



Use BlueXP classification

BlueXP classification

NetApp
July 24, 2025

Table of Contents

Use BlueXP classification	1
View governance details about the data stored in your organization with BlueXP classification	1
Review the Governance dashboard	1
Create the Data Discovery Assessment Report	4
Create the Data Mapping Overview Report	4
View compliance details about the private data stored in your organization with BlueXP classification	7
View files that contain personal data	8
View files that contain sensitive personal data	10
View files by categories	12
View files by file types	12
Categories of private data in BlueXP classification	13
Types of personal data	13
Types of sensitive personal data	18
Types of categories	19
Types of files	20
Accuracy of information found	20
Create a custom classification in BlueXP classification	21
Create a custom classification	21
Investigate the data stored in your organization with BlueXP classification	23
Filter data in the Data Investigation page	23
View file metadata	26
View users' permissions for files and directories	27
Check for duplicate files in your storage systems	28
Create the Data Investigation Report	29
Create a saved search based on selected filters	32
Manage saved searches with BlueXP classification	33
View saved searches results in the Investigation page	33
Create custom saved searches	33
Edit saved searches	35
Delete saved searches	37
Default searches	37
Change the BlueXP classification scan settings for your repositories	37
View the scan status for your repositories	38
Change the type of scanning for a repository	39
Prioritize scans	40
Stop scanning for a repository	40
Pause and resume scanning for a repository	41
View BlueXP classification compliance reports	42
Select the working environments for reports	43
Data Subject Access Request Report	43
Health Insurance Portability and Accountability Act (HIPAA) Report	45
Payment Card Industry Data Security Standard (PCI DSS) Report	46
Privacy Risk Assessment Report	47

Use BlueXP classification

View governance details about the data stored in your organization with BlueXP classification

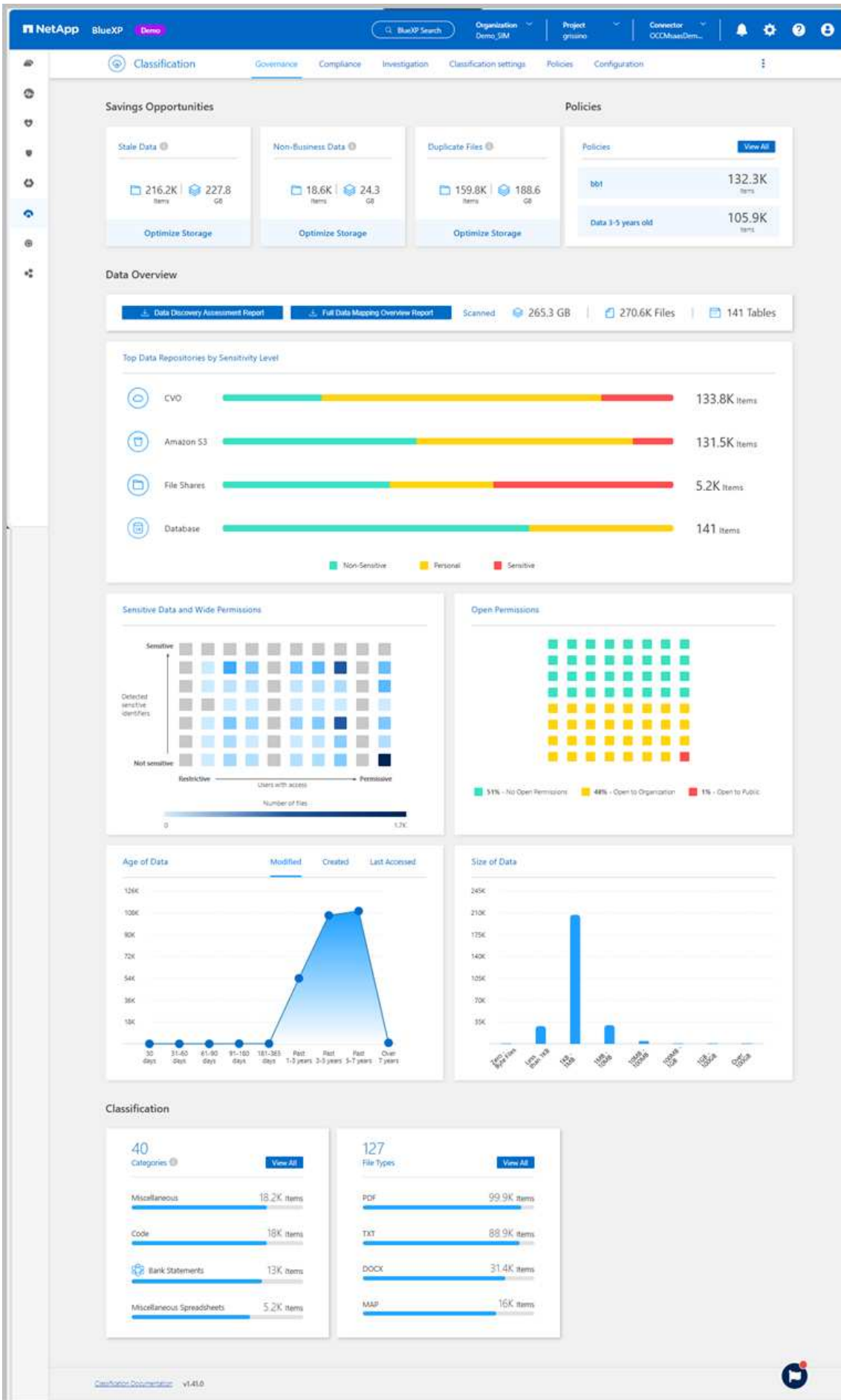
Gain control of the costs related to the data on your organization's storage resources. BlueXP classification identifies the amount of stale data, non-business data, duplicate files, and very large files in your systems so you can decide whether you want to remove or tier some files to less expensive object storage.

This is where you should begin your research. From the Governance dashboard, you can select an area for further investigation.

Additionally, if you're planning to migrate data from on-premises locations to the cloud, you can view the size of the data and whether any of the data contains sensitive information before moving it.

Review the Governance dashboard

The Governance dashboard provides information so that you can increase the efficiency and control the costs related to the data stored on your storage resources.



Steps

1. From the BlueXP menu, select **Governance > Classification**.
2. Select **Governance**.

The Governance dashboard appears.

Review savings opportunities

The *Saving Opportunities* component shows data that you can delete or tier to less expensive object storage. The data in *Saving Opportunities* update every 2 hours and can be manually updated.

Steps

1. From the BlueXP classification menu, select **Governance**.
2. Within each Savings Opportunities tile of the Governance dashboard, select **Optimize Storage** to view the filtered results in the Investigation page. To discover any data you should delete or tier to less expensive storage, investigate the the *Saving Opportunities*.
 - **Stale Data** - Data that was last modified over 3 years ago.
 - **Non-Business Data** - Data considered not to be business related, based on their Category or File Type. This includes:
 - Application Data
 - Audio
 - Executables
 - Images
 - Logs
 - Videos
 - Miscellaneous (general "other" category)
 - **Duplicate Files** - Files that are duplicated in other locations in the data sources you are scanning. [See what types of duplicate files are displayed.](#)



If any of your data sources implement data tiering, old data that already resides in object storage can be identified in the *Stale Data* category.

Review saved searches with the largest number of results

In the *Saved searches* tab, the searches with the greatest number of results appear at the top of the list. This data updates every two hours.

For details about saved searches, see [Create saved searches](#).

Steps

1. From the BlueXP classification menu, select **Governance**.
2. In the Governance dashboard, locate the Saved Searches tile. Select the name of a saved search to display the results in the Investigation page.
3. Select **View All** to view the list of all available saved searches.

In the *Saved searches* area, the searches with the greatest number of results appear at the top of the list.

Create the Data Discovery Assessment Report

The Data Discovery Assessment Report provides a high-level analysis of the scanned environment to show areas of concern and potential remediation steps. The results are based on both mapping and classifying your data. The goal of this report is to raise awareness of three significant aspects of your dataset:

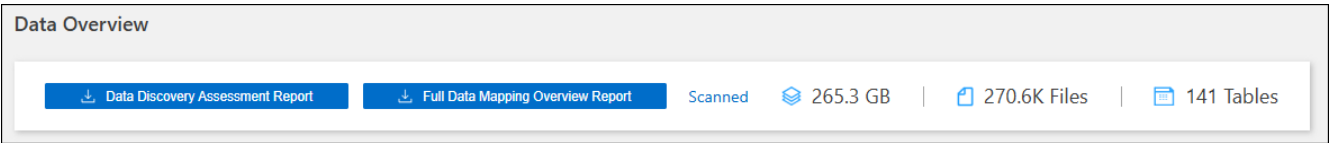
Feature	Description
Data governance concerns	A detailed picture of all the data you own and areas where you may be able to reduce the amount of data to save costs.
Data security exposures	Areas where your data is accessible to internal or external attacks because of broad access permissions.
Data compliance gaps	Where your personal or sensitive personal information is located for both security and for DSARs (data subject access requests).

Using this report, you might take the following actions:

- Reduce storage costs by changing your retention policy, or by moving or deleting certain data (stale, duplicate, or non-business data)
- Protect your data that has broad permissions by revising global group management policies
- Protect your data that has personal or sensitive personal information by moving PII to more secure data stores

Steps

1. From the BlueXP menu, select **Governance > Classification**.
2. Select **Governance**.
3. Select **Data Discovery Assessment Report**.



Result

BlueXP classification generates a PDF report that you can review and send to other groups as needed.

Create the Data Mapping Overview Report

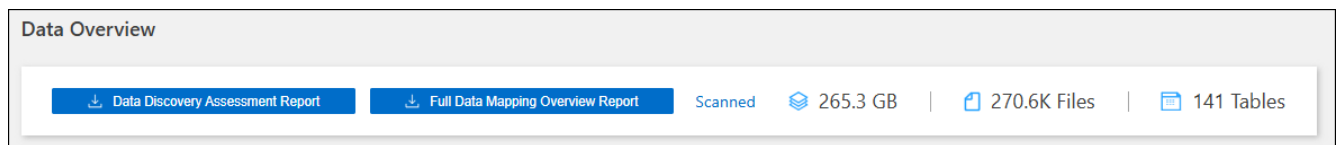
The Data Mapping Overview Report provides an overview of the data being stored in your corporate data sources to assist you with decisions of migration, back up, security, and compliance processes. The report summarizes all working environments and data sources. It also provides an analysis for each working environment.


The report includes the following information:

Category	Description
Usage Capacity	For all working environments: Lists the number of files and the used capacity for each working environment. For single working environments: Lists the files that are using the most capacity.
Age of Data	Provides three charts and graphs for when files were created, last modified, or last accessed. Lists the number of files, and their used capacity, based on certain date ranges.
Size of Data	Lists the number of files that exist within certain size ranges in your working environments.
File Types	Lists the total number of files and the used capacity for each type of file being stored in your working environments.

Steps

1. From the BlueXP menu, select **Governance > Classification**.
2. Select **Governance**.
3. Select **Full Data Mapping Overview Report**.



4. To customize the company name that appears on the first page of the report, from the top right of the BlueXP classification page, select . Then select **Change company name**. The next time you generate the report, it will include the new name.

Result

BlueXP classification generates a .pdf report that you can review and send to other groups as needed.

If the report is larger than 1 MB, the .pdf file is retained on the BlueXP classification instance and you'll see a pop-up message about the exact location. When BlueXP classification is installed on a Linux machine on your premises, or on a Linux machine you deployed in the cloud, you can navigate directly to the .pdf file. When BlueXP classification is deployed in the cloud, you'll need to SSH to the BlueXP classification instance to download the .pdf file.

Review the top data repositories listed by data sensitivity

The *Top Data Repositories by Sensitivity Level* area of the Data Mapping Overview report lists the top four data repositories (working environments and data sources) that contain the most sensitive items. The bar chart for each working environment is divided into:

- Non-Sensitive data
- Personal data
- Sensitive Personal data

This data refreshes every two hours and can be manually refreshed.

Steps

1. To see the total number of items in each category, position your cursor over each section of the bar.
2. To filter results that will appear in the Investigation page, select each area in the bar and investigate further.

Review sensitive data and wide permissions

The *Sensitive Data and Wide Permissions* area of the Data Mapping Overview report shows the percentage of files that contain sensitive data and have wide permissions. The chart shows the following types of permissions:

- From the most restrictive permissions to the most permissive restrictions on the horizontal axis.
- From the least sensitive data to the most sensitive data on the vertical axis.

Steps

1. To see the total number of files in each category, position your cursor over each box.
2. To filter results that will appear in the Investigation page, select a box and investigate further.

Review data listed by types of open permissions

The *Open Permissions* area of the Data Mapping Overview report shows the percentage for each type of permissions that exist for all files that are being scanned. The chart shows the following types of permissions:

- No Open Permissions
- Open to Organization
- Open to Public
- Unknown Access

Steps

1. To see the total number of files in each category, position your cursor over each box.
2. To filter results that will appear in the Investigation page, select a box and investigate further.

Review the age and size of data

You might want to investigate the items in the *Age* and *Size* graphs of the Data Mapping Overview report to see if there is any data you should delete or tier to less expensive object storage.

Steps

1. In the Age of Data chart, to see details about the age of the data, position your cursor over a point in the chart.
2. To filter by an age or size range, select that age or size.
 - **Age of Data graph** - Categorizes data based on the time it was created, the last time it was accessed, or the last time it was modified.
 - **Size of Data graph** - Categorizes data based on size.



If any of your data sources implement data tiering, old data that already resides in object storage might be identified in the *Age of Data* graph.

Review the most identified data classifications in your data

The *Classification* area of the Data Mapping Overview report provides a list of the most identified [Categories](#) and [File types](#) in your scanned data.

Categories can help you understand what's happening with your data by showing you the types of information that you have. For example, a category like "resumes" or "employee contracts" can include sensitive data. When you investigate the results, you might find that employee contracts are stored in a nonsecure location. You can then correct that issue.

See [Viewing files by categories](#) for more information.

Steps

1. From the BlueXP menu, select **Governance > Classification**.
2. Select **Governance** then the **Data Discovery Assessment Report** button.

Result

BlueXP classification generates a .pdf report that you can review and send to other groups as needed.

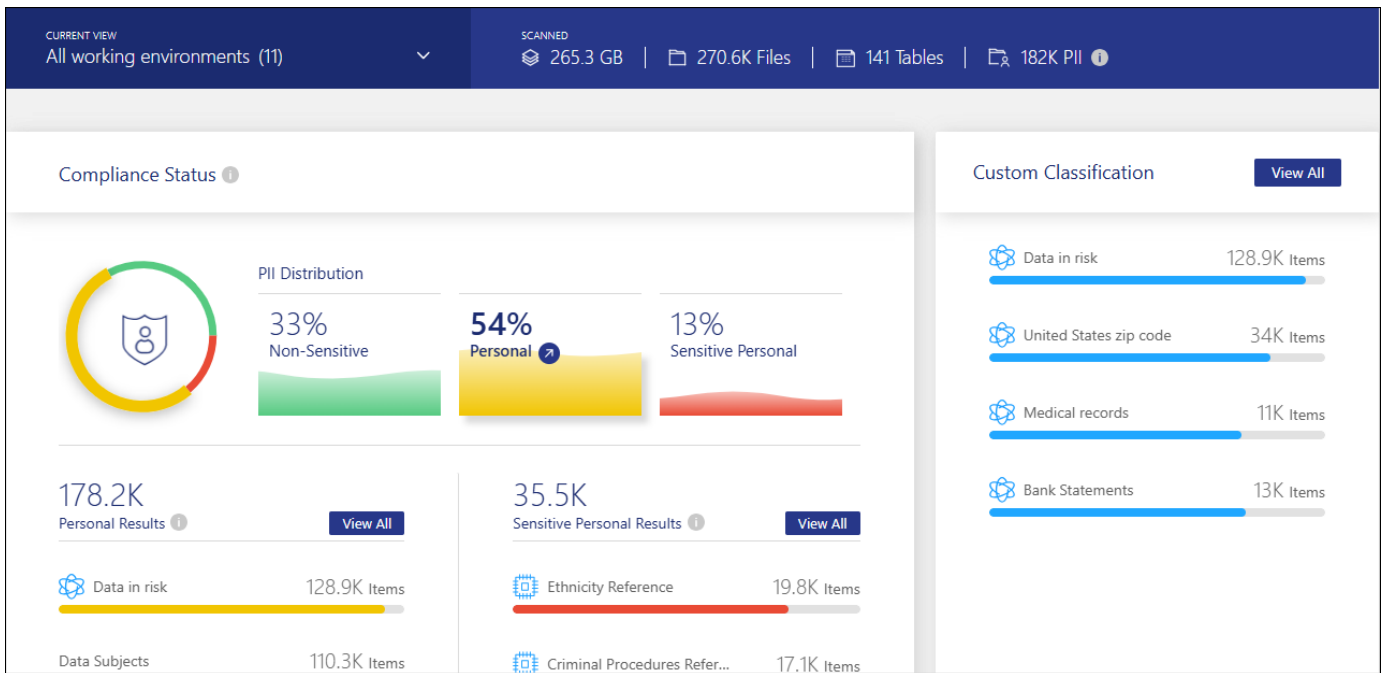
View compliance details about the private data stored in your organization with BlueXP classification

Gain control of your private data by viewing details about the personal data (PII) and sensitive personal data (SPII) in your organization. You can also gain visibility by reviewing the categories and file types that BlueXP classification found in your data.



The capabilities described in this section are available only if you have chosen to perform a full classification scan on your data sources. Data sources that have had a mapping-only scan do not show file-level details.

By default, the BlueXP classification dashboard displays compliance data for all working environments and databases. To see data for only some of the working environments, select them.



Filter the results from the Data Investigation page and download a report of the results as a CSV file. See [Filtering data in the Data Investigation page](#) for details.

View files that contain personal data

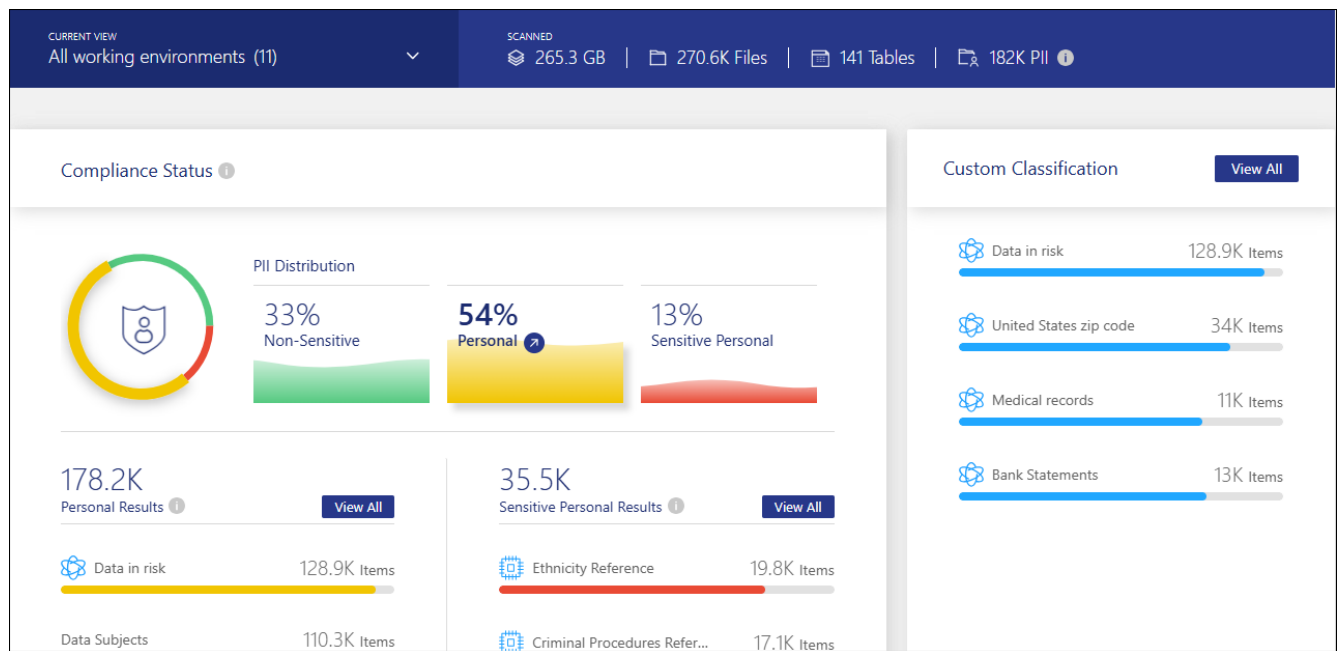
BlueXP classification automatically identifies specific words, strings, and patterns (Regex) inside the data. For example, Personal Identification Information (PII), credit card numbers, social security numbers, bank account numbers, passwords, and more. [See the full list](#). BlueXP classification identifies this type of information in individual files, in files within directories (shares and folders), and in database tables.

You can also create custom search terms to identify personal data specific to your organization. For more information, see [Create a custom classification](#).

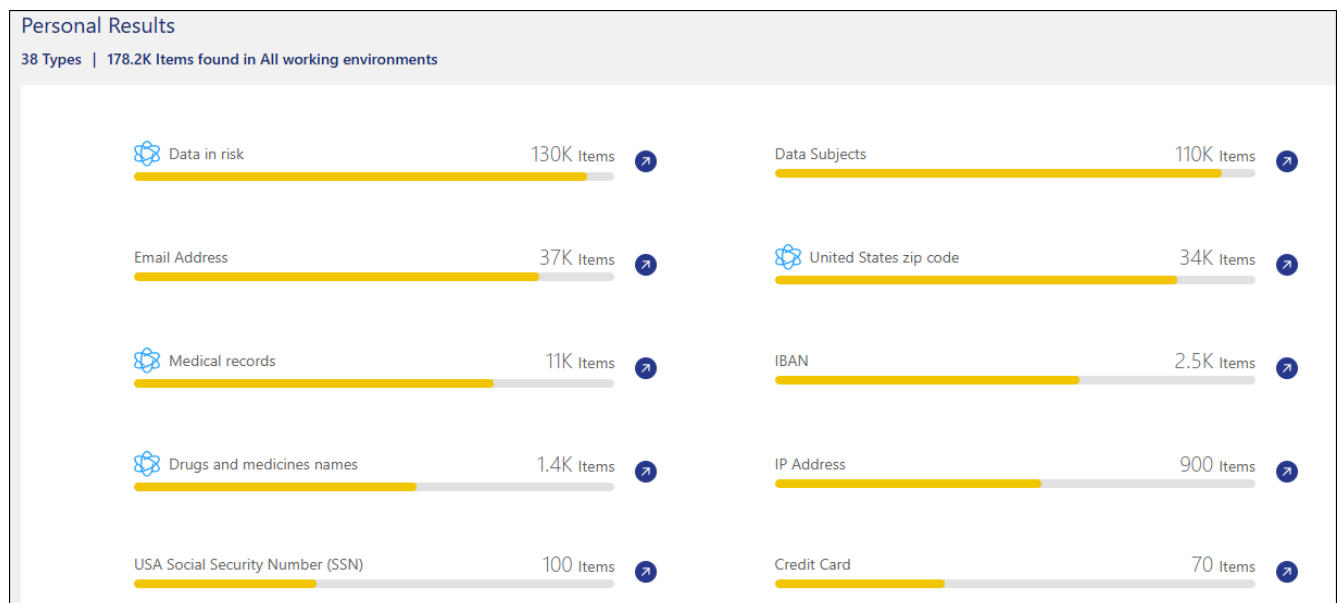
For some types of personal data, BlueXP classification uses *proximity validation* to validate its findings. The validation occurs by looking for one or more predefined keywords in proximity to the personal data that was found. For example, BlueXP classification identifies a U.S. social security number (SSN) as an SSN if it sees a proximity word next to it—for example, *SSN* or *social security*. [The table of personal data](#) shows when BlueXP classification uses proximity validation.

Steps

1. From the BlueXP classification menu, select the **Compliance** tab.
2. To investigate the details for all personal data, select the icon next to the personal data percentage.



- To investigate the details for a specific type of personal data, select **View All** and then select the **Investigate Results** arrow icon for a specific type of personal data, for example, email addresses.



- Investigate the data by searching, sorting, expanding details for a specific file, selecting the **Investigate Results** arrow to see masked information, or by downloading the file list.

The two screenshots below show personal data found in individual files, and found in files within directories (shares and folders). You can also select the **Structured** tab to view personal data found in databases.

The screenshot shows the 'Data Investigation' interface with the 'Unstructured (36.6K Files)' tab selected. On the left, there is a 'FILTERS' sidebar with expandable sections for Policies, Classification Status, Scan Analysis Event, Open Permissions, Number of Users with Access, and User / Group Permissions. Below the filters are buttons for 'Create Policy from this search' and 'Set Email Alert'. The main area displays details for a file named 'B81ALrKD.txt'. The file is 1.2K in size, has a file type of 'TXT', and is classified as 'Sensitive Personal'. It has 10 tags, including 'archivado', 'credit card', and 'Delete'. The file's metadata includes: 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), and Duplicates (None). Action buttons for 'Copy File', 'Move File', and 'Delete File' are available. The bottom status bar shows 'Total size 26.5GB | 1-20 of 36.6K'.

The screenshot shows the 'Data Investigation' interface with the 'Directories (6.1K Folders)' tab selected. The left sidebar and filter options are identical to the first screenshot. The main area displays details for a directory named '/vol_cifs_share/HR_sensitive_data/copy_100/contextual_data'. The directory is of type 'Folder' and is located in the 'vol_cifs' storage repository. Its metadata includes: Working Environment (ONTAPCluster), Storage Repository (Volume: vol_cifs), Directory Path, Discovered Time (2023-11-07 10:28), Created Time (2023-08-20 13:54), Last Modified (2023-08-20 13:54), Last Accessed (2024-10-09 19:04), Open Permissions (OPEN TO ORGANIZATION), and Directory Owner (sl). A 'Rescan' button is visible. The bottom status bar shows '1-20 of 6.1K'.

View files that contain sensitive personal data

BlueXP classification automatically identifies special types of sensitive personal information, as defined by privacy regulations such as [articles 9 and 10 of the GDPR](#). For example, information regarding a person's health, ethnic origin, or sexual orientation. [See the full list](#). BlueXP classification identifies this type of information in individual files, in files within directories (shares and folders), and in database tables.

BlueXP classification uses artificial intelligence (AI), natural language processing (NLP), machine learning (ML), and cognitive computing (CC) to understand the meaning of the content that it scans in order to extract entities and categorize it accordingly.

For example, one sensitive GDPR data category is ethnic origin. Because of its NLP abilities, BlueXP classification can distinguish the difference between a sentence that reads "George is Mexican" (indicating

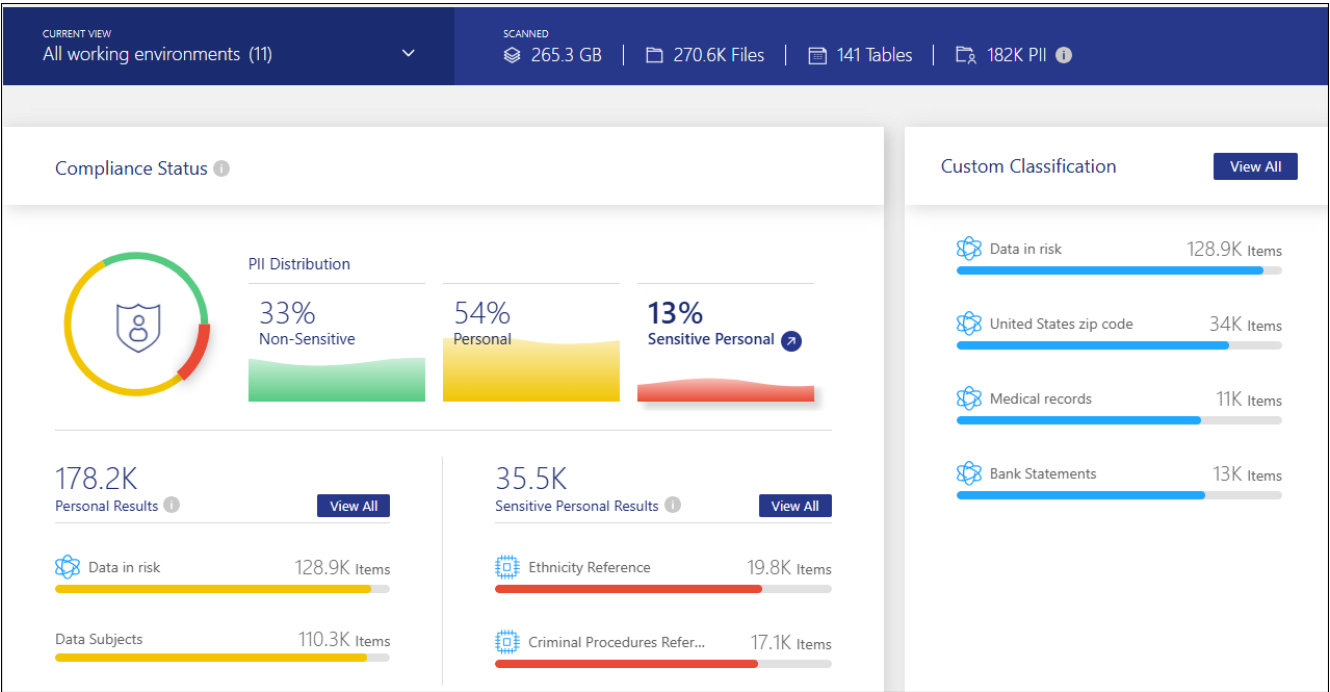
sensitive data as specified in article 9 of the GDPR), versus "George is eating Mexican food."



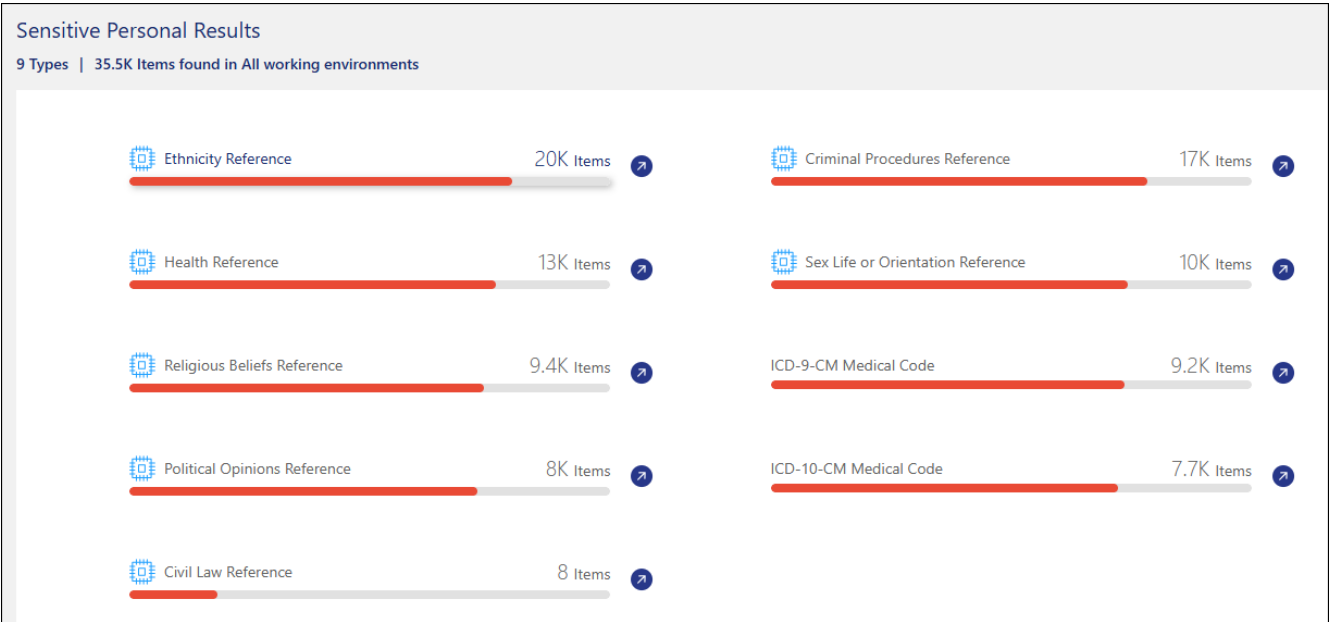
Only English is supported when scanning for sensitive personal data. Support for more languages will be added later.

Steps

- 1. From the BlueXP classification menu, select **Compliance**.
- 2. To investigate the details for all sensitive personal data, select the icon next to the sensitive personal data percentage.



- 3. To investigate the details for a specific type of sensitive personal data, select **View All** and then select the **Investigate Results** arrow icon for a specific type of sensitive personal data.



- Investigate the data by searching, sorting, expanding details for a specific file, clicking **Investigate Results** to see masked information, or by downloading the file list.

View files by categories

BlueXP classification takes the data that it scanned and divides it into different types of categories. Categories are topics based on AI analysis of the content and metadata of each file. [See the list of categories.](#)

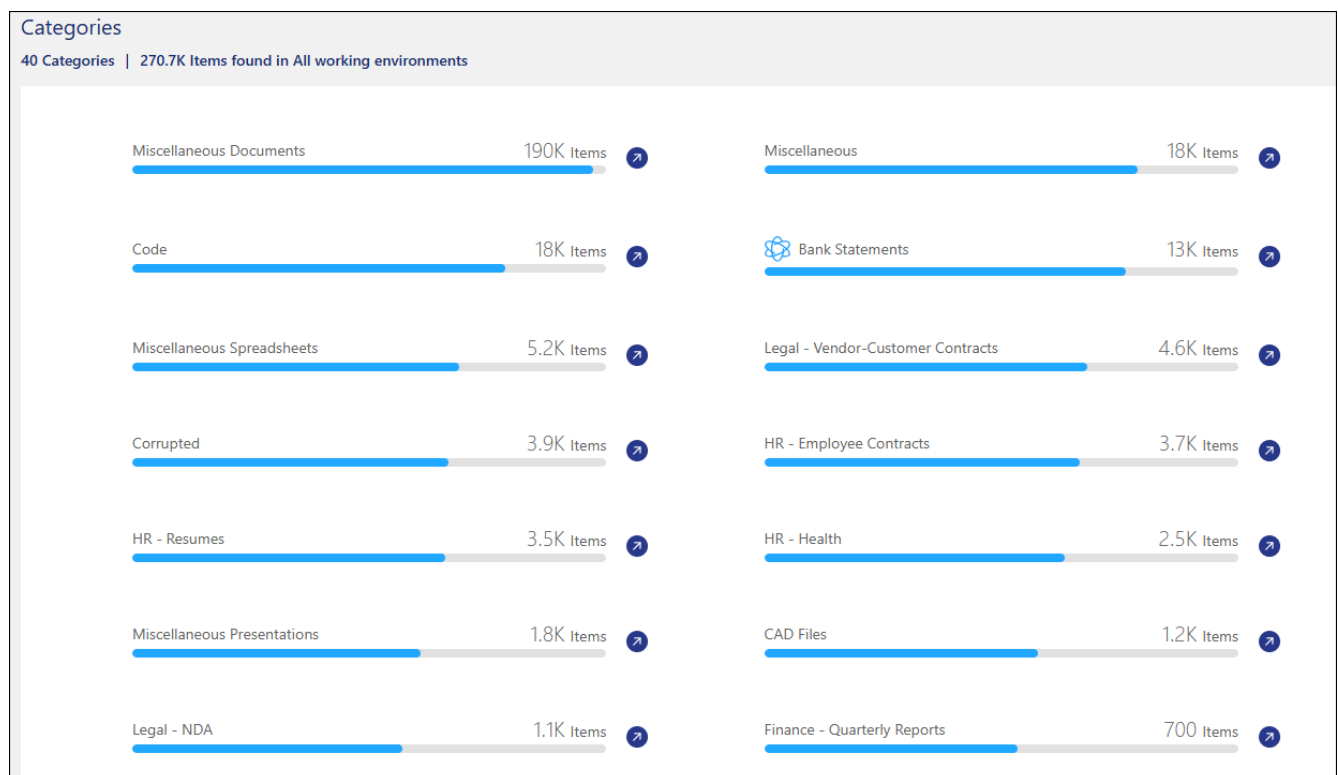
Categories can help you understand what's happening with your data by showing you the types of information that you have. For example, a category like resumes or employee contracts can include sensitive data. When you investigate the results, you might find that employee contracts are stored in a nonsecure location. You can then correct that issue.



English, German, and Spanish are supported for categories. Support for more languages will be added later.

Steps

- From the BlueXP classification menu, select the **Compliance** tab.
- Select the **Investigate Results** arrow icon for one of the top 4 categories directly from the main screen, or select **View All** and then select the icon for any of the categories.



- Investigate the data by searching, sorting, expanding details for a specific file, selecting the **Investigate Results** arrow to see masked information, or by downloading the file list.

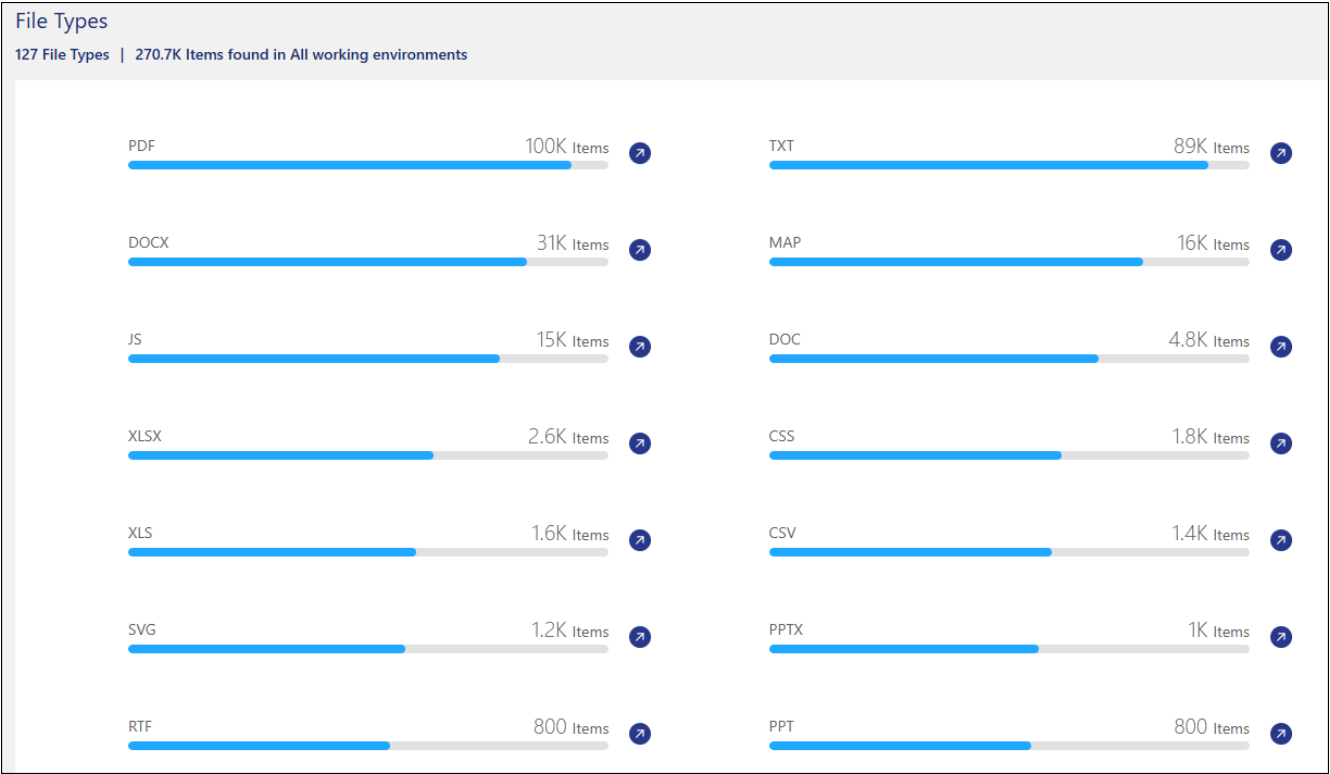
View files by file types

BlueXP classification takes the data that it scanned and breaks it down by file type. Reviewing your file types can help you control your sensitive data because you might find that certain file types are not stored correctly. [See the list of file types.](#)

For example, you might be storing CAD files that include very sensitive information about your organization. If they are unsecured, you can take control of the sensitive data by restricting permissions or moving the files to another location.

Steps

1. From the BlueXP classification menu, select the **Compliance** tab.
2. Select the **Investigate Results** arrow icon for one of the top 4 file types directly from the main screen, or select **View All** and then select the icon for any of the file types.




3. Investigate the data by searching, sorting, expanding details for a specific file, selecting the **Investigate Results** arrow to see masked information, or by downloading the file list.

Categories of private data in BlueXP classification

There are many types of private data that BlueXP classification can identify in your volumes and databases.

BlueXP classification identifies two types of personal data:

- **Personally identifiable information (PII)**
- **Sensitive personal information (SPII)**



If you need BlueXP classification to identify other private data types, such as additional national ID numbers or healthcare identifiers, contact your account manager.

Types of personal data

The personal data, or *personally identifiable information* (PII), found in files can be general personal data or

national identifiers. The third column in the table below identifies whether BlueXP classification uses [proximity validation](#) to validate its findings for the identifier.

The languages in which these items can be recognized are identified in the table.

Type	Identifier	Proximity validation?	English	German	Spanish	French	Japanese
General	Credit card number	Yes	✓	✓	✓		✓
	Data Subjects	No	✓	✓	✓		
	Email Address	No	✓	✓	✓		✓
	IBAN Number (International Bank Account Number)	No	✓	✓	✓		✓
	IP Address	No	✓	✓	✓		✓
	Password	Yes	✓	✓	✓		✓

Type	Identifier	Proximity validation?	English	German	Spanish	French	Japanese
National Identifiers							

Type	Identifier	Proximity validation?	English	German	Spanish	French	Japanese
------	------------	--------------------------	---------	--------	---------	--------	----------

Type	Identifier	Proximity validation?	English	German	Spanish	French	Japanese
------	------------	--------------------------	---------	--------	---------	--------	----------

	Corporate)						
	Latvian ID	Yes	✓	✓	✓		
Type	Lithuanian ID	Yes	✓	✓	✓		
	Luxembourg ID	Yes	✓	✓	✓		
	Maltese ID	Yes	✓	✓	✓		
	National Health Service (NHS) Number	Yes	✓	✓	✓		
	New Zealand Bank Account	Yes	✓	✓	✓		
	New Zealand Driver's License	Yes	✓	✓	✓		
	New Zealand IRD Number (Tax ID)	Yes	✓	✓	✓		
	New Zealand NHI (National Health Index) Number	Yes	✓	✓	✓		
	New Zealand Passport Number	Yes	✓	✓	✓		
	Polish ID (PESEL)	Yes	✓	✓	✓		
	Portuguese Tax Identification Number (NIF)	Yes	✓	✓	✓		
	Romanian ID (CNP)	Yes	✓	✓	✓		
	Singapore National Registration Identity Card (NRIC)	Yes	✓	✓	✓		
	Slovenian ID (EMSO)	Yes	✓	✓	✓		
	South African ID	Yes	✓	✓	✓		
	Spanish Tax Identification Number	Yes	✓	✓	✓		
	Swedish ID	Yes	✓	✓	✓		
	UK ID (NINO)	Yes	✓	✓	✓		
	USA California Driver's License	Yes	✓	✓	✓		
	USA Indiana Driver's License	Yes	✓	✓	✓		
	USA New York Driver's License	Yes	✓	✓	✓		
	USA Texas Driver's License	Yes	✓	✓	✓		
	USA Social Security Number (SSN)	Yes	✓	✓	✓		

Types of sensitive personal data

BlueXP classification can find the following sensitive personal information (SPII) in files.

The items in this category can be recognized only in English at this time.

- **Criminal Procedures Reference:** Data concerning a natural person's criminal convictions and offenses.
- **Ethnicity Reference:** Data concerning a natural person's racial or ethnic origin.
- **Health Reference:** Data concerning a natural person's health.
- **ICD-9-CM Medical Codes:** Codes used in the medical and health industry.
- **ICD-10-CM Medical Codes:** Codes used in the medical and health industry.

- **Philosophical Beliefs Reference:** Data concerning a natural person's philosophical beliefs.
- **Political Opinions Reference:** Data concerning a natural person's political opinions.
- **Religious Beliefs Reference:** Data concerning a natural person's religious beliefs.
- **Sex Life or Orientation Reference:** Data concerning a natural person's sex life or sexual orientation.

Types of categories

BlueXP classification categorizes your data as follows.

Most of these categories can be recognized in English, German, and Spanish.

Category	Type	English	German	Spanish
Finance	Balance Sheets	✓	✓	✓
	Purchase Orders	✓	✓	✓
	Invoices	✓	✓	✓
	Quarterly Reports	✓	✓	✓
HR	Background Checks	✓		✓
	Compensation Plans	✓	✓	✓
	Employee Contracts	✓		✓
	Employee Reviews	✓		✓
	Health	✓		✓
	Resumes	✓	✓	✓
Legal	NDA's	✓	✓	✓
	Vendor-Customer contracts	✓	✓	✓
Marketing	Campaigns	✓	✓	✓
	Conferences	✓	✓	✓
Operations	Audit Reports	✓	✓	✓
Sales	Sales Orders	✓	✓	
Services	RFI	✓		✓
	RFP	✓		✓
	SOW	✓	✓	✓
	Training	✓	✓	✓
Support	Complaints and Tickets	✓	✓	✓

The following metadata is also categorized and identified in the same supported languages:

- Application Data
- Archive Files

- Audio
- Breadcrumbs from BlueXP classification
Business Application Data
- CAD Files
- Code
- Corrupted
- Database and index files
- Design Files
- Email Application Data
- Encrypted (files with a high entropy score)
- Executables
- Financial Application Data
- Health Application Data
- Images
- Logs
- Miscellaneous Documents
- Miscellaneous Presentations
- Miscellaneous Spreadsheets
- Miscellaneous "Unknown"
- Password Protected files
- Structured Data
- Videos
- Zero-Byte Files

Types of files

BlueXP classification scans all files for category and metadata insights and displays all file types in the file types section of the dashboard. When BlueXP classification detects Personal Identifiable Information (PII) or when it performs a DSAR search, only the following file formats are supported:

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

Accuracy of information found

NetApp can't guarantee 100% accuracy of the personal data and sensitive personal data that BlueXP classification identifies. You should always validate the information by reviewing the data.

Based on our testing, the table below shows the accuracy of the information that BlueXP classification finds. We break it down by *precision* and *recall*:

Precision

The probability that what BlueXP classification finds has been identified correctly. For example, a precision rate of 90% for personal data means that 9 out of 10 files identified as containing personal information,

actually contain personal information. 1 out of 10 files would be a false positive.

Recall

The probability for BlueXP classification to find what it should. For example, a recall rate of 70% for personal data means that BlueXP classification can identify 7 out of 10 files that actually contain personal information in your organization. BlueXP classification would miss 30% of the data and it won't appear in the dashboard.

We are constantly improving the accuracy of our results. Those improvements will be automatically available in future BlueXP classification releases.

Type	Precision	Recall
Personal data - General	90%-95%	60%-80%
Personal data - Country identifiers	30%-60%	40%-60%
Sensitive personal data	80%-95%	20%-30%
Categories	90%-97%	60%-80%

Create a custom classification in BlueXP classification

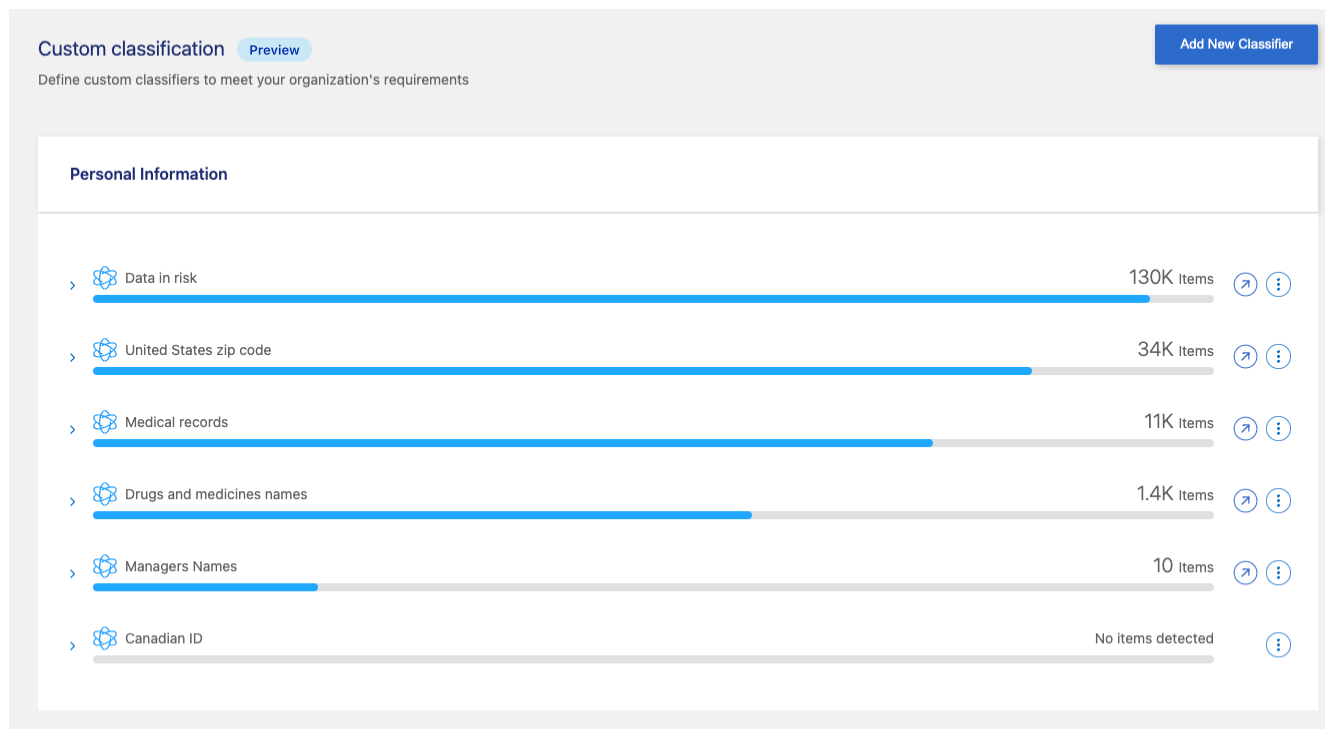
With BlueXP classification, you can create a custom search for sensitive information. The search can be scoped to a regular expression (regex).

Create a custom classification

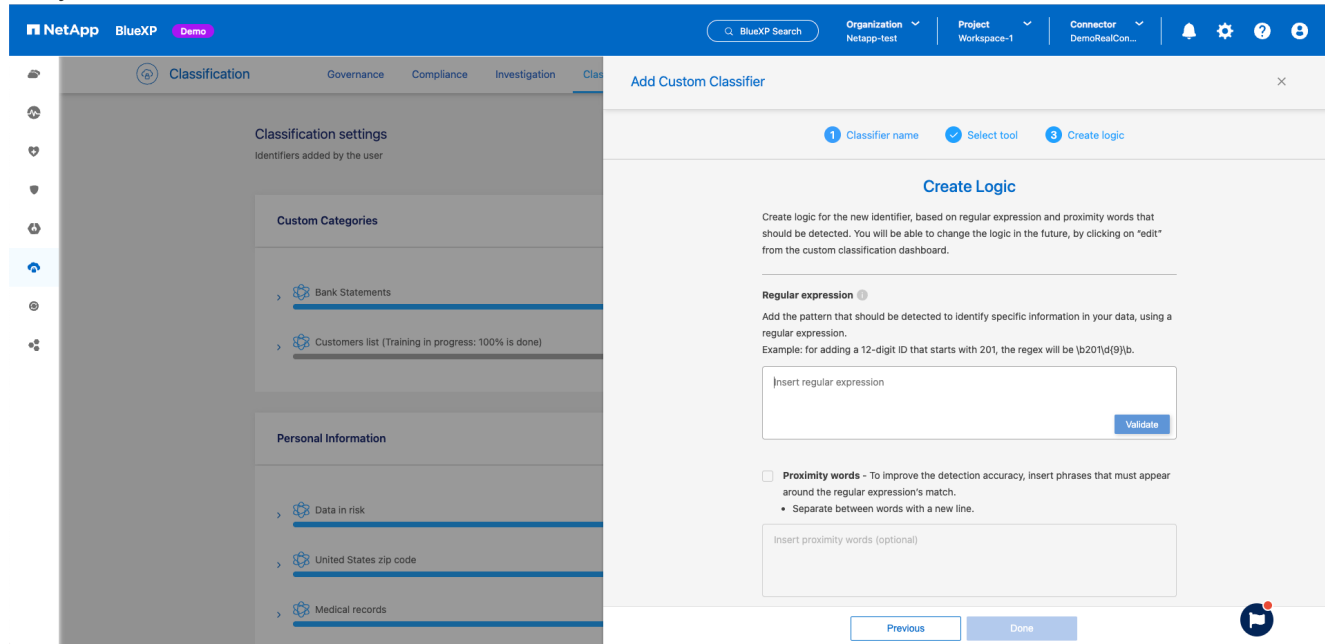
Custom classification is only available for Map & Classify scans, not mapping-only scans. This feature is currently in preview.

Steps

1. Select the **Custom classification** tab.



2. Select the **Add New Classifier** button.
3. Add a Name and Description for the new classifier.
4. To add the customization as a regular expression, select **Custom regular expression** then **Next**.
5. Add a pattern to detect the specific information of your data. Select **Validate** to confirm the syntax of your entry.



6. Select **Done** to create the custom classification.

The new customization is captured in the next scheduled scan. To view results, see [Generate compliance reports](#).

Investigate the data stored in your organization with BlueXP classification

Investigate the data from your organization by viewing details in the Data Investigation page. Here is where you can continue your research after looking at the Governance dashboard. On the Investigation page, you can filter the data using one of the many filters to show only the results you want to see. You can also view file metadata, permissions for files and directories, and check for duplicate files in your storage systems.

You can navigate to this page from many areas of the BlueXP classification UI, including the Governance and Compliance dashboards with the filters selected already on those pages. You can export the data into a CSV or JSON file for further analysis or to share with others.



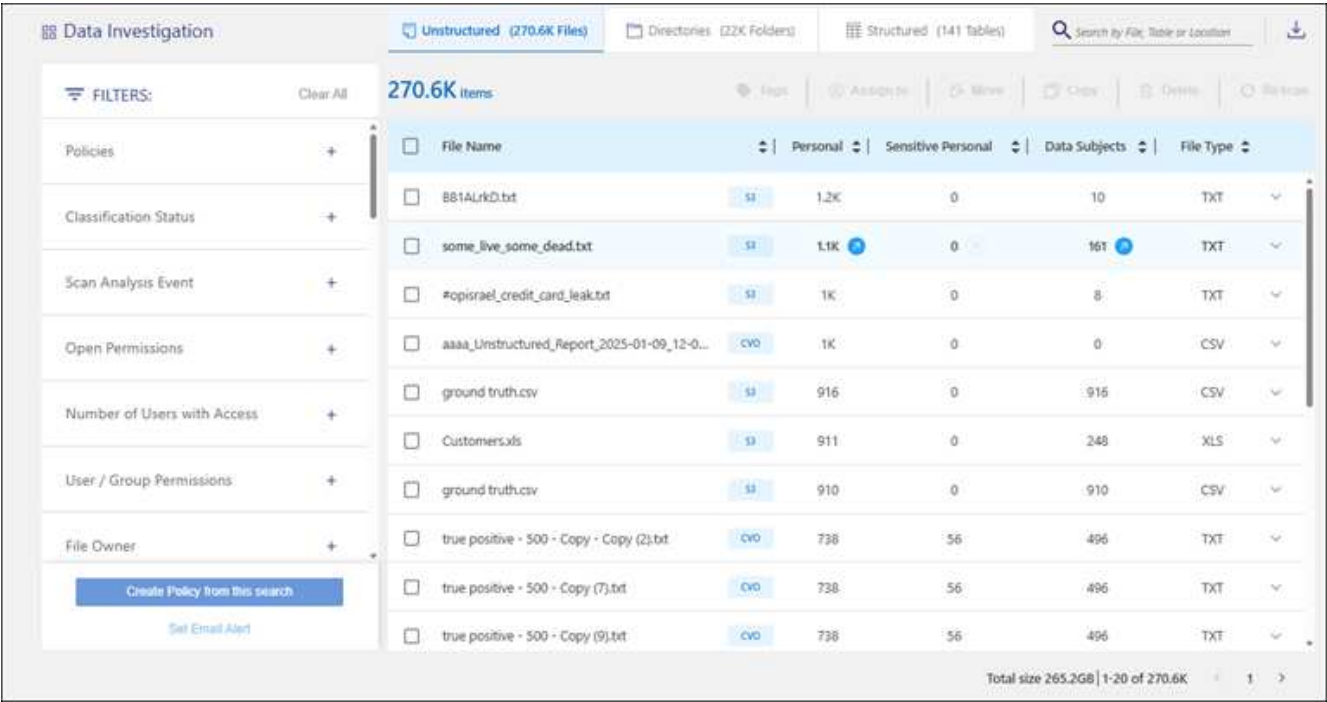
The capabilities described in this section are available only if you have chosen to perform a full classification scan on your data sources. Data sources that have had a mapping-only scan do not show file-level details.

Filter data in the Data Investigation page

You can filter the contents of the investigation page to display only the results you want to see.

Steps

1. From the BlueXP classification menu, select **Investigation**.
2. On the Data Investigation page, do any of the following:
3. To download the contents of the page as a report after you’ve refined it, select the button.



4. To view the data from files (unstructured data), directories (folders and file shares), or from databases (structured data), select one of the tabs at the top.

5. To sort the results in numerical or alphabetical order, select the control at the top of each column.
6. To refine the results even more, select one of the filters in the Filter pane.



You can only view the first 10,000 results—or 500 pages—for a scan on the Data Investigation page.

Filter data by sensitivity and content

Use the following filters to view how much sensitive information is contained in your data.

Filter	Details
Category	Select the types of categories .
Sensitivity Level	Select the sensitivity level: Personal, Sensitive personal, or Non sensitive.
Number of identifiers	Select the range of detected sensitive identifiers per file. Includes personal data and sensitive personal data. When filtering in Directories, BlueXP classification totals the matches from all files in each folder (and sub-folders). NOTE: The December 2023 (version 1.26.6) release removed the option to calculate the number of personal identifiable information (PII) data by Directories.
Personal Data	Select the types of personal data .
Sensitive Personal Data	Select the types of sensitive personal data .
Data Subject	Enter a data subject's full name or known identifier. Learn more about data subjects here .

Filter data by user owner and user permissions

Use the following filters to view file owners and permissions to access your data.

Filter	Details
Open Permissions	Select the type of permissions within the data and within folders/shares.
User / Group Permissions	Select one or multiple user names and/or group names, or enter a partial name.
File Owner	Enter the file owner name.
Number of users with access	Select one or multiple category ranges to show which files and folders are open to a certain number of users.

Filter data by time

Use the following filters to view data based on time criteria.

Filter	Details
Created Time	Select a time range when the file was created. You can also specify a custom time range to further refine the search results.

Filter	Details
Discovered Time	Select a time range when BlueXP classification discovered the file. You can also specify a custom time range to further refine the search results.
Last Modified	Select a time range when the file was last modified. You can also specify a custom time range to further refine the search results.
Last Accessed	<p>Select a time range when the file, or directory (CIFS or NFS only), was last accessed. You can also specify a custom time range to further refine the search results. For the types of files that BlueXP classification scans, this is the last time BlueXP classification scanned the file.</p> <p>BlueXP classification does not extract the "last accessed time" from the following data sources: SharePoint Online, SharePoint On-premises (SharePoint Server), OneDrive, Google Drive, and Amazon S3.</p>

Filter data by metadata

Use the following filters to view data based on location, size, and directory or file type.

Filter	Details
File Path	Enter up to 20 partial or full paths that you want to include or exclude from the query. If you enter both include paths and exclude paths, BlueXP classification finds all files in the included paths first, then it removes files from excluded paths, and then it displays the results. Note that using "*" in this filter has no effect, and that you can't exclude specific folders from the scan - all the directories and files under a configured share will be scanned.
Directory Type	Select the directory type; either "Share" or "Folder".
File Type	Select the types of files .
File Size	Select the file size range.
File Hash	Enter the file's hash to find a specific file, even if the name is different.

Filter data by storage type

Use the following filters to view data by storage type.

Filter	Details
Working Environment Type	Select the type of working environment. OneDrive, SharePoint, and Google Drive are categorized under "Apps".
Working Environment name	Select specific working environments.
Storage Repository	Select the storage repository, for example, a volume or a schema.

Filter data by saved searches

Use the following filter to view data by saved searches.

Filter	Details
Saved search	Select one saved search or multiples. Go to the saved searches tab to view the list of existing saved searches and create new ones.

Filter data by analysis status

Use the following filter to view data by the BlueXP classification scan status.

Filter	Details
Analysis Status	Select an option to show the list of files that are Pending First Scan, Completed being scanned, Pending Rescan, or that have Failed to be scanned.
Scan Analysis Event	Select whether you want to view files that were not classified because BlueXP classification couldn't revert last accessed time, or files that were classified even though BlueXP classification couldn't revert last accessed time.

See [details about the "last accessed time" timestamp](#) for more information about the items that appear in the Investigation page when filtering using the Scan Analysis Event.

Filter data by duplicates

Use the following filter to view files that are duplicated in your storage.

Filter	Details
Duplicates	Select whether the file is duplicated in the repositories.

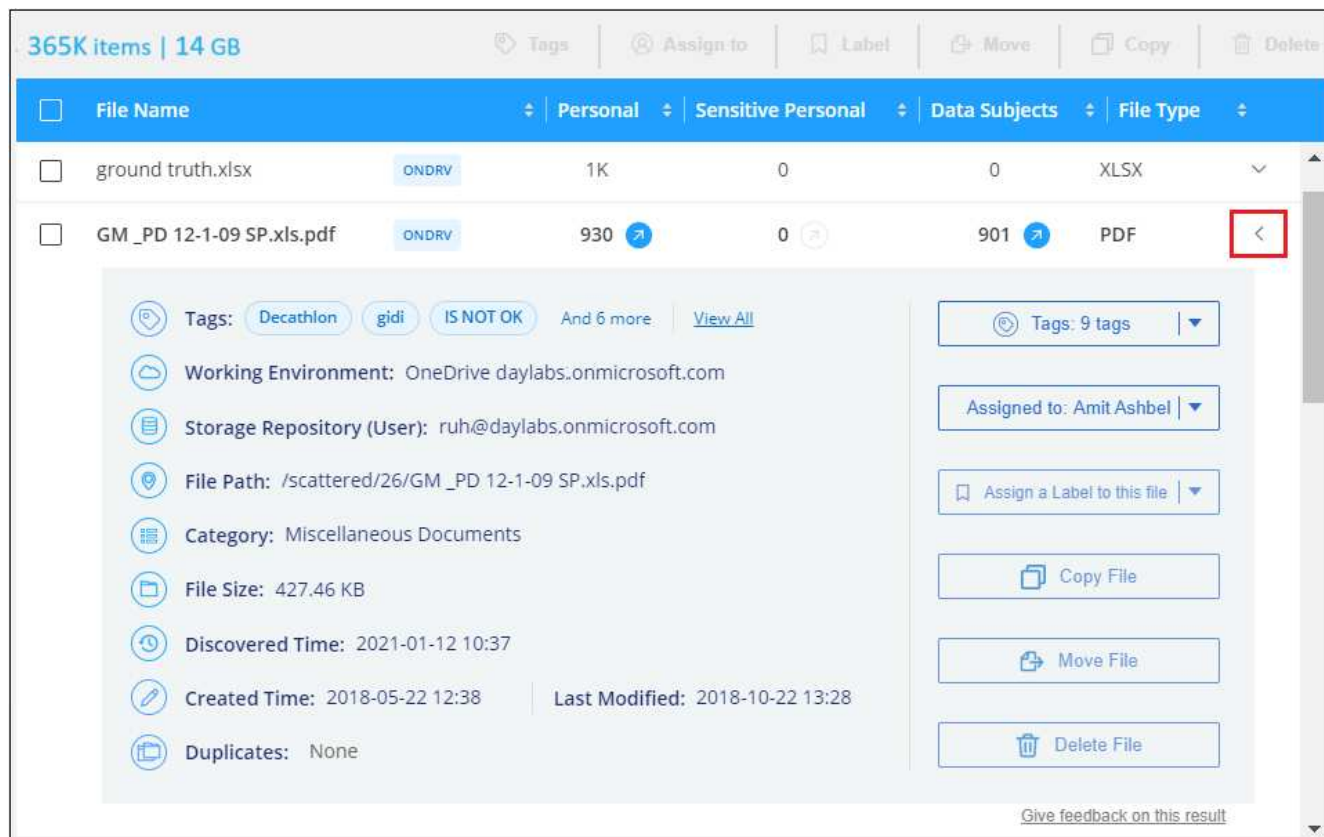
View file metadata

In addition to showing you the working environment and volume where the file resides, the metadata shows much more information, including the file permissions, file owner, and whether there are duplicates of this file. This information is useful if you're planning to [create saved searches](#) because you can see all the information that you can use to filter your data.

The availability of information depends on the data source. For example, volume name and permissions are not shared for database files.

Steps

1. From the BlueXP classification menu, select **Investigation**.
2. In the Data Investigation list on the right, select the down-caret ▼ on the right for any single file to view the file metadata.



View users' permissions for files and directories

To view a list of all users or groups who have access to a file or to a directory and the types of permissions they have, select **View all Permissions**. This button is available only for data in CIFS shares.

Note that if you see SIDs (Security IDentifiers) instead of user and group names, you should integrate your Active Directory into BlueXP classification. [See how to do this](#).

Steps

1. From the BlueXP classification menu, select **Investigation**.
2. In the Data Investigation list on the right, select the down-caret ▼ on the right for any single file to view the file metadata.
3. To view a list of all users or groups who have access to a file or to a directory and the types of permissions they have, in the Open Permissions field, select **View all Permissions**.



BlueXP classification shows up to 100 users in the list.

File Name: Expense Report TPO-1060.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

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		✓	✓

- Select the down-caret ▼ button for any group to see the list of users who are part of the group.



You can expand one level of the group to see the users who are part of the group.

- Select the name of a user or group to refresh the Investigation page so you can see all the files and directories that the user or group has access to.

Check for duplicate files in your storage systems

You can check whether duplicate files are being stored in your storage systems. This is useful if you want to identify areas where you can save storage space. It's also good to ensure certain files that have specific permissions or sensitive information are not unnecessarily duplicated in your storage systems.

All of your files (not including databases) that are 1 MB or larger, or that contain personal or sensitive personal information, are compared to see if there are duplicates.

BlueXP classification uses hashing technology to determine duplicate files. If any file has the same hash code as another file, you can be 100% sure that the files are exact duplicates—even if the file names are different.


Steps

- From the BlueXP classification menu, select **Investigation**.
- In the Investigation page Filters pane on the left, select "File Size" along with "Duplicates" ("Has duplicates") to see which files of a certain size range are duplicated in your environment.
- Optionally, download the list of duplicate files and send it to your storage admin so they can decide which files, if any, can be deleted.
- Optionally, [delete the file](#) yourself if you are confident that a specific version of the file is not needed.

View if a specific file is duplicated

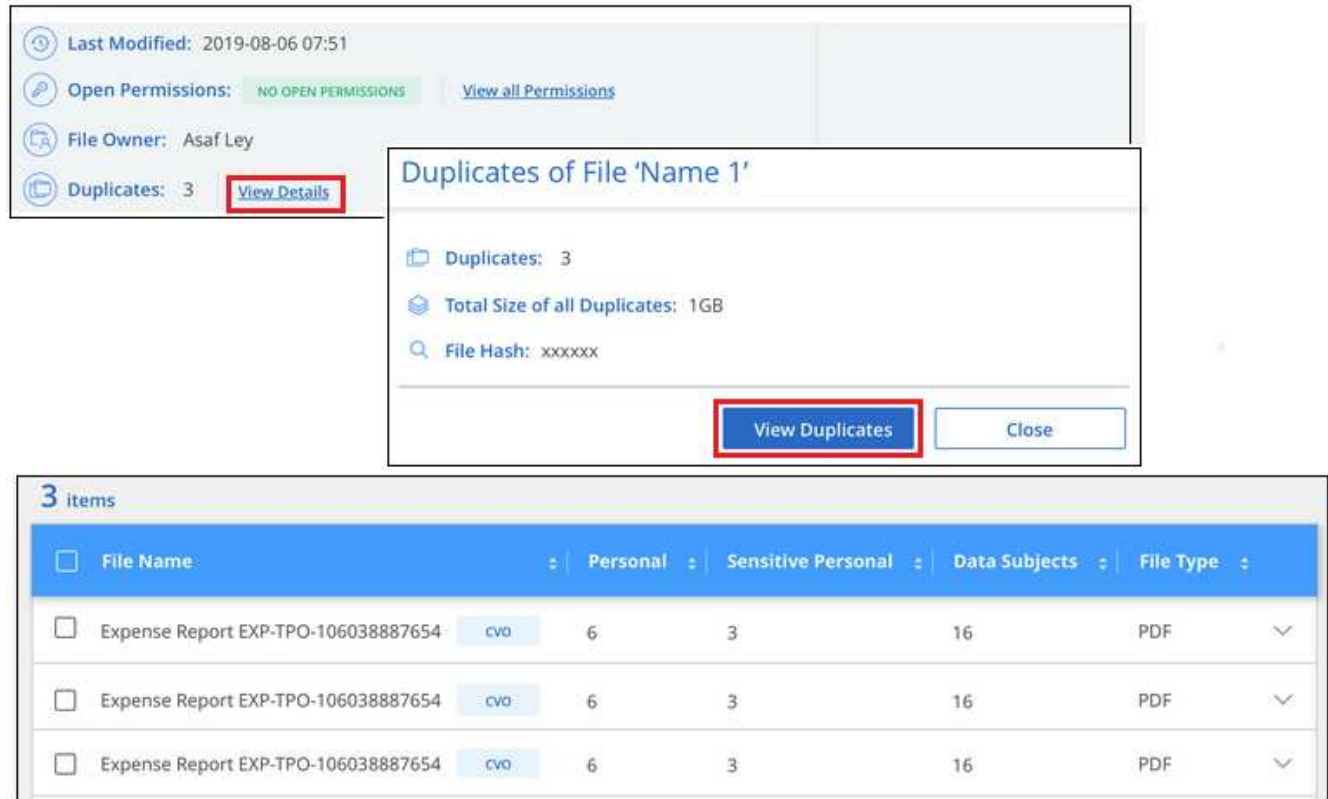
You can see if a single file has duplicates.

Steps

1. From the BlueXP classification menu, select **Investigation**.
2. In the Data Investigation list, select  on the right for any single file to view the file metadata.

If duplicates exist for a file, this information appears next to the *Duplicates* field.

3. To view the list of duplicate files and where they are located, select **View Details**.
4. In the next page select **View Duplicates** to view the files in the Investigation page.



The screenshot shows the BlueXP interface. At the top, file metadata is displayed: Last Modified: 2019-08-06 07:51, Open Permissions: NO OPEN PERMISSIONS (with a View all Permissions link), File Owner: Asaf Ley, and Duplicates: 3 (with a View Details button highlighted in a red box). Below this, a modal titled 'Duplicates of File 'Name 1'' is open. It shows Duplicates: 3, Total Size of all Duplicates: 1GB, and File Hash: xxxxxx. A View Duplicates button (highlighted in a red box) and a Close button are at the bottom of the modal. Below the modal, a table titled '3 items' displays a list of duplicate files. The table has columns for File Name, Personal, Sensitive Personal, Data Subjects, and File Type. Three identical rows are shown, each representing an 'Expense Report EXP-TPO-106038887654' PDF file.

File Name	Personal	Sensitive Personal	Data Subjects	File Type
Expense Report EXP-TPO-106038887654	6	3	16	PDF
Expense Report EXP-TPO-106038887654	6	3	16	PDF
Expense Report EXP-TPO-106038887654	6	3	16	PDF



You can use the "file hash" value provided in this page and enter it directly in the Investigation page to search for a specific duplicate file at any time - or you can use it in a saved search.

Create the Data Investigation Report

The Data Investigation Report is a download of the filtered contents of the Data Investigation page.

The report is available as a CSV or JSON file you can save to your local machine.

There can be up to three report files downloaded if BlueXP classification is scanning files (unstructured data), directories (folders and file shares), and databases (structured data).

The files are split into files with a fixed number of rows or records:

- JSON - 100,000 records per report that takes about 5 minutes to generate
- CSV - 200,000 records per report that takes about 4 minutes to generate



You can download a version of the CSV file to view in this browser. This version is limited to 10,000 records.

What's included in the Data Investigation Report

The **Unstructured Files Data Report** includes the following information about your files:

- File name
- Location type
- Working environment name
- Storage repository (for example, a volume, bucket, shares)
- Repository type
- File path
- File type
- File size (in MB)
- Created time
- Last modified
- Last accessed
- File owner
 - File owner data encompasses account name, SAM account name, and e-mail address when Active Directory is configured.
- Category
- Personal information
- Sensitive personal information
- Open permissions
- Scan Analysis Error
- Deletion detection date

The deletion detection date identifies the date that the file was deleted or moved. This enables you to identify when sensitive files have been moved. Deleted files don't contribute to the file number count that appears in the dashboard or on the Investigation page. The files only appear in the CSV reports.

The **Unstructured Directories Data Report** includes the following information about your folders and file shares:


- Working environment type
- Working environment name
- Directory name
- Storage repository (for example, a folder or file shares)
- Directory owner
- Created time
- Discovered time

- Last modified
- Last accessed
- Open permissions
- Directory type

The **Structured Data Report** includes the following information about your database tables:

- DB Table name
- Location type
- Working environment name
- Storage repository (for example, a schema)
- Column count
- Row count
- Personal information
- Sensitive personal information

Steps to generate the report

1. From the Data Investigation page, select the  button on the top, right of the page.
2. Choose the report type: CSV or JSON.
3. Enter a **Report name**.
4. To download the complete report, select **Working environment** then choose the **Working Environment** and **Volume** from the respective dropdown menus. Provide a **Destination folder path**.

To download the report in the browser, select **Local** . Note this option limits the report to the first 10,000 rows and is limited to the **CSV** format. You don't need to complete any other fields if you select **Local**.

5. Select **Download Report**.

Download Investigation Report

Report type

☒ CSV file ☐ JSON file

Report name

investigation_report

Export destination

☐ Working environment ☒ Local (limited to 10K rows)

Working environment ⓘ

Working environment ▼

Volume

Type to search... ▼

Destination folder path

/folder/subfolder

Download Report

Cancel

Result

A dialog displays a message that the reports are being downloaded.

Create a saved search based on selected filters

You can create a saved search for frequently used search filters in the Data Investigation page to easily replicate those search queries.

Steps

1. From the BlueXP classification menu, select **Investigation**.
2. On the Data Investigation page, select the filters you want to use to create a saved search.
3. At the bottom of the Filter pane, select **Create saved search from this search**.
4. Enter a name and a description for the saved search.
5. Choose any of the following:
6. Select **Create Saved Search**.



It might take up to 15 minutes for the results to appear on the Saved Searches page.

Manage saved searches with BlueXP classification

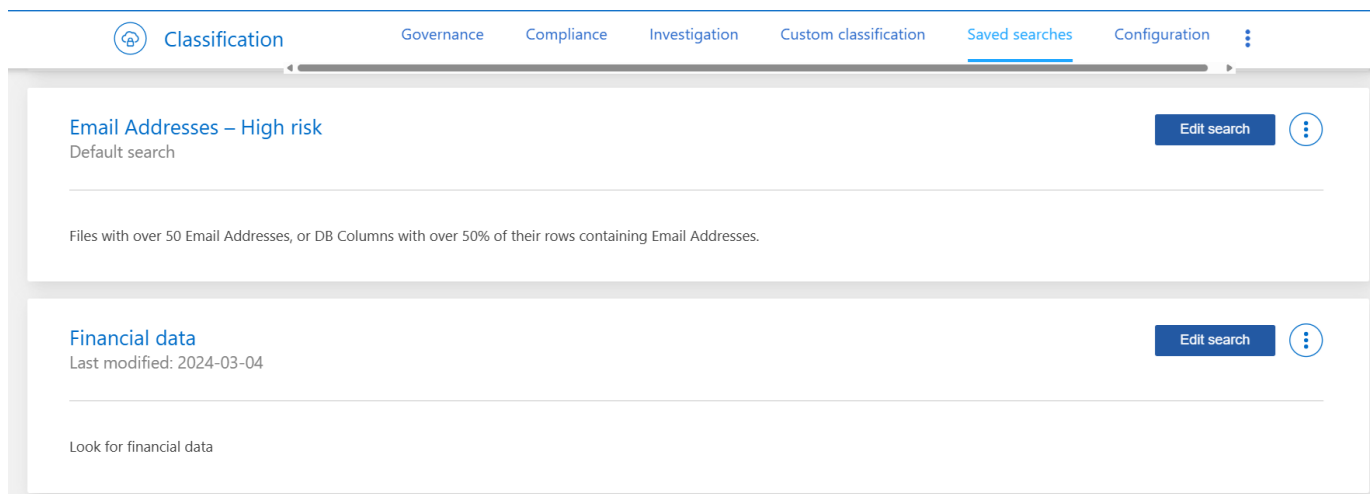
BlueXP classification supports saving your search queries. With a saved search, you can create custom filters to sort through frequent queries of your data Investigation page. BlueXP classification also includes predefined saved searches based on common requests.




In versions of BlueXP classification earlier than 1.43, saved searches were called [policies](#).

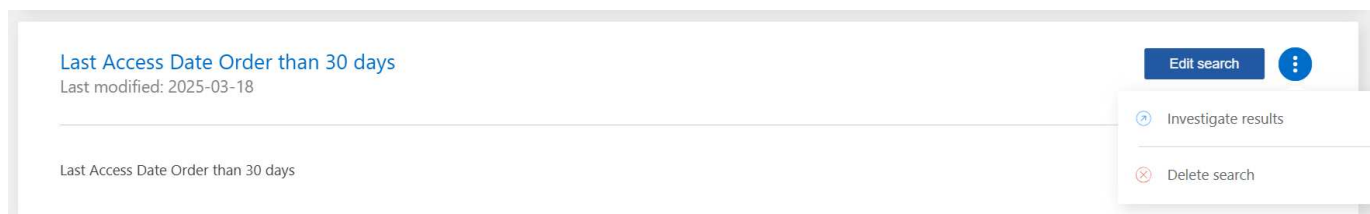
The **Saved searches** tab in the Compliance Dashboard lists all the predefined and custom saved searches available on this instance of BlueXP classification.

Saved searches also appear in the list of filters in the Investigation page.



View saved searches results in the Investigation page

To display the results for a saved search in the Investigation page, select the  button for a specific search then select **Investigate Results**.



Create custom saved searches

You can create your own custom saved searches that provide results for queries specific to your organization. Results are returned for all files and directories (shares and folders) that match the search criteria.

Steps

1. In the Investigation tab, define a search by selecting the filters you want to use. See [Filtering data in the Investigation page](#) for details.
2. Once you have all the filter characteristics set to your liking, select **Create saved search**.

Data Investigation

 FILTERS:

Clear All

Storage Repository **3**

+

File / Directory Path

+

Category

+

Sensitivity Level

+

Save this search

3. Name the saved search and add a description. The name must be unique.
4. Select **Create Saved Search**.

Create search

This will save the current selected filters and search term as a saved search. You can view or delete this later from the "Saved searches" tab.

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

Name this search

Give it a detailed description that explains what it searches for

Create search

Cancel

Once you've created the search, you can view it in the **Saved searches** tab.

Edit saved searches

You can modify the query criteria for a saved search (that is, the defined filters) to add or remove certain parameters.

You cannot modify default saved searches.

Steps

1. From the Saved searches page, select **Edit Search** for the search that you want to change.

Sensitive data
Last modified: 2024-03-04

Edit search



Look for sensitive data from 2 years ago

2. Make the changes to the name and description fields. To only change the name and description fields, select **Save search**.

To change the filters for the saved search, select **Edit query**.

Name this search

Edit query

Name this search

Sensitive data

Give it a detailed description that explains what it searches for

Look for sensitive data from 2 years ago

Save search

Cancel

3. In the Investigation page, edit the query. You can add, remove, or modify filters. To complete your changes, select **Save query for this search**.

Data Investigation

Unstructured (217.3K Files) | Directories (0 Folders) | Structured (0 Tables) | Search by File, Table or Location

FILTERS:

Clear All

Last Modified 3 +

Last Accessed +

Duplicates +

File Hash ⓘ +

Save query for search

Cancel edit query


217.3K items

File Name		Personal	Sensitive Personal	Data Subjects	File Type
customers		107	618	107	TXT
true positive.txt	cvo	107	618	107	TXT
true positive.txt	s3	107	618	107	TXT
true positive.txt	cvo	107	618	107	TXT
true positive.txt	s3	107	618	107	TXT
true positive.txt	s3	107	618	107	TXT
true positive.txt	s3	107	618	107	TXT

Total size 228.6GB | 1-20 of 1

Delete saved searches

You can delete any custom saved search if you no longer need it. You can't delete default saved searches.

To delete a saved search, select the  button for a specific search, select **Delete search**, then select **Delete search** again in the confirmation dialog.

Default searches

BlueXP classification provides the following system-defined search queries:

- **Data Subject names - High risk**

Files with more than 50 data subject names

- **Email Addresses - High risk**

Files with more than 50 email addresses or database columns with more than 50% of their rows containing email addresses

- **Personal data - High risk**

Files with more than 20 personal data identifiers or database columns with more than 50% of their rows containing personal data identifiers

- **Private data - Stale over 7 years**

Files containing personal or sensitive personal information, last modified more than 7 years ago

- **Protect - High**

Files or database columns that contain a password, credit card information, IBAN number, or social security number

- **Protect - Low**

Files that have not been accessed for more than 3 years

- **Protect - Medium**

Files that contain files or database columns with personal data identifiers including ID numbers, tax identification numbers, drivers license numbers, medicinal IDs, or passport numbers

- **Sensitive Personal data - High risk**

Files with more than 20 sensitive personal data identifiers or database columns with greater than 50% of their rows containing sensitive personal data

Change the BlueXP classification scan settings for your repositories

You can manage how your data is being scanned in each of your working environments and data sources. You can make the changes on a "repository" basis; meaning you can

make changes for each volume, schema, user, etc. depending on the type of data source you are scanning.

Some of the things you can change are whether a repository is scanned or not, and whether BlueXP classification is performing a [mapping scan](#) or a [mapping & classification scan](#). You can also pause and resume scanning, for example, if you need to stop scanning a volume for a period of time.

View the scan status for your repositories

You can view the individual repositories that BlueXP classification is scanning (volumes, buckets, etc.) for each working environment and data source. Additionally, you can see how many have been "Mapped", and how many have been "Classified". Classification takes a longer time as the full AI identification is being performed on all data.

You can view the scanning status of each work environment on the Configuration page:

- **Initializing** (light blue dot): The map or classify configuration is activated. This appears for few seconds before starting the "pending queue" status.
- **Pending queue** (orange dot): The scan task is waiting to be listed in the scanning queue.
- **Queued** (orange dot): The task was successfully added to the scanning queue. The system will start mapping or classifying the volume when its turn in the queue arrives.
- **Running** (green dot): The scan task, which was in the queue, is actively in progress on the selected storage repository.
- **Finished** (green dot): The scan of the storage repository is complete.
- **Paused** (gray dot): You selected the "Pause" option to pause scanning. While the changes in the volume are not displayed in the system, the scanned insights are still shown.
- **Error** (red dot): The scan cannot complete because it has encountered issues. If you need to complete an action, the error appears in the tooltip under the "Required action" column. Otherwise, the system shows an "error" status and tries to recover. When it finishes, the status changes.
- **Not scanning**: The volume configuration of "Off" was selected and the system is not scanning the volume.

Steps

1. From the BlueXP classification menu, select **Configuration**.

The screenshot displays the BlueXP Configuration interface. On the left, a 'Quick Navigation' sidebar lists 'Identity Services' (selected), 'Working Environments', and 'Scanner Groups'. The main content area is titled 'Identity Services' and includes a top bar with 'Active Directory Integrated' and 'Add Working Environment'. Below this, a card for 'share2scan.netapp.com' shows user 'shragaga@share2scan.netapp.com' and IP '10.128.0.188'. The '11 Working Environments' section has filters for 'S3', 'CVO', 'DB', and 'SHARES'. A card for 'S3 - 055518636490 | 50 Buckets' (Amazon S3) shows 'Scanner Group name: default' and 'Working Environment ID: S3'. A 'Scan Mode' progress bar indicates '16 Classified' (green), '16 Mapped' (blue), and '34 Not Scanned' (gray). A note states 'Continuously scanning all selected Buckets'. A 'Configuration' button and a notification icon are also visible.

2. From the Configuration tab, select the **Configuration** button for the working environment.
3. In the Scan Configuration page, view the scan settings for all repositories.

S3 - 055518636490 Scan Configuration

Buckets selected for Classification scan (16/50) Retry All

Scan	Storage Repository (Bucket)	Mapping status	Scan progress	Required Action
<input type="button" value="Off"/> <input type="button" value="Map"/> <input type="button" value="Map & Classify"/>	allenc-demo-tlveng-demo	Paused 2025-02-27 15:15 Last full cycle: 2024-10-23 08:15	Mapped 7 Classified 7	...
<input type="button" value="Off"/> <input type="button" value="Map"/> <input type="button" value="Map & Classify"/>	audit-doc-export			...

4. Hover your cursor over the chart in the *Mapping Status* column to see the number of files that remain to be mapped or classified in each repository (bucket in this example).

Change the type of scanning for a repository

You can start or stop mapping-only scans, or mapping and classification scans, in a working environment at any time from the Configuration page. You can also change from mapping-only scans to mapping and classification scans, and vice-versa.



Databases can't be set to mapping-only scans. Database scanning can be Off or On; where On is equivalent to Map & Classify.

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration tab, select the **Configuration** button for the working environment.

Quick Navigation: Identity Services, Working Environments, Scanner Groups

Identity Services

Active Directory Integrated Add Working Environment

share2scan.netapp.com Edit ...

shragaqa@share2scan.netapp.com IP 10.128.0.188

11 Working Environments

Filter by: S3 CVO DB SHARES Clear filters

S3 - 055518636490 | 50 Buckets Scanner Group name: default Working Environment ID: S3 Configuration ...

Amazon S3

Scan Mode

16 Classified 16 Mapped 34 Not Scanned

Continuously scanning all selected Buckets

3. In the Scan Configuration page, change any of the repositories (buckets in this example) to perform **Map** or **Map & Classify** scans.

S3 - 055518636490 Scan Configuration					
Buckets selected for Classification scan (16/50)					Retry All
Scan	Storage Repository (Bucket)	Mapping status	Scan progress	Required Action	
<input type="button" value="Off"/> <input type="button" value="Map"/> <input type="button" value="Map & Classify"/>	allenc-demo-tiveng-demo	Paused 2025-02-27 15:15 Last full cycle: 2024-10-23 08:15	Mapped 7 Classified 7	...	
<input type="button" value="Off"/> <input type="button" value="Map"/> <input type="button" value="Map & Classify"/>	audit-doc-export			...	

Certain types of working environments enable you to change the type of scanning globally for all repositories using a button bar at the top of the page. This is valid for Cloud Volumes ONTAP, on-premises ONTAP, Azure NetApp Files, and Amazon FSx for ONTAP systems.

The example below shows this button bar for an Azure NetApp Files system.

Azure NetApp Files Scan Configuration

3/3 Volumes selected for Data Sense scan

[Learn about the differences between Mapping and Classification →](#)

[Edit CIFS Credentials](#)

Prioritize scans

You can prioritize the most important mapping-only scans or map & classify scans to ensure high priority scans are completed first.

By default, scans are queued based on the order in which they are initiated. With the ability to prioritize scans, you can move scans to the front of the queue. Multiple scans can be prioritized. Priority is designated in a first-in, first-out order, meaning the first scan you prioritize moves to the front of the queue; the second scan you prioritize becomes second in the queue, and so forth.

Priority is granted on a one-time basis. Automatic rescans of mapping data occur in the default order.

Steps

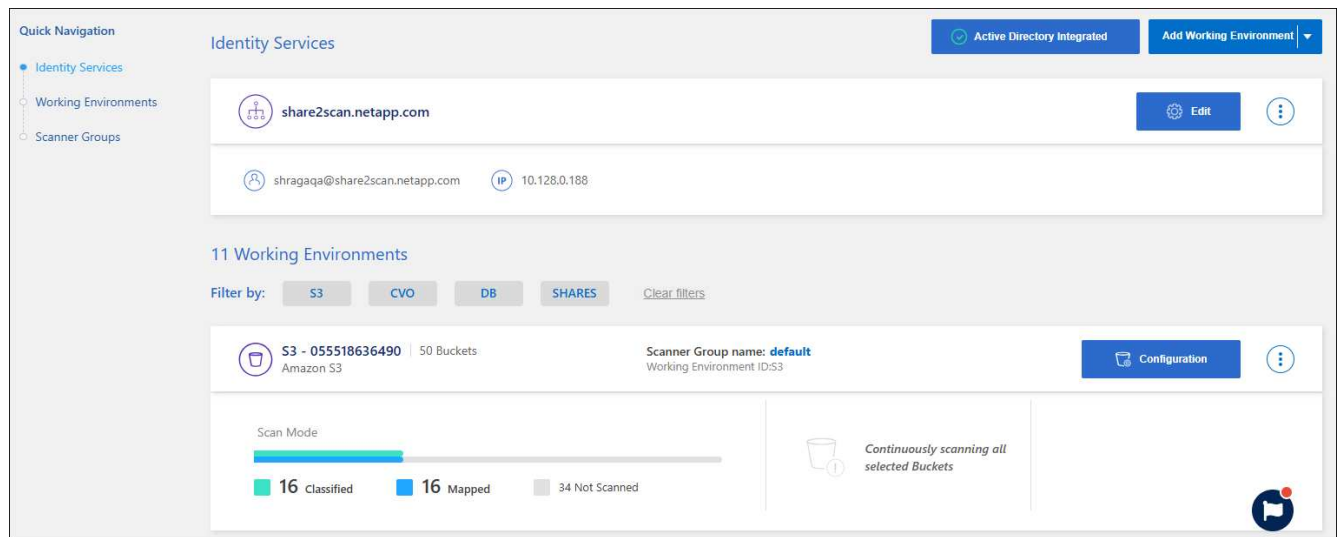
1. From the BlueXP classification menu, select **Configuration**.
2. Select the resources you want to prioritize.
3. From the Actions ... option, select **Prioritize scan**.

Stop scanning for a repository

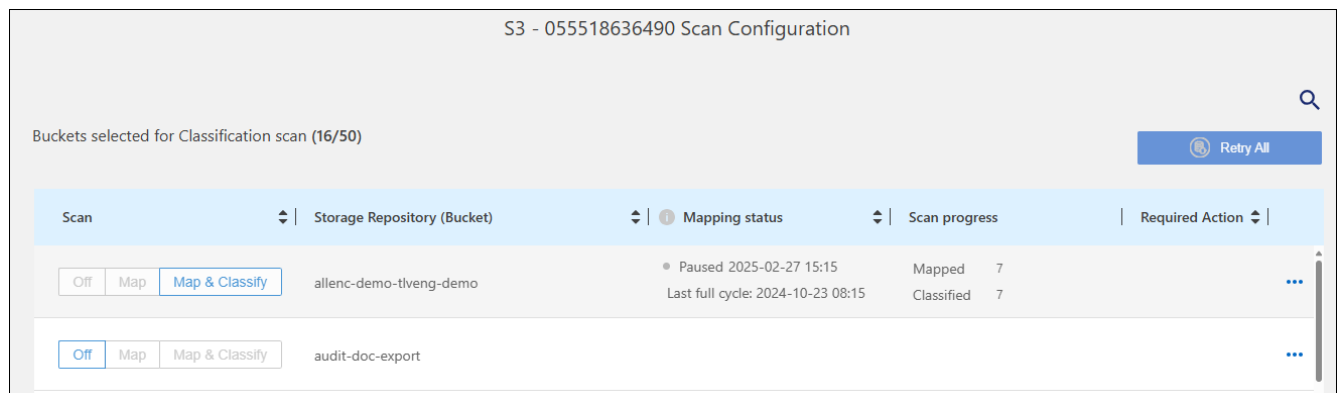
You can stop scanning a repository (for example, a volume) if you no longer need to monitor it for compliance. You do this by turning scanning "off". When scanning is turned off, all the indexing and information about that volume is removed from the system, and charging for scanning the data is stopped.

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration tab, select the **Configuration** button for the working environment.



3. In the Scan Configuration page select **Off** to stop scanning for a particular bucket.



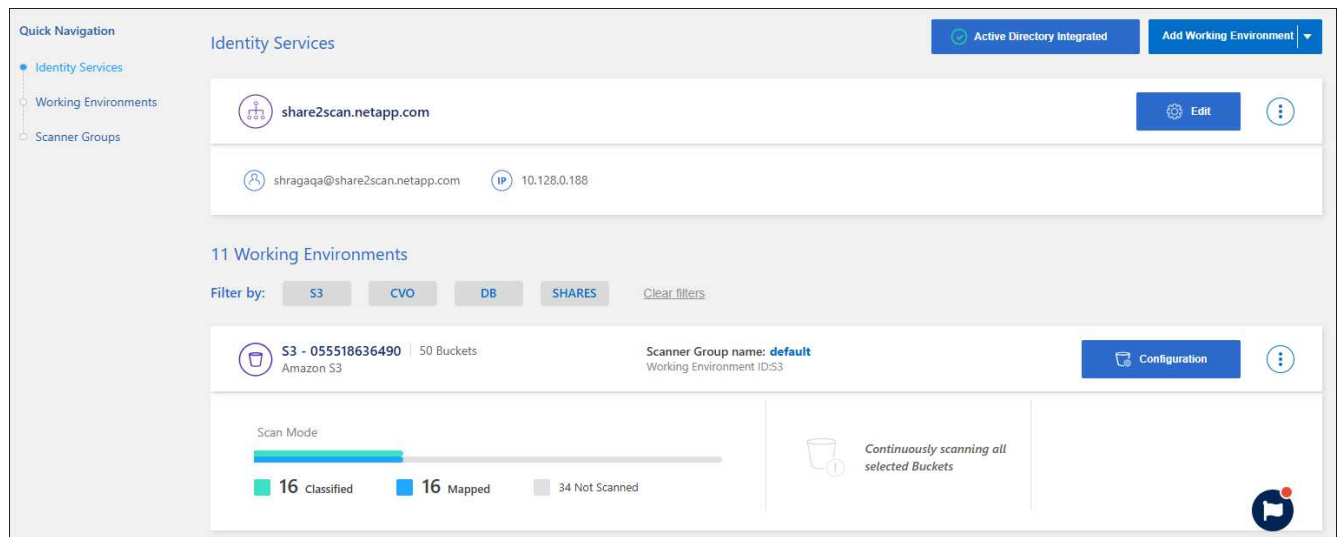
Pause and resume scanning for a repository


You can "pause" scanning on a repository if you want to temporarily stop scanning certain content. Pausing scanning means that BlueXP classification won't perform any future scans for changes or additions to the repository, but that all the current results will still be displayed in the system. Pausing scanning does not stop charging for the scanned the data because the data still exists.

You can "resume" scanning at any time.

Steps

1. From the BlueXP classification menu, select **Configuration**.
2. From the Configuration tab, select the **Configuration** button for the working environment.



3. In the Scan Configuration page, select the Actions  icon.
4. Select **Pause** to pause scanning for a volume, or select **Resume** to resume scanning for a volume that had been previously paused.

View BlueXP classification compliance reports

BlueXP classification provides reports that you can use to better understand the status of your organization's data privacy program.

By default, the BlueXP classification dashboards display compliance and governance data for all working environments, databases, and data sources. If you want to view reports that contain data for only some of the working environments, you can filter to see just them.



- The reports described in this section are available only if you have chosen to perform a full classification scan on your data sources. Data sources that have had a mapping-only scan can only generate the Data Mapping Report.
- NetApp can't guarantee 100% accuracy of the personal data and sensitive personal data that BlueXP classification identifies. You should always validate the information by reviewing the data.

The following reports are available for BlueXP classification:

- **Data Discovery Assessment report:** Provides a high-level analysis of the scanned environment to highlight the system's findings and to show areas of concern and potential remediation steps.
- **Data Mapping report:** Provides information about the size and number of files in your working environments. This includes usage capacity, age of data, size of data, and file types.
- **Data Subject Access Request report:** Enables you to extract a report of all files that contain information regarding a data subject's specific name or personal identifier.
- **HIPAA report:** Helps you identify the distribution of health information across your files.
- **PCI DSS report:** Helps you identify the distribution of credit card information across your files.
- **Privacy Risk Assessment report:** Provides privacy insights from your data and a privacy risk score.
- **Reports on a specific information type:** Reports are available that include details about the identified files that contain personal data and sensitive personal data. You can also see files broken down by

category and file type.

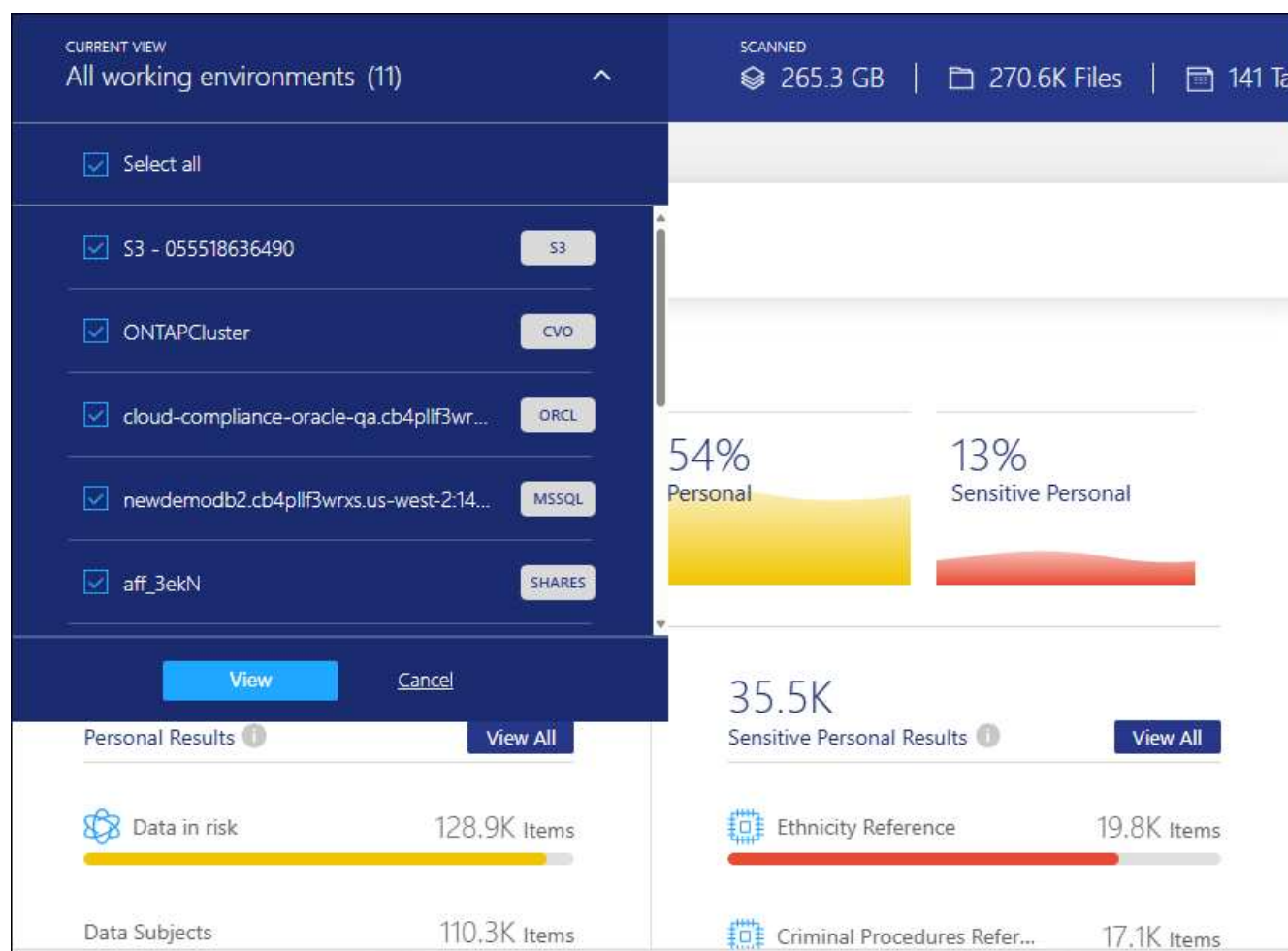
Select the working environments for reports

You can filter the contents of the BlueXP classification Compliance dashboard to see compliance data for all working environments and databases, or for just specific working environments.

When you filter the dashboard, BlueXP classification scopes the compliance data and reports to just those working environments that you selected.

Steps

1. From the BlueXP classification menu, select "**Compliance.**"
2. Select the Working environments filter drop-down and select the working environments.
3. Select **View**.



Data Subject Access Request Report

Privacy regulations such as the European GDPR grant data subjects (such as customers or employees) the right to access their personal data. When a data subject requests this information, this is known as a DSAR (data subject access request). Organizations are required to respond to these requests "without undue delay", and at the latest within one month of receipt.

You can respond to a DSAR by searching for a subject's full name or known identifier (such as an email

address) and then downloading a report. The report is designed to aid in your organization's requirement to comply with GDPR or similar data privacy laws.

How can BlueXP classification help you respond to a DSAR?

When you perform a data subject search, BlueXP classification finds all of the files, buckets, OneDrive, and SharePoint accounts that have that person's name or identifier in it. BlueXP classification checks the latest pre-indexed data for the name or identifier. It doesn't initiate a new scan.

After the search is complete, you can then download the list of files for a Data Subject Access Request report. The report aggregates insights from the data and puts it into legal terms that you can send back to the person.



Data subject search is not supported within databases at this time.

Search for data subjects and download reports

Search for the data subject's full name or known identifier and then download a file list report or DSAR report. You can search by [any personal information type](#).

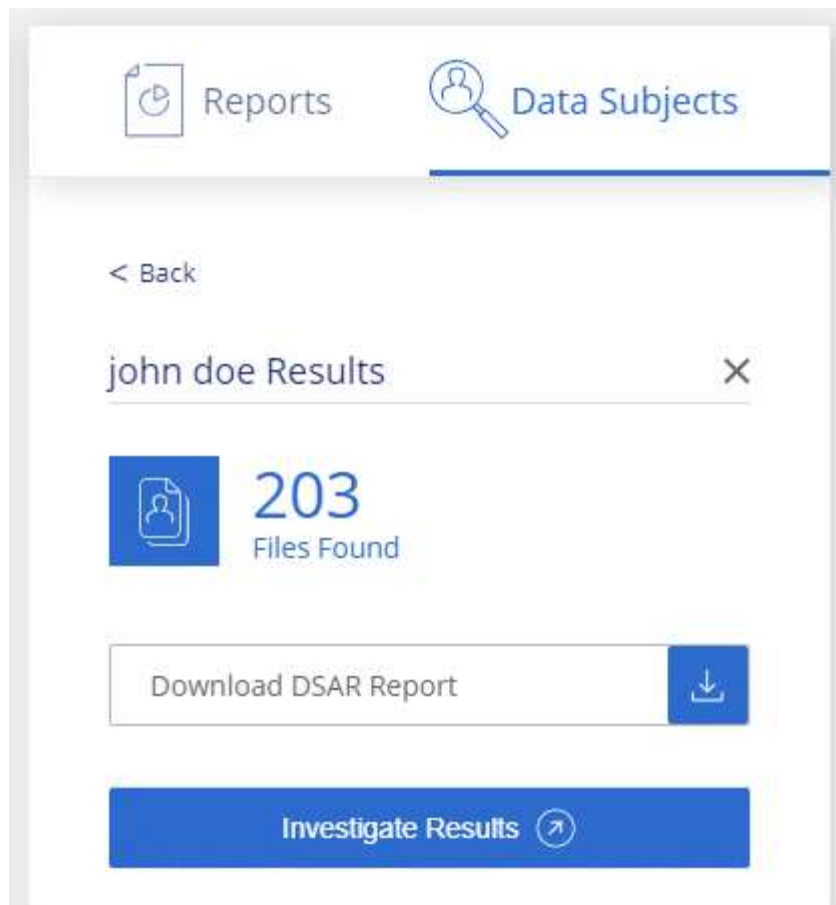


English, German, Japanese, and Spanish are supported when searching for the names of data subjects. Support for more languages will be added later.

Steps

1. From the BlueXP classification menu, select "**Compliance.**"
2. From the Compliance page, scroll down and select **Data Subjects**.
3. Search for the data subject's full name or known identifier.

Here's an example that shows a search for the name *john doe*:



4. Choose one of the available options:

- **Download DSAR Report:** A formal response to the access request that you can send to the data subject. This report contains automatically-generated information based on data that BlueXP classification found on the data subject and is designed to be used as a template. You should complete the form and review it internally before sending it to the data subject.
- **Investigate Results:** A page that enables you to investigate the data by searching, sorting, expanding details for a specific file, and by downloading the file list.



If there are more than 10,000 results, only the top 10,000 appear in the file list.

Health Insurance Portability and Accountability Act (HIPAA) Report

The Health Insurance Portability and Accountability Act (HIPAA) Report can help you identify files containing health information. It is designed to aid in your organization's requirement to comply with HIPAA data privacy laws. The information BlueXP classification looks for includes:

- Health reference pattern
- ICD-10-CM Medical code
- ICD-9-CM Medical code
- HR - Health category
- Health Application Data category

The report includes the following information:

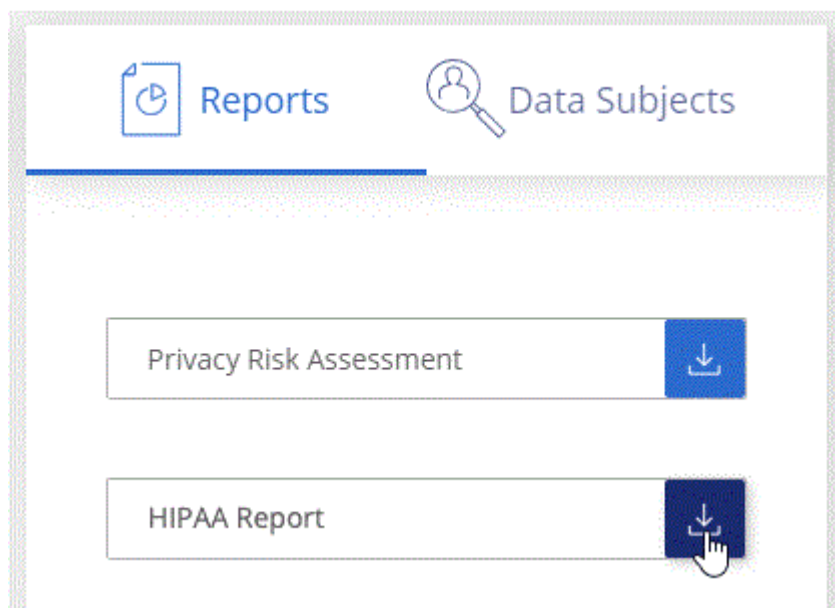
- Overview: How many files contain health information and in which working environments.
- Encryption: The percentage of files containing health information that are on encrypted or unencrypted working environments. This information is specific to Cloud Volumes ONTAP.
- Ransomware Protection: The percentage of files containing health information that are on working environments that do or don't have ransomware protection enabled. This information is specific to Cloud Volumes ONTAP.
- Retention: The timeframe in which the files were last modified. This is helpful because you shouldn't keep health information for longer than you need to process it.
- Distribution of Health Information: The working environments where the health information was found and whether encryption and ransomware protection are enabled.

Generate the HIPAA Report

Go to the Compliance tab to generate the report.

Steps

1. From the BlueXP classification menu, select "**Compliance.**"
2. Scroll down and locate the **Reports** pane.
3. Select the download icon next to **HIPAA Report**.



Result

BlueXP classification generates a PDF report that you can review and send to other groups as needed.

Payment Card Industry Data Security Standard (PCI DSS) Report

The Payment Card Industry Data Security Standard (PCI DSS) Report can help you identify the distribution of credit card information across your files.

The report includes the following information:

- Overview: How many files contain credit card information and in which working environments.

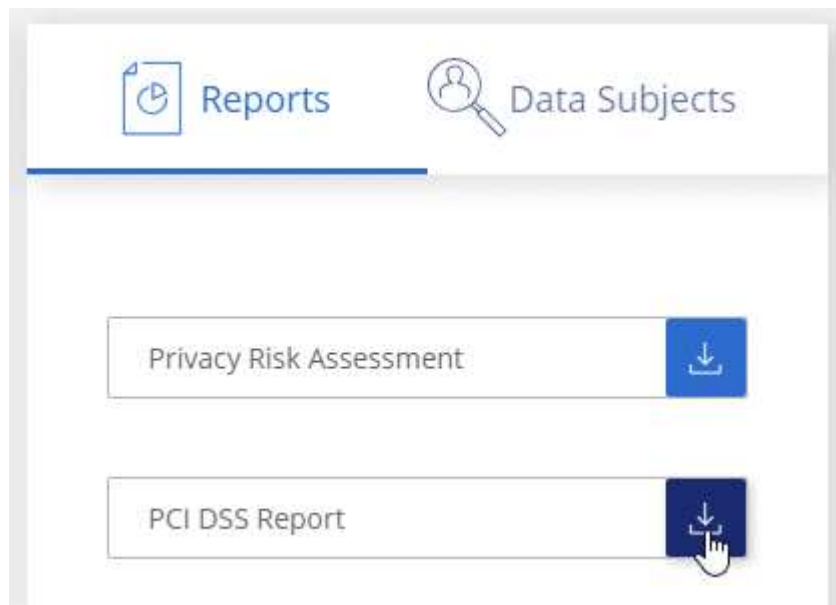
- **Encryption:** The percentage of files containing credit card information that are on encrypted or unencrypted working environments. This information is specific to Cloud Volumes ONTAP.
- **Ransomware Protection:** The percentage of files containing credit card information that are on working environments that do or don't have ransomware protection enabled. This information is specific to Cloud Volumes ONTAP.
- **Retention:** The timeframe in which the files were last modified. This is helpful because you shouldn't keep credit card information for longer than you need to process it.
- **Distribution of Credit Card Information:** The working environments where the credit card information was found and whether encryption and ransomware protection are enabled.

Generate the PCI DSS Report

Go to the Compliance tab to generate the report.

Steps

1. From the BlueXP classification menu, select "**Compliance.**"
2. Scroll down and locate the **Reports** pane.
3. Select the download icon next to **PCI DSS Report**.



Result

BlueXP classification generates a PDF report that you can review and send to other groups as needed.

Privacy Risk Assessment Report

The Privacy Risk Assessment Report provides an overview of your organization's privacy risk status, as required by privacy regulations such as GDPR and CCPA.

The report includes the following information:

- **Compliance status:** A severity score and the distribution of data, whether it's non-sensitive, personal, or sensitive personal.
- **Assessment overview:** A breakdown of the types of personal data found, as well as the categories of data.

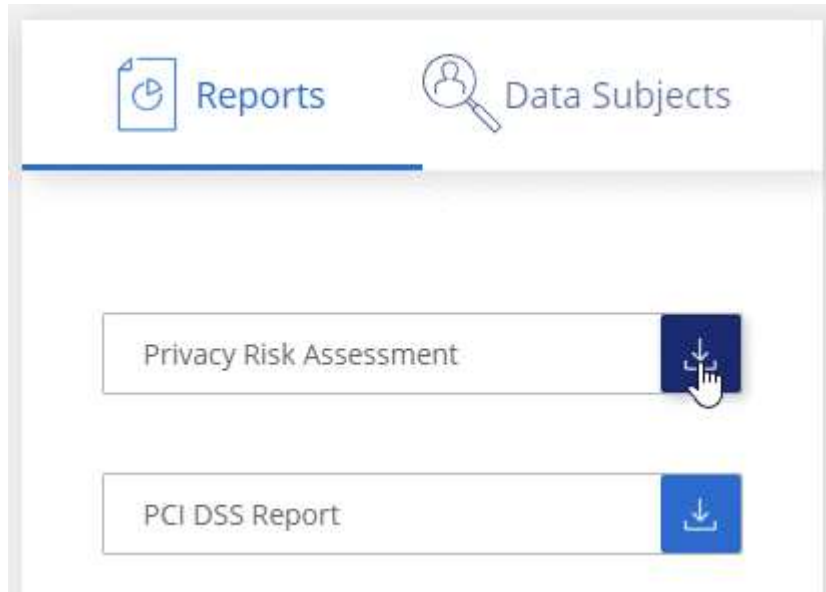
- Data subjects in this assessment: The number of people, by location, for which national identifiers were found.

Generate the Privacy Risk Assessment Report

Go to the Compliance tab to generate the report.

Steps

1. From the BlueXP classification menu, select "**Compliance.**"
2. Scroll down and locate the **Reports** pane.
3. Select the download icon next to **Privacy Risk Assessment**.



Result

BlueXP classification generates a PDF report that you can review and send to other groups as needed.

Severity score

BlueXP classification calculates the severity score for the Privacy Risk Assessment Report on the basis of three variables:

- The percentage of personal data out of all data.
- The percentage of sensitive personal data out of all data.
- The percentage of files that include data subjects, determined by national identifiers such as national IDs, Social Security numbers, and tax ID numbers.

The logic used to determine the score is as follows:

Severity score	Logic
0	All three variables are exactly 0%
1	One of the variables are larger than 0%
2	One of the variables are larger than 3%

Severity score	Logic
3	Two of the variables are larger than 3%
4	Three of the variables are larger than 3%
5	One of the variables are larger than 6%
6	Two of the variables are larger than 6%
7	Three of the variables are larger than 6%
8	One of the variables are larger than 15%
9	Two of the variables are larger than 15%
10	Three of the variables are larger than 15%

Copyright information

Copyright © 2025 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

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

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. 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.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.