Data Loss Prevention in Forcepoint Web Security Cloud

Data Loss Prevention | Forcepoint Web Security Cloud | May 2022

The Data Security (DLP Lite) feature in Forcepoint Web Security Cloud lets you monitor and prevent the loss of sensitive data and intellectual property via the web, as well as to easily assess your current level of risk exposure via reporting. You can protect intellectual property, data that is protected by national legislation or industry regulation, and data suspected to be stolen by malware or malicious activities. When DLP Lite is used for data loss prevention, basic data protection is provide by the cloud proxy.

✓ **Note**

Integration with Data Protect Service is also available for Web Security Cloud customers. With this integration, enterprise data security is handled by the Data Protection Service. For further information, please contact your account manager.

This document guides you through the steps required to get started with Data Security (DLP Lite) for your web product using the Forcepoint Cloud Security Gateway Portal, also referred to as the cloud portal.

✓ **Note**

DLP Lite is not supported with the Direct Connect endpoint or I Series appliances.

The following steps are required to configure data security for your account.

1. *Create content classifiers*

   Content classifiers are rules you can define to identify sensitive information, using custom phrases, dictionaries or regular expressions containing business specific terms or labels. This is helpful for monitoring intellectual property.

2. *Configure Data Security (DLP Lite) policy settings*

   Use the Data Security tab in your policies to define which types of data are protected, and the action to take when data loss is detected.

3. *Configure reporting permissions*

   This determines who can see data protection reports.

In addition, you can optionally:

- *Configure privacy settings*
- *Configure block pages*
- *View the dashboard*
- *View reports*
- *View the audit trail*

Create content classifiers

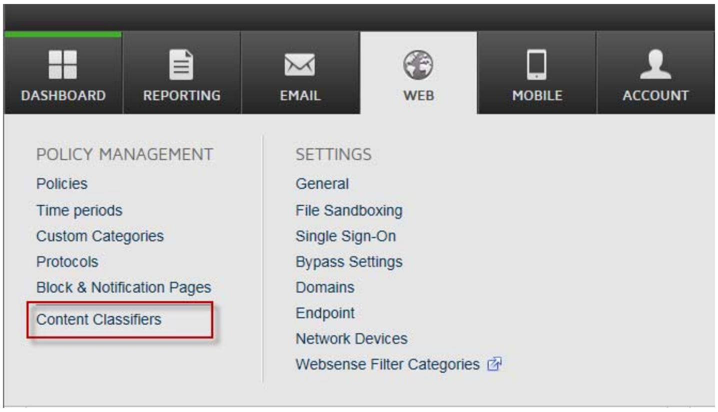
Data Loss Prevention | Forcepoint Web Security Cloud

Content classifiers can be used to identify intellectual property and data types that are not covered by the default Personally Identifiable Information (PII), Payment Card Industry (PCI), and Protected Health Information (PHI) rules. For example, a key phrase custom classifier can be created to identify a document marker, such as "Acme Corp - Internal Confidential".

The content classifiers that you create can then be used on the Data Security tab of your web policies.

If you are concerned only about data loss related to regulatory compliance, you can skip this step.

1. In the cloud portal, navigate to **Web > Policy Management > Content Classifiers**.

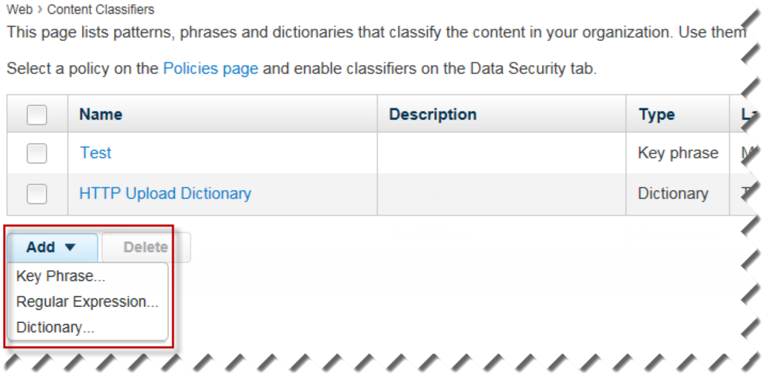


2. Click **Add** and select the type of classifier you want to create:
   - **Key Phrase**: a keyword or phrase that indicates sensitive or proprietary data (such as product code names or patents).
   - **Regular Expression**: a pattern used to describe a set of search criteria based on syntax rules.

     For example, the pattern "a\d+" detects all strings that start with the letter "a" and are followed by at least one digit, where "\d" represents any digit and "+" represents "at least one."

Regular expression patterns are detailed in the Forcepoint Web Security Cloud help: see [Regular expression content classifiers](#).

- **Dictionary**: a container for words and expressions relating to your business.



Web › Content Classifiers

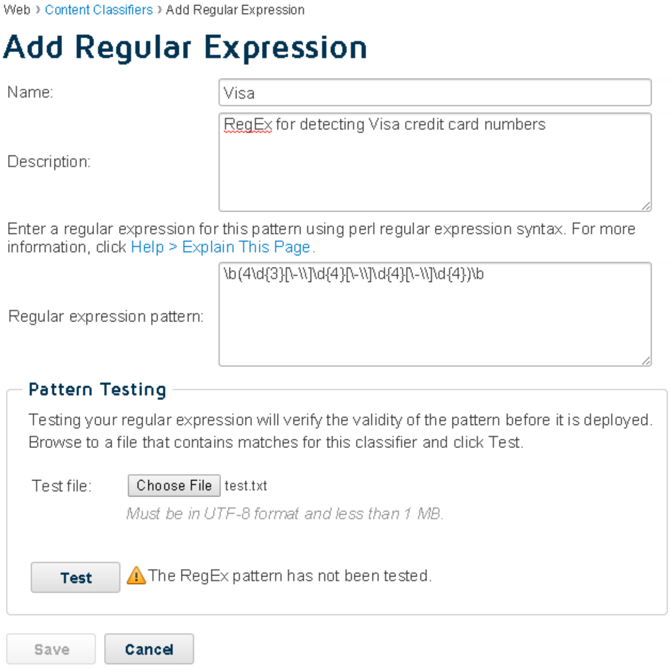
This page lists patterns, phrases and dictionaries that classify the content in your organization. Use them

Select a policy on the Policies page and enable classifiers on the Data Security tab.

□	Name	Description	Type
□	Test		Key phrase
□	HTTP Upload Dictionary		Dictionary

3. Complete the fields as described in the appropriate section, and then click **Save**.
   - *Key phrase content classifiers*, page 4
   - *Regular expression content classifiers*, page 3
   - *Dictionary content classifiers*, page 5
4. Repeat steps 2-3 until you've added all the classifiers you require.

Regular expression content classifiers



Web › Content Classifiers › Add Regular Expression

Add Regular Expression

Name: Visa

Description: RegEx for detecting Visa credit card numbers

Enter a regular expression for this pattern using perl regular expression syntax. For more information, click Help > Explain This Page.

Regular expression pattern:

```
\b(4\d{3}[\-\\]\d{4}[\-\\]\d{4}[\-\\]\d{4})\b
```

Pattern Testing

Testing your regular expression will verify the validity of the pattern before it is deployed. Browse to a file that contains matches for this classifier and click Test.

Test file: Choose File test.txt

Must be in UTF-8 format and less than 1 MB.

⚠ The RegEx pattern has not been tested.

Regular expression (regex) patterns can be detected within content, such as the pattern of an internal account number, or alphanumeric document code.

When extracted text from a transaction is scanned, the system searches for strings that match regular expression patterns and may be indicative of confidential information.

To create a regular expression classifier:

1. Enter a unique **Name** for the pattern.
2. Enter a **Description** for the pattern.
3. Enter the **Regular expression pattern** (regex) that you want the system to search for, using Perl syntax.

   For syntax and examples, click **Help > Explain This Page** within the cloud portal, or view the help page at the following link: Regular expression content classifiers.
4. Use the Pattern Testing section of the page to test your regular expression.

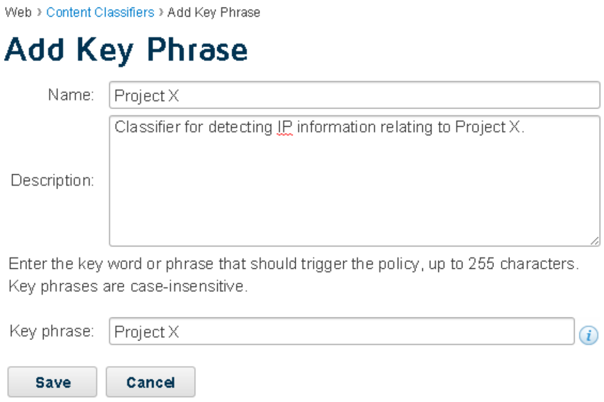
   Because regular expression patterns can be quite complex, it is important that you test the pattern before saving it. If improperly written, a pattern can create false-positive incidents.

   a. Create a .txt file (less than 1 MB) that contains values that match this regex pattern. The file must be in plain text UTF8 format.

   b. Browse to the file and click **Test** to test the validity of your pattern syntax. If the pattern you entered is invalid, you're given an opportunity to fix it. You cannot proceed until the test succeeds.

You can have up to 100 regular expression classifiers.

Key phrase content classifiers



Web › Content Classifiers › Add Key Phrase

Add Key Phrase

Name: Project X

Description: Classifier for detecting IP information relating to Project X.

Enter the key word or phrase that should trigger the policy, up to 255 characters. Key phrases are case-insensitive.

Key phrase: Project X

The presence of a keyword or phrase (such as "Top Secret" or "Project X") in a web post may indicate that classified information is being exposed. You can learn about activity like this by defining a key phrase classifier.

To create a key phrase classifier:

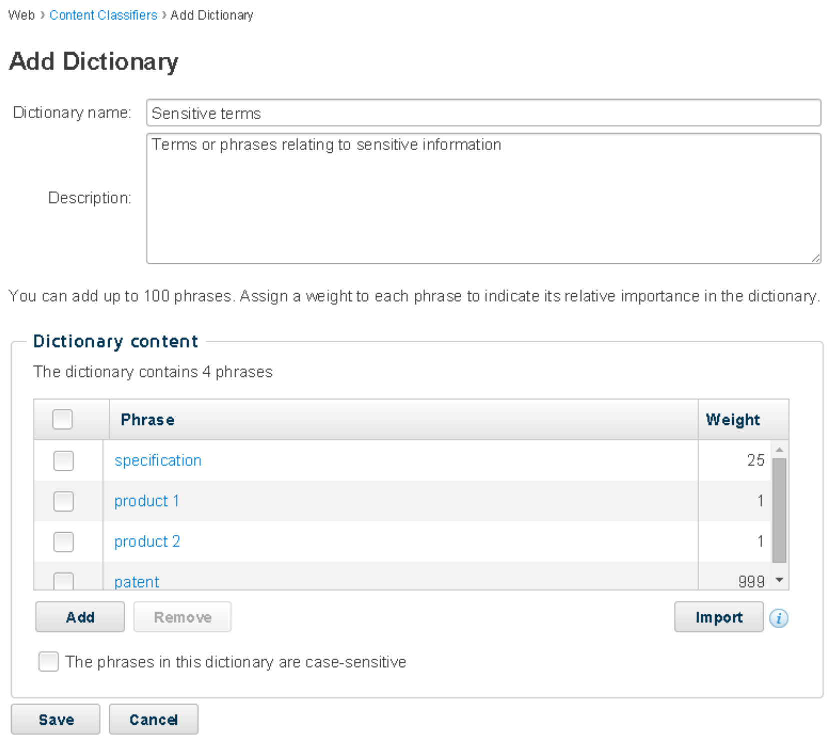
1. Enter a unique **Name** for the key phrase classifier.
2. Enter a **Description** for the key phrase.
3. Enter the key word or phrase that might indicate classified information, up to 255 characters. Key phrases are case-insensitive.

   Leading and trailing white spaces are ignored. If you need to use slashes, tabs, hyphens, underscores, or carriage returns, define a regular expression classifier rather than a key word classifier.

Key phrases also identify partial matches. For example, the key phrase “uri” reports a match for “security”. Note that wildcards are not supported for key phrases.

You can have up to 100 key phrase classifiers.

Dictionary content classifiers



Web › Content Classifiers › Add Dictionary

Add Dictionary

Dictionary name: Sensitive terms

Description: Terms or phrases relating to sensitive information

You can add up to 100 phrases. Assign a weight to each phrase to indicate its relative importance in the dictionary.

Dictionary content

The dictionary contains 4 phrases

□	Phrase	Weight
□	specification	25
□	product 1	1
□	product 2	1
□	patent	999

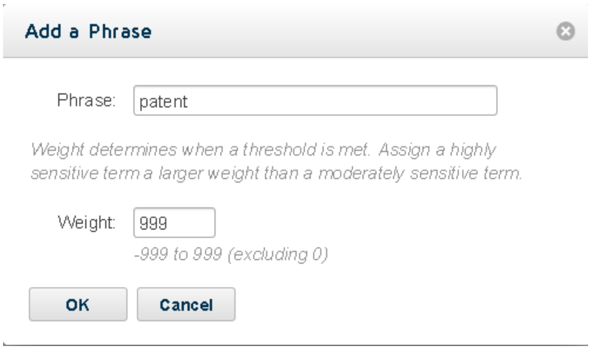
□ The phrases in this dictionary are case-sensitive

A dictionary is a container for words and expressions pertaining to your business.

To create a dictionary classifier:

1. Enter a unique **Name** for the dictionary classifier.
2. Enter a **Description** for the dictionary.

3. Dictionaries can have up to 100 phrases. To add content to the dictionary, click **Add** under Dictionary content.



Add a Phrase

Phrase: patent

Weight determines when a threshold is met. Assign a highly sensitive term a larger weight than a moderately sensitive term.

Weight: 999

-999 to 999 (excluding 0)

4. Complete the fields on the resulting dialog box as follows:
   a. **Phrase**: Enter a word or phrase to include. This phrase, when found in the content, affects whether the content is considered suspicious.
   b. **Weight**: Select a weight, from -999 to 999 (excluding 0). When matched with a threshold, weight defines how many instances of a phrase can be present, in relation to other phrases, before triggering a policy.
5. If you have many phrases to include, create a text file listing the phrases, then click **Import** and navigate to the text file.
6. Mark **The phrases in this dictionary are case-sensitive** if you want the phrases that you entered to be added to the dictionary with the same case you applied.

You can have up to 100 dictionary classifiers. Each is limited to 100 phrases.

For examples and restrictions, click **Help > Explain This Page**.

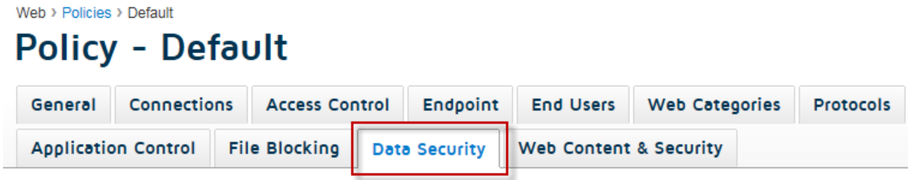
Configure Data Security (DLP Lite) policy settings

Data Loss Prevention | Forcepoint Web Security Cloud

To configure options for detecting and preventing data loss over web channels:

1. In the portal, navigate to **Account > Data Protection Settings.**
2. In the **Web Defaults** section, select **Use DLP Lite**. Save you changes.

   When **Use DLP Lite** is selected, a Data Security tab is available for new policies.
3. Navigate to the **Web > Policy Management > Policies**, page, then open the policy you want to configure.
4. Click the **Data Security** tab in the policy.



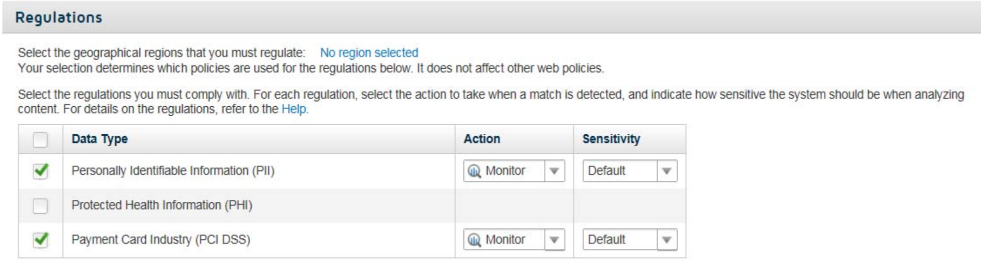
5. Complete the fields as described in the following sections:
   - *Data security regulations*, page 8
   - *Data theft detection*, page 9
   - *Custom data security classifiers*, page 10
   - *Trusted domains*, page 11
6. When you are finished, click **Save**.

The system will search for sensitive data that is being posted to HTTP and HTTPS sites, and report on it in an incident report available from the **Reporting > Report Catalog > Standard Reports > Data Security** page.

This report includes intellectual property, data that is protected by national legislation or industry regulation, and data suspected to be stolen by malware or malicious activities.

To search for data over HTTPS, be sure SSL decryption is enabled by following the instructions provided on the SSL Decryption tab.

Data security regulations



Regulations

Select the geographical regions that you must regulate: No region selected
Your selection determines which policies are used for the regulations below. It does not affect other web policies.

Select the regulations you must comply with. For each regulation, select the action to take when a match is detected, and indicate how sensitive the system should be when analyzing content. For details on the regulations, refer to the Help.

☐	Data Type	Action	Sensitivity
☑	Personally Identifiable Information (PII)	Monitor	Default
☐	Protected Health Information (PHI)		
☑	Payment Card Industry (PCI DSS)	Monitor	Default

Most countries and certain industries have laws and regulations that protect customers, patients, or staff from the loss of personal information such as credit card numbers, social security numbers, and health information.

To set up rules for the regulations that pertain to you:

1. Click **No region selected**.
2. Select the regions in which you operate.
3. Select the regulations of interest:

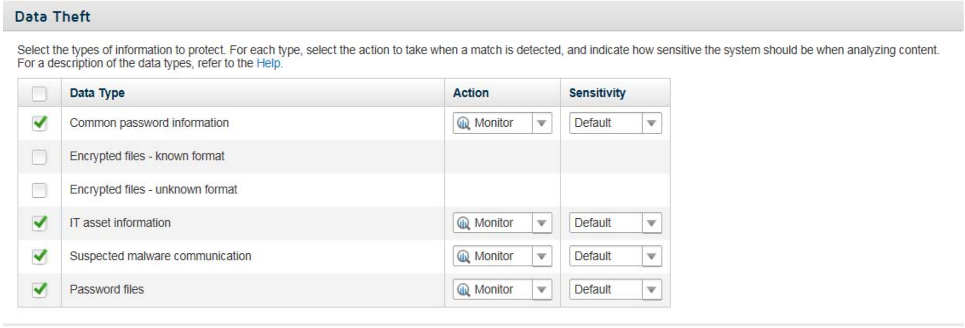
Field	Description
Personally Identifiable Information (PII)	Detects Personally Identifiable Information. For example, names, birth dates, driver license numbers, and identification numbers. This option is tailored to specific countries.
Protected Health Information (PHI)	Detects Protected Health Information. For example, terms related to medical conditions and drugs, together with identifiable information.
Payment Card Industry (PCI DSS)	Conforms to the Payment Card Industry (PCI) Data Security Standard, a common industry standard that is accepted internationally by all major credit card issuers. The standard is enforced on companies that accept credit card payments, as well as other companies and organization that process, store, or transmit cardholder data.

4. Select an action to take when matching data is detected. Select **Block** to prevent the data from being sent through the web channel. Select **Monitor** to allow it. (Incidents are created either way.) You can filter by action in the Data Security Incident Manager.
5. Select a sensitivity to indicate how narrowly or widely to conduct the search.

   Select **Wide** for the strictest security. Wide has a looser set of detection criteria than Default or Narrow, so false positives may result and performance may be affected. Select **Narrow** for tighter detection criteria. This can result in false negatives or undetected matches. **Default** is a balance between the two.

   Severity is automatically calculated for these regulations.

Data theft detection



Data Theft

Select the types of information to protect. For each type, select the action to take when a match is detected, and indicate how sensitive the system should be when analyzing content. For a description of the data types, refer to the Help.

☐	Data Type	Action	Sensitivity
☑	Common password information	Monitor	Default
☐	Encrypted files - known format		
☐	Encrypted files - unknown format		
☑	IT asset information	Monitor	Default
☑	Suspected malware communication	Monitor	Default
☑	Password files	Monitor	Default

Use this section to detect when data is being exposed due to malware or malicious transactions. When you select these options, Forcepoint Web Security Cloud searches for and reports on outbound passwords, encrypted files, network data, and other types of information that could be indicative of a malicious act.

To see if your organization is at risk for data theft:

1. Select the types of data to look for.

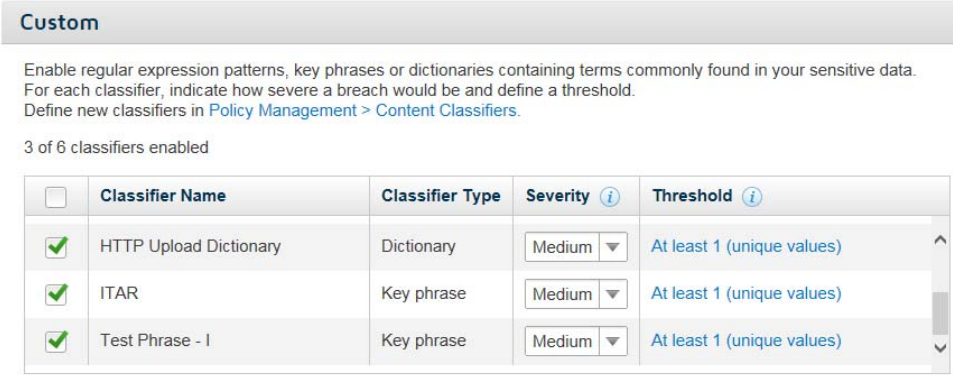
Information Type	Description
Common password information	Searches for outbound passwords in plain text
Encrypted file - known format	Searches for outbound transactions comprising common encrypted file formats
Encrypted file - unknown format	Searches for outbound files that were encrypted using unknown encryption formats
IT asset information	Searches for suspicious outbound transactions, such as those containing information about the network, software license keys, and database files.
Malware communication	Identifies traffic that is thought to be malware “phoning home” or attempting to steal information. Detection is based on the analysis of traffic patterns from known infected machines.
Password files	Searches for outbound password files, such as a SAM database and UNIX/Linux passwords files

2. Select an action to take when matching data is detected. Select **Block** to prevent the data from being sent through the web channel. Select **Monitor** to allow it. (Incidents are created either way.) You can filter by action in the Data Security Incident Manager.
3. Select a sensitivity to indicate how narrowly or widely to conduct the search.

   Select **Wide** for the strictest security. Wide has a looser set of detection criteria than Default or Narrow, so false positives may result and performance may be affected. Select **Narrow** for tighter detection criteria. This can result in false negatives or undetected matches. **Default** is a balance between the two.

Severity is automatically calculated for these types.

Custom data security classifiers



Custom

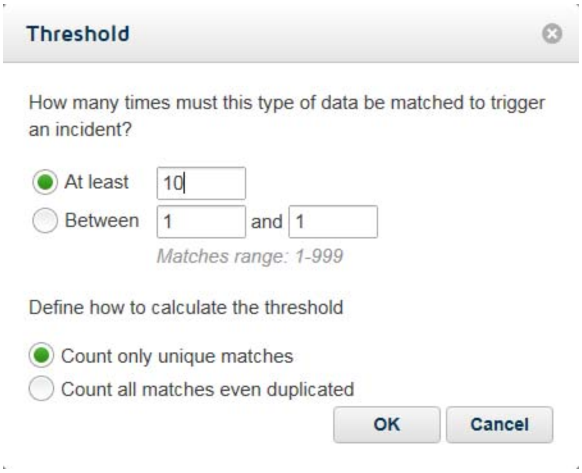
Enable regular expression patterns, key phrases or dictionaries containing terms commonly found in your sensitive data. For each classifier, indicate how severe a breach would be and define a threshold. Define new classifiers in Policy Management > Content Classifiers.

3 of 6 classifiers enabled

□	Classifier Name	Classifier Type	Severity	Threshold
☑	HTTP Upload Dictionary	Dictionary	Medium	At least 1 (unique values)
☑	ITAR	Key phrase	Medium	At least 1 (unique values)
☑	Test Phrase - I	Key phrase	Medium	At least 1 (unique values)

Use this section if you want to detect intellectual property or sensitive data using custom phrases, dictionaries, or regular expressions containing business-specific terms or data.

1. Select the classifiers that you want to enable for the policy. If you skipped the section *Create content classifiers*, page 2, go there now to populate the list.
2. Select a severity for each classifier to indicate how severe a breach would be. Select **High** for the most severe breaches. Severity is used for reporting purposes. It allows you to easily locate High, Medium, or Low severity breaches when viewing reports.
3. Configure a threshold for each classifier.



Threshold

How many times must this type of data be matched to trigger an incident?

● At least 10

○ Between 1 and 1

Matches range: 1-999

Define how to calculate the threshold

● Count only unique matches

○ Count all matches even duplicated

a. Click the link in the Threshold column.

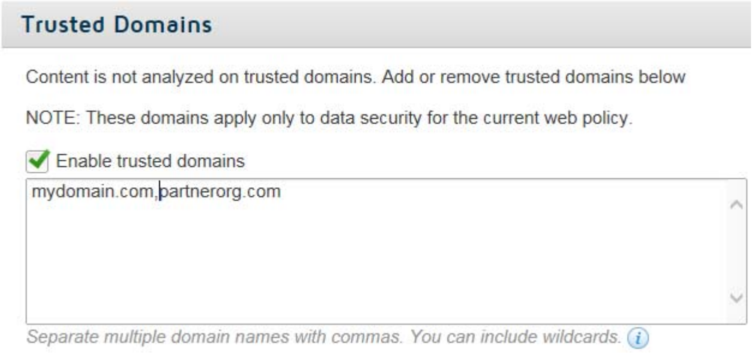
b. Indicate how many times this classifier should be matched to trigger an incident. You can indicate a range if desired, such as between 3 and 10. By default, the threshold is 1.

c. Indicate whether you want the system to count each match, even if it is a duplicate, against the threshold, or whether you'd prefer to only count unique matches.

d. Click **OK**.

Trusted domains

Select **Enable trusted domains** if you do not want certain domains to be monitored, then enter URLs for the trusted domains separated by commas.



Trusted Domains

Content is not analyzed on trusted domains. Add or remove trusted domains below

NOTE: These domains apply only to data security for the current web policy.

☑ Enable trusted domains

mydomain.com,partnerorg.com

Separate multiple domain names with commas. You can include wildcards.

The system does not analyze content passed between trusted domains. This means users can send them any type of sensitive information via HTTP, HTTPS, or other web channels from your network.

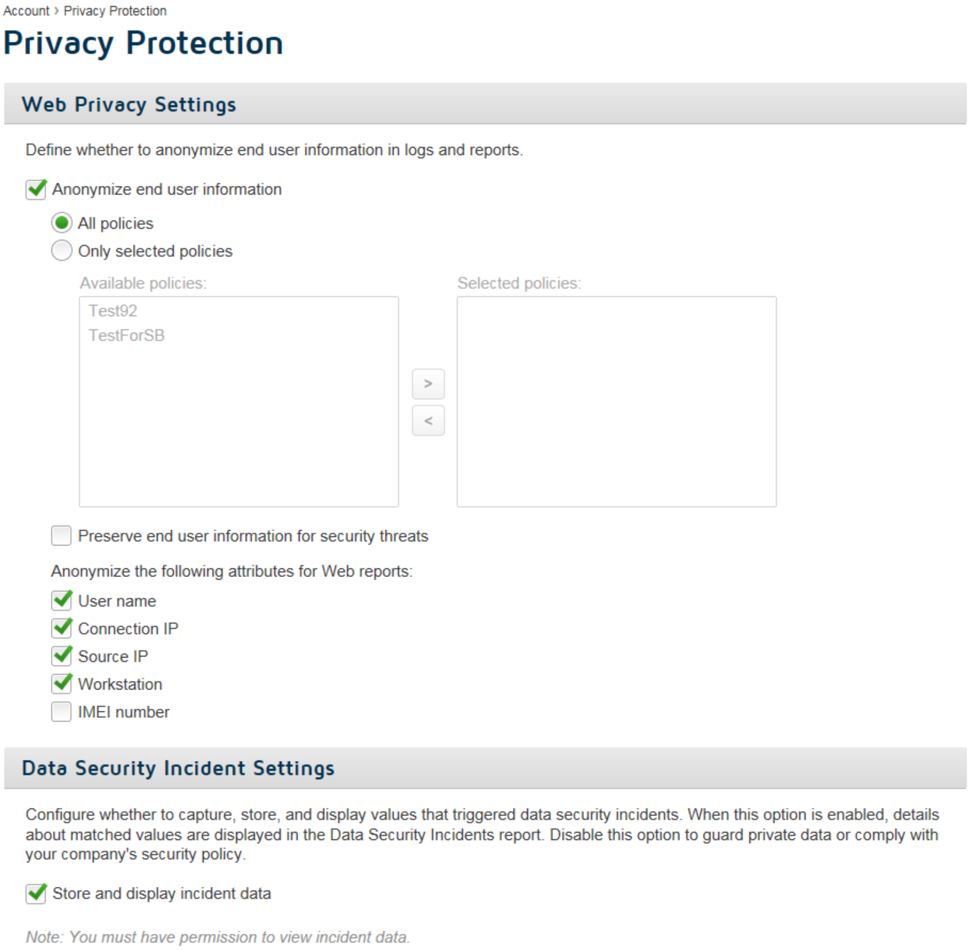
The domains you enter apply only to data security and only to the current web policy.

Duplicate URLs are not permitted. Wildcards and '?' are supported.

Configure privacy settings

Data Loss Prevention | Forcepoint Web Security Cloud

Use the **Account** > **Settings** > **Privacy Protection** page to prevent end-user identifying information, data security incident trigger values, or both from appearing in logs and web reports. If required, you can still collect this information for security threats.



Account › Privacy Protection

Privacy Protection

Web Privacy Settings

Define whether to anonymize end user information in logs and reports.

☑ Anonymize end user information

● All policies
○ Only selected policies

Available policies:
Test92
TestForSB

Selected policies:

□ Preserve end user information for security threats

Anonymize the following attributes for Web reports:

☑ User name
☑ Connection IP
☑ Source IP
☑ Workstation
□ IMEI number

Data Security Incident Settings

Configure whether to capture, store, and display values that triggered data security incidents. When this option is enabled, details about matched values are displayed in the Data Security Incidents report. Disable this option to guard private data or comply with your company's security policy.

☑ Store and display incident data

Note: You must have permission to view incident data.

By default, incident data is *not* captured, stored, or displayed. Administrators with permission to view incident data are able to see the number of matches in the report, but not the match values or context.

Select **Store and display incident data** under Data Security Incident Settings if you want the values that triggered data security incidents to be captured, stored in the incident database, and displayed in reports.

Credit card numbers, social security numbers, and email addresses are masked when they are stored, as are passwords in certain instances.

Changing this setting has no impact on incident data that has already been collected.

Configure reporting permissions

Data Loss Prevention | Forcepoint Web Security Cloud

You can control which administrators can view data security reports (and potentially sensitive information). This setting is assigned at the account level.

To give administrators these permissions:

1. Navigate to **Account > Settings > Contacts**.
2. Select the contact whose permissions you want to edit.
3. In Contact Details, click the user name (email address) to view the contact login details.
4. On the Login Details screen, click **Edit**.
5. Under Account Permissions, select **View All Reports** and **Data Security Reports**, and then click **Save**.

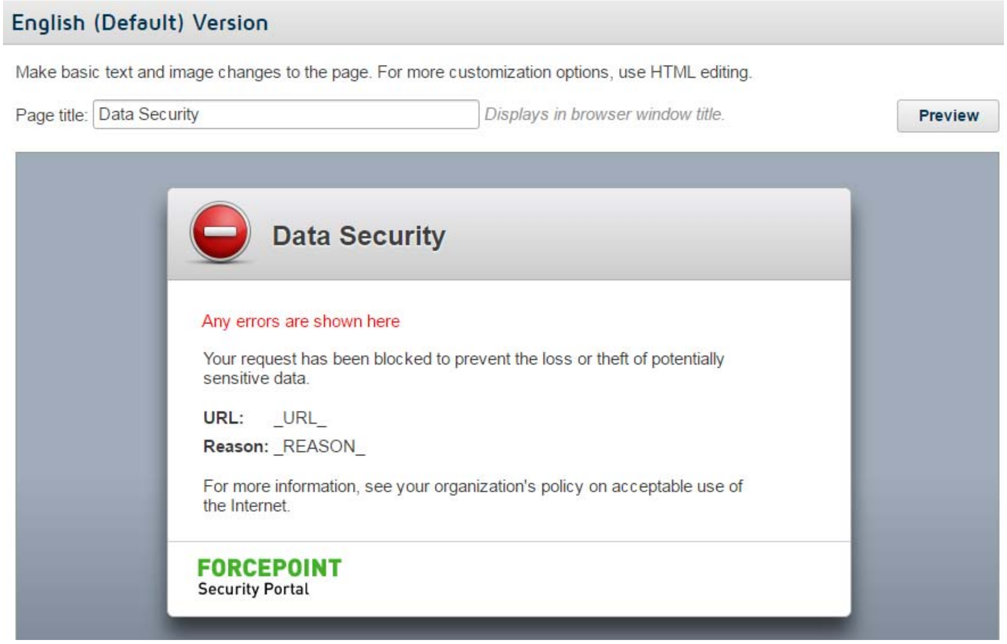
This enables users to view data security reports, which may or may not contain incident forensics and trigger data, depending on your privacy protection settings. It does not change their ability to manage data security configuration settings.

Configure block pages

Data Loss Prevention | Forcepoint Web Security Cloud

You have the option to customize the block pages that users receive when they request a web page that is blocked by a Data Security policy. To do so:

1. Go to the **Web > Policy Management > Block & Notification Pages** page.
2. Expand **General**.
3. Click **Data Security**.



English (Default) Version

Make basic text and image changes to the page. For more customization options, use HTML editing.

Page title: Data Security *Displays in browser window title.*

Data Security

Any errors are shown here

Your request has been blocked to prevent the loss or theft of potentially sensitive data.

URL: _URL_

Reason: _REASON_

For more information, see your organization's policy on acceptable use of the Internet.

4. Click in the title or body to edit the default text. You can replace logos and other images as well.
5. When you’re finished, click **OK**.

View the dashboard

Data Loss Prevention | Forcepoint Web Security Cloud

For a high-level view of activity in your organization, click **Dashboard**, and then click the **Data Security** tab. Data Security charts include:

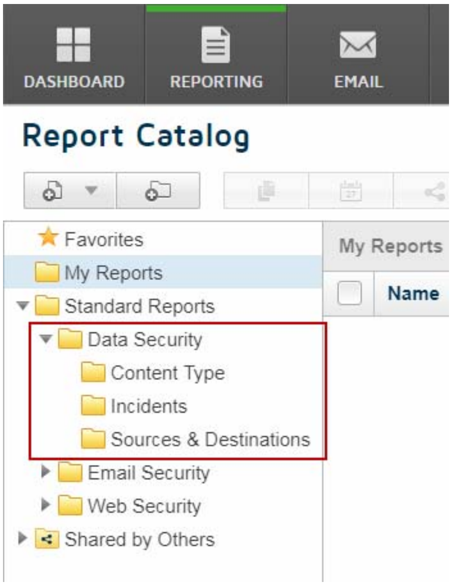
- **Incident Count Timeline** shows a daily incident count for the designated period. With it, you can quickly identify trends and make policy changes as required.
- **Incidents by Content CategoryTotal Incidents by Content Type** shows the number of regulatory incidents, data theft incidents, and custom classifier incidents in the designated period.
- **Top Sources** shows the users, machines, or IP addresses most frequently instigating data security violations as well as the severity of their incidents.
- **Top Destination Domains** shows the Internet domains most frequently targeted with sensitive data.
- **Top Web Categories** shows the website categories most frequently targeted with sensitive data. These can be custom categories or the categories classified by the URL category database.

View reports

Data Loss Prevention | Forcepoint Web Security Cloud

For a more granular view, access the data security reports.

1. Go to the **Reporting > Report Catalog** page.
2. Select **Standard Reports > Data Security** from the left navigation pane, and then select a report category: Content Type, Incidents, or Sources & Destinations.



3. Select a report from the list. The following table provides descriptions of each report

Report	Description
Content Type	
Compliance Summary	Details the compliance rules are most often violated in your organization, and provides a breakdown of the incident count for each policy or rule.
Custom Classifier Summary	Shows which custom classifiers triggered the most incidents during the designated period.
Data Theft Summary	A list of data theft classifiers that triggered the most incidents during the designated period.
Incidents	
Incident List	A list or chart of all data loss incidents that were detected during the designated period, along with incident details such as the destination, severity, and transaction size.
Sources & Destinations	
Destination Summary	The destination URLs or IP addresses involved with the most violations, broken down by severity.
Users Summary	The users, machines, or IP addresses most frequently violating data security policies and the severity of their breaches.

4. After you select a report, select a time period (last 7 days by default) and any required attributes, then click the **Update Report** button.

Tip

To view only incidents that meet a certain threshold (not every single match), filter the report using the Top Matches attribute.

Top Matches indicates the number of matches on the incident's most violated rule. For example, if rule A in MyPolicy has 2 matches, rule B has 5 matches, and rule C has 10 matches, top match equals 10.

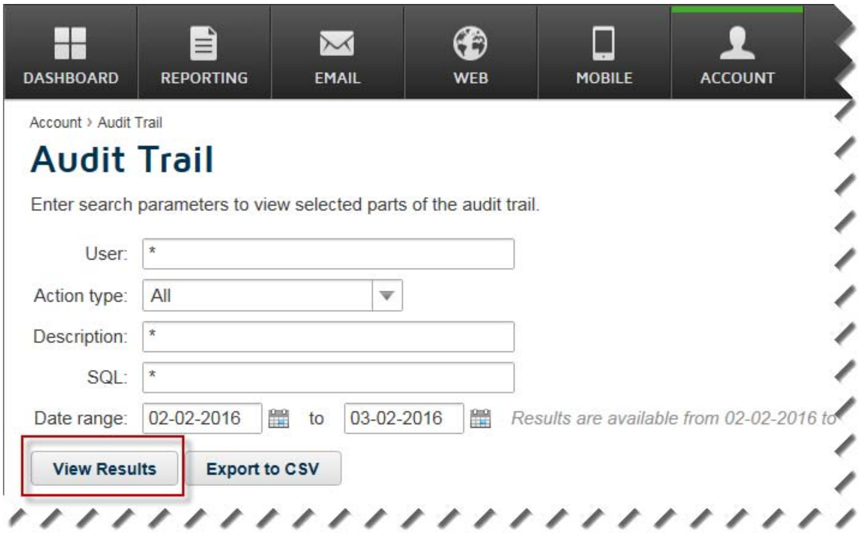
When you apply the filter, enter the threshold to include in the report, and then select the operator to use: equal to, greater than, etc.

Refer to the Forcepoint Cloud Security Gateway Portal Help for details on adding attributes to a report.

View the audit trail

Data Loss Prevention | Forcepoint Web Security Cloud

Navigate to **Account > Settings > Audit Trail**, and click **View Results** to see an audit trail of all policy configuration changes.

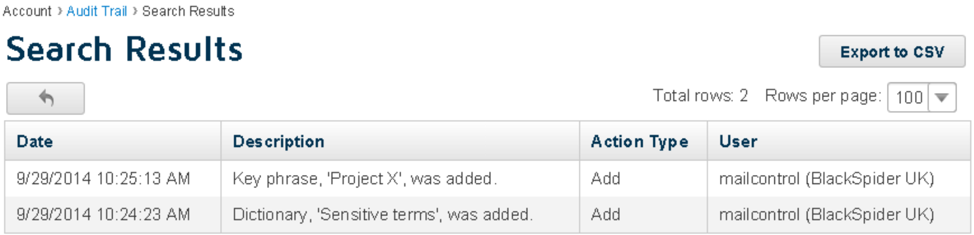


Audit Trail

Enter search parameters to view selected parts of the audit trail.

User:	*
Action type:	All
Description:	*
SQL:	*
Date range:	02-02-2016 to 03-02-2016 *Results are available from 02-02-2016 to*

You can search by user, action type, and date range.



Search Results

Total rows: 2 Rows per page: 100

Date	Description	Action Type	User
9/29/2014 10:25:13 AM	Key phrase, 'Project X', was added.	Add	mailcontrol (BlackSpider UK)
9/29/2014 10:24:23 AM	Dictionary, 'Sensitive terms', was added.	Add	mailcontrol (BlackSpider UK)

Copyright and trademarks

Data Loss Prevention | Forcepoint Web Security Cloud

Every effort has been made to ensure the accuracy of this document. However, Forcepoint makes no warranties with respect to this documentation and disclaims any implied warranties of merchantability and fitness for a particular purpose. Forcepoint shall not be liable for any error or for incidental or consequential damages in connection with the furnishing, performance, or use of this manual or the examples herein. The information in this documentation is subject to change without notice.

Published 2022