使用数据分类
NetApp Data Classification

NetApp
February 02, 2026

目录

使用数据分类

使用NetApp Data Classification查看组织中存储的数据的治理详细信息

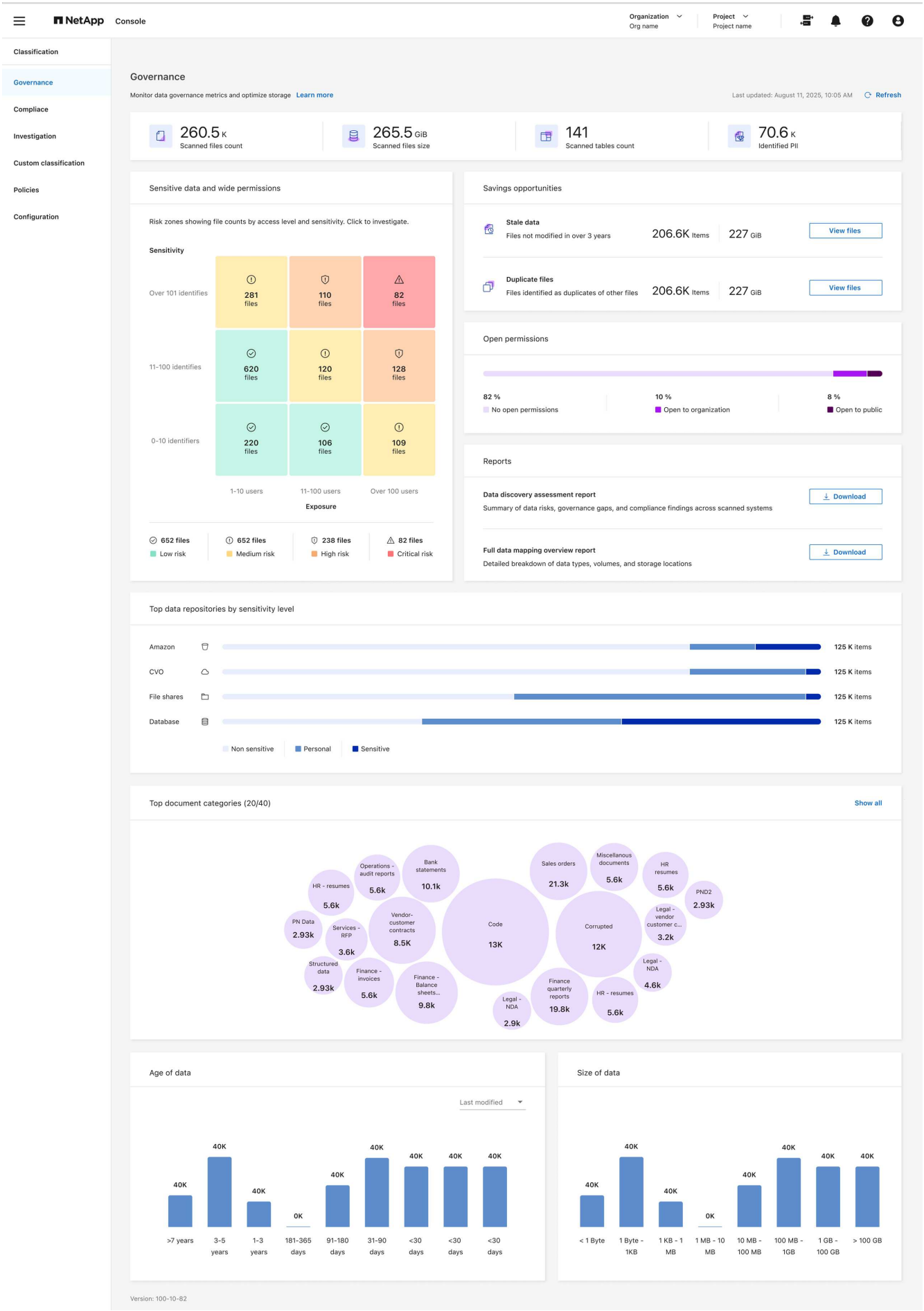
控制与组织存储资源上的数据相关的成本。 NetApp Data Classification可识别系统中陈旧数据、重复文件和超大文件的数量，以便您可以决定是否要删除某些文件或将某些文件分层到成本较低的对象存储中。

您应该从这里开始您的研究。从治理仪表板中，您可以选择一个区域进行进一步调查。

此外，如果您计划将数据从本地位置迁移到云端，则可以在移动数据之前查看数据的大小以及其中是否有任何数据包含敏感信息。

查看治理仪表板

治理仪表板提供信息，以便您可以提高效率并控制与存储在存储资源上的数据相关的成本。


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

步骤

1. 从NetApp Console菜单中，选择 治理 > 分类。
2. 选择*治理*。

   出现治理仪表板。

审查节省机会

_节省机会_组件显示您可以删除或分层到较便宜的对象存储的数据。 《节省机会》中的数据每 2 小时更新一次。您也可以手动更新数据。

步骤

1. 从数据分类菜单中，选择*治理*。
2. 在治理仪表板的每个节省机会图块中，选择*优化存储*以在调查页面中查看过滤的结果。要发现您应该删除或分层到较便宜的存储的任何数据，请调查_节省机会_。
   - 过期数据 - 默认情况下，如果数据上次修改时间超过 3 年，则该数据被视为过期数据。您可以[自定义过期数据的定义](task-stale-data.html)。
   - 重复文件 - 您正在扫描的数据源中其他位置重复的文件。"查看显示的重复文件类型" 。

如果您的任何数据源实现了数据分层，则可以在“陈旧数据”类别中识别已经驻留在对象存储中的旧数据。

创建数据发现评估报告

数据发现评估报告对扫描环境进行了高级分析，以显示关注区域和潜在的补救步骤。结果基于数据的映射和分类。本报告的目标是提高您对数据集三个重要方面的认识：

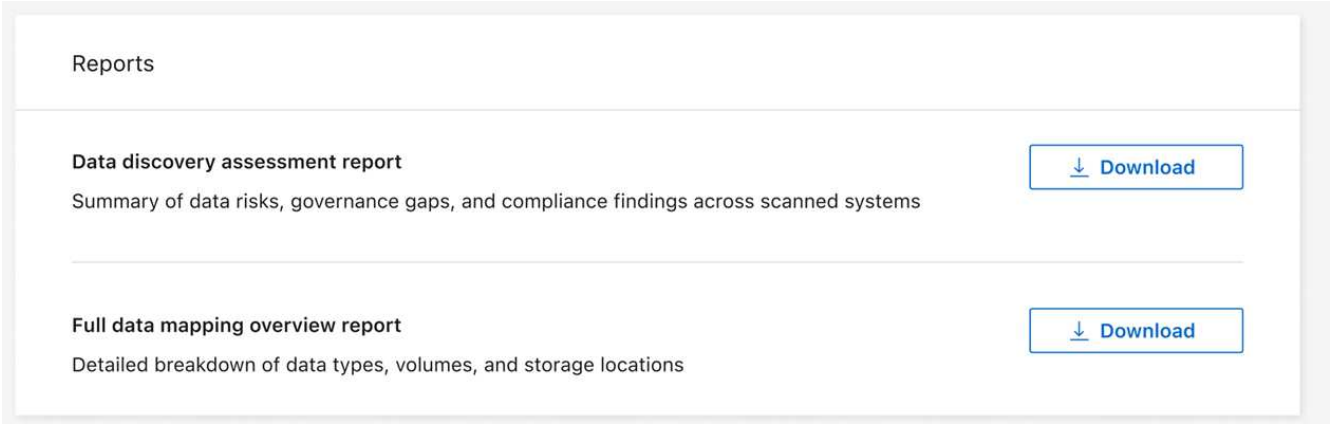
功能	描述
数据治理问题	您拥有的所有数据以及可以减少数据量以节省成本的区域的详细图片。
数据安全风险	由于访问权限广泛，您的数据可能受到内部或外部攻击的区域。
数据合规性差距	您的个人或敏感个人信息位于何处，以满足安全和 DSAR（数据主体访问请求）。

通过该报告，您可以采取以下行动：

- 通过更改保留策略或移动或删除某些数据（陈旧或重复的数据）来降低存储成本。
- 通过修改全局组管理策略来保护具有广泛权限的数据。
- 通过将 PII 移动到更安全的数据存储来保护包含个人或敏感个人信息的数据。

步骤

1. 从数据分类中，选择*治理*。
2. 在报告图块中，选择“数据发现评估报告”。



结果

数据分类会生成一份您可以查看和共享的 PDF 报告。

创建数据映射概览报告

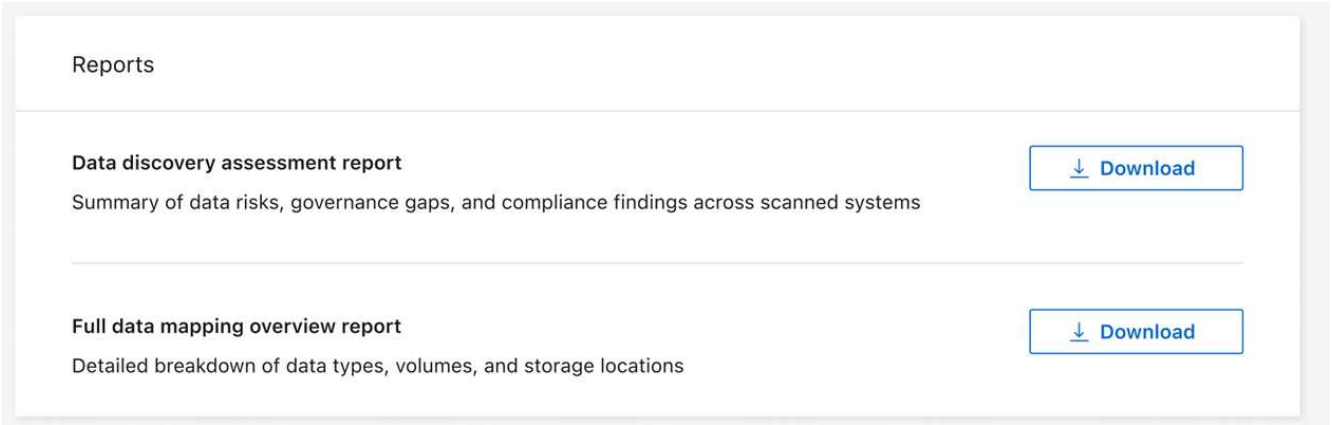
数据映射概览报告提供了存储在公司数据源中的数据的概览，以帮助您做出迁移、备份、安全和合规流程的决策。该报告总结了所有系统和数据源。它还为每个系统提供了分析。

该报告包含以下信息：

类别	描述
使用容量	对于所有系统：列出每个系统的文件数量和已用容量。对于单个系统：列出使用最多容量的文件。
数据时代	提供三个图表和图形，分别表示文件的创建时间、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	列出系统中存在于特定大小范围内的文件数。

步骤

1. 从数据分类中，选择*治理*。
2. 在报告图块中，选择*完整数据映射概览报告*。



结果

数据分类会生成一份 PDF 报告，您可以根据需要查看并发送给其他组。

如果报告大于 1 MB，则 PDF 文件将保留在数据分类实例上，您将看到有关确切位置的弹出消息。当数据分类安装在您本地的 Linux 机器上或在云中部署的 Linux 机器上时，您可以直接导航到 PDF 文件。当数据分类部署在云端时，您需要使用 SSH 授权数据分类实例下载 PDF 文件。

查看按数据敏感度列出的顶级数据存储库

数据映射概览报告中的“按敏感度级别排列的顶级数据存储库”区域列出了包含最敏感项目的前四个数据存储库（系统和数据源）。每个系统的条形图分为：

- 非敏感数据
- 个人数据
- 敏感个人数据

该数据每两小时刷新一次，可以手动刷新。

步骤

1. 要查看每个类别中的项目总数，请将光标放在栏的每个部分上。
2. 要过滤调查页面中显示的结果，请选择栏中的每个区域并进一步调查。

审查敏感数据和广泛的权限

治理仪表板的“敏感数据和广泛权限”区域显示包含敏感数据和具有广泛权限的文件的数量。该表显示以下类型的权限：

- 从横轴上最严格的权限到最宽松的限制。
- 纵轴上从最不敏感的数据到最敏感的数据。

步骤

1. 要查看每个类别中的文件总数，请将光标放在每个框上。
2. 要过滤调查页面中显示的结果，请选择一个框并进一步调查。

查看按开放权限类型列出的数据

数据映射概览报告的“打开权限”区域显示正在扫描的所有文件中每种权限的百分比。该图表显示以下类型的权限：

- 无开放权限
- 向组织开放
- 向公众开放
- 未知访问

步骤

1. 要查看每个类别中的文件总数，请将光标放在每个框上。
2. 要过滤调查页面中显示的结果，请选择一个框并进一步调查。

审查数据的年龄和大小

您可以调查数据映射概览报告的“*Age*”和“*Size*”图表中的项目，看看是否有任何数据应该删除或分层到较便宜的对象存储。

步骤

1. 在数据年龄图表中，要查看有关数据年龄的详细信息，请将光标放在图表中的某个点上。
2. 要按年龄或尺寸范围进行过滤，请选择该年龄或尺寸。
   - 数据年龄图 - 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。
   - 数据大小图 - 根据大小对数据进行分类。

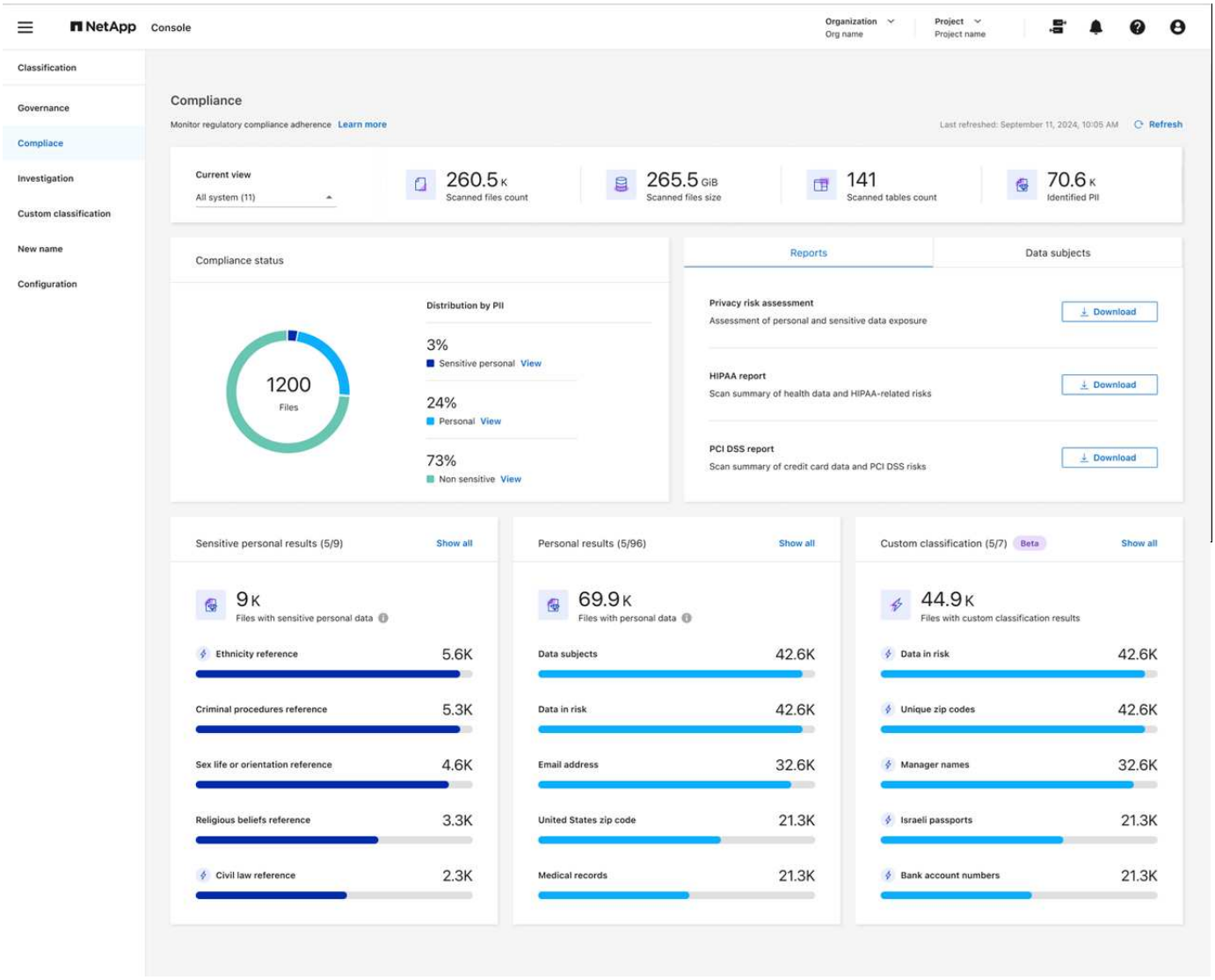
> 如果您的任何数据源实现了数据分层，则已驻留在对象存储中的旧数据可能会在“数据年龄”图中被识别。

使用NetApp Data Classification查看组织中存储的私人数据的合规性详细信息

通过查看组织中的个人数据 (PII) 和敏感个人数据 (SPII) 的详细信息来控制您的私人数据。您还可以通过查看NetApp Data Classification在您的数据中找到的类别和文件类型来获得可见性。

> 仅当您执行完整分类扫描时，才可获得文件级合规性详细信息。仅映射扫描不会产生文件级详细信息。

默认情况下，数据分类仪表板显示所有系统和数据库的合规性数据。要仅查看部分系统的数据，请选择它们。



您可以从数据调查页面过滤结果，并将结果报告下载为 CSV 文件。看"在数据调查页面中过滤数据"了解详情。

查看包含个人数据的文件

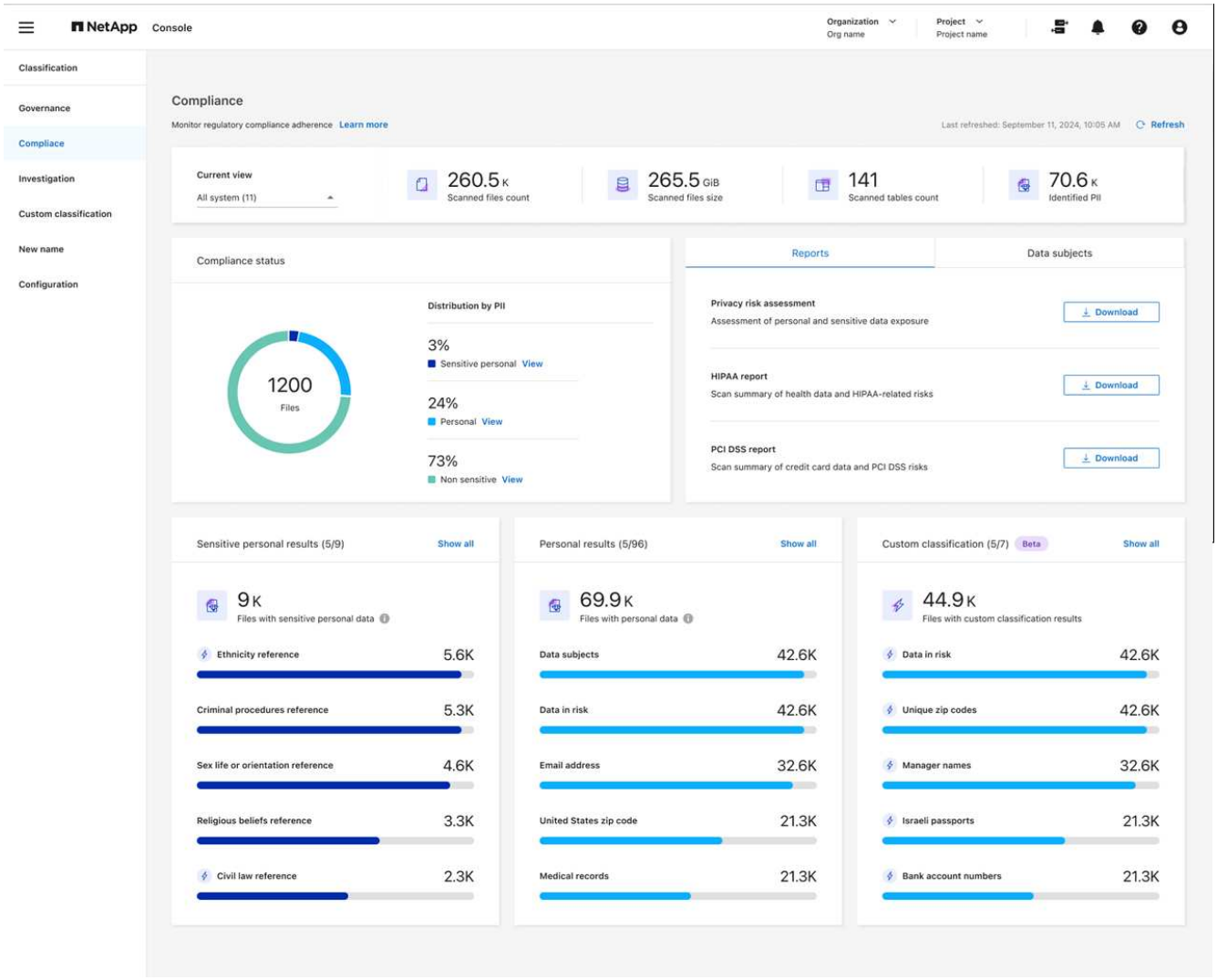
数据分类自动识别数据中的特定单词、字符串和模式（正则表达式）。"例如，信用卡号、社会保险号、银行账号、密码等等。"数据分类可在单个文件、目录（共享和文件夹）内的文件以及数据库表中识别此类信息。

您还可以创建自定义搜索词来识别特定于您组织的个人数据。有关更多信息，请参阅"创建自定义分类" 。

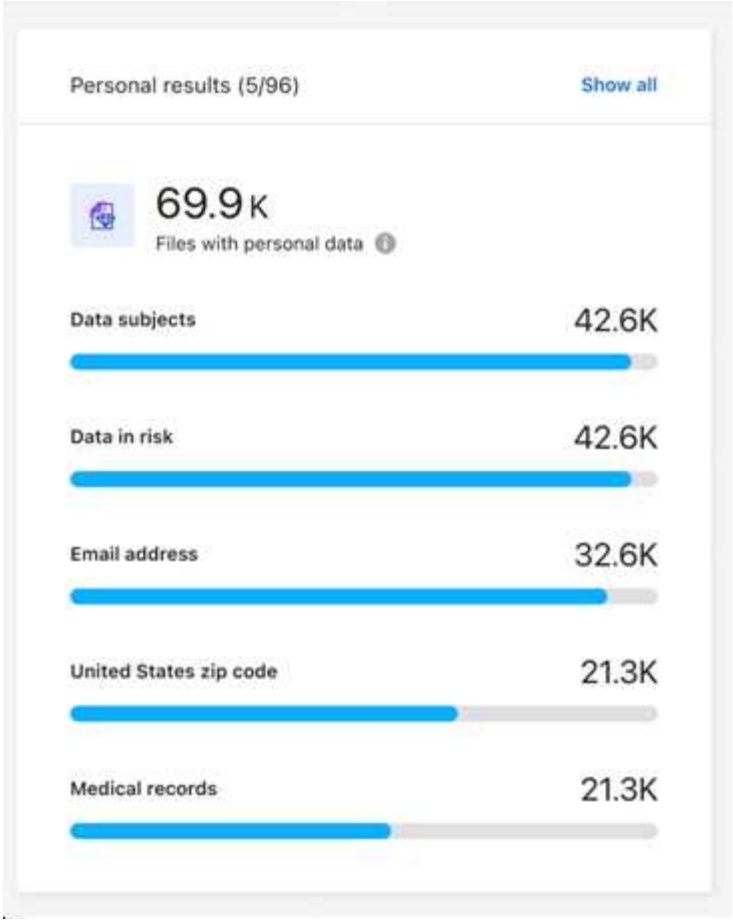
对于某些类型的个人数据，数据分类使用_邻近验证_来验证其发现。通过查找与找到的个人数据接近的一个或多个预定义关键字来进行验证。例如，如果数据分类看到旁边有一个近似词（例如，*SSN* 或 *social security*），它就会将美国社会保障号码 (SSN) 识别为 SSN。"个人资料表"显示数据分类何时使用邻近验证。

步骤

1. 从数据分类菜单中，选择“合规性”选项卡。
2. 要调查所有个人数据的详细信息，请选择个人数据百分比旁边的图标。

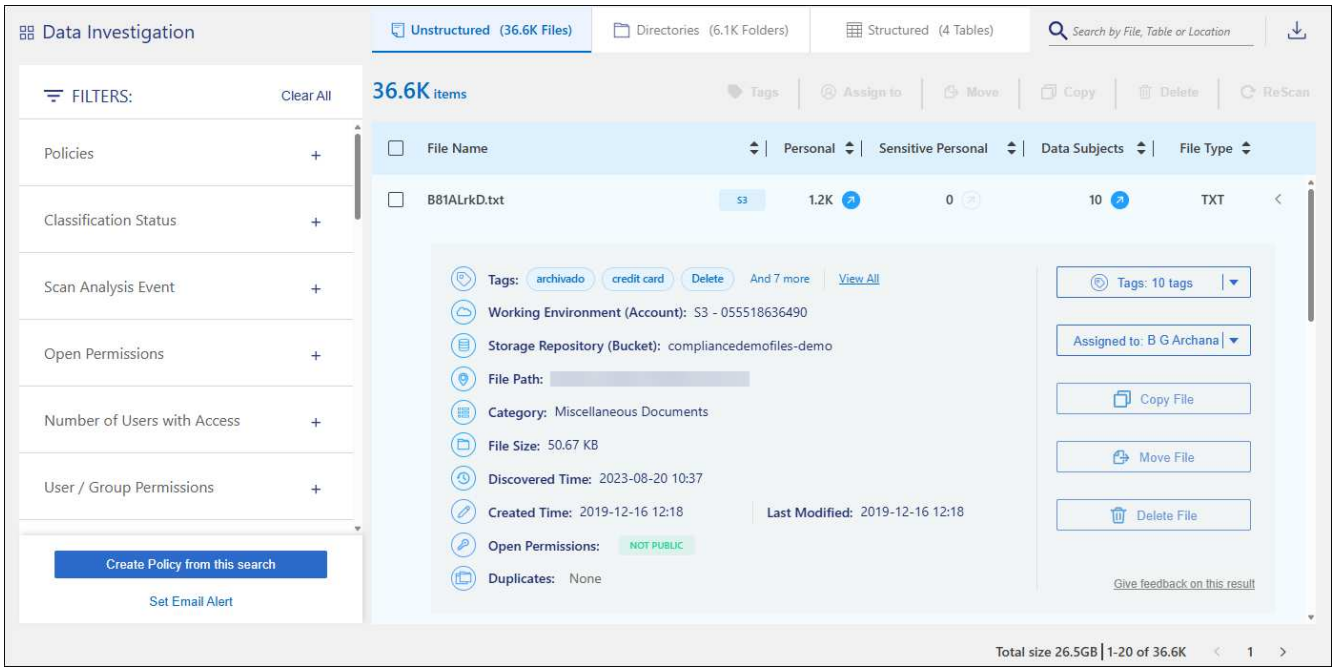


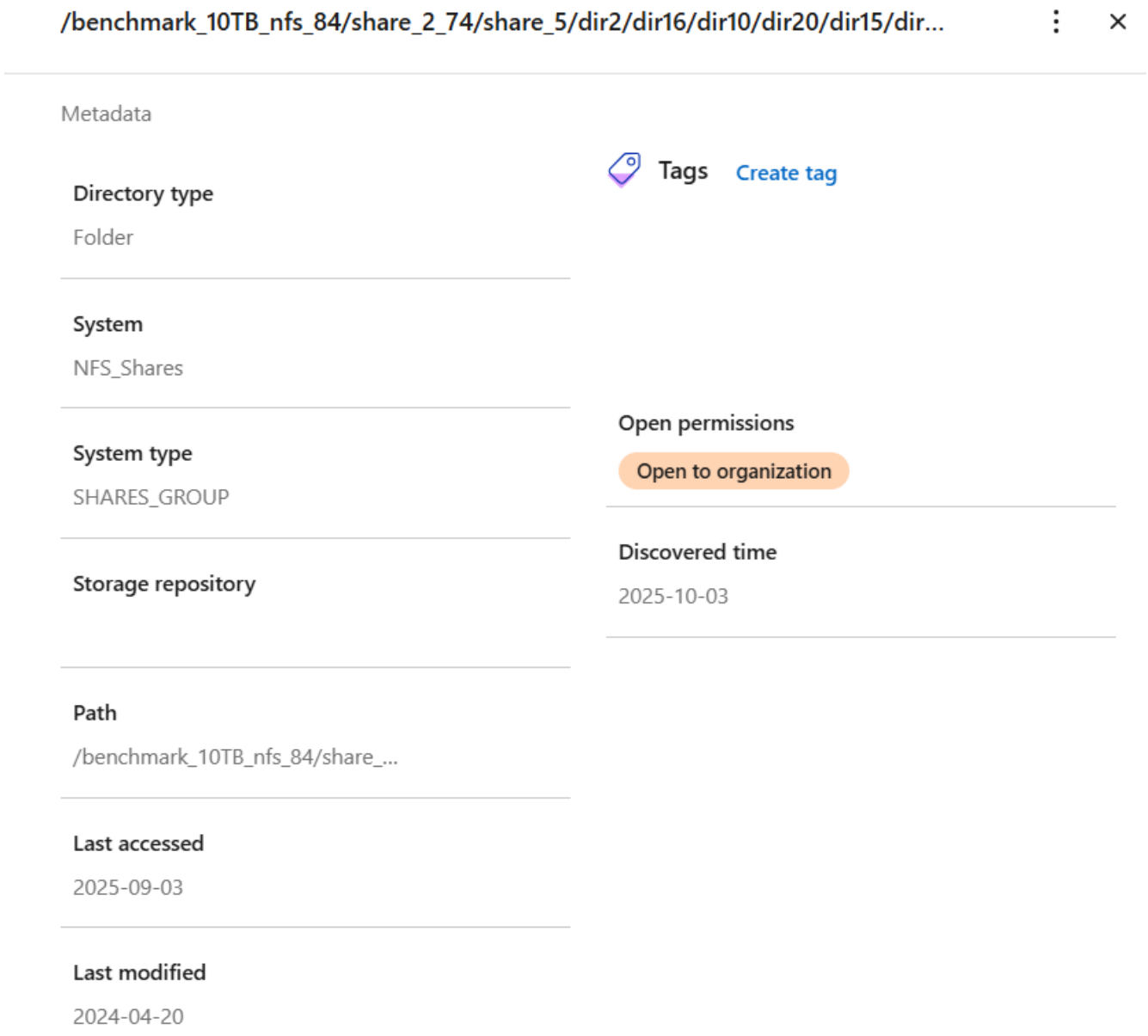
3. 要调查特定类型的个人数据的详细信息，请选择*查看全部*，然后选择特定类型的个人数据（例如电子邮件地址）的*调查结果*箭头图标。



4. 通过搜索、排序、扩展特定文件的详细信息、选择“调查结果”箭头查看屏蔽信息或下载文件列表来调查数据。

   下图显示在目录（共享和文件夹）中找到的个人数据。在“结构化”选项卡中，您可以查看数据库中的个人数据。在“非结构化”选项卡中，您可以查看文件级数据。





查看包含敏感个人数据的文件

数据分类会自动识别隐私法规所定义的特殊类型的敏感个人信息，例如 "GDPR 第 9 条和第 10 条"。例如，有关一个人的健康、种族或性取向的信息。"查看完整列表" 。数据分类可在单个文件、目录（共享和文件夹）内的文件以及数据库表中识别此类信息。

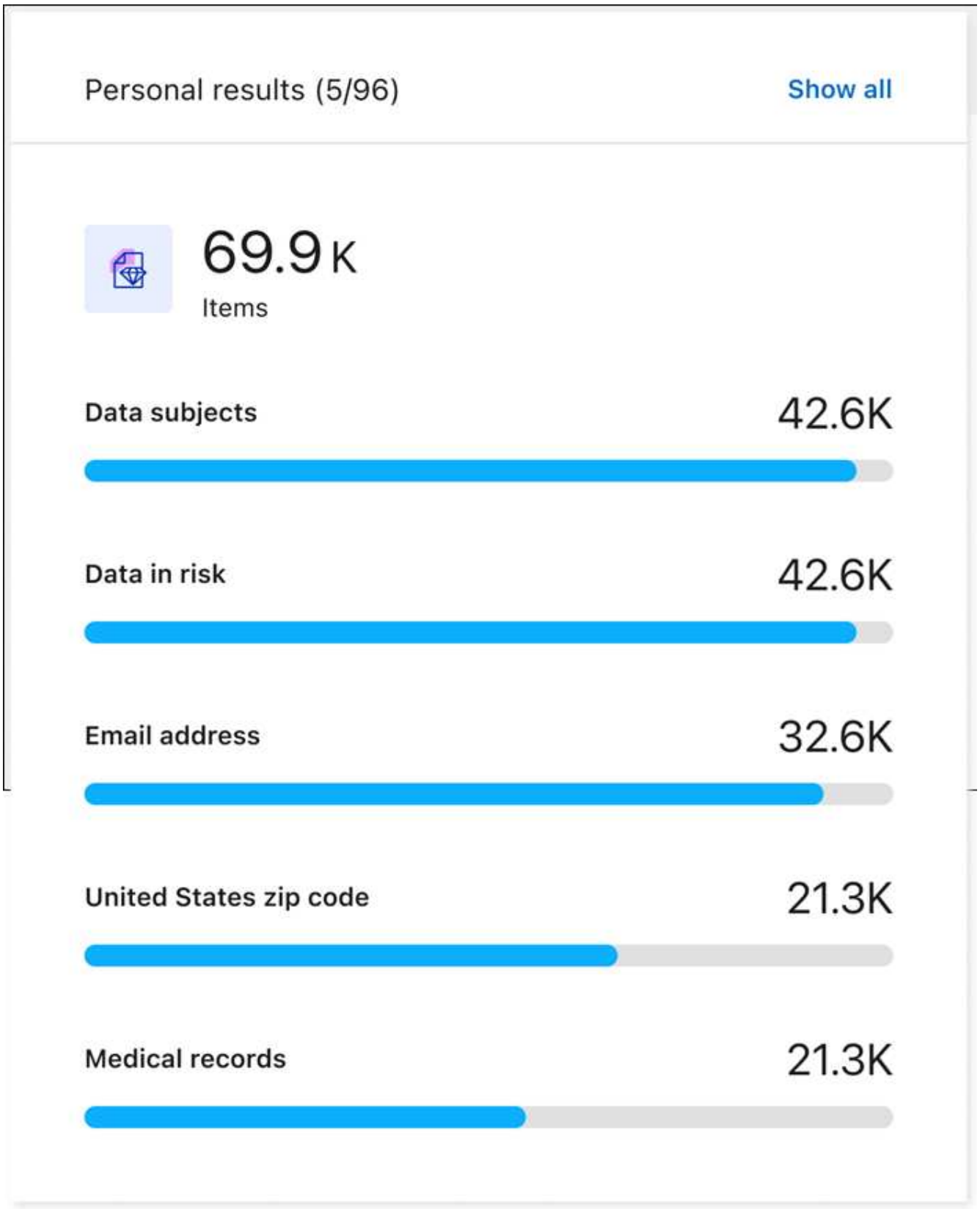
数据分类使用人工智能、自然语言处理 (NLP)、机器学习 (ML) 和认知计算 (CC) 来理解其扫描的内容的含义，以便提取实体并对其进行相应的分类。

例如，GDPR 数据的一个敏感类别是种族来源。由于其 NLP 能力，数据分类可以区分“乔治是墨西哥人”（表示 GDPR 第 9 条规定的敏感数据）和“乔治正在吃墨西哥食物”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. 从数据分类菜单中，选择*合规性*。
2. 要调查所有敏感个人数据的详细信息，请找到敏感个人信息结果卡，然后选择显示全部。



。

3. 要调查特定类型的敏感个人数据的详细信息，请选择“查看全部”，然后选择特定类型的敏感个人数据的“调查结果”箭头图标。
4. 通过搜索、排序、扩展特定文件的详细信息、单击“调查结果”查看屏蔽信息或下载文件列表来调查数据。

NetApp Data Classification中的私有数据类别

NetApp Data Classification可以在您的卷和数据库中识别多种类型的私有数据。

数据分类识别两种类型的个人数据：

- 个人身份信息（**PII**）
- 敏感个人信息（**SPII**）

如果您需要数据分类来识别其他私人数据类型，例如额外的国民身份证号码或医疗保健标识符，请联系您的客户经理。

个人数据的类型

文件中的个人数据或_个人身份信息_（PII）可以是一般个人数据或国家标识符。下表第三列标识数据分类是否使用"接近度验证"验证其对标识符的发现。

表中标明了可以识别这些项目的语言。

类型	标识符	接近度验证?	英语	德语	西班牙语	法语	日语
常规	信用卡号码	是	✓	✓	✓		✓
	数据主体	否	✓	✓	✓		
	电子邮件地址	否	✓	✓	✓		✓
	IBAN 号码（国际银行账户号码）	否	✓	✓	✓		✓
	IP 地址	否	✓	✓	✓		✓
	密码	是	✓	✓	✓		✓

类型	标识符	接近度验证?	英语	德语	西班牙语	法语	日语
国家标识符							

类型	标识符	接近度验证?	英语	德语	西班牙语	法语	日语

类型	标识符	接近度验证?	英语	德语	西班牙语	法语	日语
	希腊身份证	是	✓	✓	✓		
	匈牙利税务识别号	是	✓	✓	✓		
	爱尔兰身份证 (PPS)	是	✓	✓	✓		
	以色列身份证	是	✓	✓	✓		
	意大利税务识别号	是	✓	✓	✓		
	日本个人身份证号码（个人和公司）	是	✓	✓	✓		✓
	拉脱维亚身份证	是	✓	✓	✓		
	立陶宛身份证	是	✓	✓	✓		
	卢森堡身份证	是	✓	✓	✓		
	马耳他身份证	是	✓	✓	✓		
	国家医疗服务体系 (NHS) 号码	是	✓	✓	✓		
	新西兰银行账户	是	✓	✓	✓		
	新西兰驾驶执照	是	✓	✓	✓		
	新西兰税务局 (IRD) 号码（税号）	是	✓	✓	✓		
	新西兰 NHI（国民健康指数）号码	是	✓	✓	✓		
	新西兰护照号码	是	✓	✓	✓		
	波兰身份证 (PESEL)	是	✓	✓	✓		
	葡萄牙税务识别号（NIF）	是	✓	✓	✓		
	罗马尼亚身份证 (CNP)	是	✓	✓	✓		
	新加坡国民登记身份证（NRIC）	是	✓	✓	✓		
	斯洛文尼亚身份证 (EMSO)	是	✓	✓	✓		
	南非身份证	是	✓	✓	✓		
	西班牙税务识别号	是	✓	✓	✓		
	瑞典身份证	是	✓	✓	✓		
	英国身份证（NINO）	是	✓	✓	✓		
	美国加州驾驶执照	是	✓	✓	✓		
	美国印第安纳州驾驶执照	是	✓	✓	✓		
	美国纽约州驾驶执照	是	✓	✓	✓		
	美国德克萨斯州驾驶执照	是	✓	✓	✓		
	美国社会安全号码（SSN）	是	✓	✓	✓		

敏感个人数据的类型

数据分类可以在文件中找到以下敏感个人信息（SPII）。

以下 SPII 目前仅能以英文识别：

- 刑事诉讼参考：有关自然人的刑事定罪和犯罪的数据。

- 种族参考：有关自然人的种族或民族血统的数据。
- 健康参考：有关自然人健康的数据。
- **ICD-9-CM** 医疗代码：医疗保健行业使用的代码。
- **ICD-10-CM** 医疗代码：医疗保健行业使用的代码。
- 哲学信仰参考：有关自然人的哲学信仰的数据。
- 政治观点参考：有关自然人政治观点的数据。
- 宗教信仰参考：有关自然人的宗教信仰的数据。
- 性生活或性取向参考：有关自然人的性生活或性取向的数据。

类别类型

数据分类将您的数据分类如下。

大多数类别都可以用英语、德语和西班牙语识别。

类别	类型	英语	德语	西班牙语
金融	资产负债表	✓	✓	✓
	采购订单	✓	✓	✓
	发票	✓	✓	✓
	季度报告	✓	✓	✓
人力资源	背景调查	✓		✓
	薪酬计划	✓	✓	✓
	员工合同	✓		✓
	员工评价	✓		✓
	运行状况	✓		✓
	简历	✓	✓	✓
合法的	保密协议	✓	✓	✓
	供应商-客户合同	✓	✓	✓
营销	活动	✓	✓	✓
	会议	✓	✓	✓
操作	审计报告	✓	✓	✓
销售额	销售订单	✓	✓	
服务	射频干扰	✓		✓
	征求建议书	✓		✓
	母猪	✓	✓	✓
	培训	✓	✓	✓
支持	投诉和票务	✓	✓	✓

以下元数据也使用相同的受支持语言进行分类和识别：

- 应用程序数据
- 存档文件
- 声音的
- 数据分类业务应用数据中的面包屑
- CAD 文件
- 代码
- 腐败
- 数据库和索引文件
- 设计文件
- 电子邮件应用程序数据
- 加密（具有高熵值的文件）
- 可执行文件
- 财务应用数据
- 健康应用数据
- 图片
- 日志
- 杂项文件
- 杂项演示
- 杂项电子表格
- 杂项“未知”
- 受密码保护的文件
- 结构化数据
- 视频
- 零字节文件

文件类型

数据分类扫描所有文件的类别和元数据洞察，并在仪表板的文件类型部分显示所有文件类型。当数据分类检测个人身份信息 (PII) 或执行 DSAR 搜索时，仅支持以下文件格式：

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

所发现信息的准确性

NetApp无法保证数据分类识别的个人数据和敏感个人数据 100% 的准确性。您应该始终通过查看数据来验证信息。

根据我们的测试，下表显示了数据分类发现的信息的准确性。我们根据_精度_和_召回率_来细分它：

精确

数据分类发现的内容被正确识别的概率。例如，个人数据的准确率为 90%，意味着在被识别为包含个人信息的 10 个文件中，有 9 个实际上包含个人信息。 10 个文件中会有 1 个是误报。

记起

数据分类找到其应有内容的概率。例如，个人数据的召回率为 70%，意味着数据分类可以识别出组织中 10 个文件中实际包含个人信息的 7 个。数据分类会遗漏 30% 的数据，并且不会出现在仪表板中。

我们正在不断提高结果的准确性。这些改进将在未来的数据分类版本中自动提供。

类型	精确	记起
个人数据 - 一般	90%-95%	60%-80%
个人数据 - 国家标识符	30%-60%	40%-60%
敏感个人数据	80%-95%	20%-30%
类别	90%-97%	60%-80%

在NetApp Data Classification中创建自定义分类

NetApp Data Classification允许您创建自定义类别或个人标识符，以识别特定于您组织监管和合规要求的数据。

数据分类支持两种类型的自定义分类器：类别和个人标识符。自定义类别是根据您上传的一组文件创建的，数据分类功能会根据这些文件创建一个 AI 模型，以识别您组织中的类似数据（例如，一家健康研究公司可能会创建一个临床分析类别）。使用关键字列表或正则表达式 (regex) 创建自定义个人标识符，以识别贵组织特有的、可能构成合规风险的信息。

所有自定义分类都可以在自定义分类控制面板中找到。

创建自定义个人标识符

数据分类功能允许您使用上下文关键字或正则表达式创建自定义个人标识符，以识别贵组织特有的数据。

关键词要求

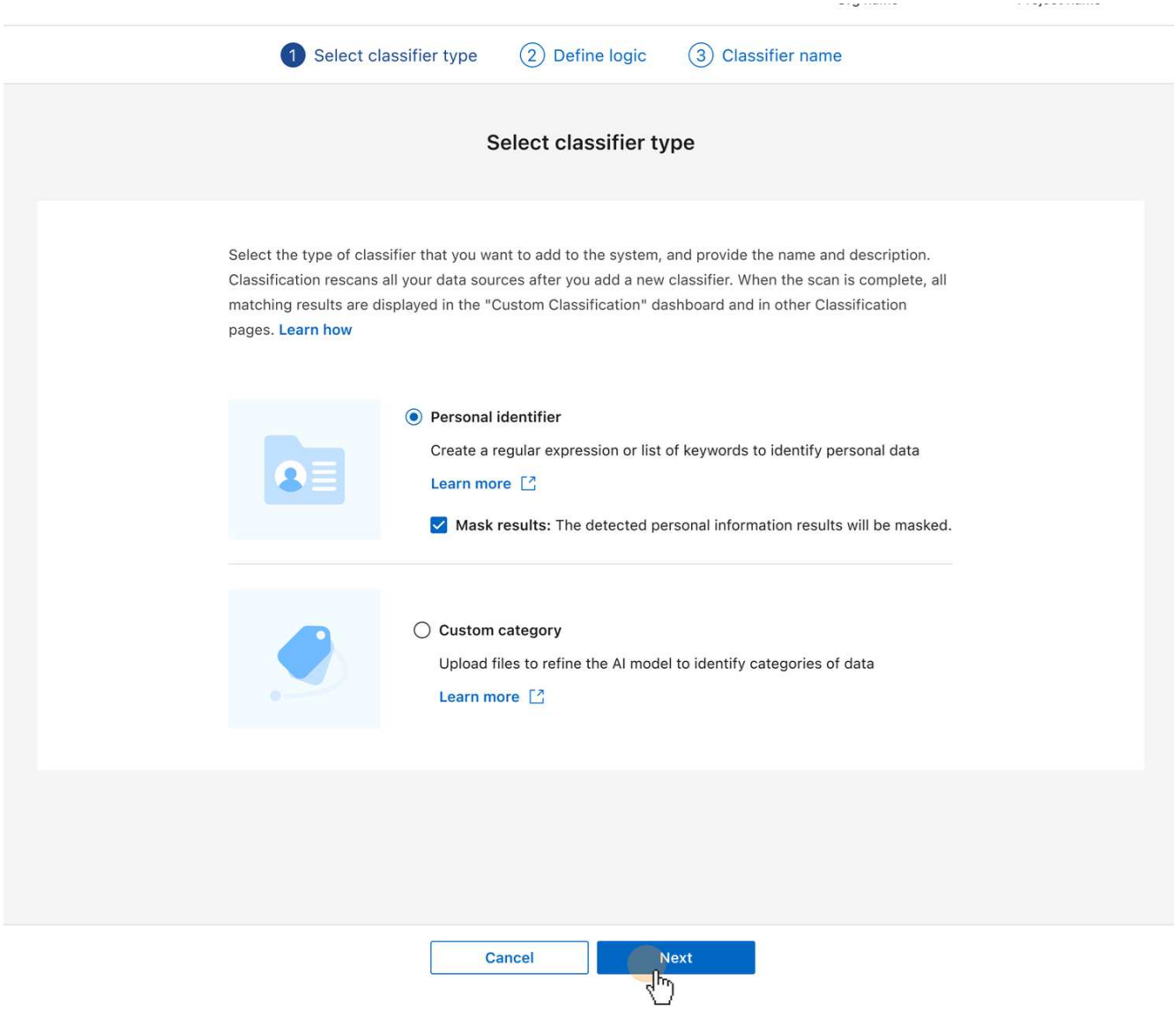
如果您使用关键词列表创建个人标识符，则该列表必须满足以下要求：

- 关键词输入不区分大小写。
- 关键词必须至少包含三个字符。长度少于三个字符的单词将被忽略。
- 重复的词语只会添加一次。
- 关键词总数不能超过 50 万个字符。列表中必须至少包含一个关键词。

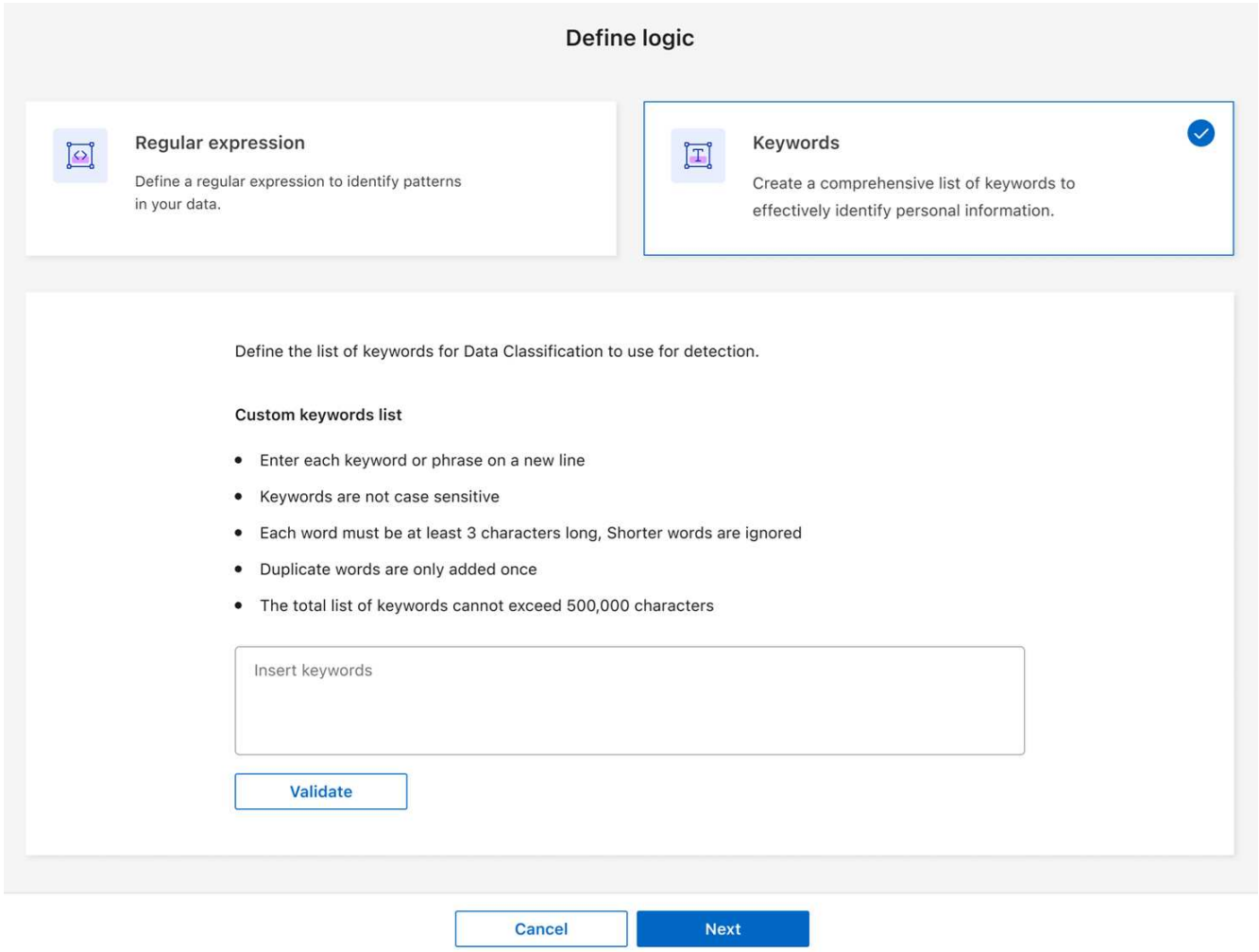
步骤

1. 选择自定义分类选项卡。
2. 选择**+** 新建分类器以创建自定义分类器。

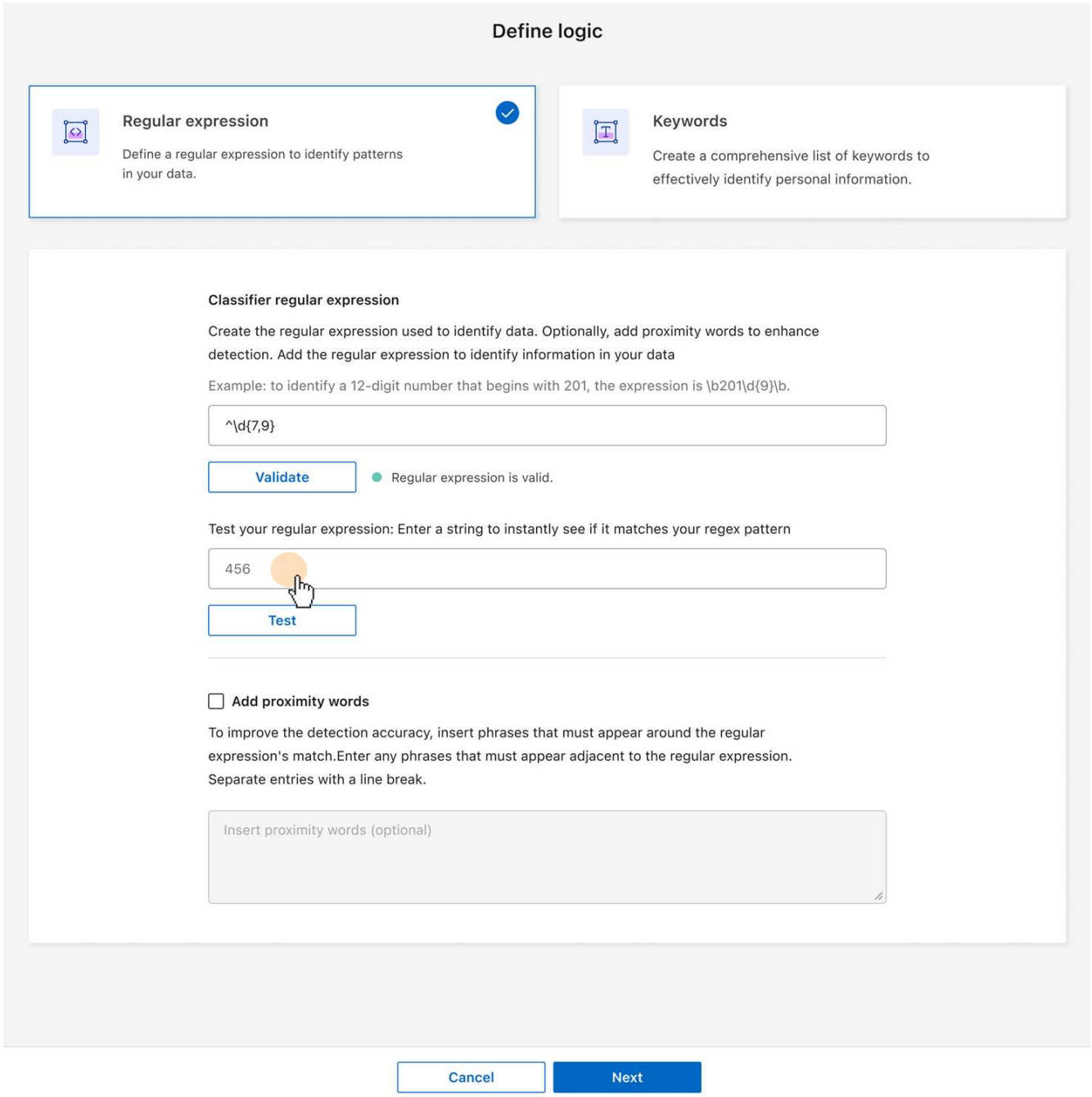
3. 请选择*个人标识符*。（可选）选择“屏蔽结果”以屏蔽检测到的个人数据。
4. 选择下一步。



5. 要添加带关键词的分类器，请选择关键词。请输入关键词列表，每个关键词占一行。请确保关键词符合要求。



要将分类器添加为正则表达式，请选择正则表达式，然后添加一个模式来检测数据的特定信息。选择验证以确认您输入的语法正确。



a. （可选）输入一个应该与正则表达式模式匹配的示例字符串，然后选择测试进行检查。

b. （可选）添加邻近词。如果添加邻近词，数据分类仅在邻近词与匹配字符串相邻时才标记正则表达式模式。

6. 选择下一步。
7. 输入分类器名称和描述，以便在仪表板中标识自定义类别。
8. 选择保存以创建自定义个人标识符。

创建自定义个人标识符后，其结果将在下次计划扫描中捕获。为了更快地获取结果，请执行按需扫描。要查看结果，请参阅 生成合规性报告。

创建自定义类别

通过自定义类别，您可以对特定于您组织的数据进行分类。自定义类别是根据您上传的文本文件创建的，数据分类功能会根据这些文件创建一个人工智能模型，以识别其他文件中的类似信息。

训练数据要求

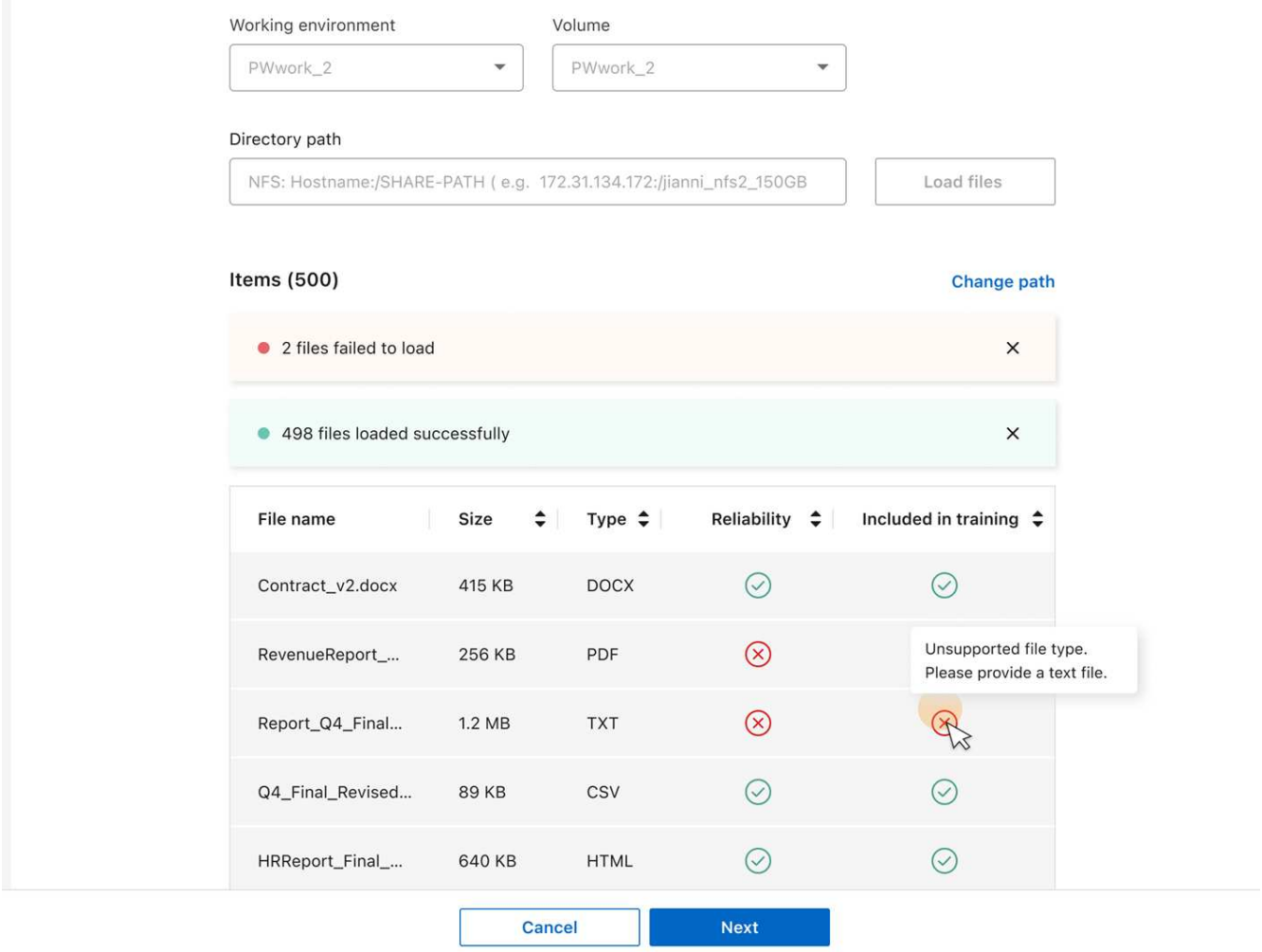
- 训练数据集必须至少包含 25 个文件。最大文件数为 1,000。
- 所有文件必须直接位于您提供的文件路径中。
- 所有文件必须大于 100 字节。
- 数据分类训练数据必须是以下文件类型之一：CSV、DOCX、DOC、GZ、JSON、PDF、PPTX、TXT、RTT、XLS 或 XLSX。您可以上传所有支持的文件类型的组合。

步骤

1. 在NetApp Data Classification中，选择“自定义分类”。
2. 选择 **+** 新建分类器。
3. 选择“自定义类别”作为分类器类型，然后下一步。
4. 使用一系列基于文本的文件来定义自定义类别的逻辑。请提供*工作地址*的IP地址，然后从下拉菜单中选择*音量*。

   输入包含训练数据的目录的目录路径。

5. 选择“加载文件”进行数据分类，以执行文件检查。您可以查看文件摘要，其中列出了文件名、大小、类型和备注（如果该文件被认为适合用于培训）。



a. 要更改文件路径或重新上传文件，请选择更改路径，然后输入数据并再次加载文件。

6. 当您对上传的文件满意后，请选择下一步。
7. 输入分类器名称和描述，以便在仪表板中标识自定义类别。
8. 选择保存以创建自定义类别。

结果

创建自定义类别后，其结果将在下次计划扫描中捕获。为了更快地获取结果，请手动启动扫描。

编辑自定义分类器

创建个人标识符后，您可以修改其逻辑。您无法更改个人标识符的类型或逻辑类型；例如，您无法将自定义类别更改为自定义个人标识符。您也不能将基于关键字的自定义标识符更改为基于正则表达式的自定义标识符。

步骤

1. 在NetApp Data Classification中，选择“自定义分类”。
2. 确定要删除的分类器，然后选择操作菜单 ... 在它那一行的末尾。
3. 选择编辑逻辑。
4. 如果要修改关键词，请添加、删除或编辑相应的关键词。如果要修改正则表达式，请输入新的正则表达式并进行验证。（可选）添加邻近关键词。

5. 选择“保存”以应用更改。

删除自定义分类器

1. 在NetApp Data Classification中，选择“自定义分类”。
2. 确定要删除的分类器，然后选择操作菜单 ... 在它那一行的末尾。
3. 选择删除分类器。

下一步

- 生成合规性报告

使用NetApp Data Classification调查组织中存储的数据

数据调查仪表板显示文件和目录级别的数据洞察，使您能够对结果进行排序和过滤。数据调查页面提供有关文件和目录元数据和权限的见解以及识别重复文件。通过文件、目录和数据库级别的洞察，您可以采取措施来提高组织的合规性并节省存储空间。数据调查页面还支持移动、复制和删除文件。

要从调查页面获得见解，您必须对数据源执行完整的分类扫描。仅进行过映射扫描的数据源不会显示文件级别的详细信息。

数据调查结构

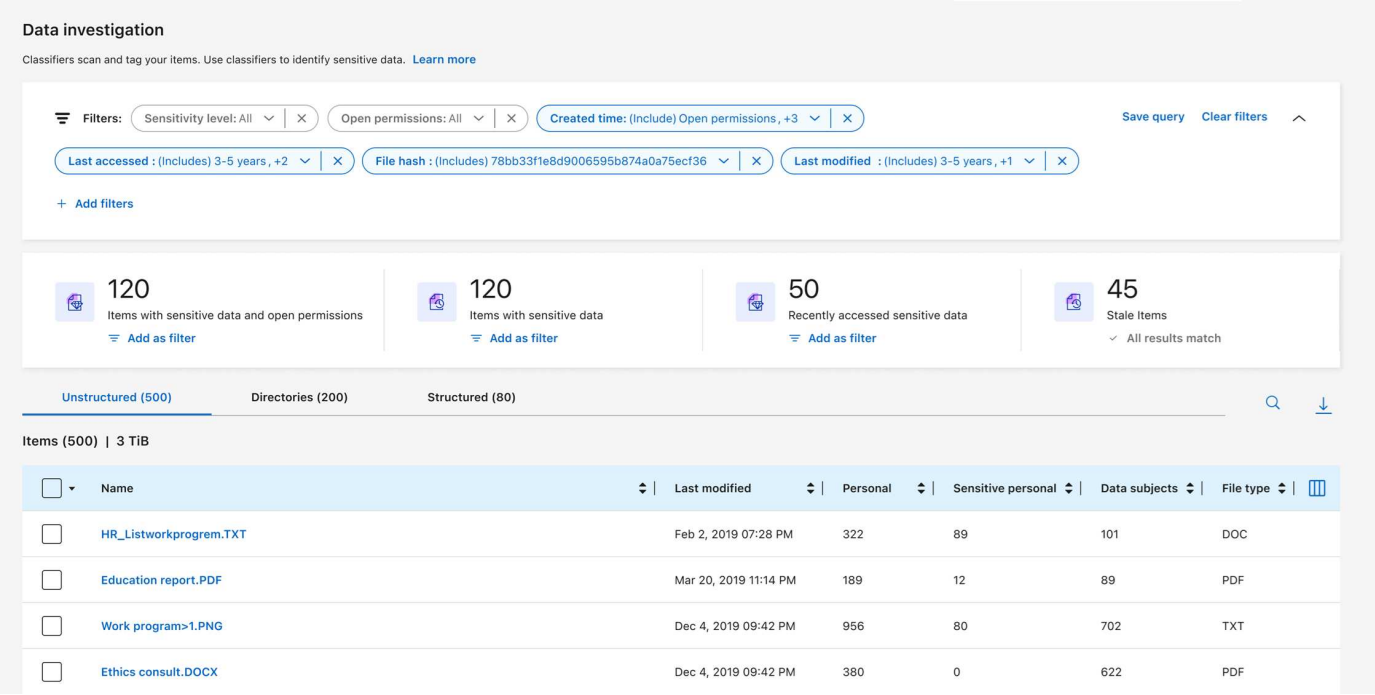
数据调查页面将数据分类到三个选项卡中：

- 非结构化数据：文件数据
- 目录：文件夹和文件共享
- 结构化：数据库

数据过滤器

数据调查页面提供了许多过滤器来对您的数据进行分类，以便您可以找到所需的数据。您可以同时使用多个过滤器。

要添加过滤器，请选择添加过滤器按钮。



过滤敏感度和内容

使用以下过滤器查看您的数据中包含多少敏感信息。

筛选器	详细信息
类别	选择"类别类型"。
敏感度等级	选择敏感度级别：个人、敏感个人或非敏感。
标识符数量	选择每个文件检测到的敏感标识符的范围。包括个人数据和敏感个人数据。在目录中过滤时，数据分类会将每个文件夹（和子文件夹）中所有文件的匹配结果汇总。注意：2023 年 12 月（版本 1.26.6）版本删除了按目录计算个人身份信息 (PII) 数据数量的选项。
个人资料	选择"个人数据的类型"。
敏感个人信息	选择"敏感个人数据的类型"。
数据主体	输入数据主体的全名或已知标识符。"在此处了解有关数据主体的更多信息" 。

过滤用户所有者和用户权限

使用以下过滤器查看文件所有者和访问数据的权限。

筛选器	详细信息
开放权限	选择数据和文件夹/共享内的权限类型。
用户/组权限	选择一个或多个用户名和/或组名，或者输入部分名称。
文件所有者	输入文件所有者名称。
有访问权限的用户数	选择一个或多个类别范围以显示哪些文件和文件夹对一定数量的用户开放。

按时间顺序过滤

使用以下过滤器根据时间标准查看数据。

筛选器	详细信息
创建时间	选择文件创建的时间范围。您还可以指定自定义时间范围来进一步优化搜索结果。
发现时间	选择数据分类发现文件的时间范围。您还可以指定自定义时间范围来进一步优化搜索结果。
上次修改	选择文件最后修改的时间范围。您还可以指定自定义时间范围来进一步优化搜索结果。
上次访问	选择文件或目录*上次被访问的时间范围。您还可以指定自定义时间范围来进一步优化搜索结果。对于数据分类扫描的文件类型，这是数据分类最后一次扫描该文件的时间。

{星号} 目录的上次访问时间仅适用于 NFS 或 CIFS 共享。

过滤元数据

使用以下过滤器根据位置、大小和目录或文件类型查看数据。

筛选器	详细信息
文件路径	输入最多 20 条要在查询中包含或排除的部分或完整路径。如果同时输入包含路径和排除路径，数据分类会首先在包含路径中找到所有文件，然后从排除路径中删除文件，然后显示结果。请注意，在此过滤器中使用“*”没有任何效果，并且您无法从扫描中排除特定文件夹 - 配置共享下的所有目录和文件都将被扫描。
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。
文件哈希	输入文件的哈希值即可查找特定文件，即使名称不同。

过滤器存储类型

使用以下过滤器按存储类型查看数据。

筛选器	详细信息
系统类型	选择系统类型。
系统环境名称	选择特定系统。
存储库	选择存储库，例如卷或模式。

过滤查询

使用以下过滤器按已保存的查询查看数据。

筛选器	详细信息
已保存的查询	选择一个或多个已保存的查询。前往"已保存的查询选项卡"查看现有已保存查询的列表并创建新查询。
标签	选择"一个或多个标签"分配给您的文件。

过滤分析状态

使用以下过滤器按数据分类扫描状态查看数据。

筛选器	详细信息
分析状态	选择一个选项来显示“等待首次扫描”、“已完成扫描”、“等待重新扫描”或“扫描失败”的文件列表。
扫描分析事件	选择是否要查看由于数据分类无法恢复上次访问时间而未分类的文件，或者即使数据分类无法恢复上次访问时间但已分类的文件。

"查看有关“上次访问时间”时间戳的详细信息"有关使用扫描分析事件进行过滤时调查页面中出现的项目的更多信息。

按重复项过滤数据

使用以下过滤器查看存储中重复的文件。

筛选器	详细信息
重复项	选择文件是否在存储库中重复。

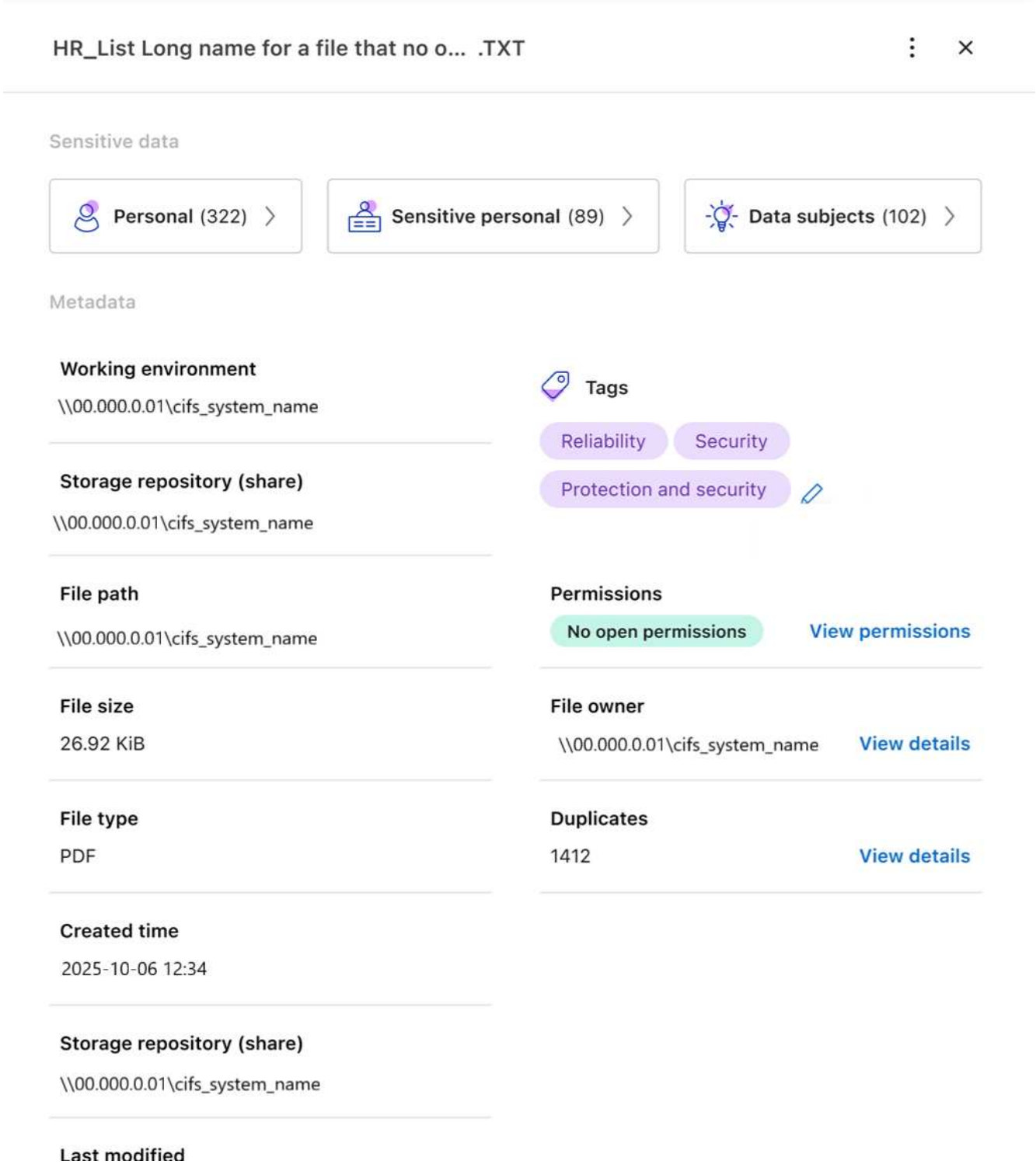
查看文件元数据

除了显示文件所在的系统和卷之外，元数据还显示更多信息，包括文件权限、文件所有者以及该文件是否有重复。如果您打算"创建已保存的查询"因为您可以看到可用于过滤数据的所有信息。

信息的可用性取决于数据来源。例如，数据库文件的卷名和权限不共享。

步骤

1. 从数据分类菜单中，选择*调查*。
2. 在右侧的数据调查列表中，选择向下插入符号 在任意单个文件的右侧查看文件元数据。



3. 或者，您可以使用*创建标签*按钮为文件创建或添加标签。从下拉菜单中选择一个现有标签或使用 + 添加 按钮添加一个新标签。标签可用于过滤数据。

查看文件和目录的用户权限

要查看有权访问文件或目录的所有用户或组的列表以及他们拥有的权限类型，请选择“查看所有权限”。此选项仅适用于 CIFS 共享中的数据。

如果您使用安全标识符 (SID) 而不是用户名和组名，则应该将 Active Directory 集成到数据分类中。有关更多信

息，请参阅"将 Active Directory 添加到数据分类" 。

步骤

1. 从数据分类菜单中，选择*调查*。
2. 在右侧的数据调查列表中，选择向下插入符号 在任意单个文件的右侧查看文件元数据。
3. 要查看有权访问文件或目录的所有用户或组的列表以及他们拥有的权限类型，请在“打开权限”字段中选择“查看所有权限”。

   数据分类在列表中显示最多 100 个用户。

4. 选择向下插入符号 任何群组的按钮即可查看属于该群组的用户列表。

   您可以展开该组的某个级别来查看属于该组的用户。

5. 选择用户或组的名称以刷新调查页面，以便您可以看到该用户或组有权访问的所有文件和目录。

检查存储系统中的重复文件

您可以检查存储系统中是否存储了重复的文件。如果您想确定可以节省存储空间的区域，这将非常有用。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不必要地重复也是很好的。

数据分类会比较所有文件（数据库除外）是否存在重复项，如果存在重复项，则进行以下操作：

- 1 MB 或更大
- 或包含个人信息或敏感个人信息

数据分类使用散列技术来确定重复文件。如果一个文件的哈希码与另一个文件相同，即使文件名不同，这两个文件也是完全相同的副本。

步骤

1. 从数据分类菜单中，选择*调查*。
2. 在“过滤器”窗格中，选择“文件大小”以及“重复项”（“有重复项”）以查看您的环境中哪些特定大小范围的文件是重复的。
3. 或者，下载重复文件的列表并将其发送给存储管理员，以便他们可以决定可以删除哪些文件（如果有）。
4. 您可以选择删除、标记或移动重复的文件。选择您想要执行操作的文件，然后选择适当的操作。

查看特定文件是否重复

您可以查看单个文件是否有重复。

步骤

1. 从数据分类菜单中，选择*调查*。
2. 在数据调查列表中，选择 在任意单个文件的右侧查看文件元数据。

   如果文件存在重复，则此信息将显示在“*Duplicates*”字段旁边。

3. 要查看重复文件的列表及其位置，请选择“查看详细信息”*。

4. 在下一页中选择“查看重复项”以查看调查页面中的文件。
5. 您可以选择删除、标记或移动重复的文件。选择您想要执行操作的文件，然后选择适当的操作。

您可以使用此页面提供的“文件哈希”值并将其直接输入到调查页面中，以便随时搜索特定的重复文件 - 或者您可以在已保存的查询中使用它。

下载您的报告

您可以以 CSV 或 JSON 格式下载过滤结果。

如果数据分类正在扫描文件（非结构化数据）、目录（文件夹和文件共享）和数据库（结构化数据），则最多可以下载三个报告文件。

文件被分割成具有固定行数或记录数的文件：

- JSON：每份报告 100,000 条记录，生成大约需要 5 分钟
- CSV：每份报告 200,000 条记录，生成大约需要 4 分钟

您可以下载 CSV 文件的版本以在此浏览器中查看。此版本限制为 10,000 条记录。

可下载报告包含的内容

*非结构化文件数据报告*包含有关您的文件的以下信息：

- 文件名
- 位置类型
- 系统名称
- 存储库（例如，卷、存储桶、共享）
- 存储库类型
- 文件路径
- 文件类型
- 文件大小（MB）
- 创建时间
- 上次修改时间
- 上次访问
- 文件所有者
  - 配置 Active Directory 时，文件所有者数据包括帐户名称、SAM 帐户名称和电子邮件地址。
- 类别
- 个人信息
- 敏感个人信息
- 开放权限

- 扫描分析错误
- 删除检测日期

  删除检测日期标识文件被删除或移动的日期。这使您能够识别敏感文件何时被移动。已删除的文件不会计入仪表板或调查页面上显示的文件数量。这些文件仅出现在 CSV 报告中。

*非结构化目录数据报告*包括有关您的文件夹和文件共享的以下信息：

- 系统类型
- 系统名称
- 目录名称
- 存储库（例如文件夹或文件共享）
- 目录所有者
- 创建时间
- 发现时间
- 上次修改时间
- 上次访问
- 开放权限
- 目录类型

*结构化数据报告*包含有关数据库表的以下信息：

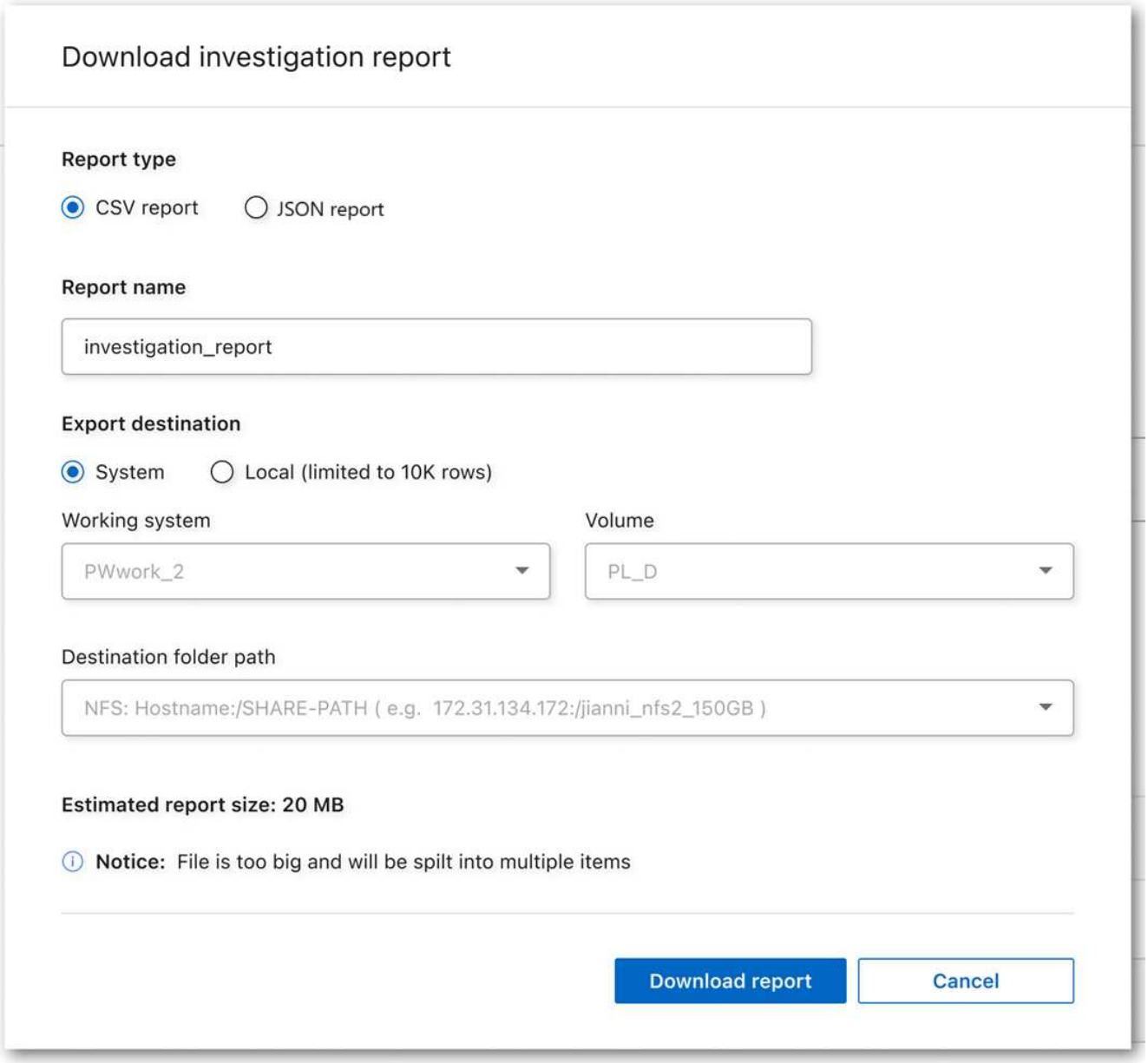
- 数据库表名称
- 位置类型
- 系统名称
- 存储库（例如，架构）
- 列数
- 行数
- 个人信息
- 敏感个人信息

生成报告的步骤

1. 从数据调查页面中，选择 页面右上方的按钮。
2. 选择报告类型：CSV 或 JSON。
3. 输入报告名称。
4. 要下载完整的报告，请选择系统，然后从相应的下拉菜单中选择系统和卷。提供目标文件夹路径。

   要在浏览器中下载报告，请选择本地。请注意，此选项将报告限制为前 10,000 行，并且仅限于 **CSV** 格式。如果您选择本地，则无需填写任何其他字段。

5. 选择下载报告。



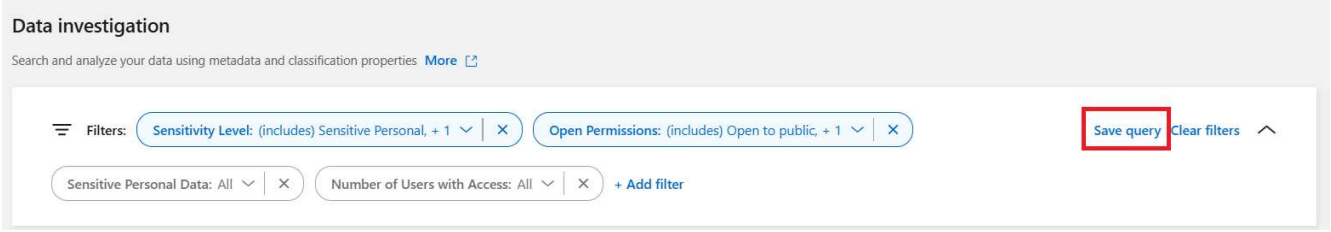
结果

对话框中将显示一条消息，提示正在下载报告。

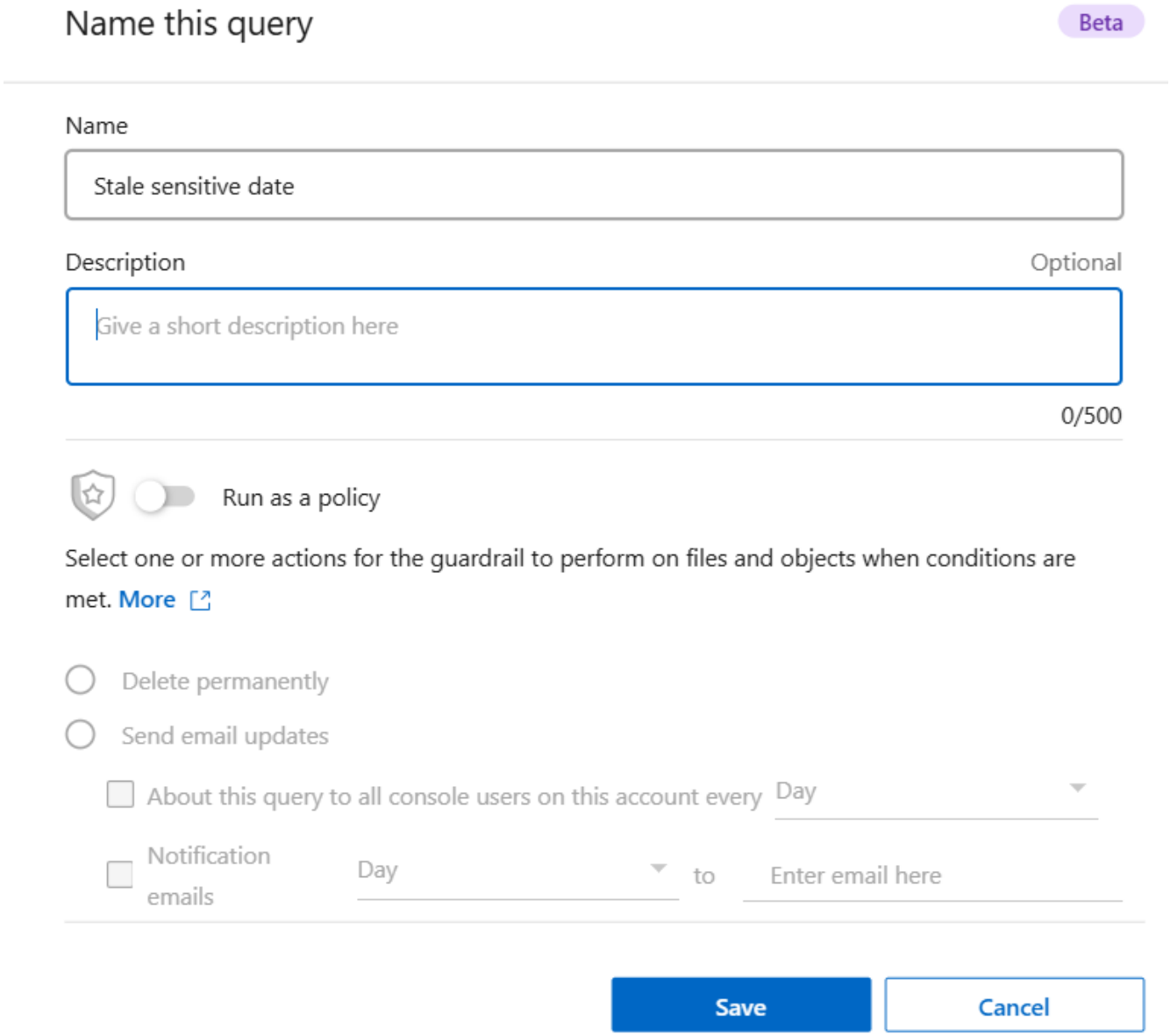
根据选定的过滤器创建已保存的查询

步骤

1. 在调查选项卡中，通过选择要使用的过滤器来定义搜索。看"在调查页面中过滤数据"了解详情。
2. 一旦您根据自己的喜好设置了所有过滤器特性，请选择*保存查询*。



3. 为保存的查询命名并添加描述。该名称必须是唯一的。
4. 您可以选择将查询保存为策略：
   a. 要将查询保存为策略，请切换*作为策略运行*开关。
   b. 选择*永久删除*或*发送电子邮件更新*。如果您选择电子邮件更新，您可以每天、每周或每月通过电子邮件将查询结果发送给所有控制台用户。或者，您可以以相同的频率将通知发送到特定的电子邮件地址。
5. 选择*保存*。



创建搜索或策略后，您可以在已保存的查询选项卡中查看它。

结果可能需要最多 15 分钟才会显示在“已保存的查询”页面上。

使用NetApp Data Classification管理已保存的查询

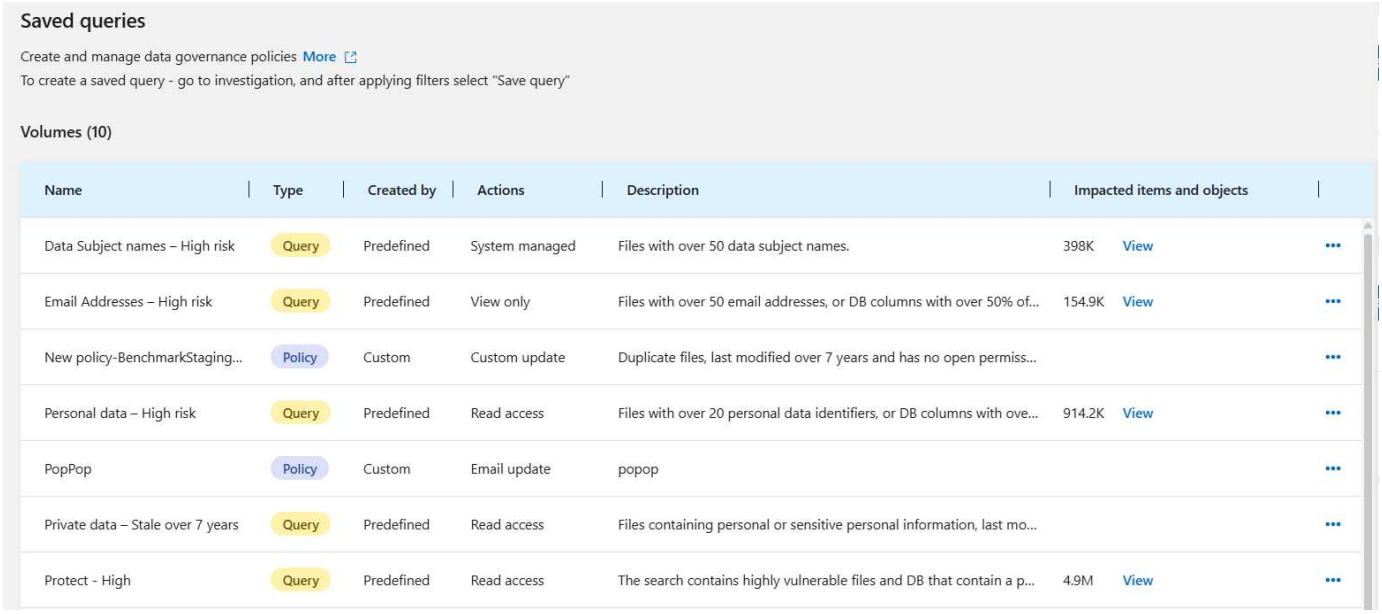
NetaApp 数据分类支持保存您的搜索查询。使用已保存的查询，您可以创建自定义过滤器来对数据调查页面的常见查询进行排序。数据分类还包括基于常见请求的预定义保存的查询。

合规性仪表板中的“已保存的查询”选项卡列出了此数据分类实例上可用的所有预定义和自定义已保存查询。

已保存的查询也可以保存为策略。查询过滤数据，而策略允许您对数据采取行动。通过策略：您可以删除发现的

数据或发送有关发现的数据的电子邮件更新。

已保存的查询也会出现在调查页面的过滤器列表中。



在调查页面中查看已保存的查询结果

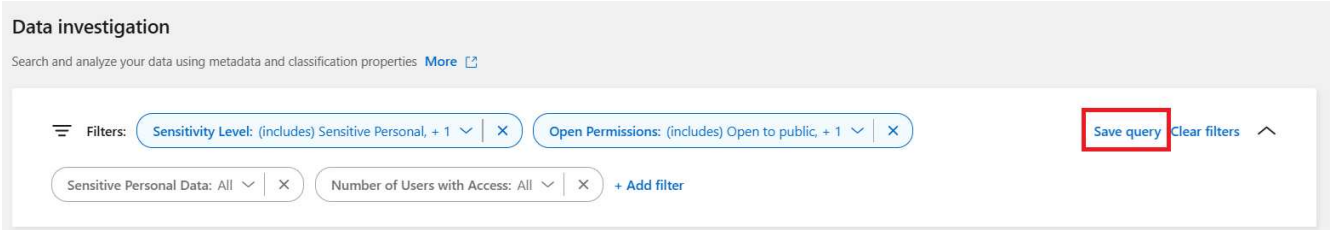
要在调查页面中显示已保存查询的结果，请选择 按钮进行特定搜索，然后选择*调查结果*。

创建已保存的查询和策略

您可以创建自己的自定义已保存查询，以提供特定于您组织的查询结果。返回符合搜索条件的所有文件和目录（共享和文件夹）的结果。

步骤

1. 在调查选项卡中，通过选择要使用的过滤器来定义搜索。看"在调查页面中过滤数据"了解详情。
2. 一旦您根据自己的喜好设置了所有过滤器特性，请选择*保存查询*。

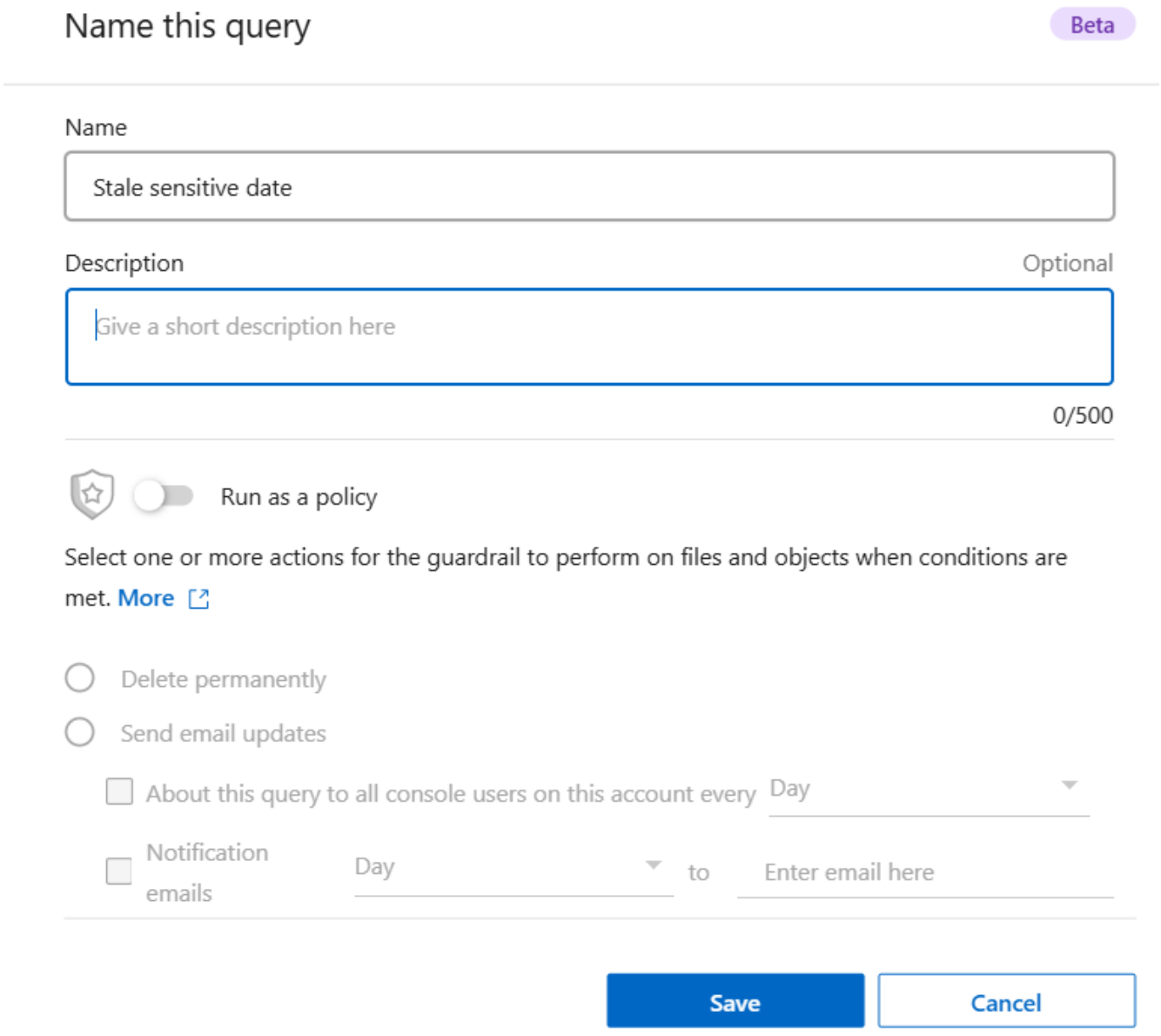


3. 为保存的查询命名并添加描述。该名称必须是唯一的。
4. 您可以选择将查询保存为策略：

a. 要将查询保存为策略，请切换*作为策略运行*开关。

b. 选择*永久删除*或*发送电子邮件更新*。如果您选择电子邮件更新，您可以每天、每周或每月通过电子邮件将查询结果发送给所有控制台用户。或者，您可以以相同的频率将通知发送到特定的电子邮件地址。

5. 选择*保存*。



创建搜索或策略后，您可以在已保存的查询选项卡中查看它。

编辑已保存的查询或策略

您可以修改已保存查询的名称和描述。您还可以将查询转换为策略，反之亦然。

您不能修改默认保存的查询。您不能修改已保存查询的过滤器。您可以交替查看已保存查询的调查结果，更改或修改过滤器，然后将其保存为新查询或策略。

步骤

1. 在“已保存的查询”页面中，选择要更改的搜索的“编辑搜索”。

2. 对名称和描述字段进行更改。仅更改名称和描述字段。

   您可以选择将查询转换为策略，或将策略转换为已保存的查询。根据需要切换*作为策略运行*开关。..如果您要将查询转换为策略，请选择*永久删除*或*发送电子邮件更新*。如果您选择电子邮件更新，您可以每天、每周或每月通过电子邮件将查询结果发送给所有控制台用户。或者，您可以以相同的频率将通知发送到特定的电子邮件地址。

3. 选择“保存”以完成更改。

删除已保存的查询

如果您不再需要任何自定义保存的查询或策略，可以将其删除。您不能删除默认保存的查询。

要删除已保存的查询，请选择 按钮进行特定搜索，选择*删除查询*，然后在确认对话框中再次选择*删除查询*。

默认查询

数据分类提供以下系统定义的搜索查询：

- 数据主体姓名 - 高风险

  包含超过 50 个数据主体名称的文件

- 电子邮件地址 - 高风险

  包含超过 50 个电子邮件地址的文件或数据库列中超过 50% 的行包含电子邮件地址

- 个人数据 - 高风险

  包含超过 20 个个人数据标识符的文件或数据库列中超过 50% 的行包含个人数据标识符

- 私人数据 - 已过期 **7** 年以上

  包含个人或敏感个人信息的文件，上次修改时间超过 7 年

- 保护 - 高

  包含密码、信用卡信息、IBAN 号码或社会安全号码的文件或数据库列

- 保护 - 低

  超过 3 年未访问的文件

- 保护 - 中等

  包含具有个人数据标识符（包括身份证号码、税务识别号、驾驶执照号码、药品 ID 或护照号码）的文件或数据库列的文件

- 敏感个人数据 - 高风险

  包含超过 20 个敏感个人数据标识符的文件或数据库列中超过 50% 的行包含敏感个人数据

更改存储库的NetApp Data Classification扫描设置

您可以管理在每个系统和数据源中如何扫描数据。您可以在“存储库”基础上进行更改；这意味着您可以根据正在扫描的数据源类型对每个卷、模式、用户等进行更改。

您可以更改的一些内容包括是否扫描存储库，以及NetApp Data Classification是否正在执行"映射扫描或映射和分类扫描"。您还可以暂停和恢复扫描，例如，如果您需要在一段时间内停止扫描某个卷。

查看存储库的扫描状态

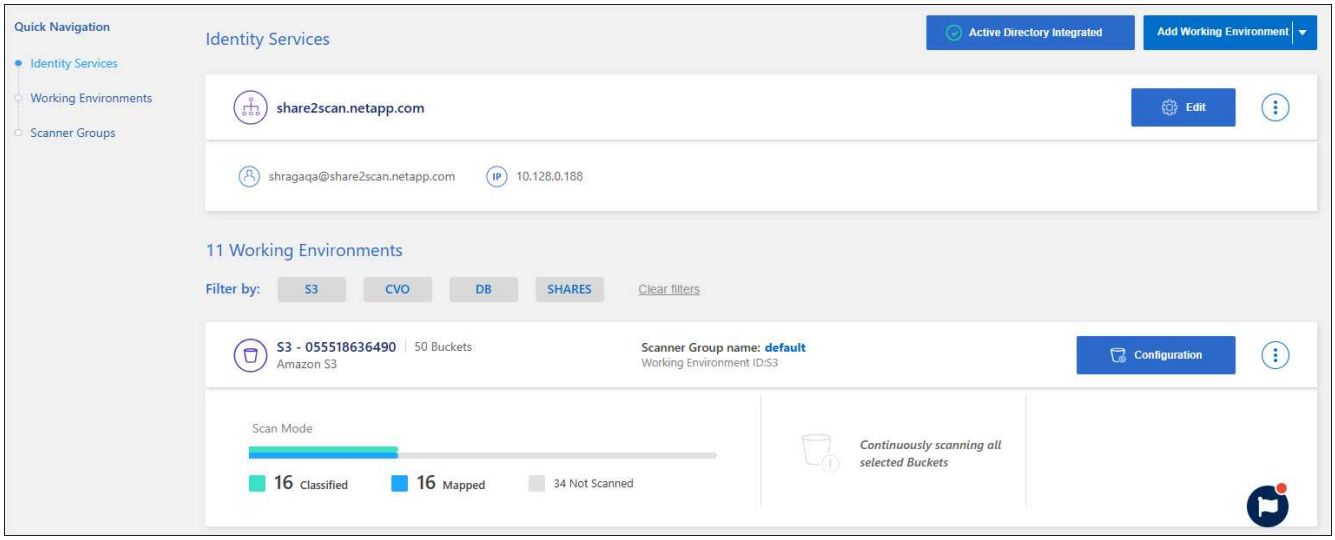
您可以查看NetApp Data Classification正在为每个系统和数据源扫描的各个存储库（卷、存储桶等）。您还可以看到有多少已被“映射”，有多少已被“分类”。分类需要更长的时间，因为所有数据都进行了完整的 AI 识别。

您可以在配置页面查看各个工作环境的扫描状态：

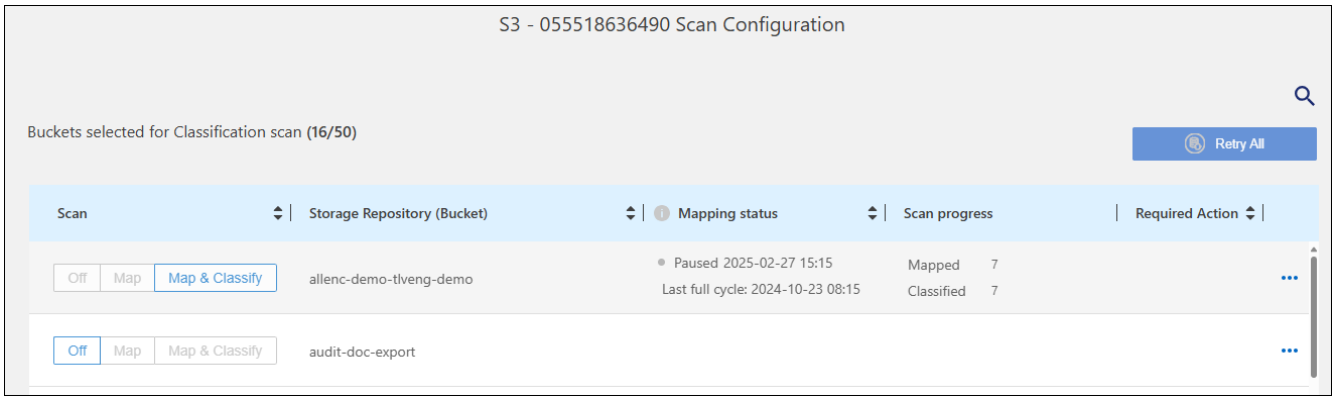
- 初始化（浅蓝色点）：地图或分类配置已激活。此状态会短暂显示，然后过渡到“待处理队列”状态。
- 待处理队列（橙色圆点）：扫描任务正在等待列入扫描队列。
- 已排队（橙色圆点）：任务已成功添加到扫描队列。当队列中的卷轮到达时，系统将开始映射或分类该卷。
- 正在运行（绿点）：队列中的扫描任务正在选定的存储库上积极进行。
- 完成（绿点）：存储库扫描已完成。
- 已暂停（灰点）：您已暂停扫描。虽然系统中未显示音量变化，但扫描结果仍然可用。
- 错误（红点）：扫描无法完成，因为遇到了问题。如果您需要完成某项操作，错误将出现在“所需操作”列下的工具提示中。否则，系统将显示“错误”状态并尝试恢复。完成后，状态就会改变。
- 未扫描：选择了“关闭”卷配置，系统未扫描该卷。

步骤

1. 从数据分类菜单中，选择*配置*。



2. 从配置选项卡中，选择系统的*配置*按钮。
3. 在扫描配置页面中，查看所有存储库的扫描设置。



4. 扫描期间，将光标悬停在“映射状态”列中的进度条上，即可查看该存储库中待映射或分类的文件数量。

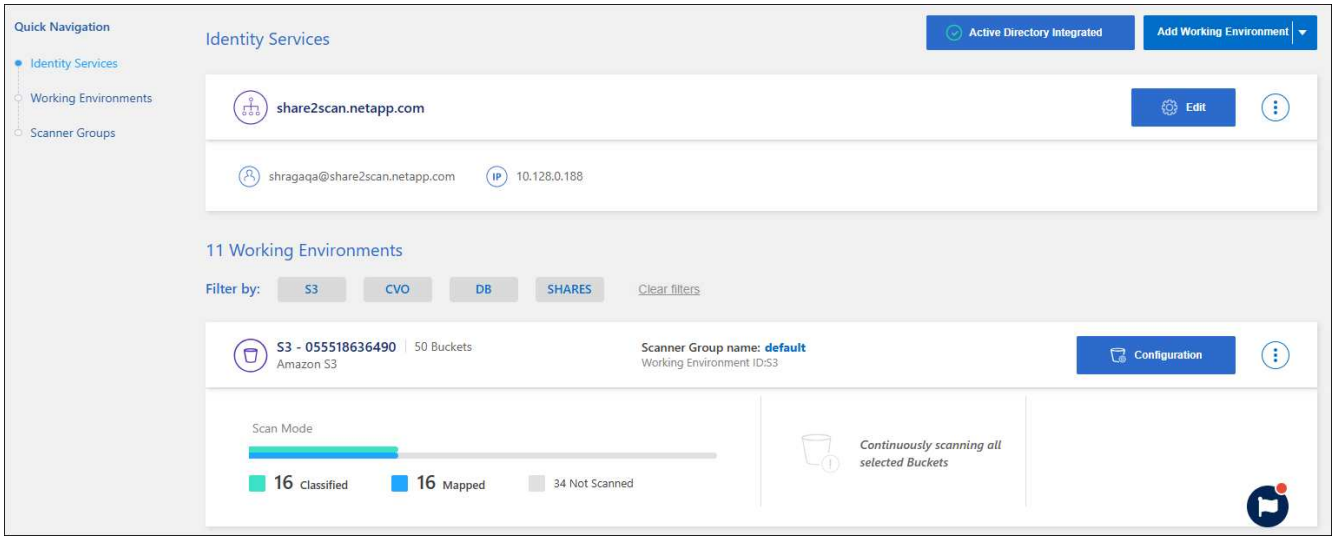
更改存储库的扫描类型

您可以随时从配置页面启动或停止系统中的仅映射扫描或映射和分类扫描。您还可以从仅映射扫描更改为映射和分类扫描，反之亦然。

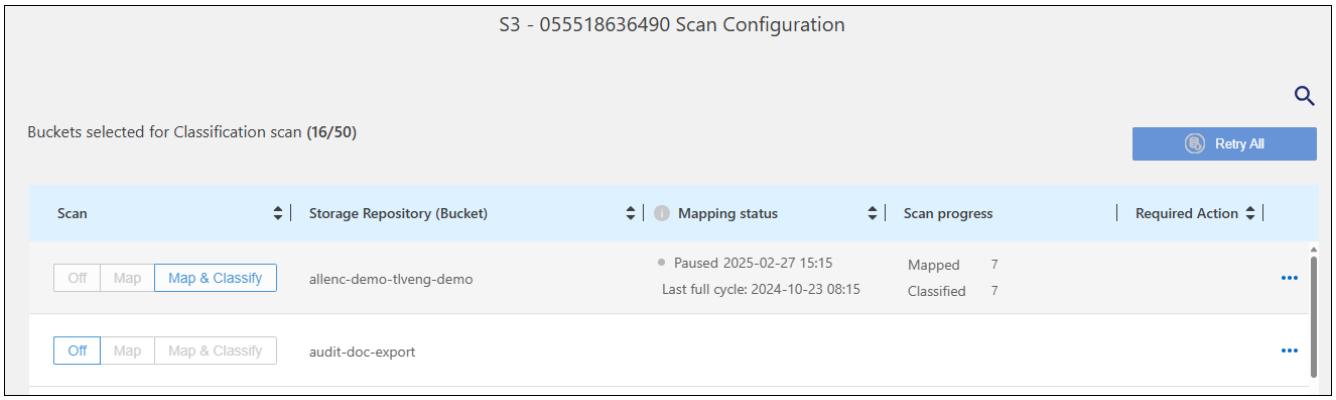
数据库不能设置为仅映射扫描。数据库扫描可以关闭或打开；其中“打开”相当于“映射和分类”。

步骤

1. 从数据分类菜单中，选择*配置*。
2. 从配置选项卡中，选择系统的*配置*按钮。

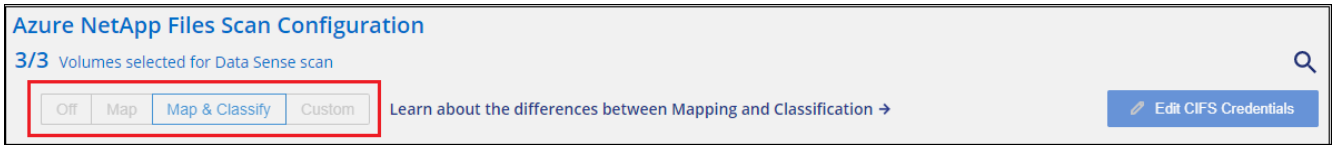


3. 在扫描配置页面中，更改任何存储库（本例中为存储桶）以执行*Map*或*Map & Classify*扫描。



某些类型的系统允许您使用页面顶部的按钮栏全局更改所有存储库的扫描类型。这对于Cloud Volumes ONTAP、本地ONTAP、 Azure NetApp Files和Amazon FSx for ONTAP系统有效。

下面的示例显示了Azure NetApp Files系统的按钮栏。



优先扫描

您可以优先考虑最重要的仅映射扫描或映射和分类扫描，以确保高优先级扫描首先完成。

默认情况下，扫描按照启动的顺序排队。通过设置扫描优先级，您可以将扫描移至队列的最前面。可以对多个扫描进行优先排序。优先级按先进先出的顺序指定，这意味着您优先考虑的第一个扫描将移至队列的最前面；您优先考虑的第二个扫描将成为队列中的第二个扫描，依此类推。

优先权是一次性授予的。映射数据的自动重新扫描按照默认顺序进行。

步骤

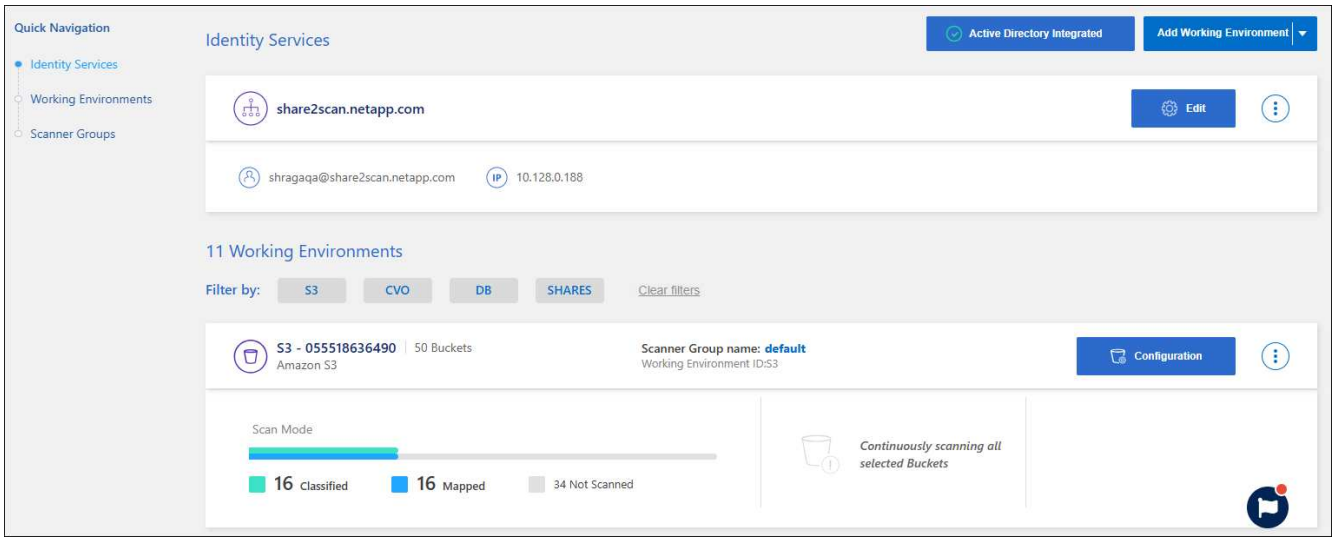
1. 从数据分类菜单中，选择*配置*。
2. 选择您想要优先考虑的资源。
3. 从行动 `…`选项，选择*优先扫描*。

停止扫描存储库

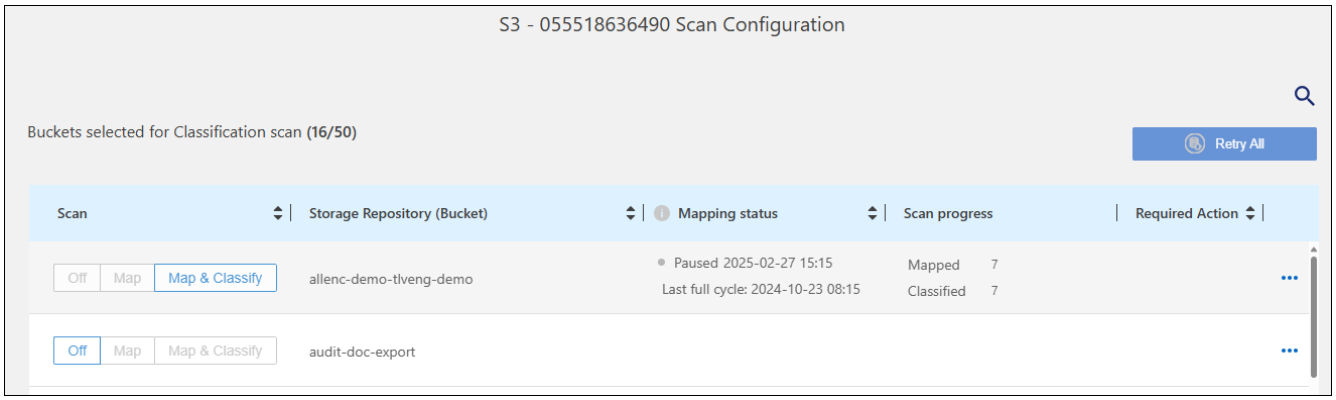
如果您不再需要监控存储库（例如卷）的合规性，则可以停止扫描它。您可以通过关闭扫描来实现此目的。当扫描关闭时，有关该卷的所有索引和信息都将从系统中删除，并且扫描数据的收费也将停止。

步骤

1. 从数据分类菜单中，选择*配置*。
2. 从配置选项卡中，选择系统的*配置*按钮。



3. 在扫描配置页面中选择“关闭”以停止扫描特定存储桶。



暂停并恢复存储库扫描

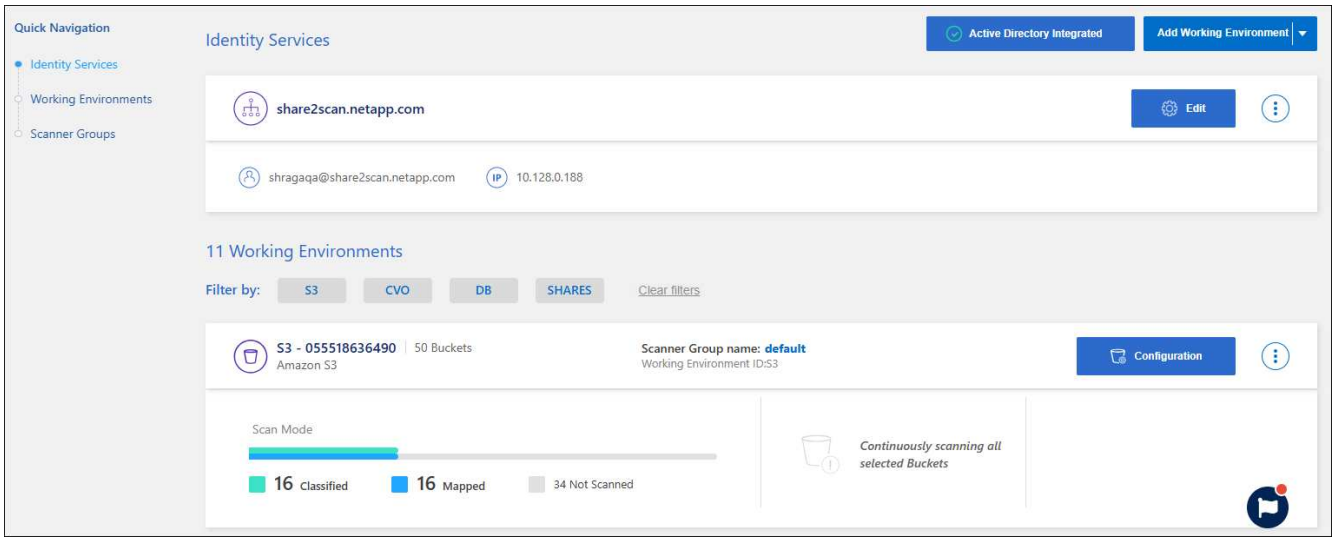
如果您想暂时停止扫描某些内容，您可以“暂停”存储库扫描。暂停扫描意味着数据分类将不再对存储库中的更改或添加执行任何未来的扫描。所有当前的扫描结果仍可在数据分类中查看。

即使暂停扫描，也不会免除计费费用，因为数据仍然保留在系统中。

您可以随时恢复扫描。

步骤

1. 从数据分类菜单中，选择*配置*。
2. 从配置选项卡中，选择系统的*配置*按钮。



3. 在扫描配置页面中，选择操作 ••• 图标。
4. 选择“暂停”暂停对卷的扫描，或选择“恢复”恢复对先前已暂停的卷的扫描。

查看NetApp Data Classification合规性报告

NetApp Data Classification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，数据分类仪表板显示所有系统、数据库和数据源的合规性和治理数据。如果您想要查看仅包含部分系统数据的报告，您可以进行筛选以仅查看这些系统的数据。

- 仅当您对数据源执行完整分类扫描时，才可获得合规性报告。已进行仅映射扫描的数据源只能生成数据映射报告。
- NetApp无法保证数据分类识别的个人数据和敏感个人数据 100% 的准确性。您应该始终通过查看数据来验证信息。

以下报告可用于数据分类：

- 数据发现评估报告：对扫描环境进行高级分析，以突出系统的发现并显示关注领域和潜在的补救步骤。此报告可在治理仪表板中找到。
- 完整数据映射概览报告：提供有关系统中文件的大小和数量的信息。这包括使用容量、数据年限、数据大小和文件类型。此报告可在治理仪表板中找到。
- 数据主体访问请求报告：使您能够提取包含有关数据主体的特定名称或个人标识符信息的所有文件的报告。此报告可在合规性仪表板中找到。
- **HIPAA** 报告：帮助您识别文件中健康信息的分布。此报告可在合规性仪表板中找到。

- **PCI DSS** 报告：帮助您识别文件中信用卡信息的分布。此报告可在合规性仪表板中找到。
- 隐私风险评估报告：提供来自您的数据的隐私见解和隐私风险评分。此报告可在合规性仪表板中找到。
- 特定信息类型的报告：可提供包含已识别文件（包含个人数据和敏感个人数据）详细信息的报告。您还可以查看按类别和文件类型细分的文件。

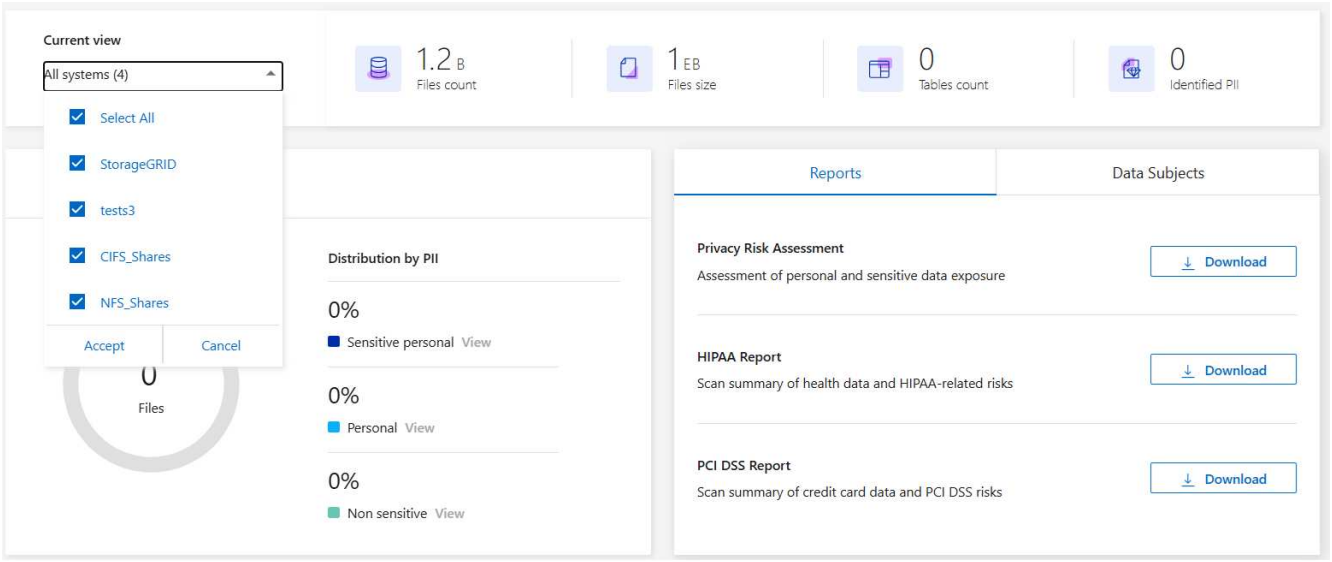
选择报告系统

您可以过滤数据分类合规性仪表板的内容，以查看所有系统和数据库的合规性数据，或仅查看特定系统的合规性数据。

当您过滤仪表板时，数据分类会将合规性数据和报告范围限定到您选择的系统。

步骤

1. 从数据分类菜单中，选择*合规性*。
2. 选择系统过滤器下拉菜单，然后选择系统。
3. 选择接受来确认您的选择。



数据主体访问请求报告

欧洲 GDPR 等隐私法规赋予数据主体（例如客户或员工）访问其个人数据的权利。当数据主体请求此信息时，这被称为 DSAR（数据主体访问请求）。各组织必须“毫不拖延”地回应这些请求，最迟不得超过收到请求后的一个月。

您可以通过搜索主题的全名或已知标识符（例如电子邮件地址）然后下载报告来回应 DSAR。该报告旨在帮助您的组织遵守 GDPR 或类似的数据隐私法。

数据分类如何帮助您响应 **DSAR**?

当您执行数据主体搜索时，数据分类会找到包含该人姓名或标识符的所有文件。数据分类检查最新的预索引数据的名称或标识符。它不会启动新的扫描。

搜索完成后，您可以下载数据主体访问请求报告的文件列表。该报告汇总了数据中的见解，并将其转化为法律术

语，以便您可以将其发送给相关人员。

目前数据库不支持数据主体搜索。

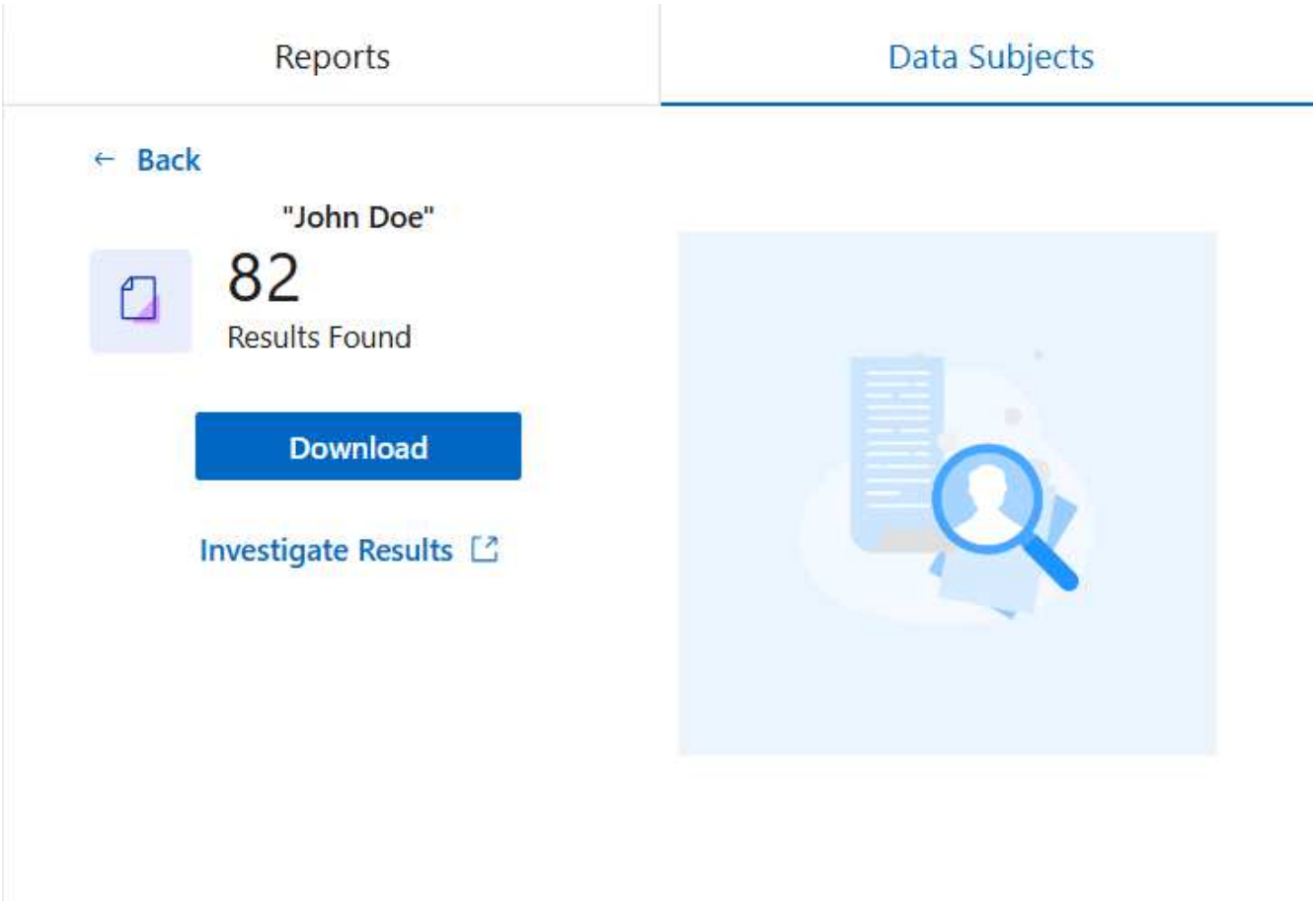
搜索数据主体并下载报告

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。您可以通过以下方式搜索"任何个人信息类型"。

搜索数据主体的姓名时支持英语、德语、日语和西班牙语。稍后将添加对更多语言的支持。

步骤

1. 从数据分类菜单中，选择*合规性*。
2. 在合规性页面中，找到数据主体选项卡。
3. 在“数据主体”部分，输入名称或已知标识符，然后选择“搜索”。
4. 搜索完成后，选择下载以访问数据主体访问请求响应。选择调查结果以在数据调查页面中查看更多信息。



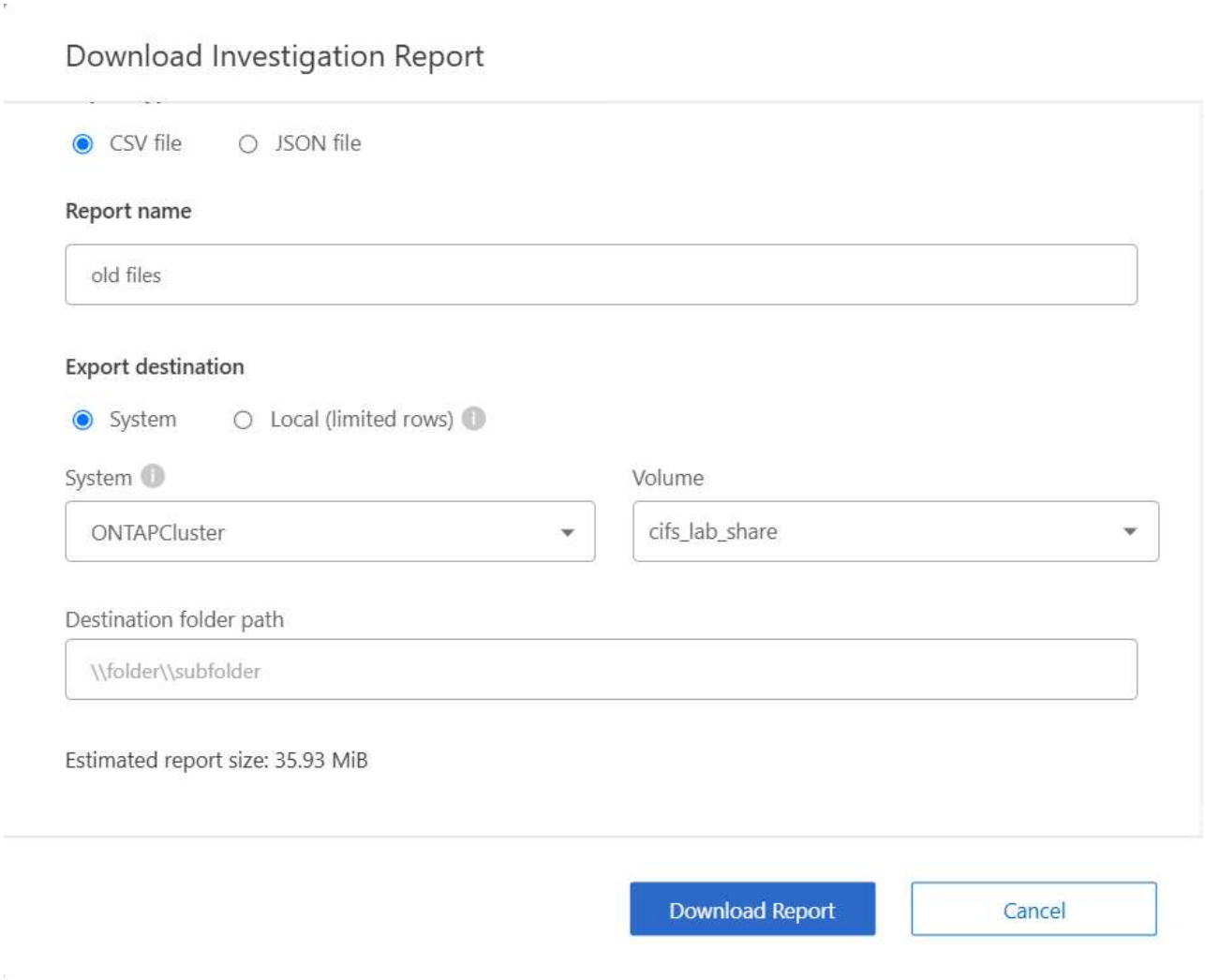
5. 查看数据分类中的结果或通过选择下载图标将其下载为报告。
   a. 选择下载图标后，配置您的下载设置：
      - 选择影片格式：CSV 或 JSON
      - 输入*报告名称*

- 选择导出目的地：*系统*或您的*本地*机器。

  如果您选择系统，则会下载所有数据。您还必须选择*系统*、卷*和*目标文件夹路径。

如果您选择*本地*，则会将报告限制为前 10,000 行非结构化数据；5,000 行非结构化数据和 1,000 行结构化数据。

a. 选择下载报告开始下载。



健康保险流通与责任法案（HIPAA）报告

健康保险流通与责任法案 (HIPAA) 报告可以帮助您识别包含健康信息的文件。它旨在帮助您的组织遵守 HIPAA 数据隐私法的要求。数据分类寻找的信息包括：

- 健康参考模式
- ICD-10-CM 医疗代码
- ICD-9-CM 医疗代码
- HR - 健康类别
- 健康应用数据类别

该报告包含以下信息：

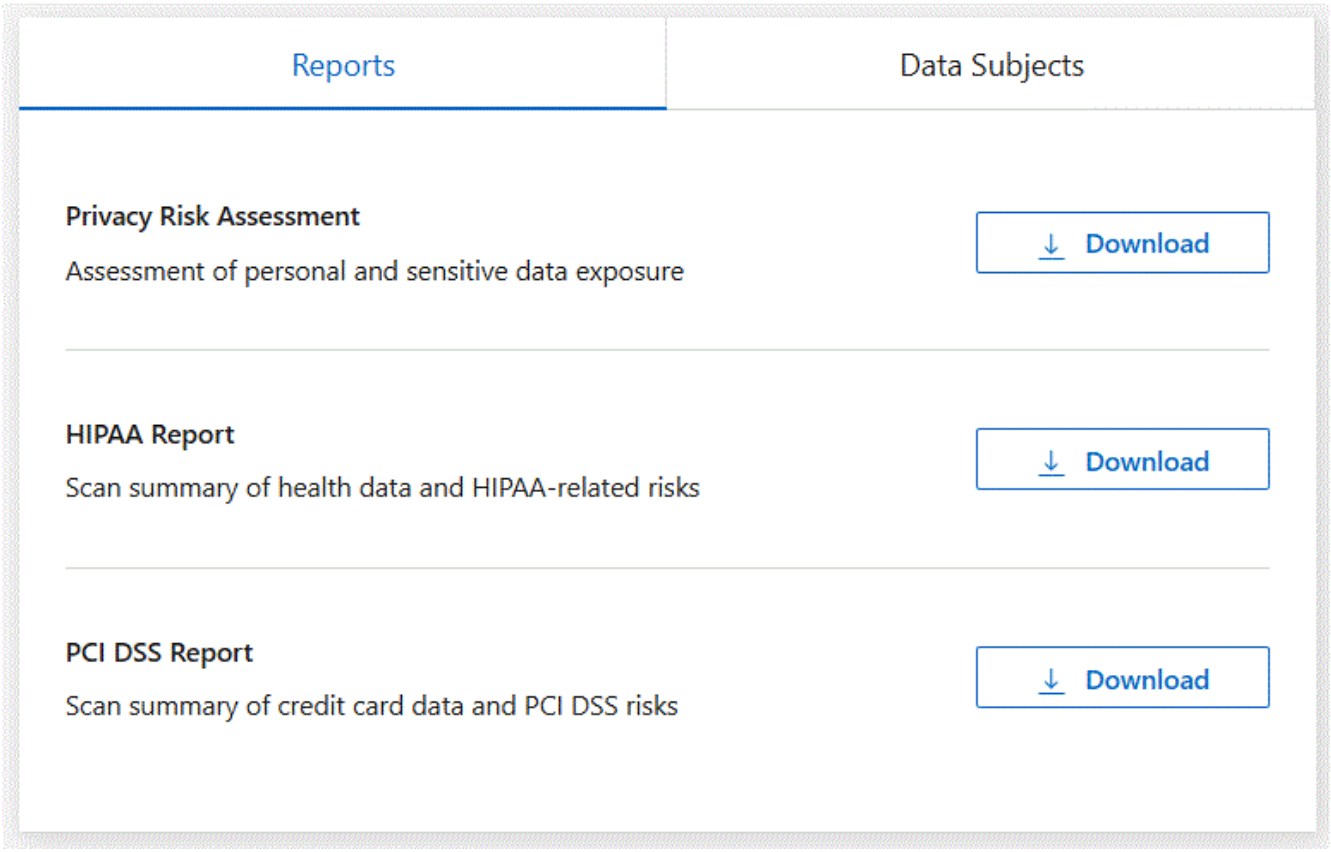
- 概述：有多少文件包含健康信息以及在哪些系统中。
- 加密：加密或未加密系统中包含健康信息的文件的百分比。此信息特定于Cloud Volumes ONTAP。
- 勒索软件防护：在启用或未启用勒索软件防护的系统上，包含健康信息的文件的百分比。此信息特定于Cloud Volumes ONTAP。
- 保留：文件最后修改的时间范围。这很有用，因为您不应该将健康信息保存超过处理所需的时间。
- 健康信息分发：发现健康信息的系统以及是否启用了加密和勒索软件保护。

生成 **HIPAA** 报告

转到“合规性”选项卡以生成报告。

步骤

1. 从数据分类菜单中，选择*合规性*。
2. 找到报告窗格。选择*HIPAA 报告*旁边的下载图标。



结果

数据分类生成 PDF 报告。

支付卡行业数据安全标准 (PCI DSS) 报告

支付卡行业数据安全标准 (PCI DSS) 报告可以帮助您识别信用卡信息在文件中的分布。

该报告包含以下信息：

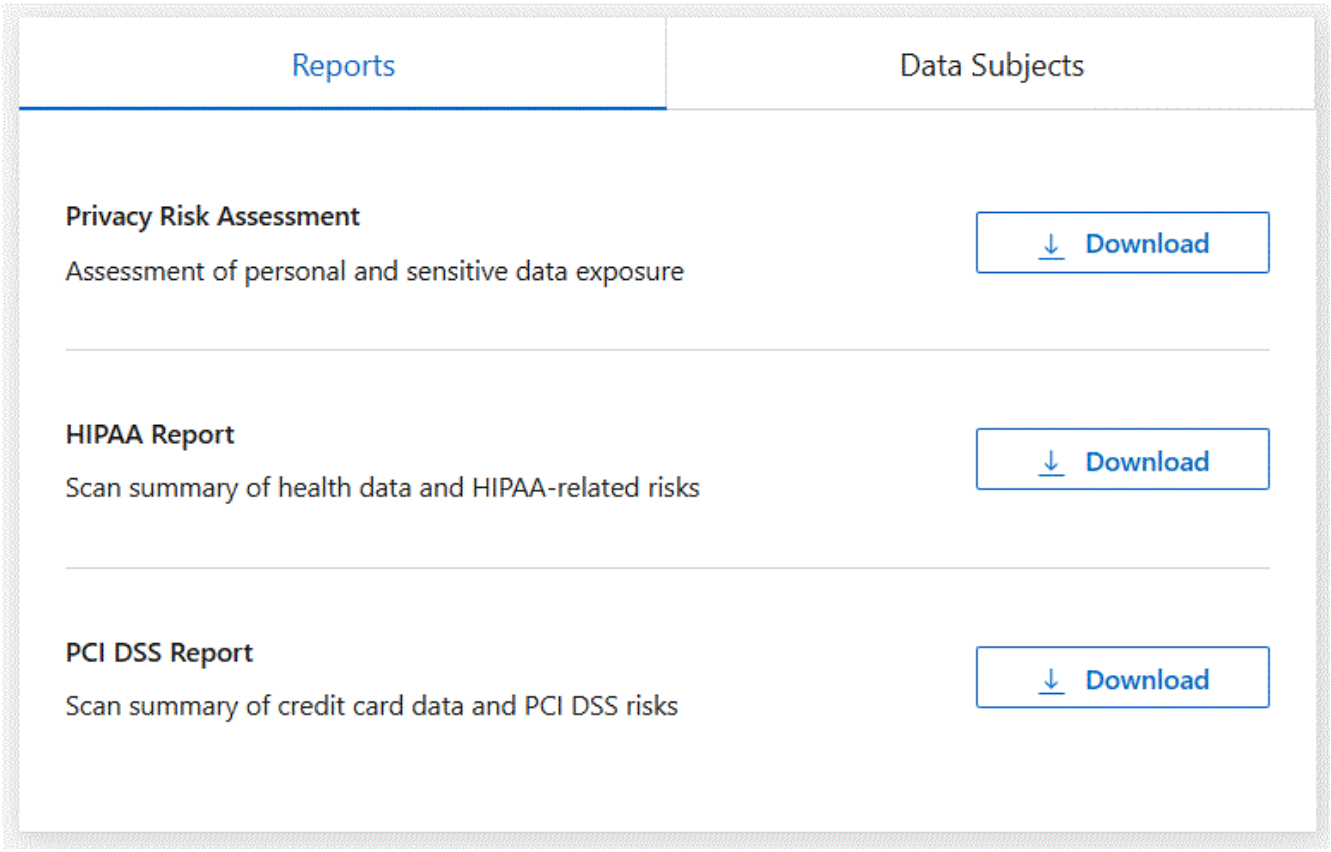
- 概述：有多少个文件包含信用卡信息以及在哪些系统中。
- 加密：加密或未加密系统中包含信用卡信息的文件的百分比。此信息特定于Cloud Volumes ONTAP。
- 勒索软件防护：在启用或未启用勒索软件防护的系统上，包含信用卡信息的文件的百分比。此信息特定于Cloud Volumes ONTAP。
- 保留：文件最后修改的时间范围。这很有用，因为您不应该将信用卡信息保存的时间超过处理所需的时间。
- 信用卡信息分发：发现信用卡信息的系统以及是否启用了加密和勒索软件保护。

生成 **PCI DSS** 报告

转到“合规性”选项卡以生成报告。

步骤

1. 从数据分类菜单中，选择*合规性*。
2. 找到报告窗格。选择*PCI DSS 报告*旁边的下载图标。



结果

数据分类会生成一份 PDF 报告，您可以根据需要查看并发送给其他组。

隐私风险评估报告

隐私风险评估报告概述了您组织的隐私风险状况，这是 GDPR 和 CCPA 等隐私法规所要求的。

该报告包含以下信息：

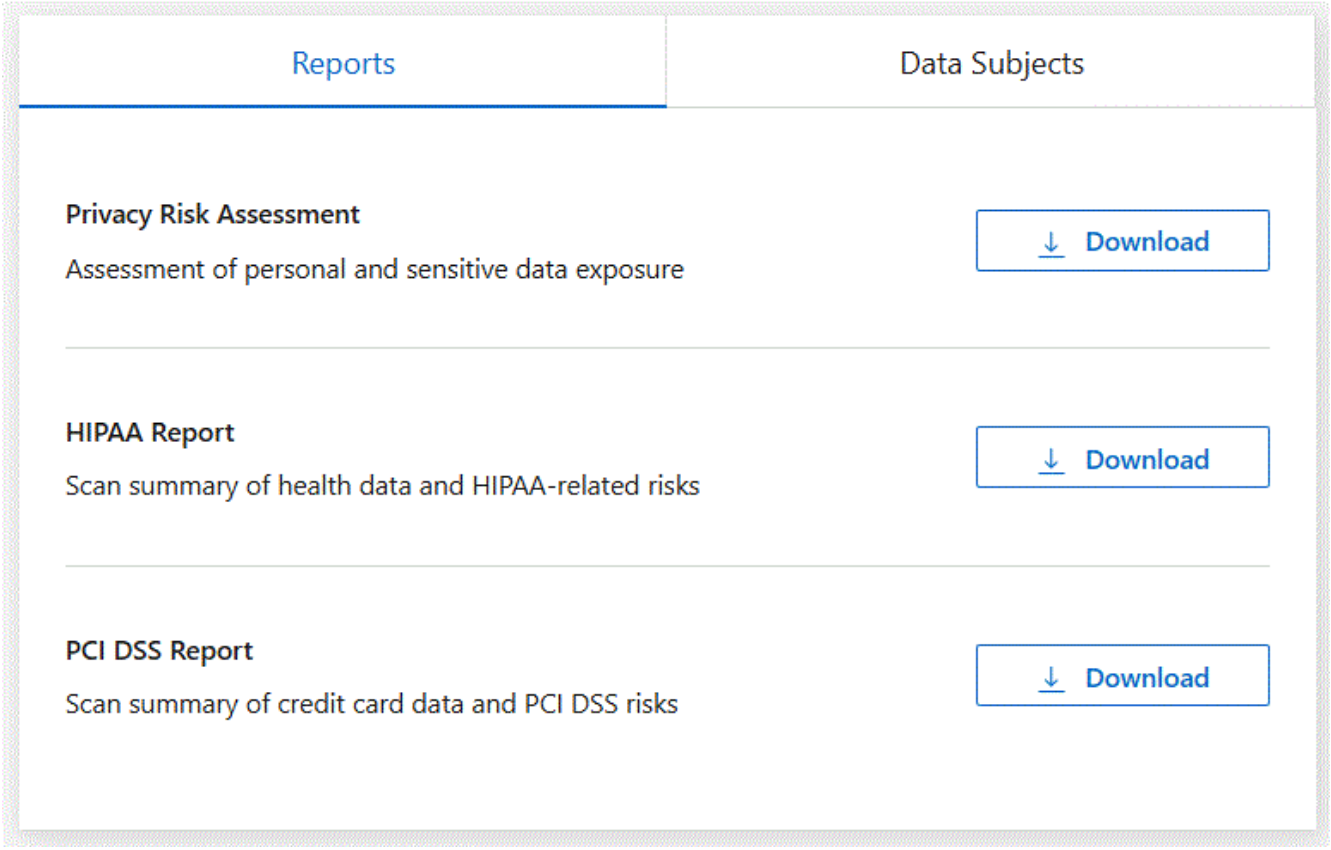
- 合规状态：严重性评分和数据分布，无论是非敏感数据、个人数据还是敏感个人数据。
- 评估概述：发现的个人数据类型以及数据类别的细分。
- 本次评估中的数据主体：按地点划分的已找到国家标识符的人数。

生成隐私风险评估报告

转到“合规性”选项卡以生成报告。

步骤

1. 从数据分类菜单中，选择*合规性*。
2. 找到报告窗格。选择*隐私风险评估报告*旁边的下载图标。



结果

数据分类会生成一份 PDF 报告，您可以根据需要查看并发送给其他组。

严重程度评分

数据分类根据三个变量计算隐私风险评估报告的严重性分数：

- 个人数据占所有数据的百分比。
- 敏感个人数据占所有数据的比例。
- 包含数据主体的文件百分比，由国家标识符（例如国民身份证、社会安全号码和税号）决定。

确定分数的逻辑如下：

严重程度评分	逻辑
0	所有三个变量都恰好为 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%

监控NetApp Data Classification的运行状况

NetApp Data Classification健康监视器仪表板提供实时监控和性能洞察。健康监视器会捕获有关您的数据分类基础架构、系统运行状况、使用指标和利用率数据的信息，使您能够识别和解决问题。

健康监测洞察

健康监测仪表盘以四类信息呈现信息。

- 基础设施状况

  查看版本状态、系统稳定性、部署类型和机器规模等信息。

- 问题容器

  查看“问题容器”字段，以了解哪些容器已停止或频繁重启。利用这些信息调查具体的容器。

- 系统信息

  系统信息面板捕获有关NetApp Console和数据分类的关键信息，例如公共和私有 IP 地址、主机名、操作系统、控制台版本和控制台 ID。

- 用途和使用方法

  查看 CPU 使用率、磁盘利用率、磁盘使用率和内存使用率。这些值以存储单位（GB）或总使用量的百分比显示。如果任何字段显示警告，请选择该警告以获取相关信息和补救建议。

访问健康监测仪表板

1. 在数据分类中，选择配置。
2. 在“配置”标题下，选择“数据分类运行状况监视器”。
3. 在健康监测仪表板中，您可以：
   - 审查使用情况和利用情况。如果任何使用情况或利用率指标显示警告，请选择该警告以获取解决问题的建议。
   - 切换图表以显示 CPU 使用率、磁盘利用率、磁盘使用率和内存使用率。您可以更改 x 轴，以按小时（6、12 或 24 小时）或天（2、7 或 14 天）显示内容。
   - 刷新仪表盘以查看最新数据指标。