



使用 **Cloud Data sense**

Cloud Data Sense

NetApp
June 15, 2022

目录

- 使用 Cloud Data sense 1
 - 查看有关存储在组织中的数据的监管详细信息 1
 - 查看有关存储在组织中的数据的合规性详细信息 4
 - 组织私有数据 15
 - 管理私有数据 31
 - 查看合规性报告 41
 - 响应数据主体访问请求 49
 - 私有数据的类别 51

使用 Cloud Data sense

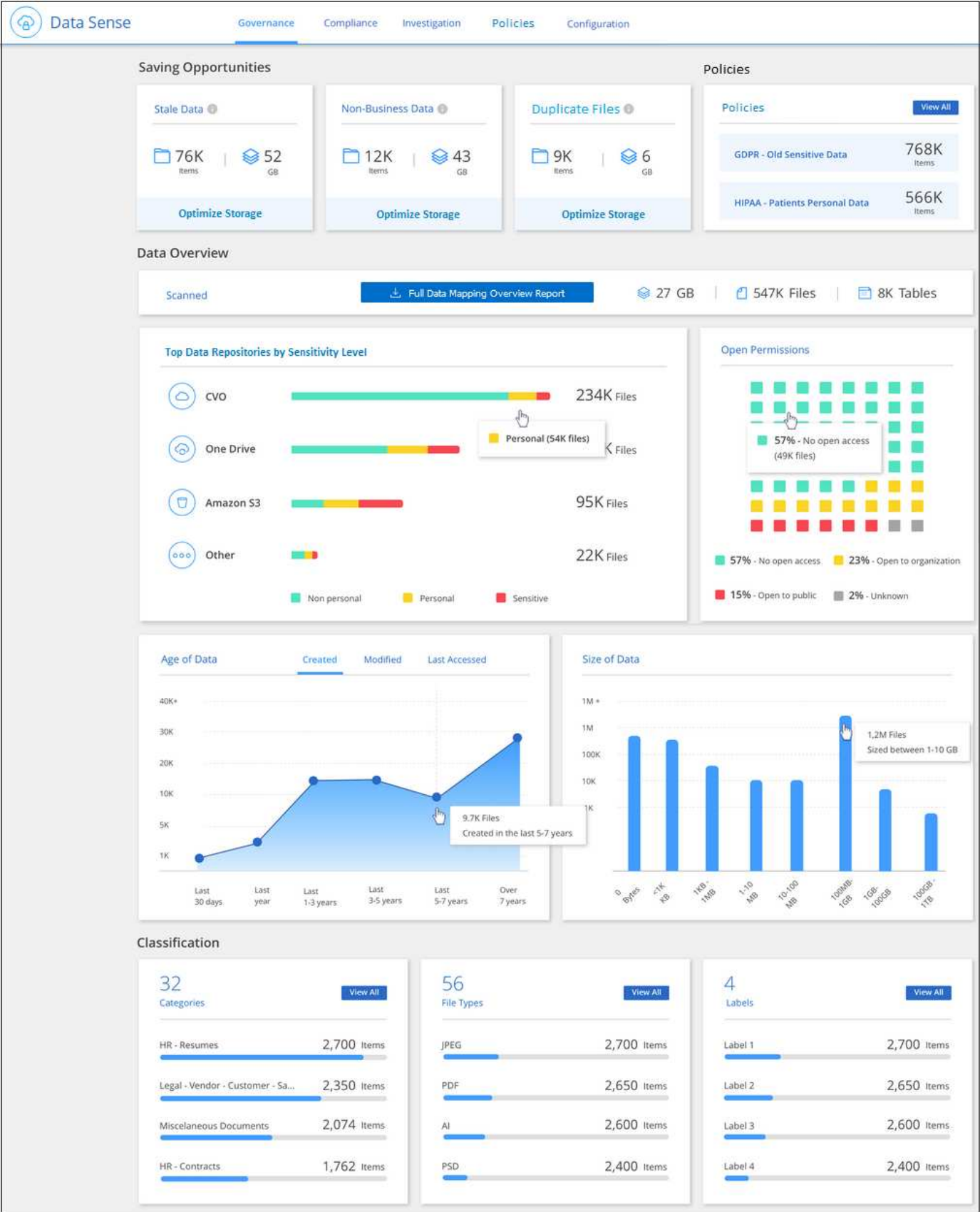
查看有关存储在组织中的数据的监管详细信息

控制与组织存储资源上的数据相关的成本。Cloud Data sense 可确定系统中陈旧数据，非业务数据，重复文件和非常大的文件的数量，以便您可以决定是要删除某些文件，还是将某些文件分层到成本较低的对象存储。

此外，如果您计划将数据从内部位置迁移到云，则可以在移动数据之前查看数据的大小以及任何数据是否包含敏感信息。

监管信息板

" 监管 " 信息板可提供相关信息，以便您提高效率并控制存储资源上存储的数据的相关成本。



节省机会

您可能需要调查 *saving opportunities* 区域中的项目，以确定是否存在任何数据，您应该删除这些数据，或者将

这些数据分层到成本较低的对象存储中。单击每个项可在 " 调查 " 页面中查看筛选后的结果。

- * 陈旧数据 * —上次修改时间超过 3 年的数据。
- * 非业务数据 * - 根据类别或文件类型视为与业务无关的数据。其中包括：
 - 应用程序数据
 - 音频
 - 可执行文件
 - 映像
 - 日志
 - 视频
 - 其他（一般 " 其他 " 类别）
- * 重复文件 * —在要扫描的数据源的其他位置复制的文件。 ["查看显示的重复文件类型"](#)。

结果数量最多的策略

在 *Policy* 区域中单击策略的名称，以在调查页面中显示结果。单击 * 查看全部 * 可查看所有可用策略的列表。

单击 ["此处"](#) 了解有关策略的更多信息。

数据概述

简要概述正在扫描的所有数据。单击按钮可下载完整的数据映射报告，其中包括所有工作环境和数据源的使用容量，数据期限，数据大小和文件类型。请参见 ["数据映射报告"](#) 了解完整的详细信息。

按数据敏感度列出的前几个数据存储库

Top Data Repository by Sensitivity level_ 区域最多可列出包含最敏感项目的前四个数据存储库（工作环境和数据源）。每个工作环境的条形图分为：

- 非个人数据
- 个人数据
- 敏感的个人数据

您可以将鼠标悬停在每个部分上以查看每个类别中的项目总数。

单击每个区域以在 " 调查 " 页面中查看筛选后的结果，以便您可以进一步调查。

按打开权限类型列出的数据

_Open Permissions_area 显示正在扫描的所有文件中存在的每种类型的权限的百分比。此图表显示了以下类型的权限：

- 无开放访问
- 对组织开放
- 打开公有

- 未知访问

您可以将鼠标悬停在每个部分上以查看每个类别中的文件总数。单击每个区域以在 " 调查 " 页面中查看筛选后的结果，以便您可以进一步调查。

数据期限和数据大小图

您可能需要调查 *age* 和 *size* 图中的项目，以确定是否存在任何数据，您应该删除这些数据，或者将这些数据分层到成本较低的对象存储中。

您可以将鼠标悬停在图表中的某个点上，以查看有关该类别中数据的使用期限或大小的详细信息。单击可查看按此期限或大小范围筛选的所有文件。

- * 数据期限图 * —根据数据创建时间，上次访问时间或上次修改时间对数据进行分类。
- 数据图形的大小 * —根据大小对数据进行分类。

最明确的数据分类

_Classification_area 提供了一个标识最多的列表 "类别"，"文件类型"，和 "AIP 标签" 扫描的数据。

类别

类别可以通过向您显示所拥有的信息类型来帮助您了解数据的变化。例如，" 恢复 " 或 " 员工合同 " 等类别可以包含敏感数据。调查结果时，您可能会发现员工合同存储在不安全的位置。然后，您可以更正此问题描述。

请参见 ["按类别查看文件"](#) 有关详细信息 ...

文件类型

查看文件类型有助于控制敏感数据，因为您可能会发现某些文件类型存储不正确。

请参见 ["查看文件类型"](#) 有关详细信息 ...

AIP 标签

如果您已订阅 Azure 信息保护（AIP），则可以通过将标签应用于内容来对文档和文件进行分类和保护。通过查看分配给文件的最常用 AIP 标签，您可以查看文件中最常使用的标签。

请参见 ["AIP 标签"](#) 有关详细信息 ...

查看有关存储在组织中的数据的合规性详细信息

通过查看有关组织中个人数据和敏感个人数据的详细信息来控制您的私有数据。您还可以查看数据中的 Cloud Data 感知的类别和文件类型来获得可见性。



只有在选择对数据源执行完整分类扫描后，才可以使用本节所述的功能。已执行仅映射扫描的数据源不会显示文件级详细信息。

默认情况下，Cloud Data sense 信息板会显示所有工作环境和数据库的合规性数据。



如果只想查看某些工作环境的数据， [选择这些工作环境](#)。

您还可以从 " 数据调查 " 页面筛选结果，并以 CSV 文件的形式下载结果报告。请参见 [筛选 " 数据调查 " 页面中的数据](#) 了解详细信息。

查看包含个人数据的文件

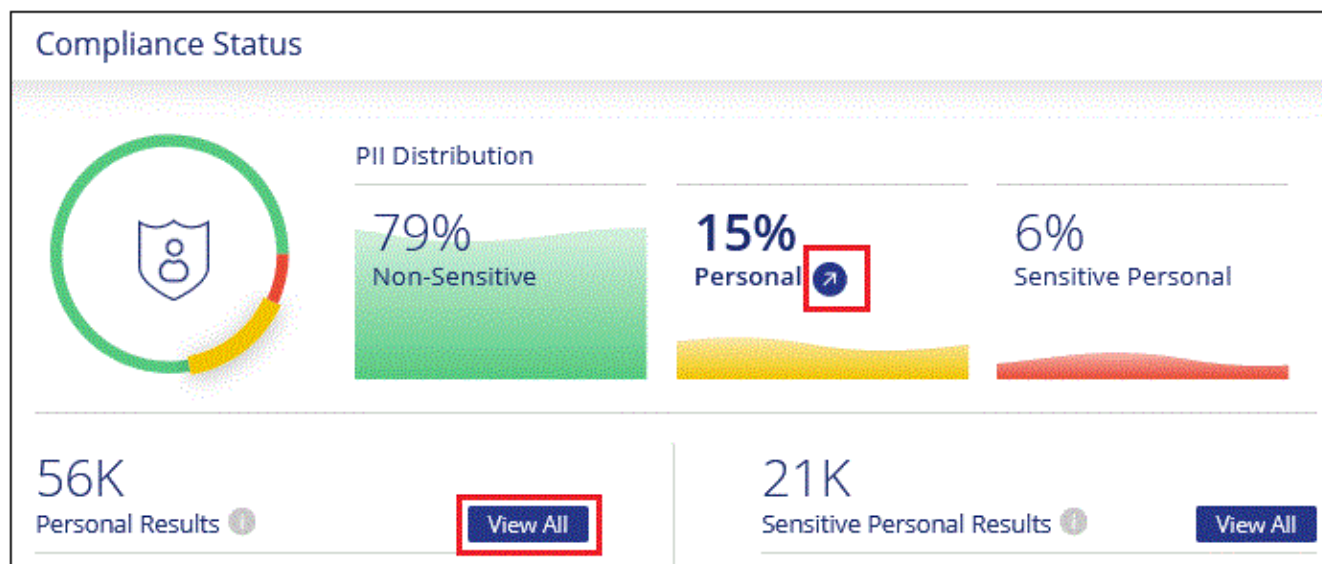
Cloud Data sense 可自动识别数据中的特定字词，字符串和模式（Reex）。例如，个人身份识别信息（PII），信用卡号，社会保险号，银行帐号，密码，等等。 ["请参见完整列表"](#)。

此外，如果您添加了要扫描的数据库服务器，则 _Data Fusion 功能允许您扫描文件，以确定这些文件或其他数据库中是否存在数据库中的唯一标识符。请参见 ["使用 Data Fusion 添加个人数据标识符"](#) 了解详细信息。

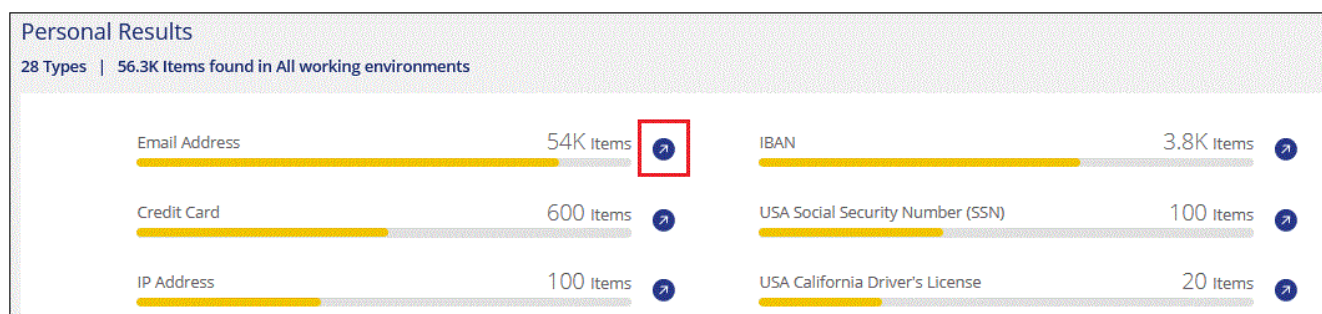
对于某些类型的个人数据，Data sense 使用 *proximity validation* 来验证其结果。通过查找与所发现的个人数据相邻的一个或多个预定义关键字来进行验证。例如，Data sense 可识别美国如果在 SSN 旁边看到一个接近词，则将社会安全号（SSN）作为 SSN，例如 _SSN 或 _社交 安全_。 ["个人数据表"](#) 显示 Data sense 何时使用接近验证。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *，然后单击 * 合规性 * 选项卡。
2. 要调查所有个人数据的详细信息，请单击个人数据百分比旁边的图标。



3. 要调查特定类型的个人数据的详细信息，请单击 * 全部查看 *，然后单击特定类型的个人数据的 * 调查结果 * 图标；例如，电子邮件地址。



4. 通过搜索，排序，展开特定文件的详细信息，单击 * 调查结果 * 以查看屏蔽信息或下载文件列表来调查数据。



查看包含敏感个人数据的文件

Cloud Data sense 可自动识别隐私法规所定义的特殊类型的敏感个人信息 "《一般数据保护条例》第 9 条和第 10 条"。例如，有关个人健康，种族或性取向的信息。"请参见完整列表"。

Cloud Data sense 利用人工智能（AI），自然语言处理（NLP），机器学习（ML）和认知计算（CC）来了解其扫描内容的含义，以便提取实体并相应地进行分类。

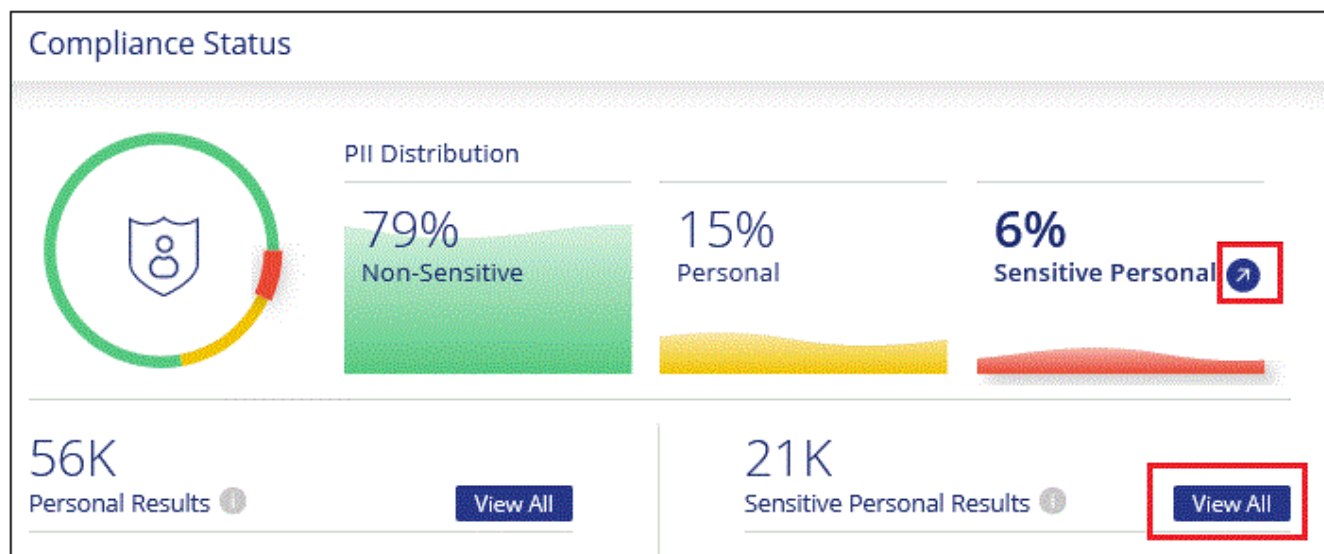
例如，一个敏感的 GDPR 数据类别是种族起源。由于其 NLP 能力，Data sense 可以区分 "George is Mexican" 一词（表示《一般数据保护条例》第 9 条规定的敏感数据）与 "George is eating Mexican food" 这两者之间的区别。



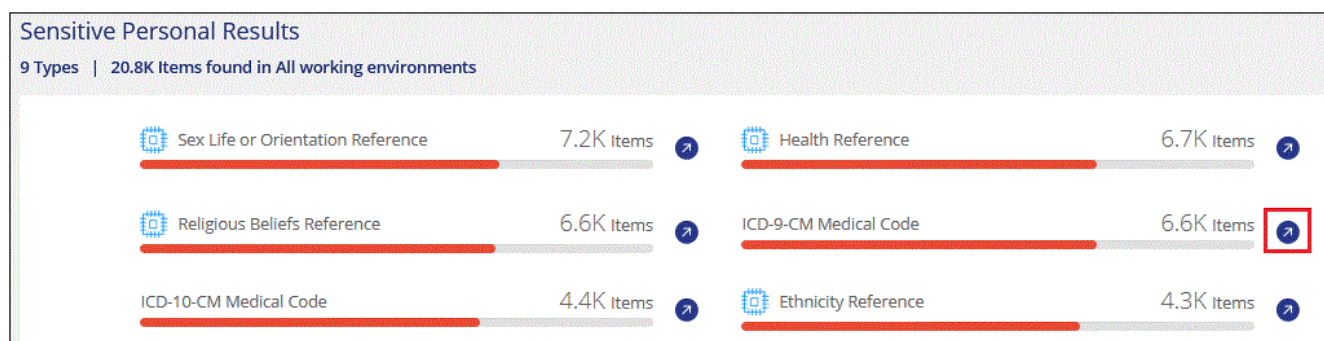
扫描敏感个人数据时，仅支持英语。稍后将添加对更多语言的支持。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *，然后单击 * 合规性 * 选项卡。
2. 要调查所有敏感个人数据的详细信息，请单击敏感个人数据百分比旁边的图标。



- 要调查特定类型的敏感个人数据的详细信息，请单击 * 查看全部 *，然后单击特定类型的敏感个人数据的 * 调查结果 * 图标。



- 通过搜索，排序，展开特定文件的详细信息，单击 * 调查结果 * 以查看屏蔽信息或下载文件列表来调查数据。

按类别查看文件

Cloud Data sense 会将扫描的数据划分为不同类型的类别。类别是基于 AI 对每个文件的内容和元数据的分析而得出的主题。"查看类别列表"。

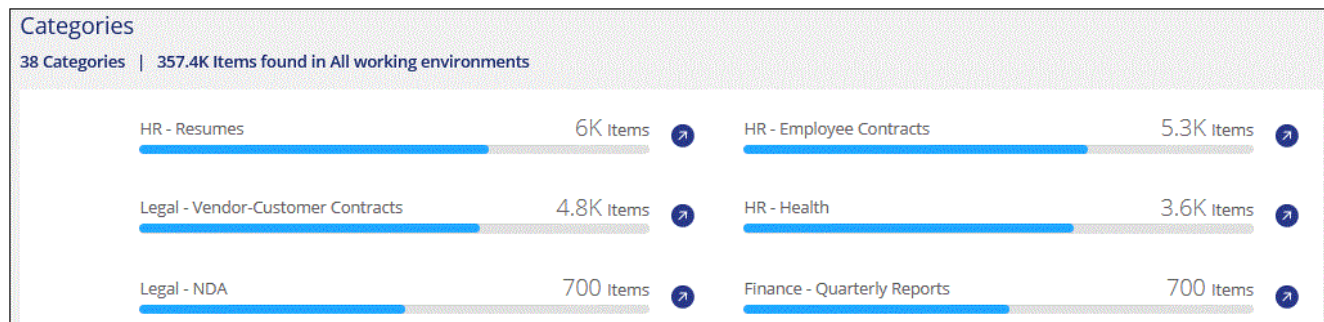
类别可以通过向您显示所拥有的信息类型来帮助您了解数据的变化。例如，恢复或员工合同等类别可以包含敏感数据。调查结果时，您可能会发现员工合同存储在不安全的位置。然后，您可以更正此问题描述。



类别支持英语，德语和西班牙语。稍后将添加对更多语言的支持。

步骤

- 在 Cloud Manager 顶部，单击 * 数据感知 *，然后单击 * 合规性 * 选项卡。
- 直接从主屏幕中单击前 4 个类别之一的 * 调查结果 * 图标，或者单击 * 查看全部 *，然后单击任意类别的图标。



3. 通过搜索，排序，展开特定文件的详细信息，单击 * 调查结果 * 以查看屏蔽信息或下载文件列表来调查数据。

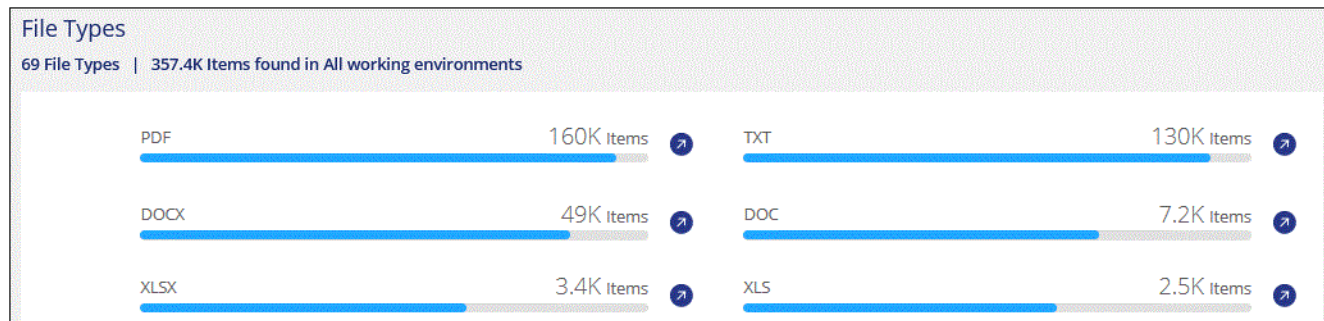
按文件类型查看文件

Cloud Data sense 会提取所扫描的数据，并按文件类型对其进行细分。查看文件类型有助于控制敏感数据，因为您可能会发现某些文件类型存储不正确。 ["请参见文件类型列表"](#)。

例如，您可能存储的 CAD 文件包含有关您的组织的非常敏感的信息。如果这些文件不安全，您可以通过限制权限或将文件移动到其他位置来控制敏感数据。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *，然后单击 * 合规性 * 选项卡。
2. 直接从主屏幕中单击前 4 种文件类型之一的 * 调查结果 * 图标，或者单击 * 查看全部 *，然后单击任意文件类型的图标。



3. 通过搜索，排序，展开特定文件的详细信息，单击 * 调查结果 * 以查看屏蔽信息或下载文件列表来调查数据。

查看文件元数据

在数据调查结果窗格中，您可以单击  用于查看文件元数据的任何单个文件。

The screenshot shows a file management interface with a top bar displaying '364.9K items' and action buttons like 'Tags', 'Assign to', 'Label', 'Move', 'Copy', and 'Delete'. Below this is a table with columns: 'File Name', 'Personal', 'Sensitive Personal', 'Data Subjects', and 'File Type'. Two files are listed: 'ground truth.xlsx' and 'GM_PD 12-1-09 SP.xls.pdf'. The second file is selected, and its details are shown in a modal window. The details include: Tags (Decathlon, gidi, IS NOT OK, And 6 more, View All), Working Environment (OneDrive daylabs.onmicrosoft.com), Storage Repository (User: ruh@daylabs.onmicrosoft.com), File Path (/scattered/26/GM_PD 12-1-09 SP.xls.pdf), Category (Miscellaneous Documents), File Size (427.46 KB), Discovered Time (2021-01-12 10:37), Created Time (2018-05-22 12:38), Last Modified (2018-10-22 13:28), and Duplicates (None). On the right side of the modal, there are buttons for 'Tags: 9 tags', 'Assigned to: Amit Ashbel', 'Assign a Label to this file', 'Copy File', 'Move File', and 'Delete File'. A red box highlights a back arrow icon in the top right corner of the modal.

除了向您显示文件所在的工作环境和卷之外，元数据还会显示更多信息，包括文件权限，文件所有者，此文件是否存在重复项以及分配的 AIP 标签（如果有）"[在云数据感知中集成 AIP](#)"。如果您计划使用此信息，此信息将非常有用 "[创建策略](#)" 因为您可以看到可用于筛选数据的所有信息。

请注意，并非所有信息都可用于所有数据源—只是适合该数据源的信息。例如，卷名称，权限和 AIP 标签与数据库文件无关。

查看单个文件的详细信息时，可以对该文件执行以下操作：

- 您可以将此文件移动或复制到任何 NFS 共享。请参见 "[将源文件移动到 NFS 共享](#)" 和 "[将源文件复制到 NFS 共享](#)" 了解详细信息。
- 您可以删除此文件。请参见 "[正在删除源文件](#)" 了解详细信息。
- 您可以为文件分配特定状态。请参见 "[正在应用标记](#)" 了解详细信息。
- 您可以将此文件分配给 Cloud Manager 用户，以负责对该文件执行任何后续操作。请参见 "[为文件分配用户](#)" 了解详细信息。
- 如果您已将 AIP 标签与 Cloud Data sense 集成在一起，则可以为此文件分配一个标签，或者如果已存在另一个标签，则可以更改为其他标签。请参见 "[手动分配 AIP 标签](#)" 了解详细信息。

查看文件的权限

要查看有权访问文件的所有用户或组的列表及其权限类型，请单击 * 查看所有权限 *。此按钮仅适用于 CIFS 共享、SharePoint 和 OneDrive 中的数据。

请注意，如果您看到的是 SID（安全标识符），而不是用户和组名称，则应将 Active Directory 集成到 Data sense 中。"[了解如何执行此操作](#)"。

File Name

Personal Sensitive Personal Data Subjects File Type

Expense Report TPO-1060.pdf

cvo 6 3 16 PDF

Working Environment: WorkingEnvironment1

Repository: Volume Name

File Path: /Prod/labs-base/Expense Report TPO-1060.pdf

Category: Legal

File Size: 22 MB

Last Modified: 2019-08-06 07:51

Open Permissions: NO OPEN PERMISSIONS View all Permissions

File Owner: Avy

Assign a Label to this file

Delete this file

Permissions list for "Expense Report TPO-1060.pdf"

User / Group	Name	Read	Write
User Name		✓	✓
Group Name		✓	✓
Group Name		✓	✓
John L		✓	✓
George H		✓	✓
Paul M		✓	✓
Ringo S		✓	✓

您可以单击  查看属于该组的用户列表。

此外、您还可以单击用户或组的名称、此时将显示调查页面、并在"用户/组权限"筛选器中填充该用户或组的名称、以便您可以查看该用户或组有权访问的所有文件。

检查存储系统中是否存在重复文件

您可以查看存储系统中是否存储了重复的文件。如果您要确定可节省存储空间区域，此功能非常有用。此外，确保具有特定权限或敏感信息的某些文件不会在存储系统中进行不必要的复制也会很有帮助。

Data sense 使用哈希技术来确定重复文件。如果任何文件与另一个文件具有相同的哈希代码，我们可以 100% 确保这些文件完全重复，即使文件名不同也是如此。

您可以下载重复文件列表并将其发送给存储管理员，以便他们确定可以删除哪些文件（如果有）。您也可以 [删除文件](#) 如果您确信不需要特定版本的文件，请自行选择。

查看所有重复的文件

如果您需要列出在工作环境中复制的所有文件以及正在扫描的数据源，则可以在 "数据调查" 页面中使用名为 * 重复项 > 具有重复项 * 的筛选器。

结果页面将显示所有文件类型（不包括数据库）中的重复文件，最小大小为 50 MB 且 / 或包含个人或敏感个人信息。

查看特定文件是否重复

如果要查看单个文件是否存在重复项，可以在数据调查结果窗格中单击  用于查看文件元数据的任何单个文件。如果某个文件存在重复项，则此信息将显示在 *Duplicates* 字段旁边。

要查看重复文件的列表及其所在位置，请单击 * 查看详细信息 *。在下一页中，单击 * 查看重复项 * 以查看调查

页面中的文件。

Last Modified: 2019-08-06 07:51

Open Permissions: NO OPEN PERMISSIONS [View all Permissions](#)

File Owner: Asaf Ley

Duplicates: 3 [View Details](#)

Duplicates of File 'Name 1'

Duplicates: 3

Total Size of all Duplicates: 1GB

File Hash: xxxxxx

[View Duplicates](#)

[Close](#)

3 items

<input type="checkbox"/>	File Name		Personal	Sensitive Personal	Data Subjects	File Type	
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼



您可以随时使用此页面中提供的 " 文件哈希 " 值并直接在 " 调查 " 页面中输入此值以搜索特定的重复文件，也可以在策略中使用此值。

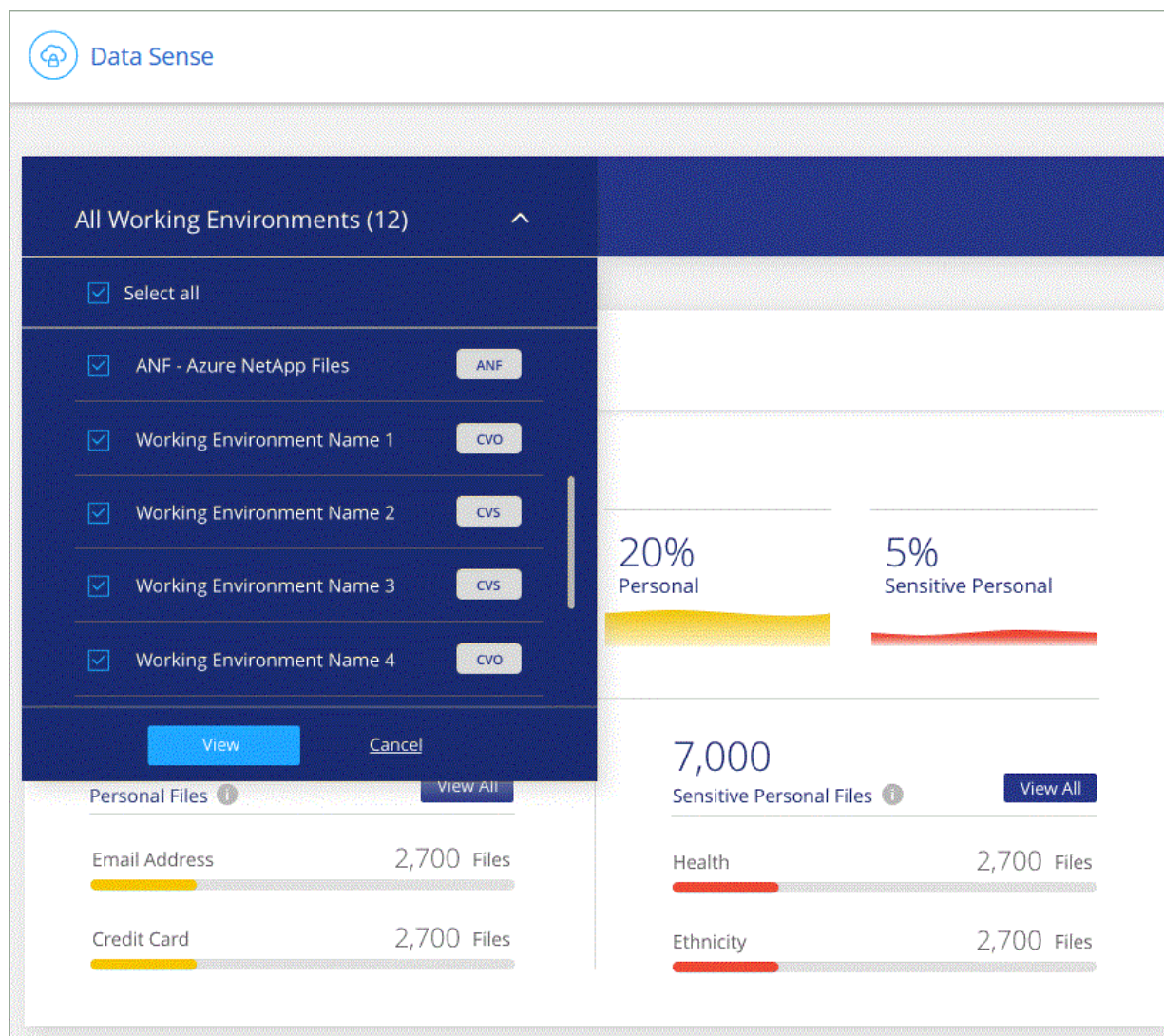
查看特定工作环境的信息板数据

您可以筛选 Cloud Data sense 信息板的内容，以查看所有工作环境和数据库的合规性数据，或者仅查看特定工作环境的合规性数据。

筛选信息板时，Data sense 会将合规性数据和报告范围仅限于您选择的工作环境。

步骤

1. 单击筛选器下拉列表，选择要查看其数据的工作环境，然后单击 * 查看 *。



筛选“数据调查”页面中的数据

您可以筛选调查页面的内容，以便仅显示要查看的结果。这是一项功能非常强大的功能、因为在对数据进行细化后、您可以使用页面顶部的按钮栏执行各种操作、包括复制文件、移动文件、向文件添加标记或AIP标签等。

如果要在优化页面内容后将其作为报告下载、请单击 [↓](#) 按钮。您可以将报告本地保存为.CSV文件(最多可包含5、000行数据)或导出到NFS共享的.json文件(最多可包含不限数量的行)。["有关数据调查报告的详细信息、请访问此处"](#)。

Data Investigation

Unstructured (364K Files)

Directories (64 Folders)

Structured (45 Tables)

Search by file or DB table

FILTERS:

Clear All

Policies

+

Open Permissions

+

File Owner

+

Label

+

Working Environment Type

2

+

Working Environment

+

Storage Repository

2

+

364K items

Tags

Assign to

Label

Move

Copy

Delete

<div><input type="checkbox"/></div>	File Name		Personal	Sensitive Personal	Data Subjects	File Type	
<input type="checkbox"/>	cgdpr_yes_adam.txt	ANF	0	797	111	TXT	
<input type="checkbox"/>	cgdpr_yes_adam.txt	ANF	0	797	111	TXT	
<input type="checkbox"/>	true positive.txt	ANF	0	611	111	TXT	
<input type="checkbox"/>	cgdpr_yes_adam.txt	ANF	0	611	111	TXT	
<input type="checkbox"/>	true positive.txt	ANF	0	611	111	TXT	
<input type="checkbox"/>	true positive.txt	ANF	0	611	111	TXT	
<input type="checkbox"/>	cgdpr_yes_adam.txt	ANF	0	611	111	TXT	
<input type="checkbox"/>	cgdpr_yes_adam.txt	ANF	0	611	111	TXT	

- 通过顶层选项卡、您可以查看文件(非结构化数据)、目录(文件夹和文件共享)或数据库(结构化数据)中的数据。
- 通过每列顶部的控件，您可以按数字或字母顺序对结果进行排序。
- 您可以通过左窗格筛选器从以下属性中进行选择来细化结果：

筛选器	详细信息
策略	选择一个或多个策略。开始 "此处" 可查看现有策略列表并创建您自己的自定义策略。
打开权限	选择数据以及文件夹/共享中的权限类型
用户 / 组权限	选择一个或多个用户名和/或组名称、或者输入部分名称
文件所有者	输入文件所有者名称
Label	选择 ... "AIP 标签" 分配给文件的
工作环境类型	选择工作环境的类型。OneDrive、SharePoint和Google Drive归类为"云应用程序"。
工作环境名称	选择特定的工作环境
存储库	选择存储库，例如卷或模式
文件路径	输入部分路径或完整路径
类别	选择 "类别类型"
敏感度级别	选择敏感度级别
标识符数量	选择每个文件检测到的敏感标识符范围。包括个人数据和敏感个人数据。
个人数据	选择 "个人数据的类型"
敏感个人数据	选择 "敏感个人数据的类型"
数据主题	输入数据主体的全名或已知标识符

筛选器	详细信息
目录类型	选择目录类型："共享"或"文件夹"
文件类型	选择 "文件类型"
文件大小	选择文件大小范围
创建时间	创建文件时选择一个范围
发现时间	选择 Data sense 发现文件时的范围
上次修改时间	选择上次修改文件的时间范围
上次访问	选择上次访问文件的时间范围。对于 Data sense 扫描的文件类型，这是 Data sense 上次扫描该文件的时间。
重复	选择是否在存储库中复制文件
文件哈希	输入文件的哈希以查找特定文件，即使名称不同也是如此
Tags	选择 ... "一个或多个标记" 分配给文件的
已分配给	选择将文件分配到的人员的姓名

请注意、"目录"级别当前不支持按钮栏和策略中提供的操作。

组织私有数据

Cloud Data sense 为您提供了多种管理和组织私有数据的方式。这样可以更轻松地查看对您最重要的数据。

- 如果您已订阅 "Azure 信息保护（AIP）" 要对文件进行分类和保护，您可以使用 Cloud Data sense 管理这些 AIP 标签。
- 您可以将标记添加到要标记用于组织或某种类型的跟进的文件中。
- 您可以将一个 Cloud Manager 用户分配给一个特定文件或多个文件，以便此用户可以负责管理此文件。
- 使用 "策略" 功能，您可以创建自己的自定义搜索查询，以便通过单击一个按钮轻松查看结果。
- 当某些关键策略返回结果时，您可以向 Cloud Manager 用户发送电子邮件警报。



只有在选择对数据源执行完整分类扫描后，才可以使用本节所述的功能。已执行仅映射扫描的数据源不会显示文件级详细信息。

是否应使用标记或标签？

以下是对数据感知标记和 Azure 信息保护标记的比较。

Tags	标签
文件标记是 Data sense 的一个集成部分。	要求您已订阅 Azure 信息保护（AIP）。
此标记仅保留在 Data sense 数据库中，不会写入到文件中。它不会更改文件或文件访问或修改时间。	此标签是文件的一部分，当此标签发生更改时，此文件将发生更改。此更改还会更改文件的访问和修改时间。

Tags	标签
一个文件可以包含多个标记。	一个文件可以有一个标签。
此标记可用于内部数据感知操作，例如复制，移动，删除，运行策略，等	可以读取文件的其他系统可以看到标签，可用于实现额外的自动化。
仅使用一个 API 调用来查看文件是否具有标记。	

使用 AIP 标签对数据进行分类

如果您已订阅，则可以管理 Cloud Data sense 正在扫描的文件中的 AIP 标签 ["Azure 信息保护（AIP）"](#)。AIP 允许您通过将标签应用于内容来对文档和文件进行分类和保护。使用 Data sense，您可以查看已分配给文件的标签，向文件添加标签以及在标签已存在时更改标签。

Cloud Data sense 支持以下文件类型中的 AIP 标签：.DOC，.docx，.PDF，.PPTX，.XLS，.XLSX。



- 您当前无法更改大于 30 MB 的文件中的标签。对于 OneDrive、SharePoint 和 Google Drive 帐户、最大文件大小为 4 MB。
- 如果某个文件的标签不再存在于 AIP 中，则 Cloud Data sense 会将其视为一个无标签的文件。
- 如果您已将 Data sense 实例部署在无法访问 Internet 的内部位置（也称为非公开站点），则 AIP 标签功能将不可用。

在工作空间中集成 AIP 标签

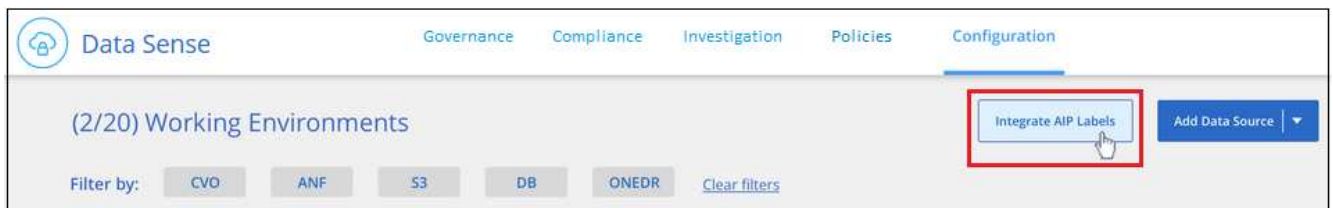
在管理 AIP 标签之前，您需要通过登录到现有 Azure 帐户将 AIP 标签功能集成到 Cloud Data sense 中。启用后，您可以管理所有文件中的 AIP 标签 ["工作环境和数据源"](#) 在 Cloud Manager 工作空间中。

要求

- 您必须拥有帐户和 Azure 信息保护许可证。
- 您必须具有 Azure 帐户的登录凭据。
- 如果您计划更改 Amazon S3 存储分段中文件的标签，请确保 IAM 角色中包含权限 `s3:PutObject`。请参见 ["设置 IAM 角色"](#)。

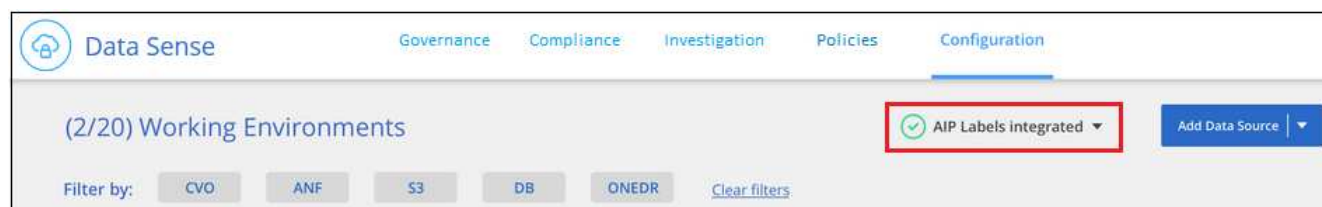
步骤

1. 在 "Cloud Data sense Configuration" 页面中，单击 * 集成 AIP 标签 *。



2. 在集成 AIP 标签对话框中，单击 * 登录到 Azure *。
3. 在显示的 Microsoft 页面中，选择帐户并输入所需的凭据。
4. 返回到 Cloud Data sense 选项卡，您将看到消息 *"AIP Labels were successfully Integrated with the account <account_name>"*。

5. 单击 * 关闭 *，您将在页面顶部看到文本 _AIP Labels Integrated _。



您可以从调查页面的结果窗格中查看和分配 AIP 标签。您还可以使用策略为文件分配 AIP 标签。

查看文件中的 **AIP** 标签

您可以查看分配给文件的当前 AIP 标签。

在数据调查结果窗格中，单击  文件以展开文件元数据详细信息。



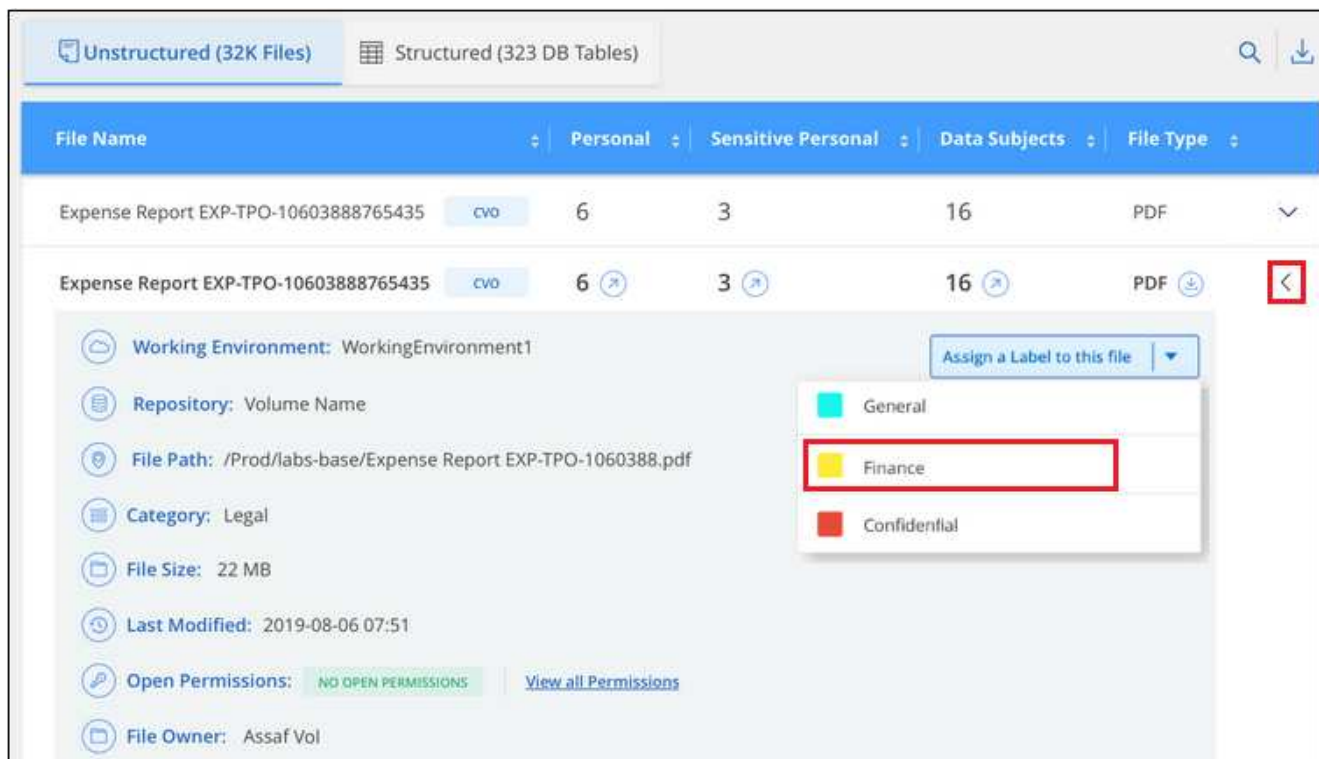
手动分配 **AIP** 标签

您可以使用 Cloud Data sense 在文件中添加，更改和删除 AIP 标签。

按照以下步骤为单个文件分配 AIP 标签。

步骤

1. 在数据调查结果窗格中，单击  文件以展开文件元数据详细信息。



2. 单击 * 为此文件分配标签 * ，然后选择此标签。

此标签将显示在文件元数据中。

为多个文件分配 AIP 标签：

步骤

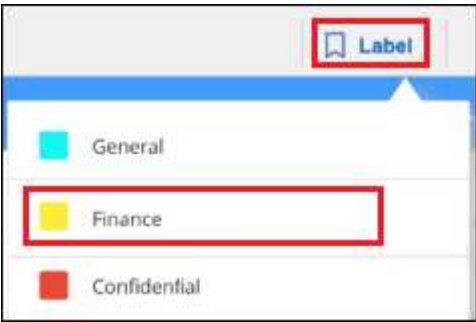
1. 在数据调查结果窗格中，选择要标记的一个或多个文件。



◦ 要选择单个文件，请选中每个文件（☒ Volume_1）。

◦ 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。

2. 从按钮栏中，单击 * 标签 * 并选择 AIP 标签：



AIP 标签将添加到所有选定文件的元数据中。

使用策略自动分配 **AIP** 标签

您可以为符合策略标准的所有文件分配 AIP 标签。您可以在创建策略时指定 AIP 标签，也可以在编辑任何策略时添加此标签。

在 Cloud Data sense 扫描文件时，标签会在文件中持续添加或更新。

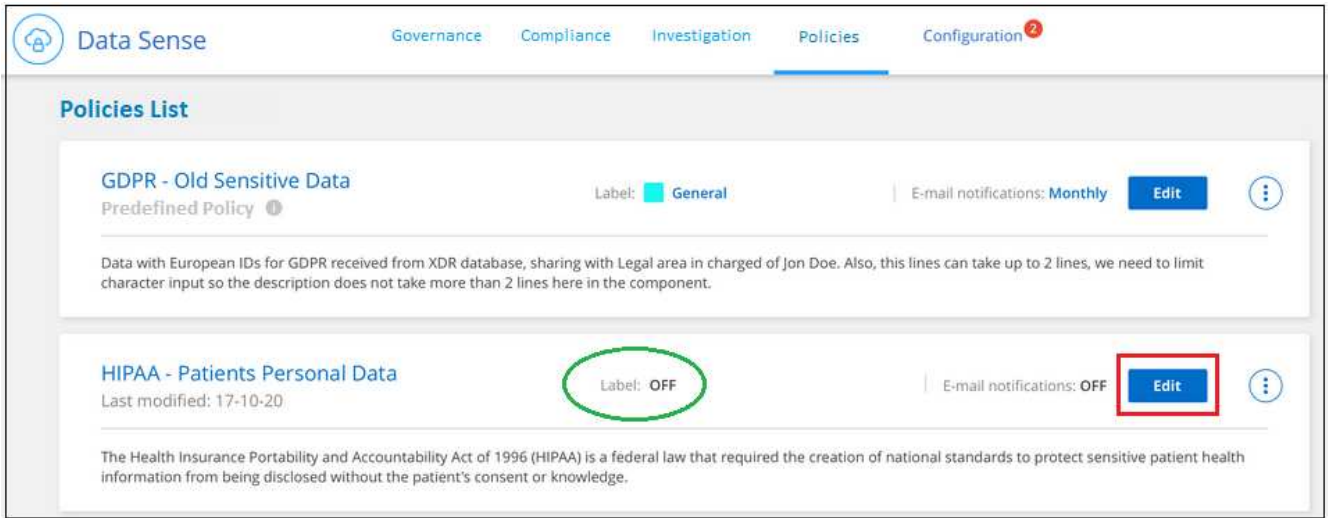
根据标签是否已应用于文件以及标签的分类级别，更改标签时会执行以下操作：

如果文件 ...	那么 ...
无标签	此时将添加此标签
具有较低分类级别的现有标签	此时将添加更高级别的标签
具有较高级别分类的现有标签	较高级别的标签将保留
手动和通过策略为其分配一个标签	此时将添加更高级别的标签
通过两个策略为其分配两个不同的标签	此时将添加更高级别的标签

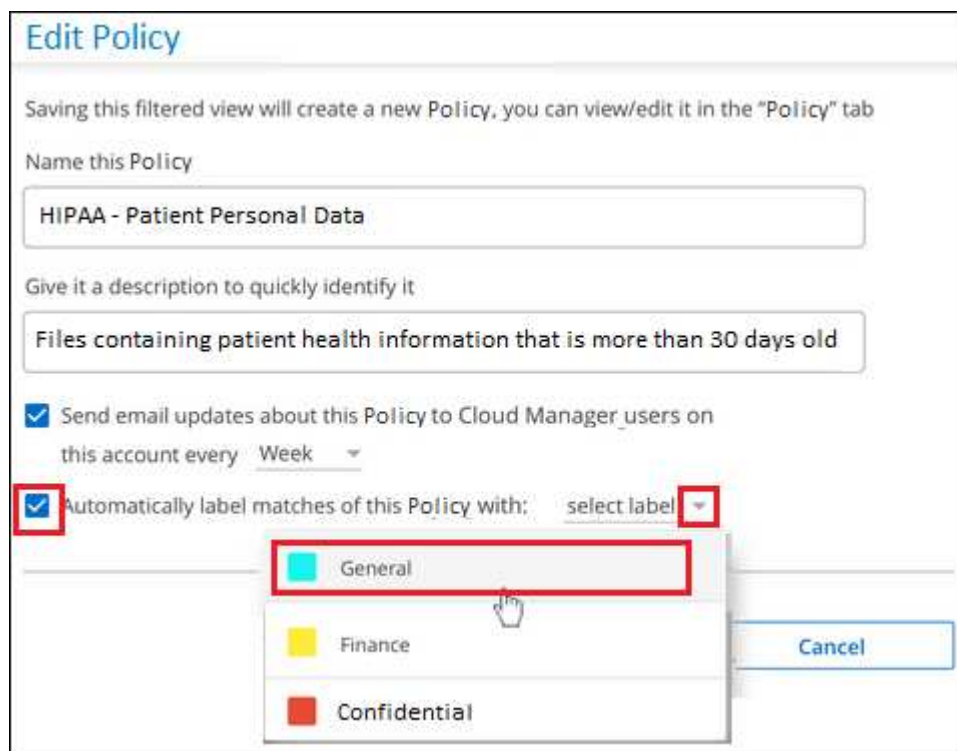
按照以下步骤向现有策略添加 AIP 标签。

步骤

1. 在策略列表页面中，单击要添加（或更改） AIP 标签的策略的 * 编辑 * 。



2. 在编辑策略页面中，选中复选框为与策略参数匹配的文件启用自动标签，然后选择标签（例如 * 常规 *）。



3. 单击 * 保存策略 *，此标签将显示在策略问题描述中。



如果为某个策略配置了标签，但此后已从 AIP 中删除了该标签，则该标签名称将变为关闭，并且不再分配该标签。

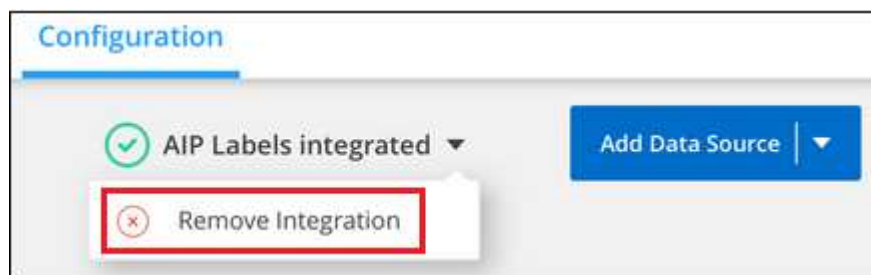
删除 AIP 集成

如果您不再希望能够管理文件中的 AIP 标签，则可以从云数据感知界面中删除 AIP 帐户。

请注意，您使用 Data sense 添加的标签不会进行任何更改。文件中存在的标签将保持当前存在的状态。

步骤

1. 从 *Configuration* 页面中，单击 * 集成 AIP 标签 > 删除集成 *。



2. 从确认对话框中单击 * 删除集成 *。

应用标记以管理扫描的文件

您可以向要标记为某种类型的跟进的文件添加标记。例如，您可能已发现一些重复文件，并且希望删除其中一个文件，但需要检查应删除哪个文件。您可以向文件添加一个标记 "Check to delete"，以使您知道此文件需要进行一些研究并在未来执行某种类型的操作。

使用 Data sense，您可以查看分配给文件的标记，在文件中添加或删除标记，以及更改名称或删除现有标记。

请注意，标记添加到文件中的方式与 AIP 标签是文件元数据的一部分不同。使用 Cloud Data sense 的 Cloud Manager 用户可以看到此标记，因此您可以查看是否需要删除文件或检查文件以进行某种类型的跟进。

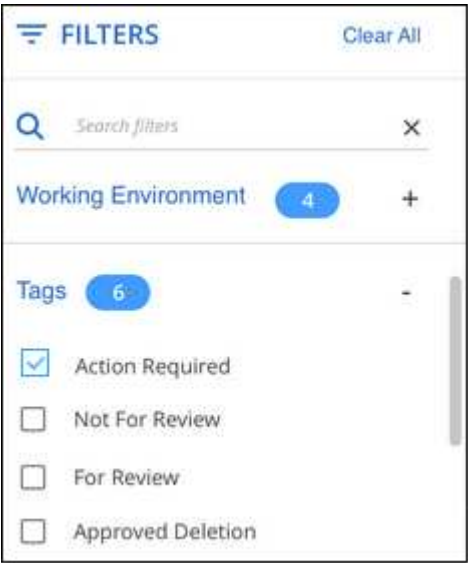


在 Cloud Data sense 中分配给文件的标记与您可以添加到资源（例如卷或虚拟机实例）的标记无关。数据感知标记在文件级别应用。

查看应用了特定标记的文件

您可以查看已分配特定标记的所有文件。

1. 单击 Cloud Data sense 中的 * 调查 * 选项卡。
2. 在数据调查页面中，单击筛选器窗格中的 * 标记 *，然后选择所需的标记。




" 调查结果 " 窗格将显示已分配这些标记的所有文件。

为文件分配标记

您可以向单个文件或一组文件添加标记。

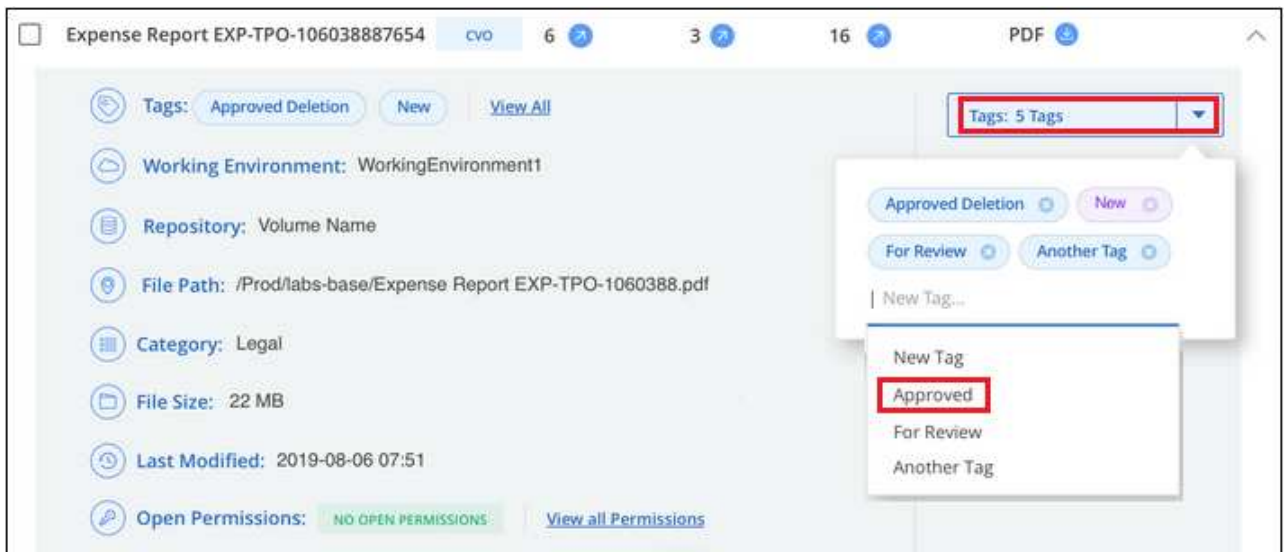
向单个文件添加标记：

步骤

1. 在数据调查结果窗格中，单击  文件以展开文件元数据详细信息。
2. 单击 * 标记 * 字段，此时将显示当前已分配的标记。

3. 添加一个或多个标记：

- 要分配现有标记，请单击 * 新标记 ...* 字段，然后开始键入标记的名称。出现要查找的标记时，请选择该标记并按 * 输入 *。
- 要创建新标记并将其分配给文件，请单击 * 新标记 ...* 字段，输入新标记的名称，然后按 * 输入 *。



此标记将显示在文件元数据中。

向多个文件添加标记：

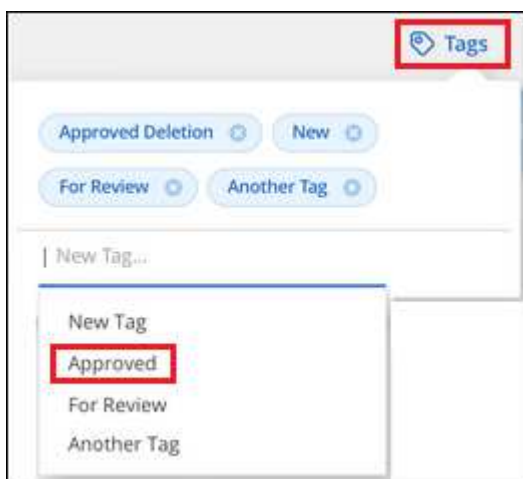
步骤

1. 在数据调查结果窗格中，选择要标记的一个或多个文件。

2345 items							Tags	Assign to	Label	Copy	Move	Delete
<input type="checkbox"/>	File Name	Personal	Sensitive Personal	Data Subjects	File Type							
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF						
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF						
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF						
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF						

- 要选择单个文件，请选中每个文件（☒ Volume_1）。
 - 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。
2. 在按钮栏中，单击 * 标记 *，此时将显示当前已分配的标记。
 3. 添加一个或多个标记：
 - 要分配现有标记，请单击 * 新标记 ...* 字段，然后开始键入标记的名称。出现要查找的标记时，请选择该标记并按 * 输入 *。

- 要创建新标记并将其分配给文件，请单击 * 新标记 ...* 字段，输入新标记的名称，然后按 * 输入 *。



4. 批准在确认对话框中添加标记，标记将添加到所有选定文件的元数据中。

从文件中删除标记

如果不再需要使用某个标记，可以将其删除。

只需单击现有标记的 * x * 即可。



如果选择了多个文件，则标记将从所有文件中删除。

分配用户以管理某些文件

您可以将一个 Cloud Manager 用户分配给一个特定文件或多个文件，以便此用户可以负责对该文件执行任何后续操作。此功能通常与功能结合使用，用于向文件添加自定义状态标记。


例如，您的文件可能包含某些个人数据，这些数据允许过多的用户进行读写访问（打开权限）。因此，您可以将状态标记 " 更改权限 " 并将此文件分配给用户 "Joan Smith"，以便用户确定如何修复问题描述。修复问题描述后，他们可以将状态标记更改为 " 已完成 "。

请注意，用户名不会作为文件元数据的一部分添加到文件中， Cloud Manager 用户只会在使用 Cloud Data sense 时看到此用户名。

通过 " 调查 " 页面中的新筛选器，您可以轻松查看 " 已分配给 " 字段中具有相同人员的所有文件。

要将用户分配给单个文件，请执行以下操作：

步骤

1. 在数据调查结果窗格中，单击  文件以展开文件元数据详细信息。
2. 单击 * 已分配给 * 字段并选择用户名。



用户名显示在文件元数据中。

要将用户分配给多个文件，请执行以下操作：

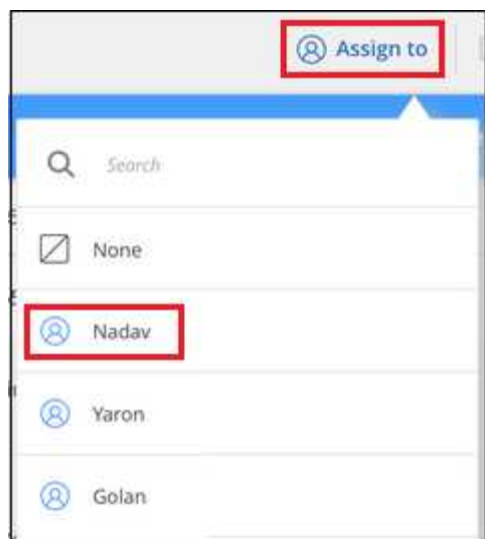
步骤

1. 在数据调查结果窗格中，选择要分配给用户的一个或多个文件。

2345 items							Tags	Assign to	Label	Copy	Move	Delete
<input type="checkbox"/>	File Name		Personal	Sensitive Personal	Data Subjects	File Type						
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-10603887654	cvo	6	3	16	PDF						
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-10603887654	cvo	6	3	6	PDF						
<input type="checkbox"/>	Expense Report EXP-TPO-10603887654	cvo	6	3	6	PDF						
<input type="checkbox"/>	Expense Report EXP-TPO-10603887654	cvo	6	3	6	PDF						

- 要选择单个文件，请选中每个文件（☒ Volume_1）。
- 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。

2. 从按钮栏中，单击 * 分配给 * 并选择用户名：



用户将添加到所有选定文件的元数据中。

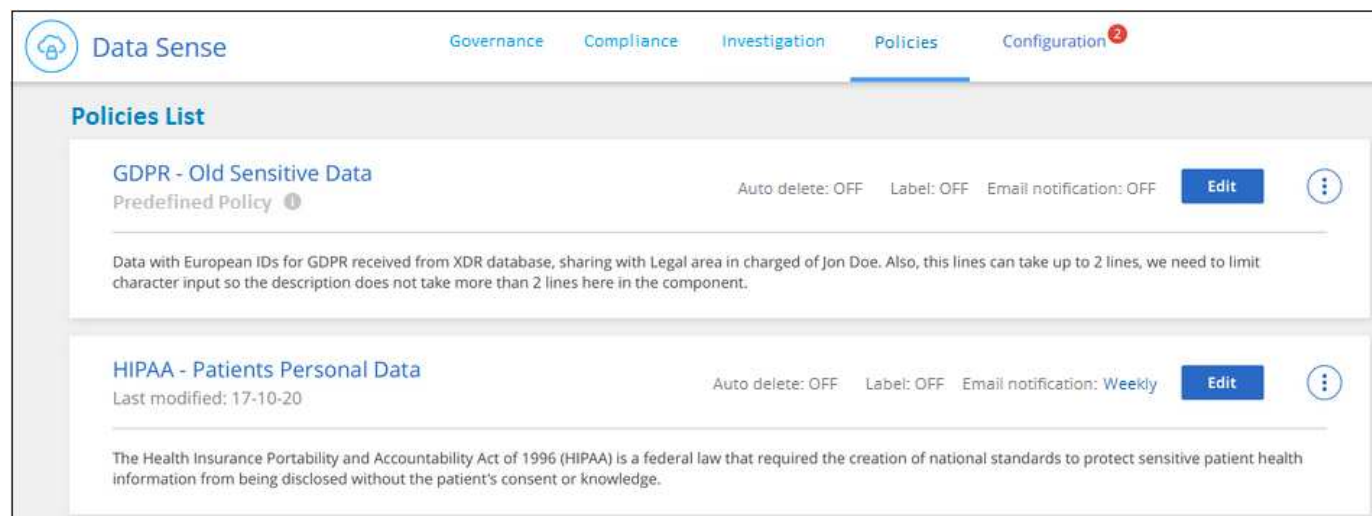
使用策略控制数据

策略类似于自定义筛选器的收藏夹列表，可在 " 调查 " 页面中为常见请求的合规性查询提供搜索结果。Cloud Data sense 可根据常见客户请求提供一组预定义策略。您可以创建自定义策略，为特定于您的组织的搜索提供结果。

策略提供以下功能：


- [预定义策略](#) 基于用户请求从 NetApp 获得
- 能够创建自己的自定义策略
- 单击一下即可启动包含策略结果的调查页面
- 当某些关键策略返回结果时，向 Cloud Manager 用户发送电子邮件警报，以便您可以获得保护数据的通知
- 将 AIP （ Azure 信息保护）标签自动分配给与策略中定义的标准匹配的所有文件
- 如果某些策略返回结果，则会自动删除文件（每天删除一次），以便您可以自动保护数据

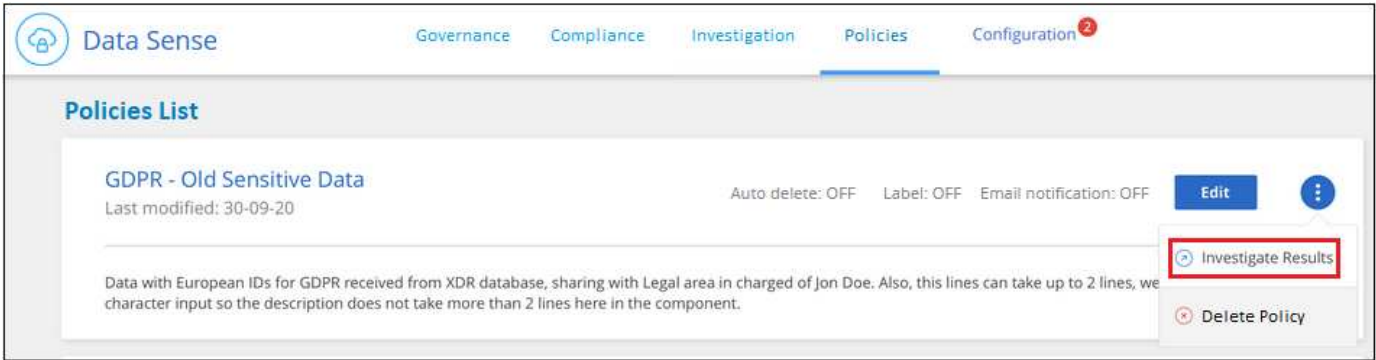
合规性信息板中的 * 策略 * 选项卡列出了此 Cloud Data sense 实例上可用的所有预定义和自定义策略。



此外，策略还会显示在 " 调查 " 页面的筛选器列表中。

在 " 调查 " 页面中查看策略结果

要在 " 调查 " 页面中显示策略的结果，请单击  按钮，然后选择 * 调查结果 *。

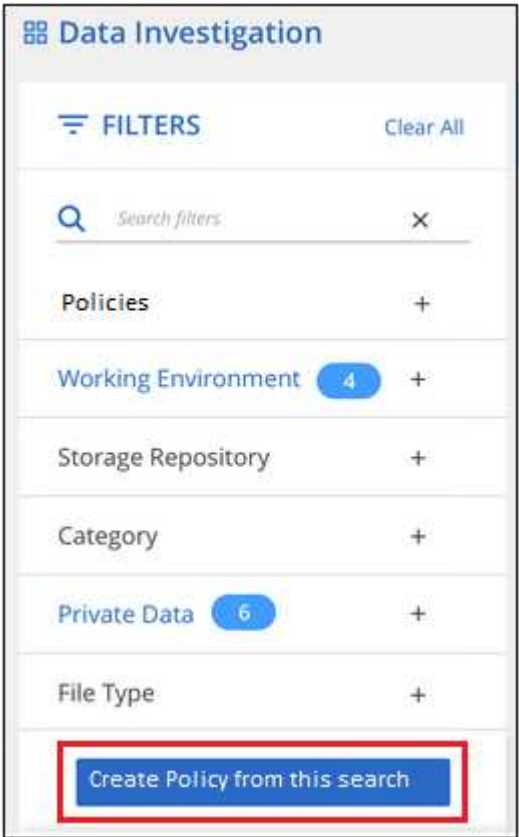


创建自定义策略

您可以创建自己的自定义策略，为特定于您的组织的搜索提供结果。

步骤

- 1. 在数据调查页面中，选择要使用的所有筛选器来定义搜索。请参见 "筛选 " 数据调查 " 页面中的数据" 了解详细信息。
- 2. 按所需方式获取所有筛选器特征后，单击 * 从此搜索创建策略 *。



3. 为策略命名，然后选择可由策略执行的其他操作：
 - a. 输入唯一名称和问题描述。
 - b. 或者，选中此框可自动删除与策略参数匹配的文件。了解更多信息 ["使用策略删除源文件"](#)。
 - c. 或者，如果您希望向 Cloud Manager 用户发送通知电子邮件，请选中此框，然后选择发送电子邮件的间隔。了解更多信息 ["根据策略结果发送电子邮件警报"](#)。
 - d. 或者，选中此框可自动为与策略参数匹配的文件分配 AIP 标签，然后选择此标签。（仅当您已集成 AIP 标签时。了解更多信息 ["AIP 标签"](#)）
 - e. 单击 * 创建策略 *。

Create Policy

This will create a new Policy according to the current selected filters and search term.
You can view or delete this later from the "Policies" tab.

Note it may take up to 15 minutes for results to be displayed for a new Policy.

Name this Policy

New Policy to view all files that were created over 60 days ago

Give it a detailed description that explains what it searches for

See if any files greater than 60 days old should be deleted from the system

☐ Automatically delete files that match this policy (Every Day)

☒ Send email updates about this Policy to Cloud Manager users on this account every Day

☐ Automatically label this Policy's matches with: Select a label

Create Policy Cancel

新策略将显示在策略选项卡中。

在发现不合规数据时发送电子邮件警报

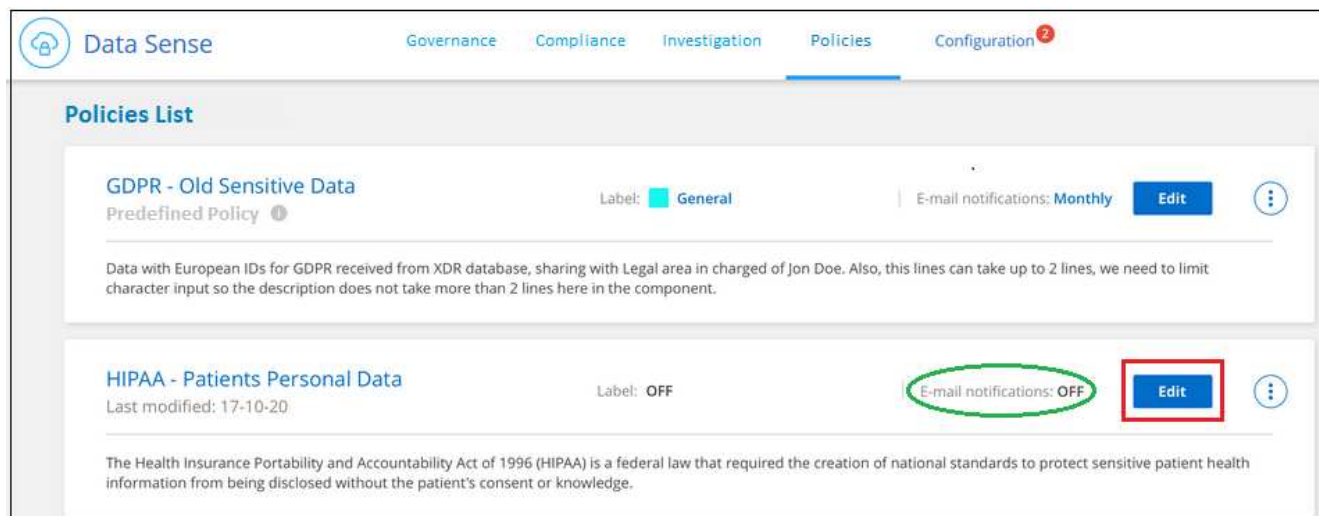
当某些关键策略返回结果时，Cloud Data sense 可以向 Cloud Manager 用户发送电子邮件警报，以便您可以获得保护数据的通知。您可以选择每天，每周或每月发送电子邮件通知。

您可以在创建策略或编辑任何策略时配置此设置。

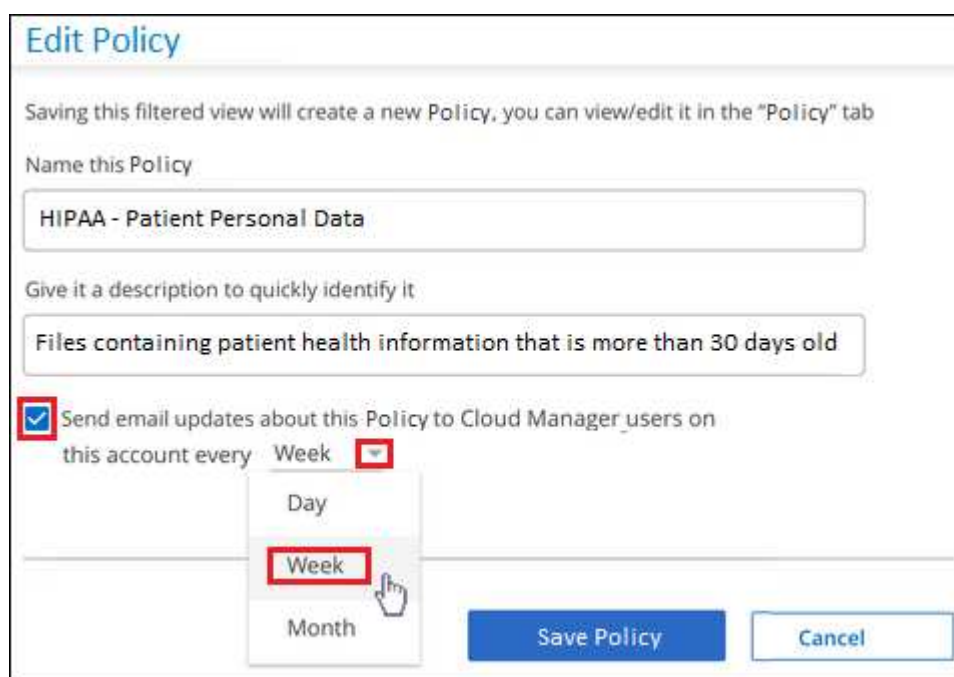
按照以下步骤向现有策略添加电子邮件更新。

步骤

1. 在策略列表页面中，单击要添加（或更改）电子邮件设置的策略的 * 编辑 *。



2. 在编辑策略页面中，如果要向 Cloud Manager 用户发送通知电子邮件，请选中此框，然后选择发送电子邮件的间隔（例如，每 * 周 *）。



3. 单击 * 保存策略 *，策略问题描述中将显示发送电子邮件的间隔。

现在，如果策略中有任何结果，则会发送第一封电子邮件，但前提是任何文件符合策略标准。通知电子邮件不会发送任何个人信息。此电子邮件指示存在与策略条件匹配的文件，并提供指向策略结果的链接。

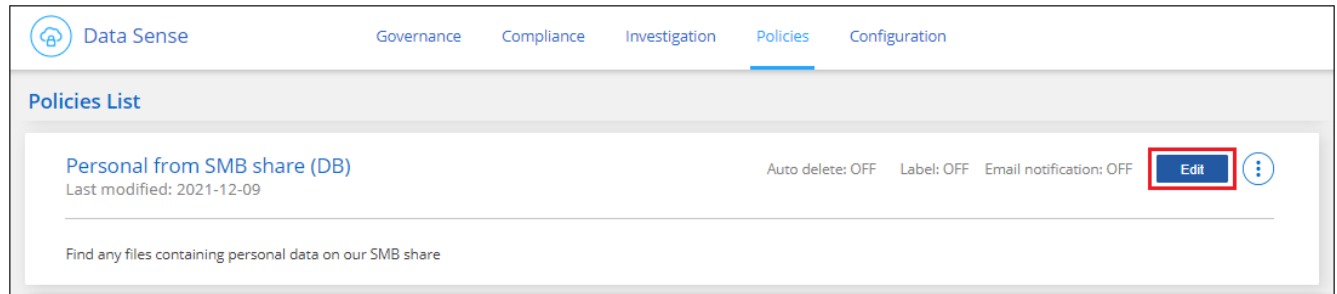
编辑策略

您可以修改先前创建的现有策略的任何条件。如果要更改查询(使用筛选器定义的项)以添加或删除某些参数、则此功能尤其有用。

请注意、对于预定义策略、您只能修改是否发送电子邮件通知以及是否添加AIP标签。不能更改任何其他值。

步骤

1. 在策略列表页面中、单击要更改的策略的*编辑*。



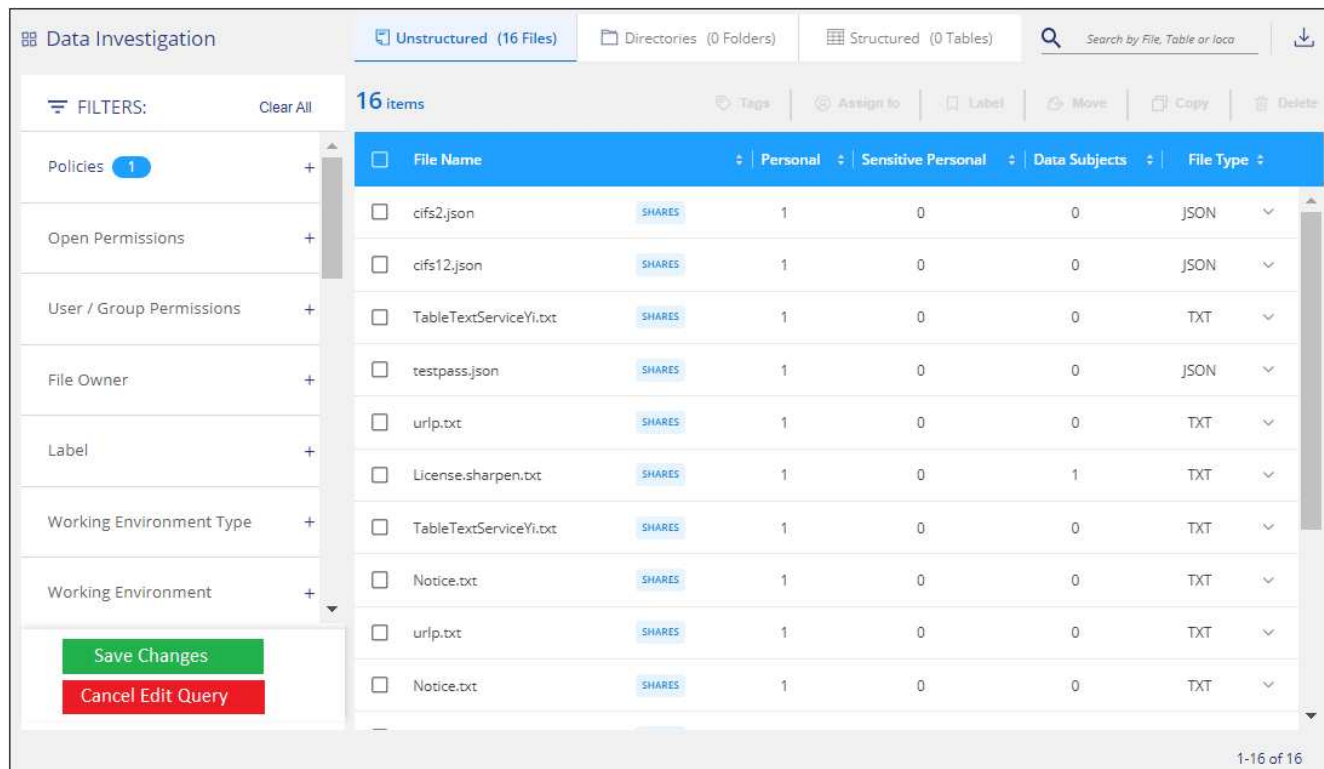
2. 如果您只想更改此页面上的项(名称、问题描述、是否发送电子邮件通知以及是否添加了AIP标签)、请进行更改并单击*保存策略*。

如果要更改已保存查询的筛选器、请单击*编辑查询*。

The screenshot shows the 'Edit Policy' form. At the top right, a blue 'Edit Query' button is highlighted with a red rectangular box. The form contains the following fields and options:

- 'Name this Policy' field with the text 'Personal from SMB share (DB)'.
- 'Give it a detailed description that explains what it searches for' field with the text 'Find any files containing personal data on our SMB share'.
- Three checkboxes with labels:
 - ☐ Automatically delete files that match this policy (Every Day)
 - ☐ Send email updates about this Policy to Cloud Manager users on this account every Day
 - ☐ Automatically label this Policy's matches with:
- At the bottom, two buttons: 'Save Policy' (blue) and 'Cancel' (light blue).

3. 在定义该查询的调查页面中、通过添加、删除或自定义筛选器来编辑查询、然后单击*保存更改*。



策略将立即更改。为该策略定义的用于发送电子邮件、添加AIP标签或删除文件的任何操作都将在下一个内部发生。

正在删除策略

如果您不再需要创建的任何自定义策略，则可以将其删除。您无法删除任何预定义策略。

要删除策略，请单击  按钮，单击 * 删除策略 *，然后在确认对话框中再次单击 * 删除策略 *。

预定义策略列表

Cloud Data sense 提供了以下系统定义的策略：

Name	Description	逻辑
S3 公开公开的私有数据	包含个人或敏感个人信息的 S3 对象，具有开放的公有读取访问权限。	S3 公有，包含个人或敏感个人信息
PCI DSS — 30 天内的陈旧数据	包含信用卡信息的文件，上次修改时间为 30 天前。	包含信用卡，最后一次修改日期为 30 天
HIPAA — 30 天内过期的数据	包含运行状况信息的文件，上次修改时间超过 30 天。	包含运行状况数据（定义方式与 HIPAA 报告相同），最后修改时间为 30 天
私有数据—过期 7 年	包含个人或敏感个人信息的文件，上次修改时间为 7 年前。	包含个人或敏感个人信息的文件，上次修改时间为 7 年前

Name	Description	逻辑
GDPR —欧洲公民	包含 5 个以上欧盟国家公民标识符的文件或包含欧盟国家公民标识符的数据库表。	包含 5 个以上的欧盟公民标识符的文件，或者包含超过 15% 列且具有一个国家或地区的欧盟标识符的行的数据库表。（欧洲国家 / 地区的任何一个国家 / 地区标识符。不包括巴西，加利福尼亚，美国 SSN，以色列，南非）
CCPA —加利福尼亚居民	包含 10 个以上加利福尼亚驱动程序许可证标识符的文件或包含此标识符的数据库表。	包含 10 个以上加利福尼亚驱动程序许可证标识符的文件或包含加利福尼亚驱动程序许可证的数据库表
数据主体名称—高风险	数据主题名称超过 50 个的文件。	数据主题名称超过 50 个的文件
电子邮件地址—高风险	电子邮件地址超过 50 个的文件，或者包含电子邮件地址的行数超过 50% 的数据库列	电子邮件地址超过 50 个的文件，或者包含电子邮件地址的行数超过 50% 的数据库列
个人数据—高风险	包含 20 个以上个人数据标识符的文件，或者包含个人数据标识符的行数超过 50% 的数据库列。	包含 20 个以上个人列的文件，或包含 50% 以上个人行的数据库列
敏感个人数据—高风险	包含 20 个以上敏感个人数据标识符的文件，或者包含敏感个人数据的行数超过 50% 的数据库列。	包含 20 多个敏感个人或数据库列的文件，其中 50% 以上的行包含敏感个人

管理私有数据

Cloud Data sense 为您提供了多种管理私有数据的方法。某些功能使您可以更轻松地为迁移数据做准备，而其他功能则允许您对数据进行更改。

- 如果要为某些数据创建副本并将其移动到其他 NFS 位置，可以将文件复制到目标 NFS 共享。
- 您可以将 ONTAP 卷克隆到新卷，同时在新克隆的卷中仅包含源卷中的选定文件。如果您要迁移数据，并且希望从原始卷中排除某些文件，则此选项非常有用。
- 您可以将文件从源存储库复制并同步到特定目标位置的目录。如果您要将数据从一个源系统迁移到另一个源系统，而源文件上仍有一些最终活动，则此功能非常有用。
- 您可以将 Data sense 正在扫描的源文件移动到任何 NFS 共享。
- 您可以删除那些看起来不安全或风险过高，无法留在存储系统中的文件，或者您已确定为重复的文件。



- 只有在选择对数据源执行完整分类扫描后，才可以使用本节所述的功能。已执行仅映射扫描的数据源不会显示文件级详细信息。
- 目前，Google Drive 帐户中的数据无法使用上述任何功能。

正在复制源文件

您可以复制 Data sense 正在扫描的任何源文件。根据您尝试完成的任务，复制操作有三种类型：

- * 将文件 * 从相同或不同的卷或数据源复制到目标 NFS 共享。

如果要为某些数据创建副本并将其移动到其他 NFS 位置，则此功能非常有用。

- * 将 ONTAP 卷 * 克隆到同一聚合中的新卷，但在新克隆的卷中仅包含源卷中的选定文件。


如果您要迁移数据，并且希望从原始卷中排除某些文件，则此选项非常有用。此操作将使用 "NetApp FlexClone" 用于快速复制卷，然后删除您 * 未 * 选择的文件的功能。

- 将文件 * 从单个源存储库（ONTAP 卷，S3 存储分段，NFS 共享等）复制并同步到特定目标（目标）位置的目录。

这对于将数据从一个源系统迁移到另一个源系统的情况非常有用。初始副本之后、服务将根据您设置的计划同步所有更改的数据。此操作将使用 "NetApp Cloud Sync" 用于将数据从源复制并同步到目标的功能。

将源文件复制到 **NFS 共享**

您可以将 Data sense 正在扫描的源文件复制到任何 NFS 共享。NFS 共享不需要与 Data sense 集成，您只需知道要复制所有选定文件的 NFS 共享的名称，格式为 `<host_name> : /<share_path>` 。



您不能复制数据库中的文件。

要求

- 要复制文件，您必须具有帐户管理员或工作空间管理员角色。
- 复制文件要求目标 NFS 共享允许从 Data sense 实例进行访问。
- 一次最多可以复制 100 , 000 个文件。

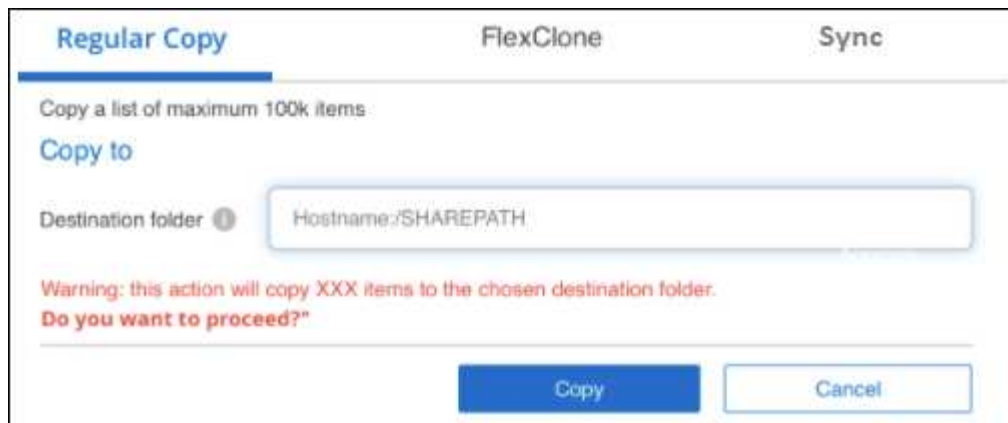
步骤

1. 在数据调查结果窗格中，选择要复制的一个或多个文件，然后单击 * 复制 * 。



- 要选择单个文件，请选中每个文件（☒ Volume_1）。
- 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。
- 要选择所有页面上的所有文件，请选中标题行（☒ File Name），然后显示在弹出消息中
All 20 Items on this page selected Select all Items in list (63K Items)下，单击 * 选择列表中的所有项（xxx 项）*。

2. 在 Copy Files 对话框中，选择 * 常规复制 * 选项卡。



3. 输入要复制所有选定文件的 NFS 共享的名称，格式为 ` : /<share_path>`，然后单击 * 复制 *。

此时将显示一个对话框，其中包含复制操作的状态。

您可以在中查看复制操作的进度 "操作状态窗格"。

请注意，您也可以在查看文件的元数据详细信息时复制单个文件。只需单击 * 复制文件 *。



将卷数据克隆到新卷

您可以使用 NetApp_FlexClone_ 功能克隆 Data sense 正在扫描的现有 ONTAP 卷。这样，您就可以快速复制卷，同时仅包含选定的文件。如果要迁移数据，并且要从原始卷中排除某些文件，或者要创建卷的副本以进行测试，则此功能非常有用。

新卷将在与源卷相同的聚合中创建。在启动此任务之前，请确保聚合中有足够的空间容纳此新卷。如有必要，请联系您的存储管理员。

- 注意：* 无法克隆 FlexGroup 卷，因为 FlexClone 不支持这些卷。

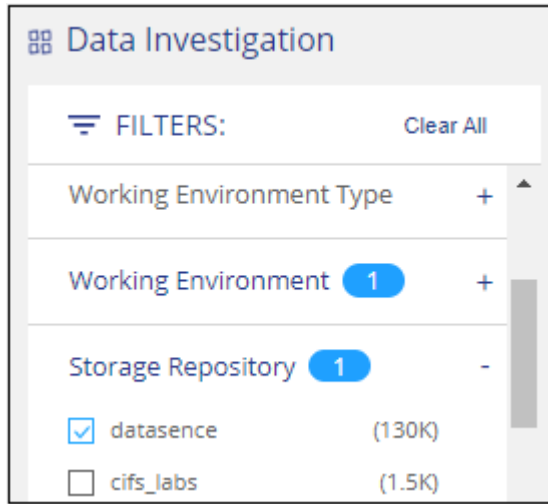
要求

- 要复制文件，您必须具有帐户管理员或工作空间管理员角色。
- 所有选定文件都必须来自同一个卷，并且该卷必须处于联机状态。
- 卷必须来自 Cloud Volumes ONTAP 或内部 ONTAP 系统。当前不支持任何其他数据源。

- 集群上必须安装 FlexClone 许可证。默认情况下，此许可证会安装在 Cloud Volumes ONTAP 系统上。

步骤

1. 在 "数据调查" 窗格中，选择一个 * 工作环境 * 和一个 * 存储库 * 来创建筛选器，以确保所有文件都来自同一个 ONTAP 卷。



应用任何其他筛选器，以便您只能看到要克隆到新卷的文件。

2. 在调查结果窗格中，选择要克隆的文件，然后单击 * 复制 *。



- 要选择单个文件，请选中每个文件 (☒ Volume_1)。
- 要选择当前页面上的所有文件，请选中标题行 (☒ File Name)。
- 要选择所有页面上的所有文件，请选中标题行 (☒ File Name)，然后显示在弹出消息中 [All 20 Items on this page selected](#) [Select all Items in list \(63K Items\)](#) 下，单击 * 选择列表中的所有项 (xxx 项) *。

3. 在 *Copy Files* 对话框中，选择 * FlexClone * 选项卡。此页面显示要从卷克隆的文件总数 (您选择的文件)，以及未从克隆的卷中包含 / 删除的文件数 (您未选择的文件)。

4. 输入新卷的名称，然后单击 * FlexClone *。

此时将显示一个对话框，其中包含克隆操作的状态。

新的克隆卷将在与源卷相同的聚合中创建。

您可以在中查看克隆操作的进度 ["操作状态窗格"](#)。

如果在源卷所在的工作环境中启用了 Data sense 后，最初选择了 * 映射所有卷 * 或 * 映射并分类所有卷 *，则 Data sense 将自动扫描新克隆的卷。如果最初未使用上述任一选项，则如果要扫描此新卷，则需要执行以下操作 ["手动对卷启用扫描"](#)。

将源文件复制并同步到目标系统

您可以将 Data sense 正在扫描的源文件从任何受支持的非结构化数据源复制到特定目标位置的目录 (["Cloud Sync 支持的目标位置"](#))。初始复制后，文件中更改的任何数据将根据您配置的计划进行同步。

这对于将数据从一个源系统迁移到另一个源系统的情况非常有用。此操作将使用 ["NetApp Cloud Sync"](#) 用于将数据从源复制并同步到目标的功能。



您不能复制和同步数据库， OneDrive 帐户或 SharePoint 帐户中的文件。

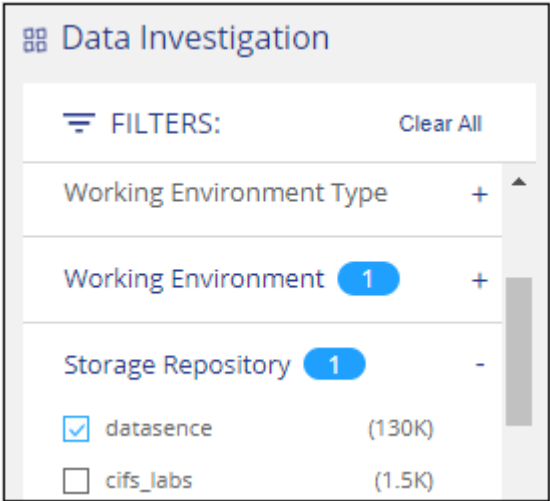
要求

- 要复制和同步文件，您必须具有帐户管理员或工作空间管理员角色。
- 所有选定文件都必须来自同一源存储库（ ONTAP 卷， S3 存储分段， NFS 或 CIFS 共享等）。
- 您需要激活 Cloud Sync 服务并至少配置一个数据代理，用于在源系统和目标系统之间传输文件。查看从开始的 Cloud Sync 要求 ["快速启动问题描述"](#)。

请注意， Cloud Sync 服务会为您的同步关系单独收取服务费用，如果您在云中部署数据代理，则会产生资源费用。

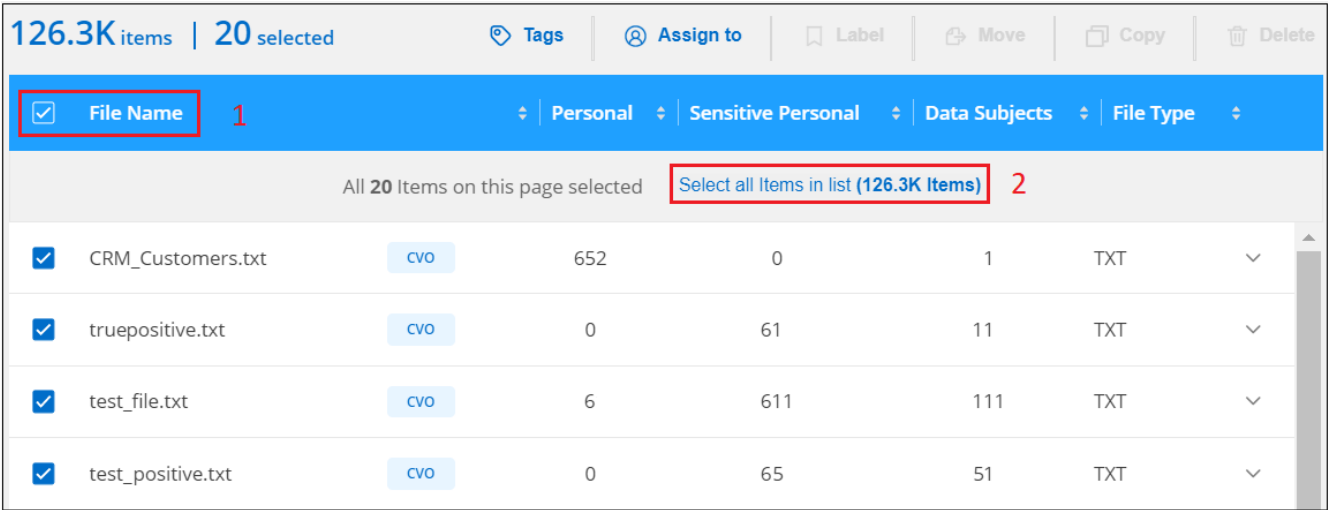
步骤

1. 在数据调查窗格中，选择一个 * 工作环境 * 和一个 * 存储库 * 来创建筛选器，以确保所有文件都来自同一个存储库。

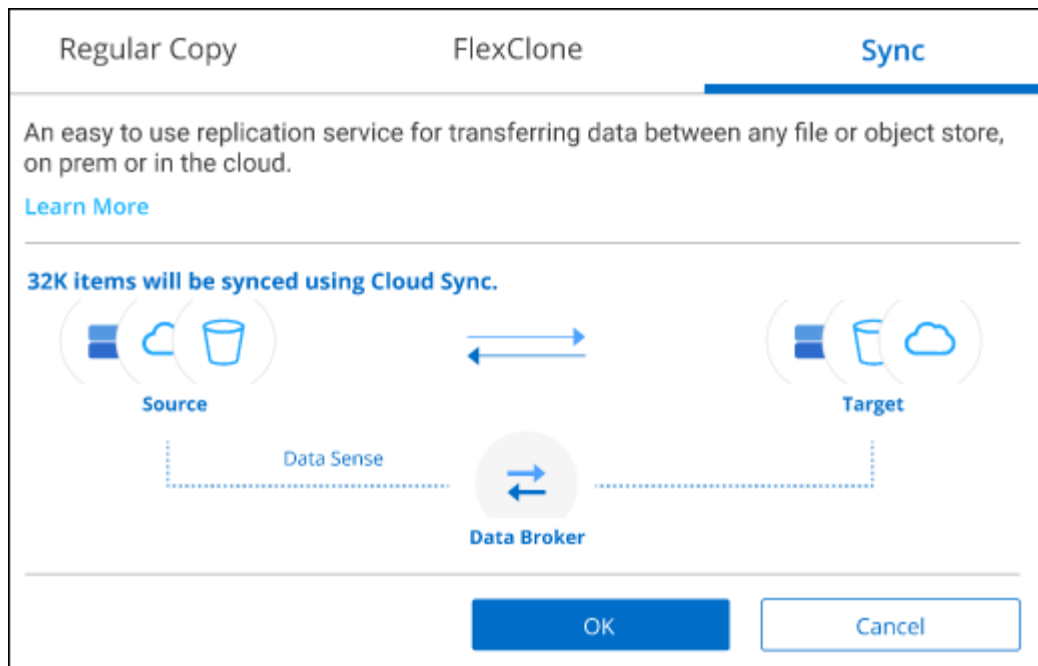


应用任何其他筛选器，以便您仅看到要复制并同步到目标系统的文件。

2. 在调查结果窗格中，选中标题行 (☒ File Name)，然后显示在弹出消息中
All 20 Items on this page selected Select all Items in list (63K Items) 单击 * 选择列表中的所有项 (xxx 项) *，然后单击 * 复制 *。



3. 在 Copy Files 对话框中，选择 * 同步 * 选项卡。



4. 如果确实要将选定文件同步到目标位置，请单击 * 确定 *。

Cloud Sync UI 将在 Cloud Manager 中打开。

系统将提示您定义同步关系。源系统会根据您在 Data sense 中选择的存储库和文件预先填充。

5. 您需要选择目标系统，然后选择（或创建）计划使用的数据代理。查看从开始的 Cloud Sync 要求 "[快速启动问题描述](#)"。

这些文件将复制到目标系统，并根据您定义的计划进行同步。如果选择一次性同步，则文件只会复制和同步一次。如果选择定期同步，则会根据计划同步文件。请注意，如果源系统添加的新文件与您使用筛选器创建的查询匹配，这些 *new* 文件将复制到目标并在将来进行同步。

请注意，从数据感知调用某些常见 Cloud Sync 操作时，这些操作将被禁用：

- 不能使用 * 删除源上的文件 * 或 * 删除目标上的文件 * 按钮。
- 已禁用运行报告。

将源文件移动到 NFS 共享

您可以将 Data sense 正在扫描的源文件移动到任何 NFS 共享。NFS 共享不需要与 Data sense 集成（请参见 "[正在扫描文件共享](#)"）。



您无法移动驻留在数据库中的文件。

要移动文件，您必须具有帐户管理员或工作空间管理员角色。

移动文件要求 NFS 共享允许从 Data sense 实例进行访问。

步骤

1. 在数据调查结果窗格中，选择要移动的一个或多个文件。

2345 items		Tags	Assign to	Label	Copy	Move	Delete
<input type="checkbox"/>	File Name		Personal	Sensitive Personal	Data Subjects	File Type	
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	16	PDF	▼
<input checked="" type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF	▼
<input type="checkbox"/>	Expense Report EXP-TPO-106038887654	cvo	6	3	6	PDF	▼

- 要选择单个文件，请选中每个文件（☒ Volume_1）。
- 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。

2. 在按钮栏中，单击 * 移动 *。

Move Multiple Files

This file will be moved to the destination folder you provide and will no longer be available at its current location.
Moving files is supported only to destination folders in NFS Shares. Any NFS Share is supported, no matter where its hosted, as long as the Share's export policy allows access from the Data Connector instance IP Address.

File Name 1
File Name 2
File Name 3

Enter the NFS destination folder path to continue

Move Files
Cancel

3. 在 *move Files* 对话框中，以 ``<host_name> : /<share_path>`` 格式输入要移动所有选定文件的 NFS 共享的名称，然后单击 * 移动文件 *。

请注意，在查看文件的元数据详细信息时，您也可以移动单个文件。只需单击 * 移动文件 *。



正在删除源文件

您可以永久删除看似不安全或风险太大，无法留在存储系统中的源文件，或者已确定为重复的源文件。此操作为永久操作，不会撤消或还原。

您可以从 " 调查 " 窗格手动删除文件，也可以使用策略自动删除文件。



您不能删除数据库中的文件。

删除文件需要以下权限：

- 对于 NFS 数据—需要使用写入权限定义导出策略。
- 对于 CIFS 数据— CIFS 凭据需要具有写入权限。
- 对于 S3 数据 - IAM 角色必须包括以下权限： s 3 : DeleteObject

手动删除源文件

要求

- 要删除文件，您必须具有帐户管理员或工作空间管理员角色。
- 一次最多可以删除 100 ， 000 个文件。

步骤

1. 在数据调查结果窗格中，选择要删除的一个或多个文件。



- 要选择单个文件，请选中每个文件（☒ Volume_1）。
- 要选择当前页面上的所有文件，请选中标题行（☒ File Name）。
- 要选择所有页面上的所有文件，请选中标题行（☒ File Name），然后显示在弹出消息中
All 20 Items on this page selected Select all Items in list (63K Items)下，单击 * 选择列表中的所有项（xxx 项）*。

2. 在按钮栏中，单击 * 删除 *。
3. 由于删除操作是永久性的，因此您必须在后续的 *Delete File* 对话框中键入 "* 永久删除 *"，然后单击 * 删除文件 *。

您可以在中查看删除操作的进度 "操作状态窗格"。


请注意，您也可以查看文件的元数据详细信息时删除单个文件。只需单击 * 删除文件 *。



使用策略自动删除源文件

您可以创建自定义策略以删除与此策略匹配的文件。例如，您可能希望删除包含敏感信息且在过去 30 天内由 Data sense 发现的文件。

只有帐户管理员可以创建一个策略来自动删除文件。

-  与此策略匹配的所有文件将每天永久删除一次。

步骤

1. 在数据调查页面中，选择要使用的所有筛选器来定义搜索。请参见 "筛选 "数据调查 " 页面中的数据" 了解详细信息。
2. 按所需方式获取所有筛选器特征后，单击 * 从此搜索创建策略 *。
3. 为策略命名，然后选择可由策略执行的其他操作：
 - a. 输入唯一名称和问题描述。
 - b. 选中 " 自动删除与此策略匹配的文件 " 复选框，然后键入 * 永久删除 * 以确认您希望此策略永久删除文件。
 - c. 单击 * 创建策略 *。

Create Policy

This will create a new Policy according to the current selected filters and search term. You can view or delete this later from the "Policies" tab.

Note it may take up to 15 minutes for results to be displayed for a new Policy.

Name this Policy

Delete files with sensitive data

Give it a detailed description that explains what it searches for

Delete files that contain sensitive information and that were discovered in the past 30 days

☒ Automatically delete files that match this policy (Every Day)

Type "permanently delete" to continue with the deletion.

permanently delete

☐ Send email updates about this Policy to Cloud Manager users on this account every Day

☐ Automatically label this Policy's matches with: Select a label

Create Policy Cancel

新策略将显示在策略选项卡中。与策略匹配的文件将在策略运行时每天删除一次。

您可以在中查看已删除的文件列表 ["操作状态窗格"](#)。

查看合规性报告

Cloud Data sense 提供报告，您可以使用这些报告更好地了解组织的数据隐私计划的状态。

默认情况下、Cloud Data sense信息板会显示所有工作环境、数据库和数据源的合规性和监管数据。如果要查看仅包含某些工作环境的数据的报告， [选择这些工作环境](#)。



- 只有在选择对数据源执行完整分类扫描后，才可以使用本节中所述的报告。已执行仅映射扫描的数据源只能生成数据映射报告。
- NetApp 无法保证 Cloud Data sense 识别的个人数据和敏感个人数据的 100% 准确性。您应始终通过查看数据来验证此信息。

隐私风险评估报告

根据 GDPR 和 CCPA 等隐私法规的要求，隐私风险评估报告概述了贵组织的隐私风险状态。此报告包含以下信息：

合规状态

答 严重性得分 以及数据的分布，无论是非敏感数据，个人数据还是敏感数据。

评估概述

所发现的个人数据类型以及数据类别的细分。

此评估中的数据主题

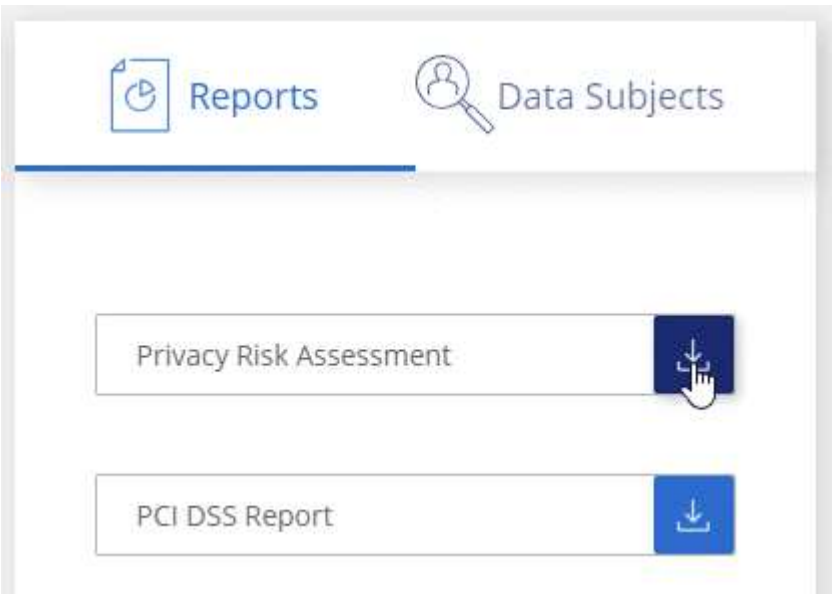
按地点列出的发现了国家标识符的人数。

生成隐私风险评估报告

转到数据感知选项卡以生成报告。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *。
2. 单击 * 合规性 *，然后单击 * 报告 * 下 * 隐私风险评估 * 旁边的下载图标。



Cloud Data sense 可生成 PDF 报告，您可以根据需要查看该报告并将其发送给其他组。

严重性得分

Cloud Data sense 会根据以下三个变量计算隐私风险评估报告的严重性得分：

- 个人数据在所有数据中所占的百分比。
- 敏感个人数据在所有数据中所占的百分比。
- 包含数据主体的文件百分比，由国家 ID，社会保障号和税务 ID 号等国家标识符确定。

用于确定得分的逻辑如下：

严重性得分	逻辑
0	所有这三个变量都恰好为 0%
1.	其中一个变量大于 0%
2.	其中一个变量大于 3%
3.	其中两个变量大于 3%
4.	其中三个变量大于 3%
5.	其中一个变量大于 6%
6.	其中两个变量大于 6%
7.	其中三个变量大于 6%
8.	其中一个变量大于 15%
9	其中两个变量大于 15%
10	其中三个变量大于 15%

PCI DSS 报告

支付卡行业数据安全标准（PCI DSS）报告可帮助您确定信用卡信息在各个文件中的分布情况。此报告包含以下信息：

概述

包含信用卡信息的文件数量以及在何种工作环境中。

加密

包含信用卡信息的文件在加密或未加密工作环境中的百分比。此信息特定于 Cloud Volumes ONTAP。

勒索软件保护

包含信用卡信息的文件在已启用或未启用勒索软件保护的工作环境中所占的百分比。此信息特定于 Cloud Volumes ONTAP。

Retention

上次修改文件的时间范围。这很有用，因为您不应将信用卡信息保留的时间超过处理信用卡信息所需的时间。

分发信用卡信息

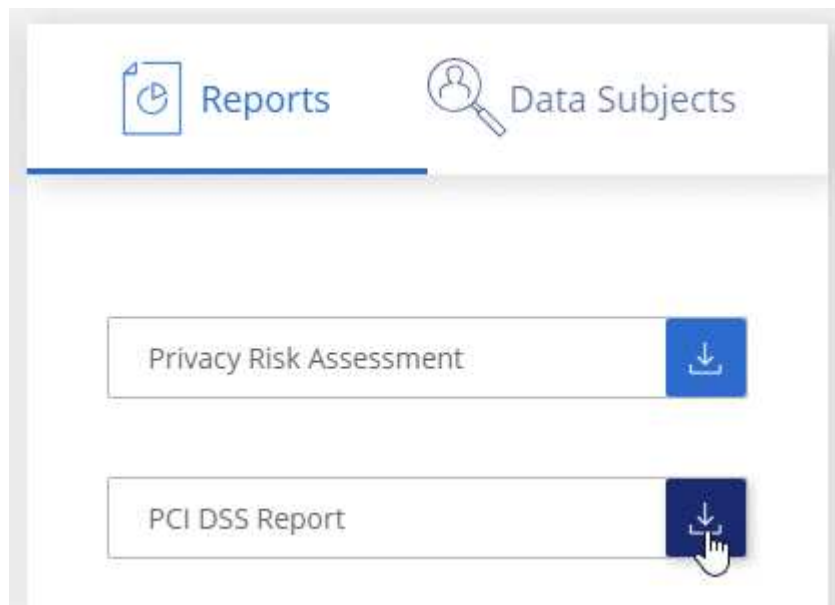
发现信用卡信息的工作环境以及是否启用了加密和勒索软件保护。

正在生成 PCI DSS 报告

转到数据感知选项卡以生成报告。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *。
2. 单击 * 合规性 *，然后单击 * 报告 * 下的 * PCI DSS 报告 * 旁边的下载图标。



Cloud Data sense 可生成 PDF 报告，您可以根据需要查看该报告并将其发送给其他组。

HIPAA 报告

《健康保险携带和责任法案》（HIPAA）报告可帮助您识别包含健康信息的文件。它旨在帮助您的组织满足 HIPAA 数据隐私法律的要求。Cloud Data sense 需要的信息包括：

- 运行状况参考模式
- ICD-10-CM 医疗代码
- ICD-9-CM 医疗代码
- HR —运行状况类别
- 运行状况应用程序数据类别

此报告包含以下信息：

概述

包含运行状况信息的文件数量以及在何种工作环境中。

加密

包含加密或未加密工作环境中运行状况信息的文件的百分比。此信息特定于 Cloud Volumes ONTAP。

勒索软件保护

包含运行状况信息的文件在未启用或未启用勒索软件保护的工作环境中所占的百分比。此信息特定于 Cloud Volumes ONTAP。

Retention

上次修改文件的时间范围。这很有用，因为运行状况信息的保留时间不应超过处理该信息所需的时间。

分发运行状况信息

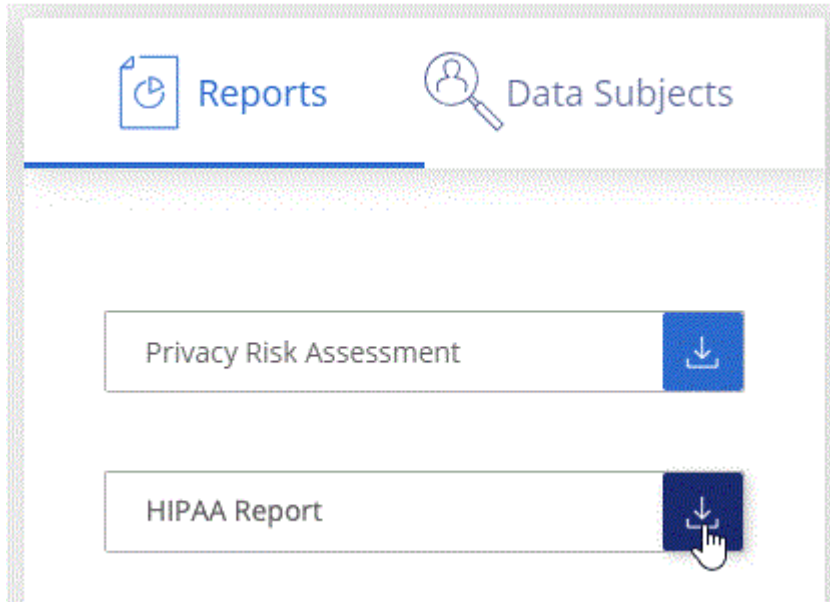
发现运行状况信息的工作环境以及是否启用了加密和勒索软件保护。

生成 HIPAA 报告

转到数据感知选项卡以生成报告。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *。
2. 单击 * 合规性 *，然后单击 * 报告 * 下的 * HIPAA 报告 * 旁边的下载图标。



Cloud Data sense 可生成 PDF 报告，您可以根据需要查看该报告并将其发送给其他组。

数据映射报告

数据映射报告概述了存储在企业数据源中的数据，可帮助您决定迁移，备份，安全性和合规性过程。该报告首先列出一份概述报告，其中汇总了您的所有工作环境和数据源，然后对每个工作环境进行了细分。

此报告包含以下信息：

使用容量

对于所有工作环境：列出每个工作环境的文件数和已用容量。对于单个工作环境：列出容量最多的文件。

数据使用期限

提供了三个图表和图形，用于显示文件的创建时间，上次修改时间或上次访问时间。根据特定日期范围列出文件数量及其已用容量。

数据大小

列出了工作环境中特定大小范围内的文件数。

文件类型

列出了工作环境中存储的每种类型的文件的总文件数和已用容量。

正在生成数据映射报告

转到数据感知选项卡以生成报告。

步骤

- 1. 在 Cloud Manager 顶部，单击 * 数据感知 *。
- 2. 单击 * 监管 *，然后单击监管信息板中的 * 完整数据映射概述报告 * 按钮。



Cloud Data sense 可生成 PDF 报告，您可以根据需要查看该报告并将其发送给其他组。

数据调查报告


数据调查报告是数据调查页面内容的下载。 ["了解有关数据调查页面的更多信息"](#)。

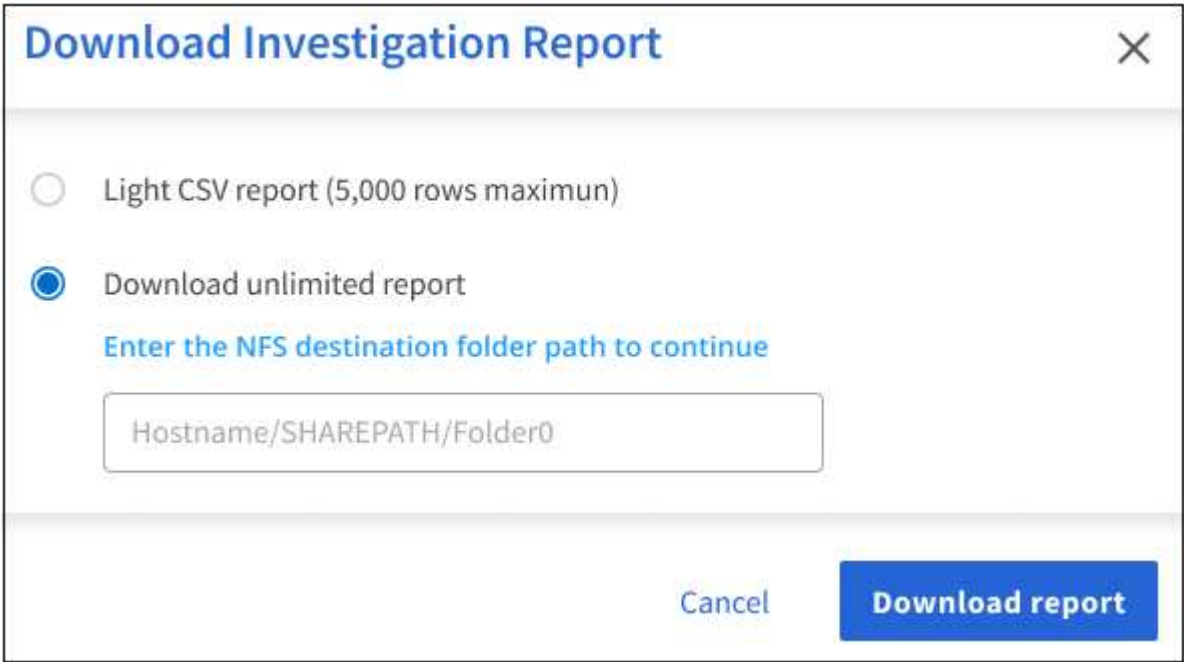
您可以将报告保存到本地计算机、作为.CSV文件(最多可包含5、000行数据)或导出到NFS共享的.json文件(最多可包含不限数量的行)。如果Data sense正在扫描文件(非结构化数据)、目录(文件夹和文件共享)或数据库(结构化数据)、则最多可以下载三个报告文件。

导出到文件共享时、请确保Data sense具有正确的导出访问权限。

生成数据调查报告

步骤

- 1. 在数据调查页面中、单击  按钮。
- 2. 选择要下载.CSV报告还是.JSON数据报告、然后单击*下载报告*。



选择.json报告时、请以`<host_name>: /<share_path>`格式输入要下载报告的NFS共享的名称。

此时将显示一条消息、指出正在下载报告。

您可以在中查看JSON报告生成的进度 [操作状态窗格](#)。

每个数据调查报告中包含的内容

*非结构化文件数据报告*包含有关文件的以下信息：

- 文件名
- 位置类型
- Working environment name
- 存储库（例如，卷，存储分段，共享）
- Working environment type
- 文件路径
- 文件类型
- 文件大小
- 创建时间
- 上次修改时间
- 上次访问
- 文件所有者
- 类别
- 个人信息
- 敏感的个人信息
- 删除检测日期

删除检测日期用于标识文件被删除或移动的日期。这样，您就可以确定何时移动了敏感文件。已删除的文件不属于信息板或 " 调查 " 页面上显示的文件编号。这些文件仅显示在 CSV 报告中。

*非结构化目录数据报告*包含有关文件夹和文件共享的以下信息：

- Working environment name
- 存储库(例如、文件夹或文件共享)
- Working environment type
- 文件路径(目录名称)
- 文件所有者
- 创建时间
- 发现时间
- 上次修改时间

- 上次访问
- 打开权限
- 目录类型

*结构化数据报告*包含有关数据库表的以下信息：

- 数据库表名称
- 位置类型
- Working environment name
- 存储库（例如模式）
- 列计数
- 行数
- 个人信息
- 敏感的个人信息

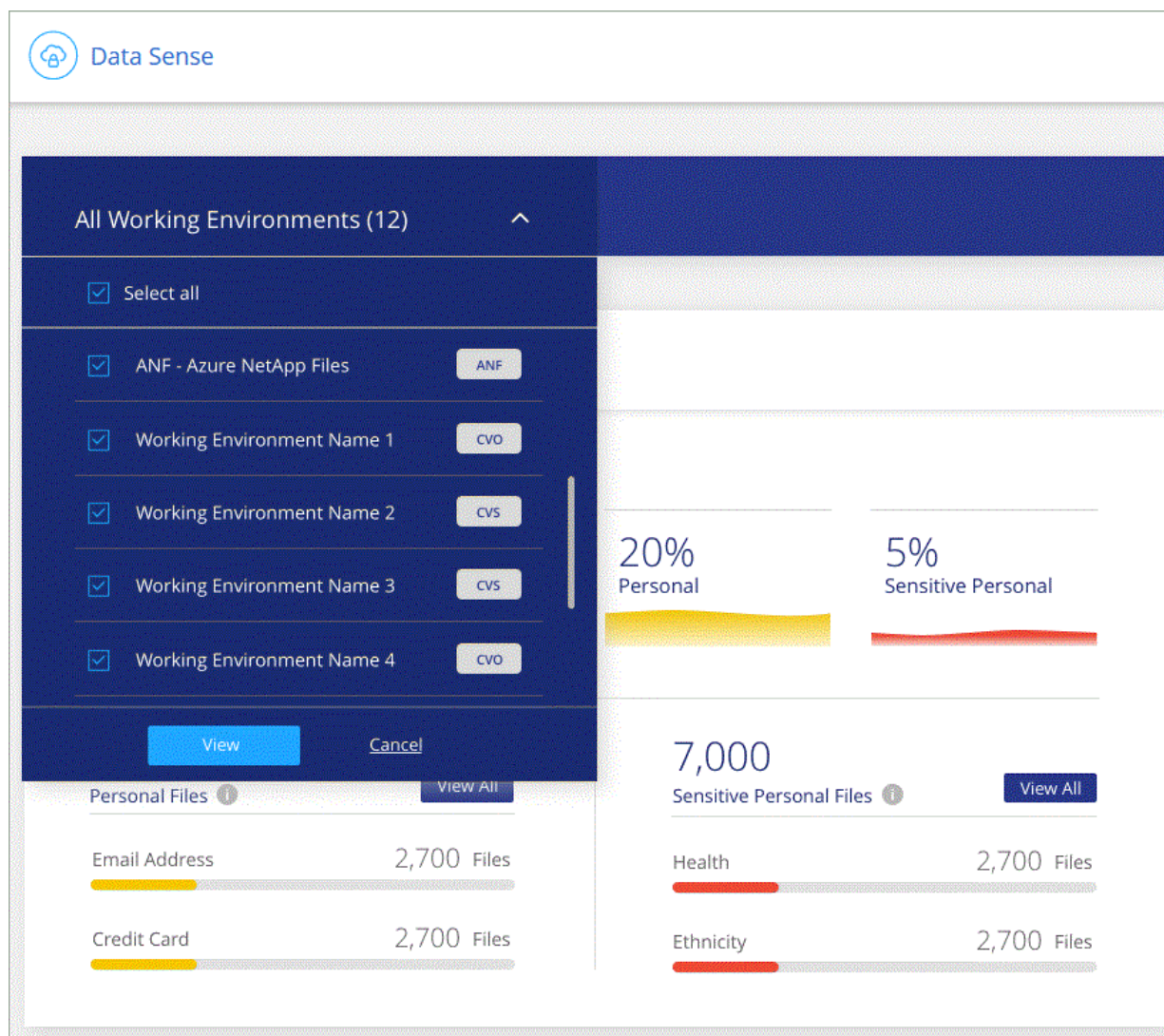
选择报告的工作环境

您可以筛选 " 云数据感知合规性 " 信息板的内容，以查看所有工作环境和数据库的合规性数据，或者仅查看特定工作环境的合规性数据。

筛选信息板时， Data sense 会将合规性数据和报告范围仅限于您选择的工作环境。

步骤

1. 单击筛选器下拉列表，选择要查看其数据的工作环境，然后单击 * 查看 *。



响应数据主体访问请求

通过搜索主题的全名或已知标识符（例如电子邮件地址）并下载报告来响应数据主题访问请求（Data Subject Access Request，DAR）。本报告旨在帮助贵组织满足 GDPR 或类似数据隐私法律的要求。



只有在选择对数据源执行完全分类扫描后，才可以使用 DSAar 功能。已执行仅映射扫描的数据源不提供文件级详细信息。



NetApp 无法保证 Cloud Data sense 识别的个人数据和敏感个人数据的 100% 准确性。您应始终通过查看数据来验证此信息。

什么是数据主体访问请求？

欧洲 GDPR 等隐私法规授予数据主体（如客户或员工）访问其个人数据的权利。当数据主体请求此信息时，此信息称为 DSAar（数据主体访问请求）。组织必须“无不当延迟”地对这些请求做出响应，并且最迟在收到请求后一个月内做出响应。

云数据感知如何帮助您应对 DAR ？

执行数据主题搜索时，Cloud Data sense 会查找包含此用户的名称或标识符的所有文件，分段，OneDrive 和 SharePoint 帐户。Data sense 会检查最新的预索引数据中的名称或标识符。它不会启动新扫描。

搜索完成后，您可以下载数据主题访问请求报告的文件列表。该报告汇总了数据中的洞察信息，并将其纳入法律条款中，您可以将其发回给相关人员。



目前，数据库不支持数据主题搜索。

搜索数据主题并下载报告

搜索数据主体的全名或已知标识符，然后下载文件列表报告或 DSAR 报告。您可以搜索依据 "[任何个人信息类型](#)"。

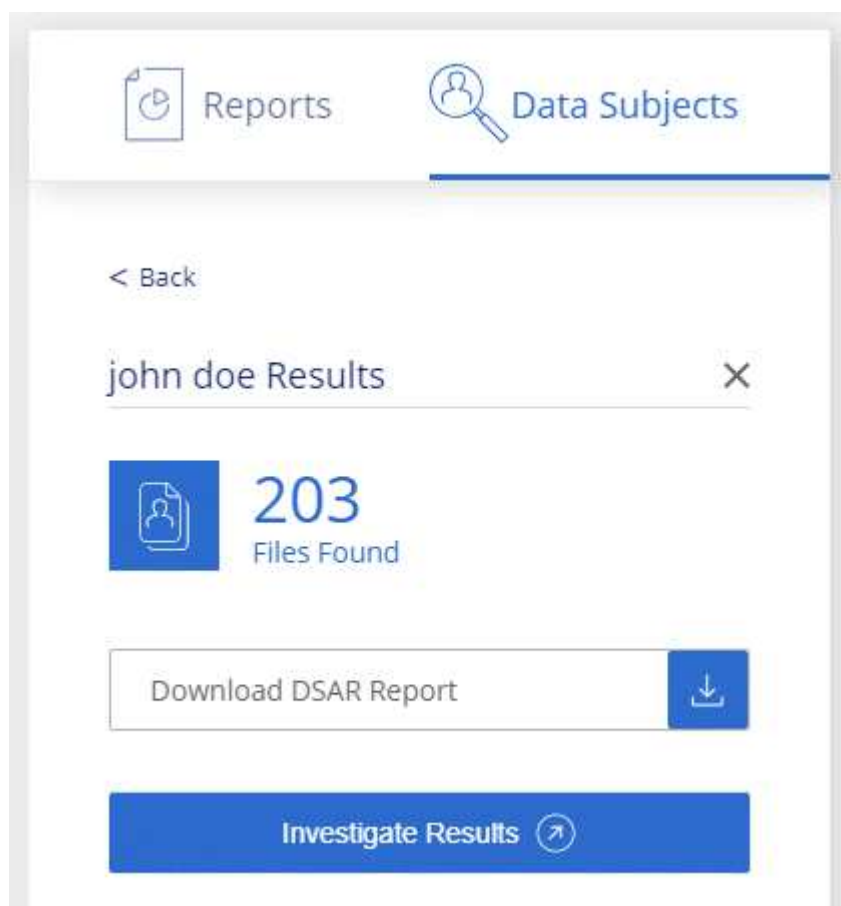


搜索数据主题的名称时，支持英语，德语和西班牙语。稍后将添加对更多语言的支持。

步骤

1. 在 Cloud Manager 顶部，单击 * 数据感知 *。
2. 单击 * 数据主题 *。
3. 搜索数据主体的全名或已知标识符。

以下示例显示了对名称 *john doe* 的搜索：



4. 选择一个可用选项：

- * 下载 DSAI 报告 *：对访问请求的正式响应，您可以将此响应发送给数据主体。此报告包含根据 Cloud Data 在数据主题上发现的数据自动生成的信息，并可用作模板。您应填写此表单并在内部进行审核，然后再将其发送给数据主体。
- * 调查结果 *：一个页面，可用于通过搜索，排序，扩展特定文件的详细信息以及下载文件列表来调查数据。



如果结果超过 10,000 个，则文件列表中仅显示前 10,000 个。

私有数据的类别

Cloud Data sense可以在您的卷、Amazon S3存储分段、数据库、OneDrive文件夹、SharePoint帐户中识别多种类型的私有数据。和Google Drive帐户。请参见以下类别。



如果您需要 Cloud Data sense 识别其他私有数据类型，例如其他国家 / 地区 ID 号或医疗保健标识符，请根据您的请求向 ng-contact-data-sense@netapp.com 发送电子邮件。

个人数据的类型

文件中的个人数据可以是常规个人数据或国家标识符。第三列标识 Cloud Data sense 是否使用 "接近验证" 以验证其对标识符的发现。

可以使用任何语言识别此类别中的项目。

请注意，如果要扫描数据库服务器，您可以将这些数据添加到文件中的个人数据列表中。Data DataSIF_ 功能允许您通过选择数据库表中的列来选择 Fusion sense 在其扫描中查找的其他标识符。请参见 "[使用 Data Fusion 添加个人数据标识符](#)" 了解详细信息。

Type	标识符	是否进行接近验证？
常规	电子邮件地址	否
	信用卡号	否
	IBAN 编号（国际银行帐号）	否
	IP 地址	否
	Password	是的。

	卢森堡 ID	是的。
	马耳他 ID	是的。
Type	国家卫生服务（NHS）编号	是否进行接近验证？
	波兰语 ID（PESEL）	是的。
	葡萄牙税务识别号（NIF）	是的。
	罗马尼亚语 ID（CNP）	是的。
	斯洛文尼亚语 ID（EMO）	是的。
	南非 ID	是的。
	西班牙语税务标识号	是的。
	瑞典语 ID	是的。
	英国ID（Nino）	是的。
	美国社会保险号（SSN）	是的。

敏感个人数据的类型

Cloud Data sense 可在文件中找到的敏感个人数据包括以下列表。目前，此类别中的项目只能识别为英语。

《刑事诉讼参考》

有关自然人的犯罪和犯罪的数据。

《种族参考》

与自然人的种族或种族有关的数据。

运行状况参考

有关自然人健康的数据。

ICD-9-CM 医疗代码

医疗和健康行业使用的代码。

ICD-10-CM 医疗代码

医疗和健康行业使用的代码。

《理念参考》

与自然人的理念相关的数据。

《政治观点参考》

与自然人的政治观点相关的数据。

《宗教信仰参考》

有关自然人的宗教信仰的数据。

性寿命或方向参考

有关自然人的性生活或性取向的数据。

类别类型

Cloud Data sense 可按如下方式对数据进行分类：其中大多数类别均可获得英语，德语和西班牙语的认可。

类别	Type	英语	德语	西班牙语
财务	资产负债表	✓	✓	✓
	采购订单	✓	✓	✓
	发票	✓	✓	✓
	季度报告	✓	✓	✓
人力资源	后台检查	✓		✓
	薪酬计划	✓	✓	✓
	员工合同	✓		✓
	员工审核	✓		✓
	运行状况	✓		✓
	恢复	✓	✓	✓
法律	NDAS	✓	✓	✓
	供应商 - 客户合同	✓	✓	✓
营销	营销活动	✓	✓	✓
	会议	✓	✓	✓
操作	审核报告	✓	✓	✓
销售	销售订单	✓	✓	
服务	RFI	✓		✓
	RFP	✓		✓
	SOW	✓	✓	✓
	培训	✓	✓	✓
支持	投诉和服务单	✓	✓	✓

此外，还会对以下元数据进行分类，并使用相同的受支持语言进行标识：

- 应用程序数据
- 归档文件
- 音频
- 业务应用程序数据
- CAD 文件
- 代码
- 已损坏
- 数据库和索引文件

- 设计文件
- 通过电子邮件发送应用程序数据
- Encrypted
- 可执行文件
- 财务应用程序数据
- 运行状况应用程序数据
- 映像
- 日志
- 其他文档
- 其他演示文稿
- 其他电子表格
- 其他 " 未知 "
- 结构化数据
- 视频
- 零字节文件

文件类型

Cloud Data sense 会扫描所有文件以获取类别和元数据洞察力，并在信息板的文件类型部分显示所有文件类型。

但是，当 Data sense 检测到个人可识别信息（ PiD ）或执行 DSAL 搜索时，仅支持以下文件格式：

.CSV、.dcm、.Dicom、.DOC、.docx、 .json、.PDF、.PPTX、.RTV、.TXT、 .XLS、.XLSX、文档、工作表和幻灯片

发现的信息准确性

NetApp 无法保证 Cloud Data sense 识别的个人数据和敏感个人数据的 100% 准确性。您应始终通过查看数据来验证此信息。

根据我们的测试，下表显示了 Data sense 所发现的信息的准确性。我们将其细分为 *precis度* _ 和 *_recall*：

精确度

正确识别数据感知所发现的可能性。例如，个人数据的精确率为 90% 意味着，在被确定包含个人信息的 10 个文件中，有 9 个文件实际上包含个人信息。每 10 个文件中就有 1 个是误报文件。

重新调用

数据感知找到应做之事的可能性。例如，个人数据的重新调用率为 70% ，这意味着 Data sense 可以识别贵组织中实际包含个人信息的 10 个文件中的 7 个文件。数据感知可能会丢失 30% 的数据，并且不会显示在信息板中。

我们不断提高结果的准确性。这些改进功能将在未来的 Data sense 版本中自动提供。

Type	精确度	重新调用
个人数据—常规	90% 到 95%	60%-80%
个人数据—国家 / 地区标识符	30% 到 60%	40%-60%
敏感的个人数据	80%-95%	20%-30%
类别	90% 到 97%	60%-80%

版权信息

版权所有©2022 NetApp、Inc.。保留所有权利。Printed in the U.S.版权所涵盖的本文档的任何部分不得以任何形式或任何手段复制、包括影印、录制、磁带或存储在电子检索系统中—未经版权所有者事先书面许可。

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

本软件由NetApp按"原样"提供、不含任何明示或默示担保、包括但不限于适销性和特定用途适用性的默示担保、特此声明不承担任何任何责任。IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

商标信息

NetApp、NetApp标识和中列出的标记 <http://www.netapp.com/TM> 是NetApp、Inc.的商标。其他公司和产品名称可能是其各自所有者的商标。