資料分類無法存取它們。這些是來自 FSx for ONTAP檔案系統的SnapMirror操作的目標磁碟區。

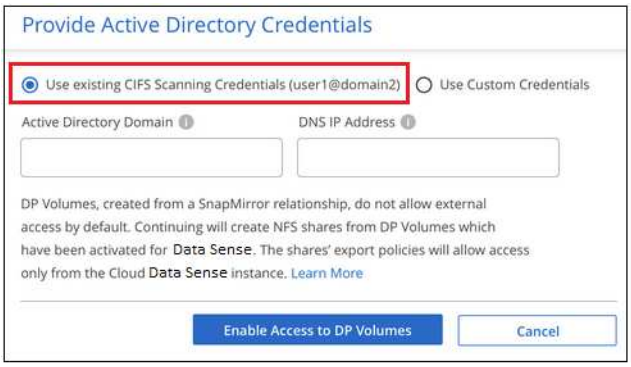
最初，磁碟區清單將這些磁碟區標識為_類型_ **DP**，其_狀態_ 未掃描*和_所需操作_ *啟用對 **DP** 磁碟區的存取。

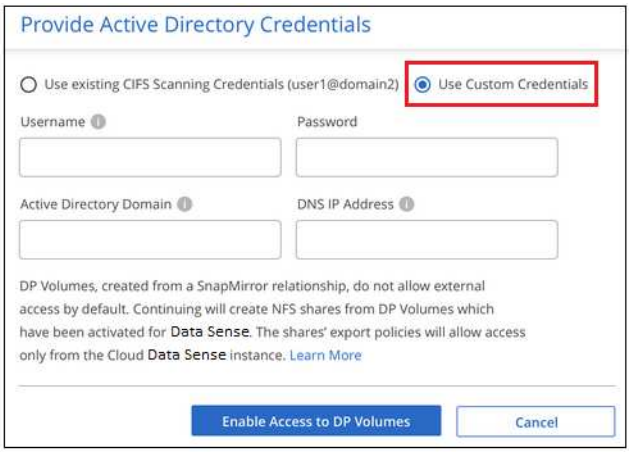


步驟

如果要掃描這些資料保護磁碟區：

1. 從資料分類選單中，選擇*配置*。
2. 選擇頁面頂部的「啟用對 DP 磁碟區的存取」*。
3. 查看確認訊息並再次選擇*啟用對 DP 磁碟區的存取*。
   - 最初在來源 FSx for ONTAP檔案系統中建立為 NFS 磁碟區的磁碟區已啟用。
   - 最初在來源 FSx for ONTAP檔案系統中建立為 CIFS 磁碟區的磁碟區會要求您輸入 CIFS 憑證來掃描這些 DP 磁碟區。如果您已經輸入了 Active Directory 憑證，以便資料分類可以掃描 CIFS 卷，您可以使用這些憑證，或者您可以指定一組不同的管理員憑證。





4. 啟動您想要掃描的每個 DP 磁碟區。

結果

一旦啟用，資料分類將從每個啟動掃描的 DP 磁碟區建立一個 NFS 共用。共享導出策略僅允許從資料分類實例進行存取。

如果您在最初啟用對 DP 磁碟區的存取權時沒有 CIFS 資料保護卷，後來又添加了一些，則按鈕 啟用對 **CIFS DP** 的存取 將出現在設定頁面的頂部。選擇此按鈕並新增 CIFS 憑證以啟用對這些 CIFS DP 磁碟區的存取。

Active Directory 憑證僅在第一個 CIFS DP 磁碟區的儲存 VM 中註冊，因此該 SVM 上的所有 DP 磁碟區都將掃描。駐留在其他 SVM 上的任何磁碟區都不會註冊 Active Directory 憑證，因此不會掃描這些 DP 磁碟區。

使用NetApp Data Classification掃描Azure NetApp Files卷

完成幾個步驟即可開始使用適用於Azure NetApp Files的NetApp Data Classification。

發現要掃描的**Azure NetApp** 檔案系統**Discover the Azure NetApp Files system that you want to scan**

如果要掃描的Azure NetApp Files系統尚未作為系統出現在NetApp Console中，"將其新增至系統頁面" 。

部署資料分類實例

"部署資料分類"如果尚未部署實例。

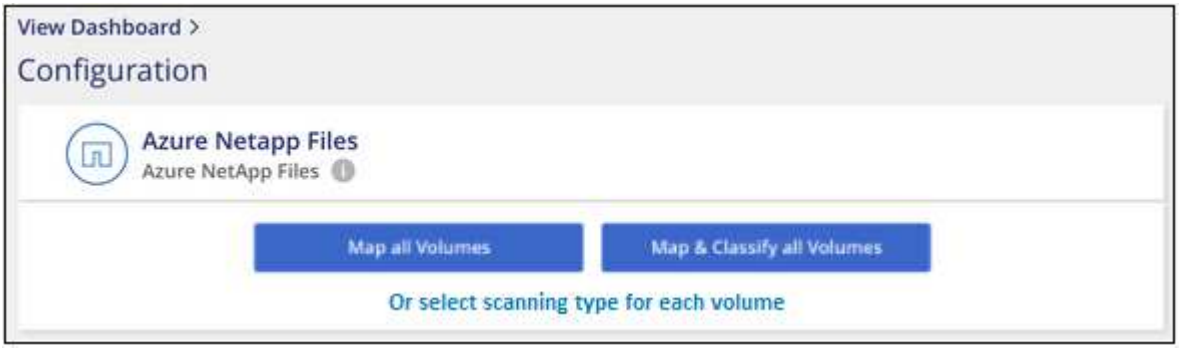
掃描Azure NetApp Files區時，資料分類必須部署在雲端中，並且必須部署在與要掃描的磁碟區相同的區域中。

*注意：*掃描Azure NetApp Files區時，目前不支援在本機位置部署資料分類。

在您的系統中啟用資料分類

您可以在Azure NetApp Files區上啟用資料分類。

1. 從資料分類選單中，選擇*配置*。



2. 選擇如何掃描每個系統中的磁碟區。"了解映射和分類掃描"：
   - 若要對應所有捲，請選擇*映射所有磁碟區*。
   - 若要對應和分類所有捲，請選擇*映射和分類所有捲*。
   - 若要自訂每個磁碟區的掃描，請選擇*或選擇每個磁碟區的掃描類型*，然後選擇要對應或對應和分類的磁碟區。

     看啟用或停用磁碟區掃描了解詳情。

3. 在確認對話方塊中，選擇*批准*。

結果

資料分類開始掃描您在系統中選擇的磁碟區。一旦資料分類完成初始掃描，結果就會顯示在合規性儀表板中。所需時間取決於資料量——可能是幾分鐘或幾小時。您可以透過導航到配置選單然後選擇系統配置來追蹤初始掃描的進度。資料分類顯示每次掃描的進度條。您可以將滑鼠懸停在進度條上，以查看相對於磁碟區中檔案總數的已掃描檔案數。

- 預設情況下，如果資料分類在 CIFS 中沒有寫入屬性權限，或者在 NFS 中沒有寫入權限，系統將不會掃描磁碟區中的文件，因為資料分類無法將「上次存取時間」還原為原始時間戳記。如果您不介意上次造訪時間是否重設，請選擇*或為每個磁碟區選擇掃描類型*。結果頁面有一個您可以啟用的設置，以便資料分類可以掃描卷，而不管權限如何。
- 資料分類僅掃描卷下的一個檔案共用。如果您的磁碟區中有多個共享，則需要將這些其他共享作為共享群組單獨掃描。"了解此資料分類限制" 。

驗證資料分類是否有權存取卷

透過檢查網路、安全群組和匯出策略，確保資料分類可以存取磁碟區。您需要為資料分類提供 CIFS 憑證，以便它可以存取 CIFS 磁碟區。

對於Azure NetApp Files，資料分類只能掃描與控制台位於同一區域的磁碟區。

清單

- 確保資料分類實例與包含Azure NetApp Files區的每個網路之間存在網路連線。
- 確保以下連接埠對資料分類實例開放：
  - 對於 NFS – 連接埠 111 和 2049。
  - 對於 CIFS – 連接埠 139 和 445。
- 確保 NFS 磁碟區匯出策略包含資料分類執行個體的 IP 位址，以便它可以存取每個磁碟區上的資料。

步驟

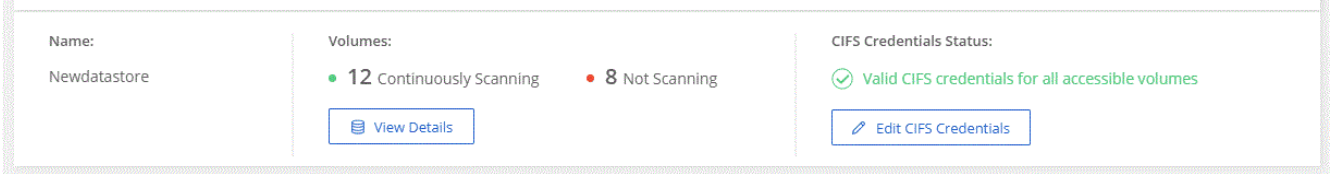
1. 從資料分類選單中，選擇*配置*。

   a. 如果您使用 CIFS（SMB），請確保 Active Directory 憑證正確。對於每個系統，選擇*編輯 CIFS 憑證*，然後輸入資料分類存取系統上的 CIFS 磁碟區所需的使用者名稱和密碼。

      憑證可以是唯讀的；提供管理員憑證可確保資料分類可以讀取任何需要提升權限的資料。憑證儲存在資料分類實例上。

      如果您想要確保檔案的「上次存取時間」不會因資料分類掃描而改變，建議使用者在 CIFS 中具有寫入屬性權限或在 NFS 中具有寫入權限。如果可能，請將 Active Directory 使用者設定為組織中具有所有檔案權限的父群組的一部分。

      輸入憑證後，您應該會看到一條訊息，表示所有 CIFS 磁碟區均已成功驗證。



2. 在設定頁面上，選擇*查看詳細資訊*以查看每個 CIFS 和 NFS 磁碟區的狀態。如有必要，請修正任何錯誤，

例如網路連線問題。

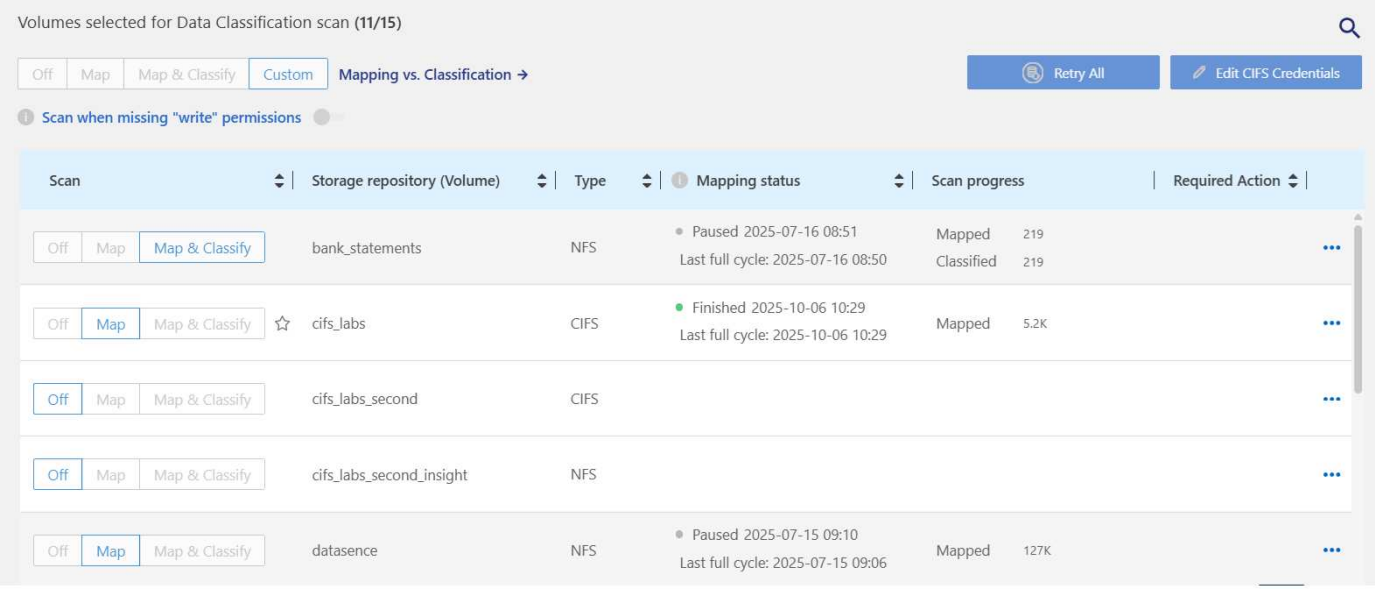
啟用或停用磁碟區掃描

您可以隨時從設定頁面啟動或停止任何系統上的掃描。您也可以將掃描從僅映射掃描切換到映射和分類掃描，反之亦然。建議您掃描系統中的所有磁碟區。

只有當您在標題區域中選擇了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增至系統的新磁碟區。當在標題區域設定為*自訂*或*關閉*時，您需要在系統中新增的每個新磁碟區上啟動對應和/或完整掃描。

頁面頂部的「缺少「寫入」權限時掃描」開關預設為停用狀態。這意味著，如果資料分類在 CIFS 中沒有寫入屬性權限或在 NFS 中沒有寫入權限，系統將不會掃描文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。如果您不介意是否重設上次造訪時間，請開啟開關，無論權限為何，都會掃描所有檔案。"了解更多"。

只有當您在標題區域中設定了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增到系統的新磁碟區。當所有磁碟區的設定都是「自訂」或「關閉」時，您需要為新增的每個新磁碟區手動啟動掃描。



步驟

1. 從資料分類選單中，選擇*配置*。
2. 選擇一個系統，然後選擇*配置*。
3. 若要啟用或停用所有磁碟區的掃描，請在所有磁碟區上方的標題中選擇對應、對應和分類或關閉。

   若要啟用或停用對單一卷的掃描，請在清單中找到該卷，然後選擇卷名稱旁的映射、映射和分類或關閉。

結果

當您啟用掃描時，資料分類將開始掃描您在系統中選擇的磁碟區。一旦資料分類開始掃描，結果就會開始出現在合規性儀表板中。掃描完成時間取決於資料量，從幾分鐘到幾小時不等。

使用NetApp Data Classification掃描Cloud Volumes ONTAP和本地ONTAP卷

完成幾個步驟即可開始使用NetApp Data Classification掃描您的Cloud Volumes ONTAP和本機ONTAP磁碟區。

先決條件

在啟用資料分類之前，請確保您具有支援的配置。

- 如果您正在掃描可透過網路存取的Cloud Volumes ONTAP和本機ONTAP系統，您可以"在雲端部署資料分類"或者"在可以存取互聯網的本地位置"。
- 如果您要掃描安裝在沒有網路存取的暗站中的本機ONTAP系統，則需要"在沒有網路存取的相同本地位置部署資料分類"。這要求將控制台代理部署在同一本地位置。

驗證資料分類是否有權存取卷

透過檢查網路、安全群組和匯出策略，確保資料分類可以存取磁碟區。您需要向資料分類提供 CIFS 憑證，以便它可以存取 CIFS 磁碟區。

清單

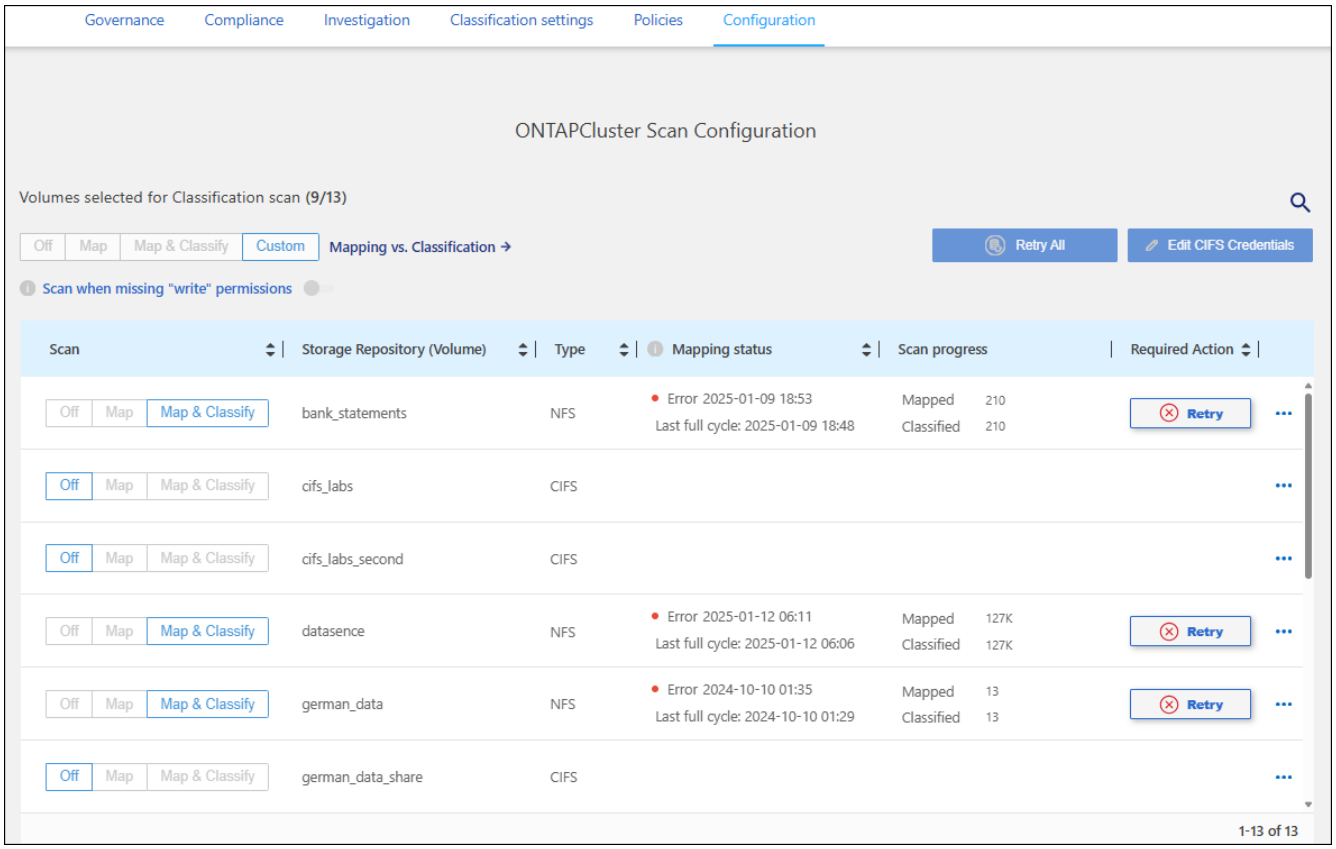
- 確保資料分類實例與包含Cloud Volumes ONTAP或本地ONTAP叢集的磁碟區的每個網路之間存在網路連線。
- 確保Cloud Volumes ONTAP的安全群組允許來自資料分類實例的入站流量。

  您可以為來自資料分類實例的 IP 位址的流量開啟安全性群組，也可以為來自虛擬網路內部的所有流量開啟安全性群組。

- 確保 NFS 磁碟區匯出策略包含資料分類執行個體的 IP 位址，以便它可以存取每個磁碟區上的資料。

步驟

1. 從資料分類選單中，選擇*配置*。



2. 如果您使用 CIFS，請向資料分類提供 Active Directory 憑證，以便它可以掃描 CIFS 磁碟區。對於每個系統，選擇*編輯 CIFS 憑證*並輸入資料分類存取系統上的 CIFS 磁碟區所需的使用者名稱和密碼。

   憑證可以是唯讀的，但提供管理員憑證可確保資料分類可以讀取任何需要提升權限的資料。憑證儲存在資料分類實例上。

   如果您想要確保檔案的「上次存取時間」不會因資料分類掃描而改變，建議使用者在 CIFS 中具有寫入屬性權限或在 NFS 中具有寫入權限。如果可能，請將 Active Directory 使用者設定為組織中具有所有檔案權限的父群組的一部分。

   如果您正確輸入了憑證，則會出現一則訊息確認所有 CIFS 磁碟區均已成功驗證。

3. 在設定頁面上，選擇*設定*以查看每個 CIFS 和 NFS 磁碟區的狀態並修正任何錯誤。

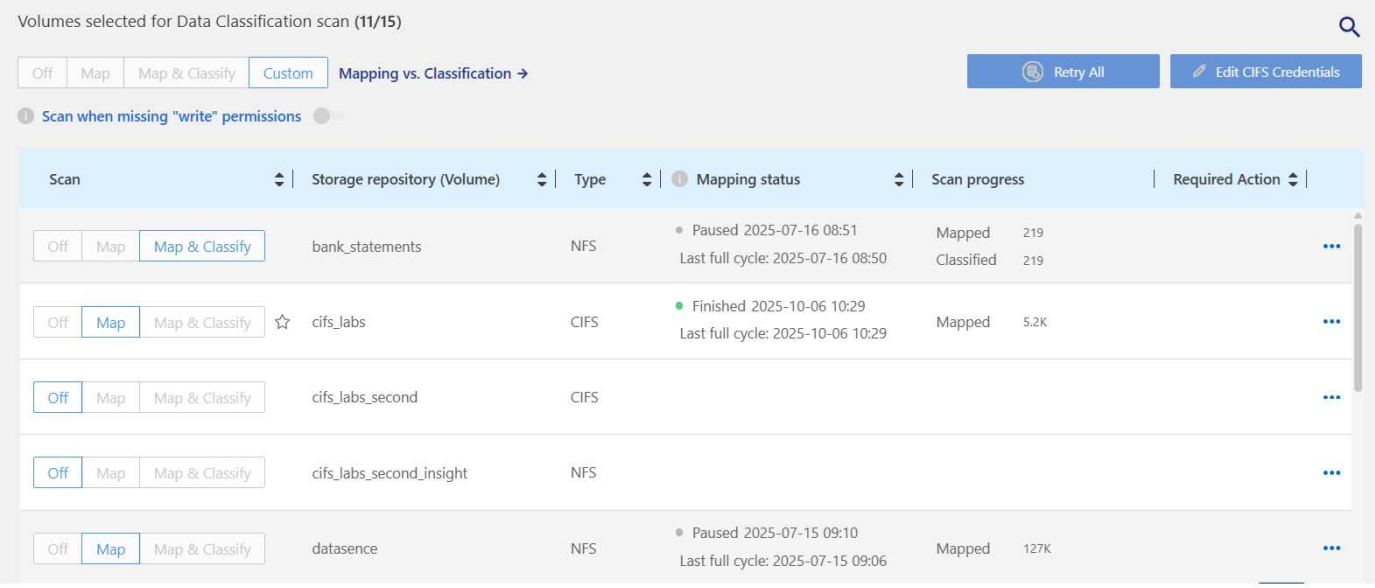
啟用或停用磁碟區掃描

您可以隨時從設定頁面啟動或停止任何系統上的掃描。您也可以將掃描從僅映射掃描切換到映射和分類掃描，反之亦然。建議您掃描系統中的所有磁碟區。

只有當您在標題區域中選擇了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增至系統的新磁碟區。當在標題區域設定為*自訂*或*關閉*時，您需要在系統中新增的每個新磁碟區上啟動對應和/或完整掃描。

頁面頂部的「缺少「寫入」權限時掃描」開關預設為停用狀態。這意味著，如果資料分類在 CIFS 中沒有寫入屬性權限或在 NFS 中沒有寫入權限，系統將不會掃描文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。如果您不介意是否重設上次造訪時間，請開啟開關，無論權限為何，都會掃描所有檔案。"了解更多"。

只有當您在標題區域中設定了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增到系統的新磁碟區。當所有磁碟區的設定都是「自訂」或「關閉」時，您需要為新增的每個新磁碟區手動啟動掃描。



步驟

1. 從資料分類選單中，選擇*配置*。
2. 選擇一個系統，然後選擇*配置*。
3. 若要啟用或停用所有磁碟區的掃描，請在所有磁碟區上方的標題中選擇對應、對應和分類或關閉。

   若要啟用或停用對單一卷的掃描，請在清單中找到該卷，然後選擇卷名稱旁的映射、映射和分類或關閉。

結果

當您啟用掃描時，資料分類將開始掃描您在系統中選擇的磁碟區。一旦資料分類開始掃描，結果就會開始出現在合規性儀表板中。掃描完成時間取決於資料量，從幾分鐘到幾小時不等。

資料分類僅掃描卷下的一個檔案共用。如果您的磁碟區中有多個共享，則需要將這些其他共享作為共享群組單獨掃描。"查看有關此數據分類限制的更多詳細信息" 。

使用NetApp Data Classification

完成幾個步驟即可開始使用NetApp Data Classification掃描資料庫模式。

審查先決條件

在啟用資料分類之前，請查看以下先決條件，以確保您具有受支援的配置。

支援的資料庫

資料分類可以掃描以下資料庫中的模式：

- 亞馬遜關係型資料庫服務 (Amazon RDS)

- MongoDB
- MySQL
- 甲骨文
- PostgreSQL
- SAP HANA
- SQL 伺服器 (MSSQL)

資料庫中必須啟用統計資訊收集功能。

資料庫要求

任何與資料分類實例連接的資料庫都可以掃描，無論它託管在何處。您只需要以下資訊即可連接到資料庫：

- IP 位址或主機名
- 港口
- 服務名稱（僅用於存取 Oracle 資料庫）
- 允許讀取架構的憑證

  選擇使用者名稱和密碼時，請務必選擇對要掃描的所有模式和表格具有完全讀取權限的使用者名稱和密碼。我們建議您為資料分類系統建立一個具有所有必需權限的專用使用者。

對於 MongoDB，需要只讀管理員角色。

部署資料分類實例

如果尚未部署實例，則部署資料分類。

如果您正在掃描可透過網際網路存取的資料庫模式，您可以"在雲端部署資料分類"或者"在可以存取網際網路的本地位置部署資料分類"。

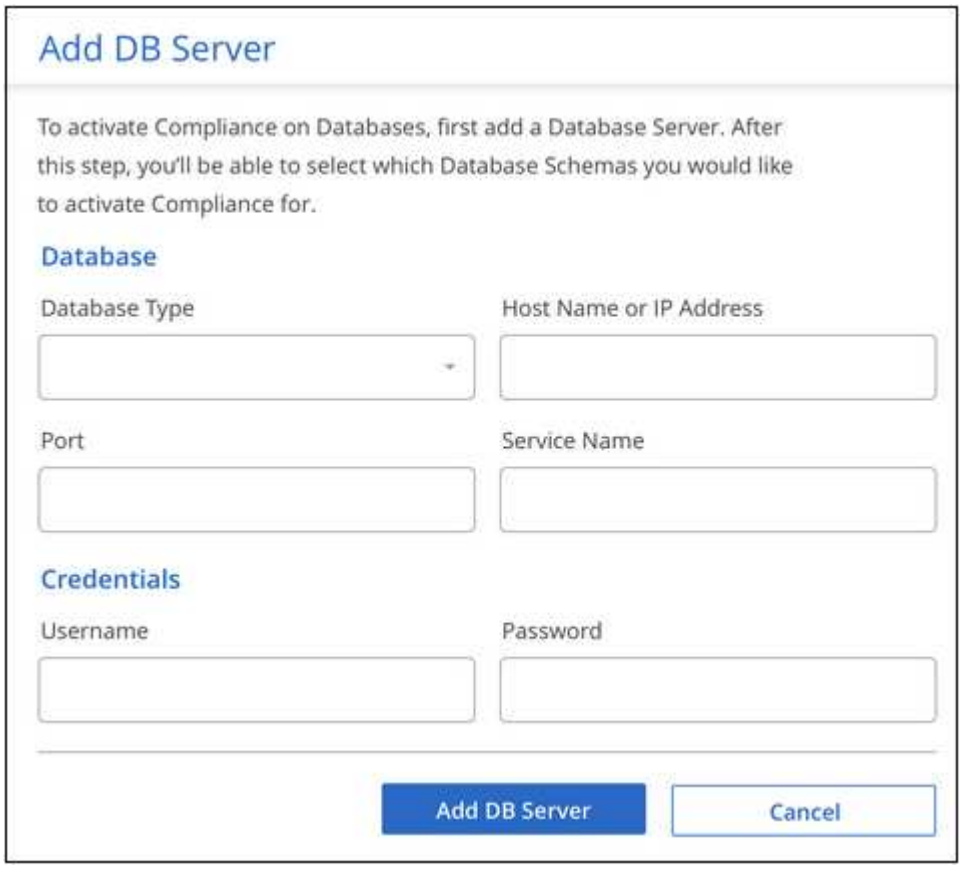
如果您正在掃描安裝在沒有網路存取的暗網中的資料庫模式，則需要"在沒有網路存取的相同本地位置部署資料分類"。這也要求將控制台代理部署在同一本地位置。

新增資料庫伺服器

新增架構所在的資料庫伺服器。

1. 從資料分類選單中，選擇*配置*。
2. 在設定頁面中，選擇*新增系統* > 新增資料庫伺服器。
3. 輸入所需資訊以識別資料庫伺服器。
   a. 選擇資料庫類型。
   b. 輸入連接埠和主機名稱或 IP 位址以連接到資料庫。
   c. 對於 Oracle 資料庫，輸入服務名稱。
   d. 輸入憑證以便資料分類可以存取伺服器。

e. 選擇*新增資料庫伺服器*。



該資料庫已新增至系統清單。

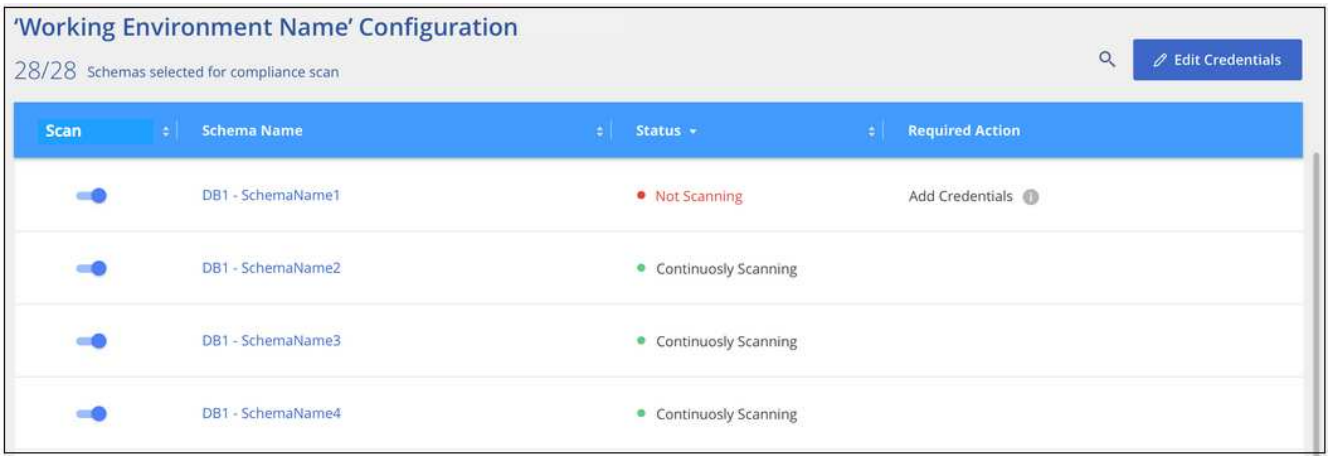
啟用和停用資料庫模式掃描

您可以隨時停止或開始對您的模式進行全面掃描。

沒有選項可以選擇僅映射資料庫模式的掃描。

1. 在配置頁面中，選擇要設定的資料庫的*配置*按鈕。

2. 透過向右移動滑桿來選擇要掃描的模式。



結果

資料分類開始掃描您啟用的資料庫模式。您可以透過導航到配置選單然後選擇系統配置來追蹤初始掃描的進度。每次掃描的進度都顯示為進度條。您也可以將滑鼠停留在進度列上，查看相對於磁碟區中檔案總數的已掃描檔案數。如果有任何錯誤，它們將出現在「狀態」列中，並顯示修復錯誤所需的操作。

資料分類每天掃描您的資料庫一次；資料庫不像其他資料來源那樣被連續掃描。

使用NetApp Data Classification掃描Google Cloud NetApp Volumes

NetApp Data Classification支援Google Cloud NetApp Volumes作為一個系統。了解如何掃描您的Google Cloud NetApp Volumes系統。

發現您要掃描的Google Cloud NetApp Volumes系統

如果您要掃描的Google Cloud NetApp Volumes系統尚未作為系統出現在NetApp Console中，"將其新增至系統頁面" 。

部署資料分類實例

"部署資料分類"如果尚未部署實例。

掃描Google Cloud NetApp Volumes時，資料分類必須部署在雲端中，並且必須部署在與您要掃描的磁碟區相同的區域。

*注意：*掃描Google Cloud NetApp Volumes時，目前不支援在本機位置部署資料分類。

在您的系統中啟用資料分類

您可以在Google Cloud NetApp Volumes系統上啟用資料分類。

1. 從資料分類選單中，選擇*配置*。
2. 選擇如何掃描每個系統中的磁碟區。"了解映射和分類掃描"：
   - 若要對應所有捲，請選擇*映射所有磁碟區*。
   - 若要對應和分類所有捲，請選擇*映射和分類所有捲*。
   - 若要自訂每個磁碟區的掃描，請選擇*或選擇每個磁碟區的掃描類型*，然後選擇要對應和/或分類的磁碟

區。

看啟用和停用磁碟區上的掃描了解詳情。

3. 在確認對話方塊中，選擇*批准*。

結果

資料分類開始掃描您在系統中選擇的磁碟區。一旦資料分類完成初始掃描，結果就會顯示在合規性儀表板中。所需時間取決於資料量：幾分鐘到幾個小時。您可以在配置選單的系統配置部分追蹤初始掃描的進度。資料分類顯示每次掃描的進度條。您也可以將滑鼠停留在進度列上，查看相對於磁碟區中總檔案數的已掃描檔案數。

- 預設情況下，如果資料分類在 CIFS 中沒有寫入屬性權限，或者在 NFS 中沒有寫入權限，系統將不會掃描磁碟區中的文件，因為資料分類無法將「上次存取時間」還原為原始時間戳記。如果您不介意上次造訪時間是否重設，請選擇*或為每個磁碟區選擇掃描類型*。結果頁面有一個您可以啟用的設置，以便資料分類可以掃描卷，而不管權限如何。
- 資料分類僅掃描卷下的一個檔案共用。如果您的磁碟區中有多個共享，則需要將這些其他共享作為共享群組單獨掃描。"了解此資料分類限制"。

驗證資料分類是否有權存取卷

透過檢查您的網路、安全群組和匯出策略，確保資料分類可以存取磁碟區。對於 CIFS 卷，您需要為資料分類提供 CIFS 憑證。

對於Google Cloud NetApp Volumes，資料分類只能掃描與控制台位於同一區域的磁碟區。

清單

- 確保資料分類實例與包含Google Cloud NetApp Volumes的每個網路之間存在網路連線。
- 確保以下連接埠對資料分類實例開放：
  - 對於 NFS – 連接埠 111 和 2049。
  - 對於 CIFS – 連接埠 139 和 445。
- 確保 NFS 磁碟區匯出策略包含資料分類執行個體的 IP 位址，以便它可以存取每個磁碟區上的資料。

步驟

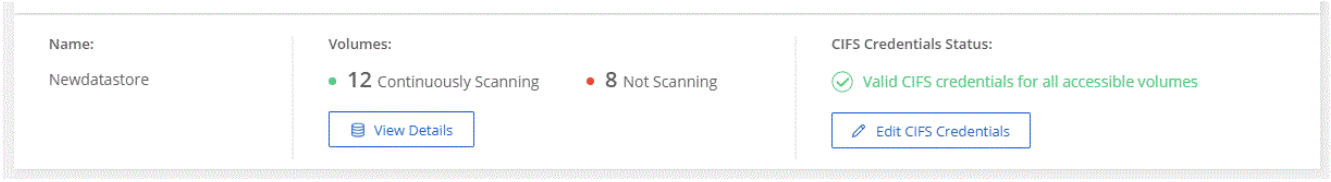
1. 從資料分類選單中，選擇*配置*。

   a. 如果您使用 CIFS（SMB），請確保 Active Directory 憑證正確。對於每個系統，選擇*編輯 CIFS 憑證*，然後輸入資料分類存取系統上的 CIFS 磁碟區所需的使用者名稱和密碼。

      憑證可以是唯讀的，但提供管理員憑證可確保資料分類可以讀取任何需要提升權限的資料。憑證儲存在資料分類實例上。

      如果您想要確保檔案的「上次存取時間」不會因資料分類掃描而改變，建議使用者在 CIFS 中具有寫入屬性權限或在 NFS 中具有寫入權限。如果可能，請將 Active Directory 使用者設定為組織中具有所有檔案權限的父群組的一部分。

      輸入憑證後，您應該會看到一條訊息，表示所有 CIFS 磁碟區均已成功驗證。



2. 在設定頁面上，選擇*查看詳細資訊*以查看每個 CIFS 和 NFS 磁碟區的狀態並修正任何錯誤。

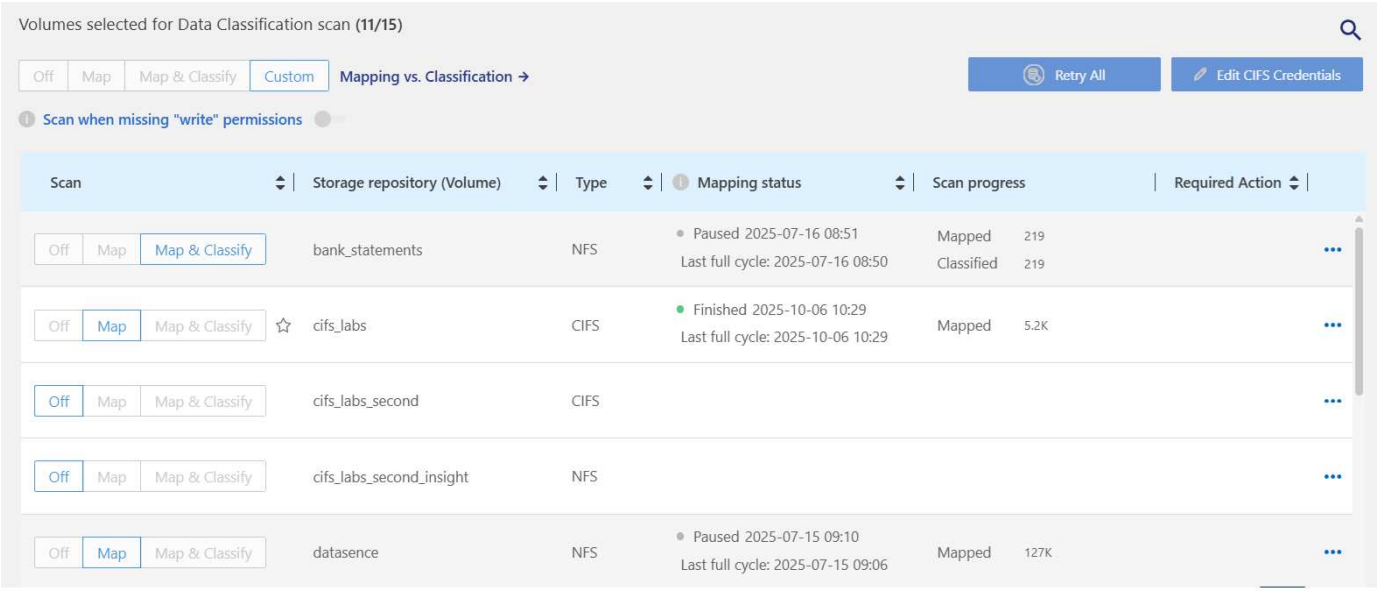
啟用和停用磁碟區上的掃描

您可以隨時從設定頁面啟動或停止任何系統上的掃描。您也可以將掃描從僅映射掃描切換到映射和分類掃描，反之亦然。建議您掃描系統中的所有磁碟區。

只有當您在標題區域中選擇了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增至系統的新磁碟區。當在標題區域設定為*自訂*或*關閉*時，您需要在系統中新增的每個新磁碟區上啟動對應和/或完整掃描。

頁面頂部的「缺少「寫入」權限時掃描」開關預設為停用狀態。這意味著，如果資料分類在 CIFS 中沒有寫入屬性權限或在 NFS 中沒有寫入權限，系統將不會掃描文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。如果您不介意是否重設上次造訪時間，請開啟開關，無論權限為何，都會掃描所有檔案。"了解更多"。

只有當您在標題區域中設定了 **Map** 或 **Map & Classify** 設定時，才會自動掃描新增到系統的新磁碟區。當所有磁碟區的設定都是「自訂」或「關閉」時，您需要為新增的每個新磁碟區手動啟動掃描。



步驟

1. 從資料分類選單中，選擇*配置*。
2. 選擇一個系統，然後選擇*配置*。
3. 若要啟用或停用所有磁碟區的掃描，請在所有磁碟區上方的標題中選擇對應、對應和分類或關閉。

   若要啟用或停用對單一卷的掃描，請在清單中找到該卷，然後選擇卷名稱旁的映射、映射和分類或關閉。

結果

當您啟用掃描時，資料分類將開始掃描您在系統中選擇的磁碟區。一旦資料分類開始掃描，結果就會開始出現在合規性儀表板中。掃描完成時間取決於資料量，從幾分鐘到幾小時不等。

使用NetApp Data Classification掃描檔案共用

若要掃描檔案共用，您必須先在NetApp Data Classification中建立檔案共用群組。檔案共用群組適用於本機或雲端中託管的 NFS 或 CIFS (SMB) 共用。

資料分類核心版本不支援掃描非NetApp檔案共享的資料。

先決條件

在啟用資料分類之前，請查看以下先決條件，以確保您具有受支援的配置。

- 共享可以託管在任何地方，包括雲端或本地端。可以將舊版NetApp 7-模式儲存系統中的 CIFS 共用掃描為檔案共用。
  - 資料分類無法從 7 模式系統中提取權限或「上次存取時間」。
  - 由於某些 Linux 版本和 7-模式系統上的 CIFS 共用之間存在已知問題，因此您必須將共用配置為僅使用啟用了 NTLM 驗證的 SMBv1。
- 資料分類實例和共用之間需要有網路連線。
- 您可以將 DFS（分散式檔案系統）共用新增為常規 CIFS 共用。由於資料分類不知道共用是基於組合為單一 CIFS 共用的多個伺服器/磁碟區所建構的，因此當訊息實際上僅適用於位於不同伺服器/磁碟區上的一個資料夾/共用時，您可能會收到有關共用的權限或連線錯誤。
- 對於 CIFS (SMB) 共用，請確保您擁有可提供共用讀取存取權限的 Active Directory 憑證。如果資料分類需要掃描任何需要提升權限的數據，則最好使用管理員憑證。

  如果您想要確保檔案的「上次存取時間」不會因資料分類掃描而改變，建議使用者在 CIFS 中具有寫入屬性權限或在 NFS 中具有寫入權限。如果可能，請將 Active Directory 使用者設定為組織中具有所有檔案權限的父群組的一部分。

- 群組中的所有 CIFS 檔案共用必須使用相同的 Active Directory 憑證。
- 您可以混合使用 NFS 和 CIFS（使用 Kerberos 或 NTLM）共用。您必須單獨將共用新增至群組。也就是說，您必須完成該過程兩次 - 每個協議一次。
  - 您無法建立混合 CIFS 驗證類型（Kerberos 和 NTLM）的檔案共用群組。
- 如果您使用具有 Kerberos 驗證的 CIFS，請確保資料分類可以存取所提供的 IP 位址。如果 IP 位址無法訪問，則無法新增檔案共用。

建立文件共享組

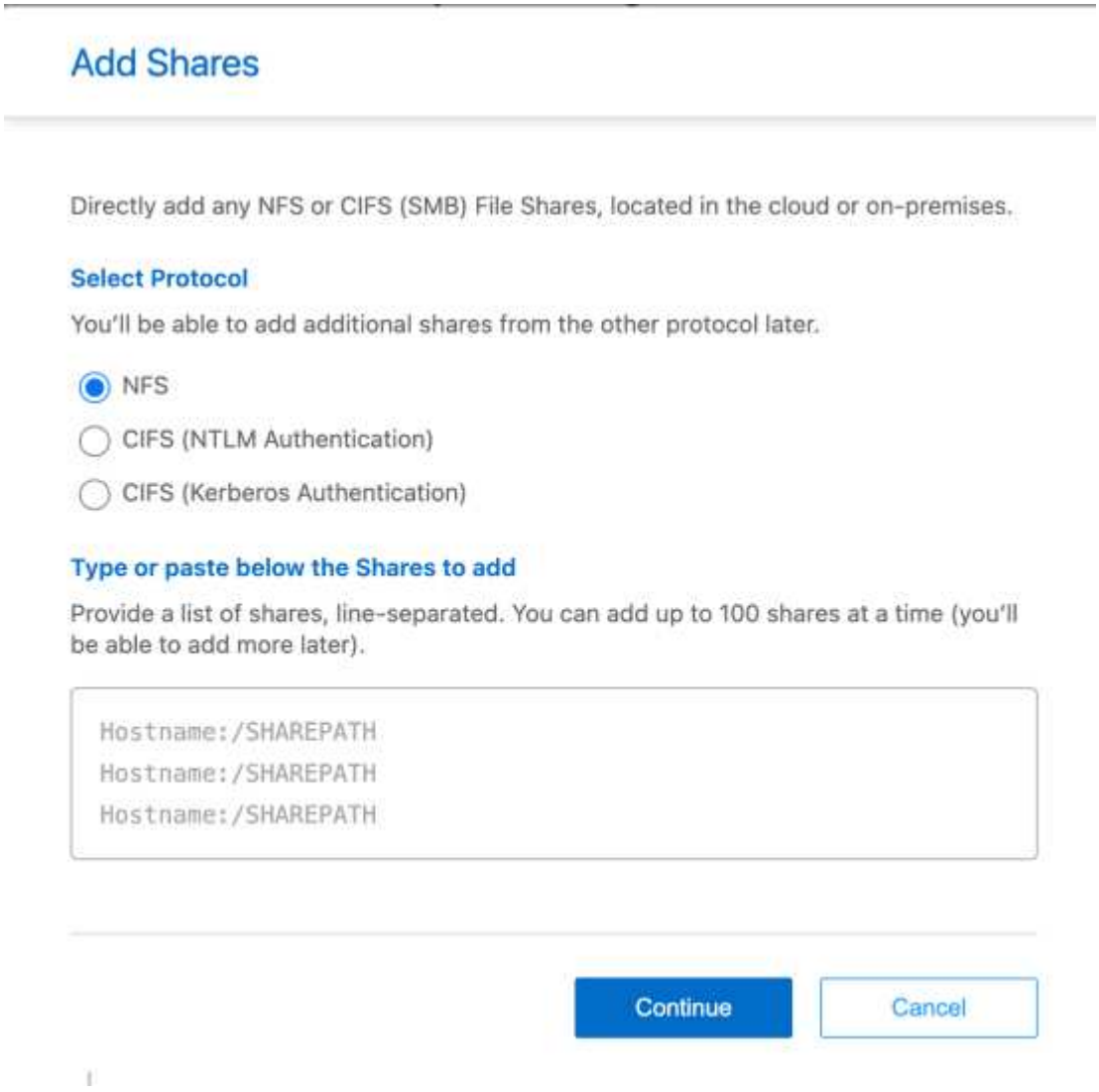
將檔案共用新增至群組時，必須使用格式 `<host_name>:/<share_path>`。

您可以單獨新增文件共享，也可以輸入要掃描的文件共享的行分隔清單。您一次最多可以添加 100 股。

步驟

1. 從資料分類選單中，選擇*配置*。

2. 從設定頁面中，選擇*新增系統*>*新增檔案共用群組*。
3. 在新增檔案共享群組對話方塊中，輸入共享群組的名稱，然後選擇*繼續*。
4. 選擇要新增的文件共享的協定。



   a. 如果您要新增具有 NTLM 驗證的 CIFS 共用，請輸入 Active Directory 憑證以存取 CIFS 磁碟區。儘管支援唯讀憑證，但建議您使用管理員憑證提供完全存取權限。選擇儲存。
5. 新增要掃描的檔案共用（每行一個檔案共用）。然後選擇繼續。
6. 確認對話方塊顯示已新增的共享數量。

   如果對話方塊列出了任何無法新增的共享，請擷取此資訊以便解決問題。如果問題與命名約定有關，您可以使用已修正的名稱重新新增共用。

7. 配置磁碟區上的掃描：
   - 若要對檔案共用啟用僅映射掃描，請選擇*映射*。
   - 若要對檔案共用啟用完整掃描，請選擇*Map & Classify*。
   - 若要停用檔案共用上的掃描，請選擇「關閉」。

頁面頂部的「缺少「寫入屬性」權限時掃描」開關預設為停用狀態。這意味著，如果資料分類在 CIFS 中沒有寫入屬性權限或在 NFS 中沒有寫入權限，系統將不會掃描文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。+ 如果將*缺少「寫入屬性」權限時掃描*切換為*開*，則掃描將重置上次存取時間並掃描所有文件，而不管權限如何。+ 要了解有關上次訪問時間戳的更多信息，請參閱"從資料分類中的資料來源收集的元資料"。

結果

資料分類開始掃描您新增的檔案共用中的檔案。你可以追蹤掃描進度並在儀表板中查看掃描結果。

如果對於使用 Kerberos 驗證的 CIFS 配置的掃描未成功完成，請檢查配置選項卡中是否有錯誤。

編輯檔案共享群組

建立檔案共用群組後，您可以編輯 CIFS 協定或新增和刪除檔案共用。

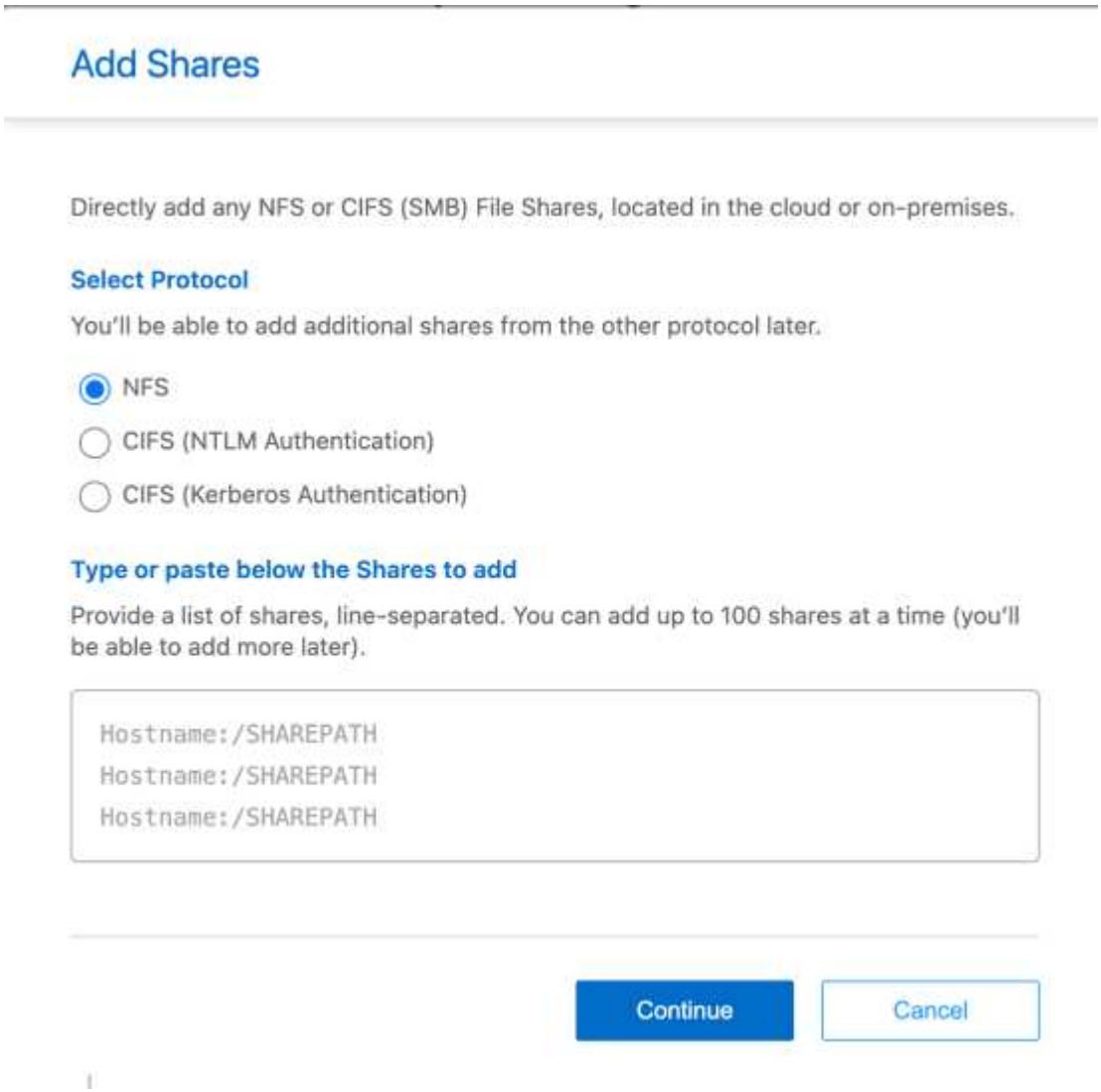
編輯 **CIFS** 協定配置

1. 從資料分類選單中，選擇*配置*。
2. 從設定頁面中，選擇要修改的檔案共用群組。
3. 選擇編輯 **CIFS** 憑證。

4. 選擇身份驗證方法：**NTLM** 或 **Kerberos**。
5. 輸入 Active Directory 使用者名稱和密碼。
6. 選擇儲存以完成該過程。

將文件共享新增至掃描

1. 從資料分類選單中，選擇*配置*。
2. 從設定頁面中，選擇要修改的檔案共用群組。
3. 選擇 **+** 新增共享。
4. 選擇要新增的文件共享的協定。



如果您要將文件共用新增至已設定的協議，則無需進行任何變更。

如果您要使用第二種協定新增檔案共用，請確保您已正確配置身份驗證，詳情請見"先決條件"。

5. 使用以下格式新增要掃描的檔案共用（每行一個檔案共用） `<host_name>:/<share_path>` 。
6. 選擇繼續以完成新增檔案共用。

從掃描中刪除檔案共享

1. 從資料分類選單中，選擇*配置*。
2. 選擇要從中刪除檔案共享的系統。
3. 選擇*配置*。
4. 在配置頁面中，選擇操作 ... 對於要刪除的檔案共用。
5. 從操作選單中，選擇*刪除共用*。

追蹤掃描進度

您可以追蹤初始掃描的進度。

1. 選擇配置選單。
2. 選擇系統配置。
3. 對於儲存庫，檢查掃描進度列以查看其狀態。

使用NetApp Data Classification掃描StorageGRID數據

完成幾個步驟即可直接使用NetApp Data Classification開始掃描StorageGRID內的資料。

查看StorageGRID要求

在啟用資料分類之前，請查看以下先決條件，以確保您具有受支援的配置。

- 您需要有端點 URL 才能連接物件儲存服務。
- 您需要擁有來自StorageGRID的存取金鑰和金鑰，以便資料分類可以存取儲存桶。

部署資料分類實例

如果尚未部署實例，則部署資料分類。

如果您正在掃描可透過網際網路存取的StorageGRID數據，您可以"在雲端部署資料分類"或者"在可以存取網際網路的本地位置部署資料分類"。

如果您要掃描安裝在沒有網路存取的暗站中的StorageGRID數據，則需要"在沒有網路存取的相同本地位置部署資料分類"。這也要求將控制台代理部署在同一本地位置。

將StorageGRID服務新增至資料分類

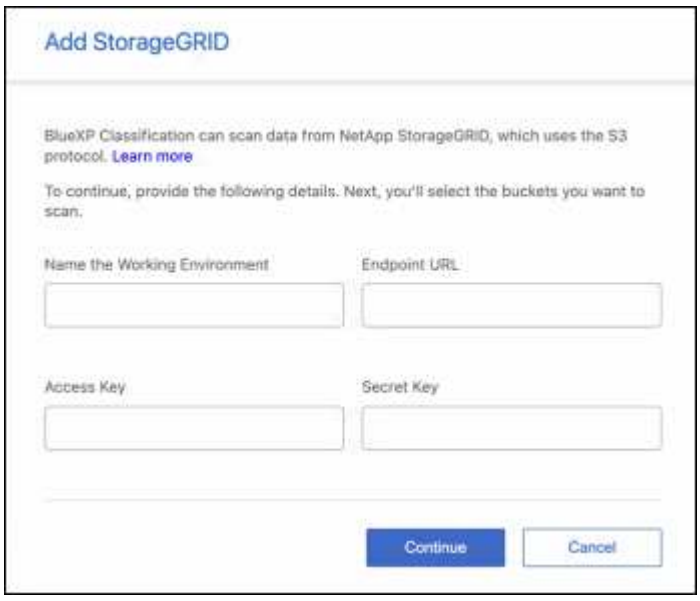
新增StorageGRID服務。

步驟

1. 從資料分類選單中，選擇*配置*選項。
2. 在設定頁面中，選擇「新增系統」>「新增StorageGRID」。
3. 在新增StorageGRID服務對話方塊中，輸入StorageGRID服務的詳細資訊並選擇*繼續*。
    a. 輸入您想要使用的系統名稱。此名稱應反映您要連線的StorageGRID服務的名稱。

b. 輸入 Endpoint URL 以存取物件儲存服務。

c. 輸入存取密鑰和密鑰，以便資料分類可以存取StorageGRID中的儲存桶。



結果

StorageGRID已新增至系統清單。

啟用並停用**StorageGRID**桶上的掃描

在StorageGRID上啟用資料分類後，下一步是設定要掃描的儲存桶。資料分類發現這些儲存桶並將它們顯示在您建立的系統中。

步驟

1. 在設定頁面中，找到StorageGRID系統。
2. 在StorageGRID系統圖塊上，選擇 配置。
3. 完成以下步驟之一來啟用或停用掃描：
   - 若要對儲存體桶啟用僅對應掃描，請選擇*Map*。
   - 若要對儲存桶啟用完整掃描，請選擇*Map & Classify*。
   - 若要停用儲存桶的掃描，請選擇「關閉」。

結果

資料分類開始掃描您啟用的儲存桶。您可以透過導航到配置選單然後選擇系統配置來追蹤初始掃描的進度。每次掃描的進度都顯示為進度條。您也可以將滑鼠停留在進度列上，查看相對於磁碟區中總檔案數的已掃描檔案數。如果有任何錯誤，它們將出現在「狀態」列中，並顯示修復錯誤所需的操作。

將您的 **Active Directory** 與**NetApp Data Classification**集成

您可以將全域 Active Directory 與NetApp Data Classification集成，以增強資料分類報告有關文件擁有者以及哪些使用者和群組有權存取您的文件的結果。

當您設定某些資料來源（如下所列）時，您需要輸入 Active Directory 憑證以便資料分類掃描 CIFS 磁碟區。此集成為資料分類提供了駐留在這些資料來源中的資料的檔案擁有者和權限詳細資訊。為這些資料來源輸入的 Active Directory 可能與您在此輸入的全域 Active Directory 憑證不同。資料分類將在所有整合的活動目錄中尋找使用者和權限詳細資訊。

此整合在資料分類的以下位置提供了附加資訊：

- 您可以使用“文件所有者”"篩選"並在調查窗格中的文件元資料中查看結果。它不是包含 SID（安全性識別碼）的檔案擁有者，而是填入實際的使用者名稱。

  您也可以查看更多關於檔案擁有者的詳細資訊：帳戶名稱、電子郵件地址和 SAM 帳戶名稱，或查看該使用者擁有的項目。

- 你可以看到"完整檔案權限"當您按一下「查看所有權限」按鈕時，每個檔案和目錄都會顯示所有權限。
- 在"治理儀表板"，開啟權限面板將顯示有關您的資料的更詳細的詳細資訊。

本機使用者 SID 和來自未知網域的 SID 不會轉換為實際使用者名稱。

支援的資料來源

Active Directory 與資料分類的整合可以識別來自下列資料來源的資料：

- 本地ONTAP系統
- Cloud Volumes ONTAP
- Azure NetApp Files
- 適用於ONTAP的 FSx

連接到您的 **Active Directory** 伺服器

部署資料分類並啟動資料來源掃描後，您可以將資料分類與 Active Directory 整合。可以使用 DNS 伺服器 IP 位址或 LDAP 伺服器 IP 位址存取 Active Directory。

Active Directory 憑證可以是唯讀的，但提供管理員憑證可確保資料分類可以讀取任何需要提升權限的資料。憑證儲存在資料分類實例上。

對於 CIFS 磁碟區/檔案共用，如果您想確保檔案的「上次存取時間」不會因資料分類掃描而改變，則使用者應該具有寫入屬性權限。如果可能的話，我們建議將 Active Directory 設定的使用者作為組織中對所有檔案具有權限的父群組的一部分。

要求

- 您必須已經為公司使用者設定了 Active Directory。
- 您必須具有 Active Directory 的資訊：

  - DNS 伺服器 IP 位址，或多個 IP 位址

    或者

  LDAP 伺服器 IP 位址，或多個 IP 位址

- 存取伺服器的使用者名稱和密碼
- 網域名稱（Active Directory 名稱）
- 您是否使用安全 LDAP (LDAPS)
- LDAP 伺服器連接埠（LDAP 通常為 389，安全 LDAP 通常為 636）

- 必須開啟以下連接埠以供資料分類實例進行出站通訊：

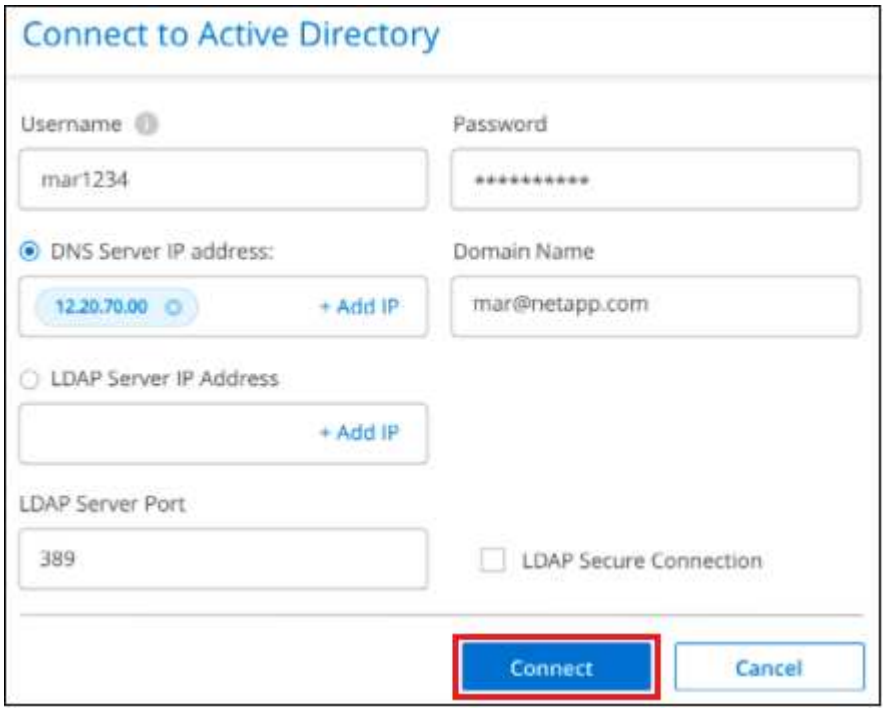
協定	港口	目的地	目的
TCP 和 UDP	389	活動目錄	LDAP
TCP	636	活動目錄	透過 SSL 的 LDAP
TCP	3268	活動目錄	全域目錄
TCP	3269	活動目錄	透過 SSL 的全域目錄

步驟

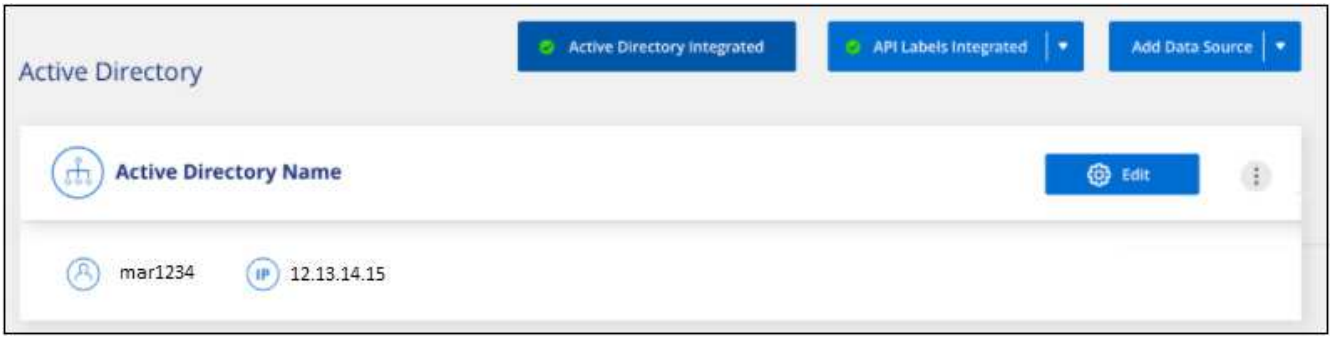
1. 在資料分類設定頁面中，按一下*新增 Active Directory*。

2. 在「連接到 Active Directory」對話方塊中，輸入 Active Directory 詳細訊息，然後按一下「連線」。

   如果需要，您可以透過選擇「新增 IP」來新增多個 IP 位址。



   資料分類整合到 Active Directory，並在設定頁面中新增一個新部分。



管理您的 **Active Directory** 集成

如果您需要修改 Active Directory 整合中的任何值，請按一下「編輯」按鈕並進行變更。

您也可以選擇 按鈕，然後*刪除 Active Directory*。

使用資料分類

使用NetApp Data Classification查看組織中儲存的資料的治理詳細信息

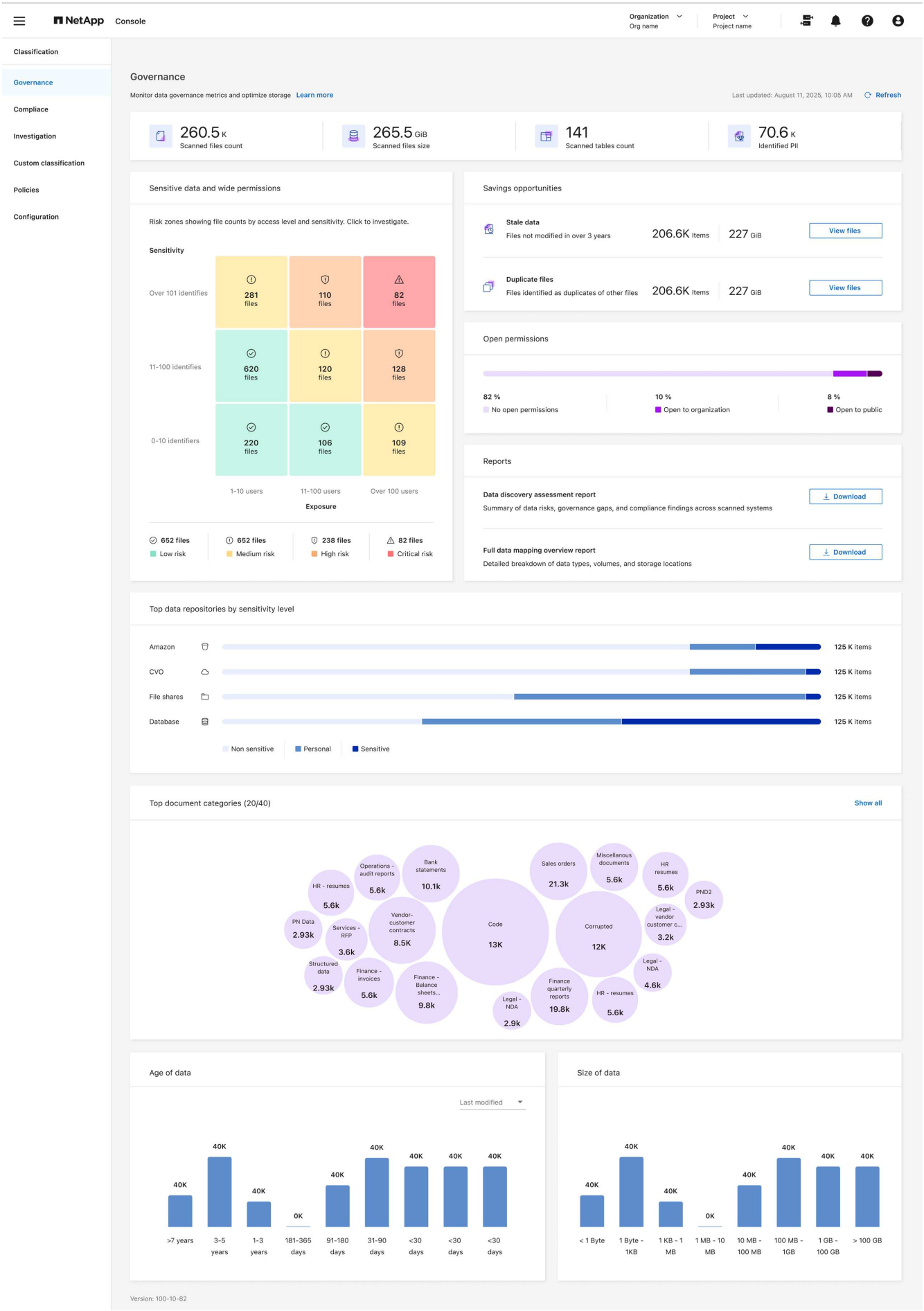
控制與組織儲存資源上的資料相關的成本。 NetApp Data Classification可識別系統中陳舊資料、重複檔案和超大檔案的數量，以便您可以決定是否要刪除某些檔案或將某些檔案分層到成本較低的物件儲存中。

您應該從這裡開始您的研究。從治理儀表板中，您可以選擇一個區域進行進一步調查。

此外，如果您打算將資料從本地位置遷移到雲端，則可以在移動資料之前查看資料的大小以及其中是否有任何資料包含敏感資訊。

查看治理儀表板

治理儀表板提供信息，以便您可以提高效率並控制與儲存在儲存資源上的資料相關的成本。


NetApp Console
Organization
Org name
Project
Project name
Classification
Governance
Compliace
Investigation
Custom classification
Policies
Configuration
Governance
Monitor data governance metrics and optimize storage Learn more
Last updated: August 11, 2025, 10:05 AM
Refresh
260.5 K
Scanned files count
265.5 GiB
Scanned files size
141
Scanned tables count
70.6 K
Identified PII
Sensitive data and wide permissions
Risk zones showing file counts by access level and sensitivity. Click to investigate.
Sensitivity
Over 101 identifies
11-100 identifies
0-10 identifiers
281 files
110 files
82 files
620 files
120 files
128 files
220 files
106 files
109 files
1-10 users
11-100 users
Over 100 users
Exposure
652 files
Low risk
652 files
Medium risk
238 files
High risk
82 files
Critical risk
Savings opportunities
Stale data
Files not modified in over 3 years
206.6K Items
227 GiB
View files
Duplicate files
Files identified as duplicates of other files
206.6K Items
227 GiB
View files
Open permissions
82 %
No open permissions
10 %
Open to organization
8 %
Open to public
Reports
Data discovery assessment report
Summary of data risks, governance gaps, and compliance findings across scanned systems
Download
Full data mapping overview report
Detailed breakdown of data types, volumes, and storage locations
Download
Top data repositories by sensitivity level
Amazon
CVO
File shares
Database
125 K items
Non sensitive
Personal
Sensitive
Top document categories (20/40)
Show all
Operations - audit reports 5.6k
Bank statements 10.1k
Sales orders 21.3k
Miscellanous documents 5.6k
HR resumes 5.6k
PND2 2.93k
HR - resumes 5.6k
PN Data 2.93k
Services - RFP 3.6k
Vendor-customer contracts 8.5K
Code 13K
Corrupted 12K
Legal - vendor customer c... 3.2k
Structured data 2.93k
Finance - invoices 5.6k
Finance - Balance sheets... 9.8k
Legal - NDA 2.9k
Finance quarterly reports 19.8k
HR - resumes 5.6k
Legal - NDA 4.6k
Age of data
Last modified
40K
>7 years
40K
3-5 years
40K
1-3 years
0K
181-365 days
40K
91-180 days
40K
31-90 days
40K
<30 days
40K
<30 days
40K
<30 days
Size of data
40K
< 1 Byte
40K
1 Byte - 1KB
40K
1 KB - 1 MB
0K
1 MB - 10 MB
40K
10 MB - 100 MB
40K
100 MB - 1GB
40K
1 GB - 100 GB
40K
> 100 GB
Version: 100-10-82

步驟

1. 從NetApp Console選單中，選擇 治理 > 分類。
2. 選擇*治理*。

   出現治理儀表板。

檢討節省機會

_節省機會_元件顯示您可以刪除或分層到較便宜的物件儲存的資料。 《節省機會》中的數據每 2 小時更新一次。您也可以手動更新資料。

步驟

1. 從資料分類選單中，選擇*治理*。
2. 在治理儀表板的每個節省機會圖塊中，選擇*最佳化儲存*以在調查頁面中查看過濾的結果。要發現您應該刪除或分層到較便宜的存儲的任何數據，請調查_節省機會_。
   - 過期資料 - 預設情況下，如果資料上次修改超過 3 年，則該資料視為過期資料。您可以[自訂過期資料的定義](task-stale-data.html)。
   - 重複檔案 - 您正在掃描的資料來源中其他位置重複的檔案。"查看顯示的重複檔案類型" 。

如果您的任何資料來源實現了資料分層，則可以在「陳舊資料」類別中識別已經駐留在物件儲存中的舊資料。

建立數據發現評估報告

數據發現評估報告對掃描環境進行了高級分析，以顯示關注區域和潛在的補救步驟。結果基於數據的映射和分類。本報告的目標是提高您對資料集三個重要面向的認識：

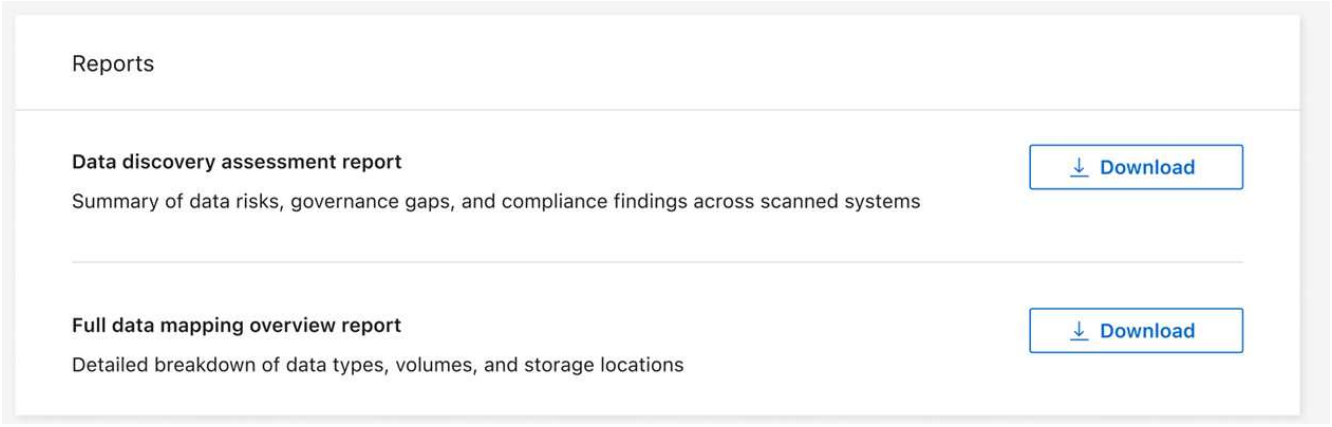
特徵	描述
資料治理問題	您擁有的所有數據以及可以減少數據量以節省成本的區域的詳細圖片。
資料安全風險	由於存取權限廣泛，您的資料可能受到內部或外部攻擊的區域。
數據合規性差距	您的個人或敏感個人資訊位於何處，以滿足安全和 DSAR（資料主體存取請求）。

透過該報告，您可以採取以下行動：

- 透過更改保留策略或移動或刪除某些資料（陳舊或重複的資料）來降低儲存成本。
- 透過修改全域群組管理策略來保護具有廣泛權限的資料。
- 透過將 PII 移至更安全的資料儲存來保護包含個人或敏感個人資訊的資料。

步驟

1. 從資料分類中，選擇*治理*。
2. 在報告圖塊中，選擇「資料發現評估報告」。



結果

資料分類會產生一份您可以檢視和分享的 PDF 報告。

建立資料映射概覽報告

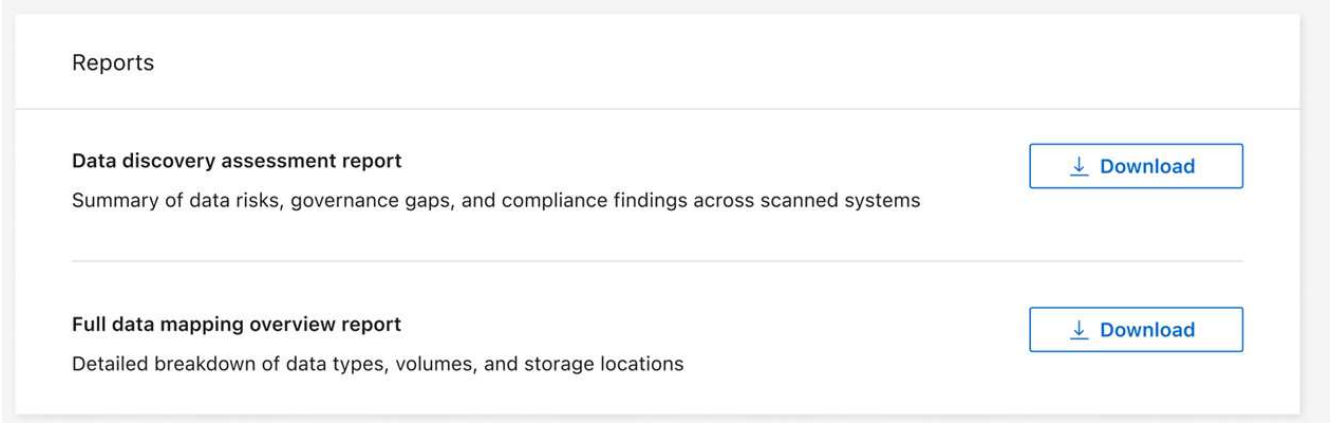
資料映射概覽報告提供了儲存在公司資料來源中的資料的概覽，以協助您做出遷移、備份、安全性和合規性流程的決策。該報告總結了所有系統和資料來源。它還為每個系統提供了分析。

該報告包含以下資訊：

類別	描述
使用容量	對於所有系統：列出每個系統的檔案數量和已使用容量。對於單一系統：列出使用最多容量的檔案。
數據時代	提供三個圖表和圖形，分別表示檔案的建立時間、上次修改時間或上次造訪時間。根據特定日期範圍列出文件數量及其已使用容量。
數據大小	列出系統中存在於特定大小範圍內的檔案數。

步驟

1. 從資料分類中，選擇*治理*。
2. 在報告圖塊中，選擇*完整資料映射概覽報告*。



結果

資料分類會產生一份 PDF 報告，您可以根據需要查看並傳送給其他群組。

如果報告大於 1 MB，則 PDF 檔案將保留在資料分類實例上，您將看到有關確切位置的彈出訊息。當資料分類安裝在您本機的 Linux 機器上或在雲端部署的 Linux 機器上時，您可以直接導覽至 PDF 檔案。當資料分類部署在雲端時，您需要使用 SSH 授權資料分類實例下載 PDF 檔案。

查看按資料敏感度列出的頂級資料儲存庫

資料映射概覽報告中的「按敏感度等級排列的頂級資料儲存庫」區域列出了包含最敏感項目的前四個資料儲存庫（系統和資料來源）。每個系統的長條圖分為：

- 非敏感數據
- 個人資料
- 敏感個人數據

數據每兩小時刷新一次，可以手動刷新。

步驟

1. 若要查看每個類別中的項目總數，請將遊標放在欄的每個部分上。
2. 若要過濾調查頁面中顯示的結果，請選擇欄中的每個區域並進一步調查。

審查敏感數據和廣泛的權限

治理儀表板的「敏感資料和廣泛權限」區域顯示包含敏感資料和具有廣泛權限的檔案的數量。此表顯示以下類型的權限：

- 從橫軸上最嚴格的權限到最寬鬆的限制。
- 縱軸上從最不敏感的資料到最敏感的資料。

步驟

1. 若要查看每個類別中的檔案總數，請將遊標放在每個方塊上。
2. 若要過濾調查頁面中顯示的結果，請選擇一個方塊並進一步調查。

查看按開放權限類型列出的數據

資料映射概覽報表的「開啟權限」區域顯示正在掃描的所有檔案中每種權限的百分比。此圖表顯示以下類型的權限：

- 無開放權限
- 向組織開放
- 對外開放
- 未知訪問

步驟

1. 若要查看每個類別中的檔案總數，請將遊標放在每個方塊上。
2. 若要過濾調查頁面中顯示的結果，請選擇一個方塊並進一步調查。

審查資料的年齡和大小

您可以調查資料對應概覽報表的「*Age*」和「*Size*」圖表中的項目，看看是否有任何資料應該刪除或分層到較便宜的物件儲存。

步驟

1. 在資料年齡圖表中，要查看有關資料年齡的詳細信息，請將遊標放在圖表中的某個點上。
2. 若要依年齡或尺寸範圍進行過濾，請選擇該年齡或尺寸。
   - 資料年齡圖 - 根據資料建立時間、上次造訪時間或上次修改時間對資料進行分類。
   - 資料大小圖 - 根據大小將資料分類。

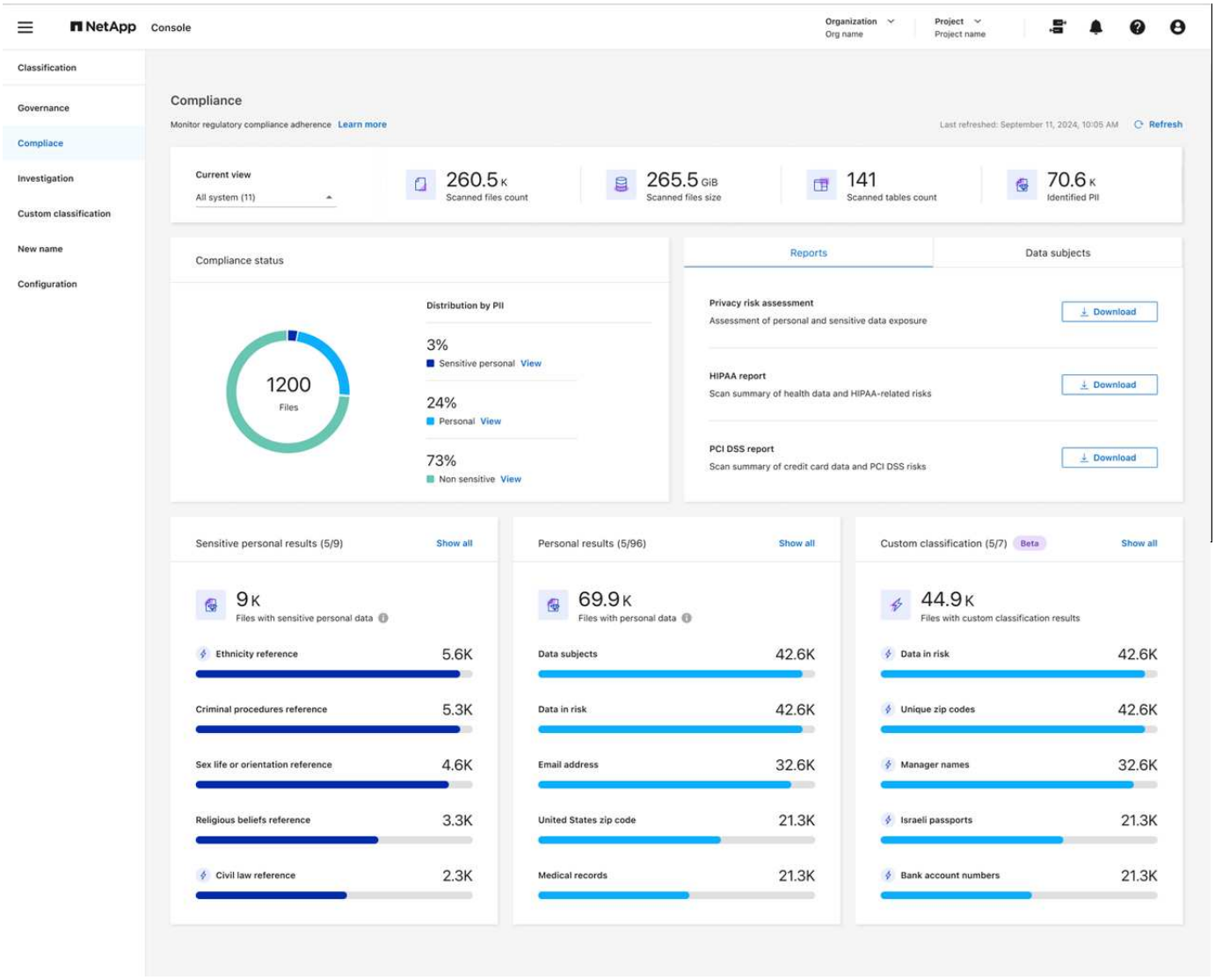
如果您的任何資料來源實現了資料分層，則已駐留在物件儲存中的舊資料可能會在「資料年齡」圖中被識別。

使用NetApp Data Classification查看組織中儲存的私人資料的合規性詳細信息

透過查看組織中的個人資料 (PII) 和敏感個人資料 (SPII) 的詳細資訊來控制您的私人資料。您也可以透過查看NetApp Data Classification在您的資料中找到的類別和檔案類型來獲得可見性。

只有當您執行完整分類掃描時，才可獲得文件級合規性詳細資訊。僅映射掃描不會產生文件級詳細資訊。

預設情況下，資料分類儀表板顯示所有系統和資料庫的合規性資料。若要僅查看部分系統的數據，請選擇它們。



您可以從資料調查頁面過濾結果，並將結果報告下載為 CSV 檔案。看"在資料調查頁面中過濾數據"了解詳情。

查看包含個人資料的文件

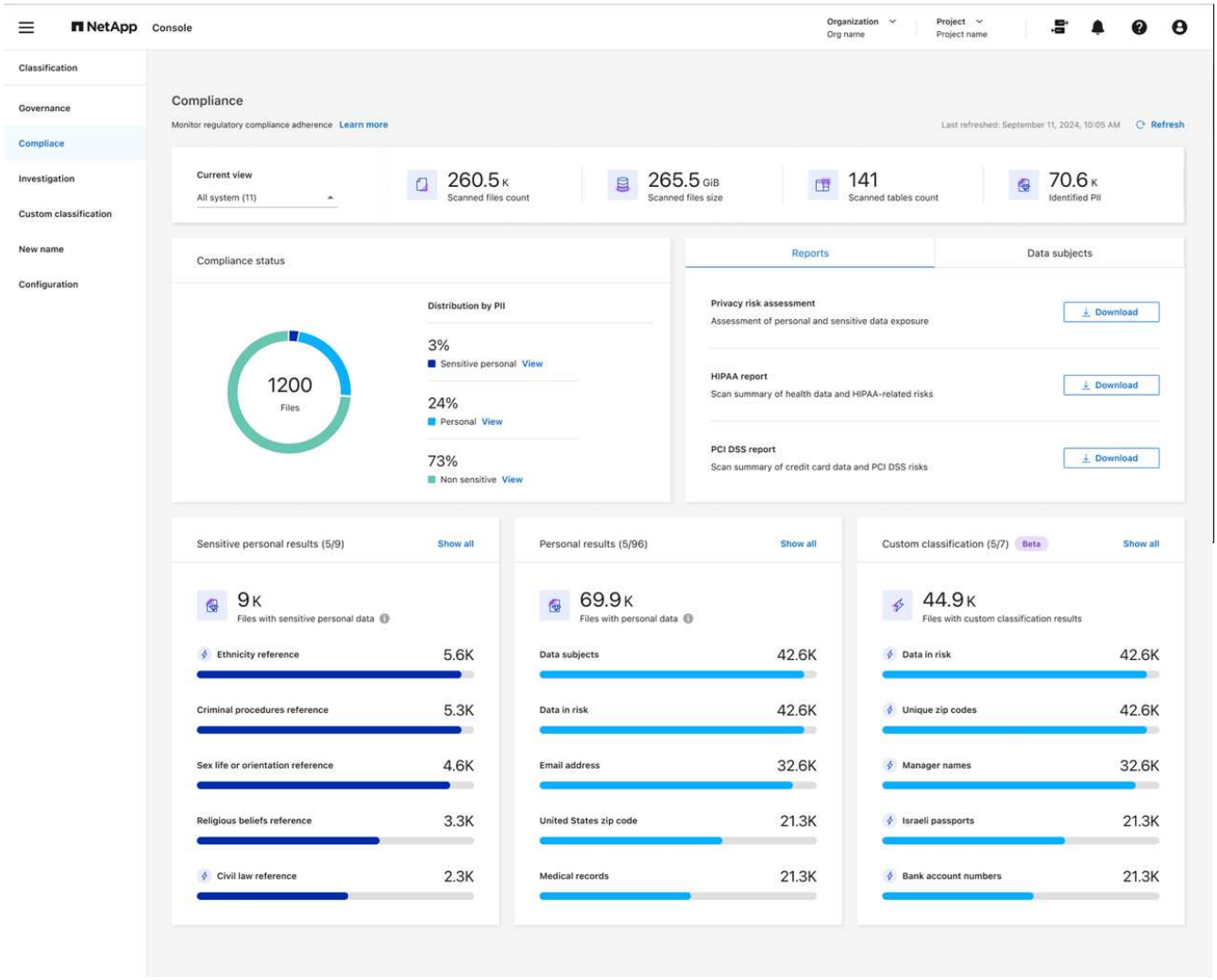
資料分類會自動識別資料中的特定單字、字串和模式（正規表示式）。"例如，信用卡號、社會安全號碼、銀行帳號、密碼等等。"資料分類可在單一檔案、目錄（共用和資料夾）內的檔案以及資料庫表中識別此類資訊。

您也可以建立自訂搜尋字詞來識別特定於您組織的個人資料。有關更多信息，請參閱"建立自訂分類" 。

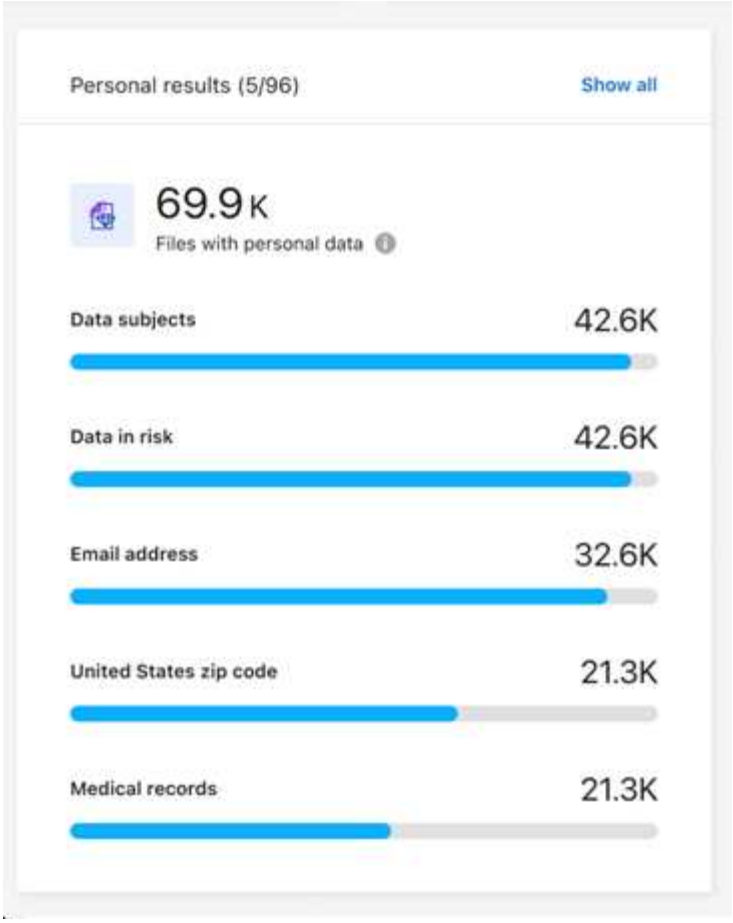
對於某些類型的個人數據，資料分類使用_鄰近驗證_來驗證其發現。透過尋找與找到的個人資料接近的一個或多個預定義關鍵字來進行驗證。例如，如果資料分類看到旁邊有一個近似詞（例如，*SSN* 或 *social security*），它就會將美國社會安全號碼 (SSN) 識別為 SSN。"個人資料表"顯示資料分類何時使用鄰近驗證。

步驟

1. 從資料分類選單中，選擇“合規性”標籤。
2. 要調查所有個人資料的詳細信息，請選擇個人資料百分比旁邊的圖示。

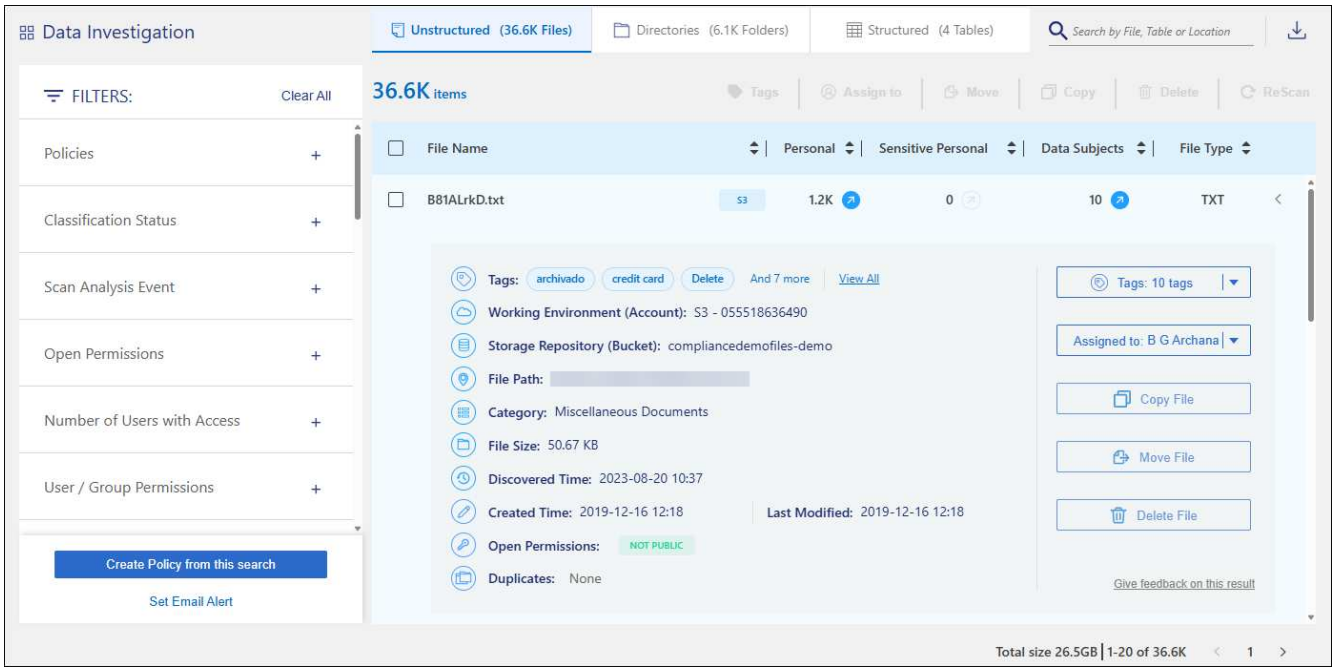


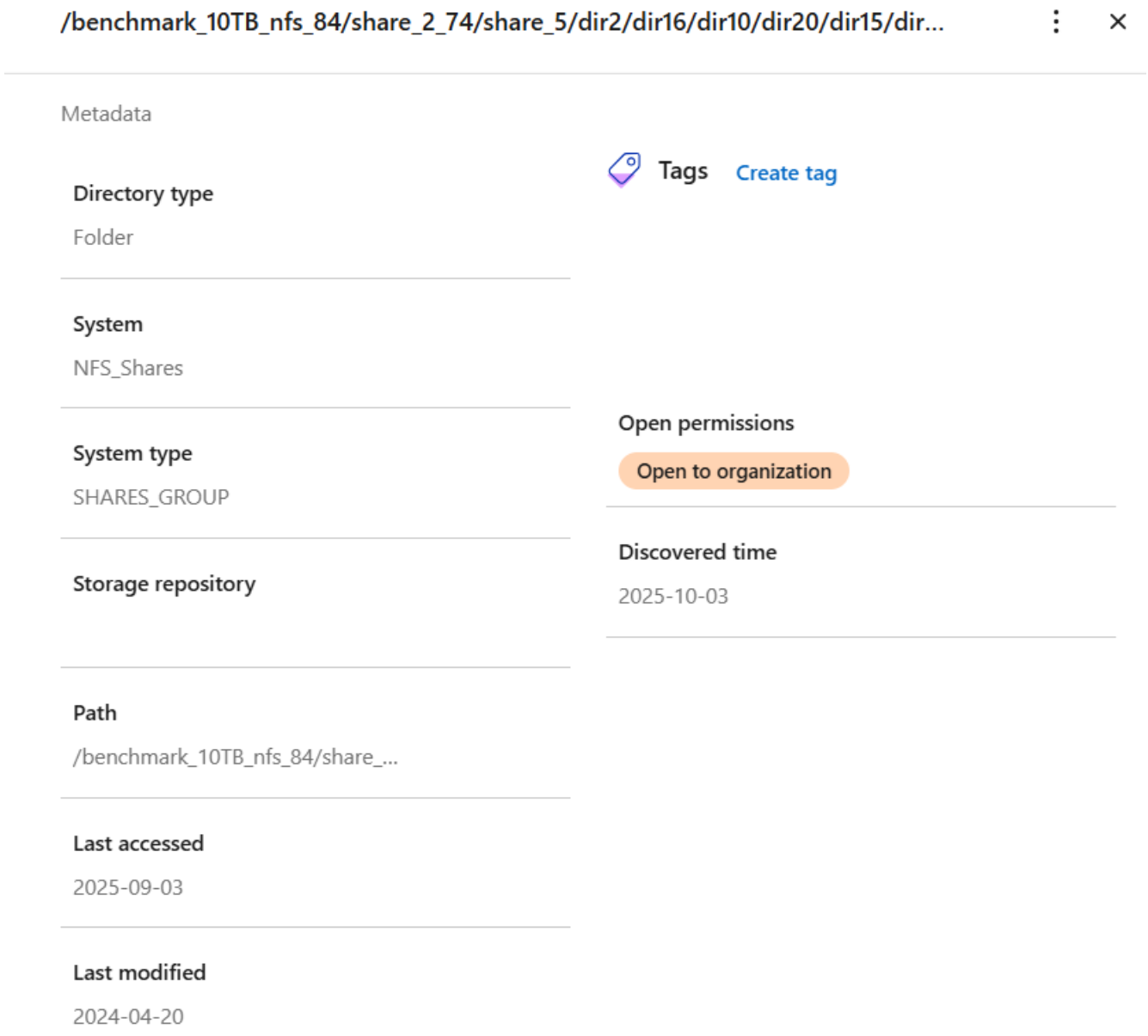
3. 要調查特定類型的個人資料的詳細信息，請選擇*查看全部*，然後選擇特定類型的個人資料（例如電子郵件地址）的*調查結果*箭頭圖示。



4. 透過搜尋、排序、擴展特定文件的詳細資訊、選擇「調查結果」箭頭查看屏蔽資訊或下載文件列表來調查資料。

   下圖顯示在目錄（共享和資料夾）中找到的個人資料。在「結構化」標籤中，您可以查看資料庫中的個人資料。在「非結構化」標籤中，您可以查看檔案層級資料。





查看包含敏感個人資料的文件

資料分類會自動識別隱私法規所定義的特殊類型的敏感個人資訊，例如 "GDPR 第 9 條和第 10 條"。例如，有關一個人的健康、種族或性取向的資訊。"查看完整列表" 。資料分類可在單一檔案、目錄（共用和資料夾）內的檔案以及資料庫表中識別此類資訊。

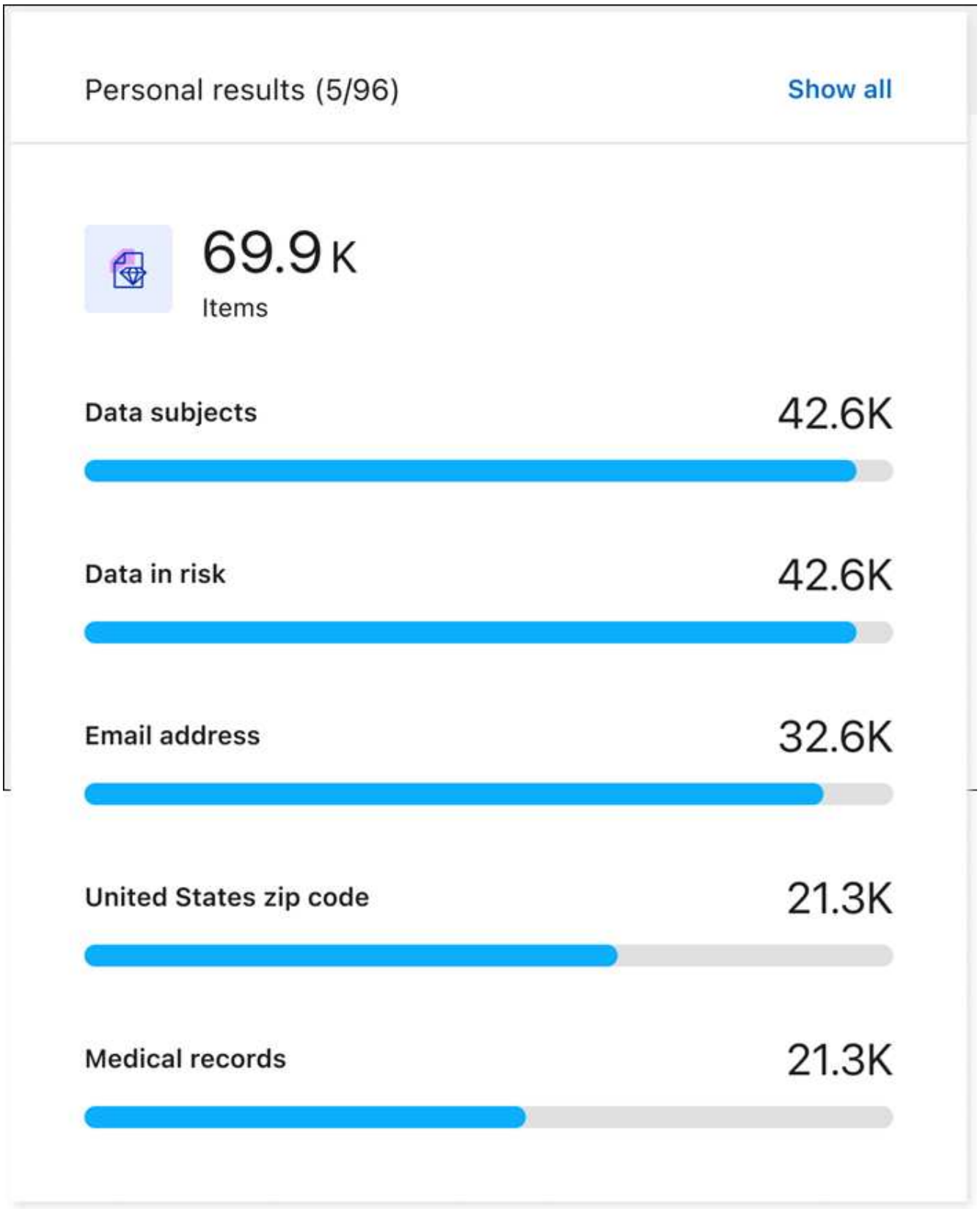
資料分類使用人工智慧、自然語言處理 (NLP)、機器學習 (ML) 和認知運算 (CC) 來理解其掃描的內容的含義，以便提取實體並對其進行相應的分類。

例如，GDPR 資料的一個敏感類別是種族來源。由於其 NLP 能力，資料分類可以區分「喬治是墨西哥人」（表示 GDPR 第 9 條規定的敏感資料）和「喬治正在吃墨西哥食物」之間的區別。

掃描敏感個人資料時僅支援英文。稍後將添加對更多語言的支援。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 要調查所有敏感個人資料的詳細信息，請找到敏感個人資訊結果卡，然後選擇顯示全部。



。

3. 要調查特定類型的敏感個人資料的詳細信息，請選擇“查看全部”，然後選擇特定類型的敏感個人資料的“調查結果”箭頭圖示。
4. 透過搜尋、排序、擴展特定文件的詳細資訊、點擊「調查結果」查看封鎖資訊或下載文件清單來調查資料。

NetApp Data Classification中的私有資料類別

NetApp Data Classification可在您的磁碟區和資料庫中識別多種類型的私有資料。

資料分類識別兩種類型的個人資料：

- 個人識別資訊（**PII**）
- 敏感個人資訊（**SPII**）

如果您需要資料分類來識別其他私人資料類型，例如額外的國民身分證號碼或醫療保健識別符，請聯絡您的客戶經理。

個人資料的類型

文件中的個人資料或_個人識別資訊_（PII）可以是一般個人資料或國家識別碼。下表第三列標識資料分類是否使用"接近度驗證"驗證其對標識符的發現。

表中標明了可以辨識這些項目的語言。

類型	識別符	接近度驗證？	英語	德文	西班牙語	法語	日本人
一般的	信用卡號碼	是的	✓	✓	✓		✓
	資料主體	不	✓	✓	✓		
	電子郵件	不	✓	✓	✓		✓
	IBAN 號碼（國際銀行帳號）	不	✓	✓	✓		✓
	IP 位址	不	✓	✓	✓		✓
	密碼	是的	✓	✓	✓		✓

類型	識別符	接近度驗證？	英語	德文	西班牙語	法語	日本人
國家識別符							

類型	識別符	接近度驗證？	英語	德文	西班牙語	法語	日本人

類型	識別符	接近度驗證的	英語	德文	西班牙語	法語	日本人
	希臘身分證	是的	✓	✓	✓		
	匈牙利稅務識別號	是的	✓	✓	✓		
	愛爾蘭身分證 (PPS)	是的	✓	✓			
	以色列身分證	是的	✓	✓	✓		
	義大利稅務識別號	是的	✓	✓	✓		
	日本個人身份證號碼（個人和公司）	是的	✓	✓	✓		✓
	拉脫維亞身分證	是的	✓	✓	✓		
	立陶宛身分證	是的	✓	✓	✓		
	盧森堡身分證	是的	✓	✓	✓		
	馬爾他身分證	是的	✓	✓	✓		
	國家醫療服務體系 (NHS) 號碼	是的	✓	✓	✓		
	紐西蘭銀行帳戶	是的	✓	✓	✓		
	紐西蘭駕駛執照	是的	✓	✓	✓		
	紐西蘭稅務局 (IRD) 號碼（稅號）	是的	✓	✓	✓		
	紐西蘭 NHI（國民健康指數）號碼	是的	✓	✓	✓		
	紐西蘭護照號碼	是的	✓	✓	✓		
	波蘭身分證 (PESEL)	是的	✓	✓	✓		
	葡萄牙稅務識別號碼（NIF）	是的	✓	✓	✓		
	羅馬尼亞身分證 (CNP)	是的	✓	✓	✓		
	新加坡國民登記身分證（NRIC）	是的	✓	✓	✓		
	斯洛維尼亞身分證 (EMSO)	是的	✓	✓	✓		
	南非身分證	是的	✓	✓	✓		
	西班牙稅務識別號	是的	✓	✓	✓		
	瑞典身分證	是的	✓	✓	✓		
	英國身分證（NINO）	是的	✓	✓	✓		
	美國加州駕駛執照	是的	✓	✓	✓		
	美國印第安納州駕照	是的	✓	✓	✓		
	美國紐約州駕駛執照	是的	✓	✓	✓		
	美國德州駕駛執照	是的	✓	✓	✓		
	美國社會安全號碼（SSN）	是的	✓	✓	✓		

敏感個人資料的類型

資料分類可以在文件中找到以下敏感個人資訊（SPII）。

以下 SPII 目前僅能以英文辨識：

- 刑事訴訟參考：有關自然人的刑事定罪和犯罪的數據。

- 種族參考：有關自然人的種族或民族血統的資料。
- 健康參考：有關自然人健康的數據。
- **ICD-9-CM** 醫療代碼：醫療保健產業使用的代碼。
- **ICD-10-CM** 醫療代碼：醫療保健產業使用的代碼。
- 哲學信仰參考：有關自然人的哲學信念的數據。
- 政治觀點參考：有關自然人政治觀點的數據。
- 宗教信仰參考：有關自然人的宗教信仰的資料。
- 性生活或性取向參考：自然人的性生活或性取向的資料。

類別類型

資料分類將您的資料分類如下。

大多數類別都可以用英語、德語和西班牙語識別。

類別	類型	英語	德文	西班牙語
金融	資產負債表	✓	✓	✓
	採購訂單	✓	✓	✓
	發票	✓	✓	✓
	季度報告	✓	✓	✓
人力資源	背景調查	✓		✓
	薪酬計劃	✓	✓	✓
	員工合約	✓		✓
	員工評價	✓		✓
	健康	✓		✓
	履歷	✓	✓	✓
合法的	保密協議	✓	✓	✓
	供應商-客戶合約	✓	✓	✓
行銷	活動	✓	✓	✓
	會議	✓	✓	✓
營運	審計報告	✓	✓	✓
銷售量	銷售訂單	✓	✓	
服務	射頻幹擾	✓		✓
	徵求建議書	✓		✓
	母豬	✓	✓	✓
	訓練	✓	✓	✓
支援	投訴和票務	✓	✓	✓

以下元資料也使用相同的支援語言進行分類和識別：

- 應用程式數據
- 存檔文件
- 聲音的
- 資料分類業務應用資料中的麵包屑
- CAD 檔案
- 程式碼
- 腐敗
- 資料庫和索引文件
- 設計文件
- 電子郵件應用程式數據
- 加密（具有高熵值的文件）
- 執行檔
- 財務應用數據
- 健康應用數據
- 圖片
- 紀錄
- 雜項文件
- 雜項演示
- 雜項電子表格
- 雜項“未知”
- 受密碼保護的文件
- 結構化資料
- 影片
- 零位元組文件

文件類型

資料分類掃描所有文件的類別和元資料洞察，並在儀表板的文件類型部分顯示所有文件類型。當資料分類偵測個人識別資訊 (PII) 或執行 DSAR 搜尋時，僅支援以下文件格式：

`.CSV, .DCM, .DOC, .DOCX, .JSON, .PDF, .PPTX, .RTF, .TXT, .XLS, .XLSX, Docs, Sheets, and Slides`

所發現資訊的準確性

NetApp無法保證資料分類識別的個人資料和敏感個人資料 100% 的準確性。您應該始終透過查看數據來驗證資訊。

根據我們的測試，下表顯示了資料分類發現的資訊的準確性。我們根據_精確度_和_召回率_來細分它：

精確

資料分類發現的內容被正確識別的機率。例如，個人資料的準確率為 90%，表示在被識別為包含個人資訊的 10 個文件中，有 9 個實際上包含個人資訊。 10 個文件中會有 1 個是誤報。

記起

資料分類找到其應有內容的機率。例如，個人資料的召回率為 70%，意味著資料分類可以識別出組織中 10 個文件中實際包含個人資訊的 7 個。數據分類會遺漏 30% 的數據，而這些數據不會出現在儀表板中。

我們正在不斷提高結果的準確性。這些改進將在未來的資料分類版本中自動提供。

類型	精確	記起
個人資料 - 一般	90%-95%	60%-80%
個人資料 - 國家識別符	30%-60%	40%-60%
敏感個人數據	80%-95%	20%-30%
類別	90%-97%	60%-80%

在NetApp Data Classification中建立自訂分類

NetApp Data Classification可讓您建立自訂類別或個人識別符，以識別特定於您組織監管和合規要求的資料。

資料分類支援兩種類型的自訂分類器：類別和個人識別碼。自訂類別是根據您上傳的一組檔案建立的，資料分類功能會根據這些檔案建立一個 AI 模型，以識別您組織中的類似資料（例如，一家健康研究公司可能會建立一個臨床分析類別）。使用關鍵字清單或正規表示式 (regex) 建立自訂個人識別符，以識別貴組織特有的、可能構成合規風險的資訊。

所有自訂分類都可以在自訂分類控制面板中找到。

建立自訂個人標識符

資料分類功能可讓您使用上下文關鍵字或正規表示式建立自訂個人識別符，以識別貴組織特有的資料。

關鍵字要求

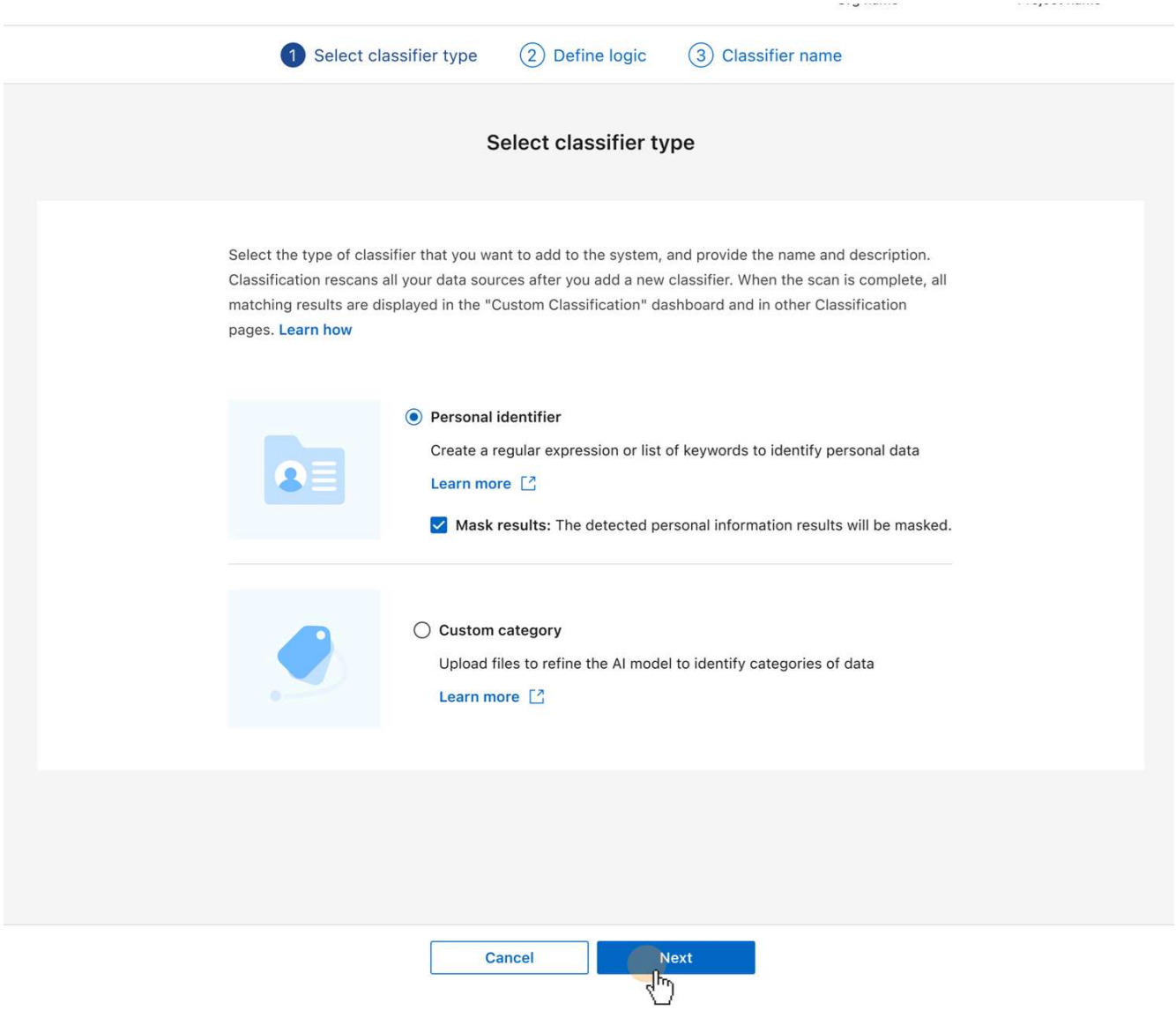
如果您使用關鍵字清單建立個人標識符，則該清單必須滿足以下要求：

- 關鍵字輸入不區分大小寫。
- 關鍵字必須至少包含三個字元。長度少於三個字元的單字將被忽略。
- 重複的字詞只會加一次。
- 關鍵字總數不能超過 50 萬個字元。清單中必須至少包含一個關鍵字。

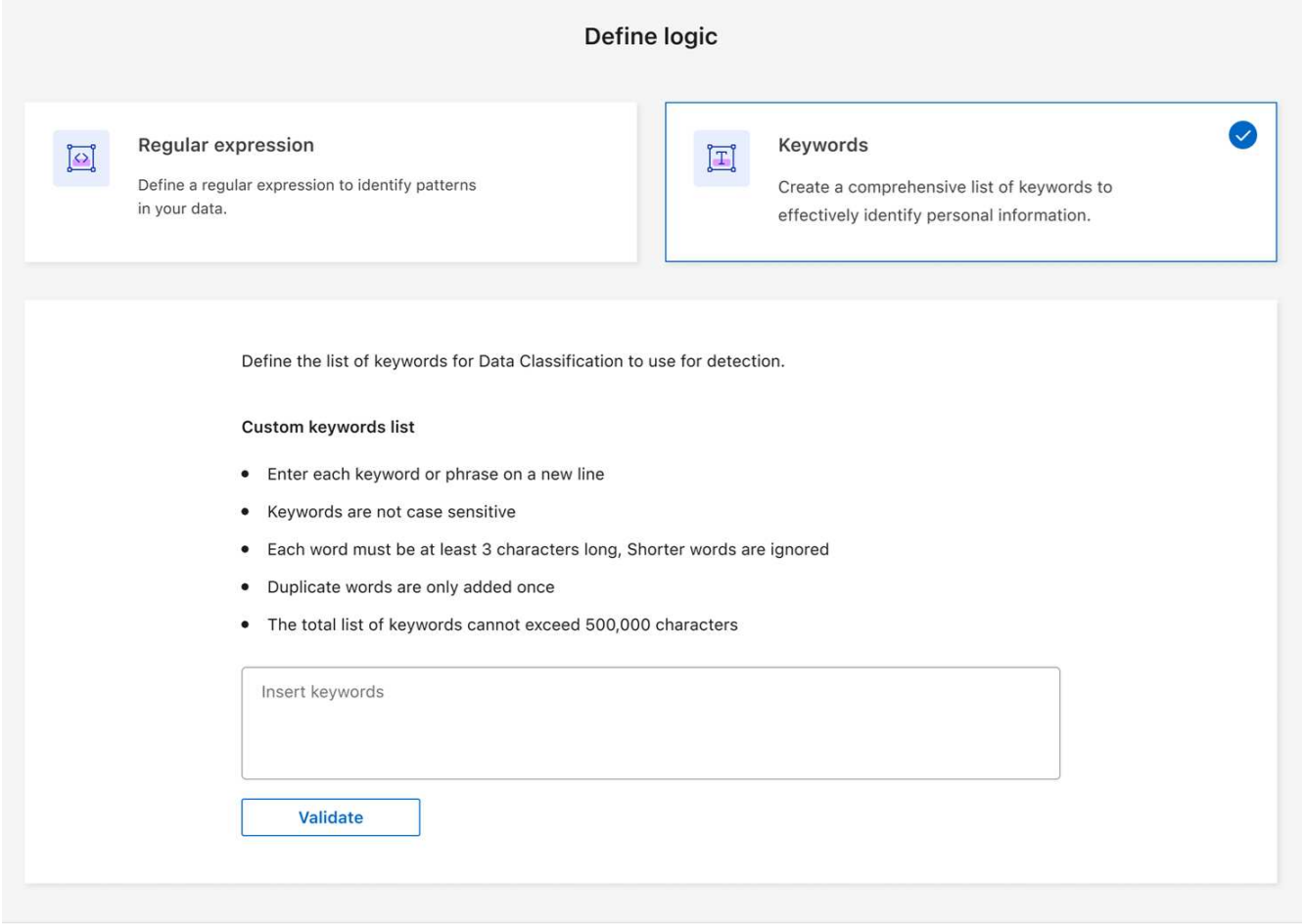
步驟

1. 選擇自訂分類選項卡。
2. 選擇**+** 新分類器以建立自訂分類器。

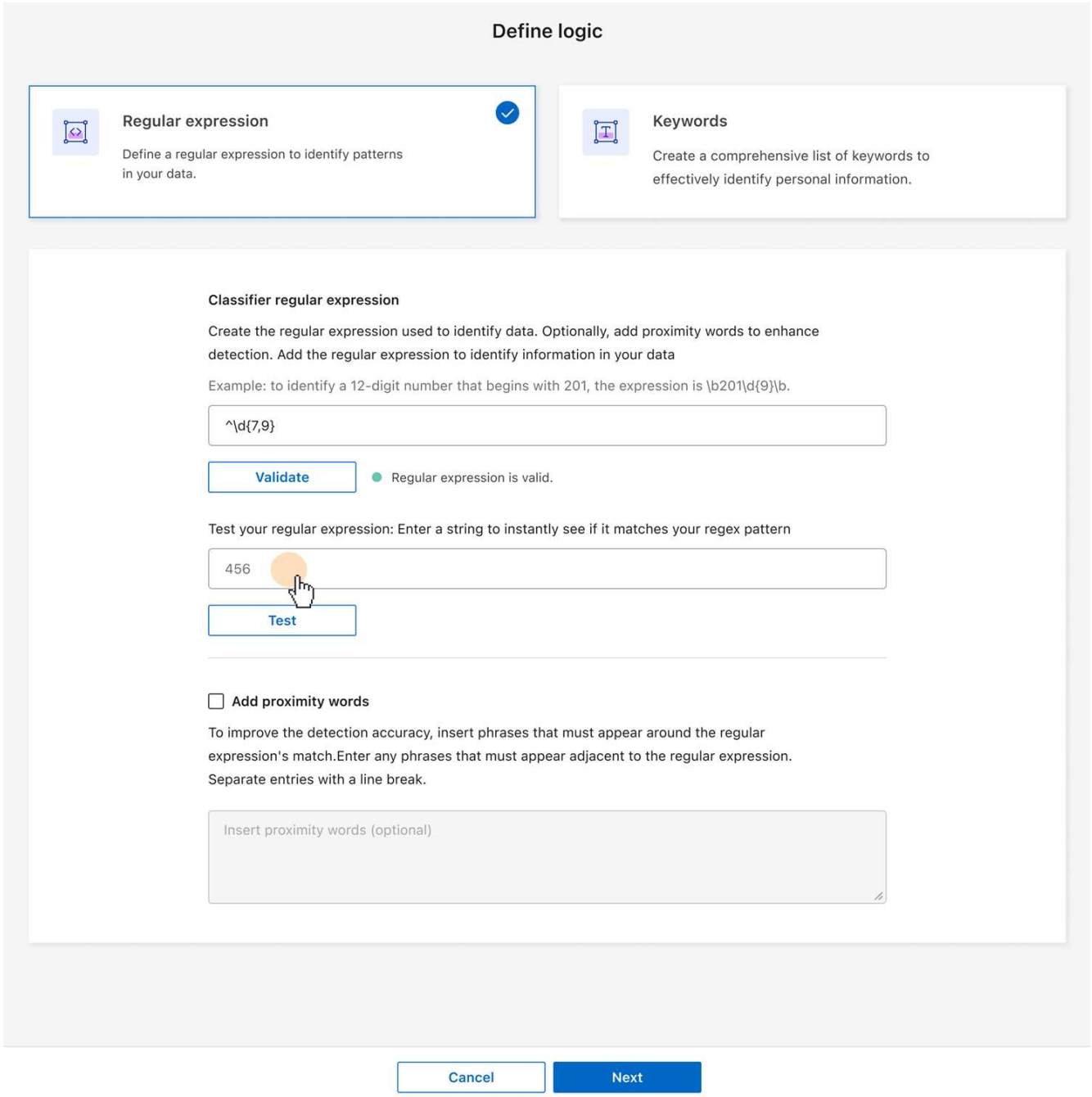
3. 請選擇*個人識別碼*。（可選）選擇「屏蔽結果」以封鎖偵測到的個人資料。
4. 選擇下一步。



5. 若要新增帶有關鍵字的分類器，請選擇關鍵字。請輸入關鍵字列表，每個關鍵字佔一行。請確保關鍵字符合要求。



若要將分類器新增為正規表示式，請選擇正規表示式，然後新增模式來偵測資料的特定資訊。選擇驗證以確認您輸入的語法正確。



a. （可選）輸入一個應該與正規表示式模式相符的範例字串，然後選擇測試進行檢查。

b. （可選）添加鄰近詞。如果新增鄰近詞，則資料分類僅在鄰近詞與匹配字串相鄰時才標記正規表示式模式。

6. 選擇下一步。

7. 輸入分類器名稱和描述，以便在儀表板中識別自訂類別。

8. 選擇儲存以建立自訂個人識別碼。

建立自訂個人標識符後，其結果將在下次計劃掃描中捕獲。為了更快地取得結果，請執行按需掃描。若要查看結果，請參閱 產生合規性報告。

建立自訂類別

透過自訂類別，您可以對特定於您組織的資料進行分類。自訂類別是根據您上傳的文字檔案建立的，資料分類功能會根據這些檔案建立一個人工智慧模型，以識別其他檔案中的類似資訊。

訓練資料要求

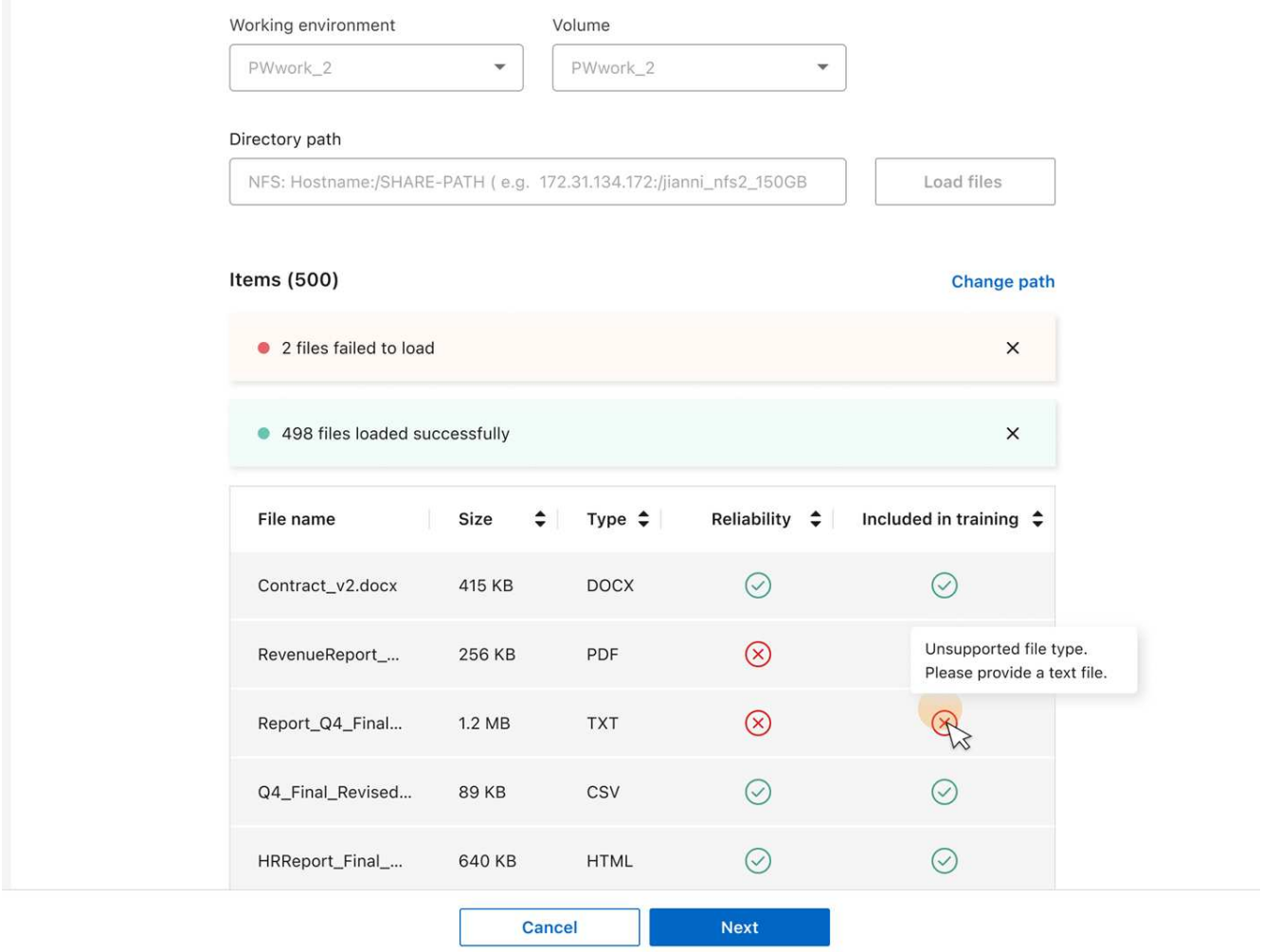
- 訓練資料集必須至少包含 25 個檔案。最大文件數為 1,000。
- 所有文件必須直接位於您提供的文件路徑中。
- 所有檔案必須大於 100 位元組。
- 資料分類訓練資料必須是下列檔案類型之一：CSV、DOCX、DOC、GZ、JSON、PDF、PPTX、TXT、RTT、XLS 或 XLSX。您可以上傳所有支援的文件類型的組合。

步驟

1. 在NetApp Data Classification中，選擇「自訂分類」。
2. 選擇 **+** 新分類器。
3. 選擇“自訂類別”作為分類器類型，然後下一步。
4. 使用一系列基於文字的文件來定義自訂類別的邏輯。請提供*工作位址*的IP位址，然後從下拉式選單中選擇*音量*。

   輸入包含訓練資料的目錄的目錄路徑。

5. 選擇“載入檔案”進行資料分類，以執行檔案檢查。您可以查看文件摘要，其中列出了文件名稱、大小、類型和備註（如果該文件被認為適合用於培訓）。



a. 若要變更檔案路徑或重新上傳文件，請選擇變更路徑，然後輸入資料並再次載入檔案。

6. 當您對上傳的文件滿意後，請選擇下一步。
7. 輸入分類器名稱和描述，以便在儀表板中識別自訂類別。
8. 選擇儲存以建立自訂類別。

結果

建立自訂類別後，其結果將在下次計畫掃描中擷取。為了更快地取得結果，請手動啟動掃描。

編輯自訂分類器

建立個人識別碼後，您可以修改其邏輯。您無法變更個人識別碼的類型或邏輯類型；例如，您無法將自訂類別變更為自訂個人識別碼。您也不能將基於關鍵字的自訂識別碼變更為基於正規表示式的自訂識別碼。

步驟

1. 在NetApp Data Classification中，選擇「自訂分類」。
2. 確定要刪除的分類器，然後選擇操作選單 ... 在它那一行的末尾。
3. 選擇編輯邏輯。
4. 如果要修改關鍵字，請新增、刪除或編輯對應的關鍵字。如果要修改正規表示式，請輸入新的正規表示式並進行驗證。（可選）加入鄰近關鍵字。

5. 選擇“儲存”以套用變更。

刪除自訂分類器

1. 在NetApp Data Classification中，選擇「自訂分類」。
2. 確定要刪除的分類器，然後選擇操作選單 ... 在它那一行的末尾。
3. 選擇刪除分類器。

下一步

- 產生合規性報告

使用NetApp Data Classification調查組織中儲存的數據

資料調查儀表板顯示文件和目錄層級的資料洞察，使您能夠對結果進行排序和過濾。數據調查頁面提供有關文件和目錄元數據和權限的見解以及識別重複文件。透過文件、目錄和資料庫層級的洞察，您可以採取措施來提高組織的合規性並節省儲存空間。資料調查頁面還支援移動、複製和刪除檔案。

要從調查頁面獲得見解，您必須對資料來源執行完整的分類掃描。僅進行過映射掃描的資料來源不會顯示檔案層級的詳細資訊。

資料調查結構

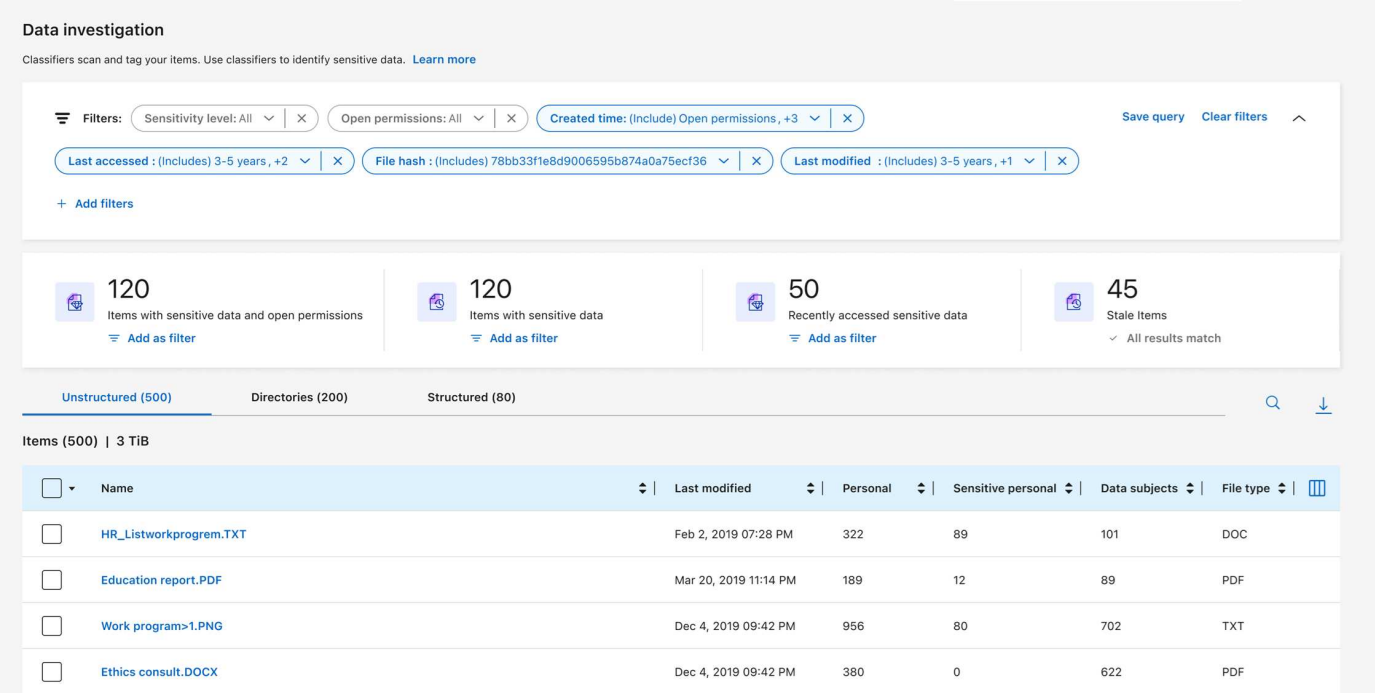
數據調查頁面將數據分類到三個選項卡：

- 非結構化資料：文件數據
- 目錄：資料夾和檔案共享
- 結構化：資料庫

數據過濾器

資料調查頁面提供了許多過濾器來對您的資料進行分類，以便您可以找到所需的資料。您可以同時使用多個過濾器。

若要新增過濾器，請選擇新增過濾器按鈕。



過濾敏感度和內容

使用以下過濾器查看您的資料包含多少敏感資訊。

篩選	細節
類別	選擇"類別類型"。
敏感度等級	選擇敏感度等級：個人、敏感個人或非敏感。
識別符數量	選擇每個檔案偵測到的敏感標識符的範圍。包括個人資料和敏感個人資料。在目錄中過濾時，資料分類會將每個資料夾（和子資料夾）中所有檔案的匹配結果匯總。注意：2023 年 12 月（版本 1.26.6）版本刪除了按目錄計算個人識別資訊 (PII) 資料數量的選項。
個人資料	選擇"個人資料的類型"。
敏感個人資訊	選擇"敏感個人資料的類型"。
資料主體	輸入資料主體的全名或已知標識符。"在此處了解有關數據主體的更多信息" 。

過濾使用者擁有者和使用者權限

使用下列過濾器查看文件擁有者和存取資料的權限。

篩選	細節
開放權限	選擇資料和資料夾/共享內的權限類型。
使用者/群組權限	選擇一個或多個使用者名稱和/或群組名，或輸入部分名稱。
文件所有者	輸入檔案所有者名稱。
有存取權限的使用者數	選擇一個或多個類別範圍以顯示哪些檔案和資料夾對一定數量的使用者開放。

按時間順序過濾

使用以下過濾器根據時間標準查看資料。

篩選	細節
創建時間	選擇文件建立的時間範圍。您也可以指定自訂時間範圍來進一步最佳化搜尋結果。
發現時間	選擇資料分類發現檔案的時間範圍。您也可以指定自訂時間範圍來進一步最佳化搜尋結果。
上次修改	選擇檔案最後修改的時間範圍。您也可以指定自訂時間範圍來進一步最佳化搜尋結果。
上次訪問	選擇檔案或目錄*上次被存取的時間範圍。您也可以指定自訂時間範圍來進一步最佳化搜尋結果。對於資料分類掃描的檔案類型，這是資料分類最後一次掃描該檔案的時間。

{星號} 目錄的上次存取時間僅適用於 NFS 或 CIFS 共用。

過濾元資料

使用下列篩選器根據位置、大小和目錄或檔案類型檢視資料。

篩選	細節
文件路徑	輸入最多 20 個要在查詢中包含或排除的部分或完整路徑。如果同時輸入包含路徑和排除路徑，資料分類會先在包含路徑中找到所有文件，然後從排除路徑中刪除文件，然後顯示結果。請注意，在此過濾器中使用“*”沒有任何效果，並且您無法從掃描中排除特定資料夾 - 配置共用下的所有目錄和檔案都將被掃描。
目錄類型	選擇目錄類型；“共享”或“資料夾”。
文件類型	選擇"文件類型"。
文件大小	選擇檔案大小範圍。
文件哈希	輸入文件的雜湊值即可找到特定文件，即使名稱不同。

過濾器儲存類型

使用以下過濾器按儲存類型查看資料。

篩選	細節
系統類型	選擇系統類型。
系統環境名稱	選擇特定係統。
儲存庫	選擇儲存庫，例如磁碟區或模式。

過濾查詢

使用下列篩選器按已儲存的查詢查看資料。

篩選	細節
已儲存的查詢	選擇一個或多個已儲存的查詢。前往"已儲存的查詢選項卡"查看現有已儲存查詢的清單並建立新查詢。
標籤	選擇"一個或多個標籤"分配給您的文件。

過濾分析狀態

使用以下過濾器按資料分類掃描狀態查看資料。

篩選	細節
分析狀態	選擇一個選項來顯示「等待首次掃描」、「已完成掃描」、「等待重新掃描」或「掃描失敗」的檔案清單。
掃描分析事件	選擇是否要查看由於資料分類無法恢復上次存取時間而未分類的文件，或即使資料分類無法恢復上次存取時間但已分類的文件。

"查看有關“上次訪問時間”時間戳的詳細信息"有關使用掃描分析事件進行過濾時調查頁面中出現的項目的詳細資訊。

按重複項過濾資料

使用以下過濾器查看儲存中重複的檔案。

篩選	細節
重複項	選擇檔案是否在儲存庫中重複。

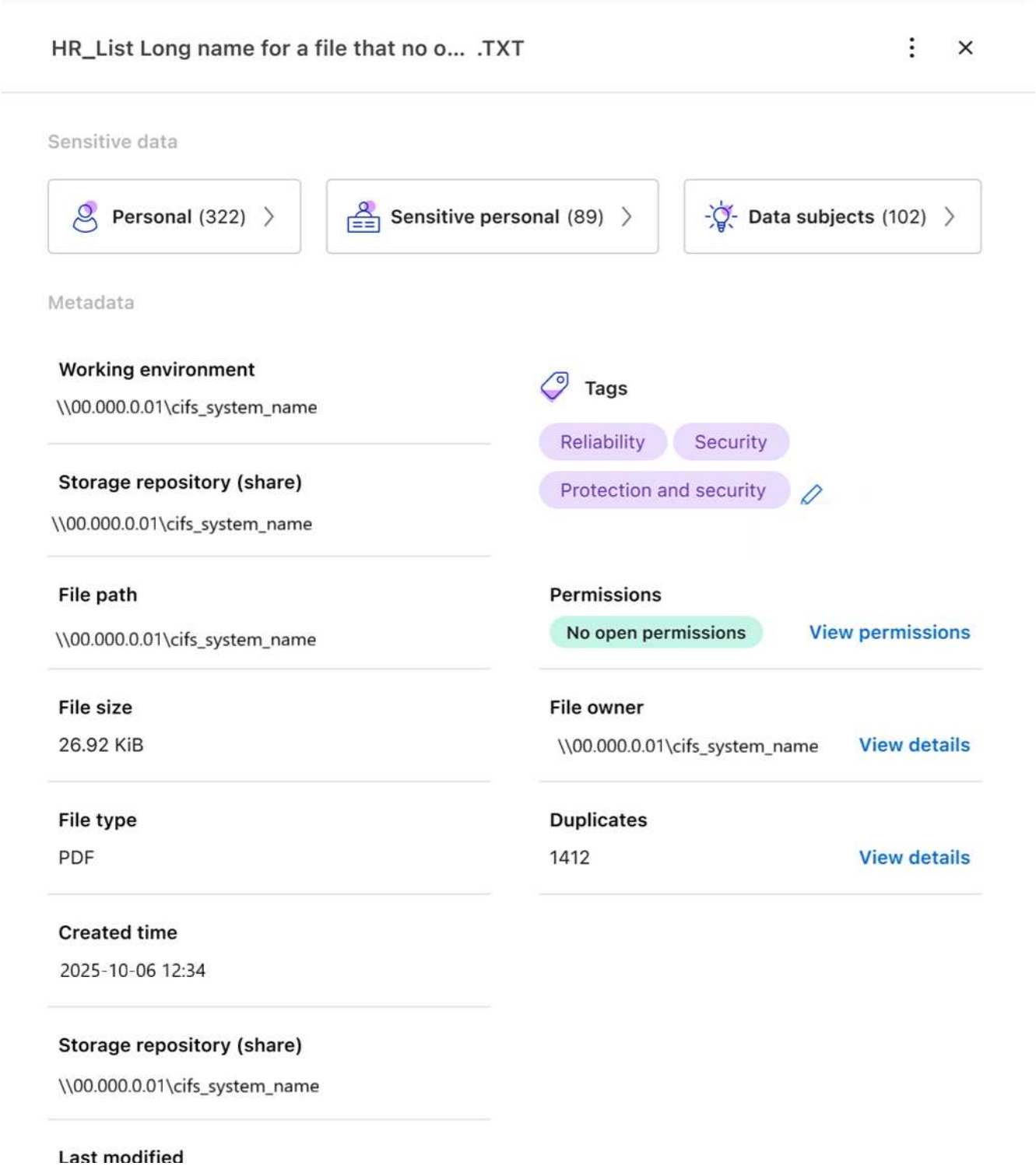
查看檔案元數據

除了顯示文件所在的系統和磁碟區之外，元資料還顯示更多信息，包括文件權限、文件擁有者以及該文件是否有重複。如果您打算"建立已儲存的查詢"因為您可以看到可用於過濾資料的所有資訊。

資訊的可用性取決於資料來源。例如，資料庫檔案的磁碟區名稱和權限不共用。

步驟

1. 從資料分類選單中，選擇*調查*。
2. 在右側的資料調查清單中，選擇向下插入符號 在任意單一文件的右側查看文件元資料。



3. 或者，您可以使用*建立標籤*按鈕為檔案建立或新增標籤。從下拉式選單中選擇一個現有標籤或使用 + 新增按鈕新增一個新標籤。標籤可用於過濾資料。

查看檔案和目錄的使用者權限

若要查看有權存取檔案或目錄的所有使用者或群組的清單以及他們擁有的權限類型，請選擇「查看所有權限」。此選項僅適用於 CIFS 共享中的資料。

如果您使用安全性識別碼 (SID) 而不是使用者名稱和群組名，則應該將 Active Directory 整合到資料分類中。有

關更多信息，請參閱"將 Active Directory 新增至資料分類" 。

步驟

1. 從資料分類選單中，選擇*調查*。
2. 在右側的資料調查清單中，選擇向下插入符號在任意單一文件的右側查看文件元資料。
3. 若要查看有權存取檔案或目錄的所有使用者或群組的清單以及他們擁有的權限類型，請在「開啟權限」欄位中選擇「查看所有權限」。

   資料分類在清單中顯示最多 100 個使用者。

4. 選擇向下插入符號任何群組的按鈕即可查看屬於該群組的使用者清單。

   您可以展開該群組的某個層級來查看屬於該群組的使用者。

5. 選擇使用者或群組的名稱以重新整理調查頁面，以便您可以看到該使用者或群組有權存取的所有檔案和目錄。

檢查儲存系統中的重複文件

您可以檢查儲存系統中是否儲存了重複的檔案。如果您想確定可以節省儲存空間的區域，這將非常有用。確保具有特定權限或敏感資訊的某些檔案不會在儲存系統中不必要地重複也是很好的。

資料分類會比較所有檔案（資料庫除外）是否有重複項，如果存在重複項，則進行以下操作：

- 1 MB 或更大
- 或包含個人資訊或敏感個人資訊

資料分類使用雜湊技術來確定重複檔案。如果一個檔案的雜湊碼與另一個檔案相同，即使檔案名稱不同，這兩個檔案也是完全相同的副本。

步驟

1. 從資料分類選單中，選擇*調查*。
2. 在「篩選器」窗格中，選擇「檔案大小」以及「重複」（「有重複」）以查看您的環境中哪些特定大小範圍的檔案是重複的。
3. 或者，下載重複檔案的清單並將其發送給儲存管理員，以便他們可以決定可以刪除哪些檔案（如果有）。
4. 您可以選擇刪除、標記或移動重複的檔案。選擇您想要執行操作的文件，然後選擇適當的操作。

查看特定檔案是否重複

您可以查看單一文件是否有重複。

步驟

1. 從資料分類選單中，選擇*調查*。
2. 在資料調查清單中，選擇在任意單一文件的右側查看文件元資料。

   如果檔案存在重複，則此資訊將顯示在「*Duplicates*」欄位旁邊。

3. 若要查看重複檔案的清單及其位置，請選擇「查看詳細資料」*。
4. 在下一頁中選擇「查看重複項」以查看調查頁面中的文件。
5. 您可以選擇刪除、標記或移動重複的檔案。選擇您想要執行操作的文件，然後選擇適當的操作。

您可以使用此頁面提供的「檔案雜湊」值並將其直接輸入到調查頁面中，以便隨時搜尋特定的重複檔案 - 或者您可以在已儲存的查詢中使用它。

下載您的報告

您可以以 CSV 或 JSON 格式下載過濾結果。

如果資料分類正在掃描檔案（非結構化資料）、目錄（資料夾和檔案共用）和資料庫（結構化資料），則最多可以下載三個報表檔案。

檔案被分割成具有固定行數或記錄數的檔案：

- JSON：每份報告 100,000 筆記錄，產生大約需要 5 分鐘
- CSV：每份報告 200,000 筆記錄，產生大約需要 4 分鐘

您可以下載 CSV 檔案的版本以在此瀏覽器中查看。此版本限制為 10,000 筆記錄。

可下載報告所包含的內容

*非結構化文件資料報告*包含有關您的文件的以下資訊：

- 檔案名稱
- 位置類型
- 系統名稱
- 儲存庫（例如，磁碟區、儲存桶、共用）
- 儲存庫類型
- 文件路徑
- 文件類型
- 文件大小（單位：MB）
- 創建時間
- 上次修改時間
- 上次訪問
- 文件所有者
  - 設定 Active Directory 時，檔案擁有者資料包含帳戶名稱、SAM 帳戶名稱和電子郵件地址。
- 類別
- 個人資訊
- 敏感個人資訊

- 開放權限
- 掃描分析錯誤
- 刪除檢測日期

  刪除檢測日期標識檔案被刪除或移動的日期。這使您能夠識別敏感文件何時被移動。已刪除的文件不會計入儀表板或調查頁面上顯示的文件數量。這些文件僅出現在 CSV 報告中。

*非結構化目錄資料報告*包括有關您的資料夾和檔案共享的以下資訊：

- 系統類型
- 系統名稱
- 目錄名稱
- 儲存庫（例如資料夾或檔案共用）
- 目錄所有者
- 創建時間
- 發現時間
- 上次修改時間
- 上次訪問
- 開放權限
- 目錄類型

*結構化資料報告*包含有關資料庫表的以下資訊：

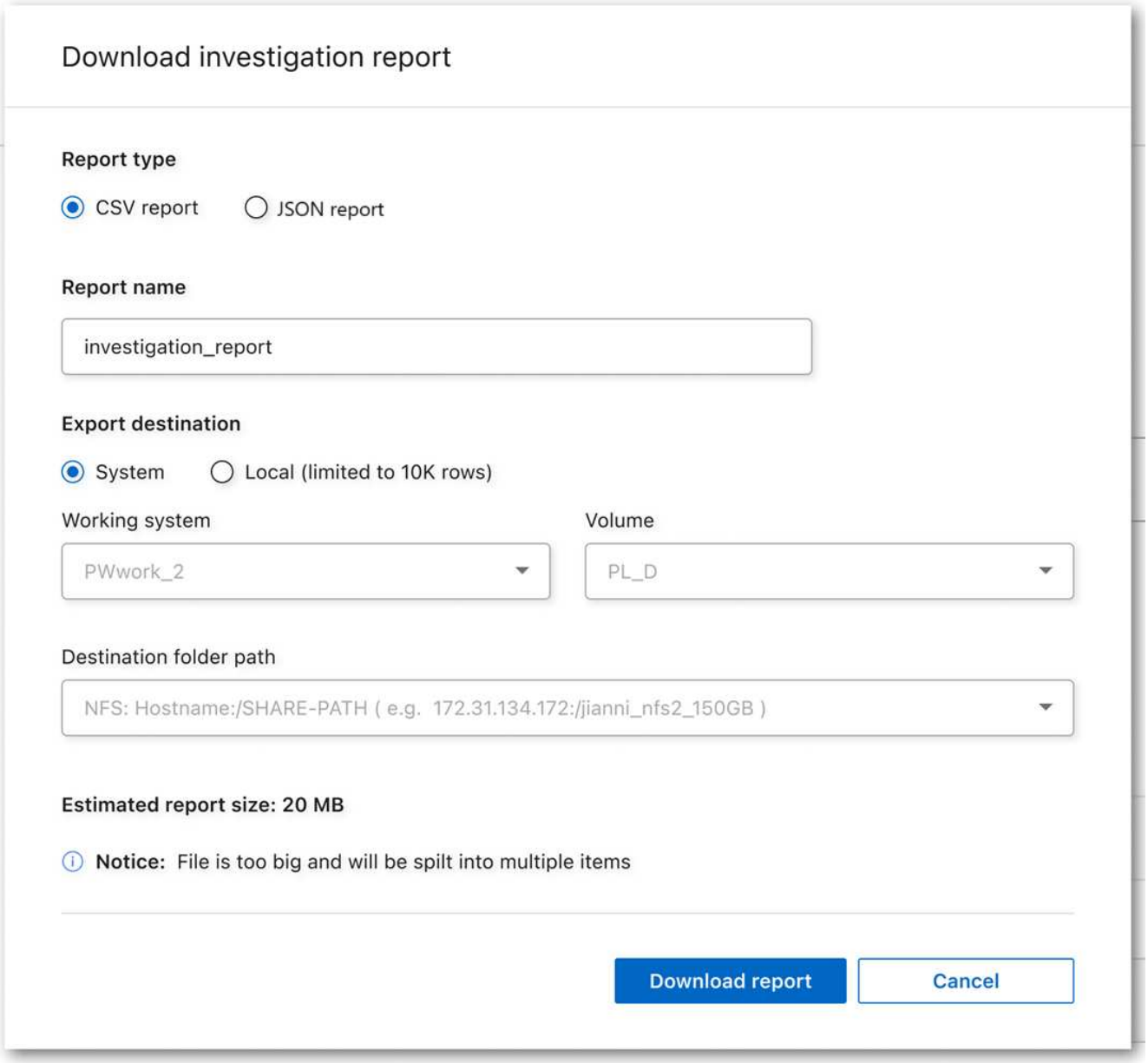
- 資料庫表名稱
- 位置類型
- 系統名稱
- 儲存庫（例如，架構）
- 列數
- 行數
- 個人資訊
- 敏感個人資訊

產生報告的步驟

1. 從資料調查頁面中，選擇  頁面右上方的按鈕。
2. 選擇報告類型：CSV 或 JSON。
3. 輸入報告名稱。
4. 若要下載完整的報告，請選擇系統，然後從對應的下拉式選單中選擇系統和磁碟區。提供目標資料夾路徑。

   若要在瀏覽器中下載報告，請選擇本機。請注意，此選項將報表限制為前 10,000 行，並且僅限於 **CSV** 格式。如果您選擇本機，則無需填寫任何其他欄位。

5. 選擇下載報告。



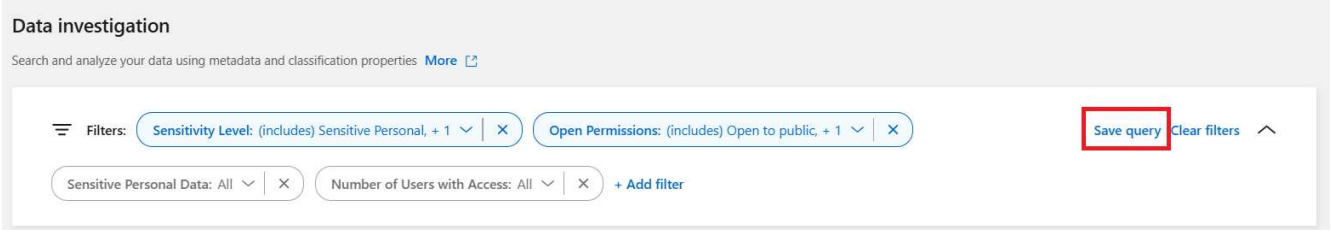
結果

對話方塊中將顯示一則訊息，提示正在下載報告。

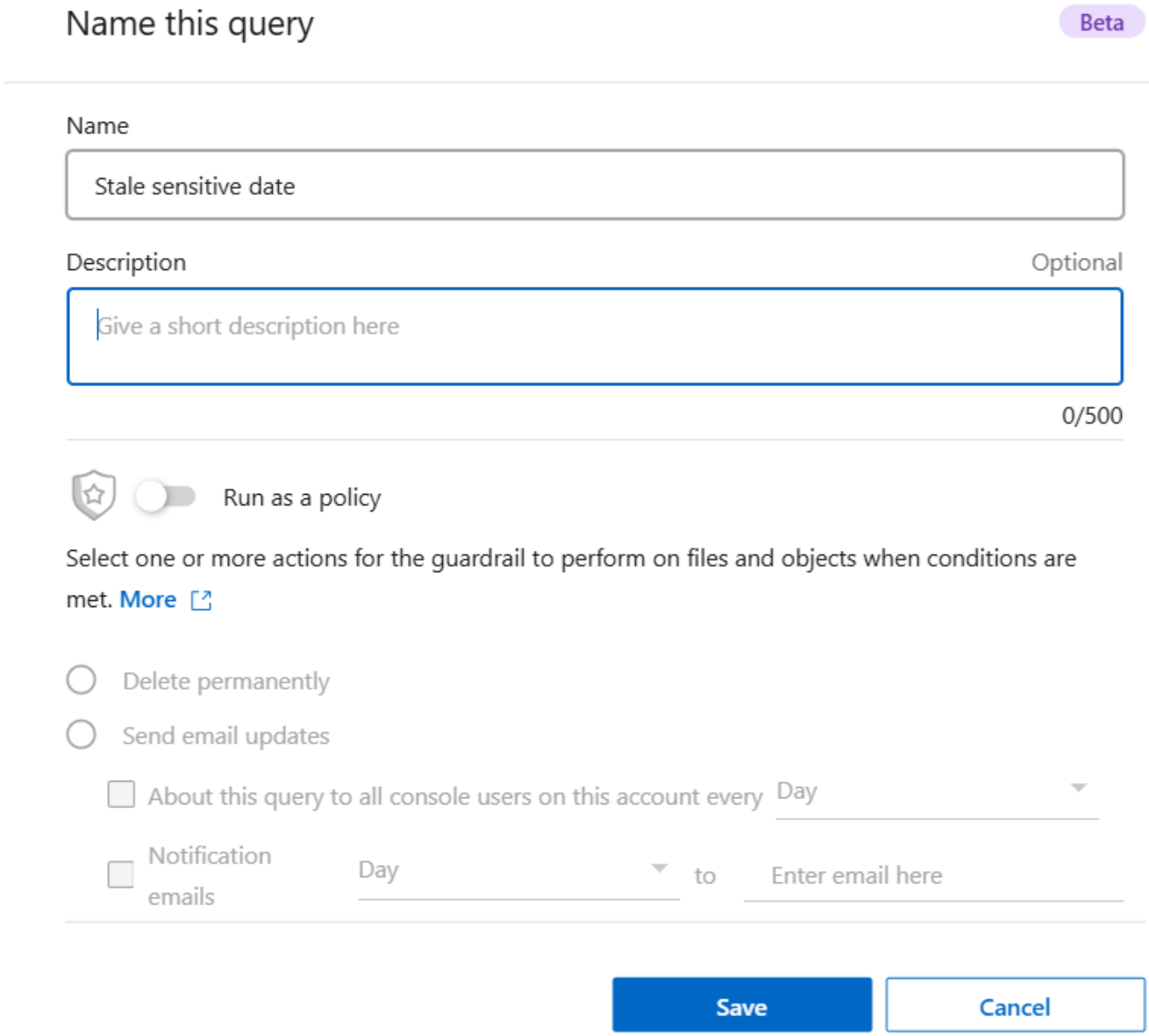
根據選定的篩選器建立已儲存的查詢

步驟

1. 在調查標籤中，透過選擇要使用的篩選器來定義搜尋。看"在調查頁面中過濾數據"了解詳情。
2. 一旦您根據自己的喜好設定了所有過濾器特性，請選擇*儲存查詢*。



3. 為已儲存的查詢命名並新增描述。該名稱必須是唯一的。
4. 您可以選擇將查詢儲存為策略：
   a. 若要將查詢儲存為策略，請切換*作為策略執行*開關。
   b. 選擇*永久刪除*或*發送電子郵件更新*。如果您選擇電子郵件更新，您可以每天、每週或每月透過電子郵件將查詢結果傳送給所有控制台使用者。或者，您可以以相同的頻率將通知傳送到特定的電子郵件地址。
5. 選擇*儲存*。



建立搜尋或策略後，您可以在已儲存的查詢標籤中查看它。

結果可能需要最多 15 分鐘才會顯示在「已儲存的查詢」頁面上。

使用NetApp Data Classification管理已儲存的查詢

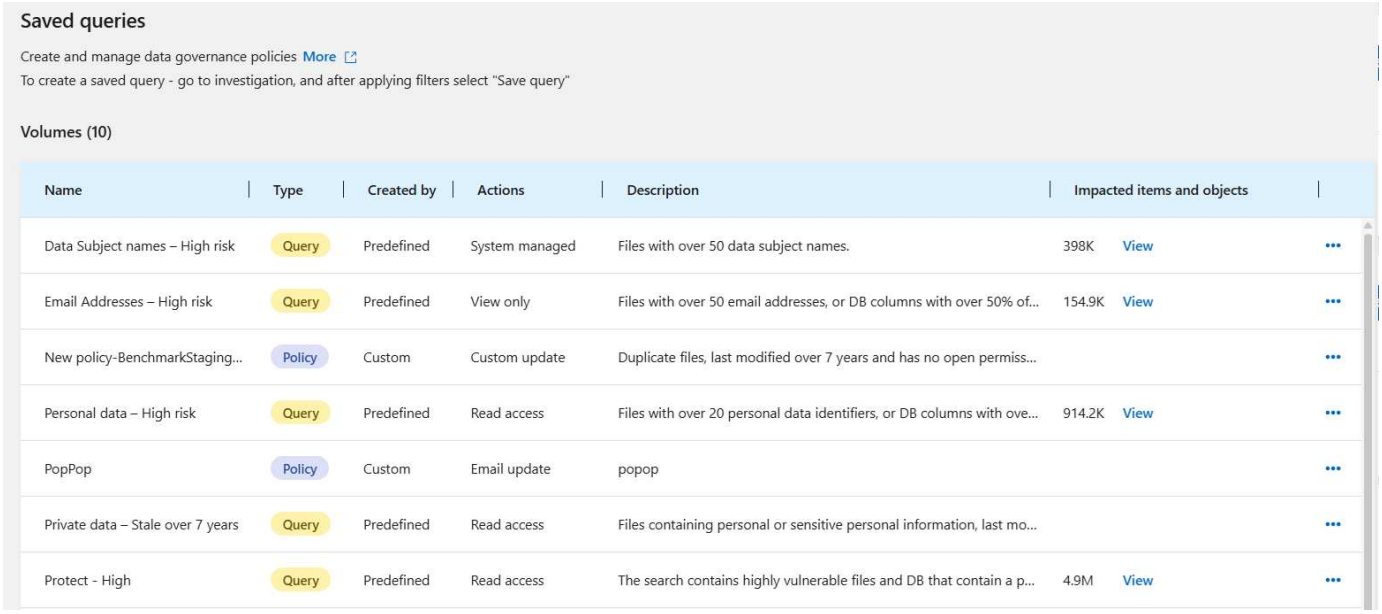
NetaApp 資料分類支援保存您的搜尋查詢。使用已儲存的查詢，您可以建立自訂篩選器來對資料調查頁面的常見查詢進行排序。資料分類還包括基於常見請求的預先定義保存的查詢。

合規性儀表板中的「已儲存的查詢」標籤列出了此資料分類實例上可用的所有預訂和自訂已儲存查詢。

已儲存的查詢也可以儲存為策略。查詢過濾數據，而策略允許您對數據採取行動。透過策略：您可以刪除發現的

資料或發送有關發現的資料的電子郵件更新。

已儲存的查詢也會出現在調查頁面的篩選器清單中。



在調查頁面中查看已儲存的查詢結果

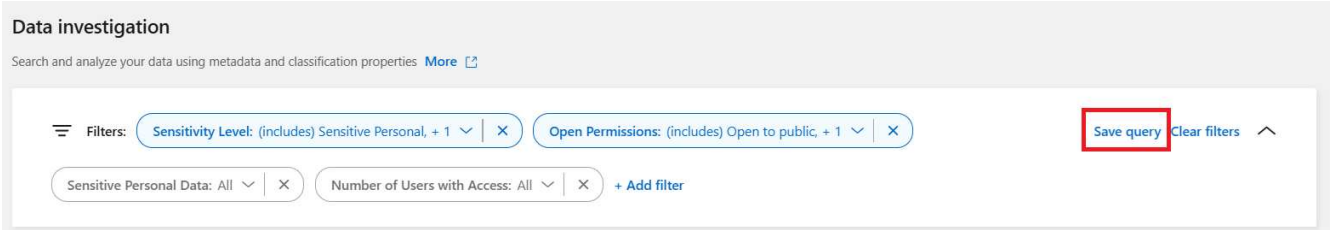
若要在調查頁面中顯示已儲存查詢的結果，請選擇 按鈕進行特定搜索，然後選擇*調查結果*。

建立已儲存的查詢和策略

您可以建立自己的自訂已儲存查詢，以提供特定於您組織的查詢結果。傳回符合搜尋條件的所有檔案和目錄（共用和資料夾）的結果。

步驟

1. 在調查標籤中，透過選擇要使用的篩選器來定義搜尋。看"在調查頁面中過濾數據"了解詳情。
2. 一旦您根據自己的喜好設定了所有過濾器特性，請選擇*儲存查詢*。

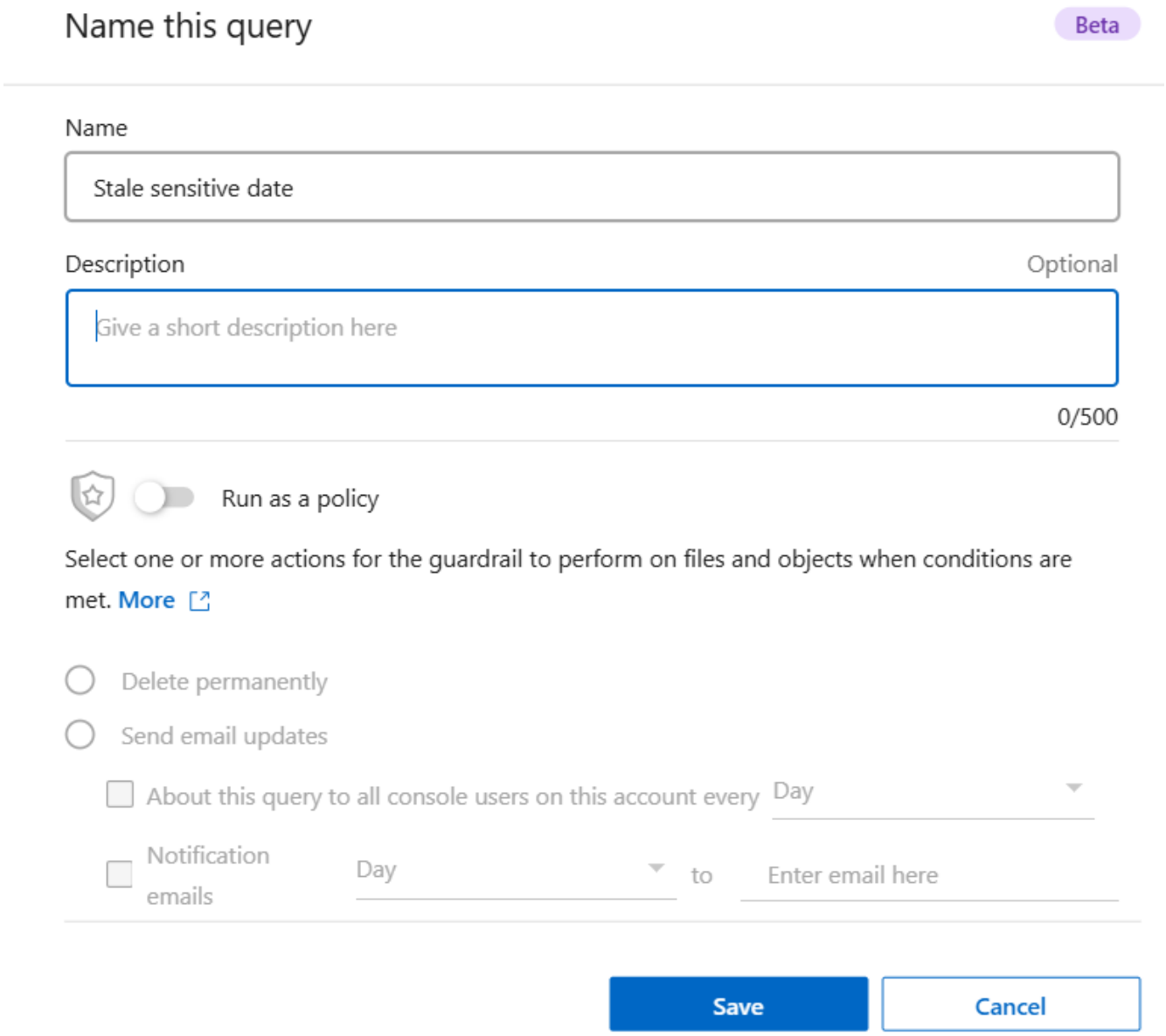


3. 為已儲存的查詢命名並新增描述。該名稱必須是唯一的。
4. 您可以選擇將查詢儲存為策略：

a. 若要將查詢儲存為策略，請切換*作為策略執行*開關。

b. 選擇*永久刪除*或*發送電子郵件更新*。如果您選擇電子郵件更新，您可以每天、每週或每月透過電子郵件將查詢結果傳送給所有控制台使用者。或者，您可以以相同的頻率將通知傳送到特定的電子郵件地址。

5. 選擇*儲存*。



建立搜尋或策略後，您可以在已儲存的查詢標籤中查看它。

編輯已儲存的查詢或策略

您可以修改已儲存查詢的名稱和描述。您也可以將查詢轉換為策略，反之亦然。

您不能修改預設儲存的查詢。您不能修改已儲存查詢的篩選器。您可以交替查看已儲存查詢的調查結果，變更或

修改篩選器，然後將其儲存為新查詢或政策。

步驟

1. 在「已儲存的查詢」頁面中，選擇要變更的搜尋的「編輯搜尋」。

2. 對名稱和描述欄位進行變更。僅更改名稱和描述欄位。

   您可以選擇將查詢轉換為策略，或將策略轉換為已儲存的查詢。根據需要切換*作為策略運作*開關。..如果您要將查詢轉換為策略，請選擇*永久刪除*或*發送電子郵件更新*。如果您選擇電子郵件更新，您可以每天、每週或每月透過電子郵件將查詢結果傳送給所有控制台使用者。或者，您可以以相同的頻率將通知傳送到特定的電子郵件地址。

3. 選擇“儲存”以完成變更。

刪除已儲存的查詢

如果您不再需要任何自訂已儲存的查詢或策略，可以將其刪除。您無法刪除預設儲存的查詢。

若要刪除已儲存的查詢，請選擇 按鈕進行特定搜索，選擇*刪除查詢*，然後在確認對話框中再次選擇*刪除查詢*。

預設查詢

資料分類提供以下系統定義的搜尋查詢：

- 資料主體姓名 - 高風險

  包含超過 50 個資料主體名稱的文件

- 電子郵件地址 - 高風險

  包含超過 50 個電子郵件地址的文件或資料庫列中超過 50% 的行包含電子郵件地址

- 個人資料 - 高風險

  包含超過 20 個個人資料標識符的文件或資料庫列中超過 50% 的行包含個人資料標識符

- 私人資料 - 已過期 7 年以上

  包含個人或敏感個人資訊的文件，上次修改超過 7 年

- 保護 - 高

  包含密碼、信用卡資訊、IBAN 號碼或社會安全號碼的文件或資料庫列

- 保護 - 低

超過 3 年未訪問的文件

- 保護 - 中等

  包含具有個人資料識別碼（包括身分證號碼、稅務識別號碼、駕駛執照號碼、藥品 ID 或護照號碼）的文件或資料庫列的文件

- 敏感個人資料 - 高風險

  包含超過 20 個敏感個人資料識別碼的文件或資料庫列中超過 50% 的行包含敏感個人數據

更改儲存庫的NetApp Data Classification掃描設置

您可以管理在每個系統和資料來源中如何掃描資料。您可以在「儲存庫」基礎上進行變更；這表示您可以根據正在掃描的資料來源類型對每個磁碟區、模式、使用者等進行變更。

您可以更改的一些內容包括是否掃描儲存庫，以及NetApp Data Classification是否正在執行"映射掃描或映射和分類掃描"。您也可以暫停和恢復掃描，例如，如果您需要在一段時間內停止掃描某個磁碟區。

查看儲存庫的掃描狀態

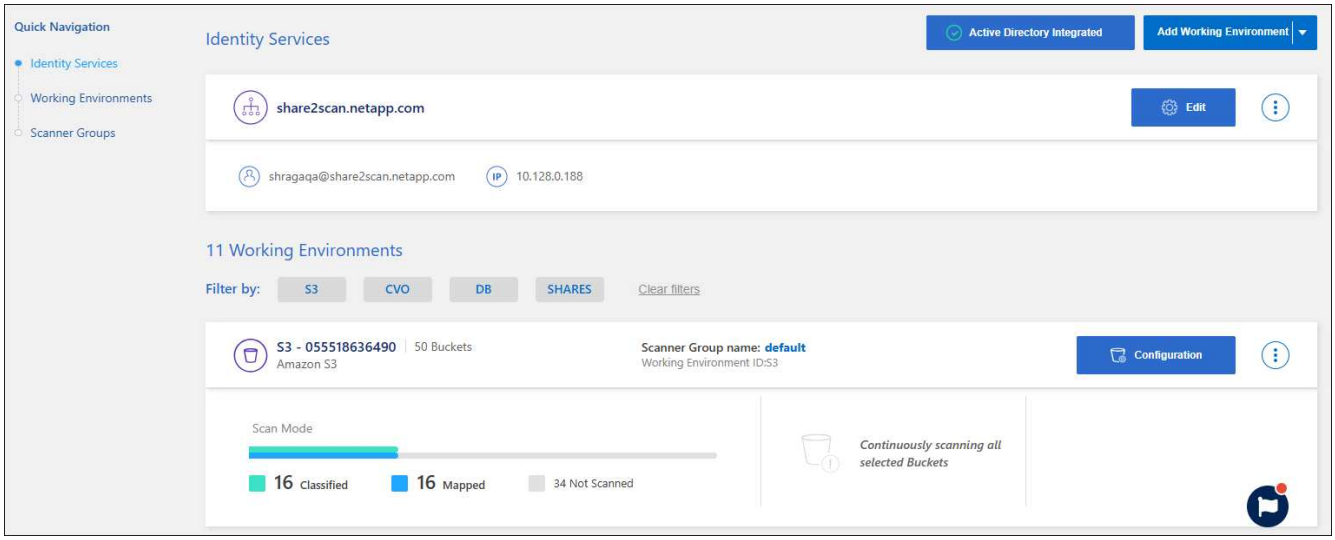
您可以查看NetApp Data Classification正在為每個系統和資料來源掃描的各個儲存庫（磁碟區、儲存桶等）。您還可以看到有多少已被“映射”，有多少已被“分類”。分類需要更長的時間，因為所有資料都進行了完整的 AI 識別。

您可以在設定頁面查看各個工作環境的掃描狀態：

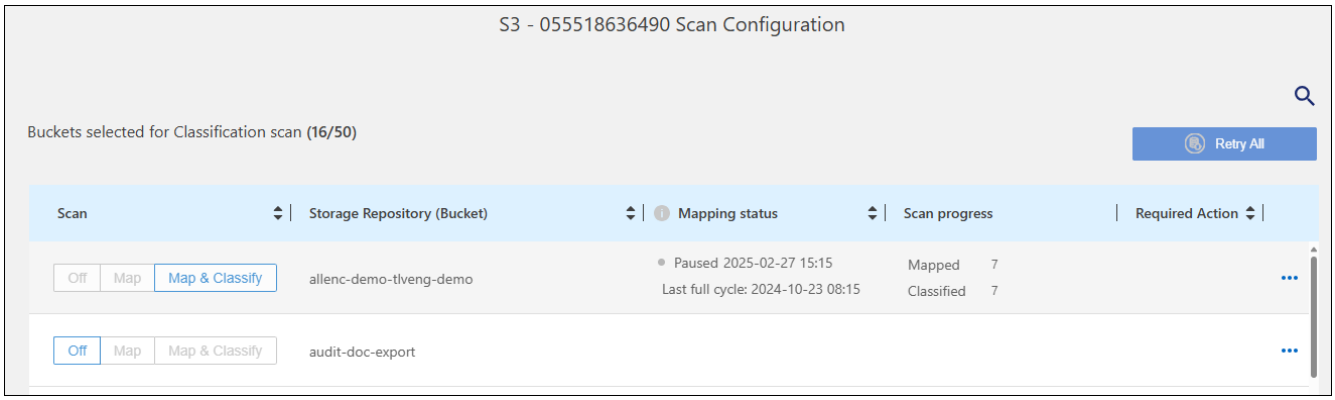
- 初始化（淺藍色點）：地圖或分類配置已啟動。此狀態會短暫顯示，然後過渡到「待處理佇列」狀態。
- 待處理佇列（橘色圓點）：掃描任務正在等待列入掃描佇列。
- 已排隊（橘色圓點）：任務已成功新增至掃描佇列。當佇列中的捲輪到達時，系統將開始映射或分類該磁碟區。
- 正在運行（綠點）：佇列中的掃描任務正在選定的儲存庫上積極進行。
- 完成（綠點）：儲存庫掃描已完成。
- 已暫停（灰點）：您已暫停掃描。雖然系統中未顯示音量變化，但掃描結果仍可使用。
- 錯誤（紅點）：掃描無法完成，因為遇到了問題。如果您需要完成某項操作，錯誤將出現在「所需操作」列下的工具提示中。否則，系統將顯示“錯誤”狀態並嘗試恢復。完成後，狀態就會改變。
- 未掃描：選擇了「關閉」磁碟區配置，系統未掃描該磁碟區。

步驟

1. 從資料分類選單中，選擇*配置*。



2. 從配置標籤中，選擇系統的*配置*按鈕。
3. 在掃描配置頁面中，查看所有儲存庫的掃描設定。



4. 掃描期間，將遊標停留在「對應狀態」列中的進度條上，即可查看該儲存庫中待對應或分類的檔案數量。

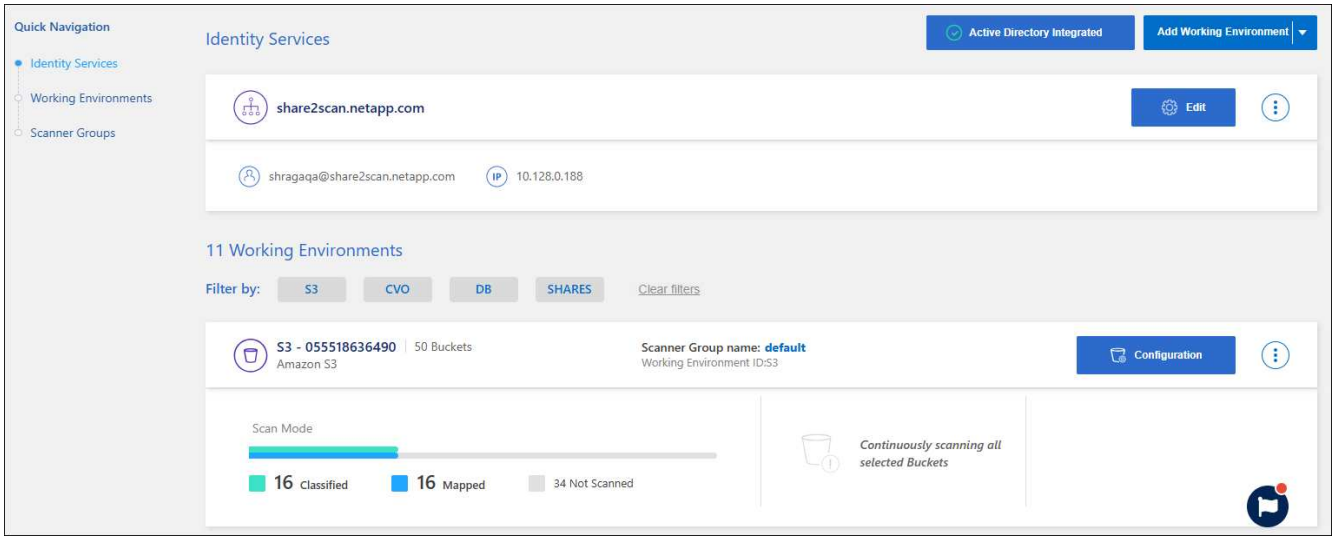
更改儲存庫的掃描類型

您可以隨時從設定頁面啟動或停止系統中的僅對應掃描或對應和分類掃描。您也可以從僅映射掃描變更為映射和分類掃描，反之亦然。

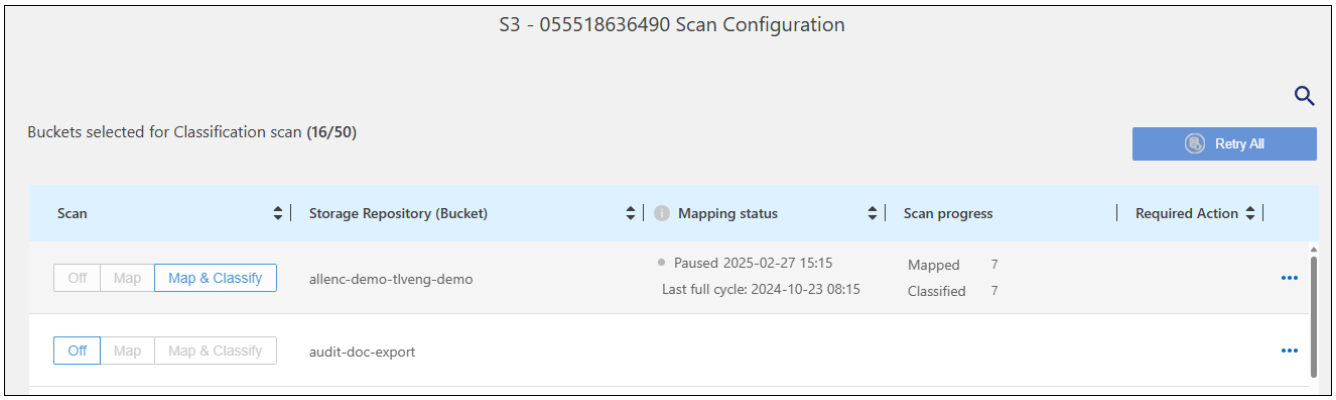
資料庫不能設定為僅映射掃描。資料庫掃描可以關閉或開啟；其中「開啟」相當於「映射和分類」。

步驟

1. 從資料分類選單中，選擇*配置*。
2. 從配置標籤中，選擇系統的*配置*按鈕。

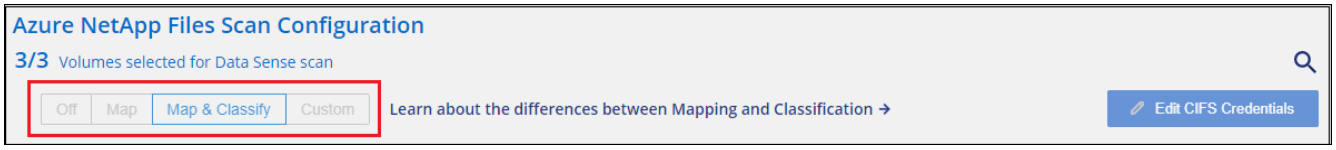


3. 在掃描設定頁面中，變更任何儲存庫（本例中為儲存桶）以執行*Map*或*Map & Classify*掃描。



某些類型的系統可讓您使用頁面頂部的按鈕列全域變更所有儲存庫的掃描類型。這對於Cloud Volumes ONTAP、本機ONTAP、 Azure NetApp Files和Amazon FSx for ONTAP系統有效。

下面的範例顯示了Azure NetApp Files系統的按鈕列。



優先掃描

您可以優先考慮最重要的僅映射掃描或映射和分類掃描，以確保高優先級掃描首先完成。

預設情況下，掃描會按照啟動的順序排隊。透過設定掃描優先權，您可以將掃描移至佇列的最前面。可以對多個掃描進行優先排序。優先權會依照先進先出的順序指定，這表示您優先考慮的第一個掃描將移至佇列的最前面；您優先考慮的第二個掃描將成為佇列中的第二個掃描，依此類推。

優先權是一次性授予的。映射資料的自動重新掃描按照預設順序進行。

步驟

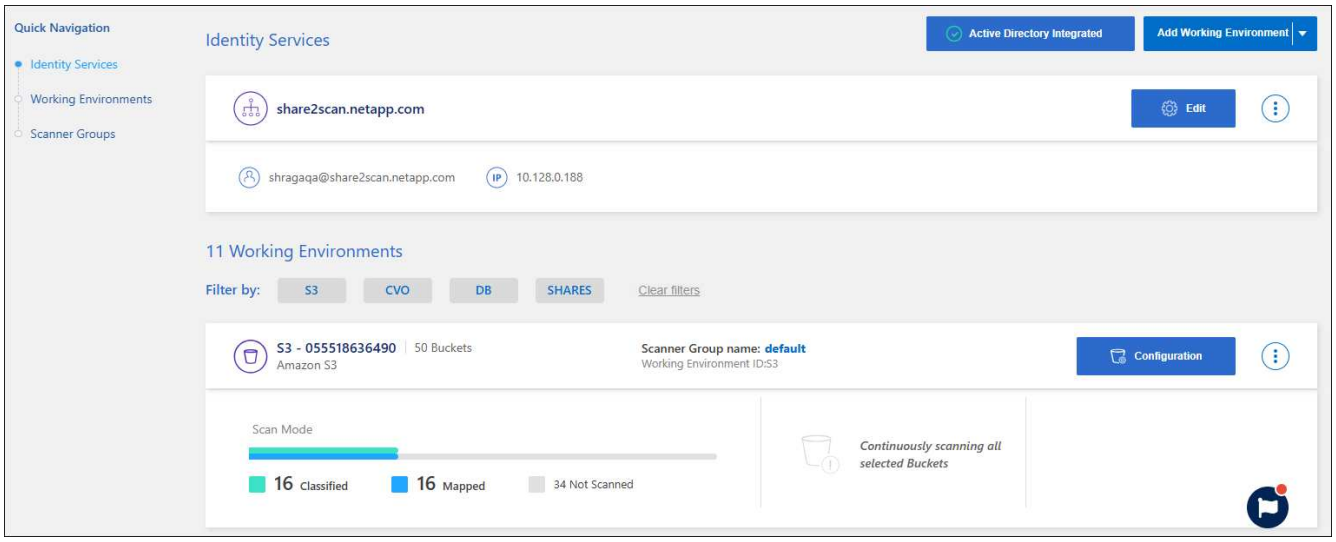
1. 從資料分類選單中，選擇*配置*。
2. 選擇您想要優先考慮的資源。
3. 從行動 `...`選項，選擇*優先掃描*。

停止掃描儲存庫

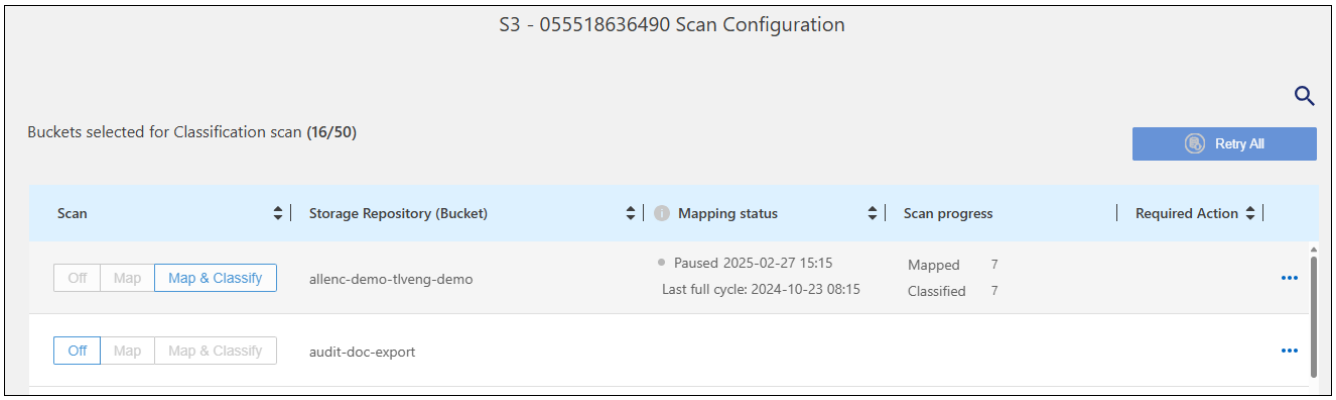
如果您不再需要監控儲存庫（例如磁碟區）的合規性，則可以停止掃描它。您可以透過關閉掃描來實現此目的。當掃描關閉時，有關該磁碟區的所有索引和資訊將從系統中刪除，並且掃描資料的收費也將停止。

步驟

1. 從資料分類選單中，選擇*配置*。
2. 從配置標籤中，選擇系統的*配置*按鈕。



3. 在掃描設定頁面中選擇「關閉」以停止掃描特定儲存桶。



暫停並恢復儲存庫掃描

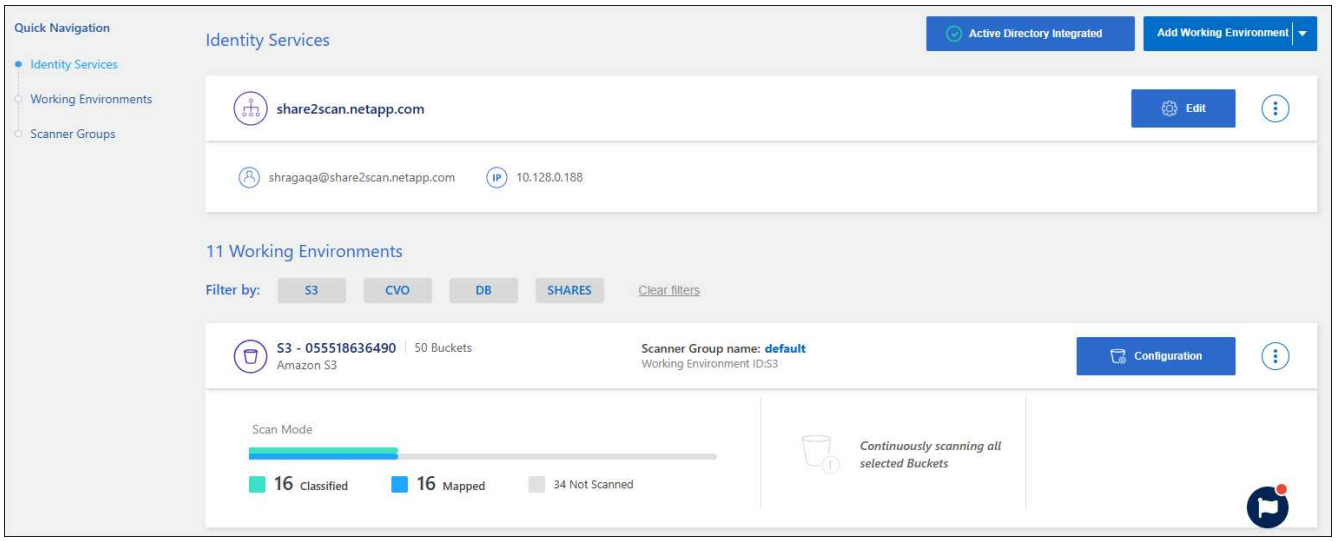
如果您想暫時停止掃描某些內容，您可以「暫停」儲存庫掃描。暫停掃描意味著資料分類將不再對儲存庫中的變更或新增執行任何未來的掃描。所有目前的掃描結果仍可在資料分類中查看。

即使暫停掃描，也不會免除計費費用，因為資料仍然保留在系統中。

您可以隨時恢復掃描。

步驟

1. 從資料分類選單中，選擇*配置*。
2. 從配置標籤中，選擇系統的*配置*按鈕。



3. 在掃描配置頁面中，選擇操作 ••• 圖示.
4. 選擇「暫停」暫停對磁碟區的掃描，或選擇「恢復」恢復先前已暫停的磁碟區的掃描。

查看NetApp Data Classification合規性報告

NetApp Data Classification提供報告，您可以使用這些報告來更好地了解組織的資料隱私計畫的狀態。

預設情況下，資料分類儀表板顯示所有系統、資料庫和資料來源的合規性和治理資料。如果您想要查看僅包含部分系統資料的報告，您可以進行篩選以僅查看這些系統的資料。

- 只有當您對資料來源執行完整分類掃描時，才可取得合規性報告。已進行僅映射掃描的資料來源只能產生資料映射報告。
- NetApp無法保證資料分類識別的個人資料和敏感個人資料 100% 的準確性。您應該始終透過查看數據來驗證資訊。

以下報告可用於資料分類：

- 資料發現評估報告：對掃描環境進行高級分析，以突出系統的發現並顯示關注領域和潛在的補救步驟。此報告可在治理儀表板中找到。
- 完整資料映射概覽報表：提供有關係統中檔案的大小和數量的資訊。這包括使用容量、資料年限、資料大小和檔案類型。此報告可在治理儀表板中找到。
- 資料主體存取請求報告：使您能夠提取包含有關資料主體的特定名稱或個人識別碼資訊的所有文件的報告。此報告可在合規性儀表板中找到。
- **HIPAA** 報告：幫助您識別文件中健康資訊的分佈。此報告可在合規性儀表板中找到。

- **PCI DSS** 報告：幫助您識別文件中信用卡資訊的分佈。此報告可在合規性儀表板中找到。
- 隱私風險評估報告：提供來自您的資料的隱私見解和隱私風險評分。此報告可在合規性儀表板中找到。
- 特定資訊類型的報告：可提供包含已識別文件（包含個人資料和敏感個人資料）詳細資訊的報告。您也可以查看按類別和文件類型細分的文件。

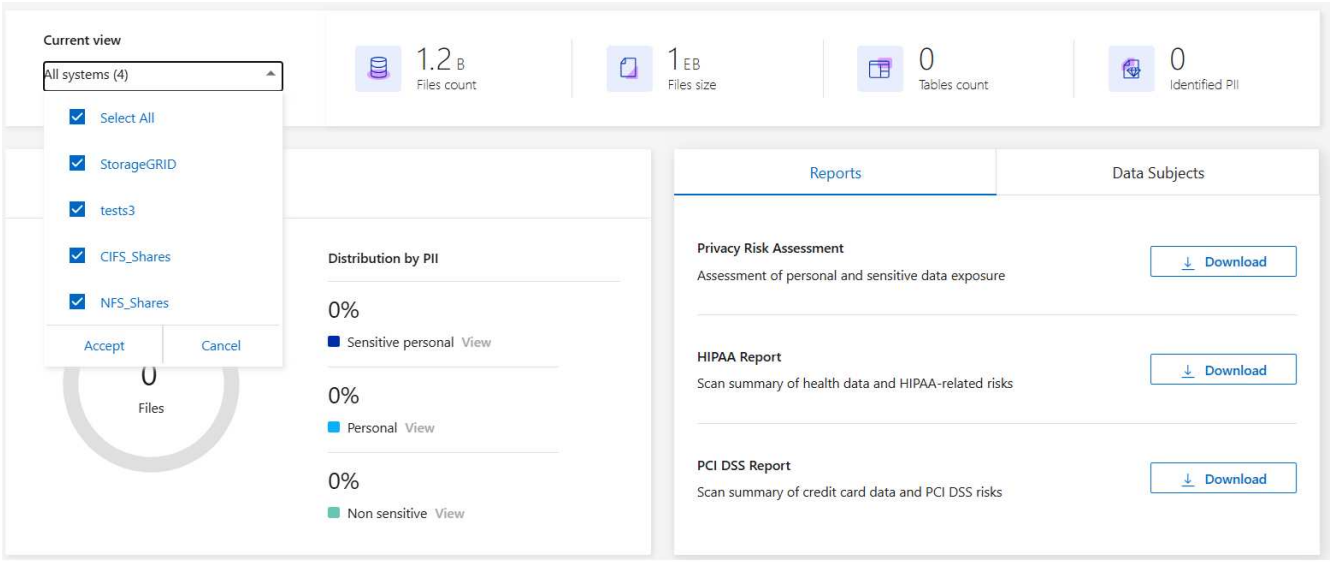
選擇報告系統

您可以過濾資料分類合規性儀表板的內容，以查看所有系統和資料庫的合規性數據，或僅查看特定係統的合規性數據。

當您篩選儀表板時，資料分類會將合規性資料和報表範圍限定到您選擇的系統。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 選擇系統過濾器下拉式選單，然後選擇系統。
3. 選擇接受來確認您的選擇。



資料主體存取請求報告

歐洲 GDPR 等隱私法規賦予資料主體（例如客戶或員工）存取其個人資料的權利。當資料主體請求此資訊時，這稱為 DSAR（資料主體存取請求）。各組織必須「毫不拖延」地回應這些請求，最晚不得超過收到請求後的一個月。

您可以透過搜尋主題的全名或已知識別碼（例如電子郵件地址）然後下載報告來回應 DSAR。該報告旨在幫助您的組織遵守 GDPR 或類似的資料隱私法。

*資料分類如何幫助您回應 DSAR？ *

當您執行資料主體搜尋時，資料分類會找到包含該人姓名或識別碼的所有檔案。資料分類檢查最新的預索引資料的名稱或識別碼。它不會啟動新的掃描。

搜尋完成後，您可以下載資料主體存取請求報告的檔案清單。該報告匯總了數據中的見解，並將其轉化為法律術

語，以便您可以將其發送給相關人員。

目前資料庫不支援資料主體搜尋。

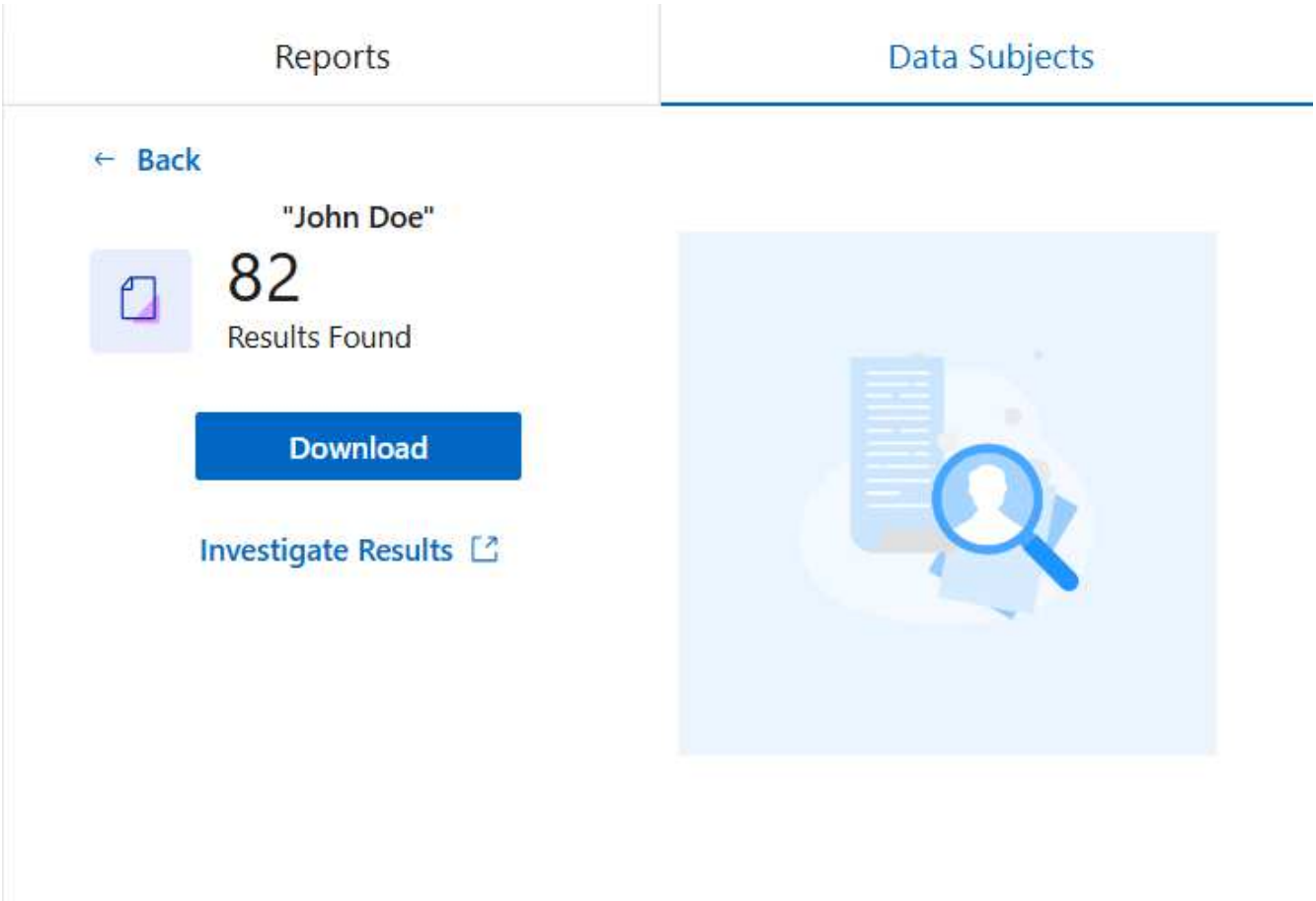
搜尋資料主體並下載報告

搜尋資料主體的全名或已知標識符，然後下載檔案清單報表或 DSAR 報表。您可以透過以下方式搜尋"任何個人資訊類型"。

搜尋資料主體的姓名時支援英語、德語、日語和西班牙語。稍後將添加對更多語言的支援。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 在合規性頁面中，找到資料主體選項卡。
3. 在「資料主體」部分，輸入名稱或已知標識符，然後選擇「搜尋」。
4. 搜尋完成後，選擇下載以存取資料主體存取請求回應。選擇調查結果以在資料調查頁面中查看更多資訊。



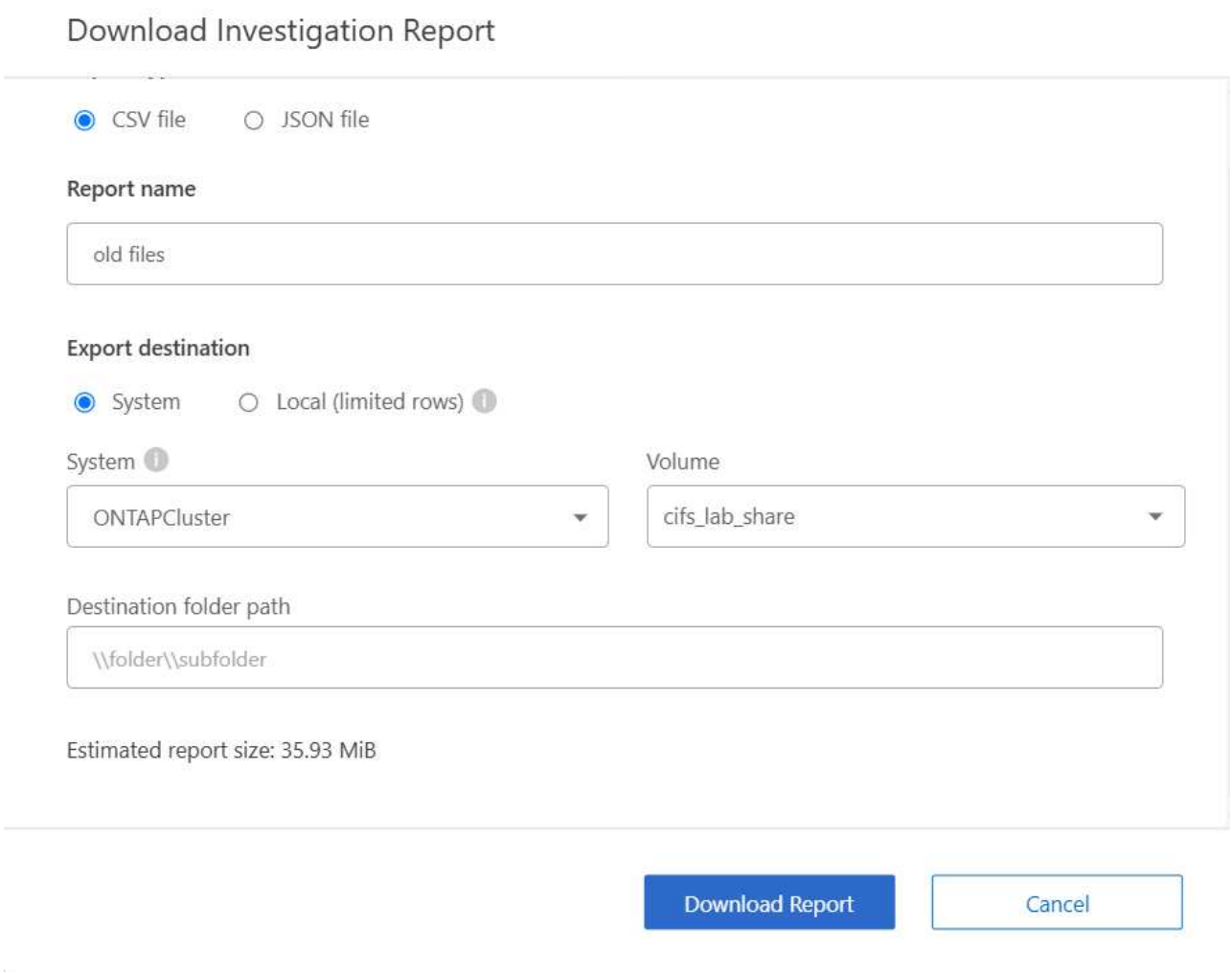
5. 查看資料分類中的結果或透過選擇下載圖示將其下載為報告。
   a. 選擇下載圖示後，配置您的下載設定：
      - 選擇影片格式：CSV 或 JSON
      - 輸入*報告名稱*

- 選擇匯出目的地：*系統*或您的*本機*機器。

  如果您選擇系統，則會下載所有資料。您還必須選擇*系統*、磁碟區*和*目標資料夾路徑。

如果您選擇*本地*，則會將報告限制為前 10,000 行非結構化資料；5,000 行非結構化資料和 1,000 行結構化資料。

a. 選擇下載報告開始下載。



健康保險流通與責任法案（HIPAA）報告

健康保險流通與責任法案 (HIPAA) 報告可以幫助您識別包含健康資訊的文件。它旨在幫助您的組織遵守 HIPAA 資料隱私法的要求。資料分類尋找的資訊包括：

- 健康參考模式
- ICD-10-CM 醫療代碼
- ICD-9-CM 醫療代碼
- HR - 健康類別
- 健康應用資料類別

該報告包含以下資訊：

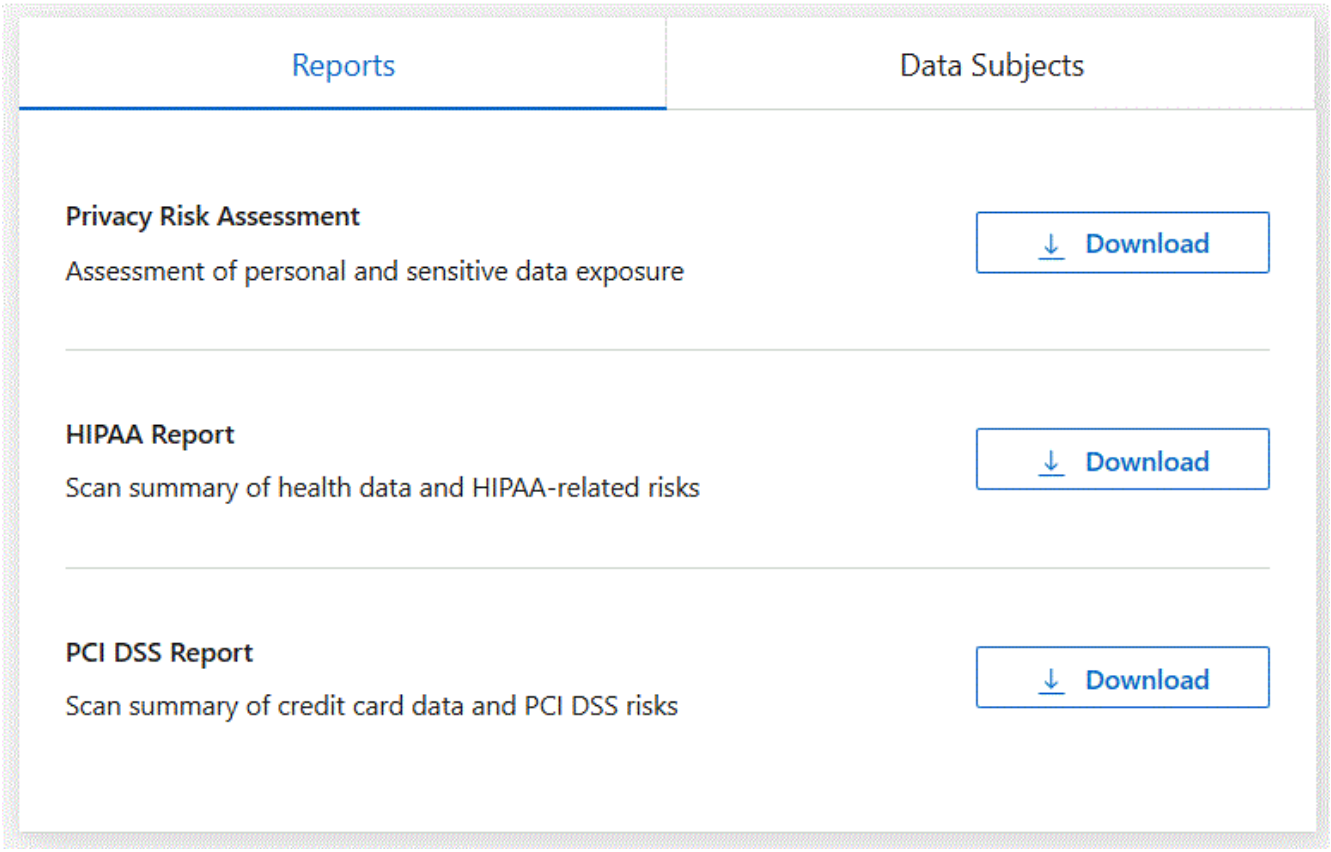
- 概述：有多少文件包含健康資訊以及在哪些系統中。
- 加密：加密或未加密系統中包含健康資訊的檔案的百分比。此資訊特定於Cloud Volumes ONTAP。
- 勒索軟體防護：在啟用或未啟用勒索軟體防護的系統上，包含健康資訊的檔案的百分比。此資訊特定於Cloud Volumes ONTAP。
- 保留：文件最後修改的時間範圍。這很有用，因為您不應該將健康資訊保存超過處理所需的時間。
- 健康資訊分發：發現健康資訊的系統以及是否啟用了加密和勒索軟體保護。

產生 **HIPAA** 報告

轉到“合規性”選項卡以產生報告。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 找到報告窗格。選擇*HIPAA 報告*旁邊的下載圖示。



結果

資料分類產生 PDF 報告。

支付卡產業資料安全標準 (PCI DSS) 報告

支付卡產業資料安全標準 (PCI DSS) 報告可以幫助您識別文件中信用卡資訊的分佈。

該報告包含以下資訊：

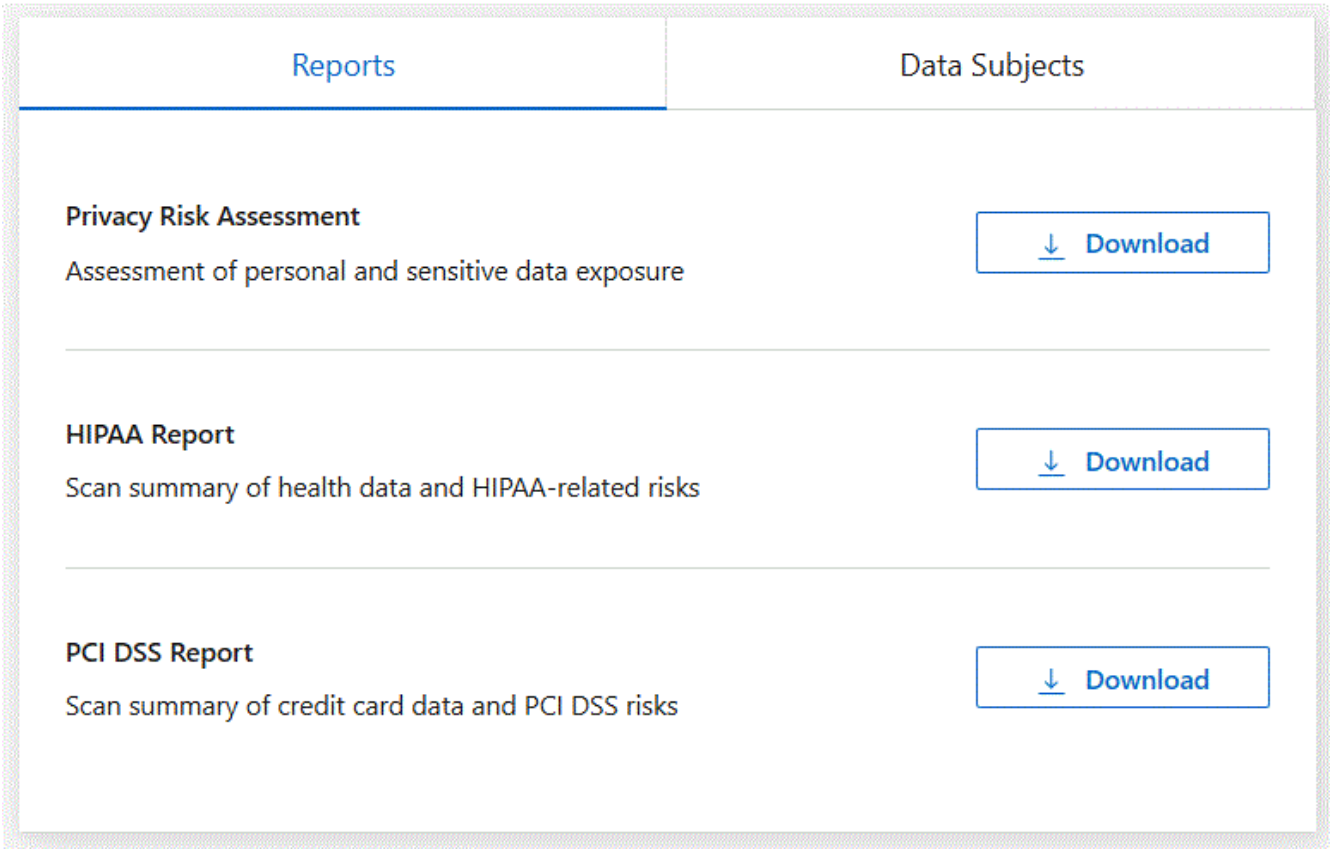
- 概述：有多少文件包含信用卡資訊以及在哪些系統中。
- 加密：加密或未加密系統中包含信用卡資訊的檔案的百分比。此資訊特定於Cloud Volumes ONTAP。
- 勒索軟體保護：在啟用或未啟用勒索軟體保護的系統上，包含信用卡資訊的檔案的百分比。此資訊特定於Cloud Volumes ONTAP。
- 保留：文件最後修改的時間範圍。這很有用，因為您不應該將信用卡資訊保存的時間超過處理所需的時間。
- 信用卡資訊分發：發現信用卡資訊的系統以及是否啟用了加密和勒索軟體保護。

產生 **PCI DSS** 報告

轉到"合規性"選項卡以產生報告。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 找到報告窗格。選擇*PCI DSS 報表*旁邊的下載圖示。



結果

資料分類會產生一份 PDF 報告，您可以根據需要查看並傳送給其他群組。

隱私風險評估報告

隱私權風險評估報告概述了您組織的隱私權風險狀況，這是 GDPR 和 CCPA 等隱私權法規所要求的。

該報告包含以下資訊：

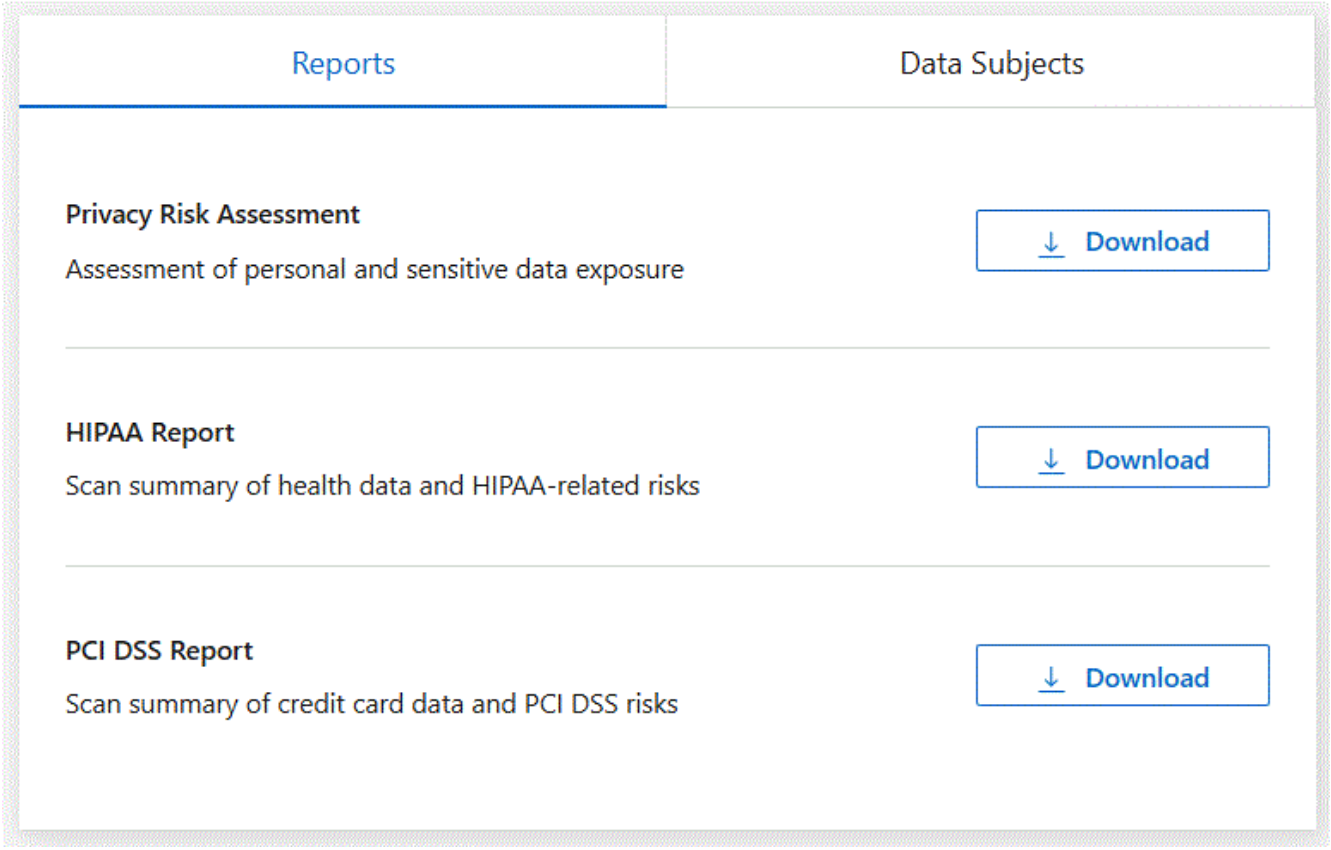
- 合規狀態：嚴重性評分和資料分佈，無論是非敏感資料、個人資料或敏感個人資料。
- 評估概述：發現的個人資料類型以及資料類別的細分。
- 本次評估中的資料主體：按地點劃分的已找到國家識別碼的人數。

產生隱私權風險評估報告

轉到“合規性”選項卡以產生報告。

步驟

1. 從資料分類選單中，選擇*合規性*。
2. 找到報告窗格。選擇*隱私風險評估報告*旁邊的下載圖示。



結果

資料分類會產生一份 PDF 報告，您可以根據需要查看並傳送給其他群組。

嚴重程度評分

資料分類根據三個變數計算隱私風險評估報告的嚴重性分數：

- 個人資料佔所有資料的百分比。
- 敏感個人資料佔所有資料的比例。
- 包含資料主體的文件百分比，由國家識別碼（例如國民身分證、社會安全號碼和稅號）決定。

決定分數的邏輯如下：

嚴重程度評分	邏輯
0	所有三個變數都恰好為 0%
1	其中一個變數大於 0%
2	其中一個變數大於3%
3	其中兩個變數大於 3%
4	其中三個變數大於 3%
5	其中一個變數大於6%
6	其中兩個變數大於 6%
7	其中三個變數大於 6%
8	其中一個變數大於15%
9	其中兩個變數大於 15%
10	其中三個變數大於 15%

監控NetApp Data Classification的運作狀況

NetApp Data Classification健康監視器儀表板提供即時監控和效能洞察。健康監視器會捕獲有關您的資料分類基礎架構、系統運行狀況、使用指標和利用率資料的信息，使您能夠識別和解決問題。

健康監測洞察

健康監測儀錶板以四類資訊呈現資訊。

- 基礎設施狀況

  查看版本狀態、系統穩定性、部署類型和機器規模等資訊。

- 問題容器

  查看「問題容器」字段，以了解哪些容器已停止或頻繁重新啟動。利用這些資訊調查具體的容器。

- 系統資訊

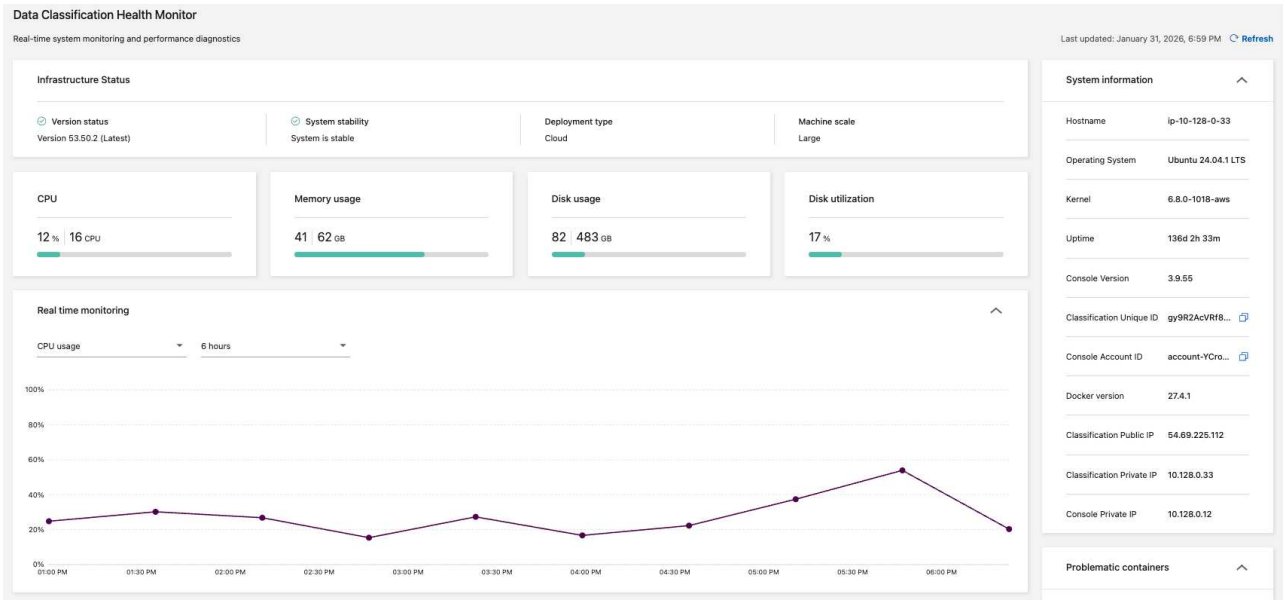
  系統資訊面板會擷取有關NetApp Console和資料分類的關鍵訊息，例如公用和私人 IP 位址、主機名稱、作業系統、控制台版本和控制台 ID。

- 用途和使用方法

  查看 CPU 使用率、磁碟使用率、磁碟使用率和記憶體使用率。這些值以儲存單位（GB）或總使用量的百分比顯示。如果任何欄位顯示警告，請選擇該警告以取得相關資訊和補救建議。

訪問健康監測儀表板

1. 在資料分類中，選擇配置。
2. 在「配置」標題下，選擇「資料分類運作狀況監視器」。
3. 在健康監測儀表板中，您可以：
   - 審查使用情況和利用情況。如果任何使用情況或使用率指標顯示警告，請選擇該警告以取得解決問題的建議。
   - 切換圖表以顯示 CPU 使用率、磁碟使用率、磁碟使用率和記憶體使用率。您可以變更 x 軸，以按小時（6、12 或 24 小時）或天（2、7 或 14 天）顯示內容。
   - 刷新儀錶板以查看最新資料指標。



管理資料分類

從NetApp Data Classification掃描排除特定目錄

如果您希望NetApp Data Classification從掃描中排除特定目錄，則可以將這些目錄名稱新增至設定檔。套用此變更後，資料分類引擎將從掃描中排除這些目錄。

預設情況下，資料分類掃描排除與磁碟區中的來源相同的磁碟區快照資料。

支援的資料來源

以下資料來源中的 NFS 和 CIFS 共用支援從資料分類掃描中排除特定目錄：

- 本地ONTAP
- Cloud Volumes ONTAP
- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- 常規文件共享

定義要排除在掃描之外的目錄

在將目錄排除在分類掃描之外之前，您需要登入資料分類系統，以便可以編輯設定檔並執行腳本。了解如何"登入資料分類系統"取決於您是否在 Linux 機器上手動安裝了該軟體，或者是否在雲端部署了該實例。

注意事項

- 每個資料分類系統最多可以排除 50 個目錄路徑。
- 排除目錄路徑可能會影響掃描時間。

步驟

1. 在資料分類系統上，前往“/opt/netapp/config/custom_configuration”，然後開啟文件 `data_provider.yaml`。
2. 在「exclude:」行下的「data_providers」部分中，輸入要排除的目錄路徑。例如：

```
exclude:
- "folder1"
- "folder2"
```

   請勿修改此文件中的任何其他內容。

3. 儲存對文件的變更。
4. 前往“/opt/netapp/Datasense/tools/customer_configuration/data_providers”並執行以下腳本：

```
update_data_providers_from_config_file.sh
```

+ 此指令將要排除在掃描範圍之外的目錄提交給分類引擎。

結果

對您的資料進行的所有後續掃描都將排除對這些指定目錄的掃描。

您可以使用相同的步驟從排除清單中新增、編輯或刪除項目。執行腳本提交更改後，修改後的排除清單將會更新。

範例

配置**1**：

名稱中包含「folder1」的每個資料夾都將被排除在所有資料來源之外。

```
data_providers:
   exclude:
   - "folder1"
```

將被排除的路徑的預期結果：

- /CVO1/資料夾1
- /CVO1/資料夾1名稱
- /CVO1/資料夾10
- /CVO1/*資料夾1
- /CVO1/+資料夾1名稱
- /CVO1/notfolder10
- /CVO22/資料夾1
- /CVO22/資料夾1名稱
- /CVO22/資料夾10

不會被排除的路徑範例：

- /CVO1/*資料夾
- /CVO1/資料夾名稱
- /CVO22/*folder20

配置**2**：

僅在名稱開頭包含“*folder1”的每個資料夾都將被排除。

```
data_providers:
   exclude:
   - "\\*folder1"
```

將被排除的路徑的預期結果：

- /CVO/*資料夾1
- /CVO/*資料夾1名稱
- /CVO/*資料夾10

不會被排除的路徑範例：

- /CVO/資料夾1
- /CVO/資料夾1名稱
- /CVO/not*folder10

配置3：

資料來源「CVO22」中名稱中包含「folder1」的每個資料夾都將被排除。

```
data_providers:
   exclude:
   - "CVO22/folder1"
```

將被排除的路徑的預期結果：

- /CVO22/資料夾1
- /CVO22/資料夾1名稱
- /CVO22/資料夾10

不會被排除的路徑範例：

- /CVO1/資料夾1
- /CVO1/資料夾1名稱
- /CVO1/資料夾10

轉義資料夾名稱中的特殊字符

如果您的資料夾名稱包含以下特殊字元之一，並且您想要排除該資料夾中的資料進行掃描，則需要在資料夾名稱前使用轉義序列 \\。

```
 ., +, *, ?, ^, $, (, ), [, ], {, }, |
例如：
```

來源中的路徑： `/project/*not_to_scan`

排除文件中的語法： `"\\*not_to_scan"`

查看目前排除列表

內容可能 `data_provider.yaml`設定檔與運行後實際提交的內容不同 `update_data_providers_from_config_file.sh`腳本。若要查看已從資料分類掃描中排除的目前目錄列表，請從「/opt/netapp/Datasense/tools/customer_configuration/data_providers」執行下列命令：

```
get_data_providers_configuration.sh
```

在NetApp Data Classification其他群組 ID 定義為對組織開放

當群組 ID (GID) 附加到 NFS 檔案共用中的檔案或資料夾時，它們定義了檔案或資料夾的權限；例如它們是否「對組織開放」。如果某些 GID 最初未設定「開放給組織」權限級別，您可以向 GID 新增該權限，以便任何附加了該 GID 的檔案和資料夾都將被視為「開放給組織」。

在您進行此變更並且NetApp Data Classification重新掃描您的檔案和資料夾後，任何附加了這些群組 ID 的檔案和資料夾都將在「調查詳細資料」頁面中顯示此權限，並且它們也會出現在您顯示檔案權限的報表中。

要啟動此功能，您需要登入資料分類系統，以便可以編輯設定檔並執行腳本。了解如何"登入資料分類系統"取決於您是否在 Linux 機器上手動安裝了該軟體，或者是否在雲端部署了該實例。

為群組 ID 新增「向組織開放」權限

在開始此任務之前，您需要有群組 ID 號碼 (GID)。

步驟

1. 在資料分類系統上，前往“/opt/netapp/config/custom_configuration”並開啟文件 `data_provider.yaml`。
2. 在「organization_group_ids: []」行中新增群組 ID。例如：

   ```
   organization_group_ids: [1014, 1015, 21, 2021, 1013, 2020, 1018, 1019]
   ```

   請勿更改此文件中的任何其他內容。

3. 儲存對文件的變更。
4. 前往“/opt/netapp/Datasense/tools/customer_configuration/data_providers”並執行以下腳本：

   ```
   update_data_providers_from_config_file.sh
   ```

   此命令將修改後的群組 ID 權限提交給分類引擎。

結果

對您的資料進行的所有後續掃描都將識別出附加有這些群組 ID 且標記為「向組織開放」的檔案或資料夾。

您可以使用相同的步驟編輯群組 ID 清單並刪除過去新增的任何群組 ID。執行腳本提交變更後，修改後的群組 ID 清單將會更新。

查看目前群組ID列表

內容可能 `data_provider.yaml`設定檔與運行後實際提交的內容不同

`update_data_providers_from_config_file.sh`腳本。若要查看已新增至資料分類的目前群組 ID 列表，請從「/opt/netapp/Datasense/tools/customer_configuration/data_providers」執行下列命令：

```
get_data_providers_configuration.sh
```

在NetApp Data Classification中自訂過期資料定義

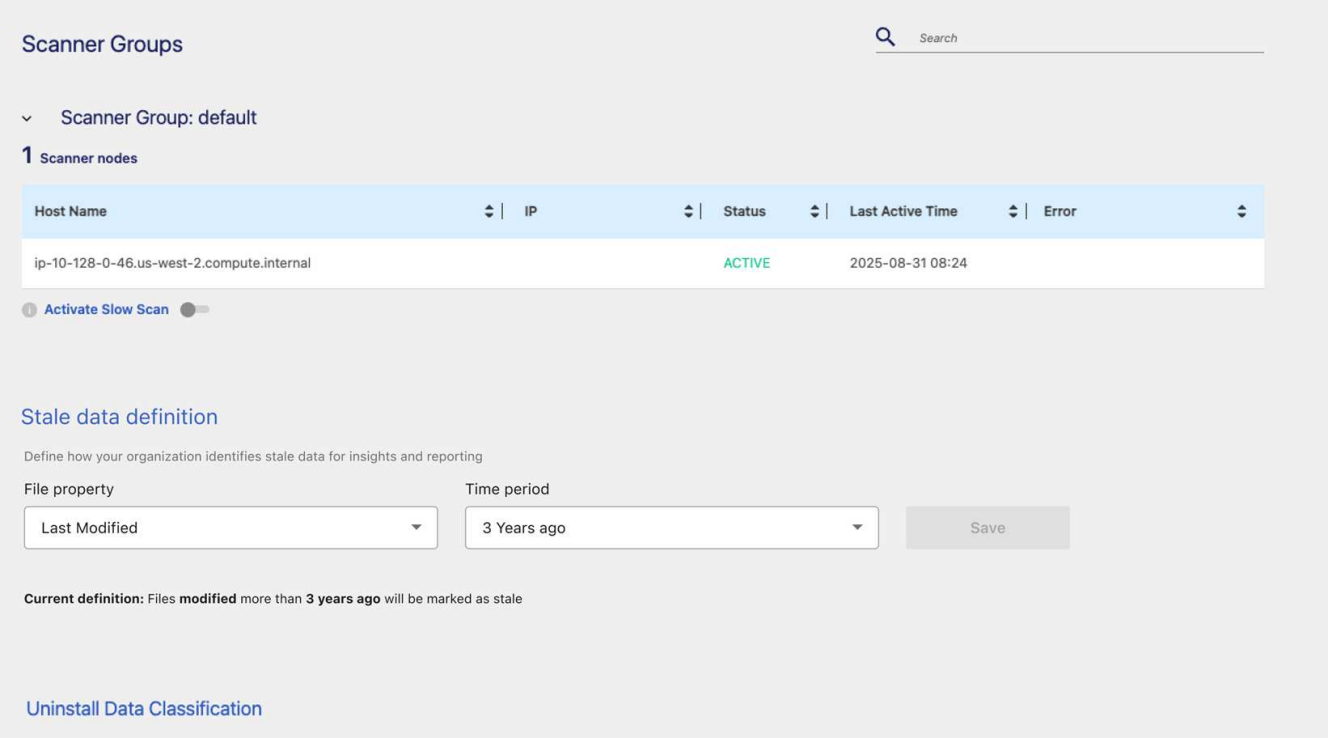
NetApp Data Classification可識別過時數據，協助您發現節省成本的機會和治理風險。由於不同組織環境中對過時資料的定義可能有所不同，因此您可以自訂資料分類如何定義過時資料。

過期資料可以根據其上次造訪時間或上次修改時間來定義。時段選擇範圍從 6 個月前到 10 年前。

預設情況下，如果資料上次修改時間距今已有三年，則該資料視為過期資料。

定義過期數據

1. 在勒索軟體復原能力中，選擇配置。
2. 在「配置」頁面中，捲動至「過期資料定義」標題。
3. 在「檔案屬性」下拉式功能表中，選擇是否要根據檔案的「上次存取時間」或「上次修改時間」來定義過期資料。
4. 選擇過期資料定義的時間段。



5. 選擇儲存。

從NetApp Data Classification中刪除資料來源

如果需要，您可以停止NetApp Data Classification掃描一個或多個系統、資料庫或檔案共用群組。

停用系統掃描

當您停用掃描時，資料分類不再掃描系統上的數據，並且它會從資料分類實例中刪除索引的見解。系統本身的資料不會被刪除。

1. 從「配置」頁面中，選擇 系統行中的按鈕，然後*停用資料分類*。

您也可以在選擇系統時從「服務」面板停用系統掃描。

從資料分類中刪除資料庫

如果您不再需要掃描某個資料庫，您可以從資料分類介面將其刪除並停止所有掃描。

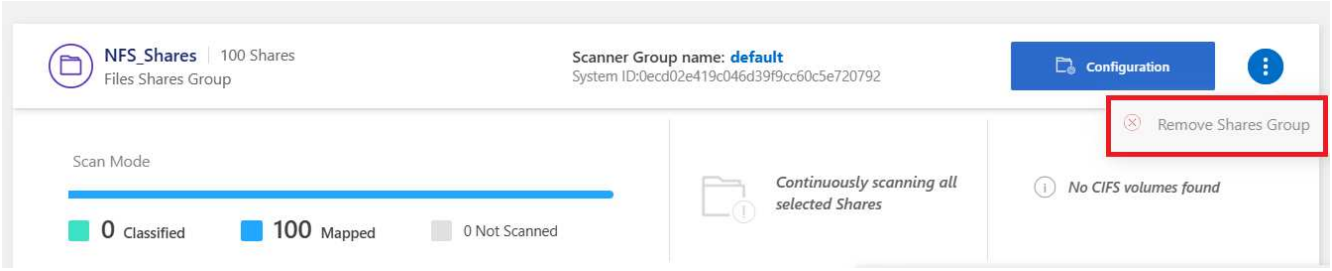
1. 從「配置」頁面中，選擇 資料庫行中的按鈕，然後*刪除資料庫伺服器*。

從資料分類中刪除一組檔案共享

如果您不再想從文件共享群組掃描使用者文件，您可以從資料分類介面刪除文件共用群組並停止所有掃描。

步驟

1. 從「配置」頁面中，選擇 檔案共用群組行中的按鈕，然後按一下*刪除檔案共用群組*。



2. 從確認對話方塊中選擇*刪除共享群組*。

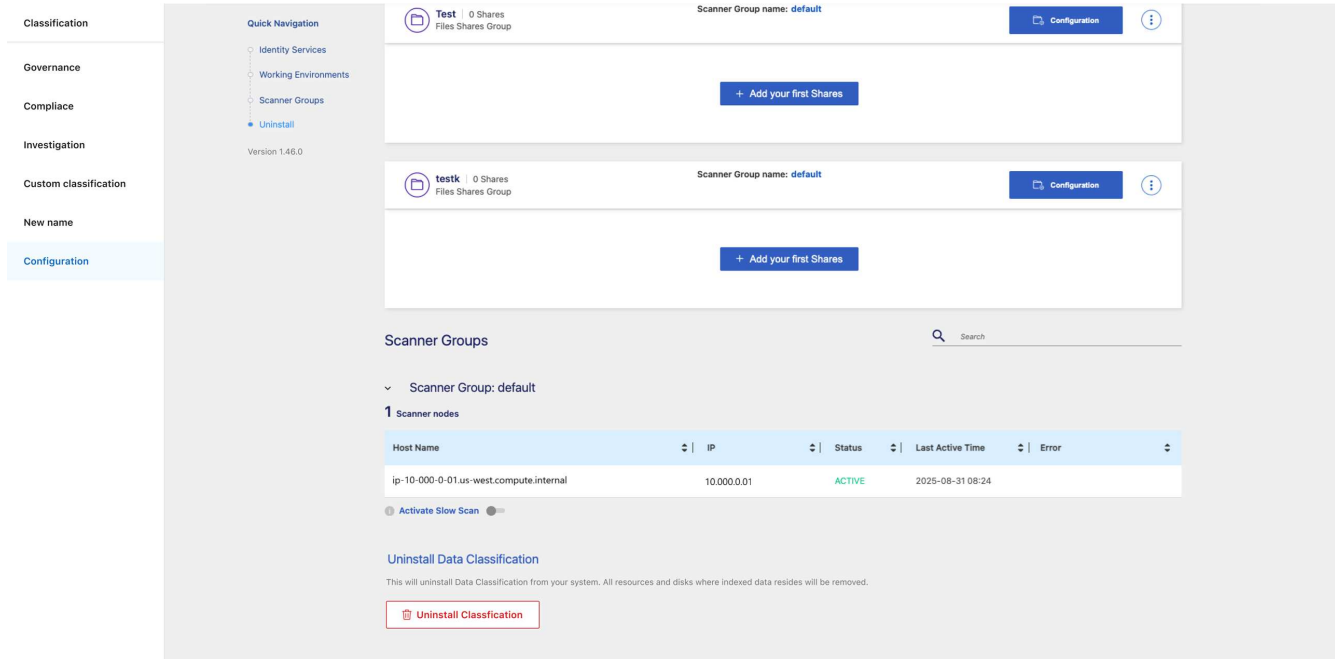
卸載NetApp Data Classification

您可以解除安裝NetApp Data Classification來解決問題或永久從主機移除該軟體。刪除實例也會刪除索引資料所在的關聯磁碟，這表示資料分類掃描的所有資訊都將永久刪除。

您需要使用的步驟取決於您是在雲端還是在本機上部署資料分類。

從雲端提供者卸載資料分類

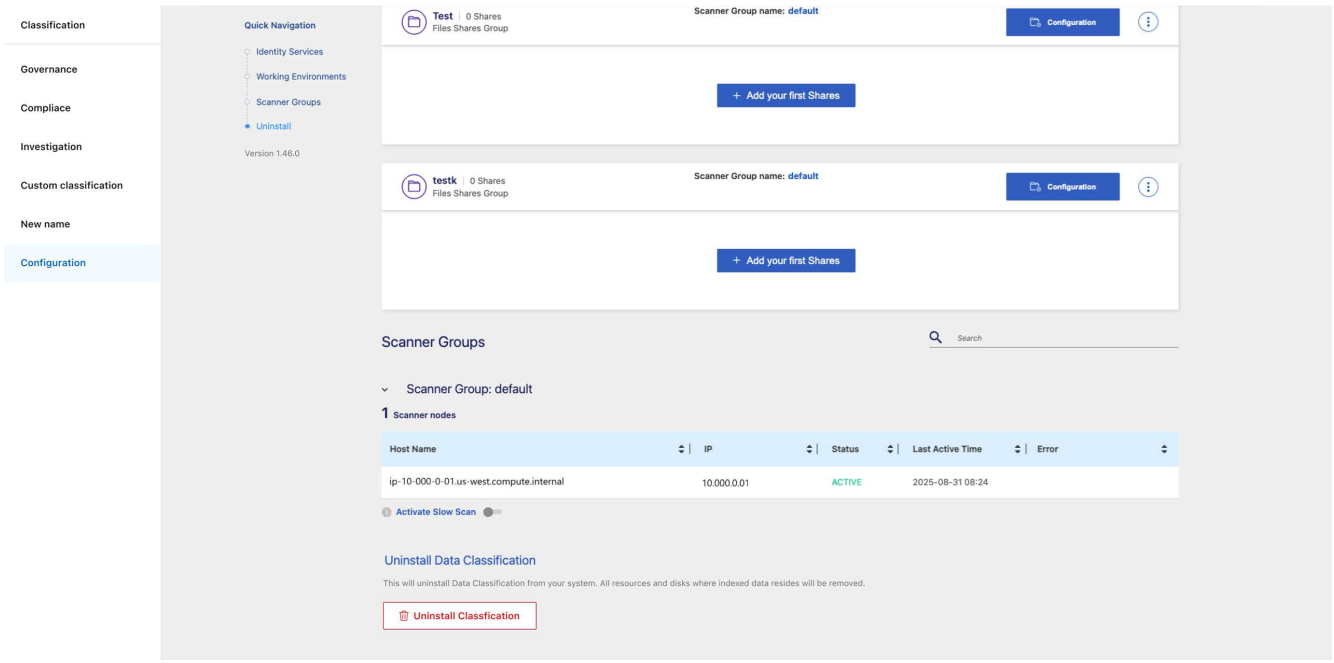
1. 從資料分類中，選擇配置。
2. 在配置頁面底部，選擇卸載分類。



3. 在對話方塊中，輸入「卸載」以繼續中斷資料分類實例與控制台代理程式的連線。選擇卸載進行確認。
4. 在「卸載分類」對話方塊中，鍵入「卸載」以確認您要中斷資料分類執行個體與控制台代理程式的連接，然後選擇「卸載」。
5. 要完成卸載過程，請前往雲端提供者的控制台並刪除資料分類實例。此實例名為 *CloudCompliance*，並帶有與之連接的產生的雜湊值（UUID）。例如：*CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

從本地部署卸載資料分類

1. 從資料分類中，選擇配置。
2. 在配置頁面底部，選擇卸載分類。



3. 在對話方塊中，輸入「卸載」以繼續中斷資料分類實例與控制台代理程式的連線。選擇卸載進行確認。
4. 若要從主機卸載軟體，請執行 `cleanup.sh`資料分類主機上的腳本，例如：

```
cleanup.sh
```

該腳本位於 `/install/light_probe/onprem_installer/cleanup.sh`目錄。了解如何"登入資料分類主機"。

參考

支援的NetApp Data Classification實例類型

NetApp Data Classification軟體必須在符合特定作業系統需求、RAM 需求、軟體需求等的主機上運作。在雲端部署資料分類時，我們建議您使用具有「大型」特性的系統以實現全部功能。

您可以在較少 CPU 和較少 RAM 的系統上部署資料分類，但在使用這些功能較弱的系統時會受到一些限制。"了解這些限制"。

在下表中，如果標記為「預設」的系統在您安裝資料分類的區域中不可用，則將部署表中的下一個系統。

AWS 執行個體類型

系統大小	規格	實例類型
特大號	32 個 CPU、128 GB RAM、1 TiB gp3 SSD	"m6i.8xlarge"(預設)
大的	16 個 CPU、64 GB RAM、500 GiB SSD	"m6i.4xlarge"（預設）m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
中等的	8 個 CPU、32 GB RAM、200 GiB SSD	"m6i.2xlarge"（預設）m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
小的	8 個 CPU、16 GB RAM、100 GiB SSD	"c6a.2xlarge"（預設）c5a.2xlarge c5.2xlarge c4.2xlarge

Azure 執行個體類型

系統大小	規格	實例類型
特大號	32 個 CPU、128 GB RAM、OS 磁碟（2,048 GiB，最小吞吐量 250 MB/s）和資料磁碟（1 TiB SSD，最小吞吐量 750 MB/s）	"Standard_D32_v3"(預設)
大的	16 個 CPU、64 GB RAM、500 GiB SSD	"Standard_D16s_v3"(預設)

GCP 實例類型

系統大小	規格	實例類型
大的	16 個 CPU、64 GB RAM、500 GiB SSD	"n2-標準-16"（預設）n2d-standard-16 n1-standard-16

從NetApp Data Classification中的資料來源收集的元數據

NetApp Data Classification在對來自資料來源和系統的資料執行分類掃描時收集某些元資料。資料分類可以存取我們對您的資料進行分類所需的大部分元數據，但有些來源我們無法存取我們需要的資料。

	元資料	CIFS	NFS
時間戳	建立時間	可用的	不可用（Linux 不支援）
	上次造訪時間	可用的	可用的
	上次修改時間	可用的	可用的
權限	開放權限	如果“EVERYONE”群組有權存取該文件，則視為“對組織開放”	如果“其他人”有權存取該文件，則視為“對組織開放”
	使用者/群組存取	使用者和群組資訊取自 LDAP	不可用（NFS 使用者通常在伺服器本地進行管理，因此同一個人在每個伺服器上可以有不同的 UID）

- 資料分類不會從資料庫資料來源提取「上次存取時間」。
- 舊版的 Windows 作業系統（例如 Windows 7 和 Windows 8）預設為停用「上次存取時間」屬性的收集，因為它會影響系統效能。當未收集此屬性時，基於「上次造訪時間」的資料分類分析將受到影響。如果需要，您可以在這些較舊的 Windows 系統上啟用上次存取時間的收集。

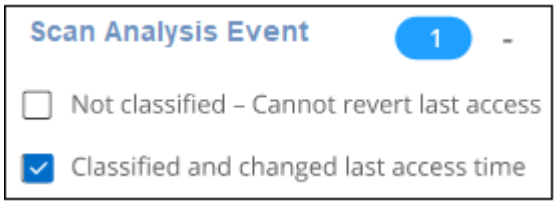
上次訪問時間時間戳

當資料分類從檔案共用中提取資料時，作業系統會將其視為存取數據，並相應地變更「上次存取時間」。掃描後，資料分類會嘗試將上次存取時間還原為原始時間戳記。如果資料分類在 CIFS 中沒有寫入屬性權限，或在 NFS 中沒有寫入權限，則係統無法將上次存取時間還原為原始時間戳記。配置了SnapLock的ONTAP磁碟區具有唯讀權限，並且無法將上次存取時間還原為原始時間戳記。

預設情況下，如果資料分類沒有這些權限，系統將不會掃描磁碟區中的這些文件，因為資料分類無法將「上次存取時間」恢復為原始時間戳記。但是，如果您不介意文件中的最後訪問時間是否重置為原始時間，則可以選擇配置頁面底部的*缺少“寫入屬性”權限時掃描*開關，以便資料分類無論權限如何都會掃描卷。

此功能適用於本機ONTAP系統、 Cloud Volumes ONTAP、 Azure NetApp Files、 Amazon FSx for NetApp ONTAP管理和第三方檔案共用。

調查頁面中有一個名為「掃描分析事件」的過濾器，它使您能夠顯示未分類的文件（因為資料分類無法恢復上次訪問時間），或者即使資料分類無法恢復上次訪問時間也已分類的文件。



過濾器選擇包括：

- 「未分類 - 無法恢復上次存取時間」 - 這顯示由於缺少寫入權限而未分類的檔案。
- 「已分類並更新的上次存取時間」 - 這顯示已分類的文件，且資料分類無法將上次存取時間重設為原始日期。此篩選器僅與您開啟*缺少「寫入屬性」權限時掃描*的環境相關。

如果需要，您可以將這些結果匯出到報告中，以便查看哪些文件因權限原因而掃描，哪些文件未被掃描。"了解有關數據調查報告的更多信息"。

登入NetApp Data Classification系統

您需要登入NetApp Data Classification系統，以便存取日誌檔案或編輯設定檔。

當資料分類安裝在您本地的 Linux 機器上或在雲端部署的 Linux 機器上時，您可以直接存取設定檔和腳本。

當資料分類部署在雲端時，您需要透過 SSH 連線到資料分類實例。您可以透過輸入使用者和密碼，或使用控制台代理安裝期間提供的 SSH 金鑰，透過 SSH 連線到系統。 SSH 指令是：

```
ssh -i <path_to_the_ssh_key> <machine_user>@<datasense_ip>
```

- `<path_to_the_ssh_key>`= ssh 驗證金鑰的位置
- `<machine_user>`:
  - 對於 AWS：使用 <ec2-user>
  - 對於 Azure：使用為控制台執行個體建立的用戶
  - 對於 GCP：使用為控制台實例建立的用戶
- `<datasense_ip>`= 虛擬機器實例的 IP 位址

需要修改安全性群組入站規則才能存取雲端的系統。有關詳細信息，請參閱：

- "AWS 中的安全群組規則"
- "Azure 中的安全性群組規則"
- "Google Cloud 中的防火牆規則"

NetApp Data ClassificationAPI

透過 Web UI 提供的NetApp Data Classification功能也可透過 REST API 提供。

資料分類中定義了四個類別，與 UI 中的選項卡相對應：

- 調查
- 遵守
- 治理
- 配置

Swagger 文件中的 API 可讓您搜尋、聚合資料、追蹤掃描以及執行複製、移動和刪除等操作。

概況

該 API 使您能夠執行以下功能：

- 匯出訊息
  - UI 中可用的所有內容都可以透過 API 匯出（報告除外）
  - 數據以 JSON 格式導出（易於解析並推送到第三方應用程序，如 Splunk）
- 使用“AND”和“OR”語句建立查詢，包括和排除資訊等等。

  例如，您可以找到沒有特定個人識別資訊 (PII) 的檔案（UI 中不可用的功能）。您也可以排除導出操作的特定欄位。
- 執行操作
  - 更新 CIFS 憑證
  - 查看和取消操作
  - 重新掃描目錄
  - 匯出數據

API 是安全的，它使用與 UI 相同的身份驗證方法。您可以在"REST API 文檔"。

存取 **Swagger API** 參考

要進入 Swagger，您需要資料分類實例的 IP 位址。對於雲端部署，您將使用公用 IP 位址。然後你需要進入這個端點：

https://<classification_ip>/documentation

使用 **API** 的範例

以下範例顯示了複製檔案的 API 呼叫。

API 請求

您最初需要取得系統的所有相關欄位和選項，以查看調查標籤中的所有篩選器。

```
curl -X GET "http://{classification_ip}/api/{classification_version}
/search/options?data_mode=ALL_EXTRACTABLE" -H "accept: application/json"
-H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… " -H "x-agent-id:
hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients"
```

回覆


```
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
      "field": "string",
      "is_rulable": true,
      "name": "string",
      "operators": [
        "EQUALS"
      ],
      "optional_values": [
        {}
      ],
      "secondary": {},
      "server_data": false,
      "type": "TEXT"
    }
  ]
}
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "POLICIES",
      "name": "Policies",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "EXTRACTION_STATUS_RANGE",
```

```
    "name": "Scan Analysis Status",
    "operators": [
      "IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "SCAN_ANALYSIS_ERROR",
    "name": "Scan Analysis Event",
    "operators": [
      "IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "PUBLIC_ACCESS",
    "name": "Open Permissions",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": true,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "USERS_PERMISSIONS_COUNT_RANGE",
    "name": "Number of Users with Access",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": true,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "USER_GROUP_PERMISSIONS",
```

```
      "name": "User / Group Permissions",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_OWNER",
      "name": "File Owner",
      "operators": [
        "EQUALS",
        "CONTAINS"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT_TYPE",
      "name": "system-type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT",
      "name": "system",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
```

```
    "field": "SCAN_TASK",
    "name": "Storage Repository",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "FILE_PATH",
    "name": "File / Directory Path",
    "operators": [
      "MULTI_CONTAINS",
      "MULTI_EXCLUDE"
    ],
    "server_data": true,
    "type": "MULTI_TEXT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_DASHBOARD_EXTRACTABLE",
    "field": "CATEGORY",
    "name": "Category",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_EXTRACTABLE",
    "field": "PATTERN_SENSITIVITY_LEVEL",
    "name": "Sensitivity Level",
    "operators": [
      "IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
```

```
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "NUMBER_OF_IDENTIFIERS",
      "name": "Number of identifiers",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_PERSONAL",
      "name": "Personal Data",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_SENSITIVE",
      "name": "Sensitive Personal Data",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DATA_SUBJECT",
      "name": "Data Subject",
      "operators": [
        "EQUALS",
        "CONTAINS"
      ],
      "server_data": true,
      "type": "TEXT"
    },
```

```
    {
      "active_directory_affected": false,
      "data_mode": "DIRECTORIES",
      "field": "DIRECTORY_TYPE",
      "name": "Directory Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "FILE_TYPE",
      "name": "File Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "FILE_SIZE_RANGE",
      "name": "File Size",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_CREATION_RANGE_RETENTION",
      "name": "Created Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
```

```
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DISCOVERED_TIME_RANGE",
      "name": "Discovered Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_LAST_MODIFICATION_RETENTION",
      "name": "Last Modified",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_LAST_ACCESS_RANGE_RETENTION",
      "name": "Last Accessed",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "IS_DUPLICATE",
      "name": "Duplicates",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
```

```
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "FILE_HASH",
      "name": "File Hash",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "USER_DEFINED_STATUS",
      "name": "Tags",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ASSIGNED_TO",
      "name": "Assigned to",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    }
  ]
}
```

我們將在請求參數中使用該回應來過濾我們想要複製的檔案。

您可以對多個項目套用一個操作。支援的操作類型包括：移動、刪除和複製。

我們將創建複製動作：

API 請求

下一個 API 是操作 API，它允許您建立多個操作。

```
curl -X POST "http://
{classification_ip}/api//{classification_version}/actions" -H "accept:
application/json" -H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… "
-H "x-agent-id: hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients " -H "Content-
Type: application/json" -d "{ \"action_type\": \"COPY\", \"data_mode\":
\"FILES\", \"policy_id\": 0, \"request_params\": { destination_nfs_path:
"{ontap_ip}:/{share_name} " },
\"requested_query\":{"condition":"AND","rules":[{"field":"ENVIRONMENT_TYPE
","operator":"IN","value":["ONPREM"]},{"field":"CATEGORY","operator":"IN",
"value":["21"]}]}}"
```

回覆

回應將傳回操作對象，因此您可以使用取得和刪除 API 來取得有關操作的狀態，或取消它。

```
{
  "action_type": "COPY",
  "creation_time": "2023-08-08T12:37:21.705Z",
  "data_mode": "FILES",
  "end_time": "2023-08-08T12:37:21.705Z",
  "estimated_time_to_complete": 0,
  "id": 0,
  "policy_id": 0,
  "policy_name": "string",
  "priority": 0,
  "request_params": {},
  "requested_query": {},
  "result": {
    "error_message": "string",
    "failed": 0,
    "in_progress": 0,
    "succeeded": 0,
    "total": 0
  },
  "start_time": "2023-08-08T12:37:21.705Z",
  "status": "QUEUED",
  "title": "string",
  "user_id": "string"
}
```

知識和支持

註冊NetApp Console支持

需要進行支援註冊才能獲得針對NetApp Console及其儲存解決方案和資料服務的技術支援。還需要支援註冊才能啟用Cloud Volumes ONTAP系統的關鍵工作流程。

註冊支援並不能使NetApp獲得雲端提供者文件服務的支援。有關雲端提供者文件服務、其基礎設施或使用該服務的任何解決方案的技術支持，請參閱該產品文件中的「取得協助」。

- "適用於ONTAP 的Amazon FSx"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

支援註冊概述

啟動支持權利的註冊方式有兩種：

- 註冊您的NetApp Console帳戶序號（您的 20 位元 960xxxxxxxx 序號位於控制台中的「支援資源」頁面上）。

  這可作為控制台內任何服務的單一支援訂閱 ID。每個控制台帳戶都必須註冊。

- 在您的雲端供應商市場中註冊與訂閱相關的Cloud Volumes ONTAP序號（這些是 20 位元 909201xxxxxxxx 序號）。

  這些序號通常稱為_PAYGO 序號_，由NetApp Console在Cloud Volumes ONTAP部署時產生。

註冊兩種類型的序號可以實現開立支援票和自動產生案例等功能。透過將NetApp支援網站 (NSS) 帳戶新增至控制台即可完成註冊，如下所述。

註冊NetApp Console以取得NetApp支持

要註冊支援並啟動支援權利，您的NetApp Console帳戶中的一名使用者必須將NetApp支援網站帳戶與其控制台登入名稱關聯。如何註冊NetApp支援取決於您是否已經擁有NetApp支援網站 (NSS) 帳號。

擁有 **NSS** 帳戶的現有客戶

如果您是擁有 NSS 帳戶的NetApp客戶，只需透過控制台註冊即可獲得支援。

步驟

1. 選擇“管理”>“憑證”。
2. 選擇*使用者憑證*。
3. 選擇*新增 NSS 憑證*並依照NetApp支援網站 (NSS) 驗證提示進行操作。
4. 若要確認註冊過程是否成功，請選擇「幫助」圖標，然後選擇「支援」。

   *資源*頁面應顯示您的控制台帳戶已註冊以獲得支援。

請注意，如果其他控制台使用者尚未將NetApp支援網站帳戶與其登入名稱關聯，他們將看不到相同的支援註冊狀態。但是，這並不意味著您的帳戶沒有註冊支援。只要組織中的一名使用者遵循了這些步驟，您的帳戶就已註冊。

現有客戶但沒有 **NSS** 帳戶

如果您是現有的NetApp客戶，擁有現有授權和序號但沒有 NSS 帳戶，則需要建立 NSS 帳戶並將其與您的控制台登入關聯。

步驟

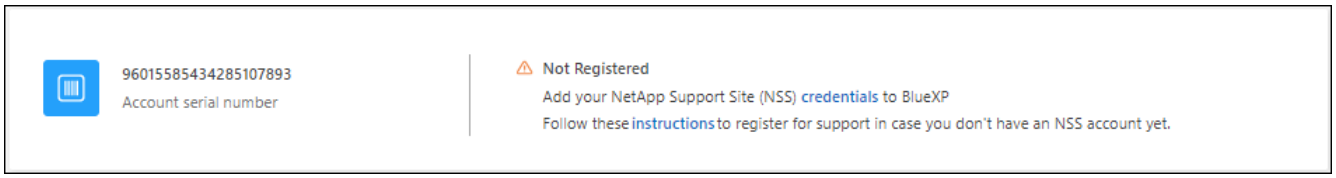
1. 透過完成以下操作建立NetApp支援網站帳戶 "NetApp支援網站使用者註冊表"
   a. 請務必選擇適當的使用者級別，通常為* NetApp客戶/最終使用者*。
   b. 請務必複製上面用於序號欄位的控制台帳戶序號（960xxxx）。這將加快帳戶處理速度。
2. 完成以下步驟，將您的新 NSS 帳戶與您的控制台登入名稱關聯擁有 NSS 帳戶的現有客戶。

NetApp全新產品

如果您是NetApp新使用者且沒有 NSS 帳戶，請依照下列步驟操作。

步驟

1. 在控制台的右上角，選擇「幫助」圖標，然後選擇「支援」。
2. 從支援註冊頁面找到您的帳戶 ID 序號。



3. 導航至 "NetApp 的支援註冊網站"並選擇*我不是註冊的NetApp客戶*。
4. 填寫必填欄位（帶有紅色星號的欄位）。
5. 在*產品線*欄位中，選擇*雲端管理員*，然後選擇適用的計費提供者。
6. 從上面的步驟 2 複製您的帳戶序號，完成安全性檢查，然後確認您已閱讀 NetApp 的全球資料隱私政策。

   一封電子郵件會立即發送到提供的郵箱以完成此安全交易。如果幾分鐘內沒有收到驗證電子郵件，請務必檢查您的垃圾郵件資料夾。

7. 從電子郵件中確認操作。

   確認向NetApp提交您的請求並建議您建立NetApp支援網站帳戶。

8. 透過完成以下操作建立NetApp支援網站帳戶 "NetApp支援網站使用者註冊表"
   a. 請務必選擇適當的使用者級別，通常為* NetApp客戶/最終使用者*。
   b. 請務必複製上面用於序號欄位的帳戶序號（960xxxx）。這將加快處理速度。

完成後

NetApp應該在過程中與您聯繫。這是針對新用戶的一次性入職培訓。

擁有NetApp支援網站帳號後，請依照下列步驟將該帳號與您的控制台登入關聯擁有 NSS 帳戶的現有客戶。

關聯 NSS 憑證以獲得Cloud Volumes ONTAP支持

需要將NetApp支援網站憑證與您的控制台帳戶關聯，才能為Cloud Volumes ONTAP啟用以下關鍵工作流程：

- 註冊即用即付Cloud Volumes ONTAP系統以獲得支持

  需要提供您的 NSS 帳戶才能啟動對您的系統的支援並獲得NetApp技術支援資源的存取權限。

- 自帶授權 (BYOL) 時部署Cloud Volumes ONTAP

  需要提供您的 NSS 帳戶，以便控制台可以上傳您的許可證金鑰並啟用您購買的期限的訂閱。這包括期限續訂的自動更新。

- 將Cloud Volumes ONTAP軟體升級至最新版本

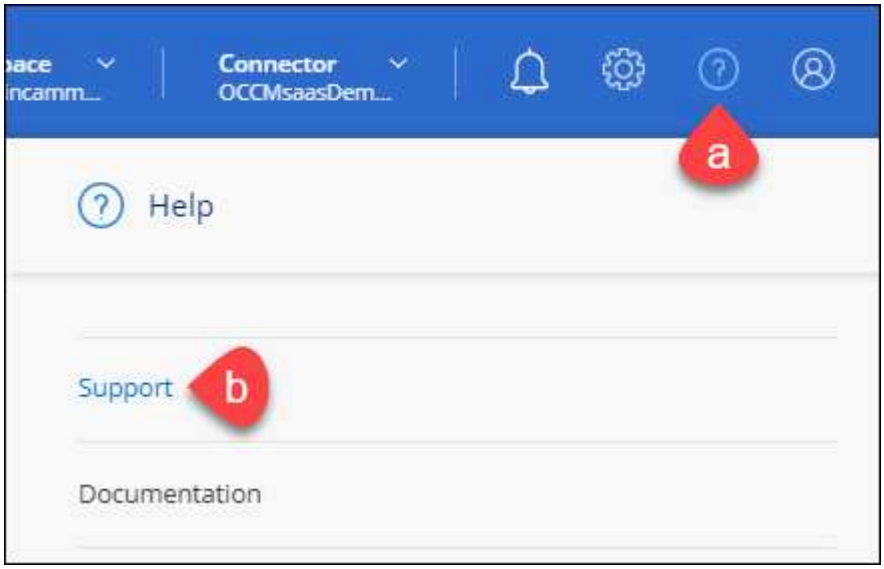
將 NSS 憑證與您的NetApp Console帳戶關聯與將 NSS 帳戶與控制台使用者登入相關聯。

這些 NSS 憑證與您的特定控制台帳戶 ID 相關聯。屬於控制台組織的使用者可以從*支援 > NSS 管理*存取這些憑證。

- 如果您有客戶級帳戶，則可以新增一個或多個 NSS 帳戶。
- 如果您有合作夥伴或經銷商帳戶，則可以新增一個或多個 NSS 帳戶，但不能與客戶級帳戶一起新增。

步驟

1. 在控制台的右上角，選擇「幫助」圖標，然後選擇「支援」。



2. 選擇*NSS 管理 > 新增 NSS 帳號*。
3. 當出現提示時，選擇「繼續」以重新導向至 Microsoft 登入頁面。

   NetApp使用 Microsoft Entra ID 作為特定於支援和授權的身份驗證服務的身份提供者。

4. 在登入頁面，提供您的NetApp支援網站註冊的電子郵件地址和密碼以執行驗證程序。

   這些操作使控制台能夠使用您的 NSS 帳戶進行許可證下載、軟體升級驗證和未來支援註冊等操作。

   請注意以下事項：

   - NSS 帳戶必須是客戶級帳戶（不是訪客或臨時帳戶）。您可以擁有多個客戶級 NSS 帳戶。
   - 如果該帳戶是合作夥伴等級帳戶，則只能有一個 NSS 帳戶。如果您嘗試新增客戶級 NSS 帳戶且合作夥伴級帳戶已存在，您將收到以下錯誤訊息：

     “此帳戶不允許使用 NSS 客戶類型，因為已經存在不同類型的 NSS 用戶。”

   如果您已有客戶級 NSS 帳戶並嘗試新增合作夥伴級帳戶，情況也是如此。

   - 成功登入後， NetApp將儲存 NSS 使用者名稱。

     這是系統產生的映射到您的電子郵件的 ID。在*NSS 管理*頁面上，您可以顯示來自 ••• 菜單。

   - 如果您需要刷新登入憑證令牌，還有一個*更新憑證*選項 ••• 菜單。

     使用此選項會提示您再次登入。請注意，這些帳戶的令牌將在 90 天後過期。我們將發布通知來提醒您此事。

取得NetApp Data Classification協助

NetApp以多種方式為NetApp Console及其雲端服務提供支援。全天候提供廣泛的免費自助支援選項，例如知識庫 (KB) 文章和社群論壇。您的支援註冊包含透過網路工單取得的遠端技術支援。

獲取雲端提供者文件服務的支持

有關雲端提供者文件服務、其基礎設施或使用該服務的任何解決方案的技術支持，請參閱該產品的文檔。

- "適用於ONTAP 的Amazon FSx"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

要獲得特定於NetApp及其儲存解決方案和資料服務的技術支持，請使用下面所述的支援選項。

使用自助選項

這些選項每週 7 天、每天 24 小時免費提供：

- 文件

  您目前正在檢視的NetApp Console文件。

- "知識庫"

搜尋NetApp知識庫以尋找有助於解決問題的文章。

- "社群"

  加入NetApp Console社區，關注正在進行的討論或創建新的討論。

向NetApp支援建立案例

除了上述自助支援選項之外，您還可以在啟動支援後與NetApp支援專家合作解決任何問題。

開始之前

- 若要使用「建立案例」功能，您必須先將您的NetApp支援網站憑證與您的控制台登入關聯。 "了解如何管理與控制台登入相關的憑證" 。
- 如果您要為具有序號的ONTAP系統開啟案例，那麼您的 NSS 帳戶必須與該系統的序號相關聯。

步驟

1. 在NetApp Console中，選擇「說明」>「支援」 。
2. 在「資源」頁面上，選擇「技術支援」下的可用選項之一：

   a. 如果您想透過電話與某人交談，請選擇「致電我們」 。您將被引導至 netapp.com 上的一個頁面，其中列出了您可以撥打的電話號碼。

   b. 選擇「建立案例」向NetApp支援專家開立票據：

      - 服務：選擇與問題相關的服務。例如，* NetApp Console* 特定於控制台內的工作流程或功能的技術支援問題。
      - 系統：如果適用於存儲，請選擇* Cloud Volumes ONTAP* 或 **On-Prem**，然後選擇相關的工作環境。

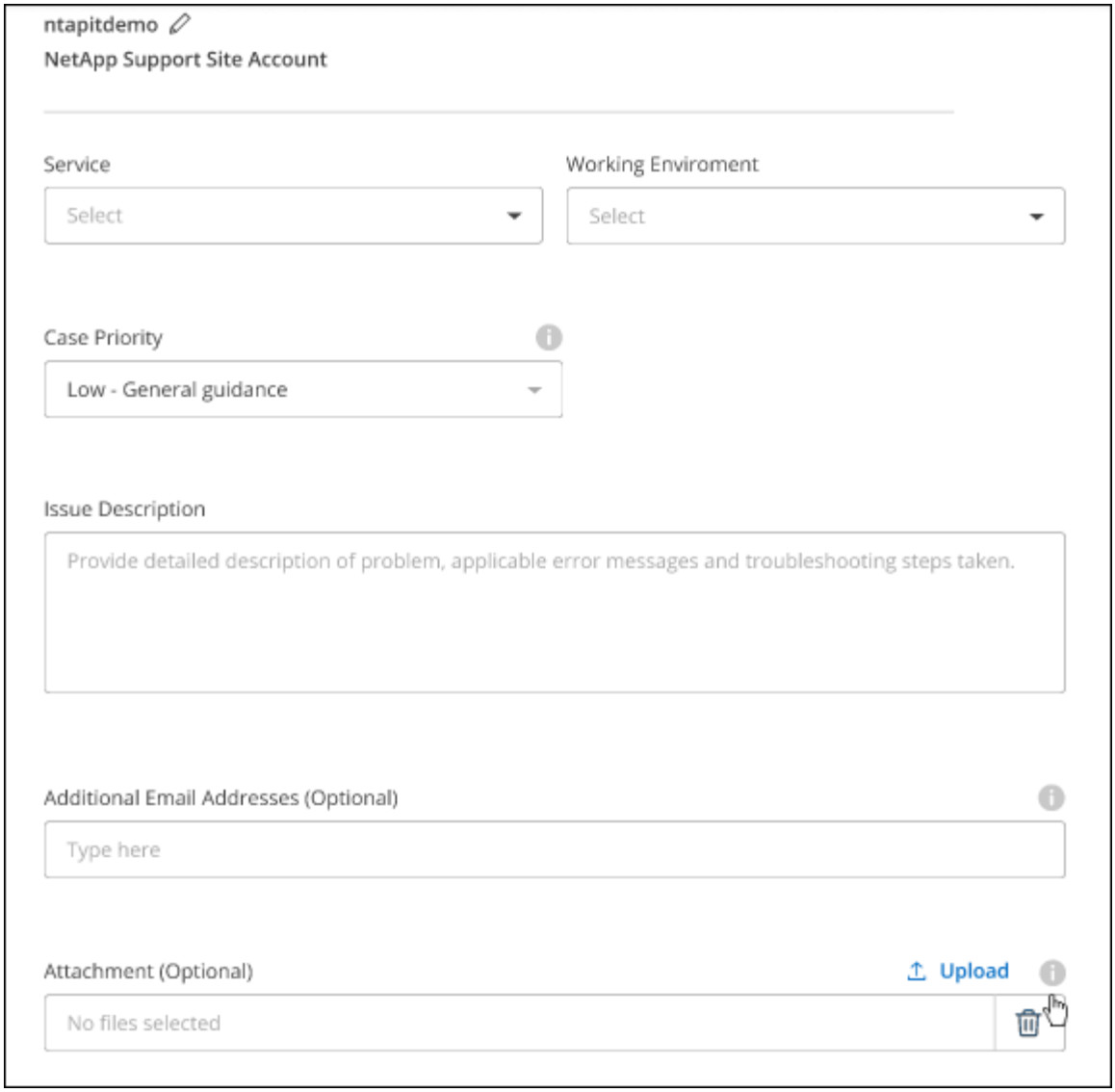
        系統清單位於控制台組織範圍內，並且您在頂部橫幅中選擇了控制台代理。

      - 個案優先級：選擇個案的優先級，可以是低、中、高或嚴重。

        要了解有關這些優先事項的更多詳細信息，請將滑鼠懸停在欄位名稱旁邊的資訊圖示上。

      - 問題描述：提供問題的詳細描述，包括任何適用的錯誤訊息或您執行的故障排除步驟。
      - 其他電子郵件地址：如果您想讓其他人知道此問題，請輸入其他電子郵件地址。
      - 附件（選購）：一次最多上傳五個附件。

        每個附件檔案大小限制為 25 MB。支援以下檔案副檔名：txt、log、pdf、jpg/jpeg、rtf、doc/docx、xls/xlsx 和 csv。



完成後

將會出現一個彈出窗口，其中顯示您的支援案例編號。NetApp支援專家將審查您的案例並儘快回覆您。

若要查看支援案例的歷史記錄，您可以選擇*設定>時間軸*並尋找名為「建立支援案例」的操作。最右邊的按鈕可讓您展開操作以查看詳細資訊。

嘗試建立案例時，您可能會遇到以下錯誤訊息：

“您無權針對所選服務建立案例”

此錯誤可能表示 NSS 帳戶及其關聯的記錄公司與NetApp Console帳戶序號的記錄公司不同（即。960xxxx）或工作環境序號。您可以使用以下選項之一尋求協助：

- 提交非技術案例 https://mysupport.netapp.com/site/help

管理您的支援案例

您可以直接從控制台檢視和管理活動和已解決的支援案例。您可以管理與您的 NSS 帳戶和公司相關的案例。

請注意以下事項：

- 頁面頂部的案例管理儀表板提供兩種視圖：
  - 左側視圖顯示了您提供的使用者 NSS 帳戶在過去 3 個月內開啟的案件總數。
  - 右側的視圖根據您的使用者 NSS 帳戶顯示了過去 3 個月內貴公司層級開設的案件總數。

  表中的結果反映了與您選擇的視圖相關的案例。
- 您可以新增或刪除感興趣的列，並且可以過濾優先順序和狀態等列的內容。其他欄位僅提供排序功能。

  請查看以下步驟以了解更多詳細資訊。
- 在每個案件級別，我們提供更新案件記錄或關閉尚未關閉或待關閉狀態的案件的功能。

步驟

1. 在NetApp Console中，選擇「說明」>「支援」。
2. 選擇*案例管理*，如果出現提示，請將您的 NSS 帳戶新增至控制台。

   案例管理*頁面顯示與您的控制台使用者帳戶關聯的 **NSS** 帳戶相關的未結案例。這與出現在 ***NSS** 管理 頁面頂部的 NSS 帳戶相同。
3. （可選）修改表中顯示的資訊：
   - 在「組織的案例」下，選擇「查看」以查看與您的公司相關的所有案例。
   - 透過選擇精確的日期範圍或選擇不同的時間範圍來修改日期範圍。
   - 過濾列的內容。
   - 透過選擇  然後選擇您想要顯示的列。
4. 透過選擇管理現有案例•••並選擇其中一個可用選項：
   - 查看案例：查看有關特定案例的完整詳細資訊。
   - 更新案例說明：提供有關您的問題的更多詳細信息，或選擇*上傳文件*以附加最多五個文件。

     每個附件檔案大小限制為 25 MB。支援以下檔案副檔名：txt、log、pdf、jpg/jpeg、rtf、doc/docx、xls/xlsx 和 csv。
   - 結案：提供有關結案原因的詳細信息，然後選擇*結案*。

關於NetApp Data Classification的常見問題解答

如果您只是想快速找到問題的答案，此常見問題解答可以為您提供協助。

NetApp Data Classification

以下問題提供了對資料分類的一般了解。

資料分類如何運作？

資料分類在NetApp Console系統和儲存系統旁邊部署了另一層 AI。然後，它會掃描磁碟區、儲存桶、資料庫和其他儲存帳戶上的數據，並對找到的數據洞察進行索引。資料分類利用人工智慧和自然語言處理，而不是通常圍繞正規表示式和模式匹配構建的替代解決方案。

資料分類使用人工智慧來提供對資料的上下文理解，以便進行準確的檢測和分類。它由人工智慧驅動，因為它是針對現代資料類型和規模而設計的。它還了解數據背景，以便提供強大、準確的發現和分類。

"詳細了解資料分類的工作原理" 。

資料分類是否有 **REST API**，它是否可以與第三方工具一起使用？

是的，資料分類有一個 REST API，用於支援控制台核心平台一部分的資料分類版本中的功能。看"API 文檔" 。

資料分類是否可以透過雲端市場取得？

資料分類是NetApp Console核心功能的一部分，因此您不需要使用此服務的市場。

資料分類掃描與分析

以下問題與資料分類掃描效能和分析有關。

資料分類多久掃描一次我的資料？

雖然初始資料掃描可能需要一點時間，但後續掃描僅檢查增量變化，從而減少系統掃描時間。資料分類以循環方式連續掃描您的數據，每次掃描六個儲存庫，以便所有更改的資料都能快速分類。

"了解掃描的工作原理" 。

資料分類每天僅掃描資料庫一次；資料庫不像其他資料來源那樣被連續掃描。

資料掃描對您的儲存系統和資料的影響可以忽略不計。

掃描性能是否有所不同？

掃描效能可能會因網路頻寬和環境中的平均檔案大小而異。它還取決於主機系統（在雲端或本地）的大小特徵。請參閱"資料分類實例"和"部署資料分類"了解更多。

在最初新增的資料來源時，您也可以選擇僅執行「對應」（Mapping only）掃描，而不是完整的「分類」（Map

& Classify）掃描。由於它不需要存取檔案來查看其中的數據，因此可以非常快速地在資料來源上完成映射。"查看映射和分類掃描之間的區別"。

我可以使用資料分類來搜尋我的資料嗎？

資料分類提供了廣泛的搜尋功能，可以輕鬆地在所有連接的來源中搜尋特定檔案或資料。資料分類使用戶能夠進行比元資料所反映的更深入的搜尋。它是一種與語言無關的服務，還可以讀取檔案並分析多種敏感資料類型，例如名稱和 ID。例如，使用者可以在結構化和非結構化資料儲存中進行搜索，以查找可能從資料庫洩漏到使用者文件的數據，從而違反公司政策。可以儲存搜尋結果以供日後使用，並且可以建立策略以設定的頻率搜尋並對結果採取行動。

一旦找到感興趣的文件，就可以列出其特徵，包括標籤、系統帳戶、儲存桶、文件路徑、類別（來自分類）、文件大小、上次修改、權限狀態、重複、敏感度等級、個人資料、文件內的敏感資料類型、擁有者、文件類型、文件大小、建立時間、文件雜湊、資料是否分配給尋求其關注的人等等。可以使用過濾器來篩選出不相關的特徵。

如果存在正確的權限，資料分類還具有基於角色的存取控制（RBAC），允許移動或刪除檔案。如果沒有正確的權限，則可以將任務指派給組織中具有正確權限的人員。

資料分類管理和隱私

以下問題提供了有關如何管理資料分類和隱私設定的資訊。

如何啟用或停用資料分類？

首先，您需要在控制台或本機系統中部署資料分類實例。執行個體執行後，您可以從「配置」標籤或透過選擇特定係統在現有系統、資料庫和其他資料來源上啟用服務。"了解如何開始"。

> ⓘ 在資料來源上啟動資料分類將立即導致初始掃描。掃描結果很快就會顯示。

您可以從資料分類設定頁面停用資料分類掃描單一系統、資料庫或檔案共用群組。看"從資料分類中刪除資料來源"。

若要完全刪除資料分類實例，請從雲端提供者的入口網站或本機位置手動刪除資料分類實例。

該服務可以排除某些目錄中的掃描資料嗎？

是的。如果您希望資料分類排除駐留在特定資料來源目錄中的掃描數據，則可以將該清單提供給分類引擎。套用變更後，資料分類將排除指定目錄中的掃描資料。"了解更多"。

是否掃描了位於ONTAP磁碟區上的快照？

否。資料分類不會掃描快照，因為其內容與磁碟區中的內容相同。

如果在ONTAP磁碟區上啟用了資料分層，會發生什麼情況？

當資料分類使用僅映射掃描掃描具有分層到物件儲存的冷資料的磁碟區時，它會掃描所有資料 - 本地磁碟上的資料和分層到物件儲存的冷資料。對於實施分層的非NetApp產品也是如此。

僅映射掃描不會使冷資料升溫—它會保持冷狀態並保留在物件儲存中。另一方面，如果您執行地圖和分類掃描，某些配置可能會使冷資料升溫。

來源系統和資料類型的類型

以下問題涉及可掃描的儲存類型以及掃描的資料類型。

在政府區域部署時有限制嗎？

當控制台代理部署在政府區域（AWS GovCloud、Azure Gov 或 Azure DoD）時，支援資料分類 - 也稱為「受限模式」。

如果我在沒有網路存取的網站安裝資料分類，我可以掃描哪些資料來源？

BlueXP私有模式（傳統BlueXP介面）通常用於沒有網路連線的本機環境和安全雲端區域，其中包括 AWS Secret Cloud、AWS Top Secret Cloud 和 Azure IL6。NetApp繼續透過傳統的BlueXP介面支援這些環境。有關舊版BlueXP介面中的私有模式文檔，請參閱"BlueXP私人模式的 PDF 文檔"。

資料分類只能掃描來自本地站點的資料來源的資料。目前，資料分類可以以「私人模式」掃描以下本地資料來源 - 也稱為「暗」網站：

- 本地ONTAP系統
- 資料庫模式
- 使用簡單存儲服務（S3）協定的對象存儲

支援哪些文件類型？

資料分類掃描所有文件的類別和元資料洞察，並在儀表板的文件類型部分顯示所有文件類型。

當資料分類偵測到個人識別資訊 (PII) 或執行 DSAR 搜尋時，僅支援以下文件格式：

```
.CSV, .DCM, .DOC, .DOCX, .JSON, .PDF, .PPTX, .RTF, .TXT, .XLS, .XLSX, Docs, Sheets, and Slides
```

資料分類擷取哪些類型的資料和元資料？

資料分類可讓您對資料來源執行常規「映射」掃描或完整「分類」掃描。映射僅提供資料的高級概述，而分類提供資料的深層掃描。由於它不需要存取檔案來查看其中的數據，因此可以非常快速地在資料來源上完成映射。

- 資料映射掃描（僅映射掃描）：資料分類僅掃描元資料。這對於整體資料管理和治理、快速專案範圍界定、大型地產和優先排序很有用。數據映射基於元數據，被認為是一種*快速*掃描。

  快速掃描後，您可以產生資料映射報告。此報告概述了您公司資料來源中儲存的數據，以幫助您做出有關資源利用率、遷移、備份、安全性和合規性流程的決策。

- 資料分類深度掃描（地圖和分類掃描）：資料分類使用標準協定和唯讀權限在整個環境中掃描資料。開啟選定的文件並掃描其中的敏感業務相關資料、私人資訊以及與勒索軟體相關的問題。

  完整掃描後，您可以將許多附加資料分類功能套用至數據，例如在資料調查頁面中查看和最佳化資料、在檔案中搜尋名稱、複製、移動和刪除來源檔案等。

資料分類擷取元數據，例如：檔案名稱、權限、建立時間、上次存取和上次修改。這包括資料調查詳情頁面和資料調查報告中顯示的所有元資料。

資料分類可以識別多種類型的私人數據，例如個人資訊（PII）和敏感個人資訊（SPII）。有關私人數據的詳細信息，請參閱資料分類掃描的私人資料類別。

我可以將資料分類資訊限制給特定使用者嗎？

是的，資料分類與NetApp Console完全整合。 NetApp Console使用者只能查看根據其權限有資格查看的系統的資訊。

此外，如果您希望允許某些使用者僅查看資料分類掃描結果而無權管理資料分類設置，則可以為這些使用者指派*分類檢視器*角色（在標準模式下使用NetApp Console時）或*合規檢視器*角色（在受限模式下使用NetApp Console時）。"了解更多"。

任何人都可以存取我的瀏覽器和資料分類之間發送的私人資料嗎？

不可以。您的瀏覽器和資料分類實例之間發送的私人資料使用 TLS 1.2 進行端對端加密保護，這表示NetApp和非NetApp方都無法讀取它。除非您要求並批准訪問，否則資料分類不會與NetApp共用任何資料或結果。

掃描的資料保留在您的環境中。

敏感資料如何處理？

NetApp無法存取敏感數據，也不會在 UI 中顯示它。敏感資料被封鎖，例如，顯示信用卡資訊的最後四位數字。

資料儲存在哪裡？

掃描結果儲存在資料分類實例內的 Elasticsearch 中。

如何存取資料？

資料分類透過 API 呼叫存取儲存在 Elasticsearch 中的數據，這些呼叫需要身份驗證並使用 AES-128 加密。直接存取 Elasticsearch 需要 root 存取權限。

許可證和費用

以下問題涉及使用資料分類的許可和成本。

資料分類的費用是多少？

資料分類是NetApp Console的核心功能。沒有充電。

控制台代理部署

以下問題與控制台代理有關。

什麼是控制台代理？

控制台代理程式是在您的雲端帳戶或本機的運算實例上執行的軟體，它使NetApp Console能夠安全地管理雲端資源。您必須部署控制台代理程式才能使用資料分類。

控制台代理需要安裝在哪裡？

掃描資料時，需要在以下位置安裝NetApp Console代理程式：

- 對於 AWS 中的Cloud Volumes ONTAP或Amazon FSx for ONTAP：控制台代理程式位於 AWS 中。
- 對於 Azure 或Azure NetApp Files中的Cloud Volumes ONTAP ：控制台代理程式位於 Azure 中。
- 對於 GCP 中的Cloud Volumes ONTAP ：控制台代理程式位於 GCP 中。
- 對於本機ONTAP系統：控制台代理程式位於本機。

如果您在這些位置有數據，您可能需要使用 "多個控制台代理"。

資料分類是否需要存取憑證？

資料分類本身不會檢索儲存憑證。相反，它們存儲在控制台代理中。

資料分類使用資料平面憑證（例如 CIFS 憑證）在掃描之前掛載共用。

服務和控制台代理之間的通訊是否使用 **HTTP**？

是的，資料分類使用 HTTP 與控制台代理進行通訊。

資料分類部署

以下問題與單獨的資料分類實例有關。

資料分類支援哪些部署模型？

NetApp Console允許使用者在幾乎任何地方掃描和報告系統，包括本機、雲端和混合環境。資料分類通常使用 SaaS 模型部署，其中服務透過控制台介面啟用，不需要安裝硬體或軟體。即使在這種點擊即運行的部署模式下，無論資料儲存是在本地還是在公有雲中，都可以進行資料管理。

資料分類需要什麼類型的實例或虛擬機器？

什麼時候"部署在雲端"：

- 在 AWS 中，資料分類在具有 500 GiB GP2 磁碟的 m6i.4xlarge 執行個體上執行。您可以在部署期間選擇較小的實例類型。
- 在 Azure 中，資料分類在具有 500 GiB 磁碟的 Standard_D16s_v3 VM 上執行。
- 在 GCP 中，資料分類在具有 500 GiB 標準持久磁碟的 n2-standard-16 VM 上運作。

"詳細了解資料分類的工作原理" 。

我可以在自己的主機上部署資料分類嗎？

是的。您可以在網路或雲端中具有網際網路存取權限的 Linux 主機上安裝資料分類軟體。一切工作正常，您可以繼續透過控制台管理掃描配置和結果。看"在本地部署資料分類"了解系統需求和安裝詳情。

沒有網路連線的安全站點怎麼樣？

是的，也支持。你可以"在沒有網路存取權限的本機站點中部署資料分類"以獲得完全安全的網站。

法律聲明

法律聲明提供對版權聲明、商標、專利等的存取。

版權

"https://www.netapp.com/company/legal/copyright/"

商標

NETAPP、NETAPP 標誌和NetApp商標頁面上列出的標誌是NetApp, Inc. 的商標。其他公司和產品名稱可能是其各自所有者的商標。

"https://www.netapp.com/company/legal/trademarks/"

專利

NetApp擁有的專利的最新清單可在以下位置找到：

https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf

隱私權政策

"https://www.netapp.com/company/legal/privacy-policy/"

開源

通知文件提供有關NetApp軟體中使用的第三方版權和許可的資訊。

- "NetApp Console通知"
- "NetApp Data Classification聲明"