



# **Manage BlueXP classification**

## **BlueXP classification**

NetApp  
April 03, 2024

# Table of Contents

- Manage BlueXP classification ..... 1
  - Add personal data identifiers to your BlueXP classification scans..... 1
  - Exclude specific directories from BlueXP classification scans..... 16
  - Viewing the status of your compliance actions ..... 19
  - Define additional group IDs as open to organization ..... 20
  - Audit the history of BlueXP classification actions..... 21
  - Reducing the BlueXP classification scan speed ..... 22
  - Removing data sources from BlueXP classification..... 23
  - Uninstalling BlueXP classification ..... 25

# Manage BlueXP classification

## Add personal data identifiers to your BlueXP classification scans

BlueXP classification provides many ways for you to add a custom list of "personal data" that BlueXP classification will identify in future scans, giving you the full picture about where potentially sensitive data resides in *all* your organizations' files.

- You can add unique identifiers based on specific columns in databases you are scanning.
- You can add custom keywords from a text file — these words are identified within your data.
- You can add a personal pattern using a regular expression (regex) — the regex is added to the existing predefined patterns.
- You can add custom categories to identify where specific categories of information are found in your data.

All of these mechanisms to add custom scanning criteria are supported in all languages.



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.

## Add custom personal data identifiers from your databases

A feature we call *Data Fusion* allows you to scan your organizations' data to identify whether unique identifiers from your databases are found in any of your other data sources. You can choose the additional identifiers that BlueXP classification will look for in its' scans by selecting a specific column, or columns, in a database table. For example, the diagram below shows how data fusion is used to scan your volumes, buckets, and databases for occurrences of all your Customer IDs from your Oracle database.

## Databases -- Structured Data

Database: Oracle  
Schema: Accounts  
Table: Customers  
Column: Customer ID

Account	Name	Customer ID	Address
1234	ABC Co	135876	125 Main St
1235	XYZ Co	213536	35A Brick R
1236	Cat Co	359264	55 Wind Av
1237	Dog Co	472637	11025 Cor
1238	Zebra Co	582455	36 Sahara
...	...	...	...

*Scan your volumes and buckets for occurrences of the Customer IDs in your Oracle database*

## Files -- Unstructured Data

File in Volume 1

```
XXXXXXXXXXXXX
xx213536xxx
XXXXXXXXXXXXX
xx472637xxx
XXXXXXXXXXXXX
XXXXXXXXXXXXX
```

File in Volume 2

```
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
xxx472637xx
```

File in Bucket 1

```
XXXXXXXXXXXXX
XXXXXXXXXXXXX
xx213536xxx
XXXXXXXXXXXXX
XXXXXXXXXXXXX
XXXXXXXXXXXXX
```

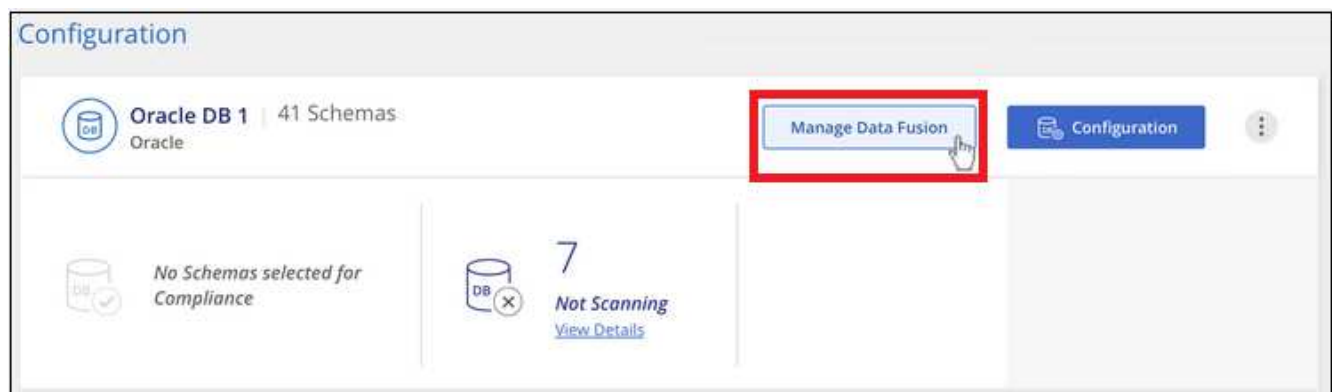
As you can see, two unique Customer IDs have been found in two volumes and in one S3 bucket. Any matches in database tables will also be identified.

Note that since you're scanning your own databases, whatever language your data is stored in will be used to identify data in future BlueXP classification scans.

### Steps

You must have [added at least one database server](#) to BlueXP classification before you can add data fusion sources.

1. In the Configuration page, click **Manage Data Fusion** in the database where the source data resides.



2. Click **Add Data Fusion source** on the next page.
3. In the *Add Data Fusion Source* page:
  - a. Select the Database Schema from the drop-down menu.

- b. Enter the Table name in that schema.
- c. Enter the Column, or Columns, that contain the unique identifiers you want to use.

When adding multiple columns, enter each column name, or table view name, on a separate line.

#### 4. Click **Add Data Fusion Source**.

Oracle DB 1 Data Fusion			<a href="#">+ Add Data Fusion source</a>
With Data Fusion, Data Sense can identify occurrences of your organization's unique identifiers found in your unstructured data stores, using structured data indexes containing those unique identifiers as a source reference. <a href="#">Learn More</a>			
Database Schema	Table	Data Fusion Source Columns	
Schema1	Table 1	Column 12, Column 4, Column 18	...
Schema2	Table 2	Column 2, Column 14, Column 8	...

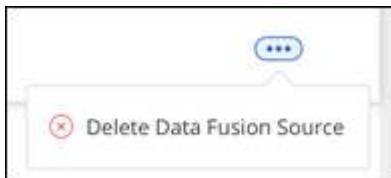
### Results

After the next scan, the results will include this new information in the Compliance Dashboard under the "Personal Results" section, and in the Investigation page in the "Personal Data" filter. The name you used for the classifier appears in the filter list, for example `Customers.CustomerID`.

Personal Results			
30 Types   96.6K Items found in All working environments			
Email Address	92K Items	IBAN	6.7K Items
Internal Product ID	6 Items	<b>Customers.CustomerID</b>	56 Items
Estonian ID	5 Items	French SPI	5 Items

### Delete a Data Fusion source

If at some point you decide not to scan your files using a certain Data Fusion source, you can select the source row from the Data Fusion inventory page and click **Delete Data Fusion Source**.



### Add custom keywords from a list of words

You can add custom keywords to BlueXP classification so that it will identify where that information is found in your data. You add the keywords just by entering each word you want BlueXP classification to recognize. The

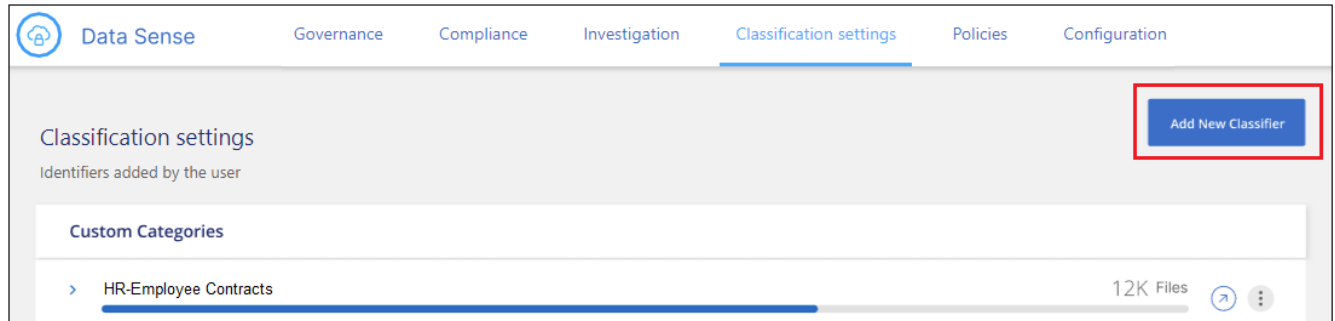
keywords are added to the existing predefined keywords that BlueXP classification already uses, and the results will be visible under the personal patterns section.

For example, you may want to see where internal Product Names are mentioned in all of your files to make sure these names are not accessible in locations that are not secure.

After updating the custom keywords, BlueXP classification will restart scanning all data sources. After the scan has completed, the new results will appear in the BlueXP classification Compliance Dashboard under the "Personal Results" section, and in the Investigation page in the "Personal Data" filter.

## Steps

1. From the *Classification settings* tab, click **Add New Classifier** to launch the *Add Custom Classifier* wizard.



2. In the *Select type* page, enter the name of the classifier, provide a brief description, select **Personal identifier**, and then click **Next**.

The name you enter will appear in the BlueXP classification UI as the heading for scanned files that match the classifier requirements, and as the name of the filter in the Investigation page.

You can also check the box to "Mask detected results in the system" so the full result won't appear in the UI. For example, you may want to do this to hide full credit card numbers or similar personal data (the mask would appear in the UI like this: "\*\*\*\* \* 3434").

1 Select type

2 Select tool

3 Create Logic

## Select type

Select the type of classifier that you want to add to the system, and provide the name and description. Data Sense re-scans all your data sources after you add a new classifier. When the scan is complete, all matching results are displayed in the "Classification Settings" dashboard and in other Data Sense pages.

---

Classifier name

Internal Product Names

Description

Identify internal product names found in all files

☒ **Personal identifier**

The classifier will be added to the system as a new personal identifier. Any matches are considered "personal data", and they are added to the results that are displayed in the Personal Results page and in the Investigation page. [See the list of personal data that Data Sense identifies by default.](#)

☐ Mask detected results in the system

☐ **Category**

The classifier will be added to the system as a new Category. Any matches are added to the results that are displayed in the Categories page and in the Investigation page. [See the list of categories that Data Sense identifies by default.](#)

Previous

Next

- In the *Select Data Analysis Tool* page, select **Custom keywords** as the method you want to use to define the classifier, and then click **Next**.

## Select Data Analysis Tool

Select the tool that will be used to build the list of words, or patterns, that Data Sense will attempt to match in your data sources.

☒

**Custom keywords** ⓘ  
Create a custom personal pattern based on a list of keywords that you provide.

☐

**Custom regular expression** ⓘ  
Create a custom personal pattern based on a regular expression that you define.

☐

**DB fusion** ⓘ  
Create a custom personal pattern based on the values found in specific columns in your scanned database tables. This allows you to identify whether unique identifiers from your databases are found in any of your other data sources.

Previous

Next

4. In the *Create Logic* page, enter the keywords you want to recognize - each word on a separate line - and click **Validate**.

The screenshot below shows internal Product Names (different types of owls). The BlueXP classification search for these items is not case sensitive.



## Create Logic

Create logic for the new identifier, based on regular expression and keywords that should be detected. You will be able to change the logic in the future, by clicking on "edit" from the custom classification dashboard.

---

### Custom keywords list <sup>1</sup>

- Maximum of 100,000 words.
- Separate between keywords with a new line
- The keywords are not case sensitive
- Each word must be at least 3 characters long. Shorter words are ignored.
- Duplicate words are only added once.

barred  
barn  
horned  
snowy  
screech

Validate

✔ Keywords list is valid.

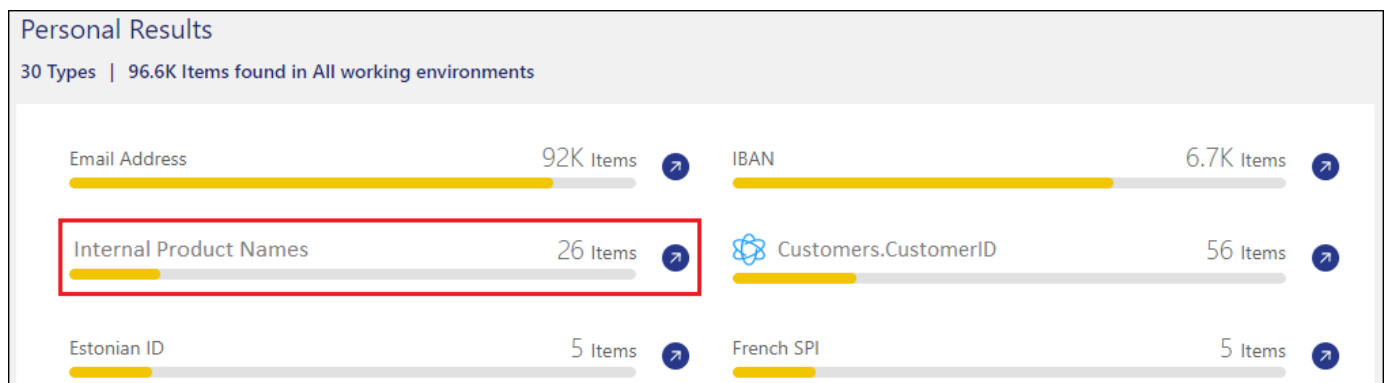
Previous

Done

5. Click **Done** and BlueXP classification starts to rescan your data.

## Results

After the scan is complete, the results will include this new information in the Compliance Dashboard under the "Personal Results" section, and in the Investigation page in the "Personal Data" filter.



As you can see, the name of the classifier is used as the name in the Personal Results panel. In this manner you can activate many different groups of keywords and see the results for each group.

## Add custom personal data identifiers using a regex

You can add a personal pattern to identify specific information in your data using a custom regular expression (regex). This allows you to create a new custom regex to identify new personal information elements that don't yet exist in the system. The regex is added to the existing predefined patterns that BlueXP classification

already uses, and the results will be visible under the personal patterns section.

For example, you may want to see where your internal Product IDs are mentioned in all of your files. If the Product ID has a clear structure, for example, it is a 12-digit number that starts with 201, you can use the custom regex feature to search for it in your files. The regular expression for this example is `\b201\d{9}\b`.

After adding the regex, BlueXP classification will restart scanning all data sources. After the scan has completed, the new results will appear in the BlueXP classification Compliance Dashboard under the "Personal Results" section, and in the Investigation page in the "Personal Data" filter.

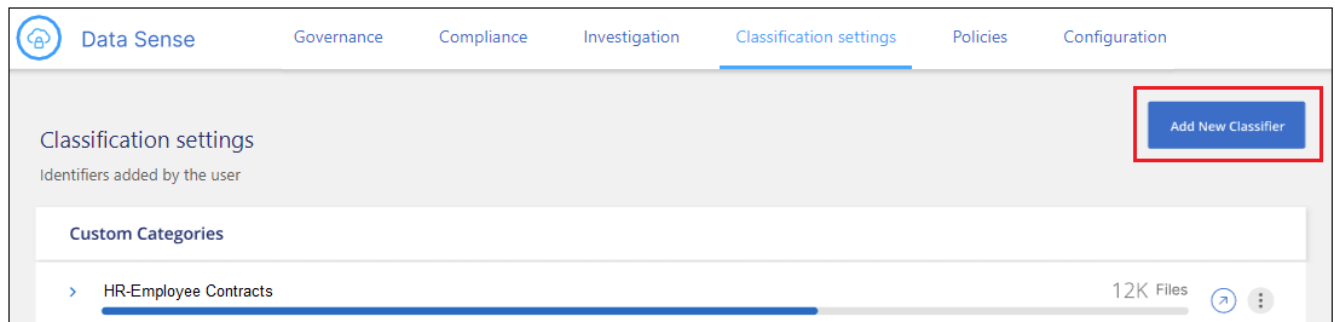
If you need assistance in building the regular expression, refer to [Regular expressions 101](#). Choose **Python** for the Flavor to see the types of results BlueXP classification will match from the regular expression. The [Python Regex Tester page](#) is also useful by displaying a graphical representation of your patterns.



Currently we do not allow the use of pattern flags when creating a regex - this means you should not use `/`.

## Steps

1. From the *Classification settings* tab, click **Add New Classifier** to launch the *Add Custom Classifier* wizard.



2. In the *Select type* page, enter the name of the classifier, provide a brief description, select **Personal identifier**, and then click **Next**.

The name you enter will appear in the BlueXP classification UI as the heading for scanned files that match the classifier requirements, and as the name of the filter in the Investigation page. You can also check the box to "Mask detected results in the system" so the full result won't appear in the UI. For example, you may want to do this to hide full credit card numbers or similar personal data.

1 Select type

2 Select tool

3 Create Logic

## Select type

Select the type of classifier that you want to add to the system, and provide the name and description. Data Sense re-scans all your data sources after you add a new classifier. When the scan is complete, all matching results are displayed in the "Classification Settings" dashboard and in other Data Sense pages.

---

Classifier name

Internal Product ID

Description

Identify internal product IDs found in all files

☒ **Personal identifier**

The classifier will be added to the system as a new personal identifier. Any matches are considered "personal data", and they are added to the results that are displayed in the Personal Results page and in the Investigation page. [See the list of personal data that Data Sense identifies by default.](#)

☐ Mask detected results in the system

☐ **Category**

The classifier will be added to the system as a new Category. Any matches are added to the results that are displayed in the Categories page and in the Investigation page. [See the list of categories that Data Sense identifies by default.](#)

Previous

Next

- In the *Select Data Analysis Tool* page, select **Custom regular expression** as the method you want to use to define the classifier, and then click **Next**.

## Select Data Analysis Tool

Select the tool that will be used to build the list of words, or patterns, that Data Sense will attempt to match in your data sources.

☐

**Custom keywords** ⓘ  
Create a custom personal pattern based on a list of keywords that you provide.

☒

**Custom regular expression** ⓘ  
Create a custom personal pattern based on a regular expression that you define.

☐

**DB fusion** ⓘ  
Create a custom personal pattern based on the values found in specific columns in your scanned database tables. This allows you to identify whether unique identifiers from your databases are found in any of your other data sources.

Previous

Next

4. In the *Create Logic* page, enter the regular expression and any proximity words, and click **Done**.
  - a. You can enter any legal regular expression. Click the **Validate** button to have BlueXP classification verify that the regular expression is valid, and that it is not too broad — meaning it will return too many results.
  - b. Optionally, you can enter some proximity words to help refine the accuracy of the results. These are words that will typically be found within 300 characters of the pattern you are searching for (either before or after the found pattern). Enter each word, or phrase, on a separate line.

## Create Logic

Create logic for the new identifier, based on regular expression and keywords that should be detected.

---

**Regular expression** ⓘ

Add the pattern that should be detected to identify specific information in your data, using a custom regular expression.

✓ **Success:** Regular expression is valid.

☒ **Proximity words** - To improve the detection accuracy, insert phrases that must appear near by the regular expression's match.

### Results

The classifier is added and BlueXP classification starts to rescan all your data sources. You are returned to the Custom Classifiers page where you can view the number of files that have matched your new classifier. Results from scanning all of your data sources will take some time depending on the number of files that need to be scanned.

[Data Sense](#) [Governance](#) [Compliance](#) [Investigation](#) [Classification settings](#) [Policies](#) [Configuration](#)

### Classification settings

Add New Classifier

Identifiers added by the user

#### Custom Categories

> HR - Employee Contracts 7.5K Files

#### Personal information

> Internal Product ID 12K Files

### Add custom categories

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

## predefined 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* may include sensitive data. When you investigate the results, you might find that employee contracts are stored in an insecure location. You can then correct that issue.

You can add custom categories to BlueXP classification so you can identify where categories of information that are unique for your data estate are found in your data. You add each category by creating "training" files that contain the categories of data that you want to identify, and then have BlueXP classification scan those files to "learn" through AI so that it can identify that data in your data sources. The categories are added to the existing predefined categories that BlueXP classification already identifies, and the results are visible under the Categories section.

For example, you may want to see where compressed installation files in .gz format are located in your files so that you can remove them, if necessary.

After updating the custom categories, BlueXP classification will restart scanning all data sources. After the scan has completed, the new results will appear in the BlueXP classification Compliance Dashboard under the "Categories" section, and in the Investigation page in the "Category" filter. [See how to view files by categories.](#)

### What you'll need

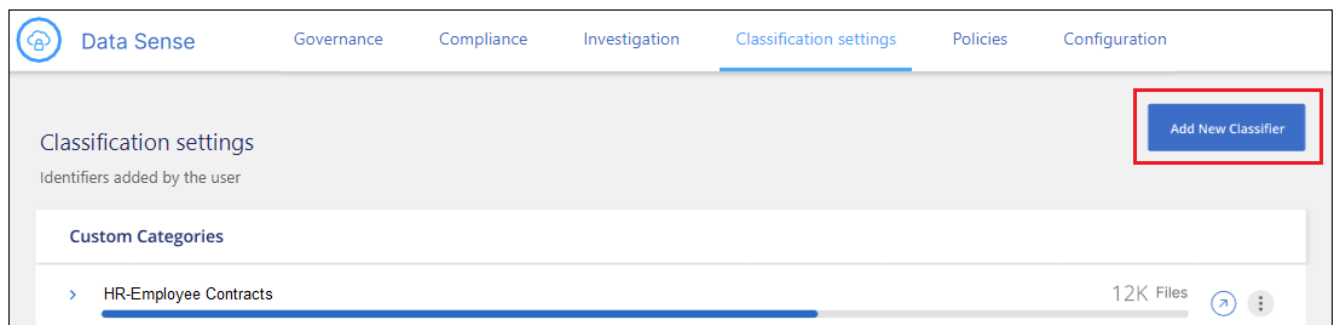
You'll need to create a minimum of 25 training files that contain samples of the categories of data that you want BlueXP classification to recognize. The following file types are supported:

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

The files must be a minimum of 100 bytes, and they must be located in a folder that is accessible by BlueXP classification.

### Steps

1. From the *Classification settings* tab, click **Add New Classifier** to launch the *Add Custom Classifier* wizard.



2. In the *Select type* page, enter the name of the classifier, provide a brief description, select **Category**, and then click **Next**.

The name you enter will appear in the BlueXP classification UI as the heading for scanned files that match the category of data you are defining, and as the name of the filter in the Investigation page.

1 Select type
2 Select tool
3 Create Logic

## Select type

Select the type of classifier that you want to add to the system, and provide the name and description. Data Sense re-scans all your data sources after you add a new classifier. When the scan is complete, all matching results are displayed in the "Classification Settings" dashboard and in other Data Sense pages.

Classifier name

Description

☐ **Personal identifier**  
The classifier will be added to the system as a new personal identifier. Any matches are considered "personal data", and they are added to the results that are displayed in the Personal Results page and in the Investigation page. [See the list of personal data that Data Sense identifies by default.](#)  
☐ Mask detected results in the system

☒ **Category**  
The classifier will be added to the system as a new Category. Any matches are added to the results that are displayed in the Categories page and in the Investigation page. [See the list of categories that Data Sense identifies by default.](#)

Previous
Next

- In the *Create Logic* page, make sure you have the learning files prepared, and then click **Select files**.

## Create Logic

**AI-based similarity training**

- Insert NFS folder path
- The folder should contain minimum 25 files and maximum 1000 files that will be used for the AI training.
- Supported file types: pdf, docx, doc, pptx, xls, xlsx, csv, txt, gz, rtf, docs, sheets, slides, json
- The keywords are not case sensitive
- Minimum file size: 100B

Compressed Installer files

- Enter the IP address of the volume, and the path where the training files are located, and click **Add**.

### Insert folder path that contains at least 25 files for the training

Enter the IP address and volume name, along with the path to the location of the training files.

IP

Training Data - Folder path

XXX.XXX.XXX.XXX:/VolumeName

folder/path/

Add

Cancel

- Verify that the training files were recognized by BlueXP classification. Click the **x** to remove any training files that do not meet the requirements. Then click **Done**.

### Create Logic

#### AI-based similarity training

- Insert NFS folder path
- The folder should contain minimum 25 files and maximum 1000 files that will be used for the AI training.
- Supported file types: pdf, docx, doc, pptx, xls, xlsx, csv, txt, gz, rtf, docs, sheets, slides, json
- The keywords are not case sensitive
- Minimum file size: 100B

Select Files

#### Compressed Installer files

Total uploaded files: 54

File name	File Size	File Type	Reliability	included in training
File1	56	File type	Sufficient	x
File2	22	File type	Sufficient	x
File3	43	File type	Sufficient	x
File4	11	File type	Sufficient	x

Previous

Done

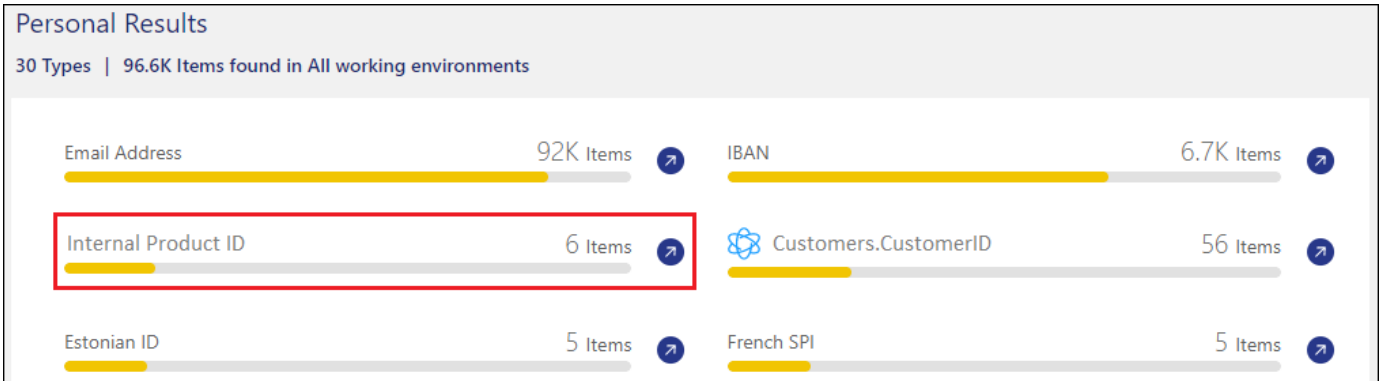
## Results

The new category is created as defined by the training files and added to BlueXP classification. Then BlueXP classification starts to rescan all your data sources to identify files that fit into this new category. You are returned to the Custom Classifiers page where you can view the number of files that have matched your new category. Results from scanning all of your data sources will take some time depending on the number of files that need to be scanned.

## View results from your custom classifiers

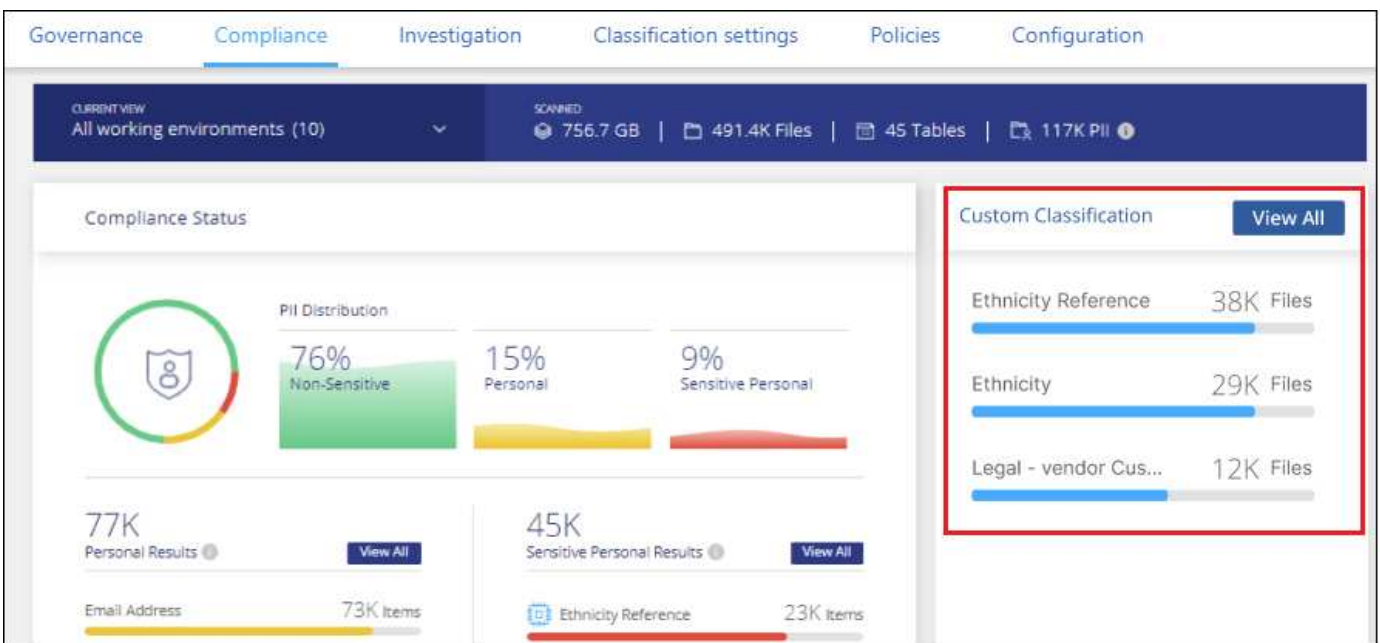
You can view the results from any of your custom classifiers in the Compliance Dashboard and in the Investigation page. For example, this screenshot shows the matched information in the Compliance Dashboard under the "Personal Results" section.





Click the  button to see the detailed results in the Investigation page.

Additionally, all of your custom classifier results appear in the Custom Classifiers tab, and the top 6 custom classifier results are displayed in the Compliance Dashboard, as shown below.



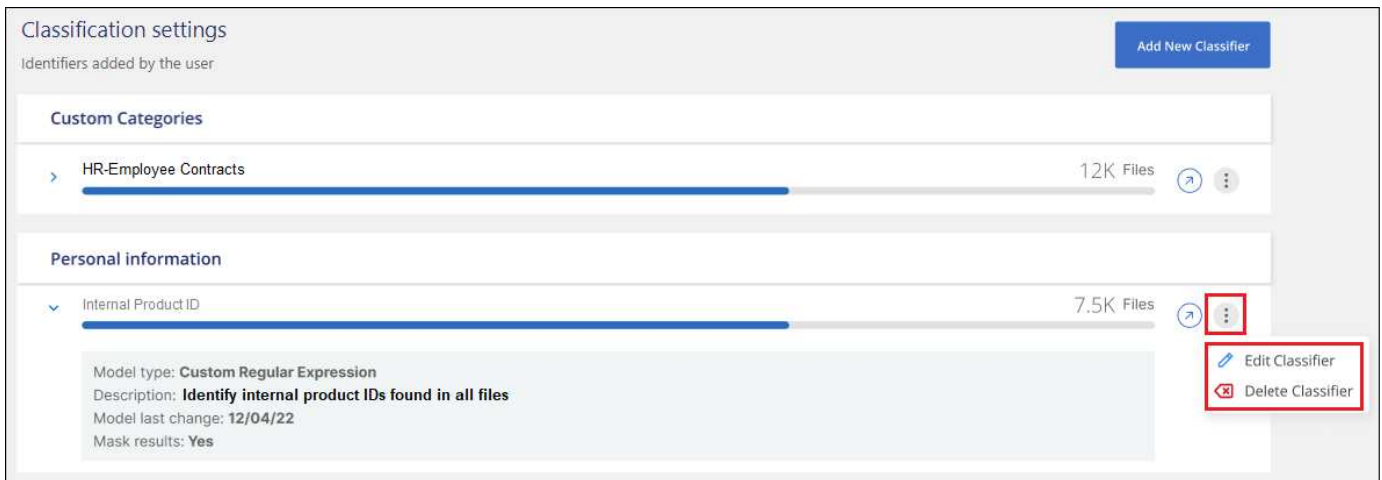
## Manage custom classifiers

You can change any of the custom classifiers that you have created by using the **Edit Classifier** button.



You can't edit Data Fusion classifiers at this time.

And if you decide at some later point that you don't need BlueXP classification to identify the custom patterns that you added, you can use the **Delete Classifier** button to remove each item.



## Exclude specific directories from BlueXP classification scans

If you want BlueXP classification to exclude scanning data that resides in certain data source directories, you can add these directory names to a configuration file. After you apply this change, the BlueXP classification engine will exclude scanning data in those directories.

Note that BlueXP classification is configured by default to exclude scanning volume snapshot data because that content is identical to the content in the volume.

This functionality is available in BlueXP classification version 1.29 and greater (starting in March 2024).

### Supported data sources

Excluding specific directories from BlueXP classification scans is supported for NFS and CIFS shares in the following data sources:

- On-premises ONTAP
- Cloud Volumes ONTAP
- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- General file shares

### Define the directories to exclude from scanning

Before you can exclude directories from classification scanning, you need to log into the BlueXP classification system so you can edit a configuration file and run a script. See how to [log in to the BlueXP classification system](#) depending on whether you manually installed the software on a Linux machine or if you deployed the instance in the cloud.



- You can exclude a maximum of 50 directory paths per BlueXP classification system.
- Excluding directory paths may affect scanning times.

## Steps

1. On the BlueXP classification system, go to `/opt/netapp/config/custom_configuration` and open the file `data_provider.yaml`.
2. In the `"data_providers"` section, under the line `"exclude:"`, enter the directory paths to exclude. For example:

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

Do not change anything else in this file.

3. Save the changes to the file.
4. Go to `/opt/netapp/Datasense/tools/customer_configuration/data_providers` and run the following script:

```
update_data_providers_from_config_file.sh
```

This command commits the directories to be excluded from scanning to the classification engine.

## Result

All subsequent scans of your data will exclude scanning of those specified directories.

You can add, edit, or delete items from the exclude list using these same steps. The revised exclude list will be updated after you run the script to commit your changes.

## Examples

### Configuration 1:

Every folder that contains `"folder1"` anywhere in the name will be excluded from all data sources.

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

### Expected results for paths that will be excluded:

- `/CVO1/folder1`
- `/CVO1/folder1name`
- `/CVO1/folder10`
- `/CVO1/*folder1`
- `/CVO1/+folder1name`
- `/CVO1/notfolder10`
- `/CVO22/folder1`

- /CVO22/folder1name
- /CVO22/folder10

**Examples for paths that will not be excluded:**

- /CVO1/\*folder
- /CVO1/foldername
- /CVO22/\*folder20

**Configuration 2:**

Every folder that contains "\*\*folder1" only at the start of the name will be excluded.

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

**Expected results for paths that will be excluded:**

- /CVO/\*folder1
- /CVO/\*folder1name
- /CVO/\*folder10

**Examples for paths that will not be excluded:**

- /CVO/folder1
- /CVO/folder1name
- /CVO/not\*folder10

**Configuration 3:**

Every folder in data source "CVO22" that contains "folder1" anywhere in the name will be excluded.

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

**Expected results for paths that will be excluded:**

- /CVO22/folder1
- /CVO22/folder1name
- /CVO22/folder10

**Examples for paths that will not be excluded:**

- /CVO1/folder1
- /CVO1/folder1name
- /CVO1/folder10

## Escaping special characters in folder names

If you have a folder name that contains one of the following special characters and you want to exclude data in that folder from being scanned, you'll need to use the escape sequence `\\` before the folder name.

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

For example:

Path in source: `/project/*not_to_scan`

Syntax in exclude file: `"\\*not_to_scan"`

## View the current exclusion list

It's possible for the contents of the `data_provider.yaml` configuration file to be different than what has actually been committed after running the `update_data_providers_from_config_file.sh` script. To view the current list of directories that you've excluded from BlueXP classification scanning, run the following command from `/opt/netapp/Datasense/tools/customer_configuration/data_providers`:

```
get_data_providers_configuration.sh
```

## Viewing the status of your compliance actions

When you run an asynchronous action from the Investigation Results pane across many files, for example, moving or deleting 100 files, the process can take some time. You can monitor the status of these actions in the *Action Status* pane so you'll know when it has been applied to all files.

This allows you to see the actions that have completed successfully, those currently in progress, and those that have failed so you can diagnose and fix any problems. Note that short operations that complete quickly, such as moving a single file, do not appear in the Actions Status pane.

The status can be:

- Success - A BlueXP classification action is finished and all items succeeded.
- Partial Success - A BlueXP classification action is finished and some items failed and some succeeded.
- In Progress - The action is still in progress.
- Queued - The action has not started.
- Canceled - The action has been canceled.
- Failed - The action failed.

Note that you can Cancel any actions that have the "Queued" or "In Progress" status.

### Steps

- 1.



In the bottom-right of the BlueXP classification UI you can see the **Actions Status** button

2. Click this button and the most recent 20 actions are listed.

You can click the name of an action to view details corresponding to that operation.

## Define additional group IDs as open to organization

When group IDs (GIDs) are attached to files or folders in NFS file shares they define the permissions for the file or folder; such as whether they are "open to the organization". If some group IDs (GIDs) are not initially set up with the "Open to Organization" permission level, you can add that permission to the GID so that any files and folders that have that GID attached will be deemed "open to the organization".

After you make this change and BlueXP classification rescans your files and folders, any files and folders that have these group IDs attached will show this permission in the Investigation Details page, and they'll also appear in reports where you are displaying file permissions.

To activate this functionality, you need to log into the BlueXP classification system so you can edit a configuration file and run a script. See how to [log in to the BlueXP classification system](#) depending on whether you manually installed the software on a Linux machine or if you deployed the instance in the cloud.

### Add the "open to organization" permission to group IDs

You need to have the group ID numbers (GIDs) before starting this task.

#### Steps

1. On the BlueXP classification system, go to `/opt/netapp/config/custom_configuration` and open the file `data_provider.yaml`.
2. In the line `"organization_group_ids: []"` add the group IDs. For example:

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

Do not change anything else in this file.

3. Save the changes to the file.
4. Go to `/opt/netapp/Datasense/tools/customer_configuration/data_providers` and run the following script:

```
update_data_providers_from_config_file.sh
```

This command commits the revised group ID permissions to the classification engine.

#### Result

All subsequent scans of your data will identify files or folders that have these group IDs attached as "open to

organization".

You can edit the list of group IDs and delete any group IDs you added in the past using these same steps. The revised list of group IDs will be updated after you run the script to commit your changes.

## View the current list of group IDs

It's possible for the contents of the `data_provider.yaml` configuration file to be different than what has actually been committed after running the `update_data_providers_from_config_file.sh` script. To view the current list of group IDs that you've added to BlueXP classification, run the following command from `/opt/netapp/Datasense/tools/customer_configuration/data_providers`:

```
get_data_providers_configuration.sh
```

## Audit the history of BlueXP classification actions

BlueXP classification logs management activities that have been performed on files from all the working environments and data sources that BlueXP classification is scanning. BlueXP classification also logs the activities when deploying the BlueXP classification instance.

You can view the contents of the BlueXP classification audit log files, or download them, to see what file changes have occurred, and when. For example, you can see what request was issued, the time of the request, and details such as source location in case a file was deleted, or source and destination location in case a file was moved.

### Log file contents

Each line in the audit log contains information in this format:

```
<full date> | <status> | ds_audit_logger | <module> | 0 | 0 | File <full file path> deleted from device <device path> - <result>
```

- Date and time - full timestamp for the event
- Status - INFO, WARNING
- Action type (delete, copy, move, create policy, update policy, rescan files, download JSON report, etc.)
- File name (if the action is relevant to a file)
- Details for the action - what was done: depends on the action
  - Policy name
  - For move - Source and destination
  - For copy - Source and destination
  - For tag - tag name
  - For assign to - user name
  - For email alert - email address / account

For example, the following lines from the log file show a successful copy operation and a failed copy operation.

```
2022-06-06 15:23:08,910 | INFO | ds_audit_logger | es_scanned_file | 237 | 49 | Copy file /CIFS_share/data/dop1/random_positives.tsv from device 10.31.133.183 (type: SMB_SHARE) to device 10.31.130.133:/export_reports (NFS_SHARE) - SUCCESS
2022-06-06 15:23:08,968 | WARNING | ds_audit_logger | es_scanned_file | 239 | 153 | Copy file /CIFS_share/data/compliance-netapp.tar.gz from device 10.31.133.183 (type: SMB_SHARE) to device 10.31.130.133:/export_reports (NFS_SHARE) - FAILURE
```

## Log file locations

The management audit log files are located on the BlueXP classification machine in:

`/opt/netapp/audit_logs/`

The installation audit log files are written to `/opt/netapp/install_logs/`

Each log file can be a maximum of 10 MB in size. When that limit is reached, a new log file is started. The log files are named "DataSense\_audit.log", "DataSense\_audit.log.1", "DataSense\_audit.log.2", and so on. A maximum of 100 log files are retained on the system - older log files are deleted automatically after the maximum has been reached.

## Access the log files

You'll need to log into the BlueXP classification system to access the log files. See how to [log in to the BlueXP classification system](#) depending on whether you manually installed the software on a Linux machine or if you deployed the instance in the cloud.

## Reducing the BlueXP classification scan speed

Data scans have a negligible impact on your storage systems and on your data. However, if you are concerned with even a very small impact, you can configure BlueXP classification to perform "slow" scans.

When enabled, slow scanning is used on all data sources - you can't configure slow scanning for a single working environment or data source.

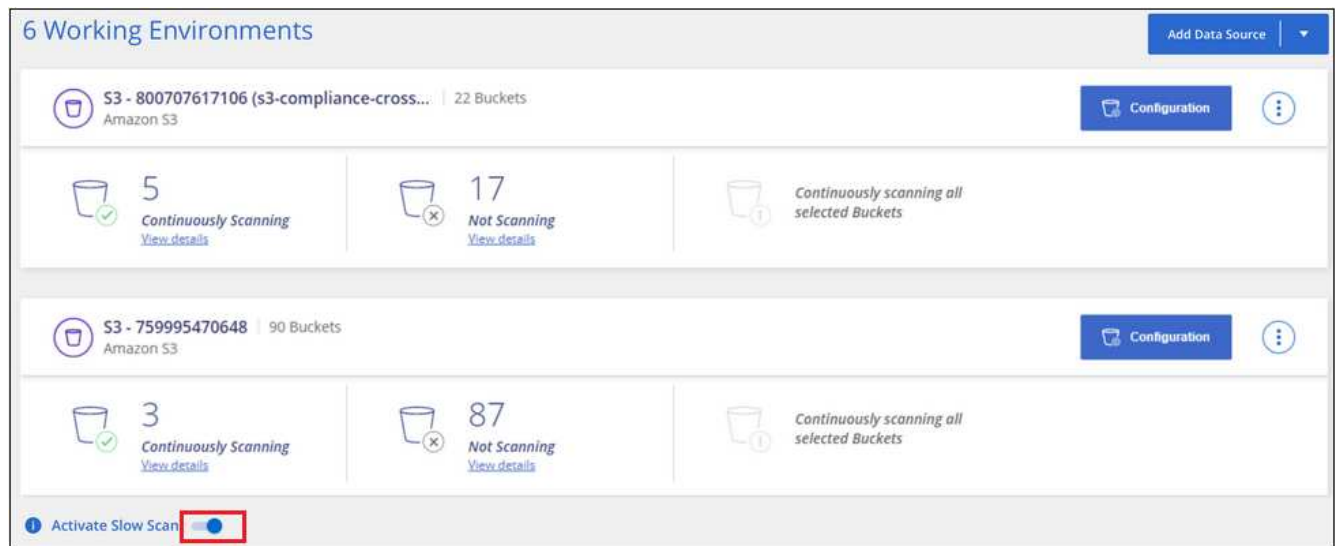


The scan speed can't be reduced when scanning databases.

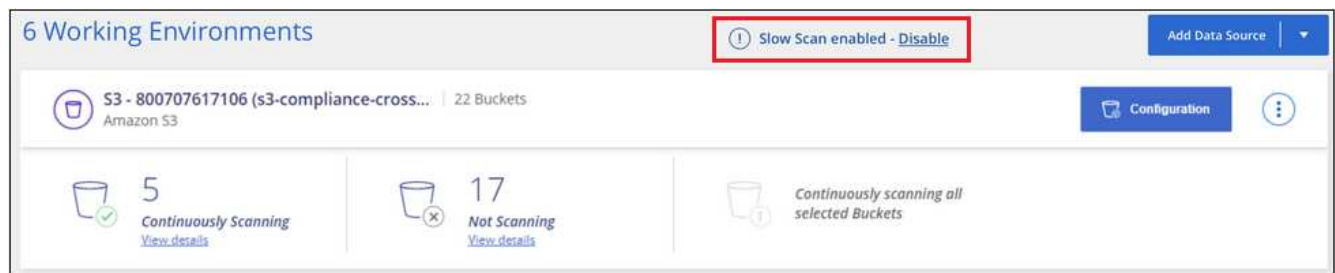
### Steps

1. From the bottom of the *Configuration* page, move the slider to the right to activate slow scanning.





The top of the Configuration page indicates that slow scanning is enabled.



2. You can disable slow scanning by clicking **Disable** from this message.


## Removing data sources from BlueXP classification

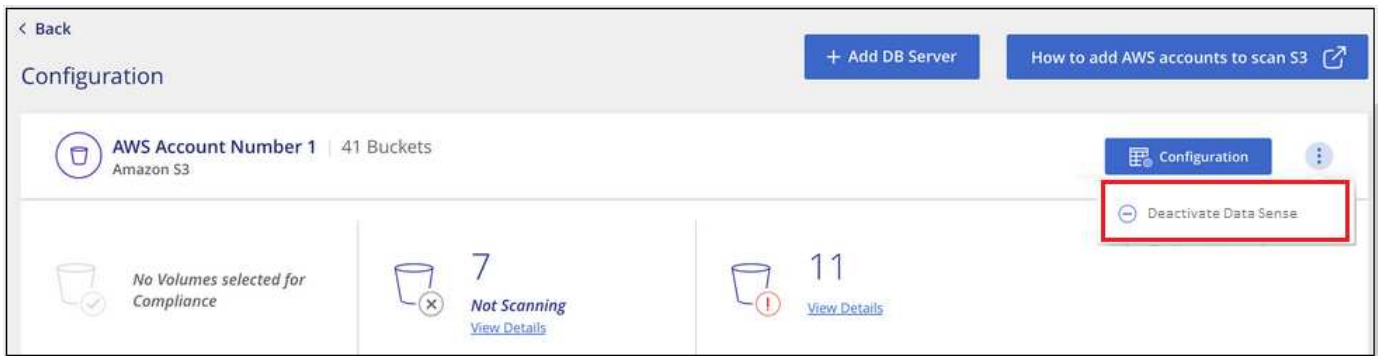
If you need to, you can stop BlueXP classification from scanning one or more working environments, databases, file share groups, OneDrive accounts, Google Drive accounts, or SharePoint accounts.

Charging for scanning the data is stopped when the data source is removed.

### Deactivating compliance scans for a working environment

When you deactivate scans, BlueXP classification no longer scans the data on the working environment and it removes the indexed compliance insights from the BlueXP classification instance (the data from the working environment itself isn't deleted).


1. From the *Configuration* page, click the  button in the row for the working environment, and then click **Deactivate Data Sense**.



You can also disable compliance scans for a working environment from the Services panel when you select the working environment.

## Removing a database from BlueXP classification

If you no longer want to scan a certain database, you can delete it from the BlueXP classification interface and stop all scans.


1. From the *Configuration* page, click the  button in the row for the database, and then click **Remove DB Server**.



## Removing a OneDrive, SharePoint, or Google Drive account from BlueXP classification

If you no longer want to scan user files from a certain OneDrive account, from a specific SharePoint account, or from a Google Drive account, you can delete the account from the BlueXP classification interface and stop all scans.

### Steps

1. From the *Configuration* page, click the  button in the row for the OneDrive, SharePoint, or Google Drive account, and then click **Remove OneDrive Account**, **Remove SharePoint Account**, or **Remove Google Drive account**.




2. Click **Delete Account** from the confirmation dialog.

## Removing a group of file shares from BlueXP classification

If you no longer want to scan user files from a file shares group, you can delete the File Shares Group from the BlueXP classification interface and stop all scans.

### Steps

1. From the *Configuration* page, click the  button in the row for the File Shares Group, and then click **Remove File Shares Group**.



2. Click **Delete Group of Shares** from the confirmation dialog.


## Uninstalling BlueXP classification

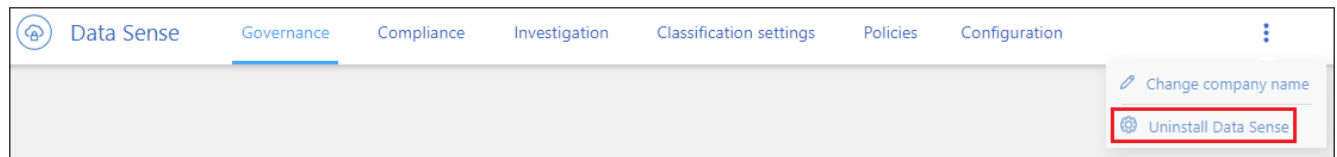
You can uninstall the BlueXP classification software to troubleshoot issues or to permanently remove the software from the host. Deleting the instance also deletes the associated disks where the indexed data resides - all the information BlueXP classification has scanned will be permanently deleted.

The steps that you need to use depend on whether you deployed BlueXP classification in the cloud or on an on-premises host.

### Uninstall BlueXP classification from a cloud deployment

You can uninstall and delete the BlueXP classification instance from the cloud provider environment if you no longer want to use BlueXP classification.

1. At the top of the BlueXP classification page, click  and then click **Uninstall Data Sense**.




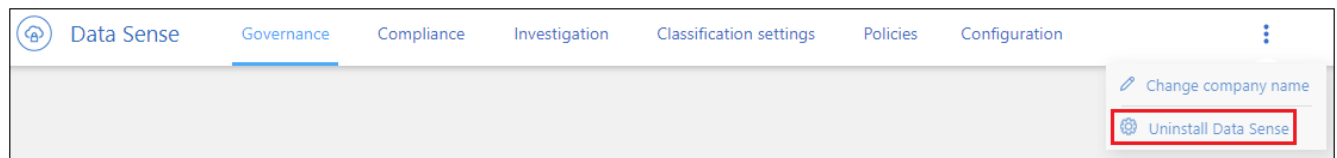
2. In the *Uninstall Data Sense* dialog, type **uninstall** to confirm that you want to disconnect the BlueXP classification instance from the BlueXP Connector, and then click **Uninstall**.
3. Go to your cloud provider's console and delete the BlueXP classification instance. The instance is named *CloudCompliance* with a generated hash (UUID) concatenated to it. For example: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

This deletes the instance and all associated data that had been collected by BlueXP classification.

## Uninstall BlueXP classification from an on-premises deployment

You can uninstall BlueXP classification from a host if you no longer want to use BlueXP classification, or if you had an issue that requires reinstallation.

1. At the top of the BlueXP classification page, click  and then click **Uninstall Data Sense**.



2. In the *Uninstall Data Sense* dialog, type **uninstall** to confirm that you want to disconnect the BlueXP classification instance from the BlueXP Connector, and then click **Uninstall**.
3. To uninstall the software from the host, run the `cleanup.sh` script on the host machine, for example:

```
cleanup.sh
```

See how to [log in to the BlueXP classification host machine](#).

## Copyright information

Copyright © 2024 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.