NetApp Data Classificationドキュメント

NetApp Data Classification

NetApp
February 06, 2026

目次

NetApp Data Classificationドキュメント

リリース ノート

NetApp Data Classificationの新機能

NetApp Data Classificationの新機能について説明します。

2026年1月14日

バージョン**1.50**

この Data Classification リリースには、バグ修正と次の更新が含まれています。

カスタム分類の改善

データ分類では、データのカスタム カテゴリの作成がサポートされるようになりました。ファイルをアップロードして、データ分類がデータにカテゴリ マーカーを適用するために使用する AI モデルを微調整できます。すべてのカスタム分類のインターフェースが改善されました。

詳細については、 "カスタム分類を作成する"。

カスタムの古いデータ定義

データ分類では、組織のニーズに合わせて古いデータの定義をカスタマイズできるようになりました。以前は、古いデータは 3 年前に最後に変更されたデータとして定義されていました。現在では、最後にアクセスされた日時、または最後に変更された日時に基づいて古いデータを識別できます。期間は 6 か月前から 10 年前までの範囲です。

詳細については、 "古いデータ定義をカスタマイズする"。

パフォーマンスの向上

データ分類、データ マッピング レポート、および調査ページのフィルターのすべてのページの読み込み時間が短縮されました。

調査報告書の所要時間の見積もり

調査レポートをダウンロードすると、データ分類ではダウンロードが完了するまでの予想時間が表示されるようになりました。

2025年12月8日

バージョン**1.49**

この Data Classification リリースには、バグ修正と次の更新が含まれています。

ヘルスモニタリングダッシュボードで指標とパフォーマンスを監視する

データ分類では、ヘルス モニタリング ダッシュボードが提供されるようになり、リソースのリアルタイム モニタリングと、メモリ使用量、ディスク使用量、ディスク使用率などの分析情報が提供されます。ヘルス モニタリング ダッシュボードから得られる分析情報を活用することで、展開のインフラストラクチャを確認し、ストレージとパフォーマンスを最適化するための分析情報を得ることができます。

詳細については、 "データ分類の健全性を監視する"。

読み込みパフォーマンスの向上

データ分類のすべてのページの読み込みパフォーマンスが向上し、より効率的なユーザー エクスペリエンスが実現しました。

2025年11月10日

バージョン**1.48**

この Data Classification リリースには、バグ修正、セキュリティの改善、パフォーマンスの強化が含まれています。

スキャンの進行状況の明確化を強化

スキャン構成には、スキャン完了に関する改善された分析情報が含まれるようになりました。以前は、スキャンの進行中にのみ進行状況バーが表示されていました。スキャンが正常に完了したことを確認できるように、完了後も進行状況バーが表示されるようになりました。マップおよびスキャンされたファイルの数を表示することもできます。

スキャン設定の詳細については、 "リポジトリのNetApp Data Classificationスキャン設定を変更する"。

2025年10月6日

バージョン**1.47**

BlueXP classificationは**NetApp Data Classification**になりました

BlueXP classificationがNetApp Data Classification に変更されました。名前の変更に加えて、ユーザー インターフェイスも強化されました。

BlueXPは**NetApp Console**になりました

BlueXP は、データ インフラストラクチャの管理における役割をより適切に反映するために、名前が変更され、再設計されました。

NetApp Consoleは、オンプレミスとクラウド環境全体にわたるエンタープライズ グレードのストレージとデータ サービスの集中管理を提供し、リアルタイムの分析情報、ワークフローの高速化、管理の簡素化を実現します。

変更内容の詳細については、 "NetApp Consoleのリリースノート" 。

強化された調査エクスペリエンス

新しい検索可能なフィルター、値ごとの結果数、主要な調査結果をまとめたリアルタイムの分析情報、カスタマイズ可能な列とスライド式の詳細ペインを備えた更新された結果テーブルを使用して、データをより速く見つけて理解できます。

詳細については、以下を参照してください。 "データを調査する" 。

新しいガバナンスとコンプライアンスのダッシュボード

直感的なウィジェット、より鮮明なビジュアル、改善された読み込みパフォーマンスにより、重要な洞察をより早く得ることができます。詳細については、"データに関するガバナンス情報を確認する"そして"データに関するコンプライアンス情報を表示する"。

保存されたクエリのポリシー（プレビュー）

データ分類により、条件付きアクションによるガバナンスを自動化できるようになりました。自動削除を含む保持ルールを作成し、定期的な電子メール通知を設定できます。これらはすべて、更新された保存済みクエリページから管理されます。

詳細については、以下を参照してください。 "ポリシーを作成" 。

アクション（プレビュー）

調査ページから直接制御し、ファイルを個別または一括で削除、移動、コピー、タグ付けして、効率的なデータ管理と修復を実現します。

詳細については、以下を参照してください。 "データを調査する" 。

Google Cloud NetApp Volumesのサポート

Data Classification は、 Google Cloud NetApp Volumesでのスキャンをサポートするようになりました。NetApp ConsoleからGoogle Cloud NetApp Volumesを簡単に追加して、シームレスなデータ スキャンと分類を実現します。詳細については、 "Google Cloud NetApp Volumesをスキャンする"。

2025年8月11日

バージョン**1.46**

このデータ分類リリースには、バグ修正と次の更新が含まれています。

監査ページでのスキャンイベントの分析情報が強化されました

監査ページでは、 BlueXP classificationのスキャン イベントに関する強化された分析がサポートされるようになりました。監査ページには、システムのスキャンが開始された日時、システムのステータス、および問題が表示されるようになりました。共有とシステムのステータスはマッピング スキャンでのみ使用できます。

監査ページの詳細については、以下を参照してください。"NetApp Consoleの操作を監視する" 。

RHEL 9.6 のサポート

このリリースでは、ダーク サイトの展開を含むBlueXP classificationの手動オンプレミス インストール用にRed Hat Enterprise Linux v9.6 のサポートが追加されました。

次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、 BlueXP classificationバージョン 1.30 以上が必要です: Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、9.4、および 9.5。

2025年7月14日

バージョン**1.45**

このBlueXP classificationリリースには、リソース使用率を最適化するコード変更が含まれており、次のようになります。

スキャン対象ファイル共有を追加するワークフローの改善

ファイル共有グループにファイル共有を追加するワークフローが簡素化されました。このプロセスでは、認証タイプ (Kerberos または NTLM) に基づいて CIFS プロトコルのサポートも区別されるようになりました。

詳細については、以下を参照してください。 "ファイル共有をスキャンする" 。

拡張ファイル所有者情報

調査タブでキャプチャされたファイルのファイル所有者に関する詳細情報を表示できるようになりました。[調査] タブでファイルのメタデータを表示するときは、ファイルの所有者を見つけて [詳細を表示] を選択し、ユーザー名、メール アドレス、SAM アカウント名を表示します。このユーザーが所有する他のアイテムも表示できます。この機能は、Active Directory が稼働している作業環境でのみ使用できます。

詳細については、以下を参照してください。 "組織内に保存されているデータを調査する" 。

2025年6月10日

バージョン**1.44**

このBlueXP classificationリリースには以下が含まれます。

ガバナンスダッシュボードの更新時間の改善

ガバナンス ダッシュボードの個々のコンポーネントの更新時間が改善されました。次の表は、各コンポーネントの更新頻度を示しています。

コンポーネント	更新時間
データの時代	24 時間
カテゴリ	24 時間
データの概要	5分
重複ファイル	2 時間
ファイルの種類	24 時間
非ビジネスデータ	2 時間
オープン権限	24 時間
保存済みの検索	2 時間
機密データと幅広い権限	24 時間
データのサイズ	24 時間
古いデータ	2 時間
機密レベル別トップデータリポジトリ	2 時間

最終更新の時刻を表示し、重複ファイル、非ビジネス データ、保存された検索、古いデータ、および機密レベル別の上位データ リポジトリ コンポーネントを手動で更新できます。ガバナンスダッシュボードの詳細については、以下を参照してください。"組織に保存されているデータに関するガバナンスの詳細を表示する" 。

パフォーマンスとセキュリティの改善

BlueXP分類のパフォーマンス、メモリ消費、セキュリティを改善するための機能強化が行われました。

バグ修正

Redis がアップグレードされ、 BlueXP classificationの信頼性が向上しました。BlueXP classificationでは、スキャン中のファイル数レポートの精度を向上させるために Elasticsearch を使用するようになりました。

2025年5月12日

バージョン**1.43**

このBlueXP分類リリースには以下が含まれます。

分類スキャンの優先順位付け

データ分類では、マッピングのみのスキャンに加えて、マップと分類のスキャンを優先順位付けする機能がサポートされており、最初に完了するスキャンを選択できます。マップと分類スキャンの優先順位付けは、スキャンの開始中および開始前にサポートされます。進行中のスキャンを優先することを選択した場合、マッピングスキャンと分類スキャンの両方が優先されます。

詳細については、以下を参照してください。 "スキャンを優先する" 。

カナダの個人識別情報（**PII**）データカテゴリのサポート

データ分類スキャンは、カナダの PII データ カテゴリを識別します。これらのカテゴリには、すべてのカナダの州および準州の銀行情報、パスポート番号、社会保険番号、運転免許証番号、健康保険証番号が含まれます。

詳細については、以下を参照してください。 "個人データのカテゴリ" 。

カスタム分類（プレビュー）

データ分類では、マップと分類スキャンのカスタム分類をサポートします。カスタム分類を使用すると、正規表現を使用して組織固有のデータを取得するようにデータ分類スキャンをカスタマイズできます。この機能は現在プレビュー段階です。

詳細については、以下を参照してください。 "カスタム分類を追加する" 。

保存した検索タブ

ポリシータブの名前が変更されました"保存された検索"。機能に変更はありません。

スキャンイベントを監査ページに送信する

データ分類は、分類イベント（スキャンの開始時と終了時）を"NetAppコンソール監査ページ"。

セキュリティアップデート

- Keras パッケージが更新され、脆弱性 (BDSA-2025-0107 および BDSA-2025-1984) が軽減されました。
- Docker コンテナの構成が更新されました。コンテナは、生のネットワーク パケットを作成するためにホストのネットワーク インターフェイスにアクセスできなくなります。このアップデートでは、不要なアクセスを減らすことで、潜在的なセキュリティ リスクを軽減します。

パフォーマンスの向上

RAM 使用量を削減し、データ分類の全体的なパフォーマンスを向上させるために、コード強化が実装されました。

バグ修正

StorageGRIDスキャンが失敗し、調査ページのフィルター オプションが読み込まれず、大量の評価でデータ検出評価がダウンロードされないというバグが修正されました。

2025年4月14日

バージョン**1.42**

このBlueXP classificationリリースには以下が含まれます。

作業環境の一括スキャン

BlueXP classificationは、作業環境の一括操作をサポートします。作業環境において、マッピング スキャンを有効にするか、マップと分類スキャンを有効にするか、スキャンを無効にするか、ボリューム全体にカスタム構成を作成するかを選択できます。個々のボリュームを選択した場合は、一括選択が上書きされます。一括操作を実行するには、[構成] ページに移動して選択を行います。

調査レポートをローカルにダウンロードする

BlueXP classificationでは、データ調査レポートをローカルにダウンロードしてブラウザーで表示する機能がサポートされています。ローカル オプションを選択した場合、データ調査は CSV 形式でのみ利用可能になり、最初の 10,000 行のデータのみが表示されます。

詳細については、以下を参照してください。 "BlueXP classificationを使用して組織内に保存されているデータを調査します" 。

2025年3月10日

バージョン**1.41**

このBlueXP classificationリリースには、一般的な改善とバグ修正が含まれています。また、次のものも含まれます:

スキャンステータス

BlueXP classificationは、ボリューム上の初期マッピングおよび分類スキャンの進行状況をリアルタイムで追跡します。個別のプログレッシブ バーでマッピング スキャンと分類スキャンが追跡され、スキャンされたファイルの合計数の割合が表示されます。進行状況バーにマウスを移動すると、スキャンされたファイルの数とファイルの合計数を表示することもできます。スキャンのステータスを追跡すると、スキャンの進行状況に関するより深い分析情報が得られ、スキャンをより適切に計画し、リソースの割り当てを把握できるようになります。

スキャンのステータスを表示するには、 BlueXP classificationの 構成 に移動し、作業環境構成 を選択します。各ボリュームごとに進行状況が一行で表示されます。

2025年2月19日

バージョン**1.40**

このBlueXP classificationリリースには、次の更新が含まれています。

RHEL 9.5 のサポート

このリリースでは、これまでサポートされていたバージョンに加えて、Red Hat Enterprise Linux v9.5 のサポートも提供されます。これは、ダーク サイトの展開を含む、 BlueXP classificationの手動オンプレミス インストールに適用されます。

次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、 BlueXP classificationバージョン 1.30 以上が必要です: Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、

9.2、9.3、9.4、および 9.5。

マッピングのみのスキャンを優先する

マッピングのみのスキャンを実行する場合、最も重要なスキャンを優先できます。この機能は、作業環境が多数あり、優先度の高いスキャンが最初に完了するようにしたい場合に役立ちます。

デフォルトでは、スキャンは開始された順序に基づいてキューに入れられます。スキャンを優先順位付けする機能を使用すると、スキャンをキューの先頭に移動できます。複数のスキャンを優先できます。優先順位は先入先出順で指定されます。つまり、最初に優先順位を指定したスキャンがキューの先頭に移動し、2 番目に優先順位を指定したスキャンはキューの 2 番目になり、以下同様に続きます。

優先権は 1 回限り付与されます。マッピング データの自動再スキャンはデフォルトの順序で実行されます。

優先順位は"マッピングのみのスキャン"; マップスキャンや分類スキャンには使用できません。

詳細については、以下を参照してください。 "スキャンを優先する" 。

すべてのスキャンを再試行する

BlueXP classificationは、失敗したすべてのスキャンを一括して再試行する機能をサポートしています。

すべて再試行 機能を使用すると、バッチ操作でスキャンを再試行できます。ネットワークの停止などの一時的な問題により分類スキャンが失敗した場合は、スキャンを個別に再試行するのではなく、1 つのボタンですべてのスキャンを同時に再試行できます。スキャンは必要に応じて何度でも再試行できます。

すべてのスキャンを再試行するには:

1. BlueXP classificationメニューから、*構成*を選択します。
2. 失敗したスキャンをすべて再試行するには、「すべてのスキャンを再試行」を選択します。

分類モデルの精度向上

機械学習モデルの精度は"定義済みカテゴリ"11%向上しました。

2025年1月22日

バージョン**1.39**

このBlueXP classificationリリースでは、データ調査レポートのエクスポート プロセスが更新されます。このエクスポートの更新は、データの追加分析を実行したり、データで追加の視覚化を作成したり、データ調査の結果を他のユーザーと共有したりするのに役立ちます。

以前は、データ調査レポートのエクスポートは 10,000 行に制限されていました。このリリースでは、制限が解除され、すべてのデータをエクスポートできるようになりました。この変更により、データ調査レポートからより多くのデータをエクスポートできるようになり、データ分析の柔軟性が向上します。

作業環境、ボリューム、保存先フォルダー、JSON または CSV 形式を選択できます。エクスポートされたファイル名には、データがいつエクスポートされたかを識別するのに役立つタイムスタンプが含まれます。

サポートされている作業環境は次のとおりです。

- Cloud Volumes ONTAP
- ONTAP向け FSx

- ONTAP
- 共有グループ

データ調査レポートからのデータのエクスポートには、次の制限があります。

- ダウンロードできるレコードの最大数は、タイプ（ファイル、ディレクトリ、テーブル）ごとに5億件です。
- 100 万件のレコードをエクスポートするには約 35 分かかると予想されます。

データ調査とレポートの詳細については、 "組織内に保存されているデータを調査する" 。

2024年12月16日

バージョン**1.38**

このBlueXP classificationリリースには、一般的な改善とバグ修正が含まれています。

2024年11月4日

バージョン**1.37**

このBlueXP classificationリリースには、次の更新が含まれています。

RHEL 8.10 のサポート

このリリースでは、これまでサポートされていたバージョンに加えて、Red Hat Enterprise Linux v8.10 のサポートも提供されます。これは、ダーク サイトの展開を含む、 BlueXP classificationの手動オンプレミス インストールに適用されます。

次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、 BlueXP classificationバージョン 1.30 以上が必要です: Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、および 9.4。

詳細はこちら "BlueXP classification"。

NFS v4.1 のサポート

このリリースでは、以前サポートされていたバージョンに加えて、NFS v4.1 のサポートも提供されます。

詳細はこちら "BlueXP classification"。

2024年10月10日

バージョン**1.36**

RHEL 9.4 のサポート

このリリースでは、これまでサポートされていたバージョンに加えて、Red Hat Enterprise Linux v9.4 のサポートも提供されます。これは、ダーク サイトの展開を含む、 BlueXP classificationの手動オンプレミス インストールに適用されます。

次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、 BlueXP classificationバージョン 1.30 以上が必要です: Red Hat Enterprise Linux バージョン 8.8、9.0、9.1、9.2、

9.3、および 9.4。

詳細はこちら "BlueXP classification展開の概要"。

スキャンパフォーマンスの向上

このリリースでは、スキャン パフォーマンスが向上しました。

2024年9月2日

バージョン**1.35**

StorageGRIDデータをスキャン

BlueXP classificationは、 StorageGRID内のデータのスキャンをサポートします。

詳細については、"StorageGRIDデータをスキャン" 。

2024年8月5日

バージョン**1.34**

このBlueXP classificationリリースには、次の更新が含まれています。

CentOSから**Ubuntu**への変更

BlueXP classificationは、 Microsoft Azure および Google Cloud Platform (GCP) 向けの Linux オペレーティング システムを CentOS 7.9 から Ubuntu 22.04 に更新しました。

展開の詳細については、 "インターネットにアクセスできる Linux ホストにインストールし、Linux ホスト システムを準備します。" 。

2024年7月1日

バージョン**1.33**

Ubuntuをサポート

このリリースは、Ubuntu 24.04 Linux プラットフォームをサポートしています。

マッピングスキャンはメタデータを収集します

次のメタデータはマッピング スキャン中にファイルから抽出され、ガバナンス、コンプライアンス、調査のダッシュボードに表示されます。

- 労働環境
- 作業環境の種類
- ストレージリポジトリ
- ファイル タイプ
- 使用済み容量
- ファイル数

- ファイル サイズ
- ファイル作成
- ファイルの最終アクセス
- ファイルの最終更新日時
- ファイル発見時刻
- 権限の抽出

ダッシュボードの追加データ

このリリースでは、マッピング スキャン中にガバナンス、コンプライアンス、調査ダッシュボードに表示されるデータが更新されます。

詳細については、 "マッピングスキャンと分類スキャンの違いは何ですか？" 。

2024年6月5日

バージョン**1.32**

構成ページの新しいマッピングステータス列

このリリースでは、構成ページに新しいマッピング ステータス列が表示されるようになりました。新しい列は、マッピングが実行中か、キューに入れられているか、一時停止中かなどを識別するのに役立ちます。

ステータスの説明については、 "スキャン設定を変更する" 。

2024年5月15日

バージョン**1.31**

分類は**BlueXP**のコアサービスとして利用可能

BlueXP classificationは、コネクタごとに最大 500 TiB のスキャン データに対して、追加料金なしでBlueXP内のコア機能として利用できるようになりました。分類ライセンスや有料サブスクリプションは必要ありません。この新しいバージョンでは、 BlueXP classification機能をNetAppストレージ システムのスキャンに重点的に使用しているため、一部の従来の機能は、以前にライセンスを支払った顧客のみが利用できるようになります。これらのレガシー機能の使用は、有料契約の終了日に達すると無効になります。

> データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 + コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

2024年4月1日

バージョン**1.30**

RHEL v8.8 および **v9.3 BlueXP classification**のサポートが追加されました

このリリースでは、Docker エンジンではなく Podman を必要とする、以前サポートされていた 9.x に加え

て、Red Hat Enterprise Linux v8.8 および v9.3 のサポートも提供されます。これは、 BlueXP classification の手動オンプレミス インストールに適用されます。

次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、 BlueXP classificationバージョン 1.30 以上が必要です: Red Hat Enterprise Linux バージョン 8.8、9.0、9.1、9.2、および 9.3。

詳細はこちら "BlueXP classification展開の概要"。

オンプレミスにある RHEL 8 または 9 ホストにコネクタをインストールする場合、 BlueXP classificationがサポートされます。RHEL 8 または 9 ホストが AWS、Azure、または Google Cloud に存在する場合はサポートされません。

監査ログ収集を有効にするオプションが削除されました

監査ログ収集を有効にするオプションが無効になっています。

スキャン速度の向上

セカンダリ スキャナー ノードでのスキャン パフォーマンスが向上しました。スキャンに追加の処理能力が必要な場合は、スキャナー ノードを追加できます。詳細については、 "インターネットにアクセスできるホストにBlueXP classificationをインストールする" 。

自動アップグレード

インターネットにアクセスできるシステムにBlueXP classificationを展開した場合、システムは自動的にアップグレードされます。以前は、最後のユーザーアクティビティから特定の時間が経過した後にアップグレードが実行されていました。このリリースでは、現地時間が午前 1 時から午前 5 時の間であれば、 BlueXP classificationが自動的にアップグレードされます。現地時間がこれらの時間外の場合、最後のユーザーアクティビティから特定の時間が経過した後にアップグレードが行われます。詳細については、 "インターネットにアクセスできる Linux ホストにインストールする" 。

インターネットにアクセスせずにBlueXP classificationを展開した場合は、手動でアップグレードする必要があります。詳細については、 "インターネットにアクセスできない Linux ホストにBlueXP classificationをインストールする" 。

2024年3月4日

バージョン**1.29**

特定のデータソースディレクトリにあるスキャンデータを除外できるようになりました

BlueXP classificationで特定のデータ ソース ディレクトリにあるスキャン データを除外する場合は、 BlueXP classificationが処理する構成ファイルにこれらのディレクトリ名を追加できます。この機能により、不要なディレクトリのスキャンや、誤った個人データ結果が返されるディレクトリのスキャンを回避できます。

"詳細情報" 。

エクストララージインスタンスのサポートが認定されました

2 億 5,000 万を超えるファイルをスキャンするためにBlueXP classificationが必要な場合は、クラウド展開またはオンプレミス インストールで特大インスタンスを使用できます。このタイプのシステムは最大 5 億個のファイルをスキャンできます。

"詳細情報" 。

2024年1月10日

バージョン**1.27**

調査ページの結果には、アイテムの合計数に加えて合計サイズが表示されます。

調査ページのフィルタリングされた結果には、ファイルの合計数に加えて、アイテムの合計サイズが表示されます。これは、ファイルを移動したり、削除したりする場合などに役立ちます。

追加のグループ**ID**を「組織に公開」として設定する

グループが最初にその権限で設定されていなかった場合、NFS 内のグループ ID をBlueXP classificationから直接「組織に公開」として設定できるようになりました。これらのグループ ID が添付されているファイルとフォルダーは、調査の詳細ページで「組織に公開」として表示されます。方法を見る"追加のグループIDを「組織に公開」として追加する"。

2023年12月14日

バージョン**1.26.6**

このリリースにはいくつかのマイナーな機能強化が含まれています。

このリリースでは、次のオプションも削除されました。

- 監査ログ収集を有効にするオプションが無効になっています。
- ディレクトリ調査中は、ディレクトリ別に個人識別情報 (PII) データの数を計算するオプションは使用できません。。 "組織内に保存されているデータを調査する" 。
- Azure Information Protection (AIP) ラベルを使用してデータを統合するオプションが無効になりました。

2023年11月6日

バージョン**1.26.3**

このリリースでは以下の問題が修正されました

- ダッシュボードでシステムによってスキャンされたファイルの数を表示する際の不一致を修正しました。
- 名前とメタデータに特殊文字が含まれるファイルとディレクトリを処理およびレポートすることにより、スキャン動作が改善されました。

2023年10月4日

バージョン**1.26**

RHEL バージョン **9** での**BlueXP classification**のオンプレミス インストールのサポート

Red Hat Enterprise Linux バージョン 8 および 9 は、 BlueXP classificationのインストールに必要な Docker エンジンをサポートしていません。コンテナ インフラストラクチャとして Podman バージョン 4 以上を使用して、RHEL 9.0、9.1、9.2 へのBlueXP classificationのインストールをサポートするようになりました。ご使用の環境で最新バージョンの RHEL を使用する必要がある場合は、Podman を使用するときにBlueXP classification(バージョン 1.26 以上) をインストールできるようになりました。

現時点では、RHEL 9.x を使用する場合、ダーク サイトのインストールまたは分散スキャン環境 (マスター ス

キャナー ノードとリモート スキャナー ノードを使用) はサポートされていません。

2023年9月5日

バージョン**1.25**

小規模および中規模の展開は一時的に利用できません

AWS でBlueXP classificationのインスタンスをデプロイする場合、デプロイ **>** 構成 を選択して小規模または中規模のインスタンスを選択するオプションは現時点では使用できません。**[**デプロイ**] > [**デプロイ**]** を選択すると、大きなインスタンス サイズを使用してインスタンスをデプロイできます。

調査結果ページから最大**10**万件のアイテムにタグを適用します

これまでは、調査結果ページで一度に 1 ページにしかタグを適用できませんでした (20 項目)。調査結果ページですべての項目を選択し、一度に最大 100,000 項目まですべての項目にタグを適用できるようになりました。

最小ファイルサイズが **1 MB** の重複ファイルを識別します

BlueXP classificationは、ファイルが 50 MB 以上の場合にのみ重複ファイルを識別します。1 MB から始まる重複ファイルを識別できるようになりました。調査ページのフィルター「ファイル サイズ」と「重複」を使用すると、環境内で重複している特定のサイズのファイルを確認できます。

2023年7月17日

バージョン**1.24**

BlueXP classificationにより、ドイツの個人データの**2**つの新しいタイプが特定されました

BlueXP classificationでは、次の種類のデータを含むファイルを識別して分類できます。

- ドイツの ID (Personalausweisnummer)
- ドイツの社会保障番号 (Sozialversicherungsnummer)

"BlueXP classificationがあなたのデータ内で識別できるすべての個人データの種類を確認します" 。

BlueXP classificationは制限モードとプライベートモードで完全にサポートされています

BlueXP classificationは、インターネット アクセスがないサイト (プライベート モード) およびインターネットからの送信アクセスが制限されているサイト (制限モード) でも完全にサポートされるようになりました。"コネクタのBlueXP展開モードの詳細" 。

BlueXP classificationのプライベートモードインストールをアップグレードするときにバージョンをスキップする機能

順次的でない場合でも、 BlueXP classificationの新しいバージョンにアップグレードできるようになりました。つまり、 BlueXP classificationを一度に 1 バージョンずつアップグレードするという現在の制限は必要なくなります。この機能はバージョン 1.24 以降で適用されます。

BlueXP classificationAPIが利用可能になりました

BlueXP classificationAPI を使用すると、アクションの実行、クエリの作成、スキャンしているデータに関する情報のエクスポートが可能になります。インタラクティブなドキュメントは Swagger を使用して利用できます。ドキュメントは、調査、コンプライアンス、ガバナンス、構成など、複数のカテゴリに分かれています。各カテゴリは、 BlueXP classificationUI のタブへの参照です。

"BlueXP classificationAPIの詳細" 。

2023年6月6日

バージョン**1.23**

データ主体名の検索時に日本語がサポートされるようになりました

データ主体アクセス要求 (DSAR) に応じて主体の名前を検索するときに、日本語の名前を入力できるようになりました。生成することができます"データ主体アクセス要求レポート"結果の情報とともに。日本語名を入力することもできます"データ調査ページの「データ主体」フィルター"対象者の名前が含まれるファイルを識別します。

Ubuntuは現在、**BlueXP classification**をインストールできる**Linux**ディストリビューションとしてサポートされています。

Ubuntu 22.04 は、 BlueXP classificationのサポート対象オペレーティング システムとして認定されました。インストーラーのバージョン 1.23 を使用する場合、ネットワーク内の Ubuntu Linux ホスト、またはクラウド内の Linux ホストにBlueXP classificationをインストールできます。 "UbuntuがインストールされているホストにBlueXP classificationをインストールする方法をご覧ください" 。

Red Hat Enterprise Linux 8.6 および **8.7** は、新しい**BlueXP classification**のインストールではサポートされなくなりました。

Red Hat は前提条件である Docker をサポートしなくなったため、これらのバージョンは新しいデプロイメントではサポートされません。 RHEL 8.6 または 8.7 で実行されている既存のBlueXP classificationマシンがある場合、 NetApp は引き続きその構成をサポートします。

BlueXP classificationは、 **ONTAP**システムから**FPolicy**イベントを受信する**FPolicy**コレクタとして設定できます。

作業環境内のボリュームで検出されたファイル アクセス イベントについて、 BlueXP classificationシステムでファイル アクセス監査ログを収集できるようにすることができます。 BlueXP classificationでは、作成、読み取り、書き込み、削除、名前の変更、所有者/権限の変更、SACL/DACL の変更といった FPolicy イベントの種類と、ファイルに対してアクションを実行したユーザーをキャプチャできます。

Data Sense BYOLライセンスがダークサイトでサポートされるようになりました

ライセンスが少なくなると通知が届くように、Data Sense BYOL ライセンスをダーク サイトのBlueXP digital walletにアップロードできるようになりました。

2023年4月3日

バージョン**1.22**

新しいデータ検出評価レポート

データ検出評価レポートでは、スキャンされた環境の高レベルの分析が提供され、システムの検出結果が強調表示され、懸念される領域と潜在的な修復手順が示されます。このレポートの目的は、データ セットのデータ ガバナンスの懸念、データ セキュリティの露出、およびデータ コンプライアンスのギャップについての認識を高めることです。 "データ検出評価レポートの生成方法と使用方法をご覧ください" 。

クラウド内の小規模なインスタンスに**BlueXP classification**を展開する機能

AWS 環境でBlueXPコネクタからBlueXP classificationを展開する場合、デフォルトのインスタンスで使用できるものよりも小さい 2 つのインスタンスタイプから選択できるようになりました。小規模な環境をスキャンする場合、クラウド コストを節約できます。ただし、小さいインスタンスを使用する場合は、いくつかの

制限があります。 "利用可能なインスタンスタイプと制限事項を確認する" 。

BlueXP classificationのインストール前に **Linux** システムを認定するためのスタンドアロン スクリプトが利用可能になりました

BlueXP classificationインストールの実行とは別に、Linux システムがすべての前提条件を満たしていることを確認したい場合は、前提条件のみをテストする別のスクリプトをダウンロードできます。 "LinuxホストがblueXP classificationをインストールする準備ができているかどうかを確認する方法をご覧ください" 。

2023年3月7日

バージョン**1.21**

BlueXP classificationUIから独自のカスタムカテゴリを追加できる新機能

BlueXP classificationでは、独自のカスタム カテゴリを追加できるようになりました。これにより、 BlueXP classificationはそれらのカテゴリに適合するファイルを識別します。 BlueXP classificationには多くの "定義済みカテゴリ"この機能を使用すると、カスタム カテゴリを追加して、組織固有の情報がデータ内のどこにあるかを識別することができます。

BlueXP classificationUIからカスタムキーワードを追加できるようになりました

BlueXP classificationには、しばらくの間、将来のスキャンでBlueXP classificationが識別するカスタム キーワードを追加する機能がありました。ただし、キーワードを追加するには、 BlueXP classificationLinux ホストにログインし、コマンド ライン インターフェイスを使用する必要がありました。このリリースでは、カスタム キーワードを追加する機能がBlueXP classificationUI に組み込まれ、これらのキーワードの追加と編集が非常に簡単になりました。

「最終アクセス時間」が変更された場合、 **BlueXP classification**でファイルをスキャンしないようにする機能

デフォルトでは、 BlueXP classificationに適切な「書き込み」権限がない場合、 BlueXP classificationは「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。ただし、ファイルの最終アクセス時刻が元の時刻にリセットされても問題がない場合は、構成ページでこの動作を上書きして、 BlueXP classificationが権限に関係なくボリュームをスキャンするようにすることができます。

この機能と連動して、「スキャン分析イベント」という新しいフィルターが追加され、 BlueXP classificationで最終アクセス時間を戻せなかったために分類されなかったファイル、またはBlueXP classificationで最終アクセス時間を戻せなかったにもかかわらず分類されたファイルを表示できるようになりました。

"「最終アクセスタイムスタンプ」とBlueXP classificationに必要な権限について詳しくは、こちらをご覧ください。" 。

BlueXP classificationにより**3**つの新しいタイプの個人データが識別される

BlueXP classificationでは、次の種類のデータを含むファイルを識別して分類できます。

- ボツワナ ID カード (オマン) 番号
- ボツワナのパスポート番号
- シンガポール国民登録身分証明書（NRIC）

"BlueXP classificationがあなたのデータ内で識別できるすべての個人データの種類を確認します" 。

ディレクトリの機能を更新しました

- データ調査レポートの「軽量 CSV レポート」オプションに、ディレクトリからの情報が含まれるように

なりました。

- 「最終アクセス」時間フィルターに、ファイルとディレクトリの両方の最終アクセス時間が表示されるようになりました。

インストールの機能強化

- インターネットにアクセスできないサイト (ダークサイト) 用のBlueXP classificationインストーラーは、インストールを正常に実行するために必要なシステムとネットワークの要件が満たされているかどうかを確認するための事前チェックを実行するようになりました。
- インストール監査ログファイルは保存され、次の場所に書き込まれます。 `/ops/netapp/install_logs` 。

2023年2月5日

バージョン**1.20**

ポリシーベースの通知メールを任意のメールアドレスに送信する機能

BlueXP classificationの以前のバージョンでは、特定の重要なポリシーが結果を返すときに、アカウント内のBlueXPユーザーに電子メールアラートを送信できました。この機能を使用すると、オンラインでないときにデータを保護するための通知を受け取ることができます。また、ポリシーから、 BlueXPアカウントに登録されていない他のユーザー (最大 20 件の電子メール アドレス) に電子メール アラートを送信できるようになりました。

"ポリシー結果に基づいてメールアラートを送信する方法の詳細" 。

BlueXP classificationUIから個人パターンを追加できるようになりました

BlueXP classificationには、しばらくの間、将来のスキャンでBlueXP classificationが識別するカスタム「個人データ」を追加する機能がありました。ただし、カスタム パターンを追加するには、 BlueXP classificationLinux ホストにログインし、コマンド ラインを使用する必要がありました。このリリースでは、正規表現を使用して個人パターンを追加する機能がBlueXP classificationUI に組み込まれ、これらのカスタム パターンの追加と編集が非常に簡単になりました。

BlueXP classificationを使用して**1500**万個のファイルを移動する能力

以前は、 BlueXP classificationによって最大 100,000 個のソース ファイルを任意の NFS 共有に移動できました。一度に最大 1500 万個のファイルを移動できるようになりました。

SharePoint Online ファイルにアクセスできるユーザーの数を確認する機能

「アクセス権を持つユーザーの数」フィルターは、SharePoint Online リポジトリに保存されているファイルをサポートするようになりました。以前は、CIFS 共有上のファイルのみがサポートされていました。現時点では、アクティブ ディレクトリ ベースではない SharePoint グループはこのフィルターではカウントされないことに注意してください。

アクションステータスパネルに新しい「部分的成功」ステータスが追加されました

新しい「部分的な成功」ステータスは、 BlueXP classificationアクションが完了し、一部の項目は失敗し、一部の項目は成功したことを示します (たとえば、100 個のファイルを移動または削除する場合)。さらに、「完了」ステータスの名前が「成功」に変更されました。以前は、「完了」ステータスに成功したアクションと失敗したアクションがリストされることがありました。ここで、「成功」ステータスは、すべてのアイテムに対するすべてのアクションが成功したことを意味します。 "アクションステータスパネルの表示方法を確認する" 。

2023年1月9日

バージョン**1.19**

機密データを含むファイルと過度に許可されているファイルのチャートを表示する機能

ガバナンス ダッシュボードに、機密データ (機密データと機密個人データの両方を含む) を含むファイルと過度に許可されているファイルのヒートマップを提供する新しい [機密データと幅広いアクセス許可] 領域が追加されました。これにより、機密データにリスクがある可能性がある場所を確認するのに役立ちます。 "詳細情報" 。

データ調査ページで**3**つの新しいフィルターが利用可能になりました

データ調査ページに表示される結果を絞り込むための新しいフィルターが利用可能になりました。

- 「アクセス権を持つユーザー数」フィルターは、特定の数のユーザーに公開されているファイルとフォルダーを表示します。数値の範囲を選択して結果を絞り込むことができます。たとえば、51 ～ 100 人のユーザーがアクセスできるファイルを確認することができます。
- 「作成時刻」、「検出時刻」、「最終更新日時」、および「最終アクセス日時」フィルターでは、事前定義された日数の範囲を選択するだけでなく、カスタムの日付範囲を作成できるようになりました。たとえば、「作成日時」が「6 か月以上前」のファイルや、「最終更新日時」が「過去 10 日以内」のファイルを検索できます。
- 「ファイル パス」フィルターを使用すると、フィルターされたクエリ結果から除外するパスを指定できるようになりました。特定のデータを含めるパスと除外するパスの両方を入力すると、 BlueXP classificationは最初に含めるパス内のすべてのファイルを検索し、次に除外するパスからファイルを削除して、結果を表示します。

"データを調査するために使用できるすべてのフィルターのリストを表示します" 。

BlueXP classificationは日本のマイナンバーを識別できる

BlueXP classificationは、日本の個人番号 (マイナンバーとも呼ばれます) を含むファイルを識別して分類できます。これには、個人および法人のマイナンバーの両方が含まれます。 "BlueXP classificationがあなたのデータ内で識別できるすべての個人データの種類を確認します" 。

NetApp Data Classificationの既知の制限

既知の制限事項では、このリリースではサポートされていない機能や正しく相互運用されない機能が特定されます。これらの制限事項をよく確認してください。

NetApp Data Classification の無効化オプション

2023 年 12 月 (バージョン 1.26.6) リリースでは、次のオプションが削除されました。

- 監査ログ収集を有効にするオプションが無効になっています。
- ディレクトリ調査中は、ディレクトリ別に個人識別情報 (PII) データの数を計算するオプションは使用できません。
- Azure Information Protection (AIP) ラベルを使用してデータを統合するオプションが無効になりました。

データ分類スキャン

データ分類スキャンでは次の制限が発生します。

データ分類はボリューム内の**1**つの共有のみをスキャンします

1 つのボリュームの下に複数のファイル共有がある場合、データ分類は最上位階層の共有をスキャンします。たとえば、次のような共有があるとします。

- /A
- /A/B
- /C
- /D/E

この構成では、/A のデータのみがスキャンされます。 /C および /D 内のデータはスキャンされません。

回避策

ボリューム内のすべての共有からデータをスキャンしていることを確認するための回避策があります。次の手順を実行します。

1. システムに、スキャンするボリュームを追加します。
2. データ分類がボリュームのスキャンを完了したら、[データ調査] ページに移動し、スキャンされている共有を確認するためのフィルターを作成します。

   「システム名」と「ディレクトリ タイプ = 共有」でデータをフィルターし、スキャンされている共有を確認します。

3. ボリューム内に存在する共有の完全なリストを取得して、スキャンされていない共有を確認できます。
4. "残りの共有を共有グループに追加する" 。

   すべての共有を個別に追加します。例:

   ```
   /C
   /D
   ```

5. 複数の共有があるシステム内の各ボリュームに対して、これらの手順を実行します。

最終アクセスタイムスタンプ

データ分類がディレクトリのスキャンを実行すると、そのスキャンはディレクトリの 最終アクセス フィールドに影響します。 最終アクセス フィールドを表示すると、そのメタデータにはスキャンの日時またはユーザーがディレクトリに最後にアクセスした時刻が反映されます。

始めましょう

NetApp Data Classificationについて学ぶ

NetApp Data Classification は、 NetApp Consoleのデータ ガバナンス サービスであり、企業のオンプレミスおよびクラウド データ ソースをスキャンしてデータをマッピングおよび分類し、個人情報を識別します。これにより、セキュリティとコンプライアンスのリスクが軽減され、ストレージ コストが削減され、データ移行プロジェクトが支援されます。

バージョン 1.31 以降、データ分類はNetApp Console内のコア機能として利用できます。追加料金はかかりません。分類ライセンスやサブスクリプションは必要ありません。 + 旧バージョン 1.30 以前を使用している場合は、サブスクリプションの有効期限が切れるまでそのバージョンを利用できます。

NetApp Console

データ分類には、NetApp Consoleからアクセスできます。

NetApp Consoleは、オンプレミスとクラウド環境全体にわたるエンタープライズ グレードのNetAppストレージとデータ サービスの集中管理を提供します。NetAppデータ サービスにアクセスして使用するには、コンソールが必要です。管理インターフェースとして、1 つのインターフェースから多数のストレージ リソースを管理できます。コンソール管理者は、企業内のすべてのシステムのストレージとサービスへのアクセスを制御できます。

NetApp Consoleの使用を開始するためにライセンスやサブスクリプションは必要ありません。ストレージ システムまたはNetAppデータ サービスへの接続を確保するためにクラウドにコンソール エージェントを展開する必要がある場合にのみ料金が発生します。ただし、コンソールからアクセスできる一部のNetAppデータ サービスは、ライセンスまたはサブスクリプションベースです。

詳細はこちら"NetApp Console"。

機能

データ分類では、人工知能 (AI)、自然言語処理 (NLP)、機械学習 (ML) を使用してスキャンしたコンテンツを理解し、エンティティを抽出してそれに応じてコンテンツを分類します。これにより、データ分類では次の機能領域が提供できるようになります。

"データ分類のユースケースについて学ぶ" 。

コンプライアンスを維持する

データ分類では、コンプライアンスの取り組みに役立ついくつかのツールが提供されます。データ分類を使用すると次のことができます。

- 個人を特定できる情報 (PII) を特定します。
- GDPR、CCPA、PCI、HIPAA のプライバシー規制で要求される、広範囲にわたる機密個人情報を特定します。
- 名前または電子メール アドレスに基づいて、データ主体アクセス要求 (DSAR) に応答します。

セキュリティを強化する

データ分類により、犯罪目的でアクセスされる危険性があるデータを識別できます。データ分類を使用すると次のことができます。

- 組織全体または一般に公開されている、オープン権限を持つすべてのファイルとディレクトリ (共有とフォルダー) を識別します。
- 最初の専用場所の外部に存在する機密データを識別します。
- データ保持ポリシーに準拠します。
- *Policies* を使用すると、新しいセキュリティ問題が自動的に検出され、セキュリティ スタッフがすぐに対処できるようになります。

ストレージ使用量を最適化する

データ分類は、ストレージの総所有コスト (TCO) の削減に役立つツールを提供します。データ分類を使用すると次のことができます。

- 重複データやビジネスに関係のないデータを識別して、ストレージ効率を向上させます。
- 非アクティブなデータを特定し、より安価なオブジェクト ストレージに階層化することで、ストレージコストを節約します。 "Cloud Volumes ONTAPシステムの階層化について詳しくは" 。 "オンプレミスのONTAPシステムからの階層化の詳細" 。

サポートされているシステムとデータソース

データ分類では、次の種類のシステムおよびデータ ソースからの構造化データと非構造化データをスキャンして分析できます。

システム

- Amazon FSx for NetApp ONTAP管理
- Azure NetApp Files
- Cloud Volumes ONTAP （AWS、Azure、または GCP にデプロイ）
- オンプレミスのONTAPクラスタ
- StorageGRID
- Google Cloud NetApp Volumes

データソース

- NetAppファイル共有
- データベース:
  - Amazon リレーショナルデータベースサービス (Amazon RDS)
  - MongoDB
  - MySQL
  - Oracle
  - PostgreSQL
  - SAP HANA

- SQL サーバー (MSSQL)

データ分類では、NFS バージョン 3.x、4.0、4.1、および CIFS バージョン 1.x、2.0、2.1、3.0 がサポートされています。

料金

データ分類は無料でご利用いただけます。分類ライセンスや有料サブスクリプションは必要ありません。

インフラコスト

- クラウドにデータ分類をインストールするには、クラウド インスタンスを展開する必要があり、展開先のクラウド プロバイダーから料金が発生します。見る各クラウドプロバイダーに展開されるインスタンスの種類 。オンプレミス システムに Data Classification をインストールする場合、料金はかかりません。
- データ分類では、コンソール エージェントを展開する必要があります。多くの場合、コンソールで使用している他のストレージやサービスがあるため、既にコンソール エージェントが存在します。コンソール エージェント インスタンスには、デプロイされているクラウド プロバイダーからの料金が発生します。参照 "各クラウドプロバイダーに展開されるインスタンスの種類"。オンプレミス システムにコンソール エージェントをインストールする場合、料金はかかりません。

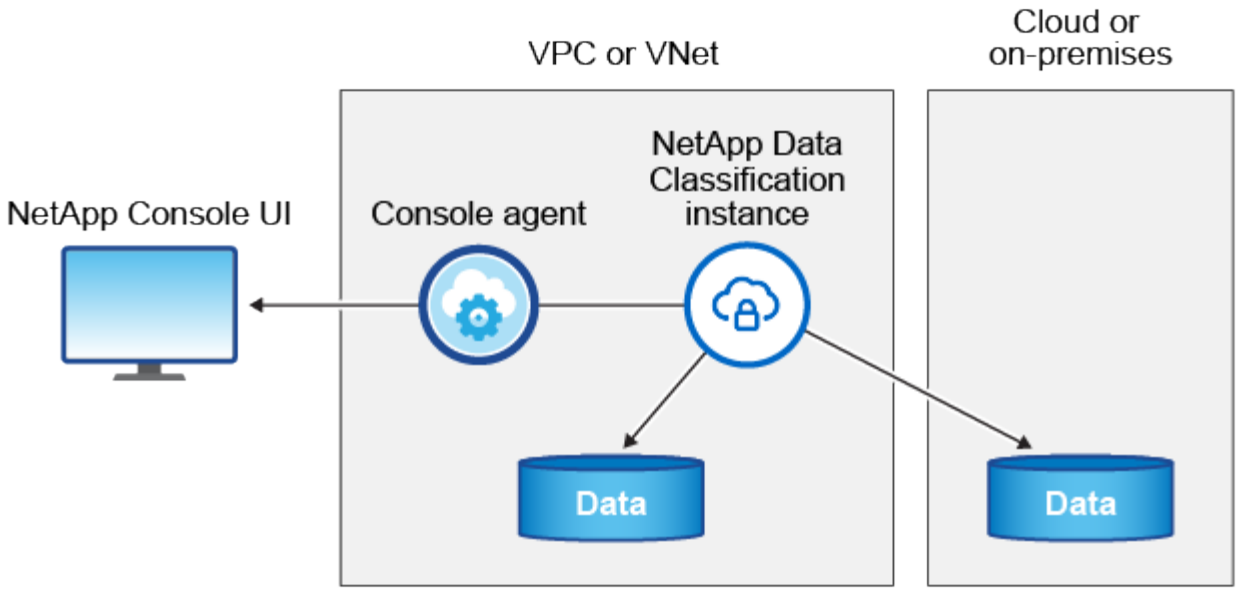
データ転送コスト

データ転送コストは設定によって異なります。データ分類インスタンスとデータ ソースが同じアベイラビリティー ゾーンとリージョンにある場合、データ転送コストは発生しません。ただし、 Cloud Volumes ONTAP システムなどのデータ ソースが別のアベイラビリティ ゾーンまたはリージョンにある場合は、クラウド プロバイダーからデータ転送コストが請求されます。詳細については、次のリンクを参照してください。

- "AWS: Amazon Elastic Compute Cloud (Amazon EC2) の料金"
- "Microsoft Azure: 帯域幅の料金詳細"
- "Google Cloud: ストレージ転送サービスの料金"

データ分類インスタンス

クラウドにデータ分類をデプロイすると、コンソールはコンソール エージェントと同じサブネットにインスタンスをデプロイします。 "コンソール エージェントの詳細について説明します。"



デフォルトインスタンスについては次の点に注意してください。

- AWSでは、データ分類は "m6i.4xlargeインスタンス"500 GiB GP2 ディスク付き。オペレーティング システム イメージは Amazon Linux 2 です。 AWS にデプロイする場合、少量のデータをスキャンする場合は、より小さいインスタンス サイズを選択できます。
- Azureでは、データ分類は"Standard_D16s_v3 VM"500 GiB のディスクを搭載。オペレーティング システム イメージは Ubuntu 22.04 です。
- GCPでは、データ分類は"n2-標準-16 VM"500 GiB の標準永続ディスクを備えています。オペレーティング システム イメージは Ubuntu 22.04 です。
- デフォルトのインスタンスが利用できないリージョンでは、データ分類は代替インスタンスで実行されます。"代替インスタンスタイプを参照" 。
- インスタンスの名前は *CloudCompliance* となり、生成されたハッシュ (UUID) が連結されます。例: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- コンソール エージェントごとに 1 つのデータ分類インスタンスのみが展開されます。

また、オンプレミスの Linux ホストまたは優先クラウド プロバイダーのホストにデータ分類を展開することもできます。どのインストール方法を選択しても、ソフトウェアはまったく同じように機能します。インスタンスがインターネットにアクセスできる限り、データ分類ソフトウェアのアップグレードは自動化されます。

データ分類は継続的にデータをスキャンするため、インスタンスは常に実行されたままにしておく必要があります。

異なるインスタンスタイプにデプロイ

インスタンス タイプの次の仕様を確認してください。

システムサイズ	仕様	制限事項
特大	32 個の CPU、128 GB の RAM、1 TiB の SSD	最大5億個のファイルをスキャンできます。

システムサイズ	仕様	制限事項
大（デフォルト）	16 CPU、64 GB RAM、500 GiB SSD	最大2億5000万個のファイルをスキャンできます。

Azure または GCP でデータ分類をデプロイするときに、より小さいインスタンス タイプを使用したい場合は、ng-contact-data-sense@netapp.com に電子メールでお問い合わせください。

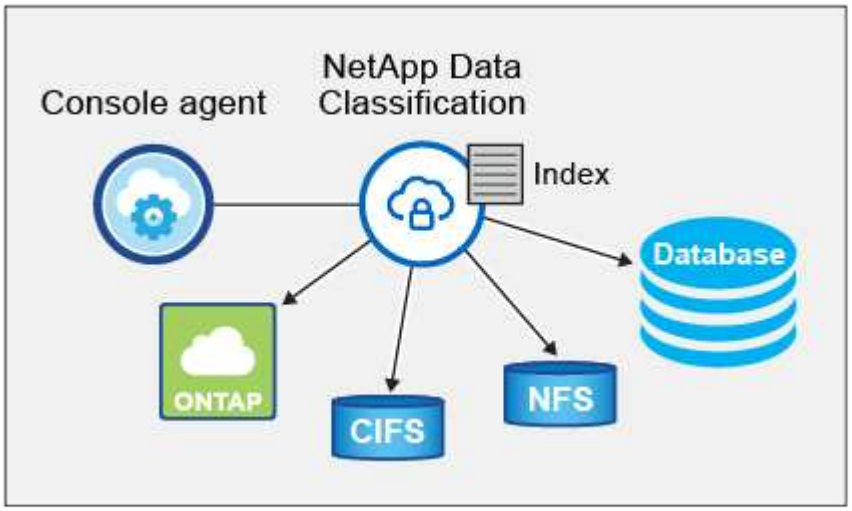
データ分類スキャンの仕組み

大まかに言えば、データ分類スキャンは次のように機能します。

1. コンソールでデータ分類のインスタンスをデプロイします。
2. 1 つ以上のデータ ソースに対して、高レベル マッピング (マッピングのみ スキャンと呼ばれる) または詳細レベル スキャン (マップと分類 スキャンと呼ばれる) を有効にします。
3. データ分類は、AI 学習プロセスを使用してデータをスキャンします。
4. 提供されているダッシュボードとレポート ツールを使用して、コンプライアンスとガバナンスの取り組みを支援します。

データ分類を有効にし、スキャンするリポジトリ (ボリューム、データベース スキーマ、またはその他のユーザー データ) を選択すると、すぐにデータのスキャンが開始され、個人データと機密データが識別されます。ほとんどの場合、バックアップ、ミラー、または DR サイトではなく、ライブの本番データのスキャンに重点を置く必要があります。次に、データ分類によって組織のデータがマッピングされ、各ファイルが分類され、データ内のエンティティと定義済みパターンが識別および抽出されます。スキャンの結果は、個人情報、機密個人情報、データ カテゴリ、およびファイル タイプのインデックスです。

Data Classification は、NFS および CIFS ボリュームをマウントすることで、他のクライアントと同様にデータに接続します。 NFS ボリュームは自動的に読み取り専用としてアクセスされますが、CIFS ボリュームをスキャンするには Active Directory の資格情報を提供する必要があります。



最初のスキャンの後、データ分類はラウンドロビン方式でデータを継続的にスキャンし、増分変更を検出します。そのため、インスタンスを実行し続けることが重要です。

ボリューム レベルまたはデータベース スキーマ レベルでスキャンを有効または無効にすることができます。

データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 + コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

マッピングスキャンと分類スキャンの違いは何ですか？

データ分類では、次の 2 種類のスキャンを実行できます。

- マッピングのみのスキャン は、データの概要のみを提供し、選択されたデータ ソースに対して実行されます。マッピングのみのスキャンでは、ファイルにアクセスして内部のデータを確認する必要がないため、マップおよび分類スキャンよりも時間がかかりません。最初にこれを実行して研究領域を特定し、次にそれらの領域に対してマップと分類のスキャンを実行することをお勧めします。
- マップと分類スキャン は、データの詳細なスキャンを提供します。

マッピングスキャンと分類スキャンの違いの詳細については、以下を参照してください。"マッピングスキャンと分類スキャンの違いは何ですか?" 。

データ分類が分類する情報

データ分類では、次のデータを収集し、インデックスを付け、カテゴリを割り当てます。

- ファイルに関する*標準メタデータ*: ファイルの種類、サイズ、作成日と変更日など。
- 個人データ: 電子メール アドレス、識別番号、クレジットカード番号などの個人を特定できる情報 (PII)。データ分類では、ファイル内の特定の単語、文字列、パターンを使用してこれを識別します。"個人データについて詳しくはこちら" 。
- 機密個人データ: 一般データ保護規則 (GDPR) やその他のプライバシー規制で定義されている、健康データ、民族的出身、政治的意見などの特別な種類の機密個人情報 (SPII)。"機密性の高い個人データについて詳しく見る" 。
- カテゴリ: データ分類は、スキャンしたデータを取得し、それをさまざまな種類のカテゴリに分割します。カテゴリは、各ファイルのコンテンツとメタデータの AI 分析に基づくトピックです。"カテゴリーについて詳しく見る"。
- 名前エンティティ認識: データ分類では、AI を使用して文書から人の自然な名前を抽出します。"データ主体のアクセス要求への対応について学ぶ" 。

ネットワークの概要

データ分類では、クラウドまたはオンプレミスなど、任意の場所に単一のサーバーまたはクラスターを展開します。サーバーは標準プロトコルを介してデータ ソースに接続し、同じサーバーにデプロイされている Elasticsearch クラスターで検出結果をインデックス化します。これにより、マルチクラウド、クロスクラウド、プライベートクラウド、オンプレミス環境のサポートが可能になります。

コンソールは、コンソール エージェントからの受信 HTTP 接続を有効にするセキュリティ グループを使用して、データ分類インスタンスをデプロイします。

コンソールを SaaS モードで使用する場合、コンソールへの接続は HTTPS 経由で提供され、ブラウザとデータ分類インスタンス間で送信されるプライベート データは TLS 1.2 を使用したエンドツーエンドの暗号化で

保護されるため、 NetAppやサードパーティが読み取ることはできません。

アウトバウンドルールは完全にオープンです。データ分類ソフトウェアをインストールおよびアップグレードし、使用状況メトリックを送信するには、インターネット アクセスが必要です。

厳しいネットワーク要件がある場合、"データ分類が接続するエンドポイントについて学習する" 。

NetApp Data Classificationにアクセス

NetApp Consoleを通じてNetApp Data Classification にアクセスできます。

コンソールにサインインするには、 NetAppサポート サイトの認証情報を使用するか、電子メールとパスワードを使用してNetApp Consoleログインにサインアップすることができます。"コンソールへのログインについて詳しくは" 。

特定のタスクには、特定のコンソール ユーザー ロールが必要です。"すべてのサービスのコンソールアクセス ロールについて学習します" 。

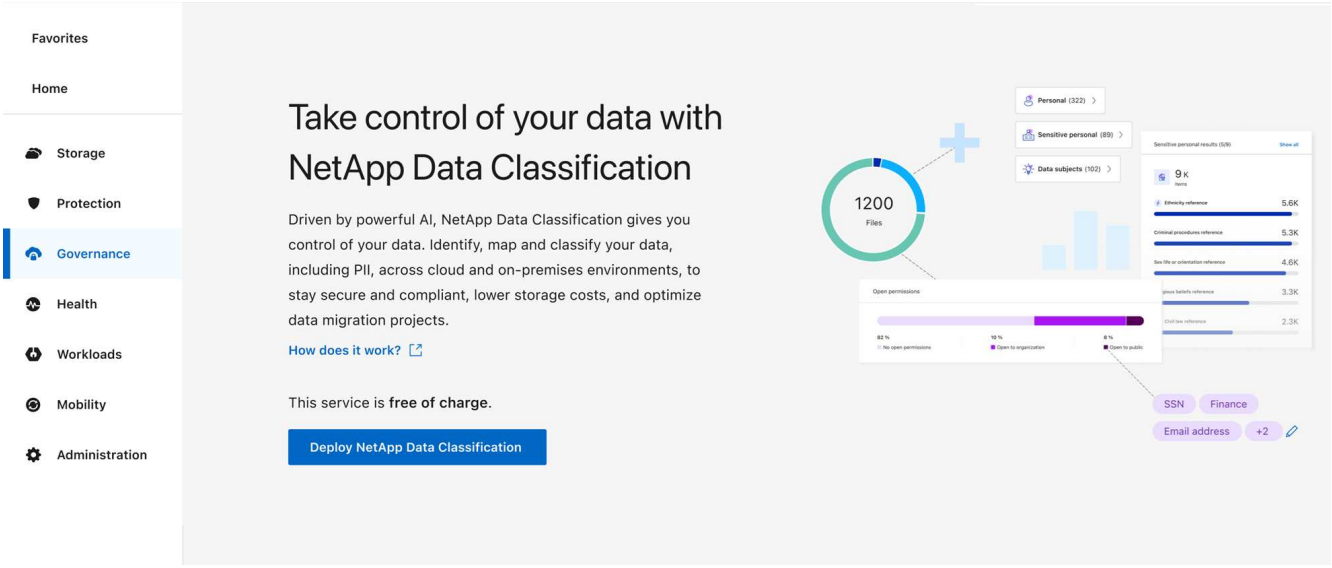
開始する前に

- "コンソール エージェントを追加する必要があります。"
- "どのデータ分類展開スタイルがワークロードに適しているかを理解します。"

手順

1. ウェブブラウザで、"コンソール" 。
2. コンソールにログインします。
3. NetApp Consoleのメイン ページから、ガバナンス > データ分類 を選択します。
4. データ分類に初めてアクセスする場合は、ランディング ページが表示されます。

   分類インスタンスの展開を開始するには、*オンプレミスまたはクラウドでの分類の展開*を選択します。詳細については、"どのデータ分類展開を使用する必要がありますか?"



   それ以外の場合は、データ分類ダッシュボードが表示されます。

データ分類を展開する

どの**NetApp Data Classification**展開を使用する必要がありますか**?**

NetApp Data Classification はさまざまな方法で導入できます。どの方法がニーズに合っているかを学びます。

データ分類は次の方法で展開できます。

- "コンソールを使用してクラウドにデプロイする" 。コンソールは、コンソール エージェントと同じクラウド プロバイダー ネットワークにデータ分類インスタンスを展開します。
- "インターネットにアクセスできる Linux ホストにインストールする" 。インターネットにアクセスできるネットワーク内の Linux ホスト、またはクラウド内の Linux ホストに Data Classification をインストールします。このタイプのインストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションですが、必須ではありません。
- "インターネットにアクセスできないオンプレミス サイトの Linux ホストにインストールする"プライベート モード とも呼ばれます。インストール スクリプトを使用するこのタイプのインストールでは、コンソール SaaS レイヤーへの接続はありません。

> BlueXPプライベート モード (レガシーBlueXPインターフェイス) は通常、インターネット接続がなく、AWS Secret Cloud、AWS Top Secret Cloud、Azure IL6 などの安全なクラウド領域があるオンプレミス環境で使用されます。NetApp は、従来のBlueXPインターフェースを使用してこれらの環境を引き続きサポートします。従来のBlueXPインターフェースのプライベートモードのドキュメントについては、"BlueXPプライベートモードの PDF ドキュメント" 。

インターネットにアクセスできる Linux ホストへのインストールと、インターネットにアクセスできない Linux ホストへのオンプレミス インストールの両方で、インストール スクリプトが使用されます。スクリプトは、システムと環境が前提条件を満たしているかどうかを確認することから始まります。前提条件が満たされている場合は、インストールが開始されます。データ分類のインストールを実行せずに前提条件を検証したい場合は、前提条件のみをテストする別のソフトウェア パッケージをダウンロードできます。

。 "Linuxホストがデータ分類をインストールする準備ができていることを確認します" 。

NetApp Consoleを使用してクラウドに**NetApp Data Classification**を導入する

NetApp Consoleを使用して、クラウドにNetApp Data Classification を導入できます。コンソールは、コンソール エージェントと同じクラウド プロバイダー ネットワークにデータ分類インスタンスを展開します。

また、"インターネットにアクセスできる Linux ホストにデータ分類をインストールする" 。このタイプのインストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションですが、必須ではありません。どのインストール方法を選択しても、ソフトウェアはまったく同じように機能します。

クイック スタート

以下の手順に従ってすぐに開始するか、残りのセクションまでスクロールして詳細を確認してください。

1 コンソールエージェントを作成する

コンソール エージェントがまだない場合は作成します。見る"AWSでコンソールエージェントを作成する" 、"Azureでコンソールエージェントを作成する" 、 または"GCP でコンソール エージェントを作成する"。

また、"オンプレミスにコンソールエージェントをインストールする"ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。

2 前提条件

お使いの環境が前提条件を満たしていることを確認してください。これには、インスタンスの外向きインターネットアクセス、ConsoleエージェントとData Classification間のポート443での接続などが含まれます。完全なリストを見る.

3 データ分類を展開する

インストール ウィザードを起動して、データ分類インスタンスをクラウドにデプロイします。

コンソールエージェントを作成する

コンソール エージェントがまだない場合は、クラウド プロバイダーにコンソール エージェントを作成します。見る "AWSでコンソールエージェントを作成する"または "Azureでコンソールエージェントを作成する"、または "GCP でコンソール エージェントを作成する"。ほとんどの場合、データ分類を有効化する前にコンソールエージェントをセットアップしておく必要があります。 "コンソール機能にはコンソールエージェントが必要です"ただし、今すぐ設定する必要がある場合もあります。

特定のクラウド プロバイダーにデプロイされたコンソール エージェントを使用する必要があるシナリオがいくつかあります。

- AWS のCloud Volumes ONTAPまたはAmazon FSx for ONTAPバケット内のデータをスキャンする場合は、AWS のコンソールエージェントを使用します。
- Azure のCloud Volumes ONTAPまたはAzure NetApp Filesでデータをスキャンする場合は、Azure のコンソール エージェントを使用します。
  - Azure NetApp Filesの場合、スキャンするボリュームと同じリージョンにデプロイする必要があります。
- GCP のCloud Volumes ONTAPでデータをスキャンする場合は、GCP のコンソール エージェントを使用します。

これらのクラウド コンソール エージェントのいずれかを使用すると、オンプレミスのONTAPシステム、NetAppファイル共有、およびデータベースをスキャンできます。

また、 "オンプレミスにコンソールエージェントをインストールする"ネットワークまたはクラウド内の Linux ホスト上。オンプレミスでデータ分類をインストールする予定のユーザーの中には、オンプレミスでコンソール エージェントをインストールすることを選択する場合もあります。

使用する必要がある状況があるかもしれません "複数のコンソールエージェント"。

データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 + コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

政府地域支援

データ分類は、コンソール エージェントが政府リージョン (AWS GovCloud、Azure Gov、または Azure DoD) にデプロイされている場合にサポートされます。この方法で展開する場合、データ分類には次の制限があります。

"政府地域におけるコンソールエージェントの導入について学習します"。

前提条件

クラウドにデータ分類を展開する前に、次の前提条件を確認して、サポートされている構成があることを確認してください。クラウドにデータ分類を展開すると、コンソール エージェントと同じサブネットに配置されます。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。プロキシは非透過である必要があります。透過プロキシは現在サポートされていません。

データ分類を AWS、Azure、GCP のいずれに展開するかに応じて、以下の適切な表を確認してください。

AWSに必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供します。
https://kinesis.us-east-1.amazonaws.com	NetApp が監査レコードからデータをストリーミングできるようにします。
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	データ分類を有効にして、マニフェストとテンプレートにアクセスしてダウンロードし、ログとメトリックを送信できるようにします。

Azureに必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
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://dseasb33srnrn.cloudfront.net/ https://production.cloudflare.docker.com/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。

GCP に必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。

エンドポイント	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://dseasb33srnrn.cloudfront.net/ https://production.cloudflare.docker.com/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。

データ分類に必要な権限があることを確認する

データ分類に、リソースをデプロイし、データ分類インスタンスのセキュリティ グループを作成するための権限があることを確認します。

- "Google Cloud の権限"
- "AWS 権限"
- "Azure のアクセス許可"

コンソールエージェントがデータ分類にアクセスできることを確認する

コンソール エージェントとデータ分類インスタンス間の接続を確認します。コンソール エージェントのセキュリティ グループは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。この接続により、データ分類インスタンスのデプロイが可能になり、コンプライアンス タブとガバナンス タブで情報を表示できるようになります。データ分類は、AWS および Azure の政府リージョンでサポートされています。

AWS および AWS GovCloud のデプロイメントには、追加の受信および送信セキュリティ グループ ルールが必要です。見る "AWS のコンソールエージェントのルール"詳細については。

Azure および Azure Government の展開には、追加の受信および送信セキュリティ グループ ルールが必要です。見る "Azure のコンソール エージェントのルール"詳細については。

データ分類を継続して実行できるようにする

データを継続的にスキャンするには、データ分類インスタンスをオンのままにしておく必要があります。

データ分類への**Web**ブラウザ接続を確保する

データ分類を有効にした後、ユーザーがデータ分類インスタンスに接続しているホストからコンソール インターフェイスにアクセスすることを確認します。

データ分類インスタンスは、インデックス付けされたデータがインターネットからアクセスできないようにするためにプライベート IP アドレスを使用します。そのため、コンソールにアクセスするために使用する Web ブラウザは、そのプライベート IP アドレスに接続している必要があります。この接続は、クラウド プロバイダーへの直接接続 (VPN など) から行うことも、データ分類インスタンスと同じネットワーク内にあるホストから行うこともできます。

vCPUの制限を確認する

クラウド プロバイダーの vCPU 制限によって、必要な数のコアを持つインスタンスのデプロイが許可されていることを確認します。コンソールが実行されているリージョン内の関連するインスタンス ファミリの vCPU 制限を確認する必要があります。"必要なインスタンスタイプを確認する" 。

vCPU 制限の詳細については、次のリンクを参照してください。

- "AWS ドキュメント: Amazon EC2 サービスクォータ"
- "Azure ドキュメント: 仮想マシンの vCPU クォータ"
- "Google Cloud ドキュメント: リソース割り当て"

クラウドでデータ分類を展開

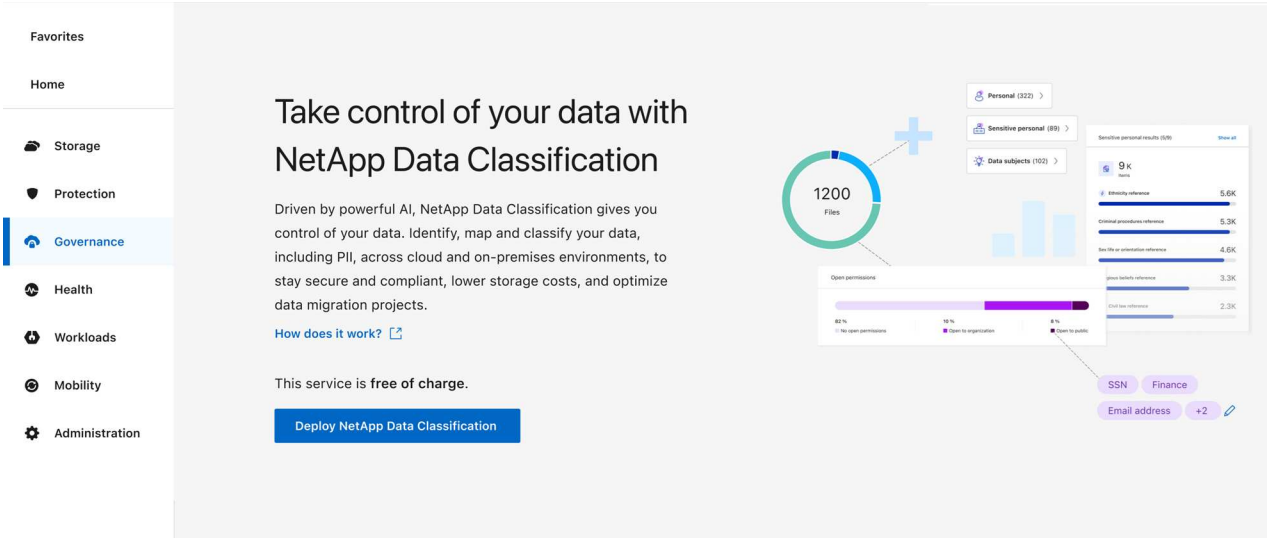
クラウドにデータ分類のインスタンスをデプロイするには、次の手順に従います。コンソール エージェントはクラウドにインスタンスを展開し、そのインスタンスにデータ分類ソフトウェアをインストールします。

デフォルトのインスタンスタイプが利用できない地域では、データ分類は"代替インスタンスタイプ"。

AWSにデプロイ

手順

1. データ分類のメイン ページから、オンプレミスまたはクラウドでの分類の展開 を選択します。

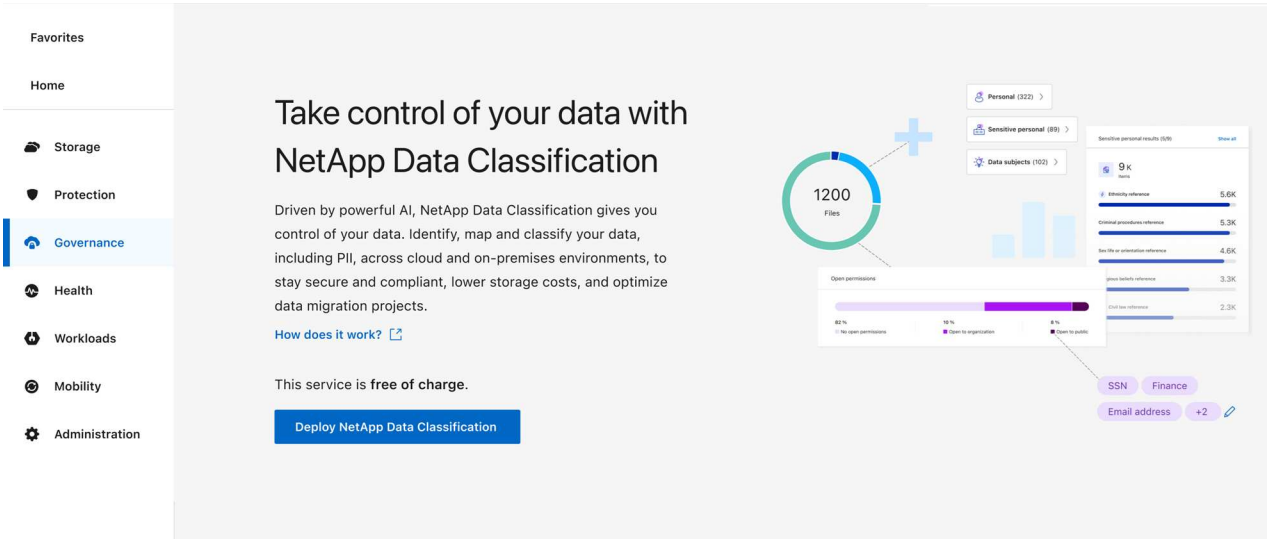


2. インストール ページで、デプロイ > デプロイ を選択して、「大」インスタンス サイズを使用し、クラウド デプロイ ウィザードを起動します。
3. ウィザードは、展開手順を実行する際の進行状況を表示します。入力が必要な場合、または問題が発生した場合には、プロンプトが表示されます。
4. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

Azureにデプロイする

手順

1. データ分類のメイン ページから、オンプレミスまたはクラウドでの分類の展開 を選択します。



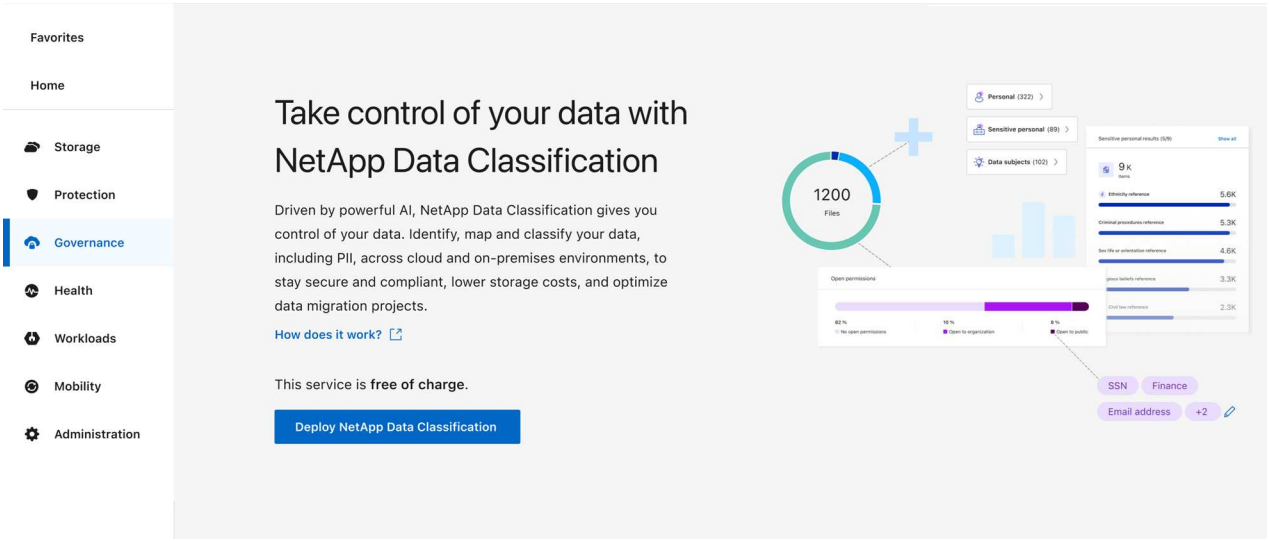
2. デプロイ を選択して、クラウド デプロイ ウィザードを起動します。

3. ウィザードは、展開手順を実行する際の進行状況を表示します。問題が発生すると停止し、入力を求められます。
4. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

Google Cloud にデプロイ

手順

1. データ分類のメイン ページから、ガバナンス > 分類 を選択します。
2. *オンプレミスまたはクラウドでの分類の展開*を選択します。



3. デプロイ を選択して、クラウド デプロイ ウィザードを起動します。
4. ウィザードは、展開手順を実行する際の進行状況を表示します。問題が発生すると停止し、入力を求められます。
5. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

結果

コンソールは、クラウド プロバイダーにデータ分類インスタンスをデプロイします。

インスタンスがインターネットに接続されている限り、コンソール エージェントとデータ分類ソフトウェアへのアップグレードは自動的に行われます。

次は何？

構成ページから、スキャンするデータ ソースを選択できます。

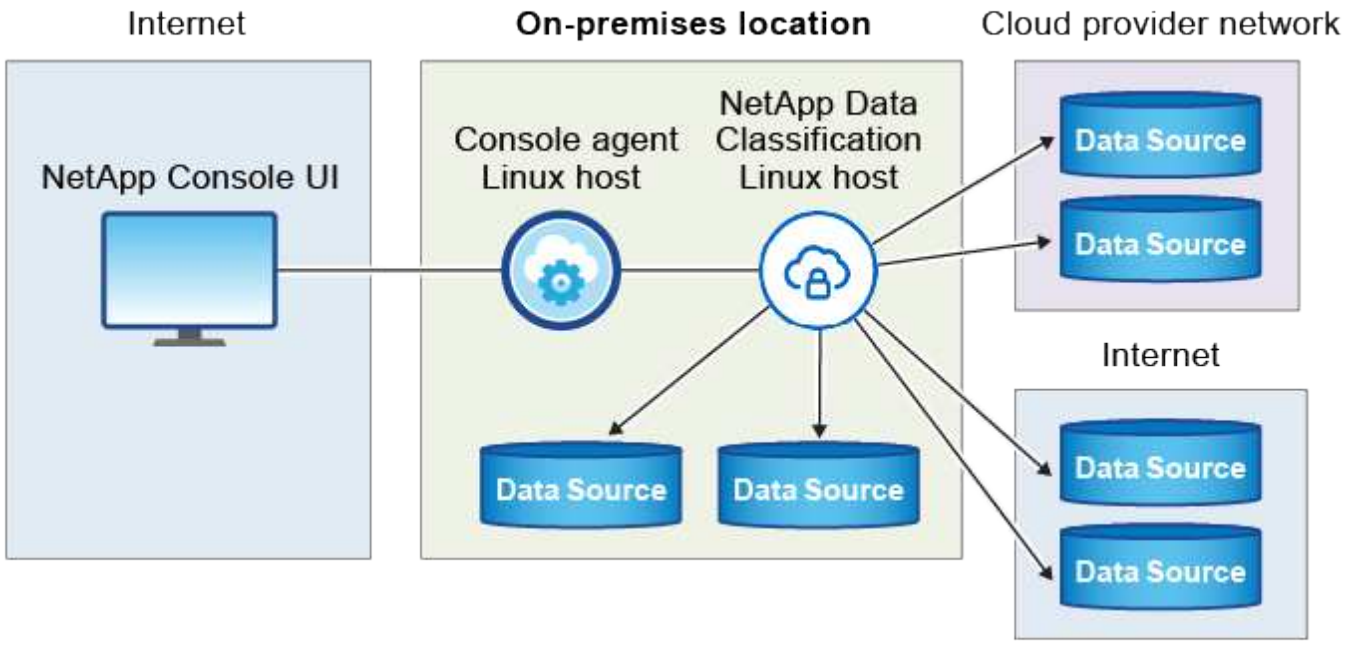
インターネットにアクセスできるホストに**NetApp Data Classification**をインストールする

ネットワーク内の Linux ホストまたはインターネットにアクセスできるクラウド内の Linux ホストにNetApp Data Classification を展開するには、ネットワークまたはクラウドに Linux ホストを手動で展開する必要があります。

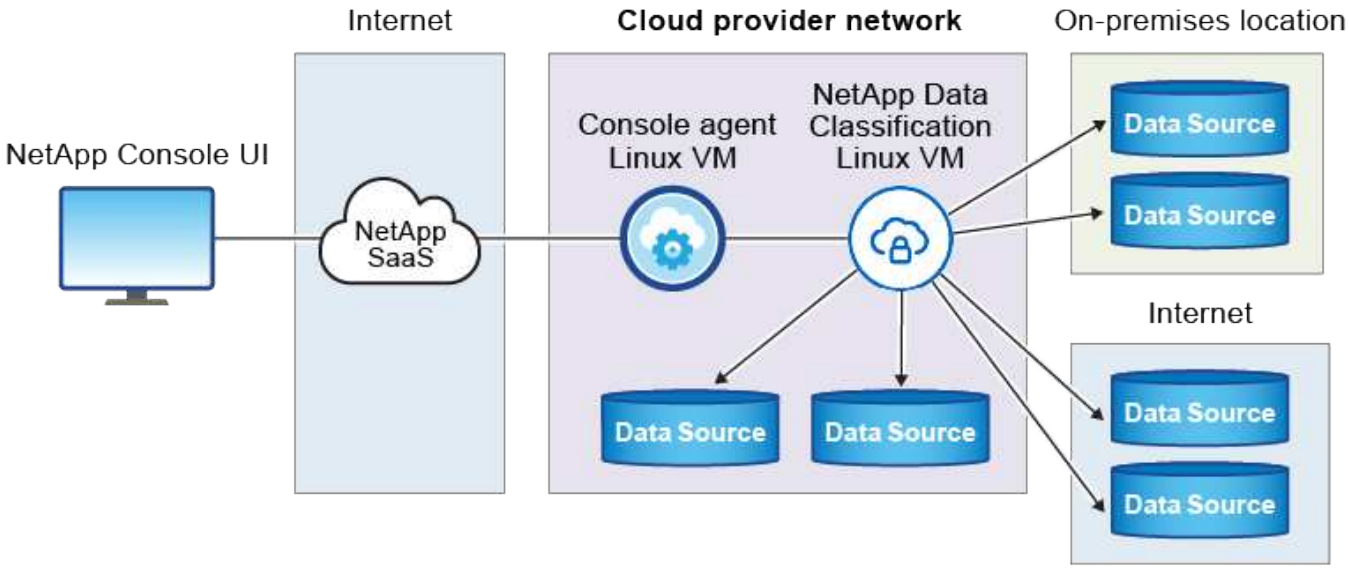
オンプレミス インストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションです。これは必須ではありません。どのインストール方法を選択しても、ソフトウェアは同じように機能します。

データ分類のインストール スクリプトは、システムと環境が必要な前提条件を満たしているかどうかを確認することから始まります。前提条件がすべて満たされている場合は、インストールが開始されます。データ分類のインストールの実行とは別に前提条件を検証したい場合は、前提条件のみをテストする別のソフトウェアパッケージをダウンロードできます。"Linuxホストがデータ分類をインストールする準備ができているかどうかを確認する方法をご覧ください" 。

社内の Linux ホストへの一般的なインストールには、次のコンポーネントと接続が含まれます。



クラウド内の Linux ホストへの一般的なインストールには、次のコンポーネントと接続が含まれます。



クイック スタート

以下の手順に従ってすぐに開始するか、残りのセクションまでスクロールして詳細を確認してください。

1 コンソールエージェントを作成する

コンソールエージェントがまだない場合は、 "コンソールエージェントをオンプレミスに展開する"ネットワーク内の Linux ホスト上、またはクラウド内の Linux ホスト上。

クラウド プロバイダーを使用してコンソール エージェントを作成することもできます。見る "AWSでコンソールエージェントを作成する"、 "Azureでコンソールエージェントを作成する" 、 または "GCP でコンソールエージェントを作成する"。

2 前提条件を確認する

ご使用の環境が前提条件を満たしていることを確認してください。これには、インスタンスのアウトバウンド インターネット アクセス、ポート 443 経由のコンソール エージェントとデータ分類間の接続などが含まれます。完全なリストを見る 。

また、以下の要件を満たすLinuxシステムも必要です。以下の要件 。

3 データ分類をダウンロードして展開する

NetAppサポート サイトから Cloud Data Classification ソフトウェアをダウンロードし、使用する予定の Linux ホストにインストーラ ファイルをコピーします。次に、インストール ウィザードを起動し、プロンプトに従ってデータ分類インスタンスをデプロイします。

コンソールエージェントを作成する

Data Classification をインストールして使用する前に、コンソール エージェントが必要です。ほとんどの場合、データ分類を有効化する前にコンソールエージェントをセットアップしておく必要があります。 "コンソール機能にはコンソールエージェントが必要です"ただし、今すぐ設定する必要がある場合もあります。

クラウドプロバイダー環境で作成するには、 "AWSでコンソールエージェントを作成する" 、 "Azureでコンソールエージェントを作成する" 、 または "GCP でコンソール エージェントを作成する"。

特定のクラウド プロバイダーにデプロイされたコンソール エージェントを使用する必要があるシナリオがいくつかあります。

- AWS のCloud Volumes ONTAPまたはAmazon FSx for ONTAPでデータをスキャンする場合は、AWS のコンソールエージェントを使用します。
- Azure のCloud Volumes ONTAPまたはAzure NetApp Filesでデータをスキャンする場合は、Azure のコンソール エージェントを使用します。

  Azure NetApp Filesの場合、スキャンするボリュームと同じリージョンにデプロイする必要があります。

- GCP のCloud Volumes ONTAPでデータをスキャンする場合は、GCP のコンソール エージェントを使用します。

オンプレミスのONTAPシステム、 NetAppファイル共有、データベース アカウントは、これらのクラウド コ

ンソール エージェントのいずれかを使用してスキャンできます。

また、 "コンソールエージェントをオンプレミスに展開する"ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。オンプレミスで Data Classification をインストールする予定のユーザーの中には、オンプレミスでコンソール エージェントをインストールすることを選択する場合もあります。

Data Classification をインストールするときは、コンソール エージェント システムの IP アドレスまたはホスト名が必要になります。オンプレミスでコンソール エージェントをインストールした場合は、この情報が得られます。コンソール エージェントがクラウドに展開されている場合は、コンソールからこの情報を確認できます。ヘルプ アイコンを選択し、次に サポート、コンソール エージェント を選択します。

Linuxホストシステムを準備する

データ分類ソフトウェアは、特定のオペレーティング システム要件、RAM 要件、ソフトウェア要件などを満たすホスト上で実行する必要があります。 Linux ホストはネットワーク内またはクラウド内に配置できます。

データ分類を実行し続けることができることを確認します。データ分類マシンは、データを継続的にスキャンするためにオンのままにしておく必要があります。

- データ分類は専用ホスト上で実行する必要があります。ホストは、他のアプリケーションやウイルス対策などのサードパーティ製ソフトウェアと共有することはできません。
- データ分類でスキャンする予定のデータ セットに合わせてサイズを選択します。

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
特大	CPU×32	128GBのRAM	• / に 1 TiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker で 895 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**
大きい	CPU×16	64GBのRAM	• / に 500 GiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker または Podman /var/lib/containers で 400 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

- データ分類インストール用にクラウドにコンピューティング インスタンスをデプロイする場合は、上記の「大規模」システム要件を満たすシステムを使用することをお勧めします。
  - **Amazon Elastic Compute Cloud (Amazon EC2)** インスタンスタイプ:「m6i.4xlarge」。"その他のAWSインスタンスタイプを見る" 。

- **Azure VM** サイズ:「Standard_D16s_v3」。"その他のAzureインスタンスタイプを見る" 。
- **GCP** マシンタイプ:「n2-standard-16」。"その他の GCP インスタンスタイプを見る" 。

- **UNIX** フォルダ権限: 次の最低限の UNIX 権限が必要です。

フォルダ	最小限の権限
/tmp	`rwxrwxrwt`
/opt	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/システム	`rwxr-xr-x`

- オペレーティング･システム：
  - 次のオペレーティング システムでは、Docker コンテナ エンジンを使用する必要があります。
    - Red Hat Enterprise Linux バージョン 7.8 および 7.9
    - Ubuntu 22.04 (データ分類バージョン 1.23 以上が必要)
    - Ubuntu 24.04 (データ分類バージョン 1.23 以上が必要)
  - 次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、データ分類バージョン 1.30 以上が必要です。
    - Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5、および 9.6。
  - ホスト システムで Advanced Vector Extensions (AVX2) を有効にする必要があります。
- **Red Hat** サブスクリプション管理: ホストは Red Hat サブスクリプション管理に登録されている必要があります。登録されていない場合、システムはリポジトリにアクセスできず、インストール中に必要なサードパーティ製ソフトウェアを更新できません。
- 追加ソフトウェア: Data Classification をインストールする前に、ホストに次のソフトウェアをインストールする必要があります。
  - 使用している OS に応じて、次のいずれかのコンテナ エンジンをインストールする必要があります。
    - Docker Engine バージョン 19.3.1 以上。 "インストール手順を見る" 。
    - Podman バージョン 4 以上。 Podmanをインストールするには、次のように入力します。(`sudo yum install podman netavark -y`）。
- Python バージョン 3.6 以上。 "インストール手順を見る" 。
  - **NTP** に関する考慮事項: NetApp、データ分類システムをネットワーク タイム プロトコル (NTP) サービスを使用するように構成することを推奨しています。データ分類システムとコンソール エージェント システムの間で時刻を同期する必要があります。
- **Firewalld**の考慮事項: 使用を計画している場合 `firewalld`、データ分類をインストールする前に有効にすることをお勧めします。設定するには次のコマンドを実行します `firewalld`データ分類と互換性があるように:

```
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
```

追加のデータ分類ホストをスキャナー ノードとして使用することを計画している場合は、この時点で次のルールをプライマリ システムに追加します。

```
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
```

有効化または更新するたびにDockerまたはPodmanを再起動する必要があることに注意してください。 `firewalld`設定。

データ分類ホスト システムの IP アドレスは、インストール後に変更することはできません。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソールとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.bluexp.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。
https://github.com/docker https://download.docker.com	docker インストールの前提条件パッケージを提供します。
http://packages.ubuntu.com/ http://archive.ubuntu.com	Ubuntu インストールの前提条件となるパッケージを提供します。

必要なポートがすべて有効になっていることを確認します

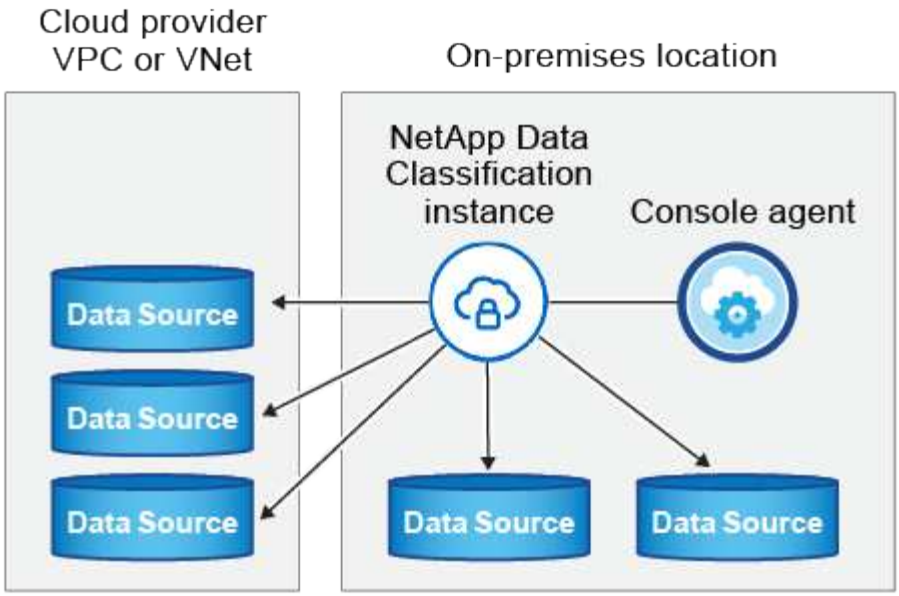
コンソール エージェント、データ分類、Active Directory、およびデータ ソース間の通信に必要なすべてのポートが開いていることを確認する必要があります。

接続タイプ	ポート	説明
コンソールエージェント <> データ分類	8080 (TCP)、443 (TCP)、および 80。9000	コンソール エージェントのファイアウォールまたはルーティング ルールでは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。コンソールでインストールの進行状況を確認できるように、ポート 8080 が開いていることを確認してください。 Linux ホストでファイアウォールが使用されている場合、Ubuntu サーバー内の内部プロセスにはポート 9000 が必要です。
コンソールエージェント <> ONTAPクラスタ (NAS)	443（TCP）	コンソールは、HTTPS を使用してONTAPクラスターを検出します。カスタム ファイアウォール ポリシーを使用する場合は、次の要件を満たす必要があります。 • コンソール エージェント ホストは、ポート 443 経由の送信 HTTPS アクセスを許可する必要があります。コンソール エージェントがクラウド内にある場合、すべての送信通信は事前定義されたファイアウォールまたはルーティング ルールによって許可されます。 • ONTAPクラスタは、ポート 443 経由の着信 HTTPS アクセスを許可する必要があります。デフォルトの「mgmt」ファイアウォール ポリシーでは、すべての IP アドレスからの受信 HTTPS アクセスが許可されます。このデフォルト ポリシーを変更した場合、または独自のファイアウォール ポリシーを作成した場合は、HTTPS プロトコルをそのポリシーに関連付け、コンソール エージェント ホストからのアクセスを有効にする必要があります。

接続タイプ	ポート	説明
データ分類 <> ONTAPクラスタ	• NFSの場合 - 111 (TCP\UDP) および 2049 (TCP\UDP) • CIFSの場合 - 139（TCP\UDP）および445（TCP\UDP）	データ分類には、各Cloud Volumes ONTAPサブネットまたはオンプレミスのONTAPシステムへのネットワーク接続が必要です。 Cloud Volumes ONTAPのファイアウォールまたはルーティング ルールは、データ分類インスタンスからの受信接続を許可する必要があります。 次のポートがデータ分類インスタンスに対して開いていることを確認します。 • NFSの場合 - 111と2049 • CIFSの場合 - 139および445 NFS ボリュームのエクスポート ポリシーでは、データ分類インスタンスからのアクセスを許可する必要があります。
データ分類 <> Active Directory	389 (TCP & UDP)、636 (TCP)、3268 (TCP)、および 3269 (TCP)	社内のユーザー用に Active Directory がすでに設定されている必要があります。さらに、データ分類では、CIFS ボリュームをスキャンするために Active Directory 資格情報が必要です。 Active Directory の情報が必要です: • DNSサーバーのIPアドレス、または複数のIPアドレス • サーバーのユーザー名とパスワード • ドメイン名（アクティブディレクトリ名） • セキュアLDAP（LDAPS）を使用しているかどうか • LDAP サーバー ポート (通常、LDAP の場合は 389、セキュア LDAP の場合は 636)

Linuxホストにデータ分類をインストールする

通常の構成では、ソフトウェアを単一のホスト システムにインストールします。ここでその手順をご覧ください 。



見るLinuxホストシステムの準備そして前提条件の確認データ分類を展開する前に、要件の完全なリストを確認してください。

インスタンスがインターネットに接続されている限り、データ分類ソフトウェアへのアップグレードは自動化されます。

> 現在、データ分類では、ソフトウェアがオンプレミスにインストールされている場合、S3 バケット、 Azure NetApp Files、または FSx for ONTAPをスキャンできません。このような場合には、クラウドに別のコンソールエージェントとデータ分類のインスタンスを展開し、 "コネクタ間の切り替え"さまざまなデータ ソース用。

一般的な構成の単一ホストインストール

単一のオンプレミス ホストにデータ分類ソフトウェアをインストールする場合は、要件を確認し、次の手順に従ってください。

"このビデオを見る"Data Classification のインストール方法を確認します。

Data Classification をインストールすると、すべてのインストール アクティビティがログに記録されることに注意してください。インストール中に問題が発生した場合は、インストール監査ログの内容を表示できます。それは、 `/opt/netapp/install_logs/` 。

開始する前に

- Linuxシステムがホスト要件。
- システムに 2 つの前提条件ソフトウェア パッケージ (Docker Engine または Podman、および Python 3) がインストールされていることを確認します。
- Linux システムでルート権限を持っていることを確認してください。
- インターネットへのアクセスにプロキシを使用している場合:
  - プロキシ サーバー情報 (IP アドレスまたはホスト名、接続ポート、接続スキーム: https または http、ユーザー名とパスワード) が必要になります。
  - プロキシが TLS インターセプションを実行している場合は、TLS CA 証明書が保存されている Data

Classification Linux システム上のパスを知っておく必要があります。

- プロキシは非透過である必要があります。データ分類は現在、透過プロキシをサポートしていません。
- ユーザーはローカル ユーザーである必要があります。ドメイン ユーザーはサポートされていません。

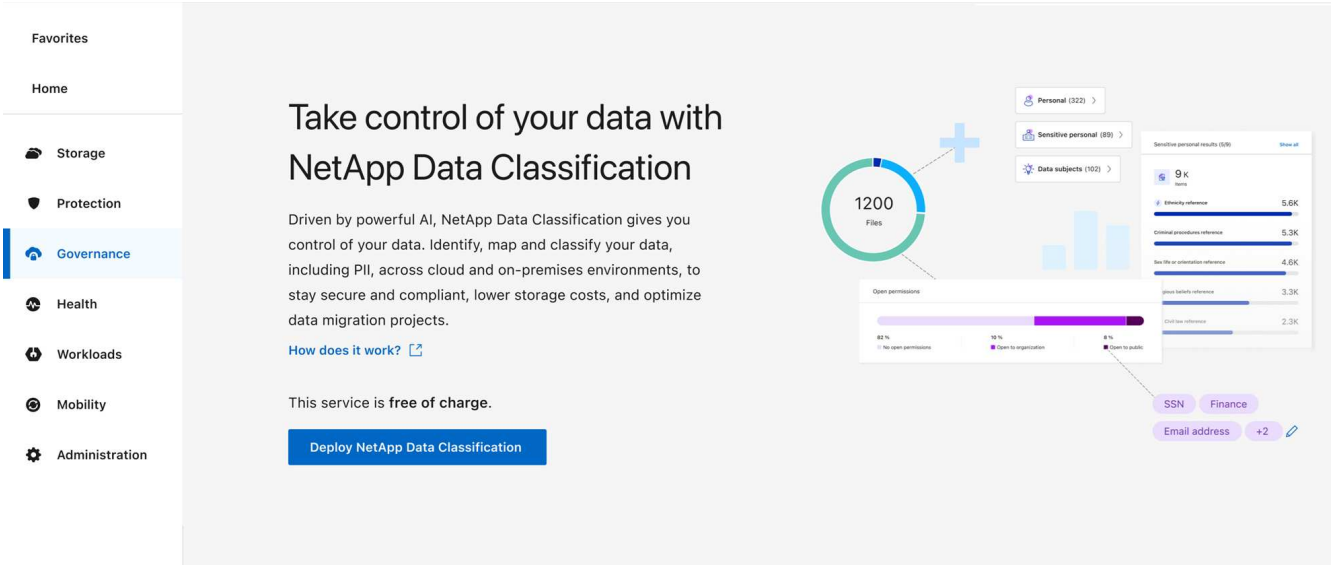
- オフライン環境が要件を満たしていることを確認する権限と接続性。

手順

1. データ分類ソフトウェアを以下からダウンロードしてください。 "NetAppサポート サイト" 。選択するファイルの名前は **DATASENSE-INSTALLER-<version>.tar.gz** です。
2. 使用する予定のLinuxホストにインストーラファイルをコピーします（ `scp`または他の方法）。
3. ホスト マシン上でインストーラ ファイルを解凍します。例:

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

4. コンソールで、*ガバナンス > 分類*を選択します。
5. *オンプレミスまたはクラウドでの分類の展開*を選択します。



6. クラウドで準備したインスタンスにデータ分類をインストールするか、オンプレミスで準備したインスタンスにデータ分類をインストールするかに応じて、適切な デプロイ オプションを選択してデータ分類のインストールを開始します。
7. オンプレミスでのデータ分類の展開 ダイアログが表示されます。提供されたコマンドをコピーします（例： `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`) を作成し、テキスト ファイルに貼り付けて、後で使用することもできます。次に、[閉じる] を選択してダイアログを閉じます。
8. ホスト マシンで、コピーしたコマンドを入力して一連のプロンプトに従うか、必要なすべてのパラメーターを含む完全なコマンドをコマンド ライン引数として指定することもできます。

   インストーラーは、インストールを正常に実行するためにシステムとネットワークの要件が満たされているかどうかを確認するための事前チェックを実行することに注意してください。 "このビデオを見る"事前チェックのメッセージとその意味を理解する。

プロンプトに従ってパラメータを入力します。	完全なコマンドを入力します。
a. 手順 7 からコピーしたコマンドを貼り付けます。 `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` クラウドインスタンス（オンプレミスではない）にインストールする場合は、 `--manual-cloud-install <cloud_provider>` 。 b. コンソール エージェント システムからアクセスできるように、データ分類ホスト マシンの IP アドレスまたはホスト名を入力します。 c. データ分類システムからアクセスできるように、コンソール エージェント ホスト マシンの IP アドレスまたはホスト名を入力します。 d. プロンプトに従ってプロキシの詳細を入力します。コンソール エージェントがすでにプロキシを使用している場合は、データ分類はコンソール エージェントが使用するプロキシを自動的に使用するため、ここでこの情報を再度入力する必要はありません。	あるいは、必要なホストとプロキシのパラメータを指定して、コマンド全体を事前に作成することもできます。 `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>`

変数値:

- *account_id* = NetAppアカウント ID
- *client_id* = コンソールエージェントのクライアントID（クライアントIDに「clients」というサフィックスがない場合は追加します）
- *user_token* = JWTユーザーアクセストークン
- *ds_host* = データ分類 Linux システムの IP アドレスまたはホスト名。
- *cm_host* = コンソール エージェント システムの IP アドレスまたはホスト名。
- *cloud_provider* = クラウドインスタンスにインストールする場合は、クラウドプロバイダーに応じて「AWS」、「Azure」、または「Gcp」を入力します。
- *proxy_host* = ホストがプロキシ サーバーの背後にある場合のプロキシ サーバーの IP またはホスト名。
- *proxy_port* = プロキシ サーバーに接続するためのポート (デフォルトは 80)。
- *proxy_scheme* = 接続スキーム: https または http (デフォルトは http)。
- *proxy_user* = 基本認証が必要な場合に、プロキシ サーバーに接続するための認証済みユーザー。ユーザーはローカル ユーザーである必要があります。ドメイン ユーザーはサポートされていません。
- *proxy_password* = 指定したユーザー名のパスワード。
- *ca_cert_dir* = 追加の TLS CA 証明書バンドルを含むデータ分類 Linux システム上のパス。プロキシが TLS インターセプションを実行している場合にのみ必要です。

結果

Data Classification インストーラーは、パッケージをインストールし、インストールを登録し、Data Classification をインストールします。インストールには10～20分かかります。

ホスト マシンとコンソール エージェント インスタンスの間にポート 8080 経由の接続がある場合は、コンソールの [データ分類] タブにインストールの進行状況が表示されます。

次は何？

構成ページから、スキャンするデータ ソースを選択できます。

インターネットにアクセスできない **Linux** ホストに**NetApp Data Classification**をインストールする

インターネットにアクセスできないオンプレミス サイトの Linux ホストにNetApp Data Classification をインストールすることを、プライベート モード と呼びます。インストール スクリプトを使用するこのタイプのインストールでは、 NetApp ConsoleSaaS レイヤーに接続できません。

> BlueXPプライベート モード (レガシーBlueXPインターフェイス) は通常、インターネット接続がなく、AWS Secret Cloud、AWS Top Secret Cloud、Azure IL6 などの安全なクラウド領域があるオンプレミス環境で使用されます。NetApp は、従来のBlueXPインターフェースを使用してこれらの環境を引き続きサポートします。従来のBlueXPインターフェースのプライベートモードのドキュメントについては、"BlueXPプライベートモードの PDF ドキュメント" 。

Linuxホストが**NetApp Data Classification**をインストールする準備ができていることを確認します

Linux ホストにNetApp Data Classification を手動でインストールする前に、オプションでホスト上でスクリプトを実行し、Data Classification をインストールするための前提条件がすべて満たされていることを確認します。このスクリプトは、ネットワーク内の Linux ホストまたはクラウド内の Linux ホストで実行できます。ホストはインターネットに接続することも、インターネットにアクセスできないサイト (ダーク サイト) に存在することもできます。

データ分類インストール スクリプトには、環境が要件を満たしていることを確認するためのテスト スクリプトが含まれています。インストール スクリプトを実行する前に、このスクリプトを個別に実行して、Linux ホストの準備状況を確認できます。

はじめに

以下のタスクを実行します。

- 必要に応じて、コンソール エージェントがまだインストールされていない場合はインストールします。コンソール エージェントがインストールされていなくてもテスト スクリプトを実行できますが、スクリプトはコンソール エージェントとデータ分類ホスト マシン間の接続をチェックするため、コンソール エージェントを使用することをお勧めします。
- ホスト マシンを準備し、すべての要件を満たしていることを確認します。
- データ分類ホスト マシンからのアウトバウンド インターネット アクセスを有効にします。

- すべてのシステムで必要なすべてのポートが有効になっていることを確認します。
- 前提条件テスト スクリプトをダウンロードして実行します。

コンソールエージェントを作成する

Data Classification をインストールして使用する前に、コンソール エージェントが必要です。ただし、コンソール エージェントなしで前提条件スクリプトを実行することは可能です。

あなたはできる "オンプレミスにコンソールエージェントをインストールする" ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。コンソール エージェントがオンプレミスにインストールされている場合は、データ分類をオンプレミスにインストールすることもできます。

クラウド プロバイダー環境でコンソール エージェントを作成するには、以下を参照してください。

- "AWSでコンソールエージェントを作成する"
- "Azureでコンソールエージェントを作成する"
- "GCP でコンソール エージェントを作成する"

前提条件スクリプトを実行するときは、コンソール エージェント システムの IP アドレスまたはホスト名が必要です。オンプレミスでコンソール エージェントをインストールした場合は、この情報が得られます。コンソール エージェントがクラウドに展開されている場合は、コンソールからこの情報を確認できます。ヘルプ アイコンを選択してから [サポート] を選択し、[エージェントと監査] セクションで [エージェントに移動] を選択します。

ホストの要件を確認する

データ分類ソフトウェアは、特定のオペレーティング システム要件、RAM 要件、およびソフトウェア要件を満たすホスト上で実行する必要があります。

- データ分類は専用ホスト上で実行する必要があります。ホストは、他のアプリケーションやウイルス対策などのサードパーティ製ソフトウェアと共有することはできません。
- データ分類でスキャンする予定のデータ セットに合わせてサイズを選択します。

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
特大	CPU×32	128GBのRAM	• / に 1 TiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker で 895 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
大きい	CPU×16	64GBのRAM	• / に 500 GiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker または Podman /var/lib/containers で 400 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

- データ分類インストール用にクラウドにコンピューティング インスタンスをデプロイする場合は、上記の「大規模」システム要件を満たすシステムを使用することをお勧めします。
  - **Amazon Elastic Compute Cloud (Amazon EC2)** インスタンスタイプ:「m6i.4xlarge」。"その他のAWSインスタンスタイプを見る" 。
  - **Azure VM** サイズ:「Standard_D16s_v3」。"その他のAzureインスタンスタイプを見る" 。
  - **GCP** マシンタイプ:「n2-standard-16」。"その他の GCP インスタンスタイプを見る" 。
- **UNIX** フォルダ権限: 次の最低限の UNIX 権限が必要です。

フォルダ	最小限の権限
/tmp	`rwxrwxrwt`
/opt	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/システム	`rwxr-xr-x`

- オペレーティング･システム：
  - 次のオペレーティング システムでは、Docker コンテナ エンジンを使用する必要があります。
    - Red Hat Enterprise Linux バージョン 7.8 および 7.9
    - Ubuntu 22.04 (データ分類バージョン 1.23 以上が必要)
    - Ubuntu 24.04 (データ分類バージョン 1.23 以上が必要)
  - 次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、データ分類バージョン 1.30 以上が必要です。
    - Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5、および 9.6。
  - ホスト システムで Advanced Vector Extensions (AVX2) を有効にする必要があります。
- **Red Hat** サブスクリプション管理: ホストは Red Hat サブスクリプション管理に登録されている必要があります。登録されていない場合、システムはリポジトリにアクセスできず、インストール中に必要なサードパーティ製ソフトウェアを更新できません。
- 追加ソフトウェア: Data Classification をインストールする前に、ホストに次のソフトウェアをインストールする必要があります。

- 使用している OS に応じて、次のいずれかのコンテナ エンジンをインストールする必要があります。
  - Docker Engine バージョン 19.3.1 以上。 "インストール手順を見る" 。
  - Podman バージョン 4 以上。 Podmanをインストールするには、次のように入力します。(`sudo yum install podman netavark -y`）。

- Python バージョン 3.6 以上。 "インストール手順を見る" 。
  - **NTP** に関する考慮事項: NetApp、データ分類システムをネットワーク タイム プロトコル (NTP) サービスを使用するように構成することを推奨しています。データ分類システムとコンソール エージェント システムの間で時刻を同期する必要があります。

- **Firewalld**の考慮事項: 使用を計画している場合 `firewalld`、データ分類をインストールする前に有効にすることをお勧めします。設定するには次のコマンドを実行します `firewalld`データ分類と互換性があるように:

```
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
```

追加のデータ分類ホストをスキャナー ノードとして使用することを計画している場合 (分散モデル)、この時点で次のルールをプライマリ システムに追加します。

```
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
```

有効化または更新するたびにDockerまたはPodmanを再起動する必要があることに注意してください。 `firewalld`設定。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。

このセクションは、インターネットに接続できないサイトにインストールされたホスト システムでは必要ありません。

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。

エンドポイント	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://dseasb33srnrn.cloudfront.net/ https://production.cloudflare.docker.com/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。
https://github.com/docker https://download.docker.com	docker インストールの前提条件パッケージを提供します。
http://packages.ubuntu.com/ http://archive.ubuntu.com	Ubuntu インストールの前提条件となるパッケージを提供します。

必要なポートがすべて有効になっていることを確認します

コンソール エージェント、データ分類、Active Directory、およびデータ ソース間の通信に必要なすべてのポートが開いていることを確認する必要があります。

接続タイプ	ポート	説明
コンソールエージェント <> データ分類	8080 (TCP)、443 (TCP)、および 80。9000	コンソール エージェントのファイアウォールまたはルーティング ルールでは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。コンソールでインストールの進行状況を確認できるように、ポート 8080 が開いていることを確認してください。 Linux ホストでファイアウォールが使用されている場合、Ubuntu サーバー内の内部プロセスにはポート 9000 が必要です。
コンソールエージェント <> ONTAPクラスタ (NAS)	443（TCP）	コンソールは、HTTPS を使用してONTAPクラスターを検出します。カスタム ファイアウォール ポリシーを使用する場合、コンソール エージェント ホストはポート 443 経由の送信 HTTPS アクセスを許可する必要があります。コンソール エージェントがクラウド内にある場合、すべての送信通信は事前定義されたファイアウォールまたはルーティング ルールによって許可されます。

データ分類の前提条件スクリプトを実行する

データ分類の前提条件スクリプトを実行するには、次の手順に従います。

"このビデオを見る"前提条件スクリプトを実行して結果を解釈する方法を確認します。

開始する前に

- Linuxシステムがホスト要件。
- システムに 2 つの前提条件ソフトウェア パッケージ (Docker Engine または Podman、および Python 3)

がインストールされていることを確認します。

- Linux システムでルート権限を持っていることを確認してください。

手順

1. データ分類の前提条件スクリプトを以下からダウンロードします。 "NetAppサポート サイト" 。選択するファイルの名前は **standalone-pre-requisite-tester-<version>** です。
2. 使用する予定のLinuxホストにファイルをコピーします（ `scp`または他の方法）。
3. スクリプトを実行するための権限を割り当てます。

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

4. 次のコマンドを使用してスクリプトを実行します。

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

インターネットにアクセスできないホストでスクリプトを実行する場合にのみ、オプション "--darksite" を追加します。ホストがインターネットに接続されていない場合、特定の前提条件テストはスキップされます。

5. スクリプトは、データ分類ホスト マシンの IP アドレスの入力を求めます。
   - IP アドレスまたはホスト名を入力します。
6. スクリプトは、コンソール エージェントがインストールされているかどうかを尋ねます。
   - コンソール エージェントがインストールされていない場合は、**N** と入力します。
   - コンソール エージェントがインストールされている場合は、「**Y**」と入力します。次に、テスト スクリプトがこの接続をテストできるように、コンソール エージェントの IP アドレスまたはホスト名を入力します。
7. スクリプトはシステム上でさまざまなテストを実行し、進行中に結果を表示します。終了すると、セッションのログを次のファイルに書き込みます。 `prerequisites-test-<timestamp>.log`ディレクトリ内 `/opt/netapp/install_logs`。

結果

すべての前提条件テストが正常に実行された場合、準備ができたらホストに Data Classification をインストールできます。

問題が発見された場合は、修正が「推奨」または「必須」として分類されます。推奨される問題は通常、データ分類のスキャンと分類のタスクの実行速度を低下させる項目です。これらの項目は修正する必要はありませんが、対処することをお勧めします。

「必須」の問題がある場合は、問題を修正して、前提条件テスト スクリプトを再度実行する必要があります。

データソースのスキャンを有効にする

NetApp Data Classificationでデータソースをスキャン

NetApp Data Classification は、選択したリポジトリ内のデータ (ボリューム、データベース スキーマ、またはその他のユーザー データ) をスキャンして、個人データや機密データを識別します。次に、データ分類により組織のデータがマッピングされ、各ファイルが分類され、データ内の定義済みパターンが識別されます。スキャンの結果は、個人情報、機密個人情報、データ カテゴリ、およびファイル タイプのインデックスです。

最初のスキャンの後、データ分類はラウンドロビン方式でデータを継続的にスキャンし、増分変更を検出します。そのため、インスタンスを実行し続けることが重要です。

ボリューム レベルまたはデータベース スキーマ レベルでスキャンを有効または無効にすることができます。

マッピングスキャンと分類スキャンの違いは何ですか？

データ分類では、次の 2 種類のスキャンを実行できます。

- マッピングのみのスキャン は、データの概要のみを提供し、選択されたデータ ソースに対して実行されます。マッピングのみのスキャンでは、ファイルにアクセスして内部のデータを確認する必要がないため、マップおよび分類スキャンよりも時間がかかりません。最初にこれを実行して研究領域を特定し、次にそれらの領域に対してマップと分類のスキャンを実行することをお勧めします。
- マップと分類スキャン は、データの詳細なスキャンを提供します。

以下の表にいくつかの違いを示します。

特徴	スキャンをマップして分類する	マッピングのみのスキャン
スキャン速度	遅い	速い
料金	空き	空き
容量	500 TiB に制限されます*	500 TiB に制限されます*
ファイルの種類と使用容量の一覧	はい	はい
ファイル数と使用容量	はい	はい
ファイルの古さとサイズ	はい	はい
実行する能力"データマッピングレポート"	はい	はい
ファイルの詳細を表示するためのデータ調査ページ	はい	いいえ
ファイル内の名前を検索する	はい	いいえ
作成する"保存されたクエリ"カスタム検索結果を提供する	はい	いいえ
他のレポートを実行する機能	はい	いいえ
ファイルのメタデータを表示する機能**	いいえ	はい

{アスタリスク} データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 +

コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

{アスタリスク}{アスタリスク} マッピング スキャン中にファイルから次のメタデータが抽出されます。

- システム
- システムタイプ
- ストレージリポジトリ
- ファイル タイプ
- 使用済み容量
- ファイル数
- ファイル サイズ
- ファイル作成
- ファイルの最終アクセス
- ファイルの最終更新日時
- ファイル発見時刻
- 権限の抽出

ガバナンス ダッシュボードの違い:

特徴	マップと分類	マップ
古いデータ	はい	はい
非ビジネスデータ	はい	はい
重複したファイル	はい	はい
定義済みの保存済みクエリ	はい	いいえ
デフォルトの保存クエリ	はい	はい
DDAレポート	はい	はい
マッピングレポート	はい	はい
感度レベル検出	はい	いいえ
幅広い権限を持つ機密データ	はい	いいえ
オープン権限	はい	はい
データの時代	はい	はい
データのサイズ	はい	はい
カテゴリ	はい	いいえ
ファイルの種類	はい	はい

コンプライアンス ダッシュボードの違い:

特徴	マップと分類	マップ
個人情報	はい	いいえ
機密個人情報	はい	いいえ
プライバシーリスク評価レポート	はい	いいえ
HIPAAレポート	はい	いいえ
PCI DSSレポート	はい	いいえ

調査フィルターの違い:

特徴	マップと分類	マップ
保存されたクエリ	はい	はい
システムタイプ	はい	はい
システム	はい	はい
ストレージリポジトリ	はい	はい
ファイル タイプ	はい	はい
ファイル サイズ	はい	はい
作成時間	はい	はい
発見された時間	はい	はい
最終更新日	はい	はい
最終アクセス	はい	はい
オープン権限	はい	はい
ファイルディレクトリパス	はい	はい
カテゴリ	はい	いいえ
感度レベル	はい	いいえ
識別子の数	はい	いいえ
個人データ	はい	いいえ
機密性の高い個人データ	はい	いいえ
データ主体	はい	いいえ
重複	はい	はい
分類ステータス	はい	ステータスは常に「限られた洞察」です
スキャン分析イベント	はい	はい
ファイルハッシュ	はい	はい
アクセス権を持つユーザーの数	はい	はい
ユーザー/グループの権限	はい	はい
ファイルの所有者	はい	はい
ディレクトリタイプ	はい	はい

NetApp Data Classificationを使用して**Amazon FSx for ONTAP**ボリュームをスキャンする

NetApp Data Classificationを使用してAmazon FSx for ONTAPボリュームをスキャンするには、いくつかの手順を実行します。

開始する前に

- データ分類を展開および管理するには、AWS にアクティブなコンソールエージェントが必要です。
- システムの作成時に選択したセキュリティ グループは、データ分類インスタンスからのトラフィックを許可する必要があります。 FSx for ONTAPファイルシステムに接続された ENI を使用して関連付けられたセキュリティグループを見つけ、AWS マネジメントコンソールを使用して編集できます。

  "Linuxインスタンス用のAWSセキュリティグループ"

  "Windowsインスタンス用のAWSセキュリティグループ"

  "AWS エラスティックネットワークインターフェース (ENI)"

- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Data Classification は、AWS のコンソールエージェントおよびスキャンする FSx ボリュームと同じ AWS ネットワークにデプロイする必要があります。

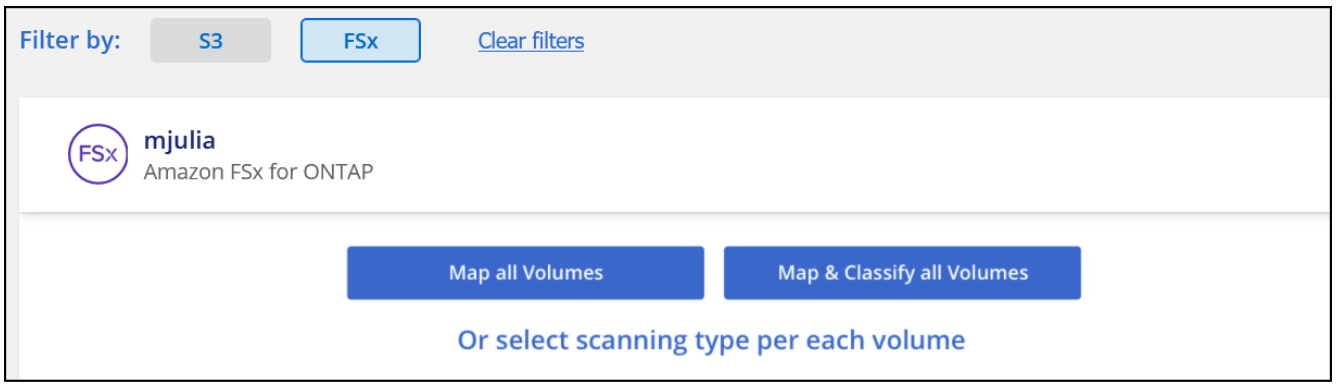
注意: FSx ボリュームをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

インスタンスがインターネットに接続されている限り、データ分類ソフトウェアへのアップグレードは自動化されます。

システムでデータ分類を有効にする

FSx for ONTAPボリュームのデータ分類を有効にすることができます。

1. NetApp Consoleから、*ガバナンス > 分類*を選択します。
2. データ分類メニューから、*構成*を選択します。



タブのスクリーンショット。"]

3. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ"

：

- すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
- すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
- 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップおよび/または分類するボリュームを選択します。

4. 確認ダイアログボックスで「承認」を選択すると、データ分類によってボリュームのスキャンが開始されます。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できるようになります。かかる時間はデータの量によって異なり、数分または数時間かかる場合があります。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。進行状況バーで各スキャンの進行状況を追跡します。進行状況バーの上にマウスを置くと、ボリューム内のファイルの合計数に対するスキャンされたファイルの数が表示されます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"このデータ分類の制限に関する詳細を見る" 。

データ分類がボリュームにアクセスできることを確認する

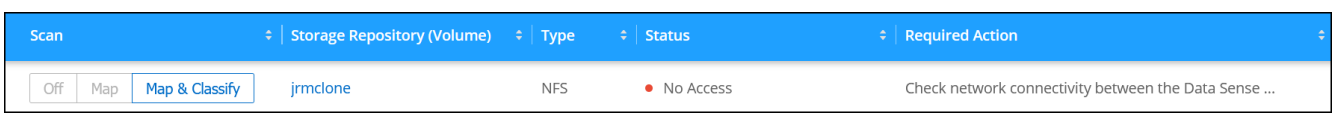
ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。

CIFS ボリュームにアクセスできるようにするには、データ分類に CIFS 資格情報を提供する必要があります。

手順

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで [詳細の表示] を選択してステータスを確認し、エラーを修正します。

   たとえば、次の画像は、データ分類インスタンスとボリューム間のネットワーク接続の問題により、データ分類がスキャンできないボリュームを示しています。



3. Data Classification インスタンスと、FSx for ONTAPのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。

FSx for ONTAPの場合、データ分類はコンソールと同じリージョン内のボリュームのみをスキャンできます。

4. NFS ボリュームのエクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていることを確認して、各ボリュームのデータにアクセスできるようにします。
5. CIFS を使用する場合は、Data Classification に Active Directory 資格情報を提供して、CIFS ボリュームをスキャンできるようにします。

   a. データ分類メニューから、*構成*を選択します。

   b. 各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

      データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

   資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。

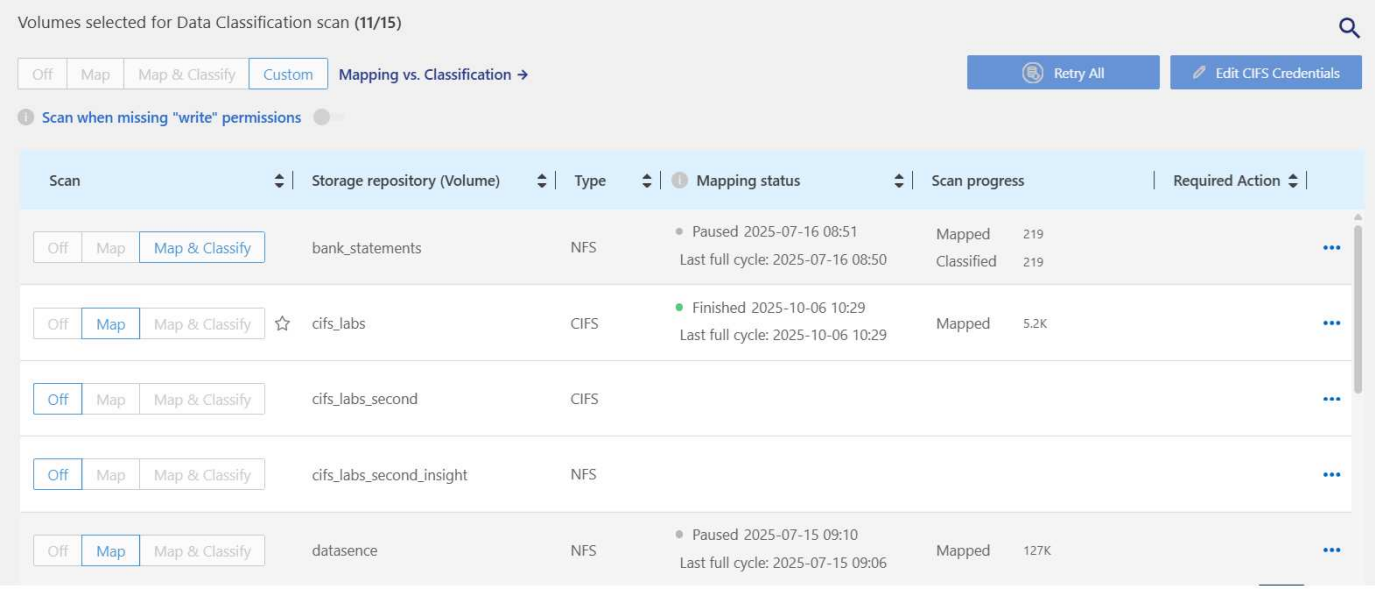
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

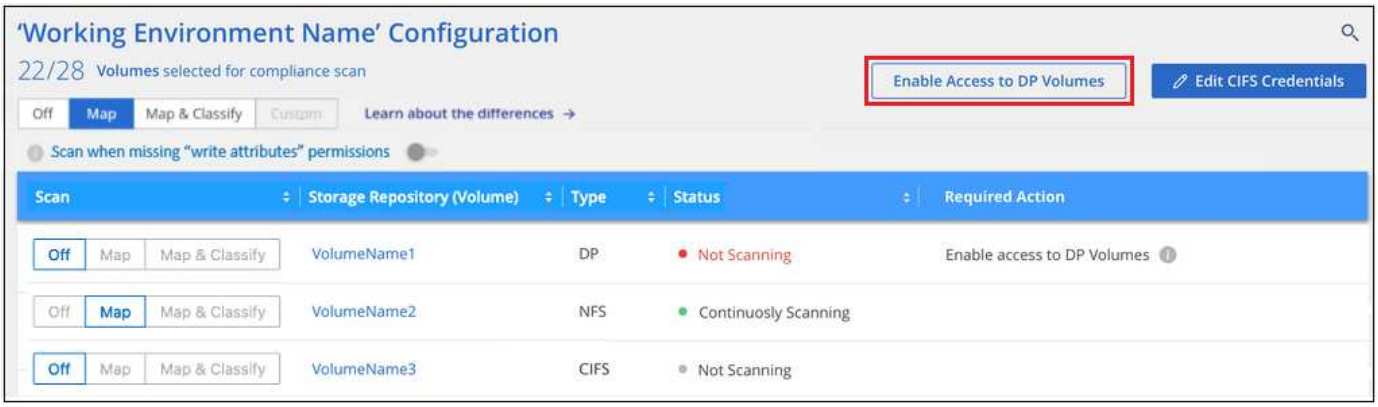
結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

データ保護ボリュームをスキャンする

デフォルトでは、データ保護 (DP) ボリュームは外部に公開されておらず、データ分類ではアクセスできないため、スキャンされません。これらは、FSx for ONTAPファイル システムからのSnapMirror操作の宛先ボリュームです。

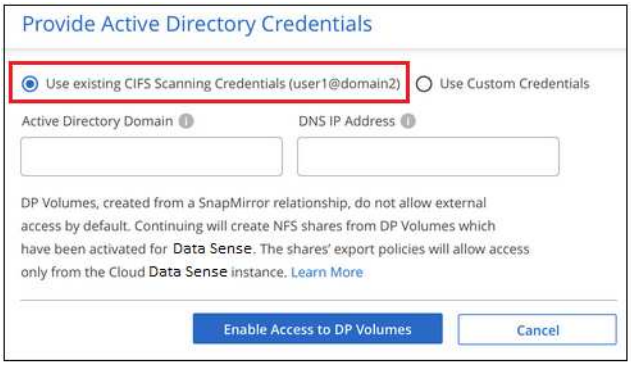
最初、ボリューム リストでは、これらのボリュームが、タイプ **DP**、ステータス スキャンなし、必要なアクション **DP** ボリュームへのアクセスを有効にする として識別されます。

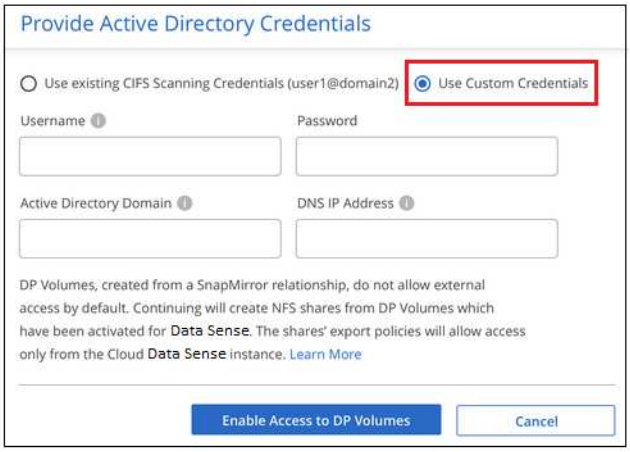


手順

これらのデータ保護ボリュームをスキャンする場合:

1. データ分類メニューから、*構成*を選択します。
2. ページの上部にある*DP ボリュームへのアクセスを有効にする*を選択します。
3. 確認メッセージを確認し、*DP ボリュームへのアクセスを有効にする*を再度選択します。
   - ソース FSx for ONTAPファイル システムで最初に NFS ボリュームとして作成されたボリュームが有効になります。
   - ソース FSx for ONTAPファイル システムで最初に CIFS ボリュームとして作成されたボリュームでは、それらの DP ボリュームをスキャンするために CIFS 認証情報を入力する必要があります。データ分類が CIFS ボリュームをスキャンできるように Active Directory 資格情報をすでに入力している場合は、その資格情報を使用することも、別の管理者資格情報セットを指定することもできます。





4. スキャンする各 DP ボリュームをアクティブ化します。

結果

有効にすると、データ分類はスキャン用にアクティブ化された各 DP ボリュームから NFS 共有を作成します。共有エクスポート ポリシーでは、データ分類インスタンスからのアクセスのみが許可されます。

最初に DP ボリュームへのアクセスを有効にしたときに CIFS データ保護ボリュームが存在せず、後でボリュームを追加すると、[構成] ページの上部に [**CIFS DP** へのアクセスを有効にする] ボタンが表示されます。このボタンを選択し、CIFS 資格情報を追加して、これらの CIFS DP ボリュームへのアクセスを有効にします。

Active Directory の資格情報は最初の CIFS DP ボリュームのストレージ VM にのみ登録されるため、その SVM 上のすべての DP ボリュームがスキャンされます。他の SVM に存在するボリュームには Active Directory 資格情報が登録されていないため、それらの DP ボリュームはスキャンされません。

NetApp Data Classificationを使用してAzure NetApp Filesボリュームをスキャンする

Azure NetApp FilesのNetApp Data Classificationを開始するには、いくつかの手順を完了します。

スキャンする**Azure NetApp Files**システムを検出します

スキャンするAzure NetApp FilesシステムがNetApp Consoleにシステムとしてまだ存在しない場合は、"システムページに追加します" 。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Azure NetApp Filesボリュームをスキャンするときは、データ分類をクラウドにデプロイする必要があり、スキャンするボリュームと同じリージョンにデプロイする必要があります。

注: Azure NetApp Filesボリュームをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

システムでデータ分類を有効にする

Azure NetApp Filesボリュームでデータ分類を有効にすることができます。

1. データ分類メニューから、*構成*を選択します。

タブのスクリーンショット。"]

2. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ" :
   - すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
   - すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
   - 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップするボリューム、またはマップして分類するボリュームを選択します。

見るボリュームのスキャンを有効または無効にする詳細については。

3. 確認ダイアログボックスで*承認*を選択します。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できます。かかる時間はデータの量によって異なり、数分または数時間かかる場合があります。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。データ分類では、スキャンごとに進行状況バーが表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数が表示されます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"データ分類の制限について学ぶ" 。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームにアクセスできるようにするには、Data Classification に CIFS 資格情報を提供する必要があります。

Azure NetApp Filesの場合、データ分類ではコンソールと同じリージョン内のボリュームのみをスキャンできます。

チェックリスト

- データ分類インスタンスと、 Azure NetApp Filesのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。
- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにします。

手順

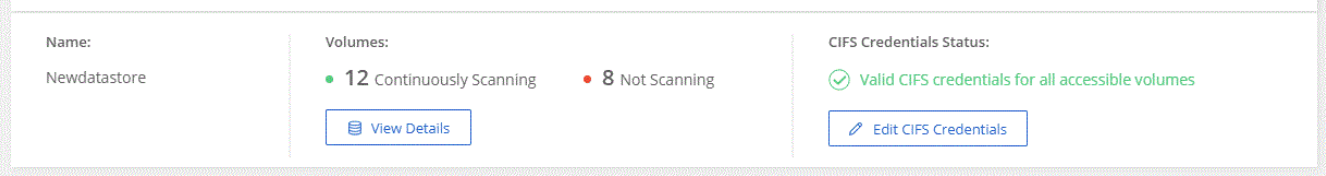
1. データ分類メニューから、*構成*を選択します。

   a. CIFS (SMB) を使用している場合は、Active Directory の資格情報が正しいことを確認してください。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができます。管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。



2. [構成] ページで [詳細の表示] を選択して、各 CIFS ボリュームと NFS ボリュームのステータスを確認します。必要に応じて、ネットワーク接続の問題などのエラーを修正します。

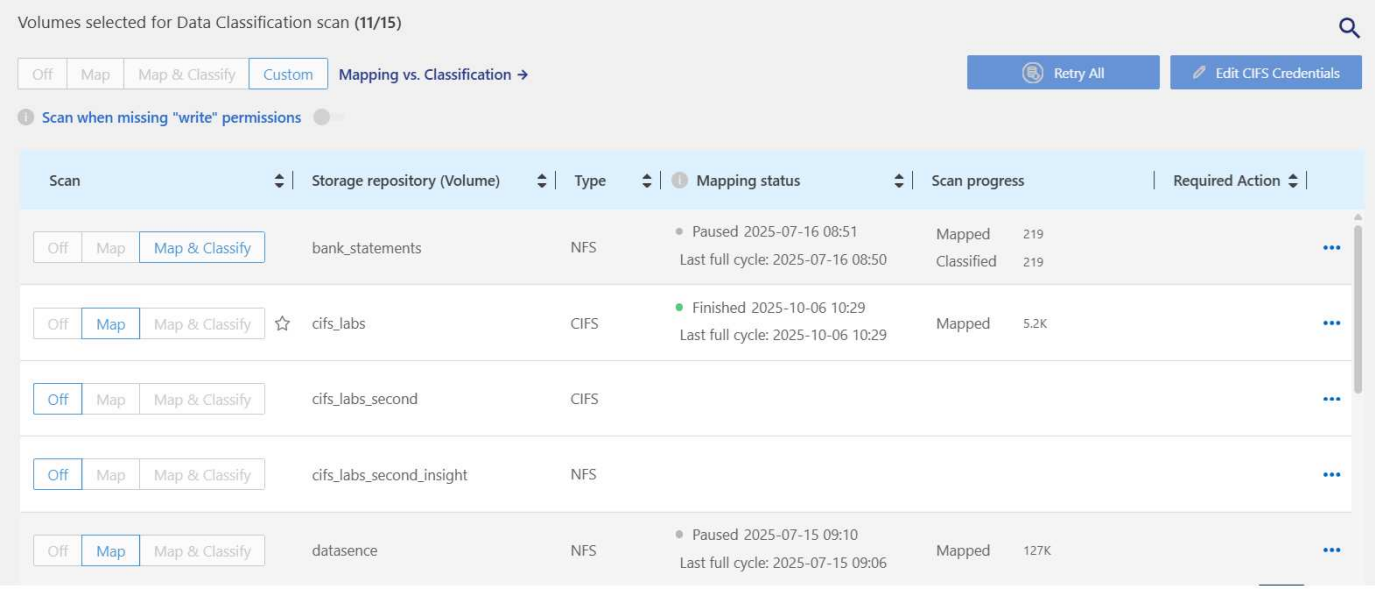
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

NetApp Data Classificationを使用してCloud Volumes ONTAPとオンプレミスのONTAPボリュームをスキャンします。

NetApp Data Classificationを使用してCloud Volumes ONTAPとオンプレミスのONTAPボリュームのスキャンを開始するには、いくつかの手順を実行します。

前提条件

データ分類を有効にする前に、サポートされている構成があることを確認してください。

- インターネット経由でアクセス可能なCloud Volumes ONTAPおよびオンプレミスのONTAPシステムをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所"。
- インターネットにアクセスできないダークサイトにインストールされているオンプレミスのONTAPシステムをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントを同じオンプレミスの場所に展開する必要があります。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームにアクセスできるようにするには、データ分類に CIFS 資格情報を提供する必要があります。

チェックリスト

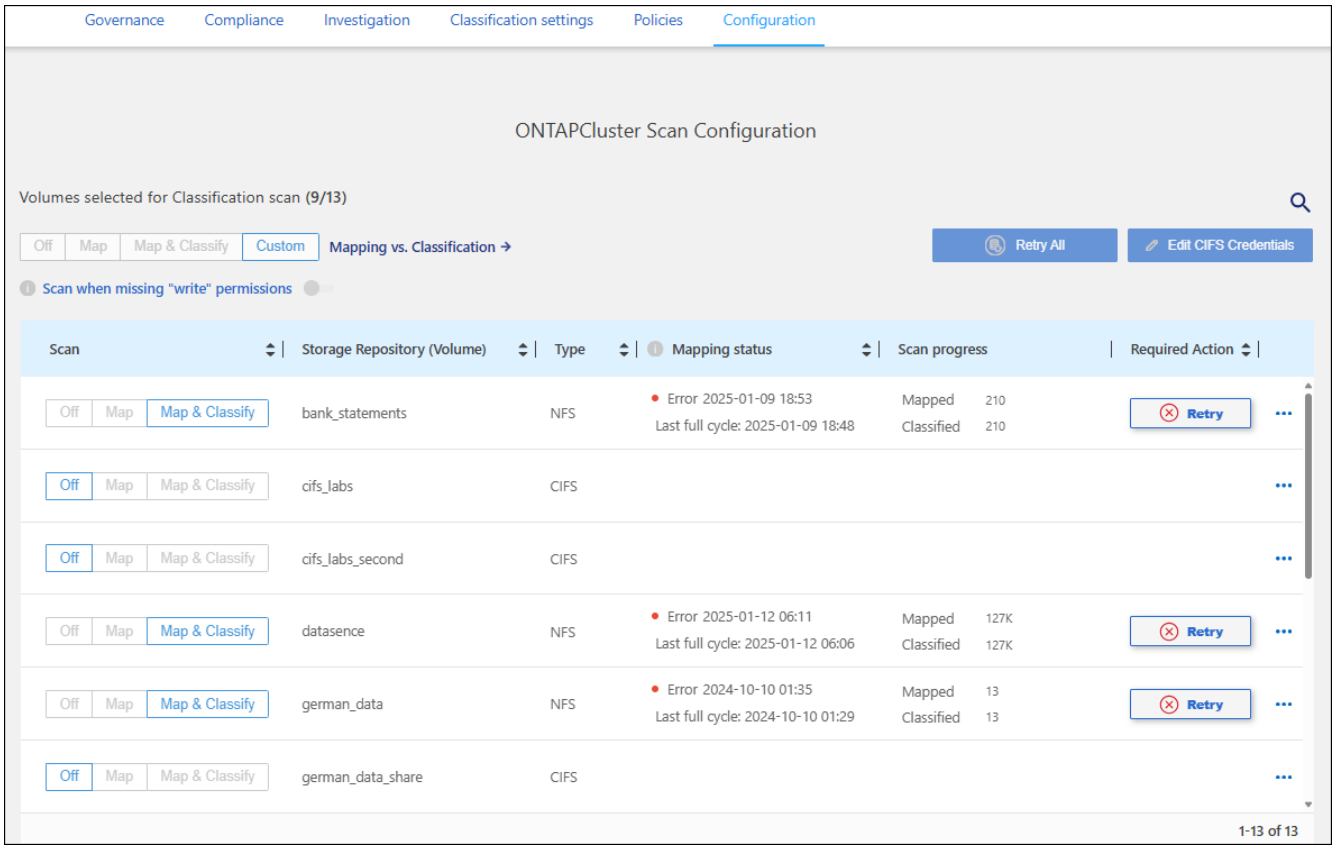
- Data Classification インスタンスと、 Cloud Volumes ONTAPまたはオンプレミスのONTAPクラスターのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- Cloud Volumes ONTAPのセキュリティ グループが、データ分類インスタンスからの受信トラフィックを許可していることを確認します。

  データ分類インスタンスの IP アドレスからのトラフィックに対してセキュリティ グループを開くことも、仮想ネットワーク内からのすべてのトラフィックに対してセキュリティ グループを開くこともできます。

- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにしていることを確認します。

手順

1. データ分類メニューから、*構成*を選択します。



タブのスクリーンショット。"]

2. CIFS を使用する場合は、Data Classification に Active Directory 資格情報を提供して、CIFS ボリュームをスキャンできるようにします。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

資格情報を正しく入力した場合、すべての CIFS ボリュームが正常に認証されたことを確認するメッセージが表示されます。

3. [構成] ページで [構成] を選択し、各 CIFS ボリュームと NFS ボリュームのステータスを確認し、エラーを修正します。

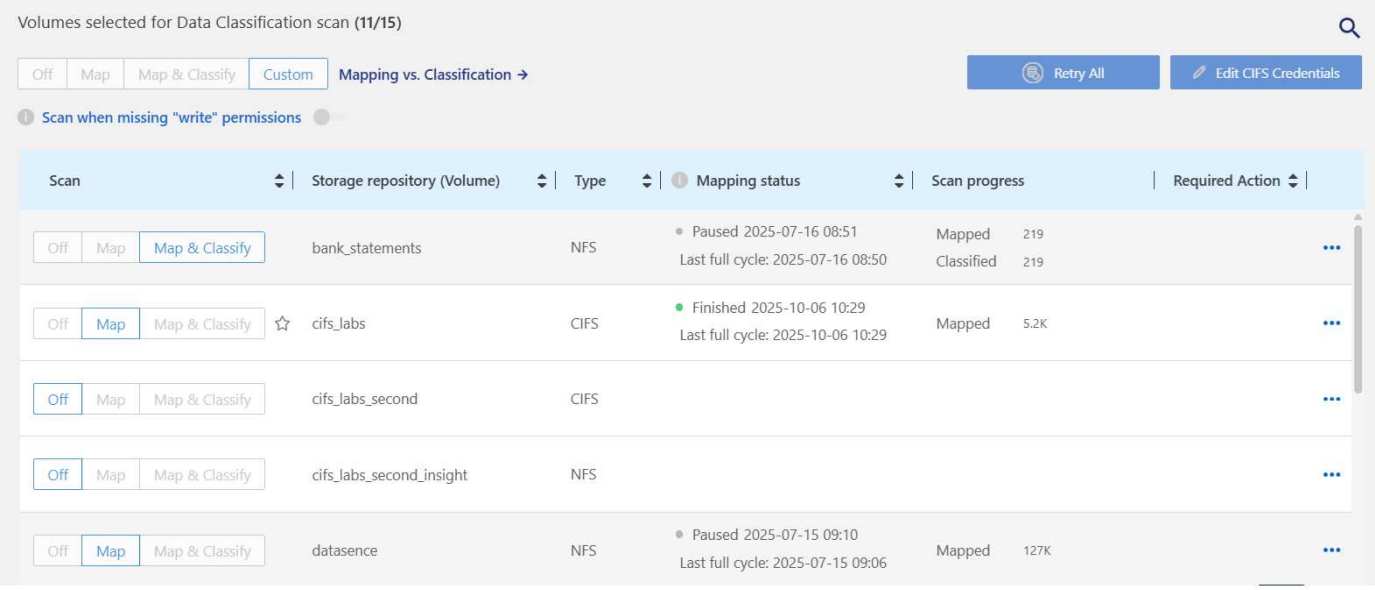
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"このデータ分類の制限に関する詳細を見る" 。

NetApp Data Classificationでデータベーススキーマをスキャンする

NetApp Data Classificationを使用してデータベース スキーマのスキャンを開始するには、いくつかの手順を実行します。

前提条件を確認する

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

サポートされているデータベース

データ分類では、次のデータベースからスキーマをスキャンできます。

- Amazon リレーショナルデータベースサービス (Amazon RDS)
- MongoDB
- MySQL
- Oracle
- PostgreSQL
- SAP HANA
- SQL サーバー (MSSQL)

データベースで統計収集機能が有効になっている必要があります。

データベース要件

ホストされている場所に関係なく、データ分類インスタンスに接続できるデータベースであればスキャンできます。データベースに接続するには、次の情報が必要です。

- IPアドレスまたはホスト名
- ポート
- サービス名（Oracle データベースへのアクセスのみ）
- スキーマへの読み取りアクセスを許可する資格情報

  ユーザー名とパスワードを選択するときは、スキャンするすべてのスキーマとテーブルに対する完全な読み取り権限を持つユーザー名とパスワードを選択することが重要です。必要なすべての権限を持つデータ分類システム専用のユーザーを作成することをお勧めします。

MongoDB の場合、読み取り専用の管理者ロールが必要です。

データ分類インスタンスをデプロイする

インスタンスがまだデプロイされていない場合は、データ分類をデプロイします。

インターネット経由でアクセス可能なデータベーススキーマをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所にデータ分類を展開する"。

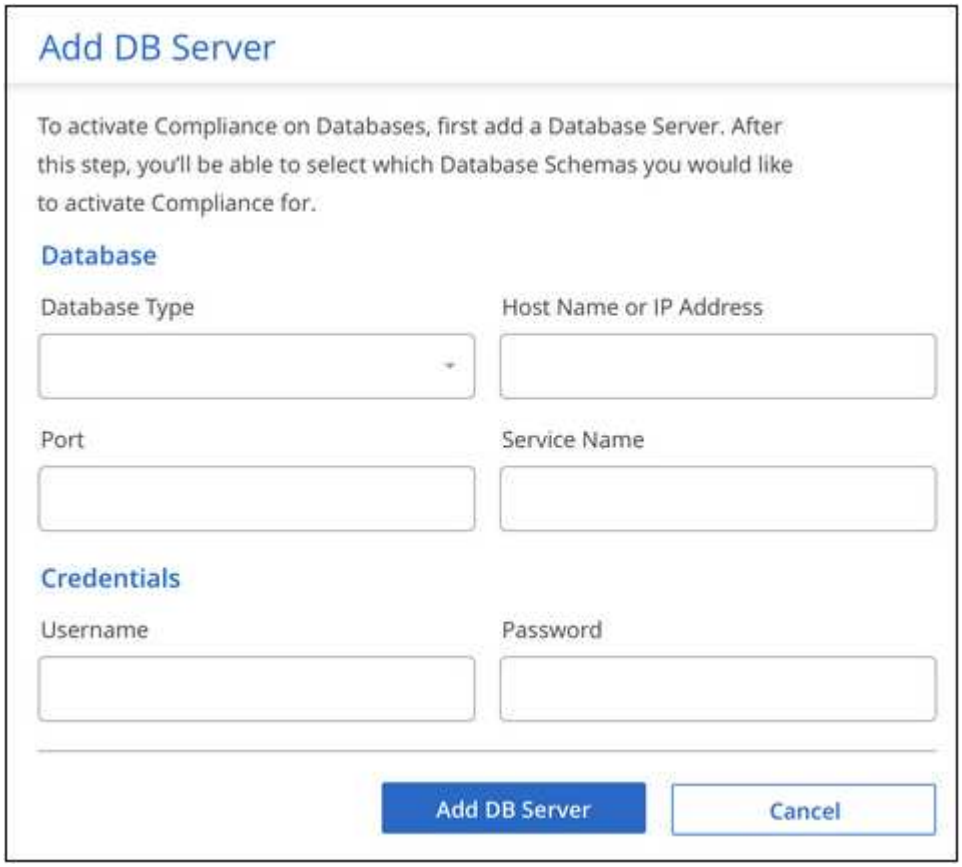
インターネットにアクセスできないダークサイトにインストールされているデータベーススキーマをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントが同じオンプレミスの場所に展開されていることも必要です。

データベースサーバーを追加する

スキーマが存在するデータベース サーバーを追加します。

1. データ分類メニューから、*構成*を選択します。
2. 構成ページで、システムの追加 > *データベース サーバーの追加*を選択します。
3. データベース サーバーを識別するために必要な情報を入力します。
   a. データベースの種類を選択します。

b. データベースに接続するためのポートとホスト名または IP アドレスを入力します。
c. Oracle データベースの場合は、サービス名を入力します。
d. データ分類がサーバーにアクセスできるように資格情報を入力します。
e. *DB サーバーの追加*を選択します。



データベースがシステムのリストに追加されます。

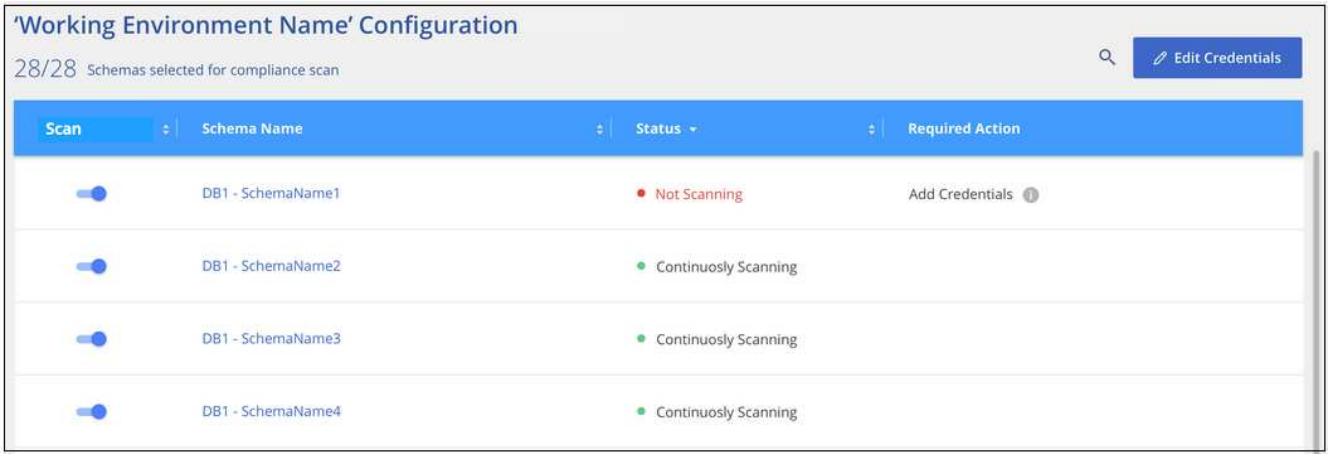
データベーススキーマのスキャンを有効または無効にする

スキーマの完全スキャンはいつでも停止または開始できます。

データベース スキーマのマッピングのみのスキャンを選択するオプションはありません。

1. 構成ページで、構成するデータベースの*構成*ボタンを選択します。

2. スライダーを右に移動して、スキャンするスキーマを選択します。



結果

データ分類は、有効にしたデータベース スキーマのスキャンを開始します。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。各スキャンの進行状況は進行状況バーとして表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。エラーがある場合は、エラーを修正するために必要なアクションとともにステータス列に表示されます。

データ分類では、データベースを 1 日に 1 回スキャンします。データベースは他のデータ ソースのように継続的にスキャンされるわけではありません。

NetApp Data Classificationを使用してGoogle Cloud NetApp Volumesをスキャンする

NetApp Data Classification は、システムとしてGoogle Cloud NetApp Volumes をサポートします。 Google Cloud NetApp Volumesシステムをスキャンする方法を学びます。

スキャンするGoogle Cloud NetApp Volumesシステムを検出します

スキャンしたいGoogle Cloud NetApp VolumesシステムがNetApp Consoleにシステムとしてまだ登録されていない場合は、"システムページに追加する" 。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Google Cloud NetApp Volumesをスキャンするときは、データ分類をクラウドにデプロイする必要があり、スキャンするボリュームと同じリージョンにデプロイする必要があります。

注: Google Cloud NetApp Volumesをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

システムでデータ分類を有効にする

Google Cloud NetApp Volumesシステムでデータ分類を有効にすることができます。

1. データ分類メニューから、*構成*を選択します。

2. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ"：
   - すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
   - すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
   - 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップおよび/または分類するボリュームを選択します。

     見るボリュームのスキャンを有効または無効にする詳細については。
3. 確認ダイアログボックスで*承認*を選択します。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できます。かかる時間はデータの量によって異なり、数分から数時間かかります。初期スキャンの進行状況は、[構成] メニューの [システム構成] セクションで追跡できます。データ分類では、スキャンごとに進行状況バーが表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"データ分類の制限について学ぶ"。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーを確認して、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームの場合、データ分類に CIFS 資格情報を提供する必要があります。

Google Cloud NetApp Volumesの場合、データ分類ではコンソールと同じリージョン内のボリュームのみをスキャンできます。

チェックリスト

- データ分類インスタンスと、 Google Cloud NetApp Volumesのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。
- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにします。

手順

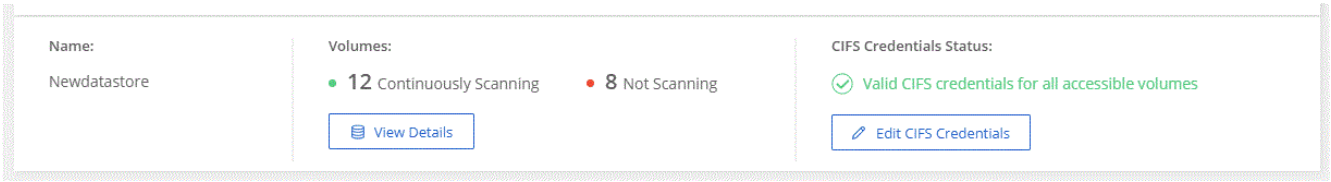
1. データ分類メニューから、*構成*を選択します。

   a. CIFS (SMB) を使用している場合は、Active Directory の資格情報が正しいことを確認してください。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

      データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

      資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。



2. [構成] ページで [詳細の表示] を選択し、各 CIFS ボリュームと NFS ボリュームのステータスを確認し、エラーを修正します。

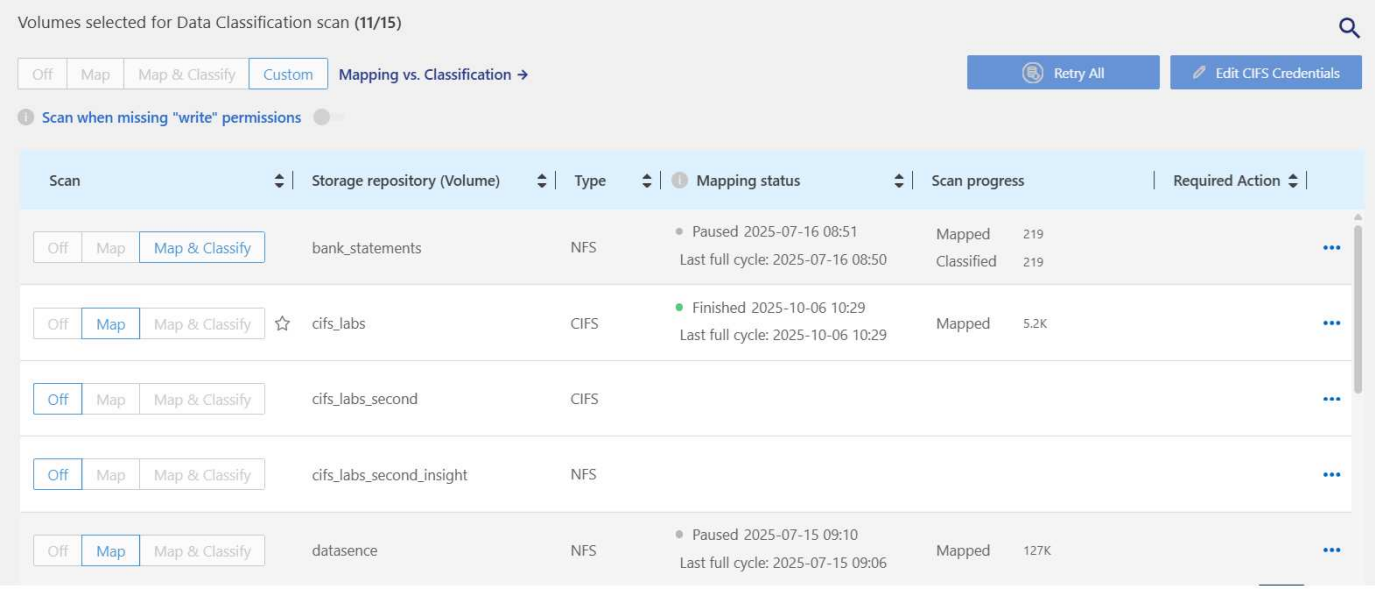
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

NetApp Data Classificationでファイル共有をスキャンする

ファイル共有をスキャンするには、まずNetApp Data Classificationでファイル共有グループを作成する必要があります。ファイル共有グループは、オンプレミスまたはクラウドでホストされる NFS または CIFS (SMB) 共有用です。

データ分類コア バージョンでは、 NetApp以外のファイル共有からのデータのスキャンはサポートされていません。

前提条件

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

- 共有は、クラウドやオンプレミスなど、どこにでもホストできます。古いNetApp 7-Mode ストレージ システムの CIFS 共有は、ファイル共有としてスキャンできます。
  - データ分類では、7-Mode システムから権限や「最終アクセス時刻」を抽出できません。

- 一部の Linux バージョンと 7-Mode システム上の CIFS 共有の間に既知の問題があるため、NTLM 認証が有効になっている SMBv1 のみを使用するように共有を構成する必要があります。

- データ分類インスタンスと共有の間にはネットワーク接続が必要です。
- DFS (分散ファイル システム) 共有を通常の CIFS 共有として追加できます。データ分類では、共有が複数のサーバー/ボリューム上に構築され、単一の CIFS 共有として結合されていることを認識しないため、メッセージが実際には別のサーバー/ボリュームにあるフォルダー/共有の 1 つにのみ適用される場合でも、共有に関するアクセス許可または接続エラーが発生する可能性があります。
- CIFS (SMB) 共有の場合は、共有への読み取りアクセスを提供する Active Directory 資格情報があることを確認します。データ分類で昇格された権限を必要とするデータをスキャンする必要がある場合は、管理者の資格情報が優先されます。

  データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

- グループ内のすべての CIFS ファイル共有では、同じ Active Directory 資格情報を使用する必要があります。
- NFS と CIFS (Kerberos または NTLM のいずれかを使用) の共有を混在させることができます。共有をグループに個別に追加する必要があります。つまり、プロトコルごとに 1 回ずつ、プロセスを 2 回完了する必要があります。
  - CIFS 認証タイプ (Kerberos と NTLM) が混在するファイル共有グループを作成することはできません。
- Kerberos 認証で CIFS を使用している場合は、提供された IP アドレスがデータ分類にアクセスできることを確認してください。 IP アドレスに到達できない場合は、ファイル共有を追加できません。

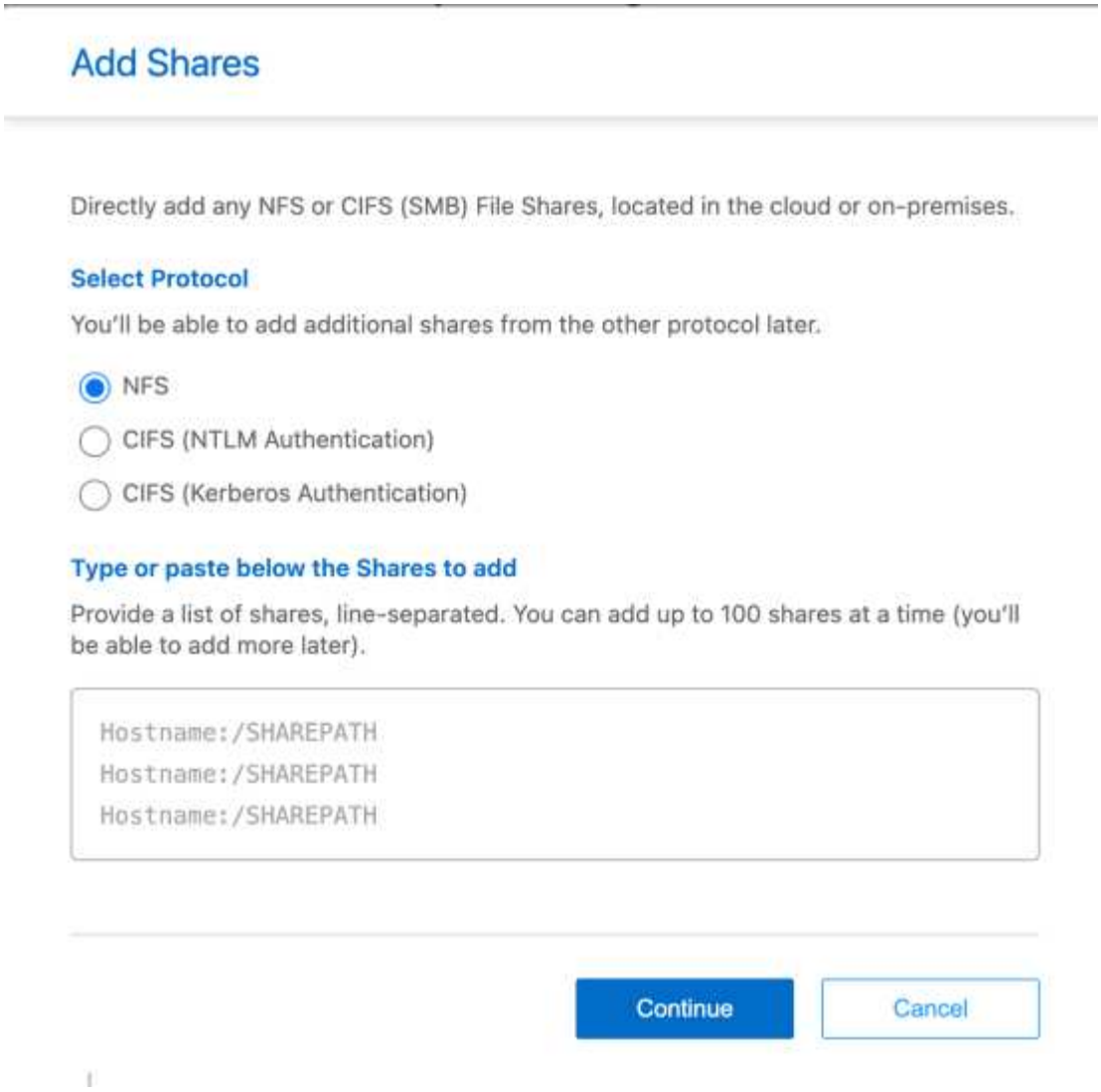
ファイル共有グループを作成する

グループにファイル共有を追加するときは、次の形式を使用する必要があります。
`<host_name>:/<share_path>` 。

ファイル共有を個別に追加することも、スキャンするファイル共有の行区切りリストを入力することもできます。一度に追加できる株式数は最大 100 です。

手順

1. データ分類メニューから、*構成*を選択します。
2. 構成ページで、システムの追加 > *ファイル共有グループの追加*を選択します。
3. [ファイル共有グループの追加] ダイアログで、共有グループの名前を入力し、[続行] を選択します。
4. 追加するファイル共有のプロトコルを選択します。



a. NTLM 認証を使用して CIFS 共有を追加する場合は、Active Directory の資格情報を入力して CIFS ボリュームにアクセスします。読み取り専用の資格情報はサポートされていますが、管理者の資格情報を使用してフルアクセスを提供することをお勧めします。 保存を選択します。

5. スキャンするファイル共有を追加します (1 行につき 1 つのファイル共有)。次に、[続行] を選択します。
6. 確認ダイアログに追加された共有数が表示されます。

   ダイアログに追加できなかった共有がリストされている場合は、問題を解決できるようにこの情報を取得します。問題が命名規則に関係する場合は、修正した名前で共有を再度追加できます。

7. ボリュームのスキャンを構成します。
   - ファイル共有でマッピングのみのスキャンを有効にするには、[マップ] を選択します。
   - ファイル共有の完全スキャンを有効にするには、[マップと分類] を選択します。
   - ファイル共有のスキャンを無効にするには、[オフ] を選択します。

ページ上部の「「属性の書き込み」権限がない場合にスキャンする」スイッチは、デフォルトでは無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。+ *「属性の書き込み」権限がない場合にスキャン*を*オン*に切り替えると、スキャンは最終アクセス時刻をリセットし、権限に関係なくすべてのファイルをスキャンします。+ 最終アクセスタイムスタンプの詳細については、以下を参照してください。"データ分類のデータソースから収集されたメタデータ" 。

結果

データ分類は、追加したファイル共有内のファイルのスキャンを開始します。あなたはできるxref:#track-the-scanning-progressダッシュボードでスキャンの結果を表示します。

Kerberos 認証を使用した CIFS 構成のスキャンが正常に完了しない場合は、[構成] タブでエラーを確認してください。

ファイル共有グループを編集する

ファイル共有グループを作成した後、CIFS プロトコルを編集したり、ファイル共有を追加および削除したりできます。

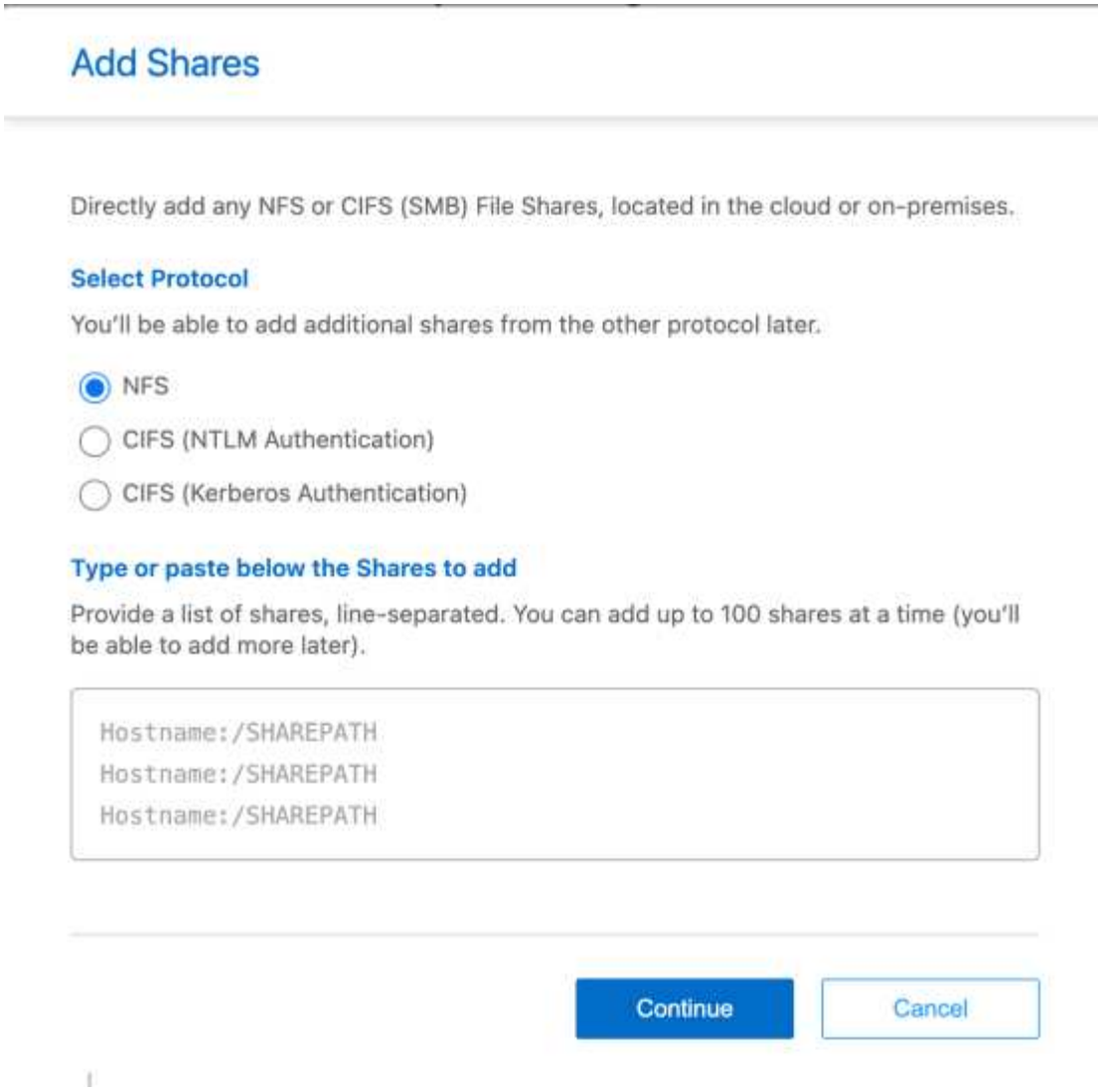
CIFSプロトコル設定を編集する

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで、変更するファイル共有グループを選択します。
3. **CIFS** 資格情報の編集 を選択します。

4. 認証方法を選択します: **NTLM** または **Kerberos**。
5. Active Directory の ユーザー名 と パスワード を入力します。
6. プロセスを完了するには、[保存] を選択します。

スキャンにファイル共有を追加する

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで、変更するファイル共有グループを選択します。
3. **+** 共有を追加 を選択します。
4. 追加するファイル共有のプロトコルを選択します。



すでに構成済みのプロトコルにファイル共有を追加する場合、変更は必要ありません。

2番目のプロトコルでファイル共有を追加する場合は、認証が適切に設定されていることを確認してください。"前提条件" 。

5. スキャンするファイル共有（1行につき1つのファイル共有）を次の形式で追加します。 `<host_name>:/<share_path>` 。
6. ファイル共有の追加を完了するには、[続行] を選択します。

スキャンからファイル共有を削除する

1. データ分類メニューから、*構成*を選択します。
2. ファイル共有を削除するシステムを選択します。
3. *構成*を選択します。
4. 設定ページからアクションを選択します ••• 削除するファイル共有の。
5. [アクション] メニューから、[共有を削除] を選択します。

スキャンの進行状況を追跡する

初期スキャンの進行状況を追跡できます。

1. 構成 メニューを選択します。
2. システム構成を選択します。
3. ストレージ リポジトリの場合は、スキャン進行状況列をチェックしてステータスを表示します。

NetApp Data ClassificationでStorageGRIDデータをスキャン

いくつかの手順を完了すると、 NetApp Data Classificationを使用してStorageGRID内のデータを直接スキャンし始めることができます。

StorageGRIDの要件を確認する

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

- オブジェクト ストレージ サービスに接続するには、エンドポイント URL が必要です。
- データ分類がバケットにアクセスできるようにするには、 StorageGRIDからのアクセス キーとシークレット キーが必要です。

データ分類インスタンスをデプロイする

インスタンスがまだデプロイされていない場合は、データ分類をデプロイします。

インターネット経由でアクセス可能なStorageGRIDからデータをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所にデータ分類を展開する"。

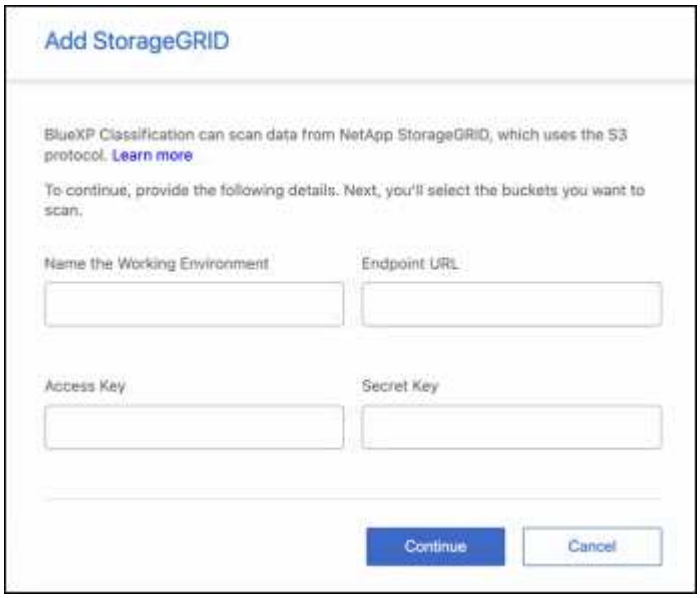
インターネットにアクセスできない暗い場所にインストールされているStorageGRIDからデータをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントが同じオンプレミスの場所に展開されていることも必要です。

データ分類にStorageGRIDサービスを追加する

StorageGRIDサービスを追加します。

手順

1. データ分類メニューから、*構成*オプションを選択します。
2. 構成ページで、システムの追加 > * StorageGRIDの追加*を選択します。
3. 「StorageGRIDサービスの追加」ダイアログで、 StorageGRIDサービスの詳細を入力し、「続行」を選択します。
   a. システムに使用する名前を入力します。この名前は、接続先のStorageGRIDサービスの名前を反映する必要があります。
   b. オブジェクト ストレージ サービスにアクセスするためのエンドポイント URL を入力します。
   c. データ分類がStorageGRID内のバケットにアクセスできるように、アクセス キーとシークレット キーを入力します。



結果

StorageGRIDがシステムのリストに追加されます。

StorageGRIDバケットのスキャンを有効または無効にする

StorageGRIDでデータ分類を有効にした後、次のステップはスキャンするバケットを構成することです。データ分類はこれらのバケットを検出し、作成したシステムに表示します。

手順

1. 構成ページで、 StorageGRIDシステムを見つけます。
2. StorageGRIDシステム タイルで、[構成] を選択します。
3. スキャンを有効または無効にするには、次のいずれかの手順を実行します。
   - バケットでマッピングのみのスキャンを有効にするには、[マップ] を選択します。
   - バケットの完全スキャンを有効にするには、[マップと分類] を選択します。
   - バケットのスキャンを無効にするには、[オフ] を選択します。

結果

データ分類は、有効にしたバケットのスキャンを開始します。 構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。各スキャンの進行状況は進行状況バーとして表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。エラーがある場合は、エラーを修正するために必要なアクションとともにステータス列に表示されます。

Active Directory とNetApp Data Classification を統合

グローバル Active Directory をNetApp Data Classificationと統合して、ファイル所有者やファイルにアクセスできるユーザーとグループに関する Data Classification のレポート結果を強化できます。

特定のデータ ソース (以下にリスト) を設定する場合、データ分類で CIFS ボリュームをスキャンするには、Active Directory の資格情報を入力する必要があります。この統合により、データ分類では、それらのデータ ソースに存在するデータのファイル所有者と権限の詳細が提供されます。これらのデータ ソースに入力された Active Directory は、ここで入力するグローバル Active Directory 資格情報とは異なる場合があります。データ分類では、統合されたすべての Active Directory でユーザーと権限の詳細を検索します。

この統合により、データ分類の次の場所に追加情報が提供されます。

- 「ファイル所有者」を使用することができます"フィルター"調査ペインのファイルのメタデータで結果を確認します。 SID (セキュリティ ID) を含むファイル所有者の代わりに、実際のユーザー名が入力されます。

  また、ファイル所有者の詳細情報（アカウント名、メール アドレス、SAM アカウント名）を表示したり、そのユーザーが所有するアイテムを表示したりすることもできます。

- 見ることができます"完全なファイル権限"「すべての権限を表示」ボタンをクリックすると、各ファイルとディレクトリに対して権限が表示されます。
- の中で"ガバナンスダッシュボード"、Open Permissions パネルにデータに関するより詳細な情報が表示されます。

ローカル ユーザーの SID および不明なドメインの SID は、実際のユーザー名に変換されません。

サポートされているデータソース

Active Directory とデータ分類の統合により、次のデータ ソース内のデータを識別できます。

- オンプレミスのONTAPシステム
- Cloud Volumes ONTAP
- Azure NetApp Files
- ONTAP向け FSx

Active Directoryサーバーに接続する

データ分類を展開し、データ ソースのスキャンを有効にしたら、データ分類を Active Directory と統合できます。 Active Directory には、DNS サーバー IP アドレスまたは LDAP サーバー IP アドレスを使用してアクセスできます。

Active Directory の資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

CIFS ボリューム/ファイル共有の場合、ファイルの「最終アクセス時刻」がデータ分類分類スキャンによって変更されていないことを確認するには、ユーザーに属性の書き込み権限が必要です。可能であれば、Active Directory で構成されたユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部にすることをお勧めします。

要件

- 社内のユーザー用に Active Directory がすでに設定されている必要があります。

- Active Directory の情報が必要です:
  - DNSサーバーのIPアドレス、または複数のIPアドレス

    または

  LDAPサーバーのIPアドレス、または複数のIPアドレス
  - サーバーにアクセスするためのユーザー名とパスワード
  - ドメイン名（アクティブディレクトリ名）
  - セキュアLDAP（LDAPS）を使用しているかどうか
  - LDAP サーバー ポート (通常、LDAP の場合は 389、セキュア LDAP の場合は 636)
- データ分類インスタンスによる送信通信のために、次のポートが開いている必要があります。

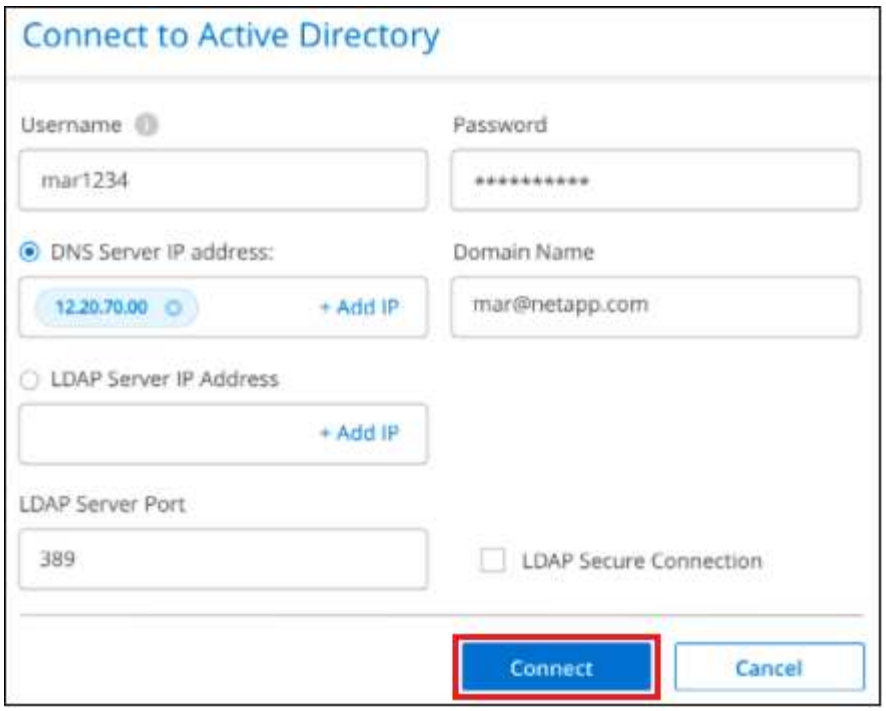
プロトコル	ポート	デスティネーション	目的
TCPとUDP	389	Active Directory	LDAP
TCP	636	Active Directory	LDAP over SSL
TCP	3268	Active Directory	グローバルカタログ
TCP	3269	Active Directory	SSL経由のグローバルカタログ

手順

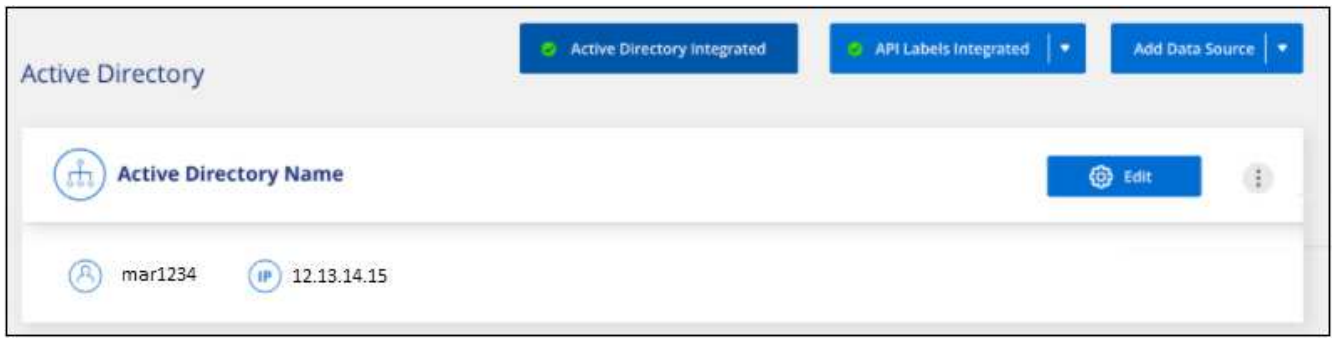
1. データ分類構成ページで、*Active Directory の追加*をクリックします。

2. 「Active Directory への接続」ダイアログで、Active Directory の詳細を入力し、「接続」をクリックします。

   必要に応じて、「**IP** を追加」を選択して複数の IP アドレスを追加できます。



データ分類は Active Directory に統合され、構成ページに新しいセクションが追加されます。



Active Directory統合を管理する

Active Directory 統合の値を変更する必要がある場合は、[編集] ボタンをクリックして変更を加えます。

統合を削除するには、 ボタンをクリックし、[Active Directory を削除] をクリックします。

データ分類を使用する

NetApp Data Classificationを使用して、組織に保存されているデータのガバナンスの詳細を表示します。

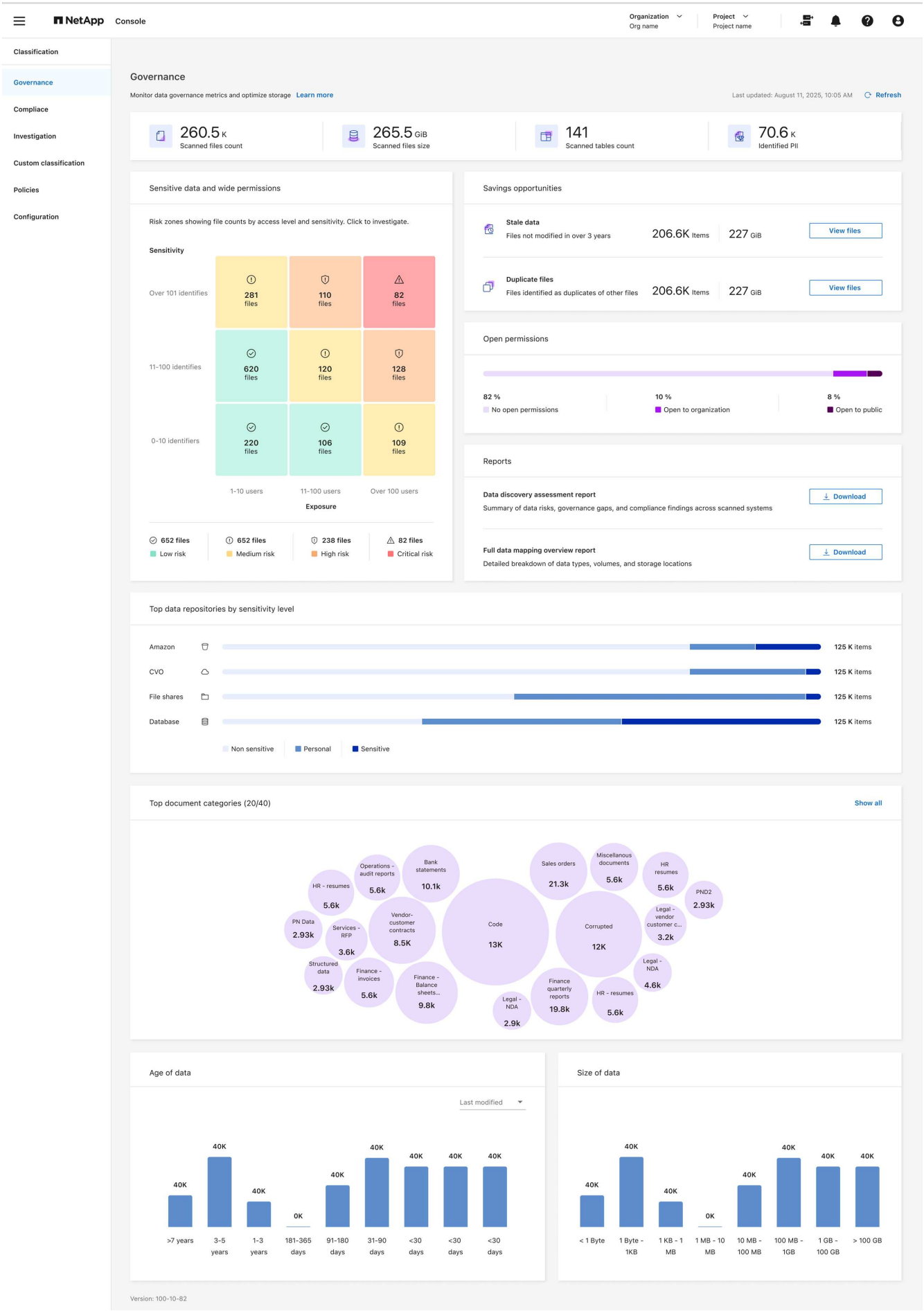
組織のストレージ リソース上のデータに関連するコストを管理します。 NetApp Data Classification は、システム内の古いデータ、重複ファイル、非常に大きなファイルの量を識別するので、一部のファイルを削除するか、より安価なオブジェクト ストレージに階層化するかを決定できます。

ここから研究を始めるべきです。ガバナンス ダッシュボードから、さらに調査する領域を選択できます。

さらに、オンプレミスの場所からクラウドにデータを移行する予定の場合は、データを移動する前に、データのサイズや、データに機密情報が含まれているかどうかを確認できます。

ガバナンスダッシュボードを確認する

ガバナンス ダッシュボードは、ストレージ リソースに保存されているデータに関連する効率を高め、コストを制御できるようにするための情報を提供します。


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

手順

1. NetApp Consoleメニューから、ガバナンス > 分類 を選択します。
2. *ガバナンス*を選択します。

   ガバナンス ダッシュボードが表示されます。

節約の機会を確認する

Saving Opportunities コンポーネントには、削除したり、より安価なオブジェクト ストレージに階層化したりできるデータが表示されます。 *Saving Opportunities* のデータは 2 時間ごとに更新されます。データを手動で更新することもできます。

手順

1. データ分類メニューから、*ガバナンス*を選択します。
2. ガバナンス ダッシュボードの各節約機会タイル内で、ストレージの最適化 を選択して、調査ページでフィルター処理された結果を表示します。削除する必要があるデータや、より安価なストレージに移動する必要があるデータを見つけるには、「*Saving Opportunities*」を調べてください。
   - 古いデータ - デフォルトでは、データが最後に変更されてから 3 年以上経過している場合、そのデータは古いものと見なされます。[古いデータの定義をカスタマイズ](task-stale-data.html)できます。
   - 重複ファイル - スキャンしているデータ ソース内の他の場所に重複しているファイル。"表示される重複ファイルの種類を確認する" 。

> いずれかのデータ ソースでデータ階層化が実装されている場合、オブジェクト ストレージに既に存在する古いデータは、古いデータ カテゴリで識別できます。

データ検出評価レポートを作成する

データ検出評価レポートでは、スキャンされた環境の高レベルの分析が提供され、懸念事項と潜在的な修復手順が示されます。結果はデータのマッピングと分類の両方に基づいています。このレポートの目的は、データセットの 3 つの重要な側面についての認識を高めることです。

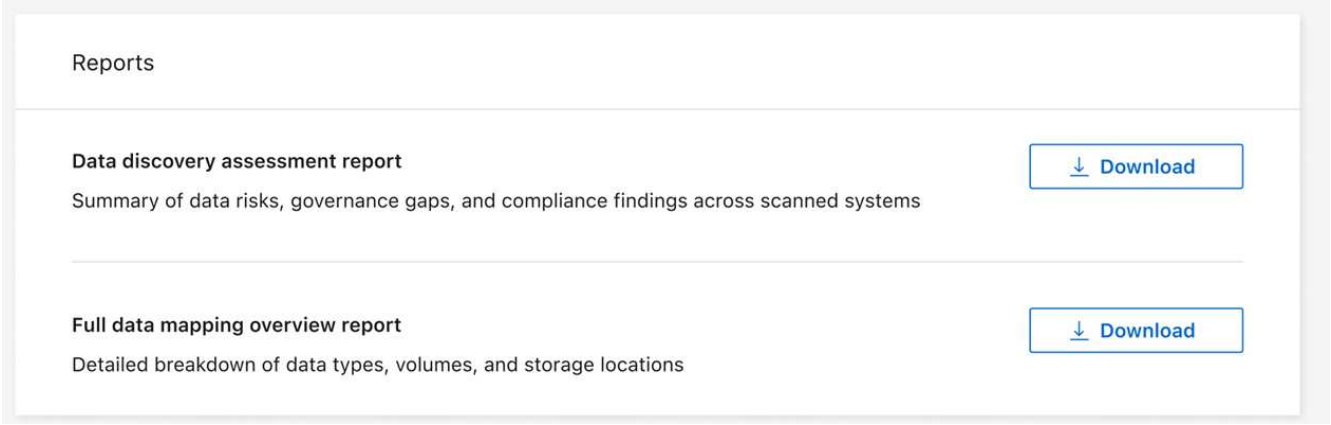
特徴	説明
データガバナンスの懸念	所有するすべてのデータの詳細な概要と、コストを節約するためにデータ量を削減できる可能性のある領域。
データセキュリティの露出	広範なアクセス権限により、内部または外部からの攻撃によってデータがアクセス可能になる領域。
データコンプライアンスのギャップ	セキュリティと DSAR (データ主体によるアクセス要求) の両方のために、個人情報または機密性の高い個人情報が保存される場所。

レポートを使用すると、次のアクションを実行できます。

- 保持ポリシーを変更するか、特定のデータ (古いデータや重複データ) を移動または削除することで、ストレージ コストを削減します。
- グローバル グループ管理ポリシーを改訂して、広範な権限を持つデータを保護します。
- PII をより安全なデータ ストアに移動することで、個人情報や機密性の高い個人情報が含まれるデータを保護します。

手順

1. データ分類から、*ガバナンス*を選択します。
2. レポート タイルで、データ検出評価レポート を選択します。



結果

データ分類では、確認および共有できる PDF レポートが生成されます。

データマッピングの概要レポートを作成する

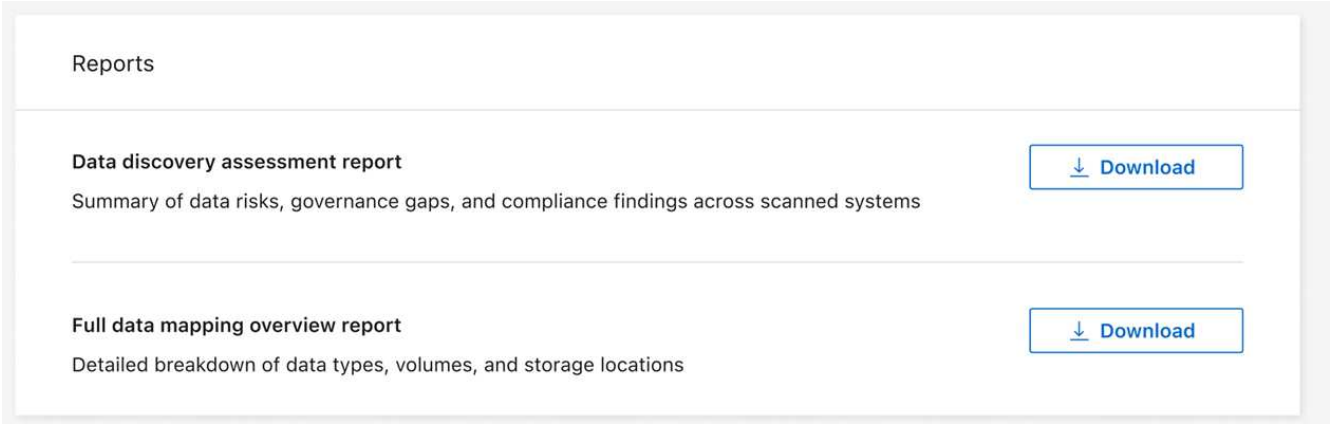
データ マッピングの概要レポートは、企業のデータ ソースに保存されているデータの概要を提供し、移行、バックアップ、セキュリティ、コンプライアンスのプロセスに関する意思決定に役立ちます。レポートでは、すべてのシステムとデータ ソースを要約します。各システムの分析も提供します。

レポートには次の情報が含まれます。

カテゴリ	説明
使用容量	すべてのシステム: 各システムのファイル数と使用容量を一覧表示します。単一システムの場合: 最も多くの容量を使用しているファイルを一覧表示します。
データの時代	ファイルが作成された日時、最後に変更された日時、最後にアクセスされた日時を示す 3 つのチャートとグラフを提供します。特定の日付範囲に基づいて、ファイルの数と使用容量を一覧表示します。
データのサイズ	システム内の特定のサイズ範囲内に存在するファイルの数を一覧表示します。

手順

1. データ分類から、*ガバナンス*を選択します。
2. レポート タイルで、*完全なデータ マッピングの概要レポート*を選択します。



結果

データ分類では、必要に応じて確認したり他のグループに送信したりできる PDF レポートが生成されます。

レポートが 1 MB より大きい場合、PDF ファイルはデータ分類インスタンスに保持され、正確な場所に関するポップアップ メッセージが表示されます。 Data Classification がオンプレミスの Linux マシンまたはクラウドに展開した Linux マシンにインストールされている場合は、PDF ファイルに直接移動できます。 Data Classification がクラウドにデプロイされている場合、PDF ファイルをダウンロードするには、Data Classification インスタンスへの SSH 認証が必要です。

データの機密性に基づいてリストされた上位のデータリポジトリを確認する

データ マッピングの概要レポートの「機密レベル別の上位データ リポジトリ」領域には、最も機密性の高い項目を含む上位 4 つのデータ リポジトリ (システムとデータ ソース) が一覧表示されます。各システムの棒グラフは次のように分かれています。

- 非機密データ
- 個人データ
- 機密性の高い個人データ

このデータは 2 時間ごとに更新され、手動で更新することもできます。

手順

1. 各カテゴリのアイテムの合計数を確認するには、バーの各セクションにカーソルを置きます。
2. 調査ページに表示される結果をフィルタリングするには、バー内の各領域を選択し、さらに調査します。

機密データと幅広い権限を確認する

ガバナンス ダッシュボードの [機密データと幅広い権限] 領域には、機密データが含まれており、幅広い権限を持つファイルの数が表示されます。表には次の種類の権限が示されています。

- 水平軸では、最も制限の厳しい権限から最も許容度の高い制限までを示します。
- 垂直軸では、最も機密性の低いデータから最も機密性の高いデータまで表示されます。

手順

1. 各カテゴリのファイルの合計数を確認するには、各ボックスの上にカーソルを置きます。

2. 調査ページに表示される結果をフィルタリングするには、ボックスを選択してさらに調査します。

オープン権限の種類別にリストされたデータを確認する

データ マッピングの概要レポートの [オープン アクセス許可] 領域には、スキャン対象のすべてのファイルに存在する各アクセス許可の種類の割合が表示されます。この表には、次の種類の権限が表示されます。

- 開く権限がありません
- 組織に開放
- 一般公開
- 不明なアクセス

手順

1. 各カテゴリのファイルの合計数を確認するには、各ボックスの上にカーソルを置きます。
2. 調査ページに表示される結果をフィルタリングするには、ボックスを選択してさらに調査します。

データの古さとサイズを確認する

データ マッピングの概要レポートの *Age* グラフと *Size* グラフの項目を調査して、削除する必要があるデータや、より安価なオブジェクト ストレージに階層化する必要があるデータがあるかどうかを確認できます。

手順

1. データの年齢グラフでデータの年齢に関する詳細を表示するには、グラフ内のポイントの上にカーソルを置きます。
2. 年齢またはサイズの範囲でフィルタリングするには、その年齢またはサイズを選択します。
    - データの年齢グラフ - データが作成された時刻、最後にアクセスされた時刻、または最後に変更された時刻に基づいてデータを分類します。
    - データ グラフのサイズ - サイズに基づいてデータを分類します。

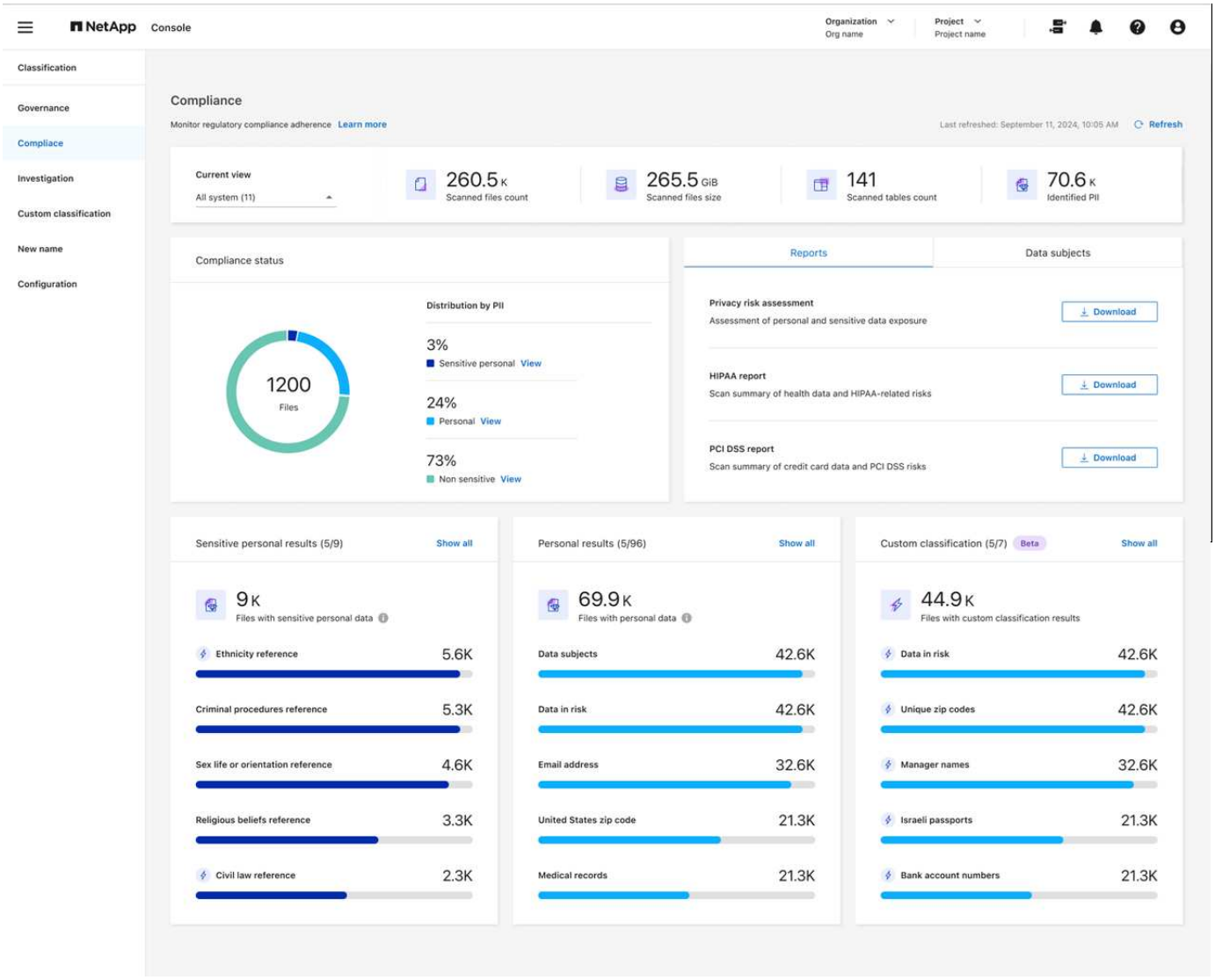
> いずれかのデータ ソースでデータ階層化が実装されている場合、オブジェクト ストレージに既に存在する古いデータが「データの年齢」グラフで識別される可能性があります。

NetApp Data Classification を使用して、組織内に保存されている個人データのコンプライアンスの詳細を表示します。

組織内の個人データ (PII) と機密個人データ (SPII) の詳細を表示して、プライベート データを制御します。また、 NetApp Data Classification がデータ内で検出したカテゴリとファイル タイプを確認することで、可視性を高めることもできます。

> ファイル レベルのコンプライアンスの詳細は、完全な分類スキャンを実行した場合にのみ利用できます。マッピングのみのスキャンでは、ファイルレベルの詳細は得られません。

デフォルトでは、データ分類ダッシュボードには、すべてのシステムとデータベースのコンプライアンス データが表示されます。一部のシステムのデータのみを表示するには、そのシステムを選択します。



データ調査ページから結果をフィルタリングし、結果のレポートを CSV ファイルとしてダウンロードできます。見る"データ調査ページでのデータのフィルタリング"詳細については。

個人データを含むファイルを表示する

データ分類は、データ内の特定の単語、文字列、パターン (正規表現) を自動的に識別します。"たとえば、クレジットカード番号、社会保障番号、銀行口座番号、パスワードなどです。"データ分類では、個々のファイル、ディレクトリ (共有とフォルダ) 内のファイル、およびデータベース テーブル内のこの種類の情報を識別します。

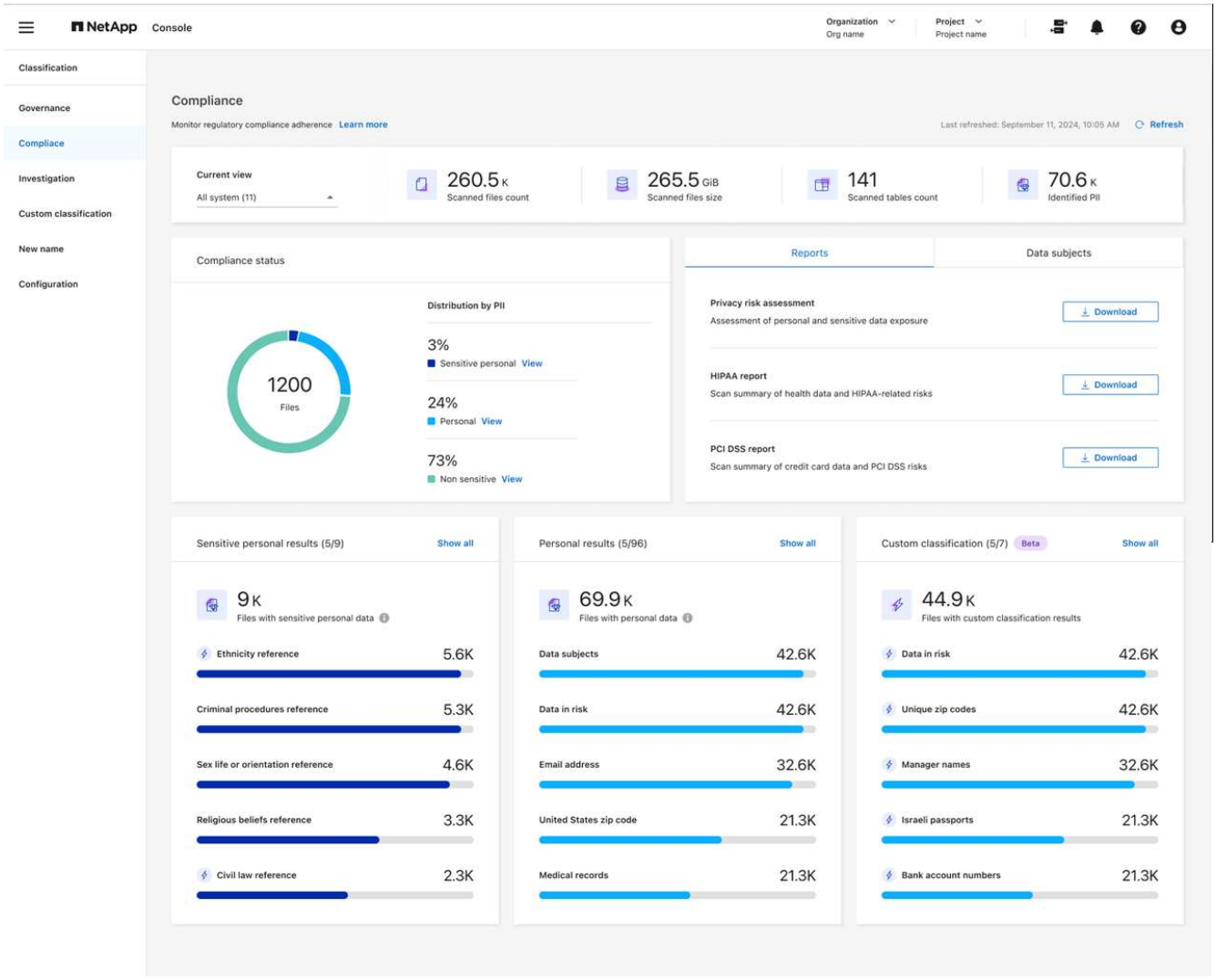
組織固有の個人データを識別するためのカスタム検索用語を作成することもできます。詳細については、以下を参照してください。 "カスタム分類を作成する" 。

一部の種類の個人データについては、データ分類は近接検証を使用して結果を検証します。検証は、見つかった個人データの近くにある 1 つ以上の定義済みキーワードを探すことによって行われます。たとえば、データ分類では、米国の社会保障番号 (SSN) の横に近接語 (たとえば、*SSN* または *social security*) がある場合、SSN として識別します。"個人データの表"データ分類で近接検証が使用されるタイミングを示します。

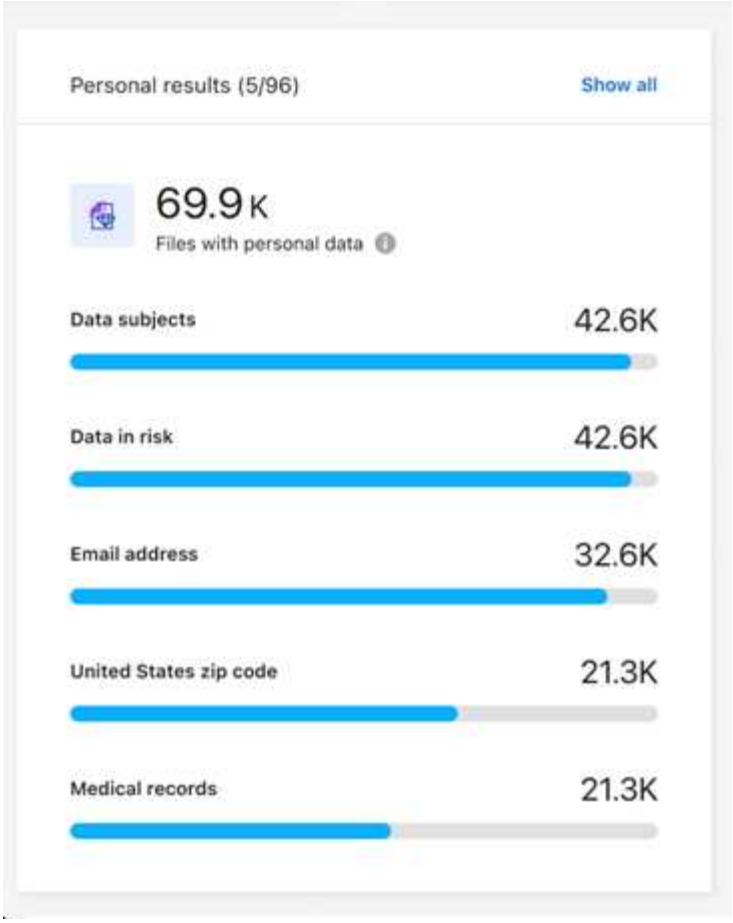
手順

1. データ分類メニューから、*コンプライアンス*タブを選択します。

2. すべての個人データの詳細を調査するには、個人データの割合の横にあるアイコンを選択します。



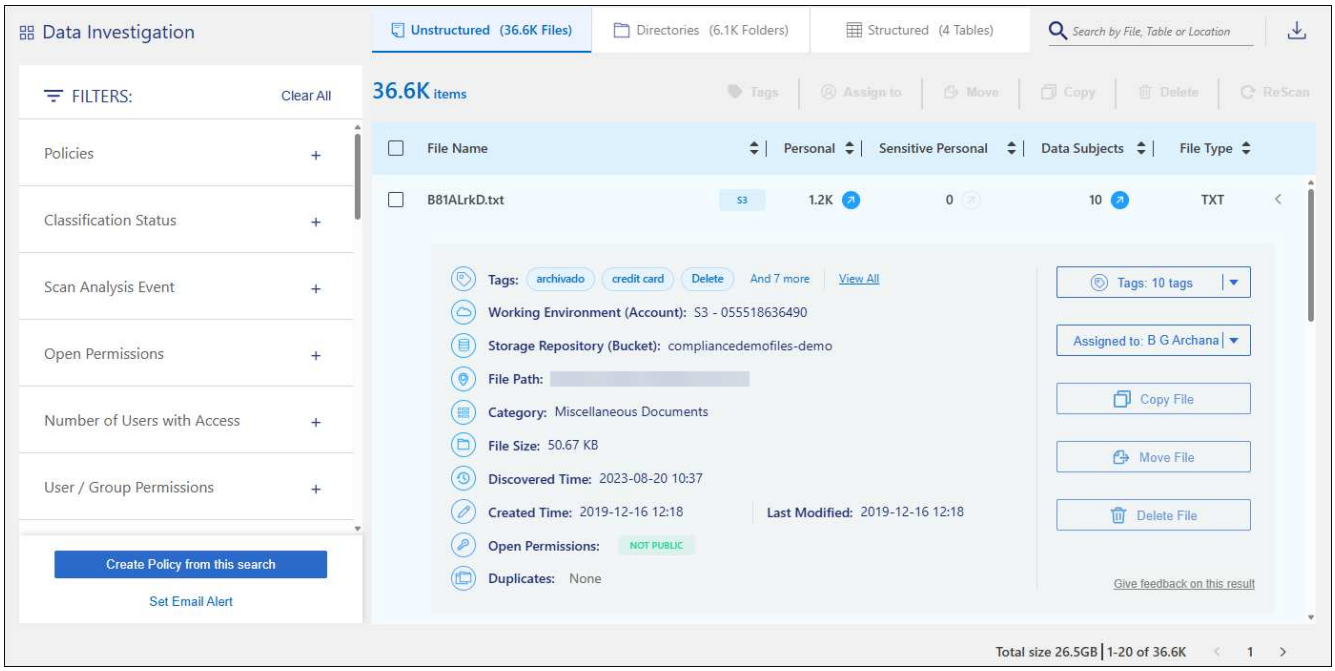
3. 特定の種類の個人データの詳細を調査するには、[すべて表示] を選択し、電子メール アドレスなどの特定の種類の個人データの [結果の調査] 矢印アイコンを選択します。

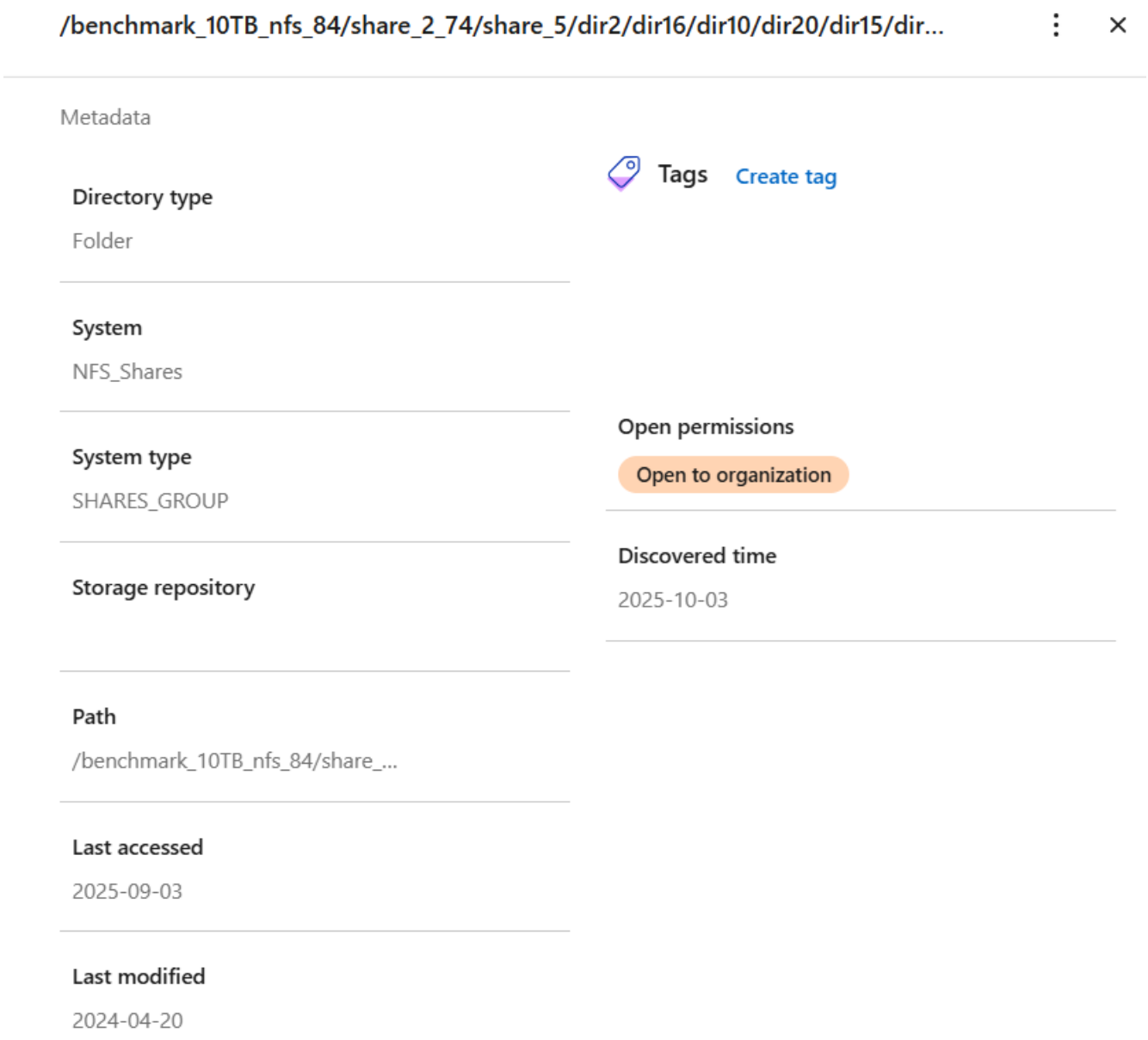


アイコンを選択できる個人ファイル ダイアログ ボックスのスクリーンショット。"]

4. 特定のファイルの詳細を検索、並べ替え、展開したり、*結果の調査*矢印を選択してマスクされた情報を表示したり、ファイル リストをダウンロードしたりして、データを調査します。

   次の画像は、ディレクトリ (共有およびフォルダー) で見つかった個人データを示しています。*構造化*タブでは、データベース内にある個人データを表示します。*非構造化*タブでは、ファイルレベルのデータを表示できます。


Data Investigation
Unstructured (36.6K Files)
Directories (6.1K Folders)
Structured (4 Tables)
Search by File, Table or Location
FILTERS:
Clear All
Policies
Classification Status
Scan Analysis Event
Open Permissions
Number of Users with Access
User / Group Permissions
Create Policy from this search
Set Email Alert
36.6K items
Tags
Assign to
Move
Copy
Delete
ReScan
File Name
Personal
Sensitive Personal
Data Subjects
File Type
B81ALrkD.txt
S3
1.2K
0
10
TXT
Tags: archivado credit card Delete And 7 more View All
Working Environment (Account): S3 - 055518636490
Storage Repository (Bucket): compliancedemofiles-demo
File Path:
Category: Miscellaneous Documents
File Size: 50.67 KB
Discovered Time: 2023-08-20 10:37
Created Time: 2019-12-16 12:18
Last Modified: 2019-12-16 12:18
Open Permissions: NOT PUBLIC
Duplicates: None
Tags: 10 tags
Assigned to: B G Archana
Copy File
Move File
Delete File
Give feedback on this result
Total size 26.5GB | 1-20 of 36.6K
1



機密性の高い個人データを含むファイルを表示する

データ分類は、プライバシー規制で定義されている特別な種類の機密個人情報を自動的に識別します。"GDPR第9条および第10条" 。たとえば、個人の健康、民族的起源、性的指向に関する情報などです。"全リストを見る" 。データ分類では、個々のファイル、ディレクトリ (共有とフォルダ) 内のファイル、およびデータベース テーブル内のこの種類の情報を識別します。

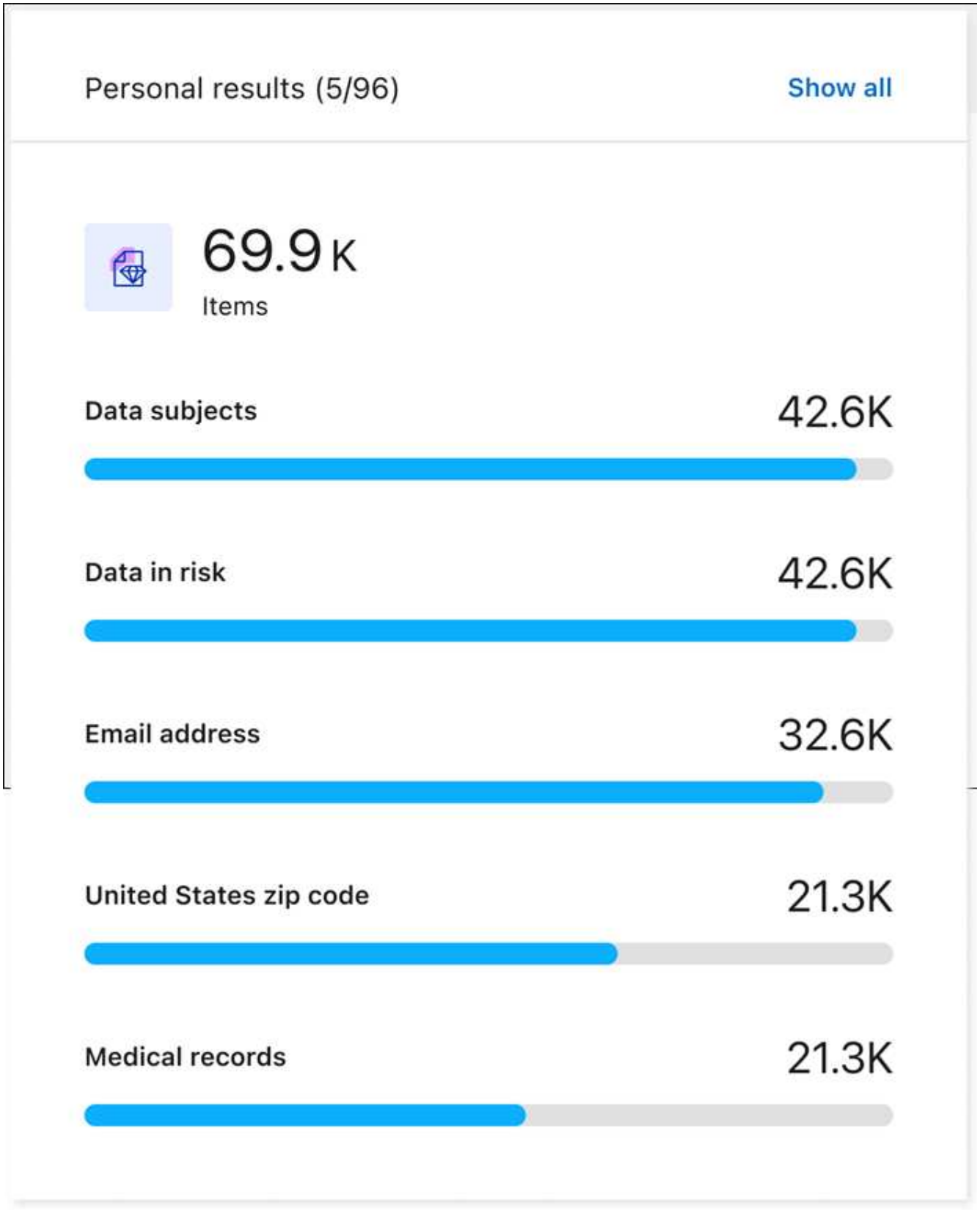
データ分類では、AI、自然言語処理 (NLP)、機械学習 (ML)、認知コンピューティング (CC) を使用して、スキャンしたコンテンツの意味を理解し、エンティティを抽出してそれに応じて分類します。

たとえば、GDPR の機密データ カテゴリの 1 つは民族的起源です。データ分類では、NLP 機能により、「George はメキシコ人です」(GDPR の第 9 条に指定されている機密データを示します) という文と、「George はメキシコ料理を食べています」という文を区別できます。

機密性の高い個人データをスキャンする場合は、英語のみがサポートされます。今後、さらに多くの言語のサポートが追加される予定です。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. すべての機密個人データの詳細を調査するには、[機密性の高い個人データの結果] カードを見つけて、[すべて表示] を選択します。



。

3. 特定の種類の機密個人データの詳細を調査するには、[すべて表示] を選択し、特定の種類の機密個人データの [結果の調査] 矢印アイコンを選択します。
4. 特定のファイルの詳細を検索、並べ替え、展開したり、「結果の調査」をクリックしてマスクされた情報を表示したり、ファイル リストをダウンロードしたりして、データを調査します。

NetApp Data Classificationにおけるプライベートデータのカテゴリ

NetApp Data Classification がボリュームやデータベース内で識別できるプライベート データには多くの種類があります。

データ分類では、次の 2 種類の個人データを識別します。

- 個人を特定できる情報（**PII**）
- 機密個人情報（**SPII**）

追加の国民 ID 番号や医療 ID 番号など、他のプライベート データの種類を識別するためにデータ分類が必要な場合は、アカウント マネージャーにお問い合わせください。

個人データの種類

ファイル内に含まれる個人データ、つまり個人を特定できる情報 (PII) は、一般的な個人データまたは国民識別子である場合があります。下の表の3番目の列は、データ分類が使用するかどうかを示しています。"近接検証"識別子に対する調査結果を検証するため。

これらの項目を認識できる言語は表に示されています。

タイプ	識別子	近接検証？	英語	ドイツ語	スペイン語	フランス語	日本語
全般	クレジットカード番号	はい	✓	✓	✓		✓
	データ主体	いいえ	✓	✓	✓		
	E メール アドレス	いいえ	✓	✓	✓		✓
	IBAN番号（国際銀行口座番号）	いいえ	✓	✓	✓		✓
	IP アドレス	いいえ	✓	✓	✓		✓
	パスワード	はい	✓	✓	✓		✓

タイプ	識別子	近接検証?	英語	ドイツ語	スペイン語	フランス語	日本語
国民識別子							

タイプ	識別子	近接検証?	英語	ドイツ語	スペイン語	フランス語	日本語

タイプ	識別子	近接検証?	英語	ドイツ語	スペイン語	フランス語	日本語

	（NRIC）						
	スロベニアID（EMSO）	はい	✓	✓	✓		
	南アフリカのID	はい	✓	✓	✓		
タイプ	識別子	近接検証	英語	ドイツ語	スペイン語	フランス語	日本語
	スペイン納税者番号	はい	✓	✓	✓		
	スウェーデンのID	はい	✓	✓	✓		
	英国ID（NINO）	はい	✓	✓	✓		
	米国カリフォルニア州運転免許証	はい	✓	✓	✓		
	米国インディアナ州運転免許証	はい	✓	✓	✓		
	米国ニューヨーク州運転免許証	はい	✓	✓	✓		
	米国テキサス州運転免許証	はい	✓	✓	✓		
	米国社会保障番号（SSN）	はい	✓	✓	✓		

機密個人データの種類

データ分類では、ファイル内の次の機密個人情報 (SPII) を見つけることができます。

以下の SPII は現在英語でのみ認識できます。

- 刑事手続きリファレンス: 自然人の刑事上の有罪判決および犯罪に関するデータ。
- 民族参照: 自然人の人種または民族的起源に関するデータ。
- 健康参考: 自然人の健康に関するデータ。
- **ICD-9-CM** 医療コード: 医療および健康業界で使用されるコード。
- **ICD-10-CM** 医療コード: 医療および健康業界で使用されるコード。
- 哲学的信念の参照: 自然人の哲学的信念に関するデータ。
- 政治的意見参照: 自然人の政治的意見に関するデータ。
- 宗教的信念の参照: 自然人の宗教的信念に関するデータ。
- 性生活または性的指向の参照: 自然人の性生活または性的指向に関するデータ。

カテゴリーの種類

データ分類では、データを次のように分類します。

これらのカテゴリのほとんどは、英語、ドイツ語、スペイン語で認識できます。

カテゴリ	タイプ	英語	ドイツ語	スペイン語
Finance	貸借対照表	✓	✓	✓
	発注書	✓	✓	✓
	請求書	✓	✓	✓
	四半期報告書	✓	✓	✓

カテゴリ	タイプ	英語	ドイツ語	スペイン語
人事	身元調査	✓		✓
	報酬プラン	✓	✓	✓
	従業員契約	✓		✓
	従業員レビュー	✓		✓
	健康	✓		✓
	履歴書	✓	✓	✓
法律上の	NDA	✓	✓	✓
	ベンダーと顧客の契約	✓	✓	✓
マーケティング	キャンペーン	✓	✓	✓
	会議	✓	✓	✓
オペレーション	監査報告書	✓	✓	✓
売り上げ	販売注文	✓	✓	
サービス	情報提供依頼	✓		✓
	提案依頼書	✓		✓
	種をまく	✓	✓	✓
	トレーニング	✓	✓	✓
サポート	苦情とチケット	✓	✓	✓

次のメタデータも同じサポートされている言語で分類および識別されます。

- アプリケーションデータ
- アーカイブファイル
- オーディオ
- データ分類ビジネスアプリケーションデータからのパンくずリスト
- CADファイル
- コード
- 破損した
- データベースとインデックスファイル
- デザインファイル
- 電子メールアプリケーションデータ
- 暗号化されたファイル（エントロピースコアの高いファイル）
- 実行可能ファイル
- 金融アプリケーションデータ
- 健康アプリケーションデータ

- イメージ
- Logs
- その他の文書
- その他のプレゼンテーション
- その他のスプレッドシート
- その他「不明」
- パスワード保護されたファイル
- 構造化データ
- ビデオ
- ゼロバイトファイル

ファイルの種類

データ分類では、すべてのファイルをスキャンしてカテゴリとメタデータの分析情報を取得し、ダッシュボードのファイル タイプ セクションにすべてのファイル タイプを表示します。データ分類が個人識別情報 (PII) を検出する場合、または DSAR 検索を実行する場合は、次のファイル形式のみがサポートされます。

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

発見された情報の正確性

NetApp は、データ分類によって識別される個人データおよび機密個人データの 100% の正確性を保証することはできません。常にデータを確認して情報を検証する必要があります。

当社のテストに基づき、以下の表はデータ分類が検出した情報の正確性を示しています。これを 精度 と 再現率 で分類します。

精度
: データ分類によって検出されたものが正しく識別された確率。たとえば、個人データの精度率が 90% の場合、個人情報が含まれていると識別された 10 個のファイルのうち 9 個に実際に個人情報が含まれていることを意味します。 10 個のファイルのうち 1 個は誤検出となります。

想起
: データ分類が必要なものを見つける確率。たとえば、個人データのリコール率が 70% の場合、データ分類では組織内の個人情報が実際に含まれているファイル 10 個のうち 7 個を識別できることになります。データ分類ではデータの 30% が失われ、ダッシュボードに表示されません。

当社は結果の精度を継続的に向上させています。これらの改善は、将来のデータ分類リリースで自動的に利用できるようになります。

タイプ	精度	想起
個人データ - 一般	90%～95%	60%～80%
個人データ - 国識別子	30%～60%	40%～60%
機密性の高い個人データ	80%～95%	20%～30%

タイプ	精度	想起
カテゴリ	90%～97%	60%～80%

NetApp Data Classificationでカスタム分類を作成する

NetApp Data Classification を使用すると、カスタム カテゴリまたは個人識別子を作成して、組織の規制およびコンプライアンス要件に固有のデータを識別できます。

データ分類では、カテゴリと個人識別子の 2 種類のカスタム分類子がサポートされています。カスタム カテゴリは、アップロードした一連のファイルに基づいて作成され、データ分類によって組織内の類似データを識別する AI モデルが作成されます (たとえば、健康研究会社では臨床分析カテゴリを作成する場合があります)。カスタム個人識別子は、キーワードリストまたは正規表現 (regex) を使用して作成され、コンプライアンスリスクをもたらす可能性のある組織固有の情報を識別します。

すべてのカスタム分類は、カスタム分類ダッシュボードで利用できます。

カスタム個人識別子を作成する

データ分類を使用すると、コンテキスト キーワードまたは正規表現を使用してカスタム個人識別子を作成し、組織固有のデータを識別できます。

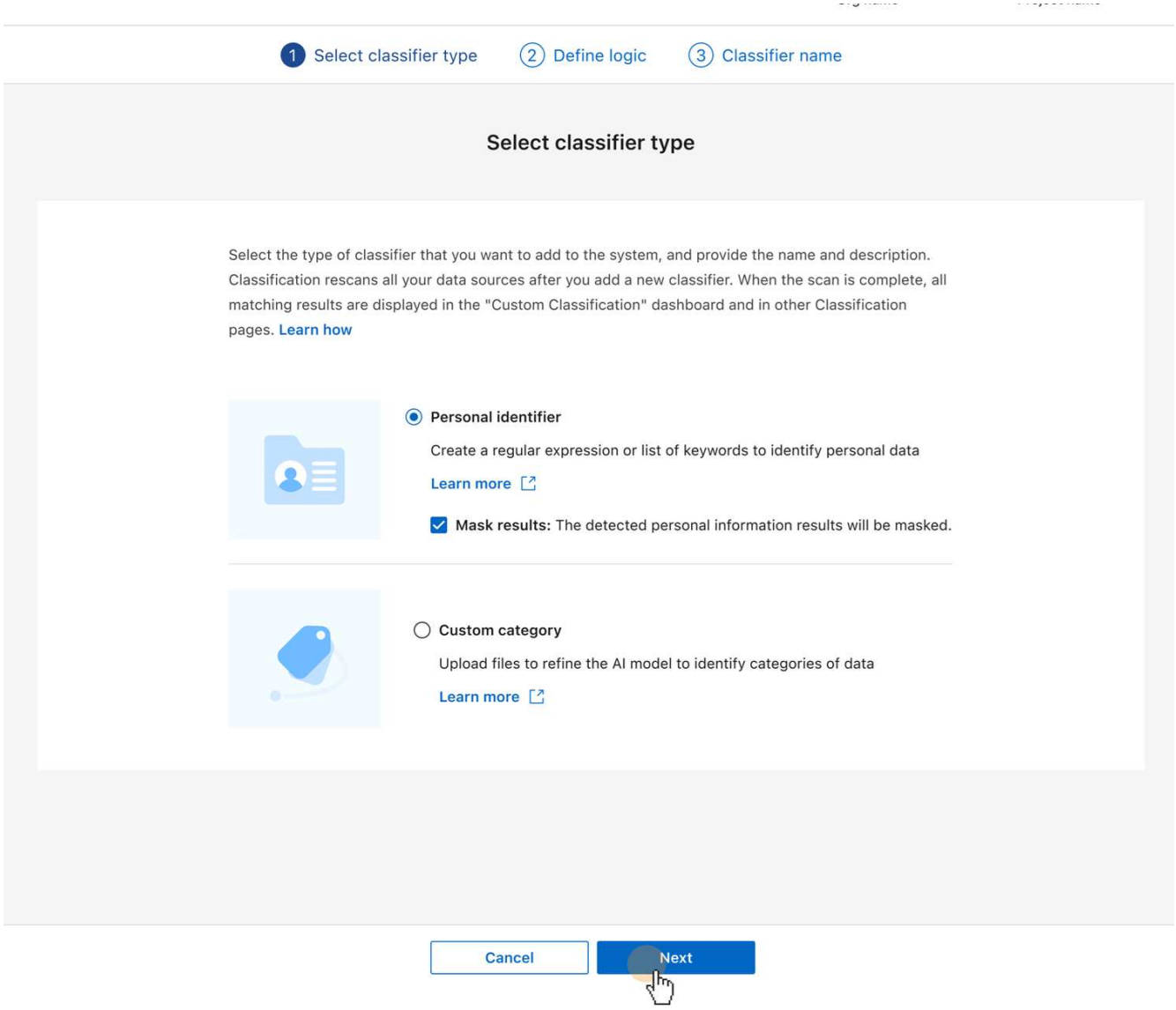
キーワードの要件

キーワードリストを使用して個人識別子を作成する場合、リストは次の要件を満たしている必要があります。

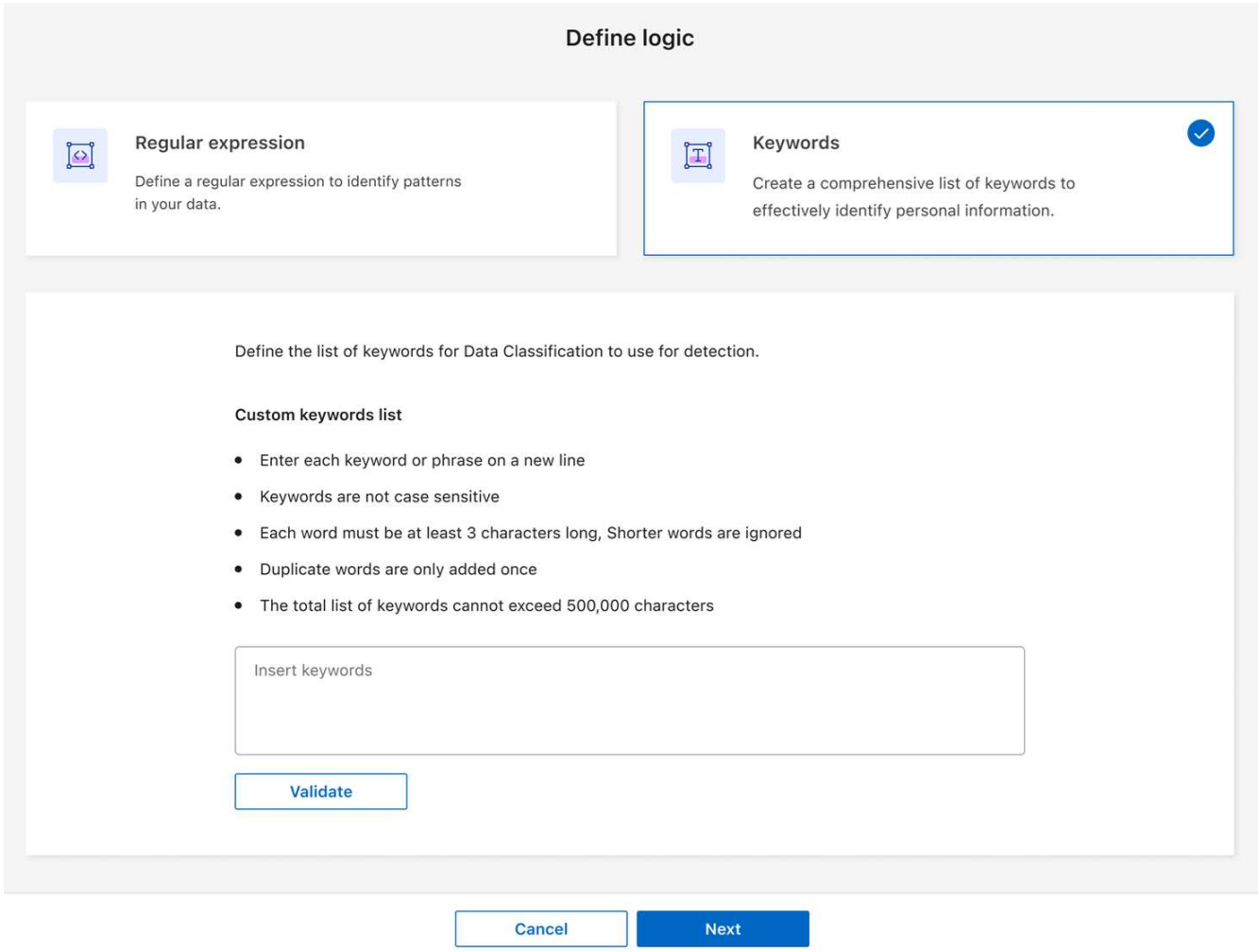
- キーワードの入力では大文字と小文字は区別されません。
- キーワードは 3 文字以上である必要があります。3 文字未満の単語は無視されます。
- 重複する単語は 1 回だけ追加されます。
- キーワードの合計リストは 500,000 文字を超えることはできません。リストには少なくとも 1 つのキーワードが含まれている必要があります。

手順

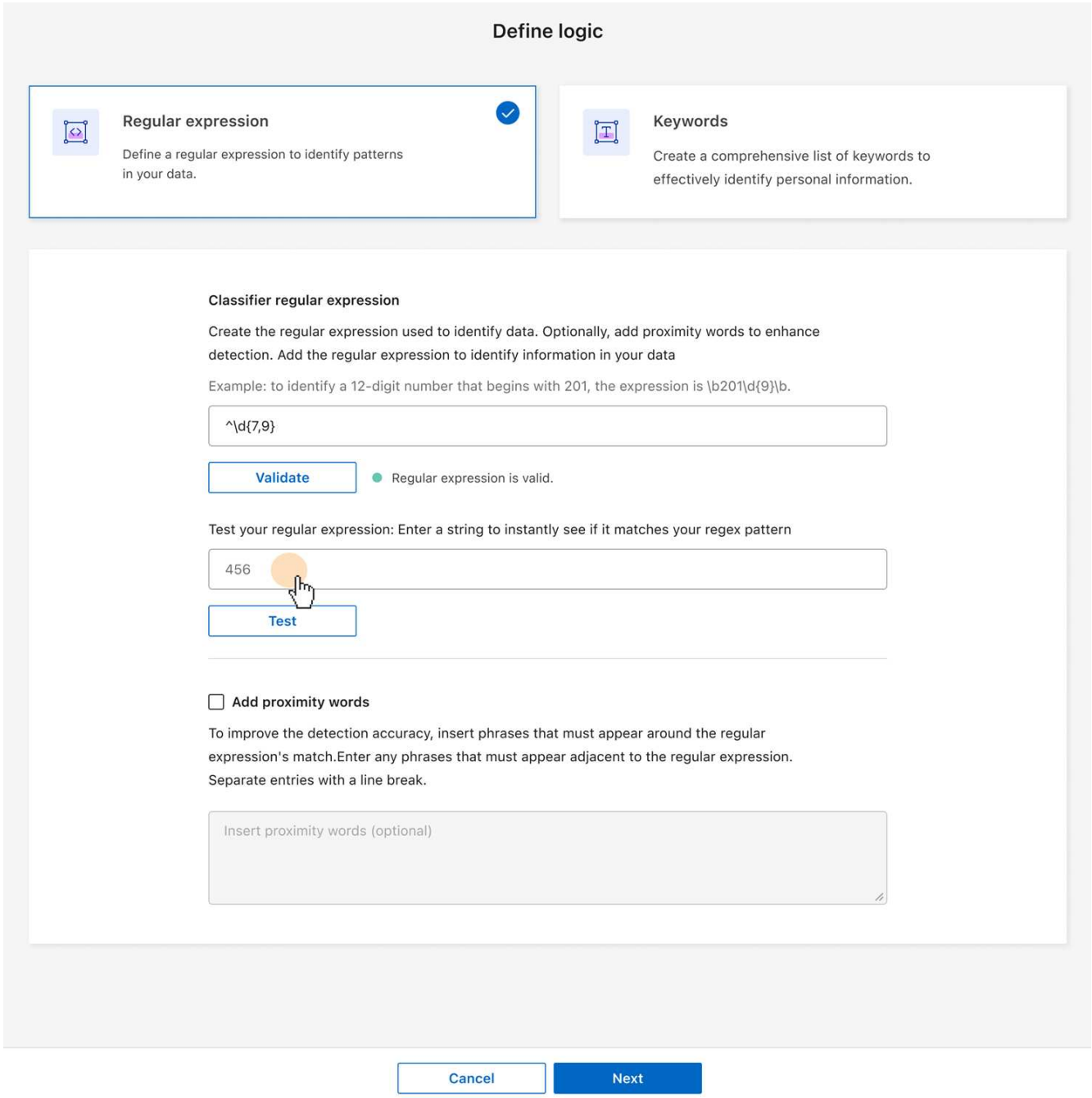
1. カスタム分類 タブを選択します。
2. カスタム分類子を作成するには、[+ 新しい分類子] を選択します。
3. *個人識別子*を選択します。必要に応じて、結果をマスク を選択して、検出された個人データをマスクします。
4. 次へを選択します。



5. キーワードを含む分類子を追加するには、キーワード を選択します。キーワードのリストを入力します。各エントリは別々の行に入力します。キーワードが要件に準拠していることを確認します。



分類子を正規表現として追加するには、正規表現 を選択し、データの特定の情報を検出するためのパターンを追加します。入力した構文を確認するには、[検証] を選択します。



a. 必要に応じて、正規表現パターンに一致するサンプル文字列を入力し、[テスト] を選択して確認します。

b. 必要に応じて、近接語を追加します。近接単語を追加すると、データ分類では、近接単語が一致する文字列に隣接している場合にのみ正規表現パターンにフラグが付けられます。

6. 次へを選択します。

7. ダッシュボードでカスタム カテゴリを識別するには、分類子名 と 概要 を入力します。

8. **[保存]** を選択して、カスタム個人識別子を作成します。

カスタム個人識別子を作成すると、その結果は次回のスケジュールされたスキャンでキャプチャされます。より早く結果を取得するには、オンデマンドスキャンを実行します。結果を見るには、 コンプライアンスレポートを生成する。

カスタムカテゴリを作成する

カスタム カテゴリを使用すると、組織固有のデータを分類できます。カスタム カテゴリは、アップロードしたテキスト ファイルに基づいて作成されます。データ分類では、このテキスト ファイルから AI モデルを作成し、他のファイル内の同様の情報を識別します。

トレーニングデータ要件

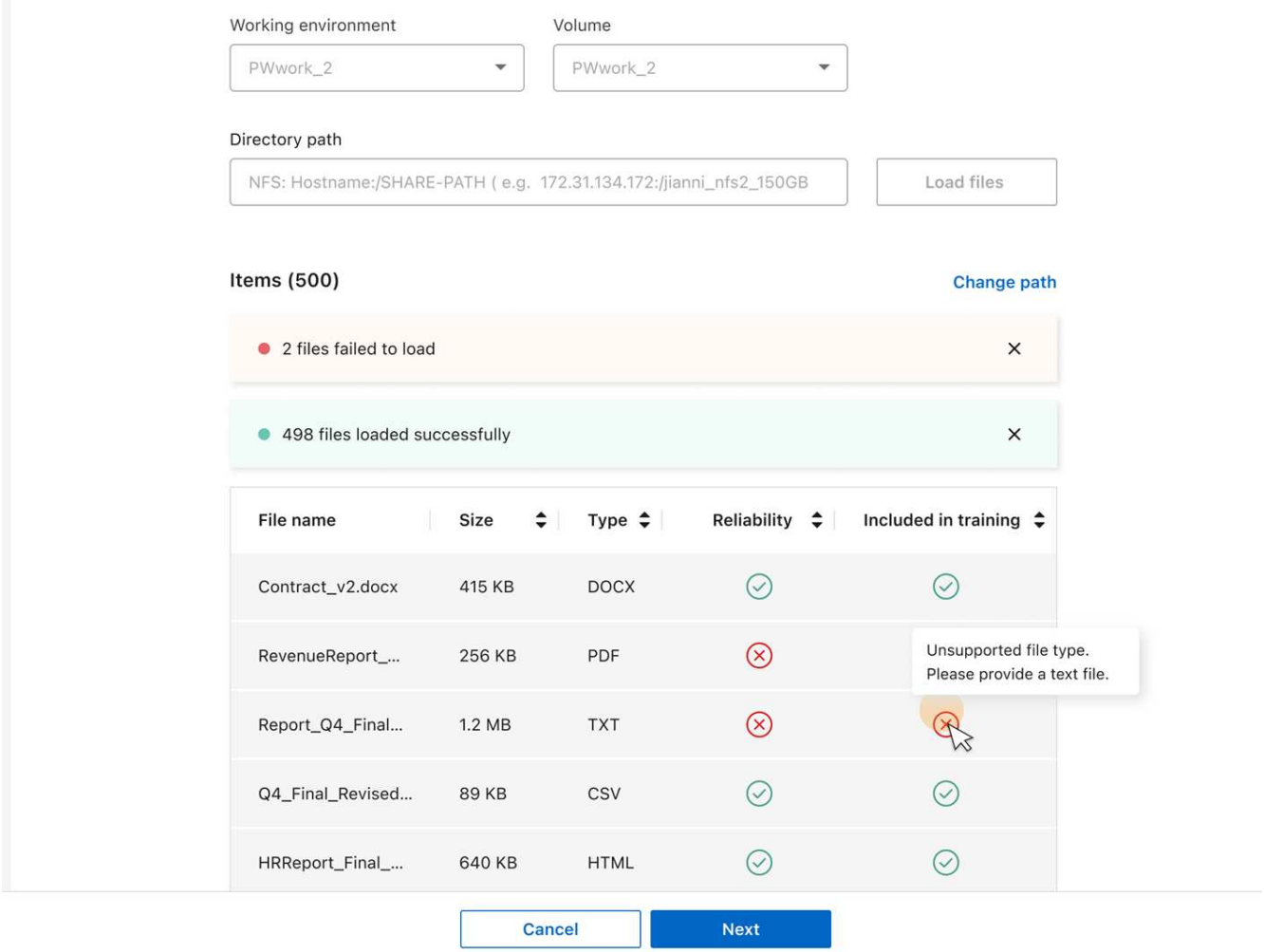
- トレーニング データセットには少なくとも 25 個のファイルが含まれている必要があります。最大ファイル数は 1,000 です。
- すべてのファイルは、指定したファイル パスに直接配置する必要があります。
- すべてのファイルは 100 バイトより大きくなければなりません。
- データ分類トレーニング データは、CSV、DOCX、DOC、GZ、JSON、PDF、PPTX、TXT、RTT、XLS、または XLSX のいずれかのファイル タイプである必要があります。サポートされているすべてのファイルタイプを組み合わせてアップロードできます。

手順

1. NetApp Data Classificationで、*カスタム分類*を選択します。
2. *+ 新しい分類子*を選択します。
3. 分類子の種類として「カスタム カテゴリ」を選択し、「次へ」をクリックします。
4. テキストベースのファイルのコレクションを使用して、カスタム カテゴリのロジックを定義します。作業アドレス*の **IP** アドレスを入力し、ドロップダウン メニューから *ボリューム を選択します。

   トレーニング データが含まれているディレクトリの ディレクトリ パス を入力します。

5. ファイルのチェックを実行するには、データ分類の [ファイルの読み込み] を選択します。ファイルの概要を確認することができます。そこには、ファイル名、サイズ、タイプ、およびファイルがトレーニングに適していると判断されたかどうかのメモが一覧表示されます。



a. ファイル パスを変更したり、ファイルを再アップロードしたりするには、[パスの変更] を選択し、データを入力してファイルを再度ロードします。

6. アップロードしたファイルに問題がなければ、[次へ] を選択します。
7. ダッシュボードでカスタム カテゴリを識別するには、分類子名 と 概要 を入力します。
8. **[保存]** を選択してカスタム カテゴリを作成します。

結果

カスタム カテゴリを作成すると、その結果は次回のスケジュールされたスキャンでキャプチャされます。より早く結果を取得するには、手動でスキャンを開始します。

カスタム分類子を編集する

個人識別子を作成した後で、そのロジックを変更できます。個人識別子のタイプまたはロジック タイプを変更することはできません。たとえば、カスタム カテゴリをカスタム個人識別子に変更することはできません。また、キーワードベースのカスタム識別子を正規表現ベースのカスタム識別子に変更することもできません。

手順

1. NetApp Data Classificationで、*カスタム分類*を選択します。
2. 削除したい分類子を特定し、アクションメニューを選択します ... 列の末尾にあります。

3. ロジックの編集を選択します。
4. キーワードを変更する場合は、適切なキーワードを追加、削除、または編集します。正規表現を変更する場合は、新しい正規表現を入力して検証します。必要に応じて、近接キーワードを追加します。
5. 変更を適用するには、[保存] を選択します。

カスタム分類子を削除する

1. NetApp Data Classificationで、*カスタム分類*を選択します。
2. 削除したい分類子を識別し、アクションメニューを選択します ... 列の末尾にあります。
3. 分類子を削除 を選択します。

次のステップ

- コンプライアンスレポートを生成する

NetApp Data Classificationを使用して組織内に保存されているデータを調査します

データ調査ダッシュボードには、データに関するファイルおよびディレクトリ レベルの分析情報が表示され、結果を並べ替えたりフィルター処理したりできます。データ調査ページでは、ファイルとディレクトリのメタデータと権限に関する詳細情報が表示され、重複したファイルも識別されます。ファイル、ディレクトリ、データベース レベルの分析情報を活用することで、組織のコンプライアンスを向上させ、ストレージ スペースを節約するためのアクションを実行できます。データ調査ページでは、ファイルの移動、コピー、削除もサポートされています。

調査ページから洞察を得るには、データ ソースに対して完全な分類スキャンを実行する必要があります。マッピングのみのスキャンが行われたデータ ソースでは、ファイル レベルの詳細は表示されません。

データ調査構造

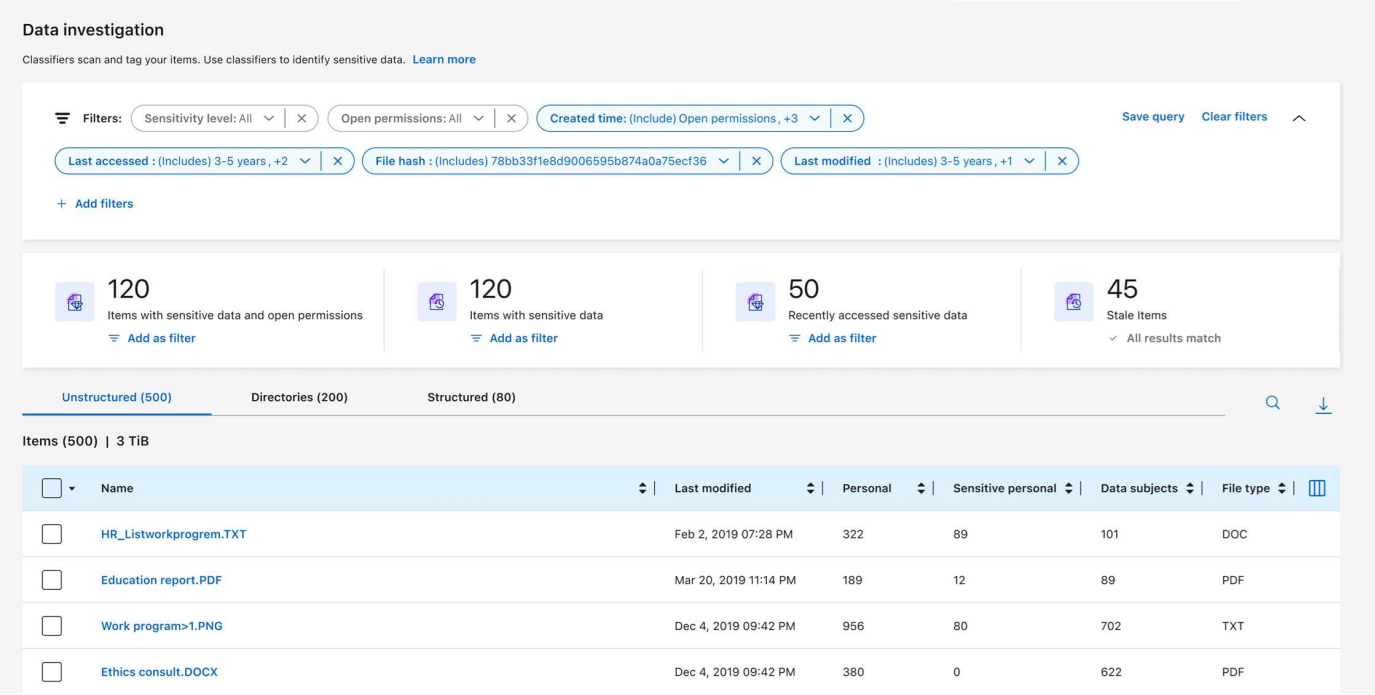
データ調査ページでは、データが 3 つのタブに分類されます。

- 非構造化データ: ファイルデータ
- ディレクトリ: フォルダとファイル共有
- 構造化: データベース

データフィルター

データ調査ページには、必要なものだけを見つけるためにデータを並べ替えるためのさまざまなフィルターが用意されています。複数のフィルターを組み合わせて使うこともできます。

フィルターを追加するには、[フィルターを追加] ボタンを選択します。



フィルターの感度とコンテンツ

次のフィルターを使用して、データに含まれる機密情報の量を確認します。

フィルタ	詳細
カテゴリ	を選択する"カテゴリーの種類"。
感度レベル	機密レベルを選択します: 個人、機密個人、または非機密。
識別子の数	ファイルごとに検出された機密識別子の範囲を選択します。個人データおよび機密性の高い個人データが含まれます。ディレクトリでフィルタリングする場合、データ分類は各フォルダー (およびサブフォルダー) 内のすべてのファイルからの一致を合計します。注: 2023 年 12 月 (バージョン 1.26.6) のリリースでは、ディレクトリ別に個人識別情報 (PII) データの数を計算するオプションが削除されました。
個人データ	を選択する"個人データの種類"。
機密性の高い個人データ	を選択する"機密個人データの種類"。
データ主体	データ主体のフルネームまたは既知の識別子を入力します。"データ主体についての詳細はこちらをご覧ください" 。

ユーザー所有者とユーザー権限をフィルタリング

次のフィルターを使用して、ファイルの所有者とデータへのアクセス許可を表示します。

フィルタ	詳細
オープン権限	データ内およびフォルダー/共有内の権限の種類を選択します。
ユーザー/グループの権限	1 つまたは複数のユーザー名またはグループ名を選択するか、名前の一部を入力します。
ファイル所有者	ファイルの所有者名を入力します。

フィルタ	詳細
アクセス権を持つユーザーの数	1 つまたは複数のカテゴリ範囲を選択して、特定の数のユーザーに公開されているファイルとフォルダーを表示します。

時系列でフィルタリング

時間基準に基づいてデータを表示するには、次のフィルターを使用します。

フィルタ	詳細
作成時刻	ファイルが作成された時間の範囲を選択します。検索結果をさらに絞り込むために、カスタムの時間範囲を指定することもできます。
発見された時間	データ分類がファイルを検出した時間範囲を選択します。検索結果をさらに絞り込むために、カスタムの時間範囲を指定することもできます。
最終変更日時	ファイルが最後に変更された時間の範囲を選択します。検索結果をさらに絞り込むために、カスタムの時間範囲を指定することもできます。
最終アクセス日時	ファイルまたはディレクトリ* が最後にアクセスされた時間の範囲を選択します。検索結果をさらに絞り込むために、カスタムの時間範囲を指定することもできます。データ分類がスキャンするファイルの種類の場合、これはデータ分類がファイルを最後にスキャンした時刻です。

{アスタリスク} ディレクトリの最終アクセス時刻は、NFS または CIFS 共有でのみ使用できます。

メタデータをフィルタリング

場所、サイズ、ディレクトリまたはファイルの種類に基づいてデータを表示するには、次のフィルターを使用します。

フィルタ	詳細
ファイルパス	クエリに含めるか除外する部分パスまたは完全パスを最大 20 個入力します。含めるパスと除外するパスの両方を入力すると、データ分類はまず含めるパス内のすべてのファイルを見つけ、次に除外するパスからファイルを削除して、結果を表示します。このフィルターで「*」を使用しても効果はなく、特定のフォルダーをスキャンから除外することはできないことに注意してください。構成された共有の下にあるすべてのディレクトリとファイルがスキャンされます。
ディレクトリタイプ	ディレクトリの種類として「共有」または「フォルダー」を選択します。
ファイル タイプ	を選択する"ファイルの種類"。
ファイル サイズ	ファイルサイズの範囲を選択します。
ファイルハッシュ	名前が異なっていても、特定のファイルを見つけるには、ファイルのハッシュを入力します。

フィルター収納タイプ

ストレージ タイプ別にデータを表示するには、次のフィルターを使用します。

フィルタ	詳細
システムタイプ	システムの種類を選択します。
システム環境名	特定のシステムを選択します。
ストレージリポジトリ	ボリュームやスキーマなどのストレージ リポジトリを選択します。

フィルタークエリ

保存されたクエリ別にデータを表示するには、次のフィルターを使用します。

フィルタ	詳細
保存されたクエリ	保存したクエリを 1 つまたは複数選択します。に行く"保存されたクエリタブ"既存の保存済みクエリのリストを表示し、新しいクエリを作成します。
タグ	選択"タグ"ファイルに割り当てられているもの。

フィルター分析ステータス

次のフィルターを使用して、データ分類スキャンのステータス別にデータを表示します。

フィルタ	詳細
分析ステータス	最初のスキャンが保留中、スキャンが完了、再スキャンが保留中、またはスキャンに失敗したファイルのリストを表示するには、オプションを選択します。
スキャン分析イベント	データ分類が最終アクセス時間を元に戻すことができなかったために分類されなかったファイルを表示するか、データ分類が最終アクセス時間を元に戻すことができなかったにもかかわらず分類されたファイルを表示するかを選択します。

"「最終アクセス時刻」のタイムスタンプの詳細を見る"スキャン分析イベントを使用してフィルタリングするときに調査ページに表示される項目の詳細については、こちらをご覧ください。

重複データによるフィルタリング

ストレージ内に重複しているファイルを表示するには、次のフィルターを使用します。

フィルタ	詳細
重複	リポジトリ内でファイルを複製するかどうかを選択します。

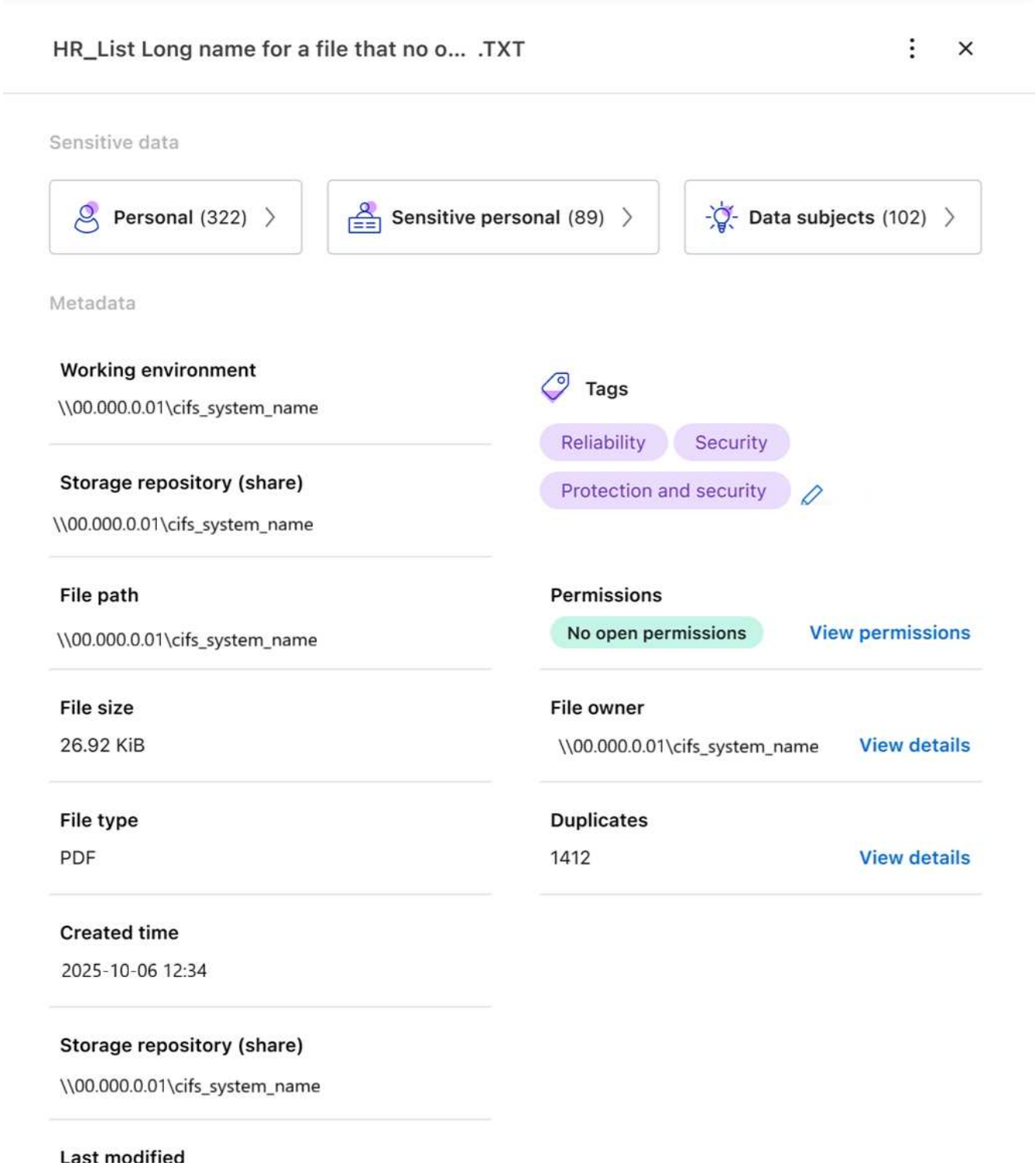
ファイルのメタデータを表示

メタデータには、ファイルが存在するシステムとボリュームが表示されるだけでなく、ファイルの権限、ファイルの所有者、このファイルの重複があるかどうかなど、さらに多くの情報が表示されます。この情報は、"保存したクエリを作成する"データをフィルタリングするために使用できるすべての情報を確認できるためです。

情報の可用性はデータ ソースによって異なります。たとえば、データベース ファイルのボリューム名とアクセス許可は共有されません。

手順

1. データ分類メニューから、*調査*を選択します。
2. 右側のデータ調査リストで、下向き矢印を選択します。ファイルのメタデータを表示するには、任意のファイルの右側にある をクリックします。



3. オプションで、[タグを作成] ボタンを使用してファイルにタグを作成または追加できます。ドロップダウン メニューから既存のタグを選択するか、[+ 追加] ボタンを使用して新しいタグを追加します。タグを使用してデータをフィルタリングできます。

ファイルとディレクトリのユーザー権限を表示する

ファイルまたはディレクトリへのアクセス権を持つすべてのユーザーまたはグループと、それらの権限の種類のリストを表示するには、[すべての権限を表示] を選択します。このオプションは、CIFS 共有内のデータに対してのみ使用できます。

ユーザー名やグループ名の代わりにセキュリティ識別子 (SID) を使用する場合は、Active Directory をデータ分類に統合する必要があります。詳細については、以下を参照してください。 "データ分類にActive Directoryを追加する" 。

手順

1. データ分類メニューから、*調査*を選択します。
2. 右側のデータ調査リストで、下向き矢印を選択します。 ファイルのメタデータを表示するには、任意のファイルの右側にある をクリックします。
3. ファイルまたはディレクトリへのアクセス権を持つすべてのユーザーまたはグループのリストと、それらの権限の種類を表示するには、[開く権限] フィールドで [すべての権限を表示] を選択します。

   データ分類では、リストに最大 100 人のユーザーが表示されます。

4. 下向き矢印を選択 任意のグループのボタンをクリックすると、そのグループに属しているユーザーのリストが表示されます。

   グループの 1 レベルを展開すると、グループに属しているユーザーを表示できます。

5. ユーザーまたはグループの名前を選択して調査ページを更新し、そのユーザーまたはグループがアクセスできるすべてのファイルとディレクトリを表示できるようにします。

ストレージシステム内の重複ファイルをチェックする

ストレージ システムに重複したファイルが保存されているかどうかを確認できます。これは、ストレージスペースを節約できる領域を特定する場合に便利です。また、特定の権限や機密情報を持つ特定のファイルが、ストレージ システム内で不必要に重複しないようにすることも重要です。

データ分類では、次の場合にすべてのファイル (データベースを除く) の重複を比較します。

- 1 MB以上
- または個人情報や機密性の高い個人情報が含まれている

データ分類では、ハッシュ テクノロジを使用して重複ファイルを判別します。いずれかのファイルに別のファイルと同じハッシュ コードがある場合、ファイル名が異なっていてもファイルは完全に重複しています。

手順

1. データ分類メニューから、*調査*を選択します。
2. フィルター ペインで、「ファイル サイズ」と「重複」(「重複あり」) を選択して、環境内で重複している特定のサイズ範囲のファイルを確認します。
3. オプションで、重複ファイルのリストをダウンロードしてストレージ管理者に送信し、削除できるファイルがあるかどうかを管理者が判断できるようにします。

4. 必要に応じて、重複ファイルを削除、タグ付け、または移動することもできます。アクションを実行するファイルを選択し、適切なアクションを選択します。

特定のファイルが重複しているかどうかを確認する

1 つのファイルに重複があるかどうかを確認できます。

手順

1. データ分類メニューから、*調査*を選択します。
2. データ調査リストで、ファイルのメタデータを表示するには、任意のファイルの右側にある をクリックします。

   ファイルに重複が存在する場合、この情報は [重複] フィールドの横に表示されます。

3. 重複ファイルのリストとその保存場所を表示するには、[詳細の表示] を選択します。
4. 次のページで「重複を表示」を選択し、調査ページでファイルを表示します。
5. 必要に応じて、重複ファイルを削除、タグ付け、または移動することもできます。アクションを実行するファイルを選択し、適切なアクションを選択します。

> このページで提供されている「ファイル ハッシュ」値を使用して、調査ページに直接入力し、いつでも特定の重複ファイルを検索できます。また、保存したクエリで使用することもできます。

レポートをダウンロードする

フィルタリングされた結果を CSV または JSON 形式でダウンロードできます。

データ分類がファイル (非構造化データ)、ディレクトリ (フォルダーとファイル共有)、およびデータベース (構造化データ) をスキャンしている場合、最大 3 つのレポート ファイルをダウンロードできます。

ファイルは、固定数の行またはレコードを持つファイルに分割されます。

- JSON: レポートあたり10万件のレコード。生成には約5分かかります。
- CSV: レポートあたり20万件のレコード、生成に約4分かかります

> このブラウザで表示するには、CSV ファイルのバージョンをダウンロードできます。このバージョンは 10,000 件のレコードに制限されています。

ダウンロード可能なレポートに含まれるもの

*非構造化ファイル データ レポート*には、ファイルに関する次の情報が含まれます。

- ファイル名
- 場所の種類
- システム名
- ストレージリポジトリ（ボリューム、バケット、共有など）
- リポジトリの種類

- ファイル パス
- ファイル タイプ
- ファイルサイズ（MB）
- 作成時間
- 最終更新日
- 最終アクセス
- ファイルの所有者
  - ファイル所有者データには、Active Directory が構成されている場合のアカウント名、SAM アカウント名、および電子メール アドレスが含まれます。
- カテゴリ
- 個人情報
- 機密個人情報
- オープン権限
- スキャン分析エラー
- 削除検出日

  削除検出日は、ファイルが削除または移動された日付を識別します。これにより、機密ファイルが移動された時期を識別できるようになります。削除されたファイルは、ダッシュボードまたは調査ページに表示されるファイル数には含まれません。ファイルは CSV レポートにのみ表示されます。

*非構造化ディレクトリ データ レポート*には、フォルダーとファイル共有に関する次の情報が含まれます。

- システムタイプ
- システム名
- ディレクトリ名
- ストレージリポジトリ（フォルダやファイル共有など）
- ディレクトリ所有者
- 作成時間
- 発見された時間
- 最終更新日
- 最終アクセス
- オープン権限
- ディレクトリタイプ

*構造化データ レポート*には、データベース テーブルに関する次の情報が含まれます。

- DBテーブル名
- 場所の種類
- システム名

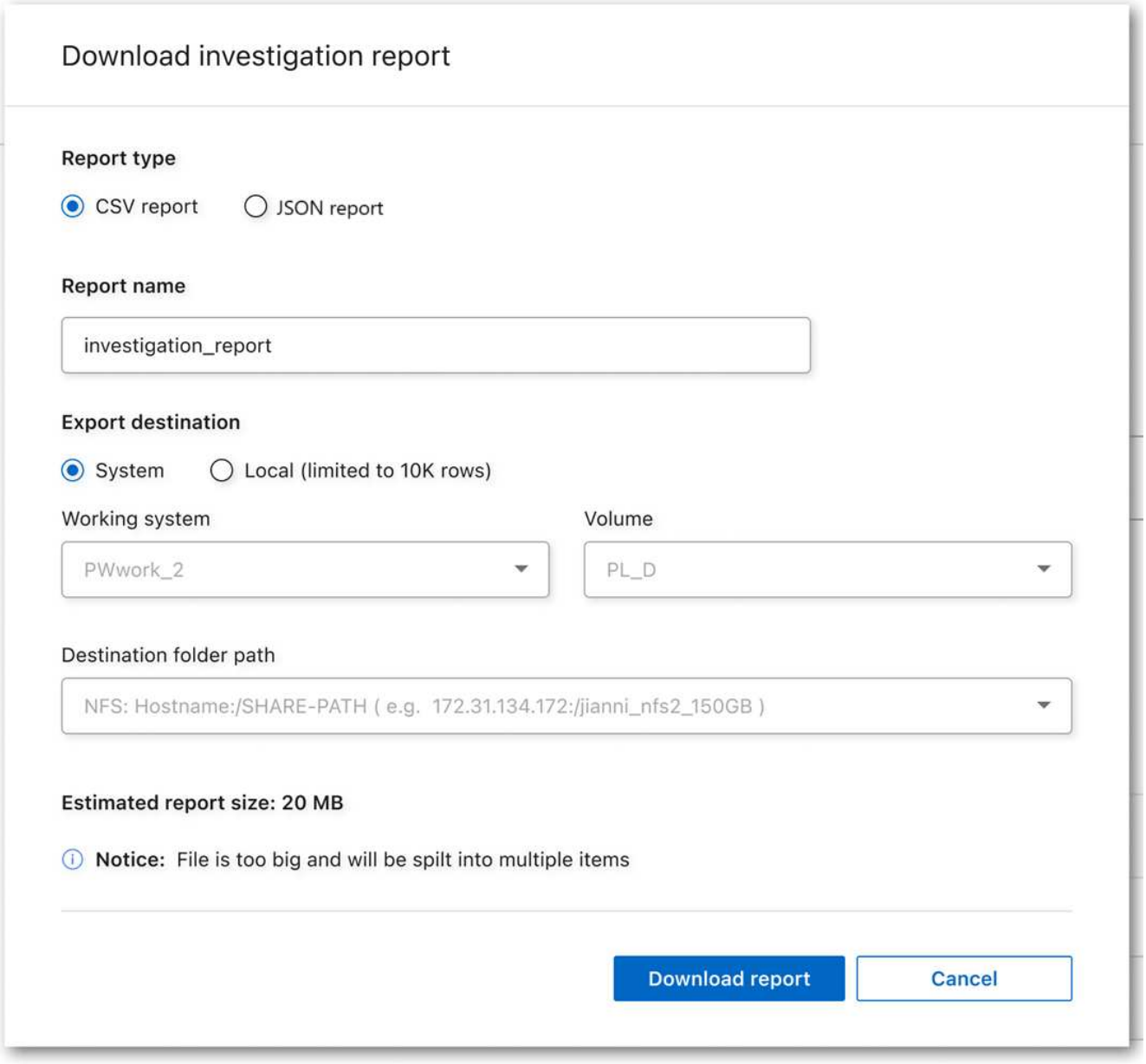
- ストレージリポジトリ（スキーマなど）
- 列数
- 行数
- 個人情報
- 機密個人情報

レポートを生成する手順

1. データ調査ページから、ページの右上にあるボタンをクリックします。
2. レポートタイプ（CSV または JSON）を選択します。
3. レポート名 を入力します。
4. 完全なレポートをダウンロードするには、[システム] を選択し、それぞれのドロップダウン メニューから [システム] と [ボリューム] を選択します。 宛先フォルダーのパス を指定します。

   ブラウザでレポートをダウンロードするには、[ローカル] を選択します。このオプションでは、レポートが最初の 10,000 行に制限され、**CSV** 形式に制限されることに注意してください。 ローカル を選択した場合は、他のフィールドを入力する必要はありません。

5. レポートのダウンロードを選択します。



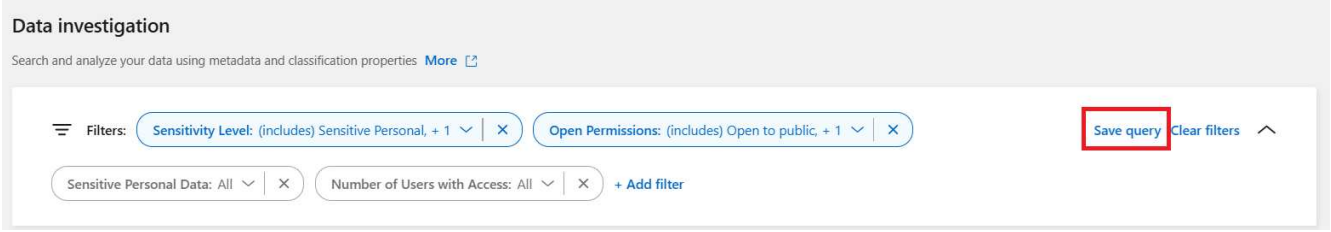
結果

レポートをダウンロード中であることを示すメッセージがダイアログに表示されます。

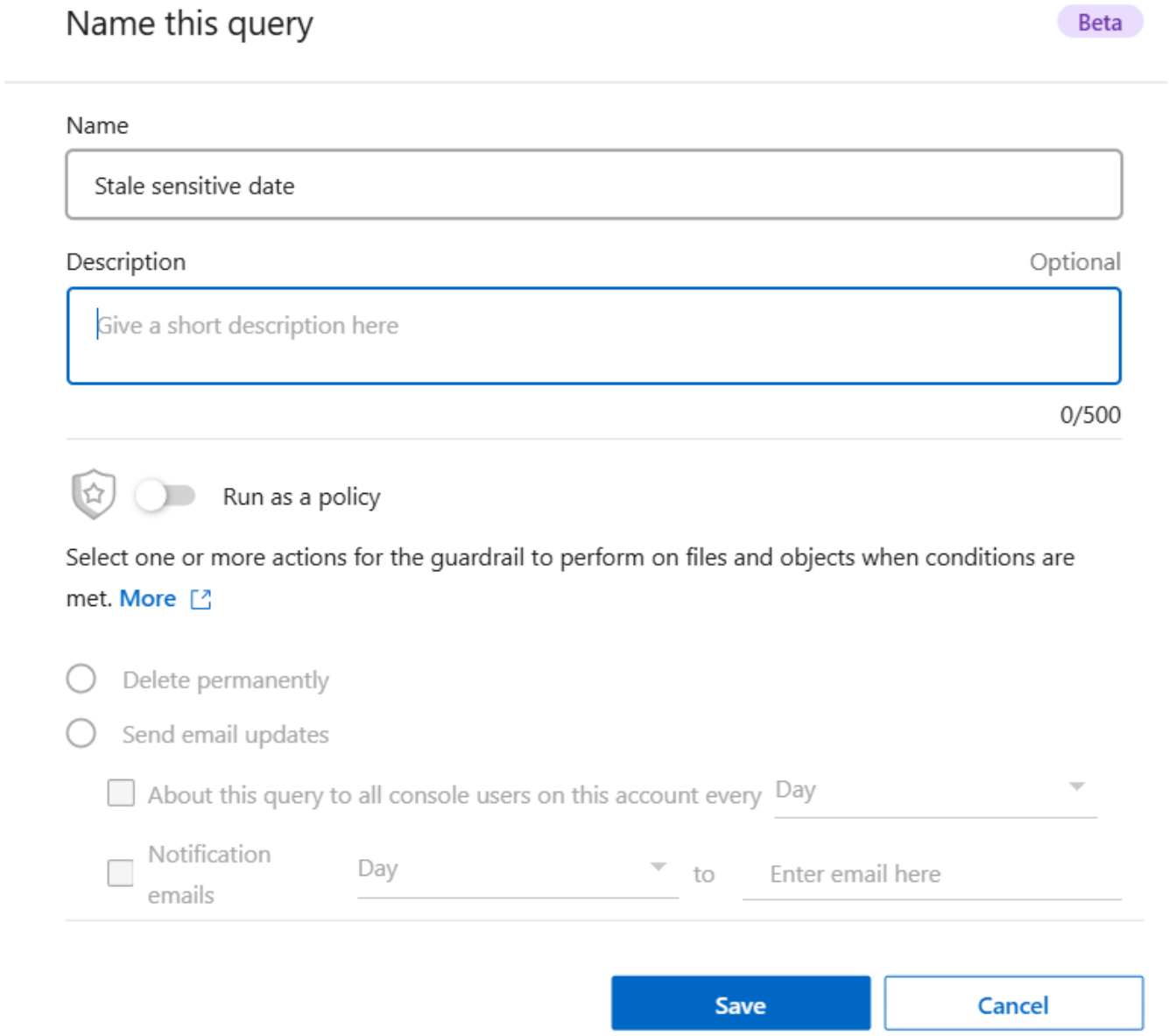
選択したフィルターに基づいて保存されたクエリを作成する

手順

1. 「調査」タブで、使用するフィルターを選択して検索を定義します。見る"調査ページでのデータのフィルタリング"詳細については。
2. すべてのフィルター特性を好みに合わせて設定したら、「クエリを保存」を選択します。



3. 保存したクエリに名前を付け、説明を追加します。名前は一意である必要があります。
4. オプションでクエリをポリシーとして保存できます。
   a. クエリをポリシーとして保存するには、[ポリシーとして実行] トグルを切り替えます。
   b. 完全に削除 または 電子メールで更新を送信 を選択します。電子メールによる更新を選択した場合は、クエリ結果を毎日、毎週、または毎月、すべてのコンソール ユーザーに電子メールで送信できます。あるいは、同じ頻度で特定の電子メール アドレスに通知を送信することもできます。
5. *保存*を選択します。



検索またはポリシーを作成したら、[保存したクエリ] タブで表示できます。

結果が「保存されたクエリ」ページに表示されるまで、最大 15 分かかる場合があります。

NetApp Data Classificationで保存したクエリを管理する

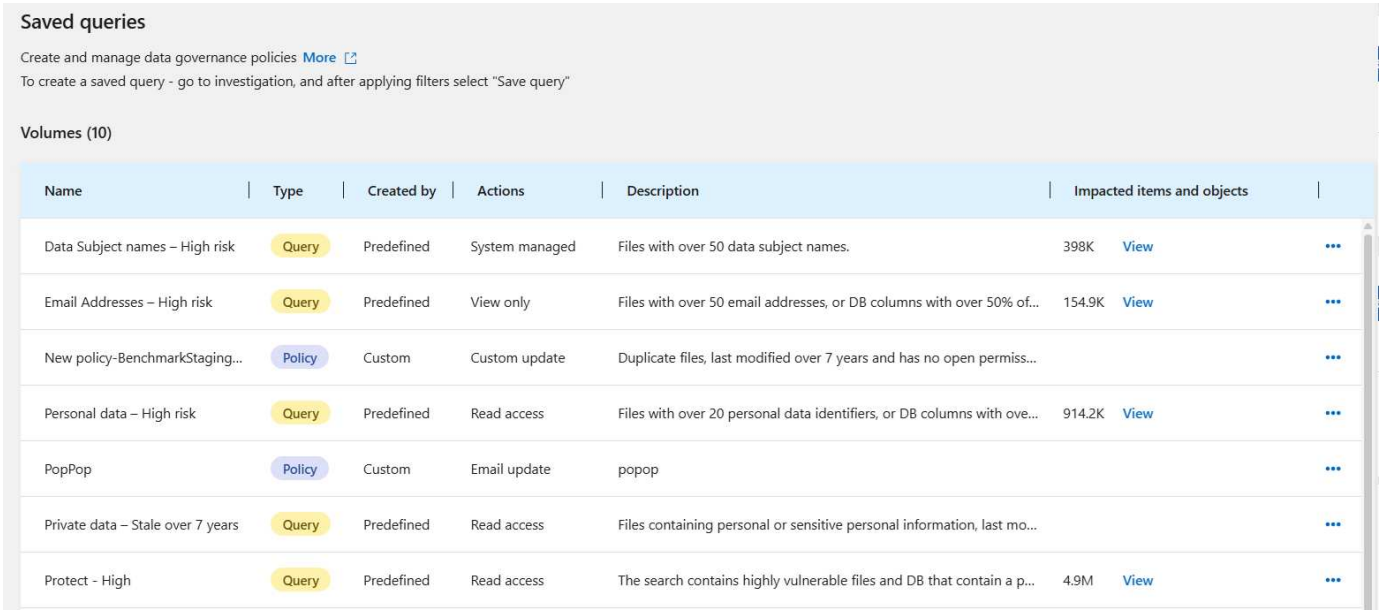
NetaApp Data Classification は、検索クエリの保存をサポートしています。保存したクエリを使用すると、カスタム フィルターを作成して、データ調査ページで頻繁に実行されるクエリを並べ替えることができます。データ分類には、一般的なリクエストに基づいて事前定義された保存済みクエリも含まれます。

コンプライアンス ダッシュボードの 保存済みクエリ タブには、データ分類のこのインスタンスで使用できる

すべての定義済みクエリとカスタム保存済みクエリが一覧表示されます。

保存されたクエリはポリシーとして保存することもできます。クエリはデータをフィルタリングしますが、ポリシーを使用するとデータに基づいて操作を行うことができます。ポリシーを使用すると、検出されたデータを削除したり、検出されたデータに関する電子メールの更新を送信したりできます。

保存されたクエリは、調査ページのフィルターのリストにも表示されます。



調査ページで保存したクエリの結果を表示する

調査ページに保存したクエリの結果を表示するには、特定の検索のボタンをクリックし、[結果の調査] を選択します。

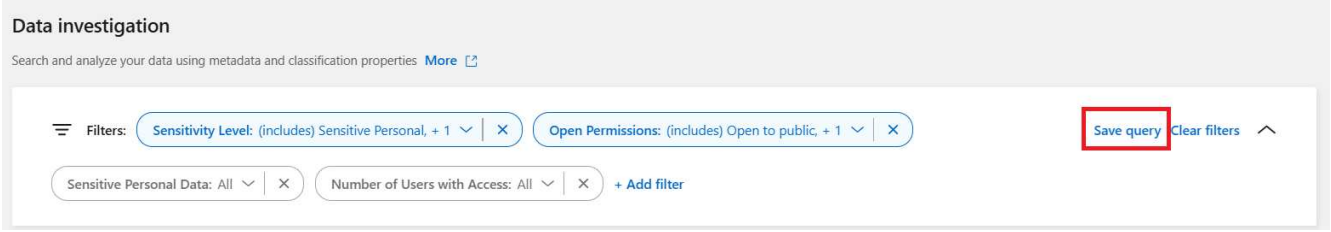
タブから特定の検索の [結果の調査] を選択するスクリーンショット。"]

保存されたクエリとポリシーを作成する

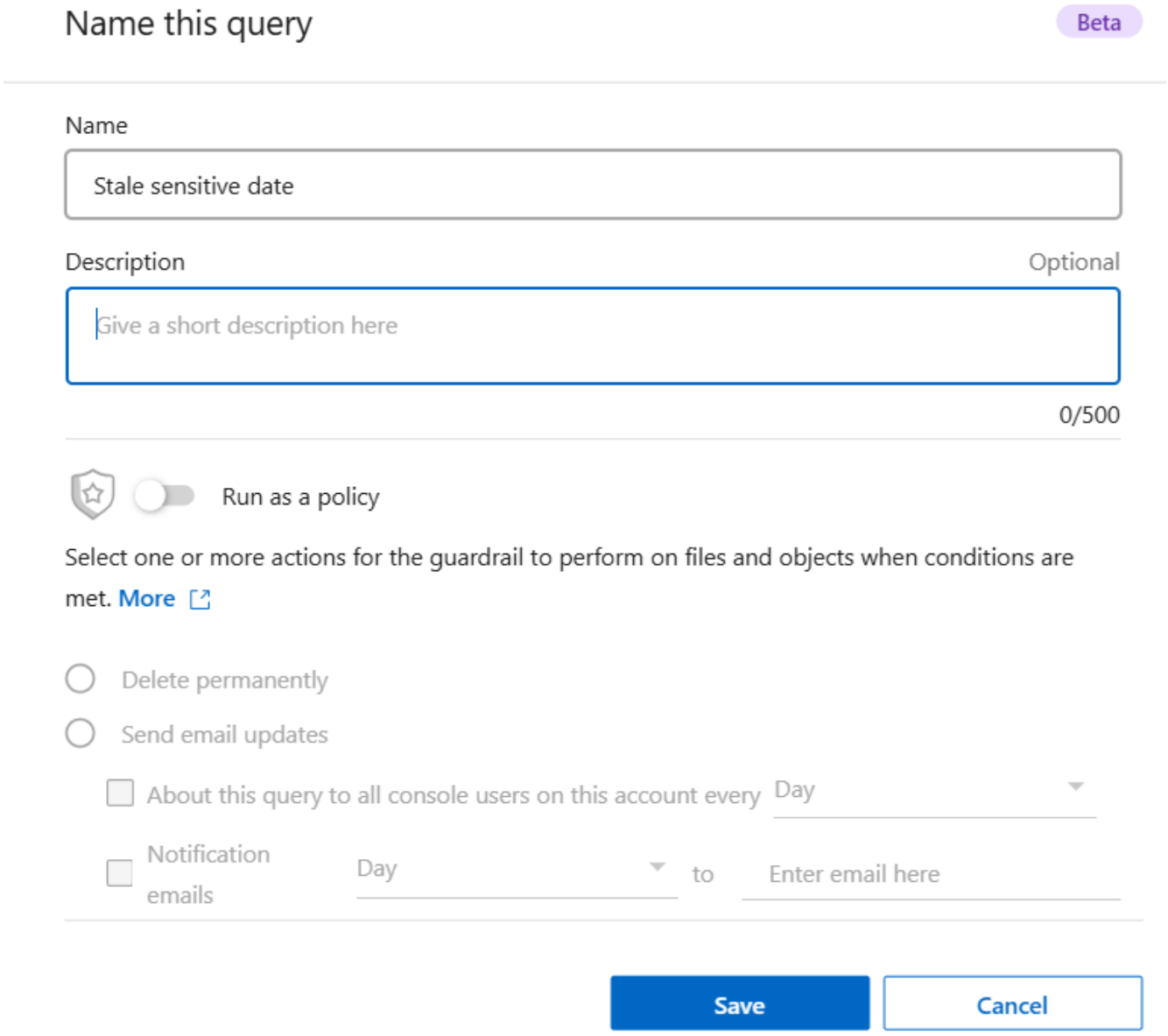
組織固有のクエリの結果を提供する独自のカスタム保存クエリを作成できます。検索条件に一致するすべてのファイルとディレクトリ (共有とフォルダー) の結果が返されます。

手順

1. 「調査」タブで、使用するフィルターを選択して検索を定義します。見る"調査ページでのデータのフィルタリング"詳細については。
2. すべてのフィルター特性を好みに合わせて設定したら、「クエリを保存」を選択します。



3. 保存したクエリに名前を付け、説明を追加します。名前は一意である必要があります。
4. オプションでクエリをポリシーとして保存できます。
   a. クエリをポリシーとして保存するには、[ポリシーとして実行] トグルを切り替えます。
   b. 完全に削除 または 電子メールで更新を送信 を選択します。電子メールによる更新を選択した場合は、クエリ結果を毎日、毎週、または毎月、すべてのコンソール ユーザーに電子メールで送信できます。あるいは、同じ頻度で特定の電子メール アドレスに通知を送信することもできます。
5. *保存*を選択します。



検索またはポリシーを作成したら、[保存したクエリ] タブで表示できます。

保存したクエリまたはポリシーを編集する

保存したクエリの名前と説明を変更できます。クエリをポリシーに変換したり、その逆を行ったりすることもできます。

デフォルトの保存済みクエリを変更することはできません。保存されたクエリのフィルターを変更することはできません。保存したクエリの調査結果を表示したり、フィルターを変更したり、修正したりして、新しいクエリまたはポリシーとして保存することもできます。

手順

1. [保存されたクエリ] ページで、変更する検索の [検索の編集] を選択します。

2. 名前と説明のフィールドを変更します。名前と説明のフィールドのみを変更します。

   オプションで、クエリをポリシーに変換したり、ポリシーを保存されたクエリに変換したりできます。必要に応じて、[ポリシーとして実行] トグルを切り替えます。..クエリをポリシーに変換する場合は、[完全に削除] または [電子メールの更新を送信] を選択します。電子メールによる更新を選択した場合は、クエリ結果を毎日、毎週、または毎月、すべてのコンソール ユーザーに電子メールで送信できます。あるいは、同じ頻度で特定の電子メール アドレスに通知を送信することもできます。

3. 変更を完了するには、[保存] を選択します。

保存したクエリを削除する

不要になった場合は、保存したカスタムクエリまたはポリシーを削除できます。デフォルトで保存されたクエリは削除できません。

保存したクエリを削除するには、 特定の検索のボタンをクリックし、[クエリの削除] を選択してから、確認ダイアログでもう一度 [クエリの削除] を選択します。

デフォルトのクエリ

データ分類では、次のシステム定義の検索クエリが提供されます。

- データ主体名 - 高リスク

  50 人以上のデータ主体名を含むファイル

- メールアドレス - 高リスク

  50 個を超える電子メール アドレスを含むファイル、またはデータベース列の 50% を超える行に電子メール アドレスが含まれているファイル

- 個人データ - 高リスク

  20 個を超える個人データ識別子を含むファイル、またはデータベース列の 50% を超える行に個人データ識別子が含まれているファイル

- 個人データ - 7年以上古い

  個人情報または機密性の高い個人情報を含むファイル（最終更新日が 7 年以上前）

- 保護 - 高

  パスワード、クレジットカード情報、IBAN 番号、社会保障番号を含むファイルまたはデータベースの列

- 保護 - 低

3年以上アクセスされていないファイル

- 保護 - 中

  ID番号、納税者番号、運転免許証番号、医薬品ID、パスポート番号などの個人データ識別子を含むファイルまたはデータベース列を含むファイル

- 機密個人データ - 高リスク

  20 個を超える機密個人データ識別子を持つファイル、または行の 50% 以上に機密個人データが含まれるデータベース列

リポジトリのNetApp Data Classificationスキャン設定を変更する

各システムおよびデータ ソースでデータがスキャンされる方法を管理できます。変更は「リポジトリ」ベースで行うことができます。つまり、スキャンするデータ ソースの種類に応じて、ボリューム、スキーマ、ユーザーなどごとに変更を加えることができます。

変更できる項目としては、リポジトリをスキャンするかどうか、 NetApp Data Classificationが"マッピングスキャンまたはマッピング＆分類スキャン"。また、一定期間ボリュームのスキャンを停止する必要がある場合など、スキャンを一時停止したり再開したりすることもできます。

リポジトリのスキャンステータスを表示する

NetApp Data Classificationがスキャンしている個々のリポジトリ (ボリューム、バケットなど) を、システムおよびデータ ソースごとに表示できます。また、いくつが「マップ」され、いくつが「分類」されたかを確認することもできます。すべてのデータに対して完全な AI 識別が実行されるため、分類には時間がかかります。

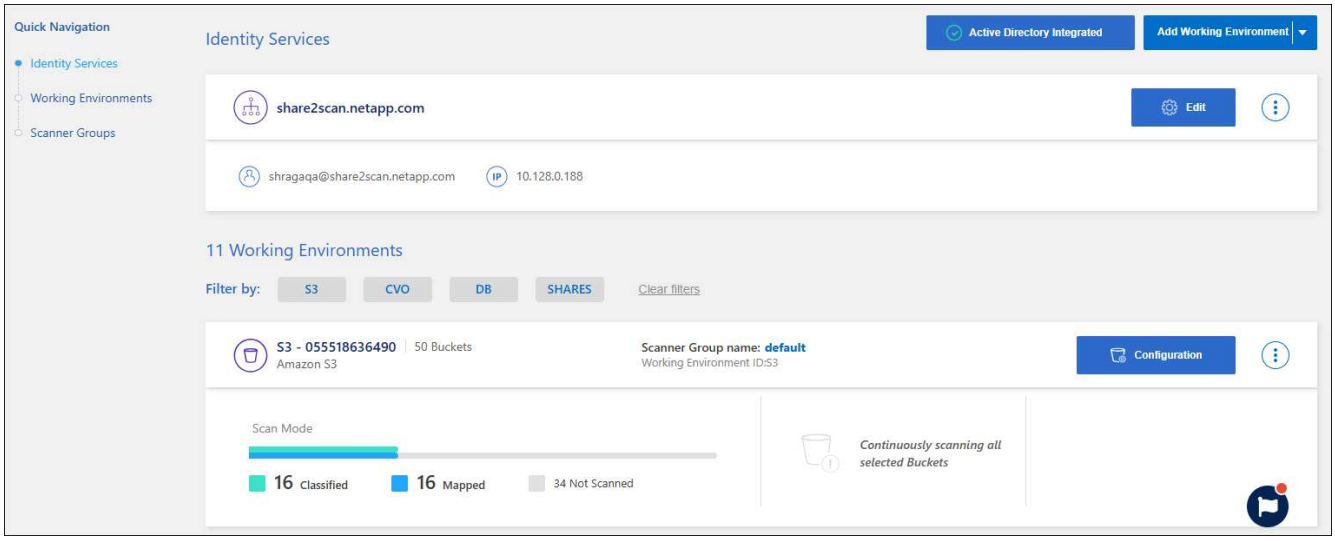
各作業環境のスキャンステータスは、構成ページで確認できます。

- 初期化中 (水色の点): マップまたは分類構成がアクティブ化されています。これは、「保留中のキュー」ステータスに移行する前に短時間表示されます。
- 保留中のキュー (オレンジ色の点): スキャン タスクはスキャン キューにリストされるのを待機しています。
- キューに追加 (オレンジ色の点): タスクがスキャン キューに正常に追加されました。システムは、キュー内の順番が来ると、ボリュームのマッピングまたは分類を開始します。
- 実行中 (緑のドット): キューにあったスキャン タスクが、選択したストレージ リポジトリでアクティブに進行中です。
- 完了 (緑の点): ストレージ リポジトリのスキャンが完了しました。
- 一時停止 (灰色の点): スキャンを一時停止しました。ボリュームの変化はシステムには表示されませんが、スキャンされたインサイトは引き続き利用できます。
- エラー (赤い点): 問題が発生したため、スキャンを完了できません。アクションを完了する必要がある場合は、「必要なアクション」列の下のツールチップにエラーが表示されます。それ以外の場合、システムは「エラー」ステータスを表示し、回復を試みます。終了するとステータスが変わります。

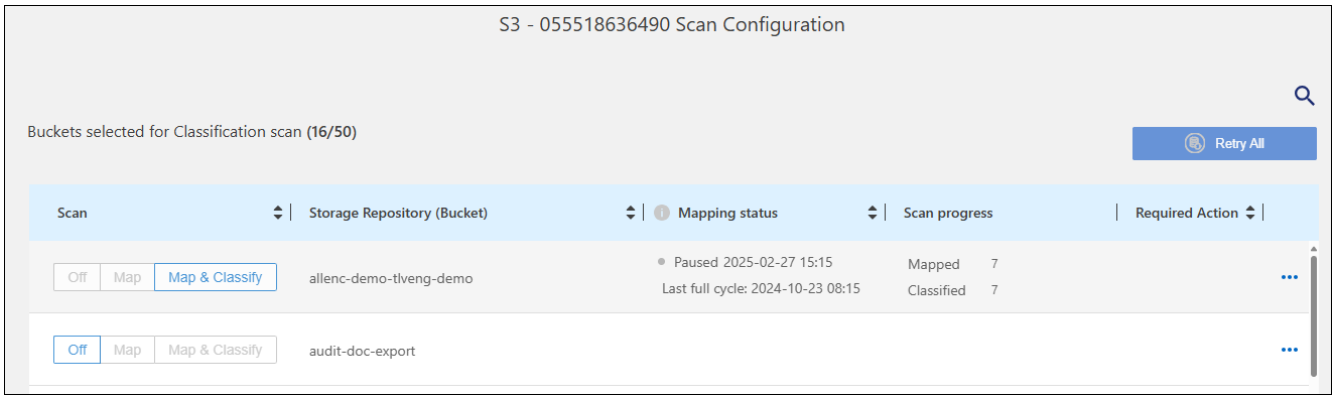
- スキャンしていません: ボリューム構成で「オフ」が選択されており、システムはボリュームをスキャンしていません。

手順

1. データ分類メニューから、*構成*を選択します。



2. [構成] タブから、システムの **[構成]** ボタンを選択します。
3. 「スキャン構成」ページで、すべてのリポジトリのスキャン設定を表示します。



4. スキャン中に、[マッピング ステータス] 列の進行状況バーにカーソルを合わせると、そのリポジトリにマッピングまたは分類されるキュー内のファイルの数が表示されます。

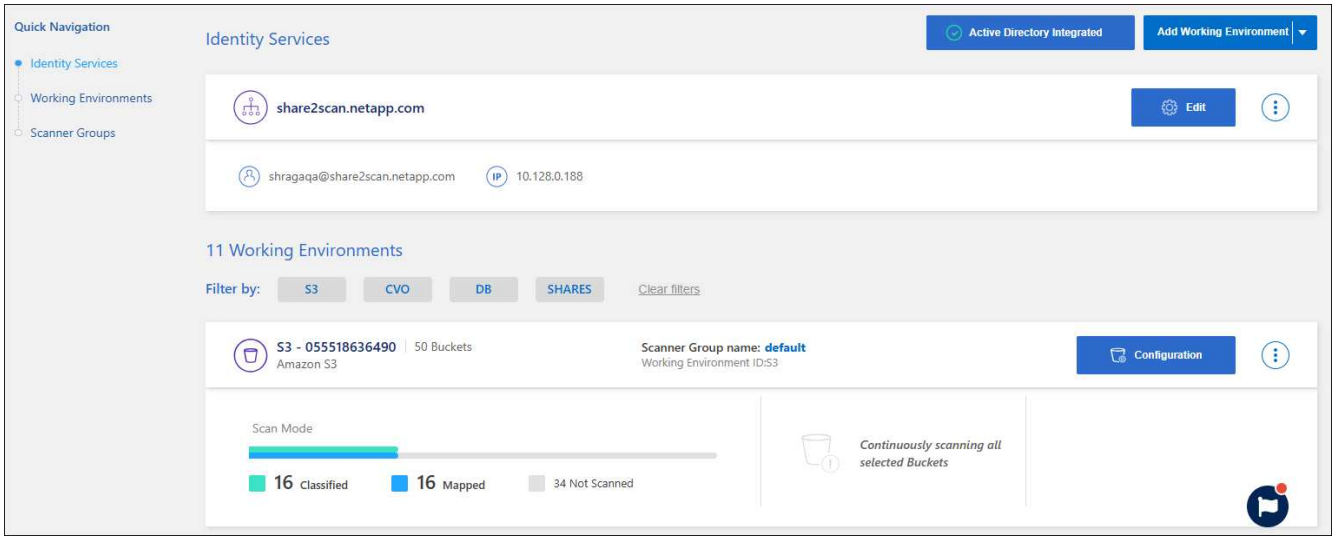
リポジトリのスキャンの種類を変更する

システム内のマッピングのみのスキャン、またはマッピングと分類のスキャンは、いつでも [構成] ページから開始または停止できます。マッピングのみのスキャンからマッピングと分類のスキャンに、またはその逆に変更することもできます。

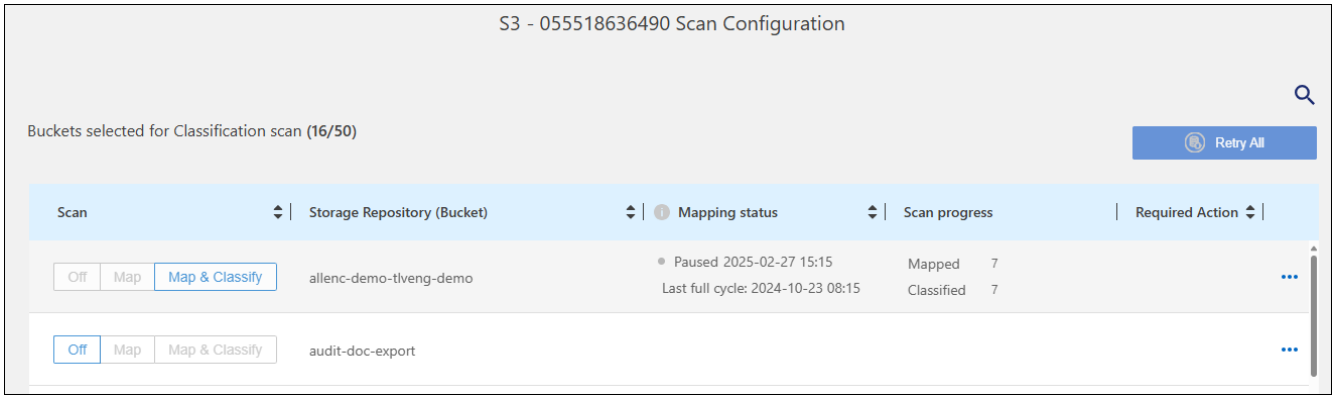
データベースをマッピングのみのスキャンに設定することはできません。データベース スキャンはオフまたはオンにすることができます。オンはマップと分類と同じです。

手順

1. データ分類メニューから、*構成*を選択します。
2. [構成] タブから、システムの [構成] ボタンを選択します。

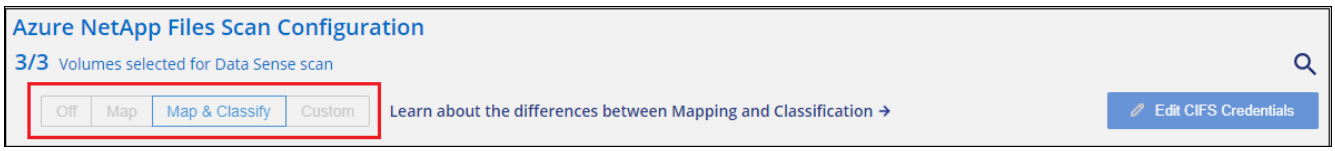


3. [スキャン構成] ページで、いずれかのリポジトリ (この例ではバケット) を変更して、マップ スキャンまたは マップと分類 スキャンを実行します。



特定の種類のシステムでは、ページ上部のボタン バーを使用して、すべてのリポジトリのスキャンの種類をグローバルに変更できます。これは、 Cloud Volumes ONTAP、オンプレミスのONTAP、 Azure NetApp Files、およびAmazon FSx for ONTAPシステムに有効です。

以下の例は、Azure NetApp Filesシステムのこのボタン バーを示しています。



スキャンを優先する

最も重要なマッピングのみのスキャンまたはマッピングと分類のスキャンを優先して、優先度の高いスキャンが最初に完了するようにすることができます。

デフォルトでは、スキャンは開始された順序に基づいてキューに入れられます。スキャンを優先順位付けする

機能を使用すると、スキャンをキューの先頭に移動できます。複数のスキャンを優先できます。優先順位は先入先出順で指定されます。つまり、最初に優先順位を指定したスキャンがキューの先頭に移動し、2 番目に優先順位を指定したスキャンはキューの 2 番目になり、以下同様に続きます。

優先権は 1 回限り付与されます。マッピング データの自動再スキャンはデフォルトの順序で実行されます。

手順

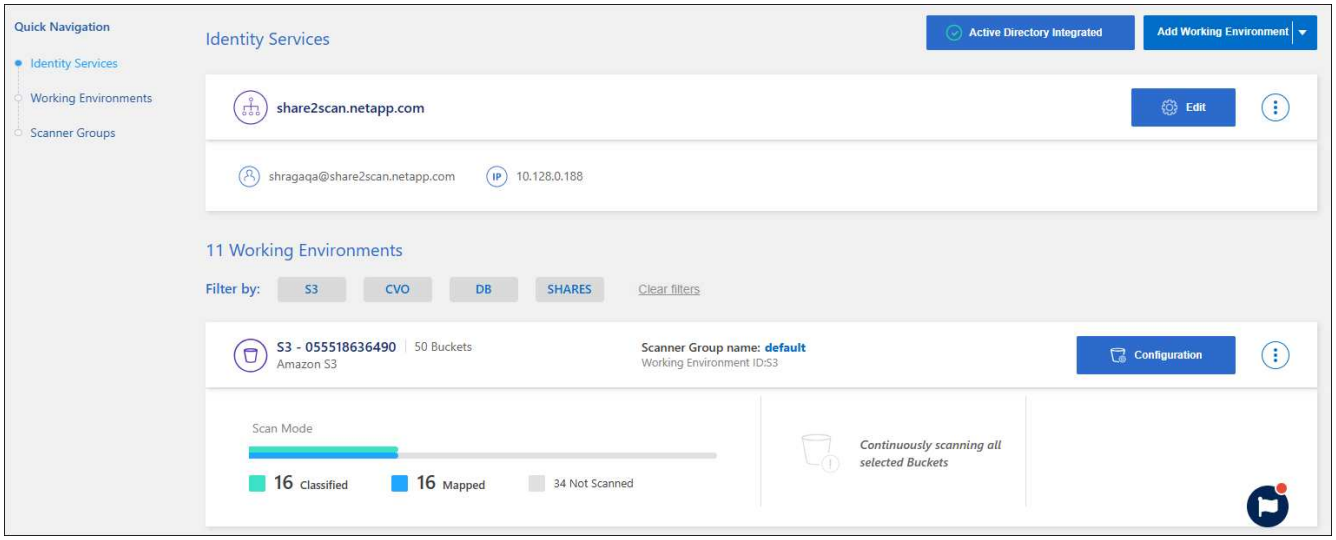
1. データ分類メニューから、*構成*を選択します。
2. 優先したいリソースを選択します。
3. 行動から `...`オプションで、[スキャンを優先] を選択します。

リポジトリのスキャンを停止する

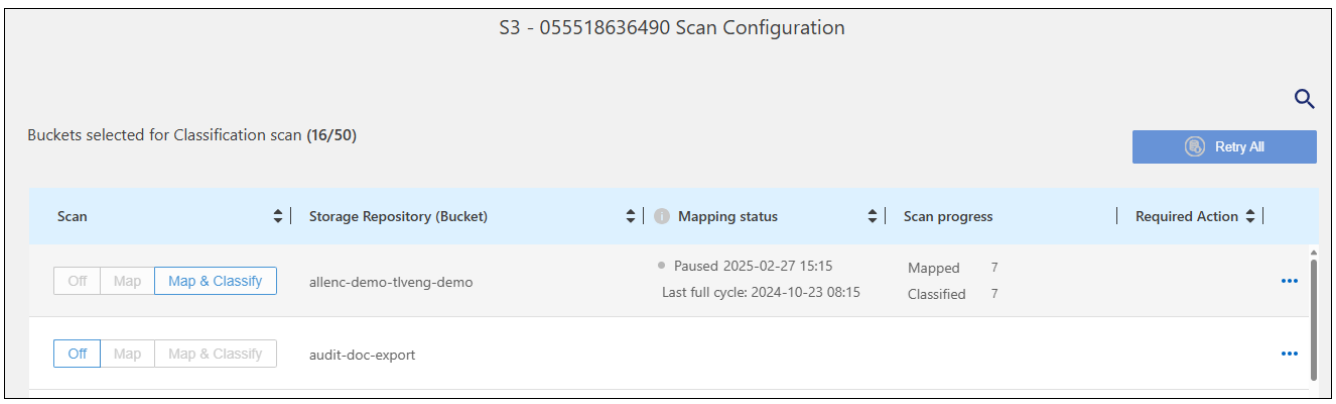
コンプライアンスを監視する必要がなくなった場合は、リポジトリ (ボリュームなど) のスキャンを停止できます。これを行うには、スキャンを「オフ」にします。スキャンをオフにすると、そのボリュームに関するすべてのインデックスと情報がシステムから削除され、データのスキャンに対する課金が停止されます。

手順

1. データ分類メニューから、*構成*を選択します。
2. [構成] タブから、システムの **[構成]** ボタンを選択します。



3. スキャン構成ページで オフ を選択して、特定のバケットのスキャンを停止します。



リポジトリのスキャンを一時停止して再開する

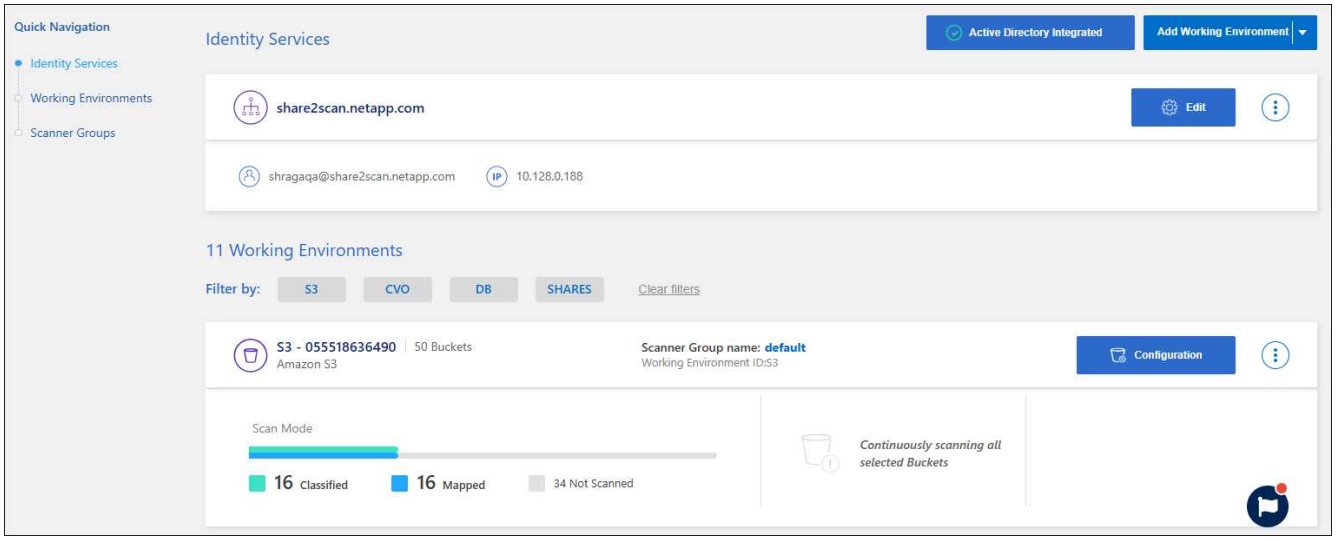
特定のコンテンツのスキャンを一時的に停止したい場合は、リポジトリのスキャンを「一時停止」することができます。スキャンを一時停止すると、データ分類はリポジトリの変更や追加について今後スキャンを実行しません。現在のすべてのスキャン結果は、データ分類で引き続きアクセスできます。

スキャンを一時停止しても、データはまだシステム上に残っているため、課金は発生しません。

いつでもスキャンを再開できます。

手順

1. データ分類メニューから、*構成*を選択します。
2. [構成] タブから、システムの [構成] ボタンを選択します。



3. スキャン設定ページで、アクションを選択します。 ••• アイコン。
4. ボリュームのスキャンを一時停止するには「一時停止」を選択し、一時停止していたボリュームのスキャンを再開するには「再開」を選択します。

NetApp Data Classificationコンプライアンスレポートを表示

NetApp Data Classification は、組織のデータ プライバシー プログラムの状態をより深く理解するために使用できるレポートを提供します。

デフォルトでは、データ分類ダッシュボードには、すべてのシステム、データベース、およびデータ ソースのコンプライアンスとガバナンスのデータが表示されます。一部のシステムのデータのみを含むレポートを表示する場合は、フィルターしてそれらのシステムのみを表示できます。

- コンプライアンス レポートは、データ ソースに対して完全な分類スキャンを実行した場合にのみ利用できます。マッピングのみのスキャンが行われたデータ ソースでは、データ マッピング レポートのみを生成できます。
- NetApp は、データ分類によって識別される個人データおよび機密個人データの 100% の正確性を保証することはできません。常にデータを確認して情報を検証する必要があります。

データ分類では次のレポートが利用可能です。

- データ検出評価レポート: スキャンされた環境の高レベルの分析を提供し、システムの検出結果を強調し、懸念事項と潜在的な修復手順を示します。このレポートはガバナンス ダッシュボードで利用できます。
- 完全なデータ マッピングの概要レポート: システム内のファイルのサイズと数に関する情報を提供します。これには、使用容量、データの古さ、データのサイズ、ファイルの種類が含まれます。このレポートはガバナンス ダッシュボードで利用できます。
- データ主体アクセス要求レポート: データ主体の特定の名前または個人識別子に関する情報が含まれるすべてのファイルのレポートを抽出できます。このレポートはコンプライアンス ダッシュボードで利用できます。
- **HIPAA** レポート: ファイル全体にわたる健康情報の分布を識別するのに役立ちます。このレポートはコンプライアンス ダッシュボードで利用できます。
- **PCI DSS** レポート: ファイル全体にわたるクレジットカード情報の分布を識別するのに役立ちます。このレポートはコンプライアンス ダッシュボードで利用できます。
- プライバシー リスク評価レポート: データから得られたプライバシーの分析情報とプライバシー リスク スコアを提供します。このレポートはコンプライアンス ダッシュボードで利用できます。
- 特定の情報タイプに関するレポート: 個人データや機密性の高い個人データを含む、特定されたファイルに関する詳細を含むレポートが利用可能です。カテゴリやファイルタイプ別に分類されたファイルを表示することもできます。

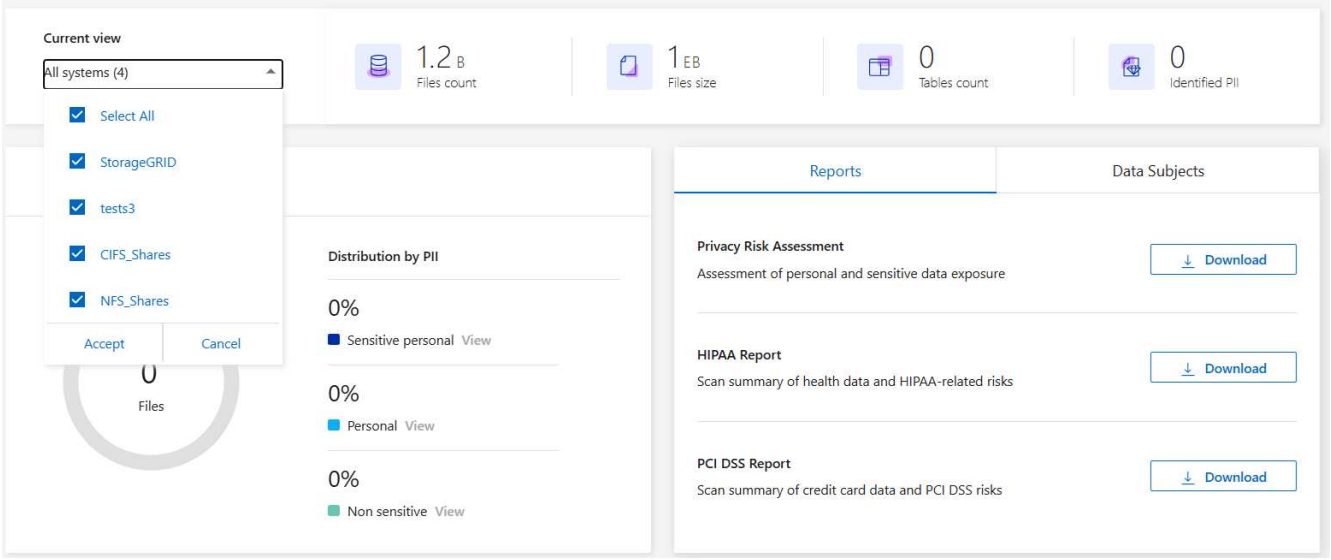
レポートのシステムを選択する

データ分類コンプライアンス ダッシュボードの内容をフィルタリングして、すべてのシステムとデータベース、または特定のシステムのコンプライアンス データを表示できます。

ダッシュボードをフィルターすると、データ分類によってコンプライアンス データの範囲が限定され、選択したシステムのみがレポートされます。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. システム フィルターのドロップダウンを選択し、システムを選択します。
3. 選択内容を確認するには、[承認] を選択します。



データ主体アクセス要求レポート

欧州の GDPR などのプライバシー規制では、データ主体 (顧客や従業員など) に個人データにアクセスする権利が付与されます。データ主体がこの情報を要求する場合、これは DSAR (データ主体アクセス要求) と呼ばれます。組織はこれらの要求に対して「不当な遅延なく」、遅くとも受領後 1 か月以内に応答する必要があります。

DSAR に応答するには、対象のフルネームまたは既知の識別子 (電子メール アドレスなど) を検索し、レポートをダウンロードします。このレポートは、GDPR または同様のデータ プライバシー法に準拠するという組織の要件を支援するために設計されています。

データ分類は **DSAR** への対応にどのように役立ちますか**?**

データ主体の検索を実行すると、データ分類によって、その人物の名前または識別子が含まれるすべてのファイルが検索されます。データ分類では、事前にインデックス付けされた最新のデータで名前または識別子を確認します。新しいスキャンは開始されません。

検索が完了したら、データ主体アクセス要求レポートのファイル リストをダウンロードできます。レポートはデータから得られた洞察を集約し、それを法的用語にまとめて相手に返送できるようにします。

現在、データベース内ではデータ主体の検索はサポートされていません。

データ主体を検索し、レポートをダウンロード

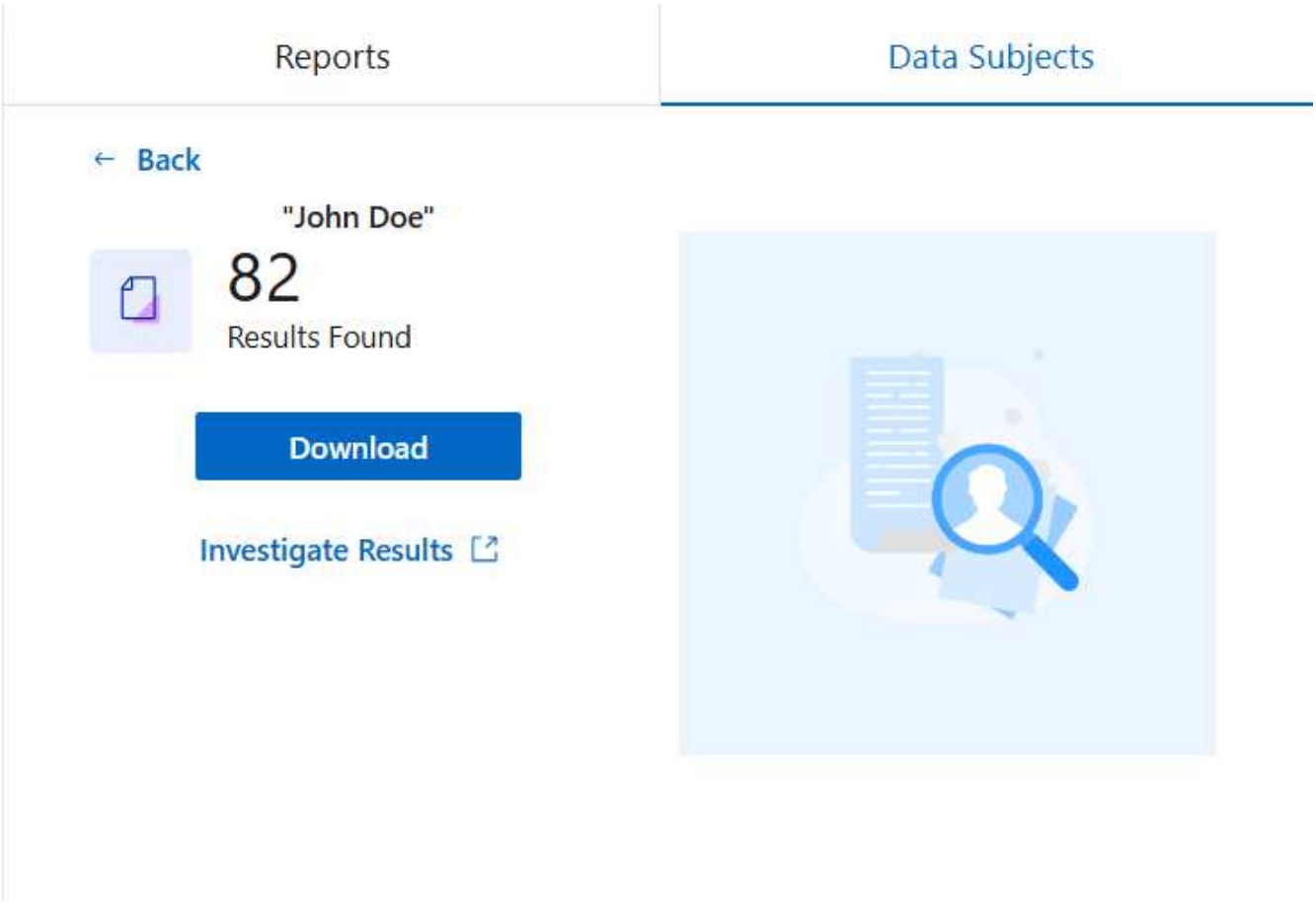
データ主体のフルネームまたは既知の識別子を検索し、ファイル リスト レポートまたは DSAR レポートをダウンロードします。検索条件"あらゆる個人情報の種類"。

データ主体の名前を検索する際には、英語、ドイツ語、日本語、スペイン語がサポートされています。今後、さらに多くの言語のサポートが追加される予定です。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. コンプライアンス ページで、データ主体 タブを見つけます。

3. *データ主体*セクションで、名前または既知の識別子を入力し、検索を選択します。
4. 検索が完了したら、[ダウンロード] を選択して、データ主体のアクセス要求応答にアクセスします。データ調査ページで詳細情報を表示するには、結果の調査 を選択しま
す。



5. データ分類で結果を確認するか、ダウンロード アイコンを選択してレポートとしてダウンロードします。

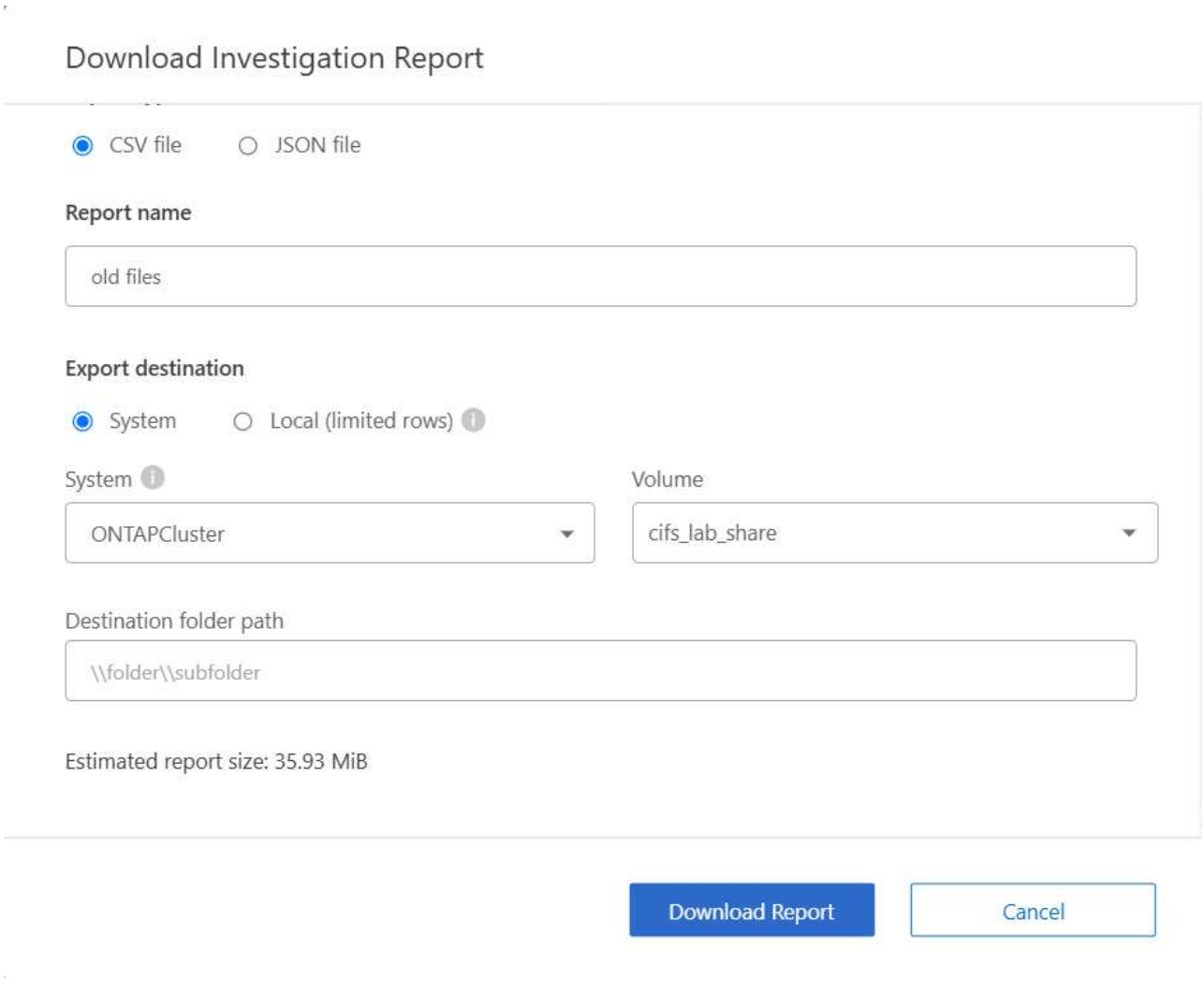
   a. ダウンロード アイコンを選択したら、ダウンロード設定を構成します。

      - フィルムフォーマットを選択してください: CSV または JSON
      - *レポート名*を入力してください
      - エクスポート先として「システム」または「ローカル」マシンを選択します。

        システムを選択した場合は、すべてのデータがダウンロードされます。 システム、ボリューム、*宛先フォルダーのパス*も選択する必要があります。

   ローカル を選択した場合、レポートは非構造化データの最初の 10,000 行、非構造化データの 5,000 行、および構造化データの 1,000 行に制限されます。

   a. ダウンロードを開始するには、[レポートのダウンロード] を選択しま
   す。



医療保険の携行性と責任に関する法律（HIPAA）に関する報告書

医療保険の携行性と責任に関する法律 (HIPAA) レポートは、健康情報を含むファイルを識別するのに役立ちます。これは、組織の HIPAA データ プライバシー法の遵守要件を支援するために設計されています。データ分類が探す情報には次のものが含まれます。

- 健康参照パターン
- ICD-10-CM医療コード
- ICD-9-CM医療コード
- HR - 健康カテゴリ
- 健康アプリケーションデータカテゴリ

レポートには次の情報が含まれます。

- 概要: 健康情報を含むファイルの数と、そのシステム。
- 暗号化: 暗号化されたシステムまたは暗号化されていないシステム上にある健康情報を含むファイルの割合。この情報はCloud Volumes ONTAPに固有のものです。
- ランサムウェア保護: ランサムウェア保護が有効になっているシステム、または有効になっていないシステム上にある健康情報を含むファイルの割合。この情報はCloud Volumes ONTAPに固有のものです。

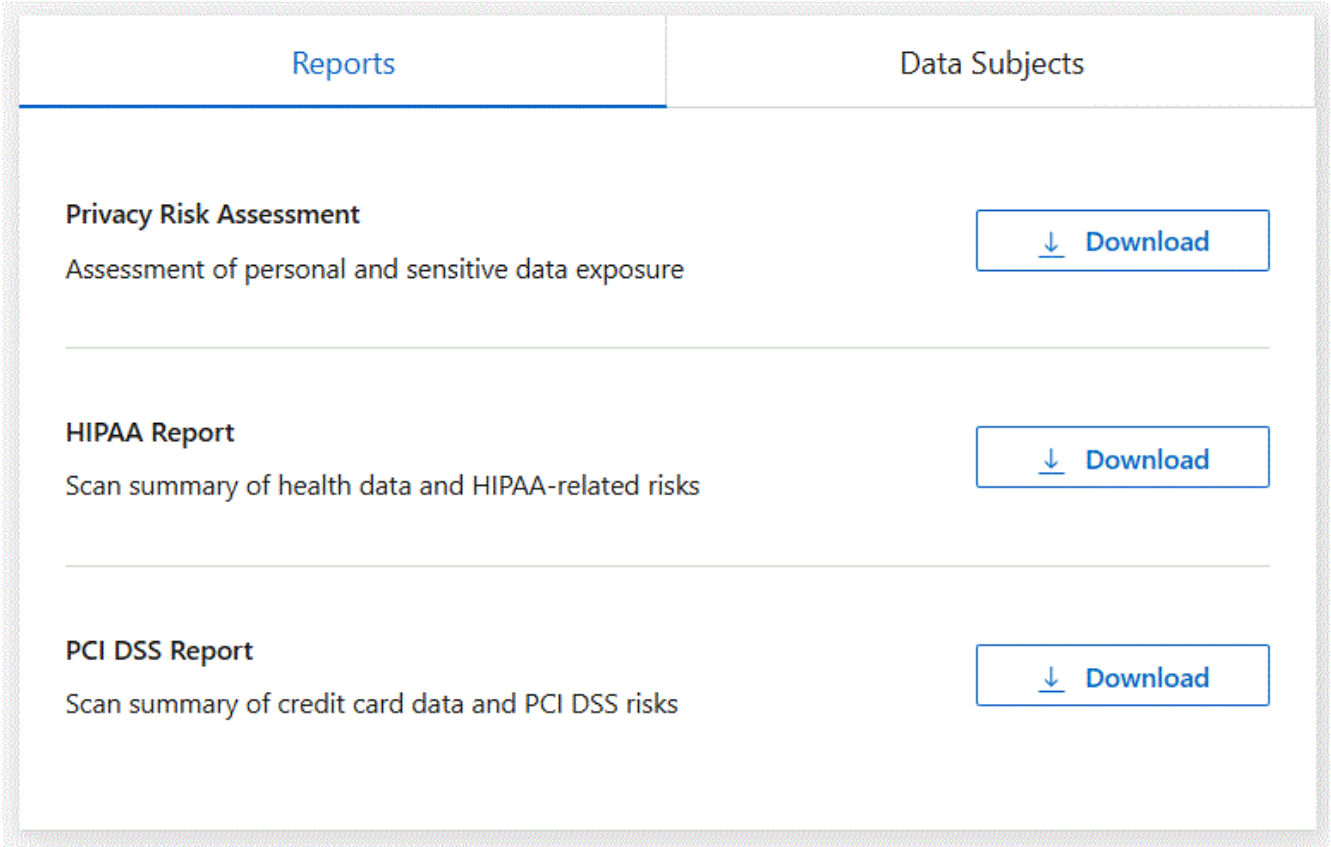
- 保持期間: ファイルが最後に変更された期間。これは、健康情報を処理に必要な期間以上保存すべきではないため、役立ちます。
- 健康情報の配布: 健康情報が見つかったシステムと、暗号化およびランサムウェア保護が有効になっているかどうか。

HIPAAレポートを生成する

レポートを生成するには、「コンプライアンス」タブに移動します。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. レポート ペイン を見つけます。 **HIPAA** レポート の横にあるダウンロード アイコンを選択します。



結果

データ分類により PDF レポートが生成されます。

ペイメントカード業界データセキュリティ基準（**PCI DSS**）レポート

ペイメント カード 業界データ セキュリティ標準 (PCI DSS) レポートは、ファイル全体にわたるクレジットカード情報の分布を識別するのに役立ちます。

レポートには次の情報が含まれます。

- 概要: クレジットカード情報が含まれているファイルの数と、そのシステム。
- 暗号化: 暗号化されたシステムまたは暗号化されていないシステム上にあるクレジットカード情報を含む

ファイルの割合。この情報はCloud Volumes ONTAPに固有のものです。

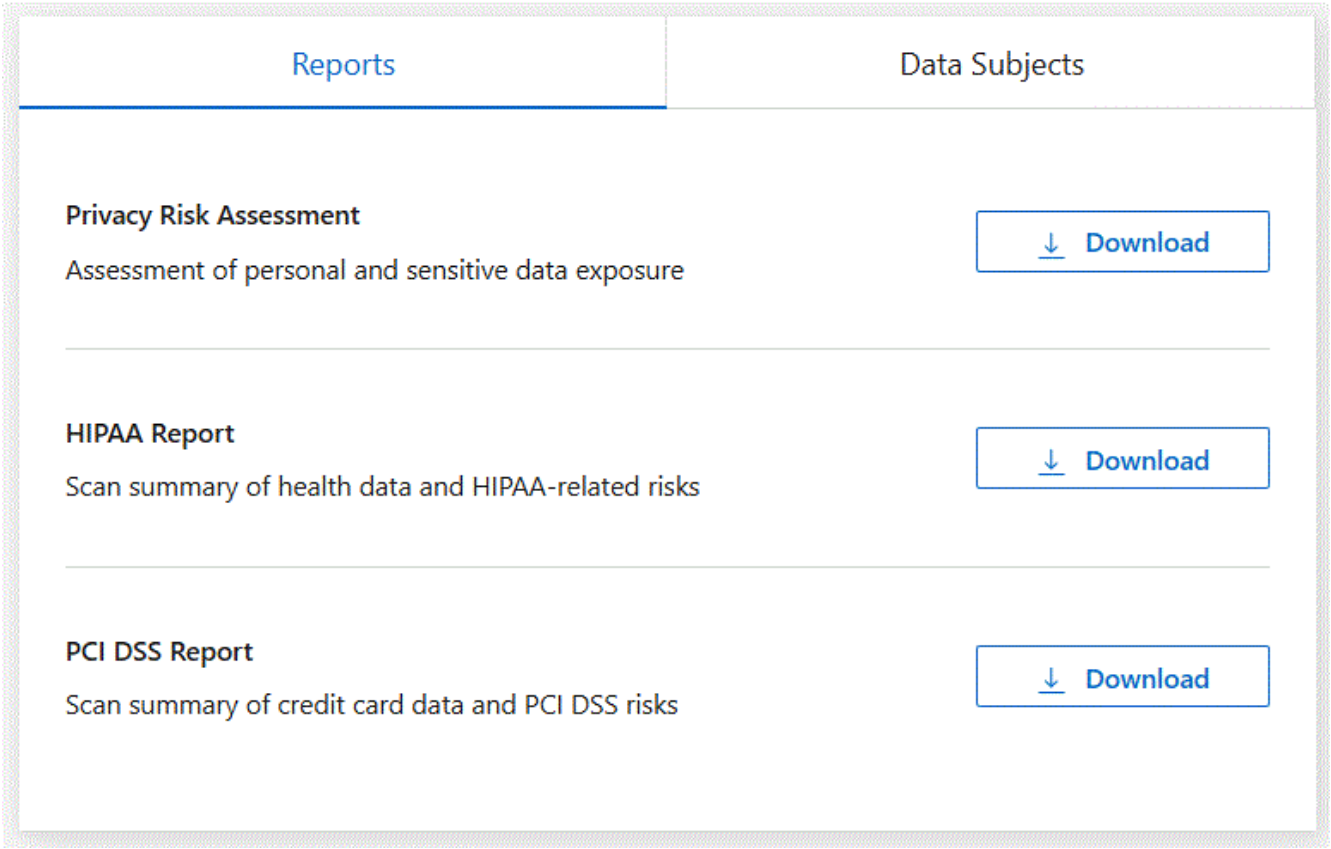
- ランサムウェア保護: ランサムウェア保護が有効になっているシステム、または有効になっていないシステム上にある、クレジットカード情報を含むファイルの割合。この情報はCloud Volumes ONTAPに固有のものです。
- 保持期間: ファイルが最後に変更された期間。これは、クレジットカード情報を処理に必要な期間以上保持するべきではないため、役立ちます。
- クレジットカード情報の配布: クレジットカード情報が見つかったシステムと、暗号化およびランサムウェア保護が有効になっているかどうか。

PCI DSSレポートを生成する

レポートを生成するには、「コンプライアンス」タブに移動します。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. レポート ペイン を見つけます。 **PCI DSS** レポート の横にあるダウンロード アイコンを選択します。



結果

データ分類では、必要に応じて確認したり他のグループに送信したりできる PDF レポートが生成されます。

プライバシーリスク評価レポート

プライバシー リスク評価レポートでは、GDPR や CCPA などのプライバシー規制の要件に従って、組織のプライバシー リスク状態の概要が提供されます。

レポートには次の情報が含まれます。

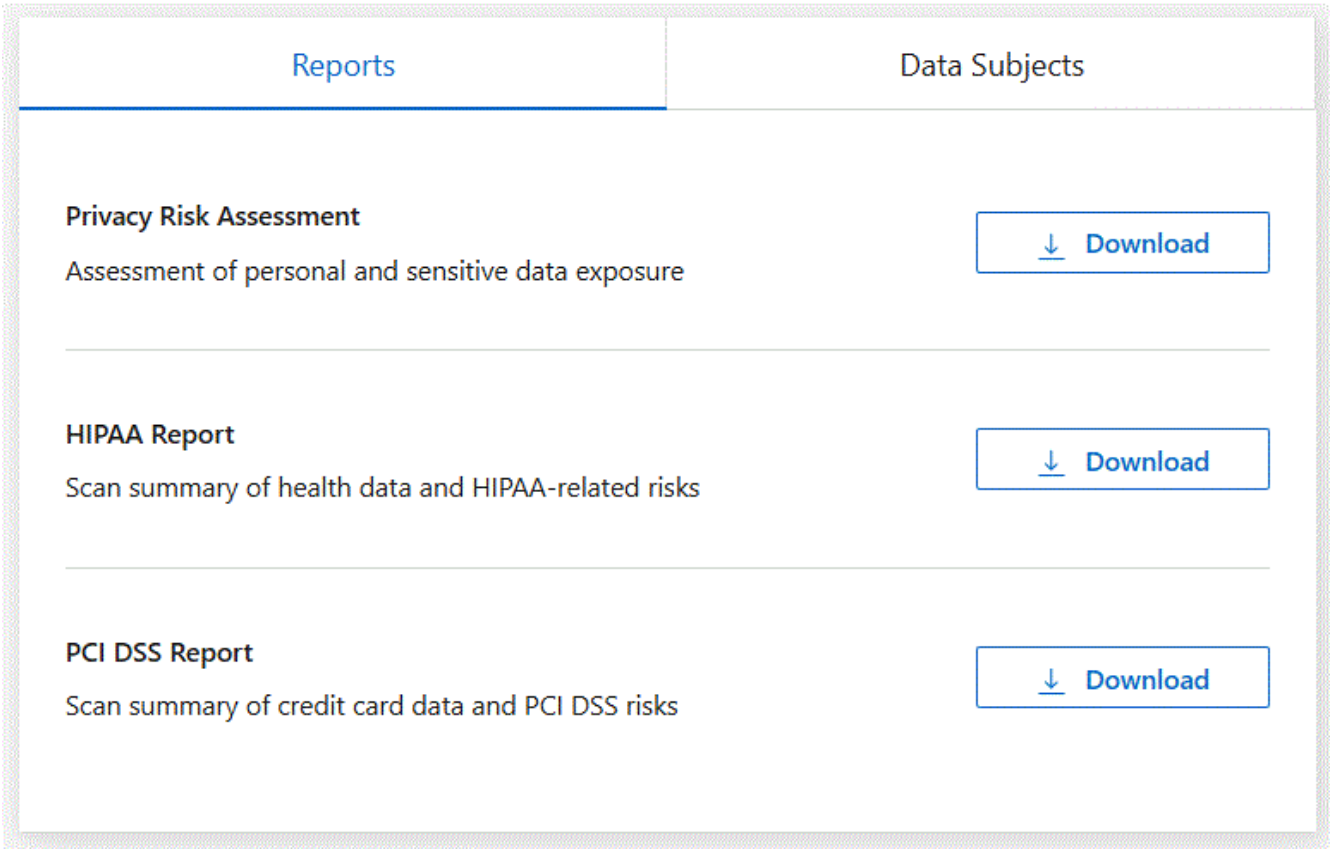
- コンプライアンス ステータス: 重大度スコアと、データが非機密データ、個人情報、または機密個人情報であるかどうかの分布。
- 評価の概要: 見つかった個人データの種類とデータのカテゴリの内訳。
- この評価におけるデータ主体: 国民識別子が見つかった場所別の人数。

プライバシーリスク評価レポートを生成する

レポートを生成するには、「コンプライアンス」タブに移動します。

手順

1. データ分類メニューから、*コンプライアンス*を選択します。
2. レポート ペイン を見つけます。 プライバシー リスク評価レポート の横にあるダウンロード アイコンを選択します。



結果

データ分類では、必要に応じて確認したり他のグループに送信したりできる PDF レポートが生成されます。

重症度スコア

データ分類では、次の 3 つの変数に基づいてプライバシー リスク評価レポートの重大度スコアを計算します。

- すべてのデータのうち個人データが占める割合。

- すべてのデータのうちの機密個人データの割合。
- 国民 ID、社会保障番号、納税者番号などの国民識別子によって決定されるデータ主体を含むファイルの割合。

スコアを決定するために使用されるロジックは次のとおりです。

重症度スコア	論理
0	3つの変数はすべて0%
1	変数の1つが0%より大きい
2	変数の1つが3%より大きい
3	変数のうち2つは3%より大きい
4	変数のうち3つは3%より大きい
5	変数の1つが6%より大きい
6	変数のうち2つは6%より大きい
7	変数のうち3つは6%より大きい
8	変数の1つが15%より大きい
9	変数のうち2つは15%より大きい
10	変数のうち3つは15%より大きい

NetApp Data Classificationの健全性を監視する

NetApp Data Classification Health Monitor ダッシュボードは、パフォーマンスのリアルタイム監視と分析情報を提供します。ヘルス モニターは、データ分類インフラストラクチャ、システムの健全性、使用状況メトリック、使用率データに関する情報を取得し、問題を特定して修復できるようにします。

ヘルスモニターの洞察

ヘルス モニター ダッシュボードには、4 つのカテゴリで情報が表示されます。

- インフラストラクチャの状態

  バージョンの状態、システムの安定性、展開の種類、マシンのスケールなどの情報を表示します。

- 問題のあるコンテナ

  問題のあるコンテナ フィールドを確認して、頻繁に停止または再起動されるコンテナに関する情報を取得します。この情報を使用して、特定のコンテナーを調査します。

- システム情報

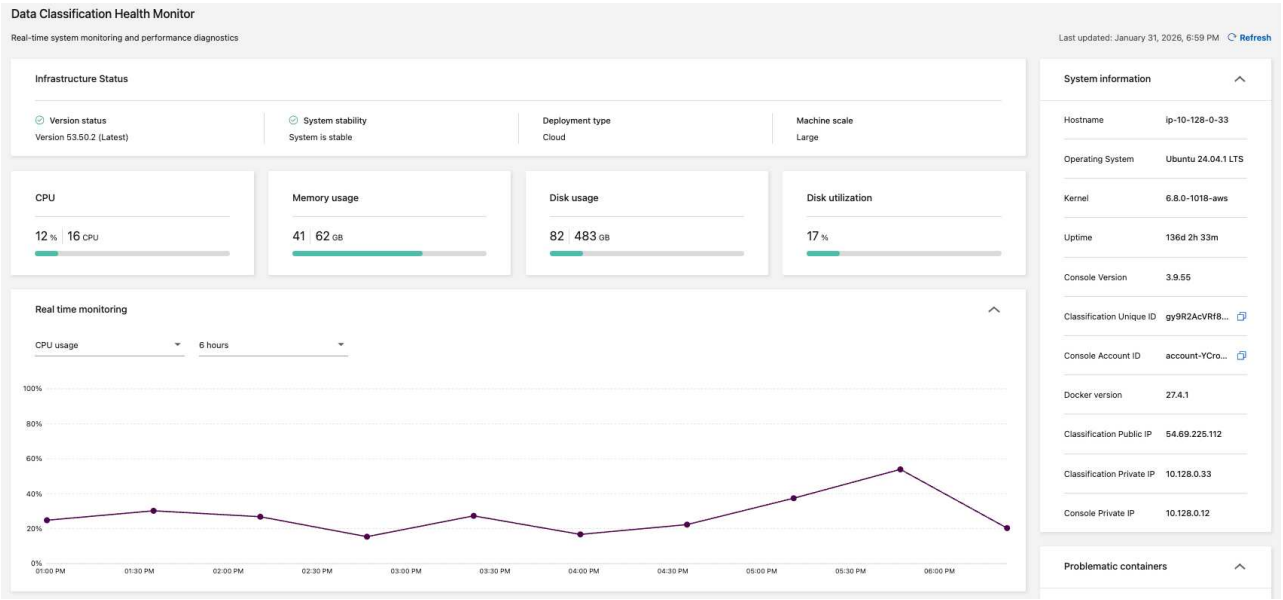
  システム情報パネルには、パブリックおよびプライベート IP アドレス、ホスト名、オペレーティング システム、コンソール バージョン、コンソール ID など、 NetApp Consoleとデータ分類に関する重要な情報が記録されます。

- 使用と活用

  CPU 使用率、ディスク使用率、ディスク使用率、メモリ使用率を確認します。これらの値は、ストレージ単位 (GB) または合計使用量のパーセンテージで表示されます。いずれかのフィールドに警告が表示された場合は、情報と修復の推奨事項については警告を選択してください。

ヘルスモニターダッシュボードにアクセスする

1. データ分類で、構成 を選択します。
2. 構成 の見出しの下で、データ分類ヘルスモニター を選択します。
3. ヘルスモニターダッシュボードでは、次のことができます。
   - 使用状況と利用状況を確認します。使用状況または使用率のメトリックに警告が表示される場合は、問題を解決するための推奨事項の警告を選択してください。
   - グラフを切り替えると、CPU 使用率、ディスク使用率、ディスク使用率、メモリ使用率が表示されます。x 軸を変更して、コンテンツを時間単位 (6、12、または 24) または日単位 (2、7、または 14) で表示できます。
   - 最新のデータ メトリックを表示するには、ダッシュボードを更新します。



データ分類の管理

NetApp Data Classificationスキャンから特定のディレクトリを除外する

NetApp Data Classification で特定のディレクトリをスキャンから除外する場合は、これらのディレクトリ名を構成ファイルに追加できます。この変更を適用すると、データ分類エンジンはそれらのディレクトリをスキャンから除外します。

デフォルトでは、データ分類スキャンでは、ボリューム内のソースと同一のボリューム スナップショット データが除外されます。

サポートされているデータソース

データ分類スキャンから特定のディレクトリを除外することは、次のデータ ソースの NFS および CIFS 共有でサポートされています。

- オンプレミスのONTAP
- Cloud Volumes ONTAP
- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- 一般的なファイル共有

スキャンから除外するディレクトリを定義する

ディレクトリを分類スキャンから除外する前に、データ分類システムにログインして、構成ファイルを編集し、スクリプトを実行する必要があります。方法を見る"データ分類システムにログインする"Linux マシンにソフトウェアを手動でインストールしたか、インスタンスをクラウドに展開したかによって異なります。

考慮事項

- データ分類システムごとに最大 50 個のディレクトリ パスを除外できます。
- ディレクトリ パスを除外すると、スキャン時間に影響する可能性があります。

手順

1. データ分類システムで、「/opt/netapp/config/custom_configuration」に移動し、ファイルを開きます。`data_provider.yaml` 。
2. 「exclude:」行の下の「data_providers」セクションに、除外するディレクトリ パスを入力します。例えば：

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

このファイル内の他の部分は変更しないでください。

3. ファイルへの変更を保存します。
4. 「/opt/netapp/Datasense/tools/customer_configuration/data_providers」に移動し、次のスクリプトを実行します。

```
update_data_providers_from_config_file.sh
```

+ このコマンドは、スキャンから除外するディレクトリを分類エンジンにコミットします。

結果

以降のすべてのデータスキャンでは、指定されたディレクトリのスキャンは除外されます。

同じ手順を使用して、除外リストの項目を追加、編集、または削除できます。スクリプトを実行して変更をコミットすると、修正された除外リストが更新されます。

例

構成1:

名前のどこかに「folder1」が含まれるすべてのフォルダーは、すべてのデータ ソースから除外されます。

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

除外されるパスの予想される結果:

- /CVO1/フォルダ1
- /CVO1/フォルダ1名
- /CVO1/フォルダ10
- /CVO1/*フォルダ1
- /CVO1/+フォルダ1名
- /CVO1/notfolder10
- /CVO22/フォルダ1
- /CVO22/フォルダ1名
- /CVO22/フォルダ10

除外されないパスの例:

- /CVO1/*フォルダ
- /CVO1/フォルダ名
- /CVO22/*フォルダ20

構成2:

名前の先頭にのみ「*folder1」が含まれるすべてのフォルダーが除外されます。

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

除外されるパスの予想される結果:

- /CVO/*フォルダ1
- /CVO/*フォルダ1名
- /CVO/*フォルダ10

除外されないパスの例:

- /CVO/フォルダ1
- /CVO/フォルダ1名
- /CVO/not*folder10

構成3:

データ ソース「CVO22」内の、名前のどこかに「folder1」が含まれるすべてのフォルダーが除外されます。

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

除外されるパスの予想される結果:

- /CVO22/フォルダ1
- /CVO22/フォルダ1名
- /CVO22/フォルダ10

除外されないパスの例:

- /CVO1/フォルダ1
- /CVO1/フォルダ1名
- /CVO1/フォルダ10

フォルダ名の特殊文字をエスケープする

フォルダー名に次のいずれかの特殊文字が含まれており、そのフォルダー内のデータをスキャン対象から除外する場合は、フォルダー名の前にエスケープ シーケンス \\ を使用する必要があります。

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

例えば：

ソース内のパス: `/project/*not_to_scan`

除外ファイルの構文: "*not_to_scan"

現在の除外リストを表示する

内容は `data_provider.yaml`設定ファイルは、実行後に実際にコミットされたものと異なる可能性があります。 `update_data_providers_from_config_file.sh`スクリプト。データ分類スキャンから除外したディレクトリの現在のリストを表示するには、「/opt/netapp/Datasense/tools/customer_configuration/data_providers」から次のコマンドを実行します。

```
get_data_providers_configuration.sh
```

NetApp Data Classificationで組織に公開されている追加のグループIDを定義します。

グループ ID (GID) が NFS ファイル共有内のファイルまたはフォルダーに関連付けられると、そのファイルまたはフォルダーのアクセス許可 (組織に公開されているかどうかなど) が定義されます。一部の GID が最初に「組織に公開」権限レベルで設定されていない場合は、その権限を GID に追加して、その GID が添付されているすべてのファイルとフォルダーが「組織に公開」されているとみなされるようにすることができます。

この変更を行った後、 NetApp Data Classificationによってファイルとフォルダが再スキャンされると、これらのグループ ID が添付されているすべてのファイルとフォルダの [調査の詳細] ページにこの権限が表示され、ファイル権限を表示するレポートにも表示されます。

この機能を有効にするには、データ分類システムにログインして、構成ファイルを編集し、スクリプトを実行する必要があります。方法を見る"データ分類システムにログインする"Linux マシンにソフトウェアを手動でインストールしたか、インスタンスをクラウドに展開したかによって異なります。

グループIDに「組織に公開」権限を追加する

このタスクを開始する前に、グループ ID 番号 (GID) が必要です。

手順

1. データ分類システムで、「/opt/netapp/config/custom_configuration」に移動し、ファイルを開きます。 `data_provider.yaml` 。
2. 「organization_group_ids: []」の行にグループIDを追加します。例えば：

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

   このファイルの他の部分は変更しないでください。

3. ファイルへの変更を保存します。
4. 「/opt/netapp/Datasense/tools/customer_configuration/data_providers」に移動し、次のスクリプトを実行します。

```
update_data_providers_from_config_file.sh
```

このコマンドは、変更されたグループ ID 権限を分類エンジンにコミットします。

結果

データのその後のすべてのスキャンでは、これらのグループ ID が添付されているファイルまたはフォルダーが「組織に公開」として識別されます。

同じ手順を使用して、グループ ID のリストを編集し、過去に追加したグループ ID を削除できます。スクリプトを実行して変更をコミットすると、修正されたグループ ID のリストが更新されます。

現在のグループIDのリストを表示する

内容は `data_provider.yaml`設定ファイルは、実行後に実際にコミットされたものと異なる可能性があります。 `update_data_providers_from_config_file.sh`スクリプト。データ分類に追加したグループ ID の現在のリストを表示するには、「/opt/netapp/Datasense/tools/customer_configuration/data_providers」から次のコマンドを実行します。

```
get_data_providers_configuration.sh
```

NetApp Data Classificationで古いデータ定義をカスタマイズする

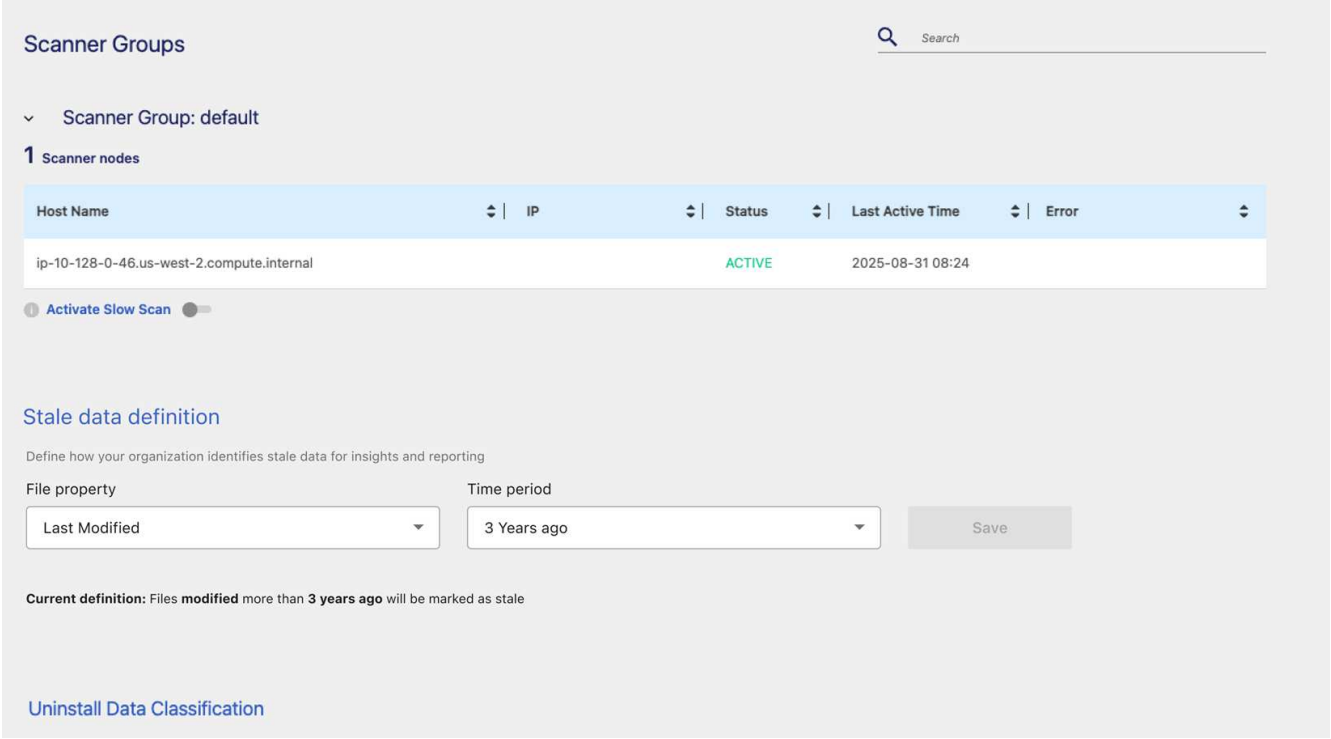
NetApp Data Classification は古いデータを識別し、節約の機会とガバナンスのリスクを特定するのに役立ちます。古いデータの定義は組織のコンテキストによって異なる可能性があるため、データ分類で古いデータを定義する方法をカスタマイズできます。

古いデータは、最終アクセス または 最終変更 された日時に基づいて定義できます。選択できる期間は 6 か月前から 10 年前までです。

デフォルトでは、データが最後に変更されてから 3 年前であれば古いデータとみなされます。

古いデータを定義する

1. ランサムウェア耐性で、[構成] を選択します。
2. 構成ページで、古いデータ定義 の見出しまでスクロールします。
3. ファイル プロパティ ドロップダウン メニューで、最終アクセス または 最終変更 に基づいて古いデータを定義するかどうかを選択します。
4. 古いデータ定義の期間を選択します。



5. 保存を選択します。

NetApp Data Classificationからデータソースを削除する

必要に応じて、 NetApp Data Classification による 1 つ以上のシステム、データベース、またはファイル共有グループのスキャンを停止できます。

システムのスキャンを無効にする

スキャンを非アクティブ化すると、データ分類はシステム上のデータをスキャンしなくなり、インデックス付けされた分析情報がデータ分類インスタンスから削除されます。システム自体のデータは削除されません。

1. _Configuration_ページから、 システムの行のボタンをクリックし、*データ分類を非アクティブ化*します。

システムを選択するときに、[サービス] パネルからシステムのスキャンを無効にすることもできます。

データ分類からデータベースを削除する

特定のデータベースをスキャンする必要がなくなった場合は、データ分類インターフェースからそのデータベースを削除し、すべてのスキャンを停止できます。

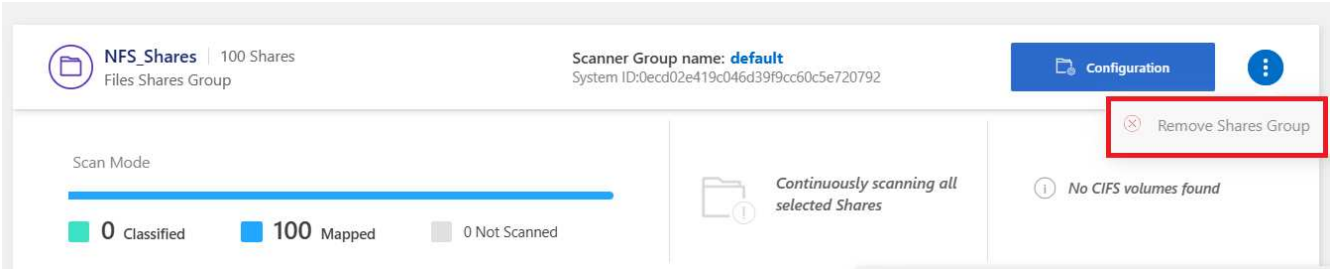
1. _Configuration_ページから、 データベースの行のボタンをクリックし、*DB サーバーの削除*をクリックします。

データ分類からファイル共有のグループを削除する

ファイル共有グループからユーザー ファイルをスキャンする必要がなくなった場合は、データ分類インターフェースからファイル共有グループを削除し、すべてのスキャンを停止できます。

手順

1. _Configuration_ページから、ファイル共有グループの行にあるボタンをクリックし、*ファイル共有グループの削除*をクリックします。



2. 確認ダイアログから*共有グループの削除*を選択します。

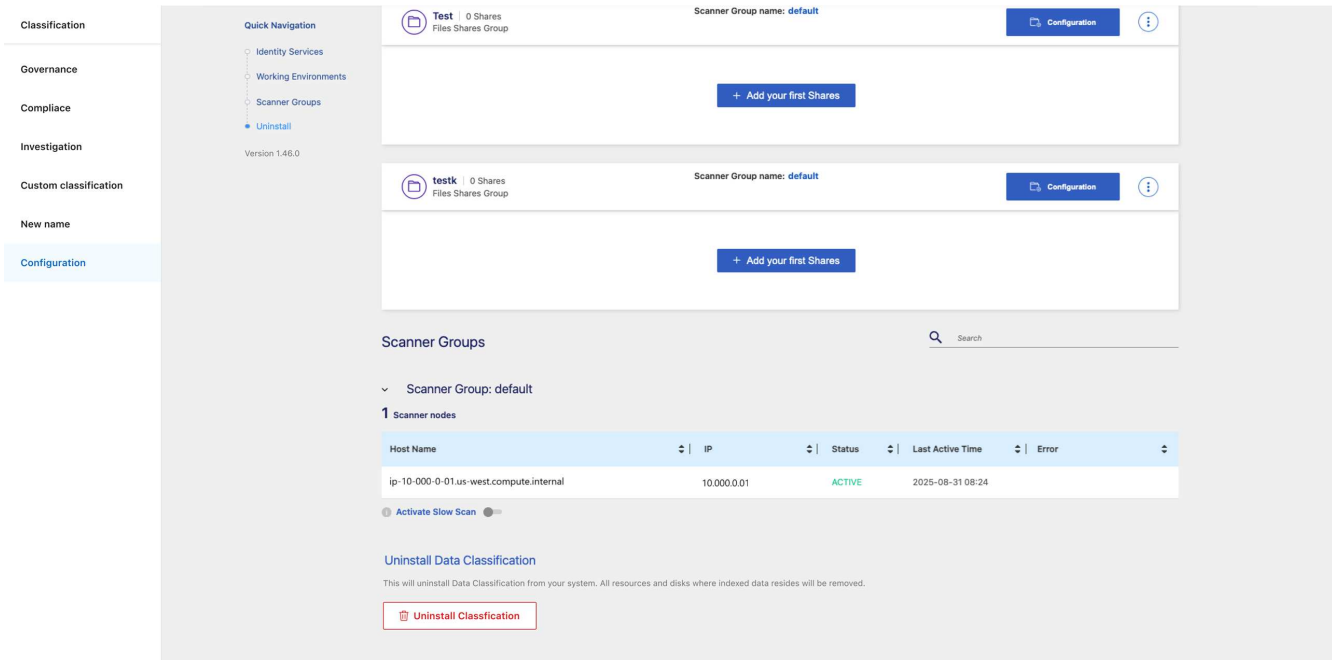
NetApp Data Classificationをアンインストールする

問題のトラブルシューティングを行ったり、ホストからソフトウェアを完全に削除したりするために、 NetApp Data Classificationをアンインストールすることができます。インスタンスを削除すると、インデックス付けされたデータが存在する関連ディスクも削除されるため、データ分類によってスキャンされたすべての情報が完全に削除されます。

使用する必要がある手順は、データ分類をクラウドに展開したか、オンプレミス ホストに展開したかによって異なります。

クラウドプロバイダーからデータ分類をアンインストールする

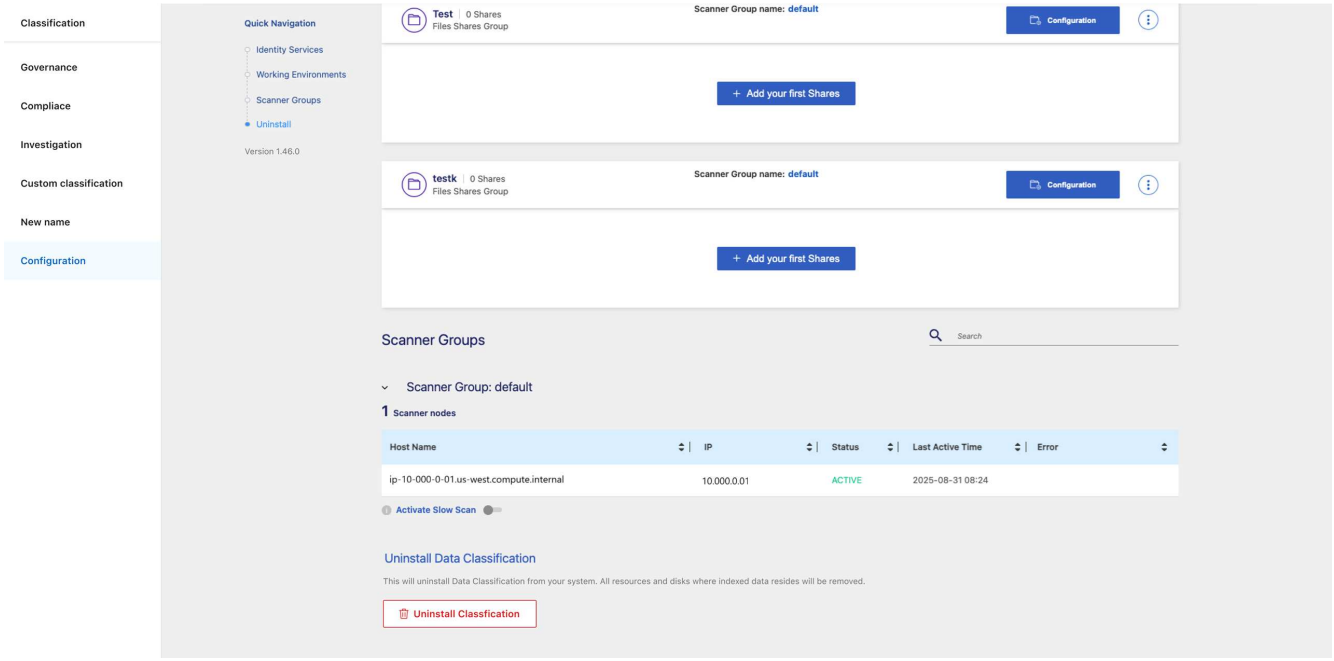
1. データ分類から構成を選択します。
2. 構成ページの下部で、[**Uninstall Classification**] を選択します。



3. ダイアログで「uninstall」と入力して、コンソール エージェントからデータ分類インスタンスの切断を続行します。確認するにはアンインストールを選択します。
4. [*Uninstall Classification*] ダイアログで、「**uninstall**」と入力して、コンソール エージェントからデータ分類インスタンスを切断することを確認し、[**Uninstall**] を選択します。
5. アンインストール プロセスを完了するには、クラウド プロバイダーのコンソールに移動し、データ分類インスタンスを削除します。インスタンスの名前は *CloudCompliance* となり、生成されたハッシュ (UUID) が連結されます。例: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

オンプレミス展開からデータ分類をアンインストールする

1. データ分類から構成を選択します。
2. 構成ページの下部で、[**Uninstall Classification**] を選択します。



3. ダイアログで「uninstall」と入力して、コンソール エージェントからデータ分類インスタンスの切断を続行します。確認するにはアンインストールを選択します。
4. ホストからソフトウェアをアンインストールするには、 `cleanup.sh`データ分類ホスト マシン上のスクリプト。例:

```
cleanup.sh
```

スクリプトは `/install/light_probe/onprem_installer/cleanup.sh`ディレクトリ。方法を見る"データ分類ホストマシンにログインする"。

参照

サポートされている**NetApp Data Classification**インスタンスタイプ

NetApp Data Classificationソフトウェアは、特定のオペレーティング システム要件、RAM 要件、ソフトウェア要件などを満たすホスト上で実行する必要があります。クラウドでデータ分類を展開する場合は、完全な機能を実現するために「大規模」特性を持つシステムを使用することをお勧めします。

CPU 数と RAM 数が少ないシステムでもデータ分類を展開できますが、これらの非力なシステムを使用する場合はいくつかの制限があります。"これらの制限について学ぶ" 。

次の表で、「デフォルト」とマークされているシステムが、Data Classification をインストールするリージョンで使用できない場合は、表の次のシステムが展開されます。

AWSインスタンスタイプ

システムサイズ	仕様	インスタンスタイプ
特大	32 個の CPU、128 GB の RAM、1 TiB の gp3 SSD	"m6i.8xlarge"（デフォルト）
大規模	16 CPU、64 GB RAM、500 GiB SSD	"m6i.4xlarge"(デフォルト) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
中	8 CPU、32 GB RAM、200 GiB SSD	"m6i.2xlarge"(デフォルト) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
小規模	8 CPU、16 GB RAM、100 GiB SSD	"c6a.2xlarge"(デフォルト) c5a.2xlarge c5.2xlarge c4.2xlarge

Azureインスタンスの種類

システムサイズ	仕様	インスタンスタイプ
特大	32 個の CPU、128 GB の RAM、OS ディスク (2,048 GiB、最小 250 MB/秒のスループット)、およびデータ ディスク (1 TiB SSD、最小 750 MB/秒のスループット)	"標準_D32_v3"（デフォルト）
大規模	16 CPU、64 GB RAM、500 GiB SSD	"標準_D16s_v3"（デフォルト）

GCPインスタンスタイプ

システムサイズ	仕様	インスタンスタイプ
大規模	16 CPU、64 GB RAM、500 GiB SSD	"n2-標準-16"(デフォルト) n2d-standard-16 n1-standard-16

NetApp Data Classificationのデータソースから収集されたメタデータ

NetApp Data Classification は、データ ソースおよびシステムからのデータの分類スキャンを実行するときに、特定のメタデータを収集します。データ分類では、データを分類するために必要なメタデータのほとんどにアクセスできますが、必要なデータにアクセスできないソースもいくつかあります。

	メタデータ	CIFS	NFS
タイムスタンプ	作成時間	利用可能	利用できません（Linuxではサポートされていません）
	最終アクセス時間	利用可能	利用可能
	最終更新日時	利用可能	利用可能
権限	オープン権限	「EVERYONE」グループがファイルにアクセスできる場合、そのファイルは「組織に公開」されているとみなされます。	「その他」がファイルにアクセスできる場合、そのファイルは「組織に公開」されているとみなされます。
	ユーザー/グループアクセス	ユーザーとグループの情報はLDAPから取得されます	利用できません (NFS ユーザーは通常、サーバー上でローカルに管理されるため、同じユーザーが各サーバーで異なる UID を持つことができます)

- データ分類では、データベース データ ソースから「最終アクセス時刻」を抽出しません。
- 古いバージョンの Windows OS (Windows 7 や Windows 8 など) では、システム パフォーマンスに影響を及ぼす可能性があるため、「最終アクセス時刻」属性の収集がデフォルトで無効になっています。この属性が収集されない場合、「最終アクセス時刻」に基づくデータ分類分析に影響が出ます。必要に応じて、これらの古い Windows システムでの最終アクセス時刻の収集を有効にすることができます。

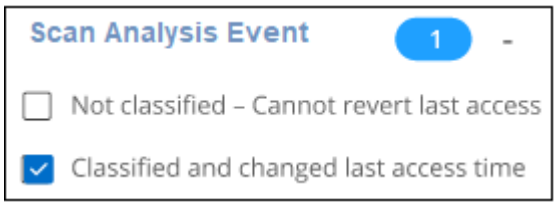
最終アクセス時間のタイムスタンプ

データ分類がファイル共有からデータを抽出すると、オペレーティング システムはそれをデータへのアクセスと見なし、それに応じて「最終アクセス時刻」を変更します。スキャン後、データ分類は最終アクセス時刻を元のタイムスタンプに戻そうとします。データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、システムは最終アクセス時刻を元のタイムスタンプに戻すことができません。 SnapLockが設定されたONTAPボリュームには読み取り専用権限があり、最終アクセス時刻を元のタイムスタンプに戻すこともできません。

デフォルトでは、データ分類にこれらの権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のこれらのファイルをスキャンしません。ただし、ファイルの最終アクセス時刻が元の時刻にリセットされても構わない場合は、構成ページの下部にある「属性の書き込み」権限がない場合にスキャンする スイッチを選択して、権限に関係なくデータ分類がボリュームをスキャンするようにすることができます。

この機能は、オンプレミスのONTAPシステム、 Cloud Volumes ONTAP、 Azure NetApp Files、 Amazon FSx for NetApp ONTAP管理、およびサードパーティのファイル共有に適用できます。

調査ページには、「スキャン分析イベント」というフィルターがあり、これを使用すると、データ分類で最終アクセス時刻を戻せなかったために分類されなかったファイル、またはデータ分類で最終アクセス時刻を戻せなかったにもかかわらず分類されたファイルを表示できます。



フィルターの選択肢は次のとおりです。

- 「未分類 - 最終アクセス時間を戻すことができません」 - 書き込み権限がないため分類されなかったファイルが表示されます。
- 「分類され、更新された最終アクセス時刻」 - 分類されたファイルが表示され、データ分類では最終アクセス時刻を元の日付にリセットできませんでした。このフィルターは、*「属性の書き込み」権限がない場合にスキャン*をオンにした環境にのみ関連します。

必要に応じて、これらの結果をレポートにエクスポートして、権限によりスキャンされているファイルとスキャンされていないファイルを確認できます。"データ調査レポートの詳細"。

NetApp Data Classificationシステムにログインする

ログ ファイルにアクセスしたり、構成ファイルを編集したりするには、 NetApp Data Classificationシステムにログインする必要があります。

Data Classification がオンプレミスの Linux マシンまたはクラウドに展開した Linux マシンにインストールされている場合は、構成ファイルとスクリプトに直接アクセスできます。

Data Classification をクラウドにデプロイする場合は、Data Classification インスタンスに SSH で接続する必要があります。ユーザー名とパスワードを入力するか、コンソール エージェントのインストール時に指定し

た SSH キーを使用して、システムに SSH 接続します。 SSH コマンドは次のとおりです。

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

- `<path_to_the_ssh_key>`= SSH認証キーの場所
- `<machine_user>`:
  - AWSの場合: <ec2-user> を使用します
  - Azureの場合: コンソールインスタンス用に作成されたユーザーを使用します
  - GCPの場合: コンソールインスタンス用に作成されたユーザーを使用します
- `<datasense_ip>`= 仮想マシンインスタンスのIPアドレス

クラウド内のシステムにアクセスするには、セキュリティ グループの受信ルールを変更する必要があります。詳細については、以下を参照してください。

- "AWSのセキュリティグループルール"
- "Azure のセキュリティ グループ ルール"
- "Google Cloud のファイアウォール ルール"

NetApp Data ClassificationAPI

Web UI を通じて利用できるNetApp Data Classification機能は、REST API を通じても利用できます。

データ分類内には、UI のタブに対応する 4 つのカテゴリが定義されています。

- 調査
- コンプライアンス
- ガバナンス
- 構成

Swagger ドキュメントの API を使用すると、検索、データの集約、スキャンの追跡、コピー、移動、削除などのアクションを実行できます。

概要

API を使用すると、次の機能を実行できます。

- 輸出情報
  - UI で利用できるものはすべて API 経由でエクスポートできます (レポートを除く)
  - データは JSON 形式でエクスポートされます (簡単に解析でき、Splunk などのサードパーティ アプリケーションにプッシュできます)
- 「AND」および「OR」ステートメントを使用してクエリを作成し、情報を含めたり除外したりします。

  たとえば、特定の個人識別情報 (PII) を含まないファイルを見つけることができます (UI では機能は使用で

きません)。エクスポート操作から特定のフィールドを除外することもできます。

- アクションを実行する
  - CIFSクレデンシャルの更新
  - アクションの表示とキャンセル
  - ディレクトリを再スキャン
  - データをエクスポートする

API は安全で、UI と同じ認証方法を使用します。認証に関する情報は、"REST API ドキュメント" 。

Swagger APIリファレンスへのアクセス

Swagger にアクセスするには、データ分類インスタンスの IP アドレスが必要です。クラウド展開の場合は、パブリック IP アドレスを使用します。次に、このエンドポイントにアクセスする必要があります。

https://<分類IP>/documentation

APIの使用例

次の例は、ファイルをコピーするための API 呼び出しを示しています。

API 要求

調査タブですべてのフィルターを表示するには、最初にシステムの関連するフィールドとオプションをすべて取得する必要があります。

```
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"
```

応答

```
{
  "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"
    }
  ]
}
```

リクエストパラメータでその応答を使用して、コピーする目的のファイルをフィルタリングします。

複数のアイテムにアクションを適用できます。サポートされているアクション タイプには、移動、削除、コピーなどがあります。

コピーアクションを作成します。

API 要求

次の API はアクション API であり、複数のアクションを作成できます。

```
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"]}]}}"
```

応答

応答ではアクション オブジェクトが返されるため、get および delete API を使用してアクションのステータスを取得したり、アクションをキャンセルしたりできます。

```
{
  "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"
}
```

知識とサポート

NetApp Consoleサポートに登録する

NetApp Consoleとそのストレージ ソリューションおよびデータ サービスに固有のテクニカル サポートを受けるには、サポート登録が必要です。Cloud Volumes ONTAPシステムの主要なワークフローを有効にするには、サポート登録も必要です。

サポートに登録しても、クラウド プロバイダー ファイル サービスに対するNetAppサポートは有効になりません。クラウド プロバイダーのファイル サービス、そのインフラストラクチャ、またはサービスを使用するソリューションに関連するテクニカル サポートについては、その製品のドキュメントの「ヘルプの取得」を参照してください。

- "Amazon FSx for ONTAP"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

サポート登録の概要

サポート資格を有効にするには、次の 2 つの登録形式があります。

- NetApp Consoleアカウントのシリアル番号 (コンソールの [サポート リソース] ページにある 20 桁の 960xxxxxxxxx シリアル番号) を登録します。

  これは、コンソール内のすべてのサービスに対する単一のサポート サブスクリプション ID として機能します。各コンソール アカウントを登録する必要があります。

- クラウド プロバイダーのマーケットプレイスで、サブスクリプションに関連付けられたCloud Volumes ONTAPシリアル番号を登録します (これらは 20 桁の 909201xxxxxxxx シリアル番号です)。

  これらのシリアル番号は一般に *PAYGO* シリアル番号 と呼ばれ、 Cloud Volumes ONTAP の導入時にNetApp Consoleによって生成されます。

両方のタイプのシリアル番号を登録すると、サポート チケットの開設やケースの自動生成などの機能が有効になります。登録は、以下の説明に従ってNetAppサポート サイト (NSS) アカウントをコンソールに追加することで完了します。

NetAppサポートのためにNetApp Consoleを登録する

サポートに登録し、サポート資格を有効にするには、 NetApp Consoleアカウントの 1 人のユーザーがNetAppサポート サイト アカウントをコンソール ログインに関連付ける必要があります。NetAppサポートに登録する方法は、 NetAppサポート サイト (NSS) アカウントをすでにお持ちかどうかによって異なります。

NSSアカウントをお持ちの既存顧客

NSS アカウントをお持ちのNetApp のお客様の場合は、コンソールからサポートに登録するだけです。

手順

1. 管理 > *資格情報*を選択します。
2. *ユーザー資格情報*を選択します。
3. **NSS** 資格情報の追加 を選択し、 NetAppサポート サイト (NSS) の認証プロンプトに従います。
4. 登録プロセスが成功したことを確認するには、[ヘルプ] アイコンを選択し、[サポート] を選択します。

   リソース ページには、コンソール アカウントがサポートに登録されていることが表示されます。

   他のコンソール ユーザーは、ログインにNetAppサポート サイト アカウントを関連づけていない場合、同じサポート登録ステータスを表示しないことに注意してください。ただし、これはあなたのアカウントがサポートに登録されていないことを意味するものではありません。組織内の 1 人のユーザーがこれらの手順を実行していれば、アカウントは登録済みになります。

既存の顧客だが**NSS**アカウントがない

既存のNetApp顧客であり、既存のライセンスとシリアル番号を持っているものの、NSS アカウントを持っていない場合は、NSS アカウントを作成し、それをコンソール ログインに関連付ける必要があります。

手順

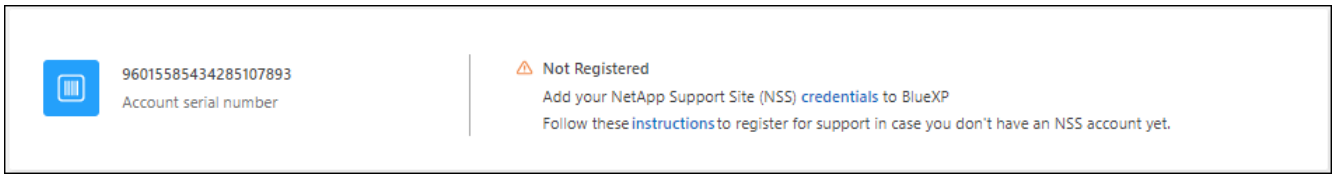
1. NetAppサポートサイトのアカウントを作成するには、 "NetAppサポートサイト ユーザー登録フォーム"
   a. 適切なユーザー レベル (通常は * NetApp顧客/エンド ユーザー*) を選択してください。
   b. 上記で使用したコンソール アカウントのシリアル番号 (960xxxx) を必ずシリアル番号フィールドにコピーしてください。これにより、アカウント処理が高速化されます。
2. 以下の手順を実行して、新しいNSSアカウントをコンソールログインに関連付けます。NSSアカウントをお持ちの既存顧客 。

NetAppの新着情報

NetAppを初めて使用し、NSS アカウントをお持ちでない場合は、以下の手順に従ってください。

手順

1. コンソールの右上にあるヘルプ アイコンを選択し、サポート を選択します。
2. サポート登録ページからアカウント ID シリアル番号を見つけます。



3. 移動先 "NetAppのサポート登録サイト"私は登録済みの**NetApp**顧客ではありません を選択します。
4. 必須フィールド（赤いアスタリスクが付いているフィールド）に入力します。
5. 製品ライン フィールドで、**Cloud Manager** を選択し、該当する請求プロバイダーを選択します。
6. 上記の手順 2 からアカウントのシリアル番号をコピーし、セキュリティ チェックを完了して、NetApp のグローバル データ プライバシー ポリシーを読んだことを確認します。

   この安全な取引を完了するために、指定されたメールボックスに電子メールが直ちに送信されます。検証

メールが数分以内に届かない場合は、必ずスパム フォルダーを確認してください。

7. メール内からアクションを確認します。

   確認すると、リクエストがNetAppに送信され、 NetAppサポート サイトのアカウントを作成することが推奨されます。

8. NetAppサポートサイトのアカウントを作成するには、 "NetAppサポートサイト ユーザー登録フォーム"
   a. 適切なユーザー レベル (通常は * NetApp顧客/エンド ユーザー*) を選択してください。
   b. 上記で使用したアカウントのシリアル番号 (960xxxx) を必ずシリアル番号フィールドにコピーしてください。これにより処理速度が向上します。

終了後の操作

このプロセス中に、 NetAppから連絡が来るはずです。これは、新規ユーザー向けの 1 回限りのオンボーディング演習です。

NetAppサポートサイトのアカウントを取得したら、以下の手順を実行して、アカウントをコンソールログインに関連付けます。NSSアカウントをお持ちの既存顧客 。

Cloud Volumes ONTAPサポートに NSS 認証情報を関連付ける

Cloud Volumes ONTAPの次の主要なワークフローを有効にするには、 NetAppサポート サイトの認証情報をコンソール アカウントに関連付ける必要があります。

- 従量課金制のCloud Volumes ONTAPシステムをサポート対象として登録する

  システムのサポートを有効にし、 NetAppテクニカル サポート リソースにアクセスするには、NSS アカウントを提供する必要があります。

- BYOL（個人ライセンス使用）時にCloud Volumes ONTAP を導入する

  コンソールがライセンス キーをアップロードし、購入した期間のサブスクリプションを有効にするには、NSS アカウントを提供する必要があります。これには、期間更新の自動更新が含まれます。

- Cloud Volumes ONTAPソフトウェアを最新リリースにアップグレードする

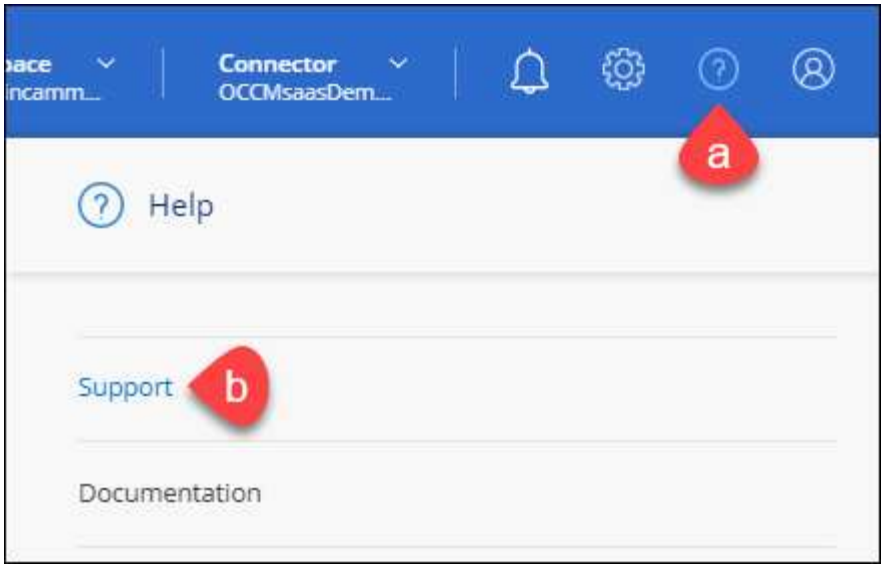
NSS 資格情報をNetApp Consoleアカウントに関連付けることは、コンソール ユーザー ログインに関連付けられている NSS アカウントとは異なります。

これらの NSS 資格情報は、特定のコンソール アカウント ID に関連付けられています。コンソール組織に属するユーザーは、サポート **> NSS** 管理 からこれらの資格情報にアクセスできます。

- 顧客レベルのアカウントをお持ちの場合は、1 つ以上の NSS アカウントを追加できます。
- パートナー アカウントまたは再販業者アカウントをお持ちの場合は、1 つ以上の NSS アカウントを追加できますが、顧客レベルのアカウントと一緒に追加することはできません。

手順

1. コンソールの右上にあるヘルプ アイコンを選択し、サポート を選択します。



2. *NSS管理 > NSSアカウントの追加*を選択します。
3. プロンプトが表示されたら、[続行] を選択して、Microsoft ログイン ページにリダイレクトします。

   NetApp は、サポートとライセンスに固有の認証サービスの ID プロバイダーとして Microsoft Entra ID を使用します。

4. ログイン ページで、 NetAppサポート サイトに登録した電子メール アドレスとパスワードを入力して、認証プロセスを実行します。

   これらのアクションにより、コンソールはライセンスのダウンロード、ソフトウェア アップグレードの検証、将来のサポート登録などに NSS アカウントを使用できるようになります。

   次の点に注意してください。

   - NSS アカウントは顧客レベルのアカウントである必要があります (ゲスト アカウントや一時アカウントではありません)。顧客レベルの NSS アカウントを複数持つことができます。
   - パートナー レベルのアカウントの場合、NSS アカウントは 1 つだけ存在できます。顧客レベルの NSS アカウントを追加しようとしたときに、パートナー レベルのアカウントが存在する場合は、次のエラー メッセージが表示されます。

     「異なるタイプの NSS ユーザーがすでに存在するため、このアカウントでは NSS 顧客タイプは許可されません。」

   既存の顧客レベルの NSS アカウントがあり、パートナー レベルのアカウントを追加しようとする場合も同様です。

   - ログインが成功すると、 NetApp はNSS ユーザー名を保存します。

     これは、メールにマッピングされるシステム生成の ID です。*NSS管理*ページでは、 ••• メニュー。

   - ログイン認証トークンを更新する必要がある場合は、 ••• メニュー。

     このオプションを使用すると、再度ログインするよう求められます。これらのアカウントのトークンは 90 日後に期限切れになることに注意してください。これを知らせる通知が投稿されます。

NetApp Data Classificationのサポートを受ける

NetApp は、 NetApp Consoleとそのクラウド サービスをさまざまな方法でサポートします。ナレッジ ベース (KB) 記事やコミュニティ フォーラムなど、豊富な無料のセルフ サポート オプションが 24 時間 365 日ご利用いただけます。サポート登録には、Webチケットによるリモートテクニカルサポートも含まれます。

クラウドプロバイダーのファイルサービスのサポートを受ける

クラウド プロバイダーのファイル サービス、そのインフラストラクチャ、またはサービスを使用するソリューションに関連するテクニカル サポートについては、その製品のドキュメントを参照してください。

- "Amazon FSx for ONTAP"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

NetAppとそのストレージ ソリューションおよびデータ サービスに固有のテクニカル サポートを受けるには、以下に説明するサポート オプションを使用してください。

セルフサポートオプションを使用する

以下のオプションは、24 時間 365 日無料でご利用いただけます。

- ドキュメント

  現在表示しているNetApp Consoleのドキュメント。

- "ナレッジベース"

  NetAppナレッジベースを検索して、問題のトラブルシューティングに役立つ記事を見つけます。

- "コミュニティ"

  NetApp Consoleコミュニティに参加して、進行中のディスカッションをフォローしたり、新しいディスカッションを作成したりできます。

NetAppサポートでケースを作成する

上記のセルフ サポート オプションに加えて、サポートを有効にした後は、 NetAppサポート スペシャリストと協力して問題を解決することもできます。

始める前に

- *ケースの作成*機能を使用するには、まずNetAppサポート サイトの資格情報をコンソール ログインに関連付ける必要があります。 "コンソールログインに関連付けられた資格情報を管理する方法を学びます" 。
- シリアル番号を持つONTAPシステムのケースを開く場合は、NSS アカウントがそのシステムのシリアル番号に関連付けられている必要があります。

手順

1. NetApp Consoleで、[ヘルプ] > [サポート] を選択します。
2. *リソース*ページで、テクニカル サポートの下にある利用可能なオプションのいずれかを選択します。

    a. 電話で誰かと話したい場合は、「電話する」を選択してください。電話をかけることができる電話番号をリストした netapp.com のページに移動します。

    b. NetAppサポート スペシャリストとのチケットを開くには、[ケースを作成] を選択します。

        - サービス: 問題が関連付けられているサービスを選択します。たとえば、* NetApp Console* は、コンソール内のワークフローまたは機能に関するテクニカル サポートの問題に固有の場合です。
        - システム: ストレージに該当する場合は、* Cloud Volumes ONTAP* または **On-Prem** を選択し、関連する作業環境を選択します。

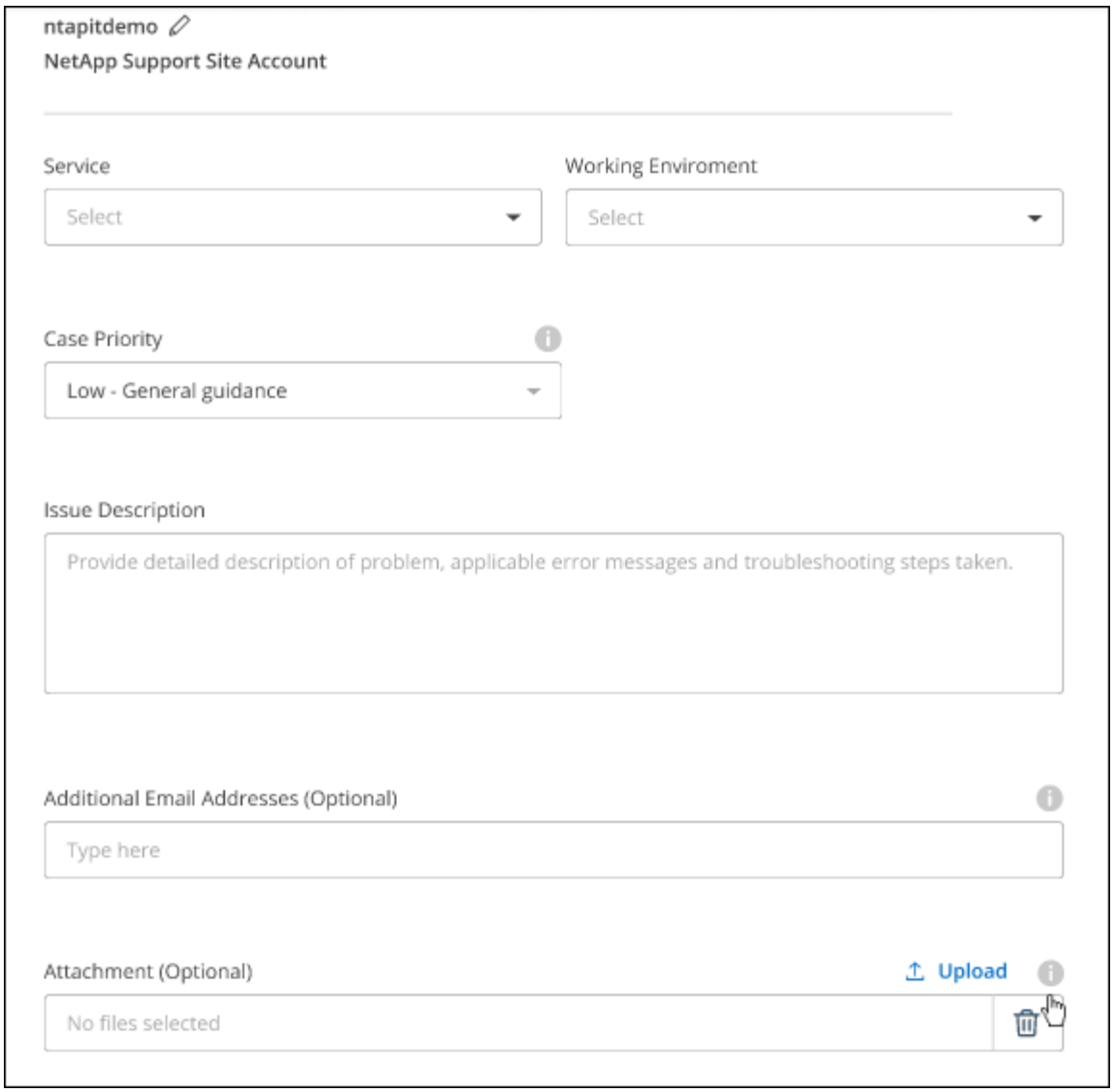
          システムのリストは、コンソール組織と、上部のバナーで選択したコンソール エージェントの範囲内にあります。

        - ケースの優先度: ケースの優先度 (低、中、高、重大) を選択します。

          これらの優先順位の詳細を確認するには、フィールド名の横にある情報アイコンの上にマウスを置きます。

        - 問題の説明: 該当するエラー メッセージや実行したトラブルシューティング手順など、問題の詳細な説明を入力します。
        - 追加のメールアドレス:この問題を他の人に知らせたい場合は、追加のメールアドレスを入力してください。
        - 添付ファイル **(**オプション**)**: 一度に 1 つずつ、最大 5 つの添付ファイルをアップロードします。

          添付ファイルはファイルごとに 25 MB までに制限されます。サポートされているファイル拡張子は、txt、log、pdf、jpg/jpeg、rtf、doc/docx、xls/xlsx、csv です。



終了後の操作

サポート ケース番号を示すポップアップが表示されます。NetAppサポート スペシャリストがお客様のケースを確認し、すぐにご連絡いたします。

サポート ケースの履歴については、設定 > タイムライン を選択し、「サポート ケースの作成」というアクションを探します。右端のボタンを使用すると、アクションを展開して詳細を表示できます。

ケースを作成しようとすると、次のエラー メッセージが表示される場合があります。

「選択したサービスに対してケースを作成する権限がありません」

このエラーは、NSS アカウントとそれに関連付けられているレコード会社が、NetApp Consoleアカウントのシリアル番号のレコード会社と同じではないことを意味している可能性があります (つまり、960xxxx) または作業環境のシリアル番号。次のいずれかのオプションを使用してサポートを求めることができます。

- 非技術的なケースを提出する https://mysupport.netapp.com/site/help

サポートケースを管理する

アクティブなサポート ケースと解決済みのサポート ケースをコンソールから直接表示および管理できます。NSS アカウントおよび会社に関連付けられたケースを管理できます。

次の点に注意してください。

- ページ上部のケース管理ダッシュボードには、次の 2 つのビューがあります。
  - 左側のビューには、指定したユーザー NSS アカウントによって過去 3 か月間に開かれたケースの合計が表示されます。
  - 右側のビューには、ユーザーの NSS アカウントに基づいて、会社レベルで過去 3 か月間に開かれたケースの合計が表示されます。

  表の結果には、選択したビューに関連するケースが反映されます。
- 関心のある列を追加または削除したり、優先度やステータスなどの列の内容をフィルタリングしたりできます。その他の列は並べ替え機能のみを提供します。

  詳細については、以下の手順をご覧ください。
- ケースごとに、ケースメモを更新したり、まだ「クローズ」または「クローズ保留中」ステータスになっていないケースをクローズしたりする機能を提供します。

手順

1. NetApp Consoleで、[ヘルプ] > [サポート] を選択します。
2. *ケース管理*を選択し、プロンプトが表示されたら、NSS アカウントをコンソールに追加します。

   ケース管理 ページには、コンソール ユーザー アカウントに関連付けられている NSS アカウントに関連するオープン ケースが表示されます。これは、**NSS** 管理 ページの上部に表示される NSS アカウントと同じです。
3. 必要に応じて、テーブルに表示される情報を変更します。
   - *組織のケース*の下で*表示*を選択すると、会社に関連付けられているすべてのケースが表示されます。
   - 正確な日付範囲を選択するか、別の期間を選択して日付範囲を変更します。
   - 列の内容をフィルタリングします。
   - 表に表示される列を変更するには、次に、表示する列を選択します。
4. 既存のケースを管理するには、利用可能なオプションのいずれかを選択します。
   - ケースを表示: 特定のケースに関する詳細をすべて表示します。
   - ケースノートを更新: 問題に関する追加の詳細を入力するか、*ファイルのアップロード*を選択して最大 5 つのファイルを添付します。

     添付ファイルはファイルごとに 25 MB までに制限されます。サポートされているファイル拡張子は、txt、log、pdf、jpg/jpeg、rtf、doc/docx、xls/xlsx、csv です。
   - ケースを閉じる: ケースを閉じる理由の詳細を入力し、[ケースを閉じる] を選択します。

NetApp Data Classificationに関するよくある質問

質問に対する簡単な回答を探している場合は、この FAQ が役立ちます。

NetApp Data Classification

次の質問は、データ分類に関する一般的な理解を提供します。

データ分類はどのように機能しますか?

データ分類は、NetApp Consoleシステムおよびストレージ システムとともに AI の別のレイヤーを展開します。次に、ボリューム、バケット、データベース、その他のストレージ アカウント上のデータをスキャンし、見つかったデータの分析情報をインデックス化します。データ分類では、正規表現とパターン マッチングを中心に構築される一般的な代替ソリューションとは対照的に、人工知能と自然言語処理の両方を活用します。

データ分類では AI を使用してデータのコンテキストを理解し、正確な検出と分類を実現します。最新のデータ タイプとスケールに合わせて設計されているため、AI によって駆動されます。また、強力で正確な検出と分類を提供するために、データのコンテキストも理解します。

"データ分類の仕組みについて詳しく見る" 。

データ分類には **REST API** がありますか? また、サードパーティのツールと連携できますか?

はい、データ分類には、コンソール コア プラットフォームの一部であるデータ分類バージョンでサポートされている機能用の REST API があります。見る"APIのドキュメント" 。

データ分類はクラウド マーケットプレイスを通じて利用できますか?

データ分類はNetApp Consoleのコア機能の一部であるため、このサービスにマーケットプレイスを使用する必要はありません。

データ分類のスキャンと分析

次の質問は、データ分類のスキャン パフォーマンスと分析に関するものです。

データ分類はどのくらいの頻度でデータをスキャンしますか?

データの最初のスキャンには多少時間がかかりますが、後続のスキャンでは増分変更のみが検査されるため、システムスキャン時間が短縮されます。データ分類では、一度に 6 つのリポジトリをラウンドロビン方式で継続的にスキャンするため、変更されたすべてのデータが非常に迅速に分類されます。

"スキャンの仕組みを学ぶ" 。

データ分類では、データベースは 1 日に 1 回だけスキャンされます。データベースは他のデータ ソースのように継続的にスキャンされるわけではありません。

データ スキャンは、ストレージ システムとデータにほとんど影響を与えません。

スキャンのパフォーマンスは変化しますか?

スキャンのパフォーマンスは、ネットワーク帯域幅と環境内の平均ファイル サイズによって異なります。また、ホスト システム (クラウドまたはオンプレミス) のサイズ特性によっても異なります。。 "データ分類インスタンス"そして"データ分類の導入"詳細についてはこちらをご覧ください。

新しいデータ ソースを最初に追加するときに、完全な「分類」(マップと分類) スキャンではなく、「マッピング」(マッピングのみ) スキャンのみを実行するように選択することもできます。ファイルにアクセスして内部のデータを確認する必要がないため、データ ソースのマッピングを非常に迅速に実行できます。"マッピングスキャンと分類スキャンの違いを確認する" 。

データ分類を使用してデータを検索できますか?

データ分類には広範な検索機能が用意されており、接続されているすべてのソースから特定のファイルやデータを簡単に検索できます。データ分類により、ユーザーはメタデータが反映するものよりもさらに深く検索できるようになります。これは言語に依存しないサービスであり、ファイルを読み取って、名前や ID などの多数の機密データ タイプを分析することもできます。たとえば、ユーザーは構造化データ ストアと非構造化データ ストアの両方を検索して、企業ポリシーに違反してデータベースからユーザー ファイルに漏洩した可能性のあるデータを見つけることができます。検索は後で使用するために保存することができ、設定された頻度で検索して結果に対してアクションを実行するためのポリシーを作成することもできます。

関心のあるファイルが見つかったら、タグ、システム アカウント、バケット、ファイル パス、カテゴリ (分類から)、ファイル サイズ、最終更新日時、アクセス許可の状態、重複、機密レベル、個人データ、ファイル内の機密データの種類、所有者、ファイルの種類、ファイル サイズ、作成時刻、ファイル ハッシュ、データが注意を求める誰かに割り当てられたかどうかなどの特性を一覧表示できます。フィルターを適用して、関係のない特性を除外することができます。

データ分類には、適切な権限がある場合にファイルを移動または削除できるようにするロールベースのアクセス制御 (RBAC) も備わっています。適切な権限がない場合、タスクは組織内で適切な権限を持つユーザーに割り当てることができます。

データ分類管理とプライバシー

次の質問は、データ分類とプライバシー設定を管理する方法に関する情報を提供します。

データ分類を有効または無効にするにはどうすればよいですか?

まず、コンソールまたはオンプレミス システムにデータ分類のインスタンスをデプロイする必要があります。インスタンスが実行されると、[構成] タブまたは特定のシステムを選択して、既存のシステム、データベース、およびその他のデータ ソースでサービスを有効にできます。"始める方法を学ぶ" 。

> データ ソースでデータ分類をアクティブ化すると、すぐに初期スキャンが実行されます。スキャン結果はすぐに表示されます。

データ分類構成ページから、データ分類による個々のシステム、データベース、またはファイル共有グループのスキャンを無効にすることができます。見る"データ分類からデータソースを削除する" 。

データ分類インスタンスを完全に削除するには、クラウド プロバイダーのポータルまたはオンプレミスの場所からデータ分類インスタンスを手動で削除します。

サービスは特定のディレクトリ内のデータのスキャンを除外できますか?

○データ分類で特定のデータ ソース ディレクトリにあるスキャン データを除外する場合は、そのリストを分類エンジンに提供できます。変更を適用すると、データ分類では指定されたディレクトリ内のスキャン データが除外されます。"詳細情報" 。

ONTAPボリューム上に存在するスナップショットはスキャンされますか?

いいえ。データ分類では、コンテンツがボリューム内のコンテンツと同一であるため、スナップショットはスキャンされません。

ONTAPボリュームでデータ階層化を有効にするとどうなりますか?

データ分類がマッピングのみのスキャンを使用してオブジェクト ストレージに階層化されたコールド データを持つボリュームをスキャンする場合、ローカル ディスク上のデータとオブジェクト ストレージに階層化されたコールド データを含むすべてのデータがスキャンされます。これは、階層化を実装するNetApp以外の製品にも当てはまります。

マッピングのみのスキャンでは、コールド データはヒートアップされず、コールドのままオブジェクト ストレージに残ります。一方、マップと分類スキャンを実行する場合、構成によってはコールド データがヒートアップする可能性があります。

ソースシステムの種類とデータ型

次の質問は、スキャンできるストレージの種類と、スキャンされるデータの種類に関するものです。

政府地域に展開する場合、何か制限はありますか?

データ分類は、コンソール エージェントが政府リージョン (AWS GovCloud、Azure Gov、または Azure DoD) にデプロイされている場合にサポートされます (「制限モード」とも呼ばれます)。

インターネットにアクセスできないサイトにデータ分類をインストールする場合、どのデータ ソースをスキャンできますか?

> BlueXPプライベート モード (レガシーBlueXPインターフェイス) は通常、インターネット接続がなく、AWS Secret Cloud、AWS Top Secret Cloud、Azure IL6 などの安全なクラウド領域があるオンプレミス環境で使用されます。NetApp は、従来のBlueXPインターフェースを使用してこれらの環境を引き続きサポートします。従来のBlueXPインターフェースのプライベートモードのドキュメントについては、"BlueXPプライベートモードの PDF ドキュメント" 。

データ分類では、オンプレミス サイトのローカルにあるデータ ソースからのデータのみをスキャンできます。現時点では、データ分類では、次のローカル データ ソースを「プライベート モード」(「ダーク」サイトとも呼ばれます) でスキャンできます。

- オンプレミスのONTAPシステム
- データベーススキーマ
- Simple Storage Service (S3) プロトコルを使用するオブジェクト ストレージ

どのようなファイル形式がサポートされていますか?

データ分類では、すべてのファイルをスキャンしてカテゴリとメタデータの分析情報を取得し、ダッシュボードのファイル タイプ セクションにすべてのファイル タイプを表示します。

データ分類が個人識別情報 (PII) を検出する場合、または DSAR 検索を実行する場合は、次のファイル形式のみがサポートされます。

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

データ分類ではどのような種類のデータとメタデータをキャプチャしますか?

データ分類を使用すると、データ ソースに対して一般的な「マッピング」スキャンまたは完全な「分類」スキャンを実行できます。マッピングではデータの概要のみが提供されますが、分類ではデータの詳細なスキャンが提供されます。ファイルにアクセスして内部のデータを確認する必要がないため、データ ソースのマッピングを非常に迅速に実行できます。

- データ マッピング スキャン (マッピングのみのスキャン): データ分類ではメタデータのみをスキャンします。これは、全体的なデータ管理とガバナンス、プロジェクトの迅速なスコープ設定、非常に大規模な資産、優先順位付けに役立ちます。データ マッピングはメタデータに基づいており、高速スキャンと見なされます。

  高速スキャンの後、データ マッピング レポートを生成できます。このレポートは、企業のデータ ソースに保存されているデータの概要であり、リソースの使用率、移行、バックアップ、セキュリティ、コンプライアンス プロセスに関する意思決定に役立ちます。

- データ分類ディープ スキャン (マップと分類スキャン): データ分類では、環境全体で標準プロトコルと読み取り専用権限を使用してデータをスキャンします。選択されたファイルが開かれ、ビジネス関連の機密データ、個人情報、およびランサムウェアに関連する問題がないかスキャンされます。

  完全スキャン後には、データ調査ページでのデータの表示と絞り込み、ファイル内の名前の検索、ソース ファイルのコピー、移動、削除など、データに適用できる追加のデータ分類機能が多数あります。

データ分類では、ファイル名、権限、作成時間、最終アクセス、最終変更などのメタデータを取得します。これには、データ調査の詳細ページとデータ調査レポートに表示されるすべてのメタデータが含まれます。

データ分類では、個人情報 (PII) や機密個人情報 (SPII) など、さまざまな種類のプライベート データを識別できます。個人データの詳細については、データ分類がスキャンする個人データのカテゴリ 。

データ分類情報を特定のユーザーに制限できますか?

はい、データ分類はNetApp Consoleと完全に統合されています。 NetApp Consoleユーザーは、権限に応じて表示資格のあるシステムの情報のみを表示できます。

さらに、特定のユーザーにデータ分類設定の管理権限を与えずにデータ分類スキャン結果の表示のみを許可する場合は、そのユーザーに 分類閲覧者 ロール (NetApp Consoleを標準モードで使用する場合) または コンプライアンス閲覧者 ロール (NetApp Consoleを制限モードで使用する場合) を割り当てることができます。"詳細情報" 。

ブラウザとデータ分類の間で送信されるプライベートデータに誰でもアクセスできますか?

いいえ。ブラウザとデータ分類インスタンス間で送信されるプライベート データは、TLS 1.2 を使用したエンドツーエンドの暗号化によって保護されるため、 NetAppおよびNetApp以外の関係者はそれを読み取ることができません。データ分類は、アクセスを要求して承認しない限り、データや結果をNetAppと共有しません。

スキャンされたデータは環境内に残ります。

機密データはどのように扱われますか?

NetApp は機密データにアクセスできず、UI に表示しません。機密データはマスクされます。たとえば、クレジットカード情報の場合は最後の 4 桁が表示されます。

データはどこに保存されますか?

スキャン結果は、データ分類インスタンス内の Elasticsearch に保存されます。

データにはどのようにアクセスしますか?

データ分類は、API 呼び出しを通じて Elasticsearch に保存されているデータにアクセスします。このアクセスには認証が必要であり、AES-128 を使用して暗号化されます。 Elasticsearch に直接アクセスするには、ルート アクセスが必要です。

ライセンスとコスト

次の質問は、データ分類を使用するためのライセンスとコストに関するものです。

データ分類にはどれくらいの費用がかかりますか?

データ分類は、 NetApp Consoleのコア機能です。料金はかかりません。

コンソールエージェントの展開

次の質問はコンソール エージェントに関するものです。

コンソールエージェントとは何ですか?

コンソール エージェントは、クラウド アカウント内またはオンプレミスのコンピューティング インスタンスで実行されるソフトウェアであり、 NetApp Consoleがクラウド リソースを安全に管理できるようにします。データ分類を使用するには、コンソール エージェントを展開する必要があります。

コンソール エージェントはどこにインストールする必要がありますか?

データをスキャンする場合、 NetApp Consoleエージェントを次の場所にインストールする必要があります。

- AWS のCloud Volumes ONTAPまたはAmazon FSx for ONTAPの場合: コンソールエージェントは AWS にあります。

- Azure またはAzure NetApp FilesのCloud Volumes ONTAPの場合: コンソール エージェントは Azure にあります。
- GCP のCloud Volumes ONTAPの場合: コンソール エージェントは GCP にあります。
- オンプレミスのONTAPシステムの場合: コンソール エージェントはオンプレミスです。

これらの場所にデータがある場合は、 "複数のコンソールエージェント" 。

データ分類には資格情報へのアクセスが必要ですか?

データ分類自体はストレージ資格情報を取得しません。代わりに、コンソール エージェント内に保存されます。

データ分類では、スキャン前に共有をマウントするための CIFS 資格情報などのデータ プレーン資格情報を使用します。

サービスとコンソール エージェント間の通信には **HTTP** が使用されますか?

はい、データ分類は HTTP を使用してコンソール エージェントと通信します。

データ分類の展開

次の質問は、個別のデータ分類インスタンスに関連しています。

データ分類はどのような展開モデルをサポートしていますか?

NetApp Consoleを使用すると、オンプレミス、クラウド、ハイブリッド環境など、事実上あらゆる場所のシステムをスキャンしてレポートできます。データ分類は通常、SaaS モデルを使用して展開されます。このモデルでは、サービスはコンソール インターフェイスを介して有効化され、ハードウェアやソフトウェアのインストールは必要ありません。このクリックアンド実行の展開モードでも、データ ストアがオンプレミスにあるかパブリック クラウドにあるかに関係なく、データ管理を行うことができます。

データ分類にはどのようなタイプのインスタンスまたは **VM** が必要ですか?

いつ"クラウドに展開":

- AWS では、データ分類は 500 GiB GP2 ディスクを備えた m6i.4xlarge インスタンスで実行されます。デプロイ時により小さいインスタンス タイプを選択できます。
- Azure では、データ分類は 500 GiB のディスクを備えた Standard_D16s_v3 VM 上で実行されます。
- GCP では、データ分類は 500 GiB の標準永続ディスクを備えた n2-standard-16 VM 上で実行されます。

"データ分類の仕組みについて詳しく見る" 。

データ分類を自分のホストに展開できますか?

○ネットワークまたはクラウドでインターネットにアクセスできる Linux ホストにデータ分類ソフトウェアをインストールできます。すべてが同じように機能し、コンソールを通じてスキャン構成と結果を引き続き管理できます。見る"オンプレミスでのデータ分類の導入"システム要件とインストールの詳細については、こちらをご覧ください。

インターネットにアクセスできない安全なサイトはどうなりますか?

はい、それもサポートされています。あなたはできる"インターネットにアクセスできないオンプレミスサイトにデータ分類を展開する"完全に安全なサイトのために。

法律上の表示

法的通知から、著作権情報、商標、特許などを確認できます。

著作権

"https://www.netapp.com/company/legal/copyright/"

商標

NetApp、NetAppのロゴ、NetAppの商標一覧のページに掲載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。

"https://www.netapp.com/company/legal/trademarks/"

特許

現在NetAppが所有する特許の一覧は以下のページから閲覧できます。

https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf

プライバシー ポリシー

"https://www.netapp.com/company/legal/privacy-policy/"

オープンソース

通知ファイルには、 NetAppソフトウェアで使用されるサードパーティの著作権およびライセンスに関する情報が提供されます。

- "NetApp Consoleに関するお知らせ"
- "NetApp Data Classificationに関する通知"

著作権に関する情報

商標に関する情報