Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

This topic is 1 of 7 in a series 1 2 3 4 5 6 7

Introduction

This solution demonstrates how to protect sensitive data that is stored in Office 365 services. It includes prescriptive recommendations for discovering, classifying, protecting, and monitoring personal data. This solution uses General Data Protection Regulation (GDPR) as an example, but you can apply the same process to achieve compliance with many other regulations.

GDPR regulates the collection, storage, processing, and sharing of personal data. Personal data is defined very broadly under the GDPR as any data that relates to an identified or identifiable natural person that is a resident of the European Union (EU). See Topic 2 to review the Article 4 definition.

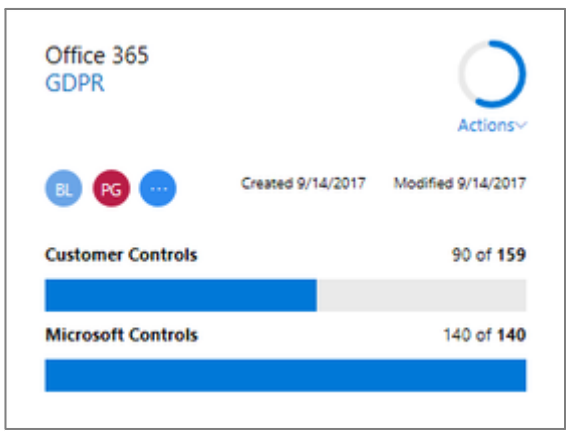
This solution content is intended to help organizations discover and protect personal data in Office 365 that might be subject to the GDPR. It is not offered as a GDPR compliance attestation. Organizations are responsible for ensuring their own GDPR compliance and are advised to consult their legal and compliance teams or to seek guidance and advice from third parties that specialize in compliance.

GDPR Assessment is a quick, online self-evaluation tool available at no cost to help your organization review its overall level of readiness to comply with the GDPR (http://aka.ms/gdprassessment).

Assess and manage your compliance risk

The first step towards GDPR compliance is to assess whether the GDPR applies to your organization, and, if so, to what extent. This analysis includes understanding the data your organization processes and where it resides.

1 Use Compliance Manager to view the regulation requirements and track your progress



Compliance Manager provides tools to track, implement, and manage the auditing controls to help your organization reach compliance against various standards, including GDPR.

For more information, see topic 7 in this guide — **Use Compliance Manager in the Service Trust Portal**.

2 Use Content Search and sensitive information types to find personal data

Discover personal data in your environment that is subject to the GDPR. Use Content Search together with sensitive information types to:

- Find and report on where personal data resides.
- Optimize sensitive data types and other queries to find all personal data in your environment.

Sensitive information types define how the automated process recognizes specific information types such as health service numbers and credit card numbers.

This guide includes a set you can use as a starting point. Many more sensitive information types are coming soon for personal data in EU countries.

For more information, see topic 2 in this guide — **Search for and find personal data**.

Classify, protect, and monitor personal data in Office 365 and other SaaS apps

Some of the capabilities used for information protection in Office 365 can also be used to protect sensitive data in other SaaS applications.

3 Decide if you want to use labels in addition to sensitive information types

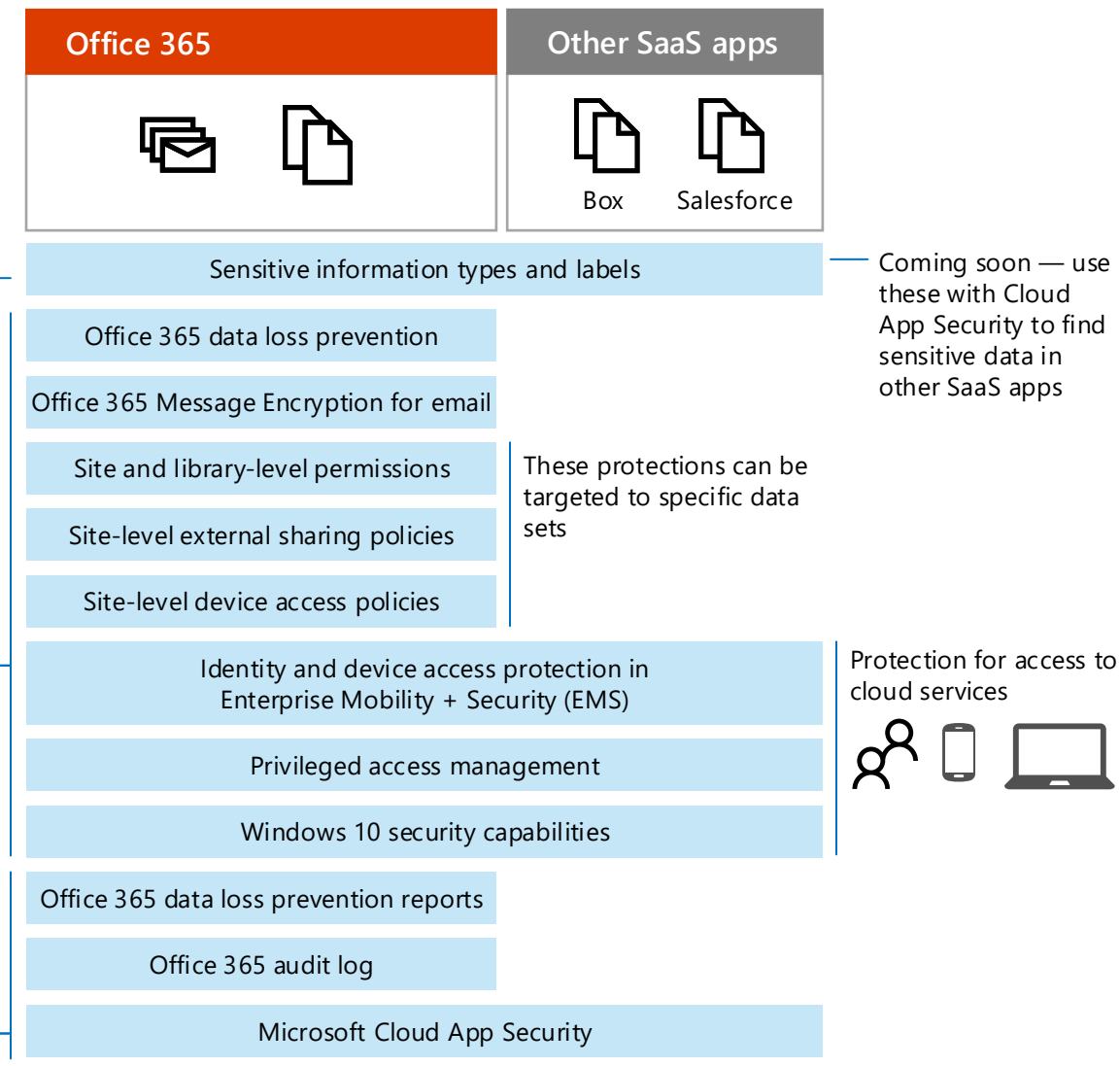
Sensitive information types are a form of classification. See topic 3, **Architect a classification schema for personal data**, to decide if you also want to implement labels. To apply labels, see topic 4, **Apply labels to personal data in Office 365**.

4 Protect personal data in Office 365

See topic 5, **Apply protection to sensitive data in Office 365**, for more information about configuring data loss prevention and other protections for sensitive data.

5 Monitor for leaks of personal data

Office 365 data loss prevention reports provide the greatest level of detail for monitoring sensitive data. Cloud App Security extends the ability to find and monitor sensitive data to other SaaS providers. See topic 6, **Monitor for breaches of personal data**.



See topics 2-7 for more information and resources.

Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

Search for and find personal data

Personal data is defined very broadly under the GDPR as any data that relates to an identified or identifiable natural person that is a resident of the European Union (EU).

Article 4 – Definitions

'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

This topic demonstrates how to find personal data stored in SharePoint Online and OneDrive for Business (which includes the sites for all Office 365 groups and Microsoft Teams).

Finding personal data that is subject to GDPR relies on using *sensitive information types* in Office 365. These define how the automated process recognizes specific information types such as health service numbers and credit card numbers. At this time these cannot be used to find data in Exchange mailboxes at rest. However, sensitive information types can be used with data loss prevention policies to find personal data in mail while it is in transit.

So, while you can't currently use Content Search to find personal data at rest in Exchange Online mailboxes, you can use the sensitive information types you curate for GDPR to find and protect personal information as it is sent through email.

Use Content Search to find personal data

Microsoft recommends a three-stage approach to finding personal data in Office 365. The rest of this topic provides guidance for each of these stages.

1 **Search for sensitive information types**	Start by using sensitive information types to find personal data. Create a Content Search query for each sensitive information type. Run the query and analyze the results.	If needed, add parameters to the query to reduce false positives: • Count range • Confidence range • Other properties or operators for more complex queries	If necessary, modify a sensitive information type to improve accuracy for your organization. • Adjust the confidence level directly in the XML. • Add key words. • Adjust the proximity requirements for keywords.
2 **Use Keyword Query Language (KQL) to find additional personal data in your environment**	To find data not included in sensitive information types, use the KQL query language to develop custom queries.	Test the results of these searches and adjust the KQL query string until you achieve the expected result.	
3 **Create new custom sensitive information types using the KQL queries**	After optimizing KQL queries to find target data, create new custom sensitive information types using these queries. You can then use these custom sensitive information types with Content Search, in DLP policies and other tools, and within other KQL queries.	Coming soon — You'll be able to create and modify sensitive information types in a new user interface in the Security and Compliance Center. You can dynamically see matching results and tune sensitive information types to meet your needs.	

Search for sensitive information types using Content Search

Begin searching for personal data by using the sensitive information types that are included with Office 365. These are listed in the Security and Compliance Center under Classification.

The next page of this topic includes a list of current sensitive information types that apply to citizens in the European Union. Use these as a starting point. Check back frequently for new additions that can help with GDPR compliance.

Also see this article: List of sensitive information types and what each one looks for.

Sensitive information types

Defines how the automated process recognizes specific information types such as bank account numbers, health service numbers, and credit card numbers. Sensitive information types are also referred to as conditions. A sensitive information type is defined by a pattern that can be identified by a regular expression or a function. In addition, corroborative evidence such as keywords and checksums can be used to identify a sensitive information type. Confidence level and proximity are also used in the evaluation process.

At this time sensitive information types cannot be used to find data at rest in mailboxes.

Continued on next page

Using Content Search with sensitive information types

1 Go to Content Search in the Security and Compliance Center

In the left pane of the Security & Compliance Center, click **Search & investigation** > **Content search**.

Run a Content Search in the Office 365 Security & Compliance Center

2 Create a new search item for each sensitive information type

Use the following syntax:
SensitiveType:"<type>"
For example:
SensitiveType:"France Passport Number"

Scope the search to SharePoint (includes OneDrive for Business). Make sure the syntax is exact and there are no extra spaces or typos.

Form a query to find sensitive data stored on sites

3 Review the results for each search

Look for these types of issues to determine if the query accuracy is on target:

- Many false positives
- Missing known instances of data

Export Content Search results from the Office 365 Security & Compliance Center

Note: if you're using Mozilla Firefox or Chrome, you might need to first download reports using Internet Explorer or Edge in order to install the required add-in.

Sensitive information types for EU citizen data

Note: Many more sensitive information types are coming soon for personal data in EU countries.

	Sensitive information type	Category
Customer Data	Belgium National Number	Personally Identifiable Information (PII)
	Credit Card Number	Personal Financial Information
	Croatia Identity Card Number	Personally Identifiable Information (PII)
	Croatia Personal Identification (OIB) Number	Personally Identifiable Information (PII)
	Czech National Identity Card Number	Personally Identifiable Information (PII)
	Denmark Personal Identification Number	Personally Identifiable Information (PII)
	EU Debit Card Number	Personal Financial Information
	Finland National ID	Personally Identifiable Information (PII)
	Finland Passport Number	Personally Identifiable Information (PII)
	France Driver's License Number	Personally Identifiable Information (PII)
	France National ID Card (CNI)	Personally Identifiable Information (PII)
	France Passport Number	Personally Identifiable Information (PII)
	France Social Security Number (INSEE)	Personally Identifiable Information (PII)
	German Driver's License Number	Personally Identifiable Information (PII)
	Germany Identity Card Number	Personally Identifiable Information (PII)
	German Passport Number	Personally Identifiable Information (PII)
	Greece National ID Card	Personally Identifiable Information (PII)
	International Banking Account Number (IBAN)	Personal Financial Information
	IP Address	Personally Identifiable Information (PII)
	Ireland Personal Public Service (PPS) Number	Personally Identifiable Information (PII)
	Italy's Driver's License Number	Personally Identifiable Information (PII)
	Netherlands Citizen's Service (BSN) Number	Personally Identifiable Information (PII)
	Norway Identity Number	Personally Identifiable Information (PII)
	Poland Identity Card	Personally Identifiable Information (PII)
	Poland National ID (PESEL)	Personally Identifiable Information (PII)
	Poland Passport	Personally Identifiable Information (PII)
	Portugal Citizen Card Number	Personally Identifiable Information (PII)
	Spain Social Security Number (SSN)	Personally Identifiable Information (PII)
	Sweden National ID	Personally Identifiable Information (PII)
	Sweden Passport Number	Personally Identifiable Information (PII)
	U.K. Driver's License Number	Personally Identifiable Information (PII)
	U.K. Electoral Roll Number	Personally Identifiable Information (PII)
	U.K. National Health Service Number	Personal Health Information
	U.K. National Insurance Number (NINO)	Personally Identifiable Information (PII)
	U.S./U.K. Passport Number	Personally Identifiable Information (PII)

Continued on next page

Add parameters to a sensitive information type query to hone the results

You can add these parameters to a sensitive information type query:

- Count range — define the number of occurrences of sensitive information a document needs to contain before it's included in the query results.
- Confidence range — the level of confidence that the detected sensitive type is actually a match, such as 85 (85%).

You can also use properties, and operators to illustrate how you can refine your queries. For more information and examples, see Form a query to find sensitive data stored on sites.

SensitiveType:"<type>|<count range>|<confidence range>"

SensitiveType:"Credit Card Number|5"
(return only documents that contain exactly five credit card numbers)

SensitiveType:"Credit Card Number|*|85.."
(confidence range is 85 percent or higher)

Note: "SensitiveType" is case sensitive, but the rest of the query is not.

Modify a sensitive information type to improve accuracy

If you're still not returning the expected results or the query returns too many false positives, consider modifying the sensitive information type to work better with your environment.

The best practice when creating or customizing a sensitive information type is to create a new sensitive information type based on an existing one, giving it a unique name and identifiers. For example, if you wish to adjust the parameters of the "EU Debit Card Number" sensitive information type, you could name your copy of that rule "EU Debit Card Enhanced" to distinguish it from the original.

In your new sensitive information type, simply modify the values you wish to change to improve its accuracy. Once complete, you will upload your new sensitive information type and create a new DLP rule (or modify an existing one) to use the new sensitive information type you just added. Modifying the accuracy of sensitive information types could require some trial and error, so maintaining a copy of the original type allows you to fall back to it if required in the future.

See:

Customize a built-in sensitive information type

Customize or create new sensitive information types for GDPR for detailed steps. <this article coming soon>

1. Export the existing Microsoft Rule Package of built in sensitive information types in Office 365
2. Rename this XML file and open it in your favorite XML editor
3. Isolate the sensitive information type and remove all others
4. Use PowerShell to generate two new GUIDs for the sensitive information type you are modifying
5. Modify the ID and other basic elements so the sensitive information type is unique (this includes replacing two GUIDs with the new ones you generated)
6. Tune the match requirements to improve accuracy

   Proximity modifications

   Modify the character pattern proximity to expand or shrink the window in which keywords must be found around the sensitive information type.

   Keyword modifications

   Add keywords to one of the <Keywords> element in order to provide our sensitive information type more specific corroborative evidence to search for in order to signal a match on this rule. Or remove keywords that are causing false positives.

   Confidence modifications

   Modify the confidence with which the sensitive information type must match the criteria specified in its definition before a match is signaled and reported.
7. Upload the new sensitive information type
8. Recrawl your content to identify the sensitive information

   Manually request crawling and re-indexing of a site

Also see Customize a built-in sensitive information type.

Example: modify the 'EU Debit Card Number' sensitive information type

Improving the accuracy of DLP rules in any system requires testing against a sample data set, and may require fine tuning through repetitive modifications and tests. This example demonstrates modifications to the 'EU Debit Card Number' sensitive information type to improve its accuracy.

When searching for an EU Debit Card Number in our example, the definition of that number is strictly defined as 16 digits using a complex pattern, and being subject to the validation of a checksum. We cannot alter this pattern due to the string definition of this sensitive information type. However, we can make the following adjustments in order to improve the accuracy of how Office 365 DLP finds this sensitive information type in our content within Office 365.

Proximity modifications

We'll shrink the window by modifying the patternProximity value in our <Entity> element from 300 to 150 characters. This means that our corroborative evidence, or our keywords, must be closer to our sensitive information type in order to signal a match on this rule.

```
<Entity id="48da7072-821e-4804-9fab-
72ffb48f6f78" patternsProximity="150"
recommendedConfidence="85">
```

Keyword modifications

Some keywords might cause false positives to occur. As a result you might want to remove keywords. Here are the keywords for this example::

```
<Keyword id="Keyword_card_terms_dict">
    <Group>
      <Term>corporate card</Term>
      <Term>organization card</Term>
      <Term>acct nbr</Term>
      <Term>acct num</Term>
      <Term>acct no</Term>
      ...
    </Group>
</Keyword>
```

Confidence modifications

If you remove keywords from the definition, you would typically want to adjust how confident you are that this sensitive information type was found by lowering this value. The default level for EU Debit Card Number type is 85.

```
<Entity id="48da7072-821e-4804-9fab-
72ffb48f6f78" patternsProximity="150"
recommendedConfidence="85">
    <Pattern confidenceLevel="85">
    ...
     </Pattern>
</Entity>
```

Continued on next page

Create custom KQL queries to find additional data in your environment

You might need to create additional queries to find personal data that is subject to GDPR. Content Search uses Keyword Query Language (KQL) to find data. Most sensitive data can't be accurately detected using just KQL without sensitive information types. So the goal is to test and optimize KQL strings using Content Search and then use these to create and tune new sensitive information types where you can achieve even greater accuracy.

Use these resources to formulate and optimize queries using KQL:

Keyword Query Language (KQL) syntax reference (DMC)

Run a Content Search in the Office 365 Security & Compliance Center

Content Search provides another resource to help you develop KQL queries and sensitive information types — keywords. Why use the keyword list? You can get statistics that show how many items match each keyword. This can help you quickly identify which keywords are the most (and least) effective. For more information about search statistics, see View keyword statistics for Content Search results.

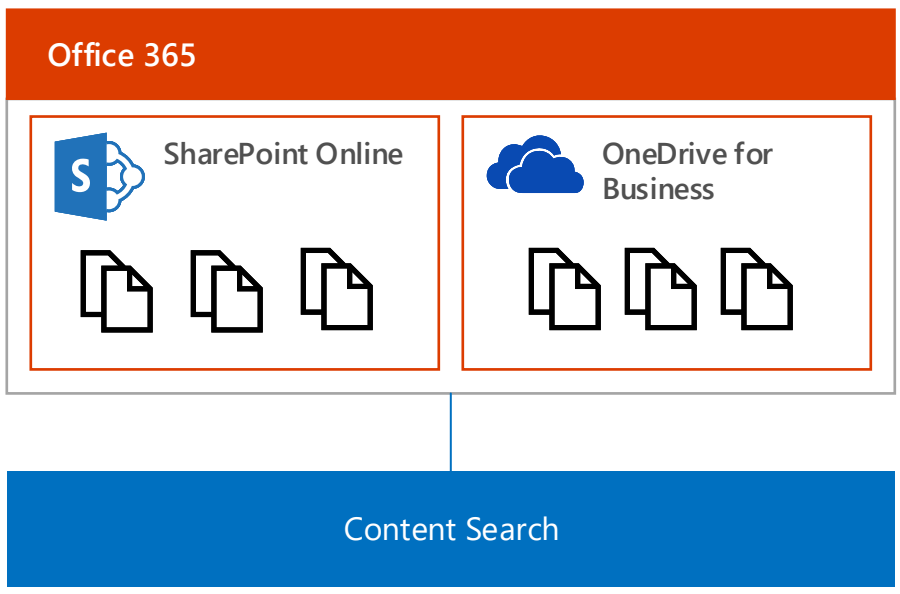
Keywords on each row are connected by the OR operator in the search query that's created. You can also use a keyword phrase (surrounded by parentheses) in a row.

Keyword queries and search conditions for Content Search

Example — Using Content Search to identify email addresses

Email addresses are considered sensitive information related to data subjects. This is a simple example to demonstrate how Content Search can help.

KQL and keywords can't be used together. Use these tools separately to hone your query and determine keywords that might be useful in sensitive information types.



KQL Query

```
(^|\b)([a-zA-Z0-9_\-\.]+)@([a-zA-Z0-9_\-\.]+)\.([a-zA-Z]{2,5})($|\b)
```

Notes:

- You can use NEAR and ONEAR for proximity searches.
- Unfortunately KQL doesn't support queries with the Regex Class (ex: IdRef="Regex_email_address")

— OR —

Keywords

Keywords
email address
mail
contact
sender
recipient
cc
bcc

In this example, you might learn the keywords are not necessary and produce a lot of false positive results.

Create new custom sensitive information types

After using KQL queries and keywords to identify sensitive information, use these to create new custom sensitive information types. In many cases, you'll require the sophistication of sensitive information types to achieve the right level of accuracy. You can then use these custom sensitive information types with Content Search, in DLP policies and other tools, and within other KQL queries.

The best practice is to create a new sensitive information type based on an existing one. Use the same process described earlier in this topic.

We'll use email address as an example because it's simple to illustrate.

Example modifications for a new email sensitive information type

1 Set the IdRef property

Within the <Entity> element, modify the <IdMatch> element so that its idRef property is = to a unique value, for example IdRef="Regex_email_address". This value will point to an element that defines our regular expression to find email addresses.

IdRef="Regex_email_address"

2 Proximity attribute

We'll start with a patternProximity value in our <Entity> element of 300.

patternsProximity="300"

3 Confidence level

Set the recommendedConfidence property to a value you feel will represent the confidence of finding an accurate match. This will likely require testing with a representative data set to get an accurate result. As an initial setting, set this value to 75.

recommendedConfidence="75">

Continued on next page

Entity element

The resulting XML for these first three elements combined looks like this:

```
<Entity id="42e6348e-27f0-4774-9604-d470cb3e219a" patternsProximity="300"
recommendedConfidence="75">
    <Pattern confidenceLevel="75">
      <IdMatch idRef="Regex_email_address" />
      <Any minMatches="1">
        <Match idRef="Keyword_email_terms" />
      </Any>
    </Pattern>
  </Entity>
```

4 Regex element

Add a new <Regex> element immediately be below the <Entity> elements that defines the regular expression used to identify email addresses. This element should appear as follows:

```
<Regex id="Regex_email_address">(^|\b)([a-zA-Z0-9_\-\.]+)@([a-zA-
Z0-9_\-\.]+)\.([a-zA-Z]{2,5})($|\b)</Regex>
```

5 Keywords

Add a new <Keyword> element below the <Regex> element that defines list of email address related keywords. Ensure that the id value for the <Keyword> element matches the <Match idRef> value in the <Entity><Pattern> element. You may continue to add your own keywords if needed.

Keywords are likely not necessary to include in an email sensitive information type. These are provided as an example.

```
<Keyword id="Keyword_email_terms">
  <Group>
    <Term>email</Term>
                <Term>email address</Term>
                <Term>contact</Term>
  </Group>
</Keyword>
```

6 LocalizedStrings element

In the <LocalizedStrings><Resource> element ensure that you have a unique name that identifies your sensitive information type, as follows:

```
  <LocalizedStrings>
    <Resource idRef="42e6348e-27f0-4774-9604-d470cb3e219a">
      <Name default="true" langcode="en-us">Email Address</
Name>
      <Description default="true" langcode="en-us">Detects email
addresses.</Description>
    </Resource>
  </LocalizedStrings>
```

Additional example of using KQL and creating a custom sensitive information type

Contoso uses a Contoso Customer Number (CCN) to identify each customer in their customer database. A CCN consists of the following taxonomy:

- Two digits to represent the year that the record was created. Contoso was founded in 2002; therefore, the earliest possible value would be 02
- Three digits to represent the partner agency that created the record. Possible agency values range from 000 to 999.
- An alpha character to represent the line of business. Possible values are a-z and should be case insensitive.
- A four digit serial number. Possible serial number values range from 0000 to 9999.

Contoso always refers to customers by using a CCN in internal correspondence, external correspondence, documents, etc. They would like to create a custom sensitive information type to detect the use of CCN in Office 365 so that they may apply protection to the use of this form of personal data.

Example CCNs: 15080P9562
14040O1119
15020J8317
14050E2330
16050E2166
17040O1118

1 Contoso uses PowerShell and Content Search to find documents that match an example set of CCNs

```
#Connect to Office 365 Security & Compliance Center
$adminUser = "alland@contoso.com"
Connect-IPPSSession -UserPrincipalName $adminUser

#Create & start search for sample data
$searchName = "Sample Customer Information Search"
$searchQuery = "15080P9562 OR 14040O1119 OR 15020J8317 OR 14050E2330 OR 16050E2166 OR 17040O1118"
New-ComplianceSearch -Name $searchName -SharePointLocation All -ExchangeLocation All -ContentMatchQuery $searchQuery
Start-ComplianceSearch -Identity $searchName
```

KQL Query

2 Contoso analyzes the results

Every time the CCN was used, an EU formatted date was used and one of the following keywords were also used within a proximity of 300 characters:

customer number, customer no, customer #, customer#, Contoso customer

3 Contoso developed the following Regular Expression (RegEx) pattern to identify their CCN

```
[0-1][0-9][0-9]{3}[A-Za-z][0-9]{4}
```

4 Contoso developed the following Regular Expression (RegEx) pattern to identify EU dates in the formats used by their various subsidiaries.

```
(0?[1-9]|[12][0-9]|3[0-1])[\/-](0?[1-9]|1[0-2]|j\x00e4n(uar)?|jan(uary|uari|uar|eiro|vier|v)?|ene(ro)?|genn(aio)?|
feb(ruary|ruari|rero|braio|ruar|br)?|f\x00e9vr(ier)?|fev(ereiro)?|mar(zo|o|ch|s)?|m\x00e4rz|maart|
apr(ile|il)?|abr(il)?|avril|may(o)?|magg(io)?|mai|mei|mai(o)?|jun(io|i|e|ho)?|giugno|juin|jul(y|io|i|ho)?|lu(glio)?|j
uil(let)?|ag(o|osto)?|aug(ustus|ust)?|ao\x00fbt|sep|sept(ember|iembre|embre)?|sett(embre)?|set(embro)?|
oct(ober|ubre|obre)?|ott(obre)?|okt(ober)?|out(ubro)?|nov(ember|iembre|embre|embro)?|dec(ember)?|
dic(iembre|embre)?|dez(ember|embro)?|d\x00e9c(embre)?)[ \/-](19|20)?[0-9]{2}
```

Continued on next page

5 Contoso uses PowerShell to generate three unique GUIDs

```
#Generate a unique GUID for RulePack Id, Publisher Id, and Entity Id
[guid]::NewGuid().Guid
[guid]::NewGuid().Guid
[guid]::NewGuid().Guid
```

6 Contoso defines the following parameters for their sensitive item type rule:

Name: Contoso Customer Number (CCN)

Description: Contoso Customer Number (CCN) that looks for additional keywords and EU formatted date

7 Contoso creates an XML file for a new sensitive information type to detect a Contoso Customer Number (CCN) and saves this to a local file system as C:\Scripts\ContosoCCN.xml in with UTF-8 encoding

Ensure encoding specified here matches saved file encoding

```
<?xml version="1.0" encoding="utf-8"?>
<RulePackage xmlns="http://schemas.microsoft.com/office/2011/mce">
<RulePack id="130ae63b-a91e-4a12-9e02-a90e36a83d7f">
<Version major="1" minor="0" build="0" revision="0" />
<Publisher id="47148982-defd-42a1-890a-7b9472099f1f" />
<Details defaultLangCode="en">
<LocalizedDetails langcode="en">
<PublisherName>Contoso Ltd.</PublisherName>
<Name>Contoso Rule Package</Name>
<Description>Defines Contoso's custom set of classification rules</Description>
</LocalizedDetails>
</Details>
</RulePack>
<Rules>
<!-- Contoso Customer Number (CCN) -->
<Entity id="a91f9a2e-6cfc-4622-8c5d-954875aa5b2b" patternsProximity="300"
recommendedConfidence="85">
<Pattern confidenceLevel="85">
<IdMatch idRef="Regex_contoso_ccn" />
<Match idRef="Keyword_contoso_ccn" />
<Match idRef="Regex_eu_date" />
</Pattern>
</Entity>
<Regex id="Regex_contoso_ccn">[0-1][0-9][0-9]{3}[A-Za-z][0-9]{4}</Regex>
<Keyword id="Keyword_contoso_ccn">
<Group matchStyle="word">
<Term caseSensitive="false">customer number</Term>
<Term caseSensitive="false">customer no</Term>
<Term caseSensitive="false">customer #</Term>
<Term caseSensitive="false">customer#</Term>
<Term caseSensitive="false">Contoso customer</Term>
</Group>
</Keyword>
<Regex id="Regex_eu_date"> (0?[1-9]|[12][0-9]|3[0-1])[\/-](0?[1-9]|1[0-2]|j\
x00e4n(uar)?|jan(uary|uari|uar|eiro|vier|v)?|ene(ro)?|genn(aio)? |feb(ruary|ruari|rero|braio|ruar|br)?|f\
x00e9vr(ier)?|fev(ereiro)?|mar(zo|o|ch|s)?|m\x00e4rz|maart|apr(ile|il)?|abr(il)?|avril
|may(o)?|magg(io)?|mai|mei|mai(o)?|jun(io|i|e|ho)?|giugno|juin|jul(y|io|i|ho)?|lu(glio)?|juil(let)?|ag(o|
osto)?|aug(ustus|ust)?|ao\
x00fbt|sep|sept(ember|iembre|embre)?|sett(embre)?|set(embro)?|oct(ober|ubre|obre)?|ott(obre)?|o
kt(ober)?|out(ubro)?
|nov(ember|iembre|embre|embro)?|dec(ember)?|dic(iembre|embre)?|dez(ember|embro)?|d\
x00e9c(embre)?)[ \/-](19|20)?[0-9]{2}</Regex>
<LocalizedStrings>
<Resource idRef="a91f9a2e-6cfc-4622-8c5d-954875aa5b2b">
<Name default="true" langcode="en-us">Contoso Customer Number (CCN)</Name>
<Description default="true" langcode="en-us">Contoso Customer Number (CCN) that looks for
additional keywords and EU formatted date</Description>
</Resource>
</LocalizedStrings>
</Rules>
</RulePackage>
```

RulePack GUID from Step 5

Publisher GUID from Step 5

Entity GUID from Step 5

IdMatch

CCN Regex from Step 3

IdMatch

EU date Regex from Step 4

Entity GUID from Step 5

Name from Step 6

Description from Step 6

8 Contoso creates the custom sensitive information type with the following PowerShell:

```
#Connect to Office 365 Security & Compliance Center
$adminUser = "alland@contoso.com"
Connect-IPPSSession -UserPrincipalName $adminUser

#Create new Sensitive Information Type
New-DlpSensitiveInformationTypeRulePackage -FileData (Get-Content -Path "C:\Scripts\
ContosoCCN.xml" -Encoding Byte -ReadCount 0)
```

Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

Architect a classification schema for personal data

Up to this point we've focused on using sensitive information types to identify personal data that is subject to GDPR. Sensitive information types are a form of classification. This might be all the classification you need. However, many organizations implement a broader data governance strategy using labels. Use this topic to decide if you also want to implement labels as part of your GDPR plan. If you do, this topic provides some guidance and examples.

Note: Defining a classification schema for an organization and configuring policies, labels, and conditions requires careful planning and preparation. It is important to realize that this is not an IT driven process. Be sure to work with your legal and compliance team to develop an appropriate classification and labeling schema for your organization's data.

Decide if you are using labels in addition to sensitive data types

You can take one of two approaches for classification in Office 365 for personal information. Either of these can be used for GDPR protection.

If you're using only sensitive information types for classification, you can skip the rest of this topic.

Use only Office 365 sensitive information types

- Sensitive information types work well to identify and protect personal data subject to GDPR and other types of regulations.
- These are simpler to use if your organization doesn't already have or plan to implement a broader data governance plan using labels.
- These work with DLP rules (so do Office labels).
- In the future these will work with Cloud App Security so you can detect sensitive information in other SaaS apps.

— OR —

Use sensitive information types + Office labels

- You'll need sensitive information types to automatically apply labels to personal data that is subject to GDPR, so these are a prerequisite.
- Using Office labels allows you to include personal data that is subject to GDPR into a broader data governance plan for your organization.
- Later, Office labels will converge with Azure Information Protection labels into a unified classification and labeling engine.

Develop a label schema that includes personal data

Before using technical capabilities to apply labels and protection, first work across your organization to define a classification schema. Your organization might already have a classification schema, which makes it easier to add personal data. This topic includes an example classification schema. You can use this as a starting point, if needed.

Getting started

Begin by deciding on the number and names of labels to implement. Do this activity without worrying about which technology to use and how labels will be applied. Apply this schema universally throughout your organization, including data that resides on premises and in other cloud services.

Recommendations

When designing and implementing policies, labels, and conditions, consider following these recommendations:

- **Use existing classification schema (if any)** — Many organizations already are using data classification in some form. Carefully evaluate the existing label schema and if possible use it as is. Using familiar labels that are recognizable to the end-user will drive adoption.
- **Start with default policies and labels** — All solutions come with a set of predefined policies and labels. Carefully evaluate these against the organizations legal and business requirements and consider using them instead of creating new ones.
- **Start small** — There is virtually no limit to the number of labels that can be created. However, large numbers of labels and sub-labels will negatively impact the adoption. Too many choices often means no choice at all.
- **Use scenarios and use cases** — Identify common use cases within the organization and use scenarios derived from the GDPR to verify if the envisioned label and classification configuration will work in practice.

Example classification schema

Label name	Description
Personal	Non-business data, for personal use only.
Public	Business data that is specifically prepared and approved for public consumption.
Customer data	Business data that contains personal identifiable information. Examples are credit card numbers, bank account numbers, and social security numbers.
HR data	Human Resource data about Contoso employees, such as employee number and salary data.
Confidential	Sensitive business data that could cause damage to the business if shared with unauthorized people. Examples include contracts, security reports, forecast summaries, and sales account data.
Highly confidential	Very sensitive business data that would cause damage to the business if it was shared with unauthorized people. Examples include employee and customer information, passwords, source code, and pre-announced financial reports.

- **Question every request for a new label**, does every scenario or use case really need a new label or can we use what we already have? Keeping the number of labels to a minimum improves adoption.
- **Use sub-labels for key departments**, some departments will have specific needs that require specific labels. Define these labels as sub-labels to an existing label and consider using scoped policies that are assigned to user groups instead of globally.
- **Consider scoped policies**, polices targeted at subsets of users will prevent "label overload". A scoped policy enables assigning role or department specific (sub-)labels to just employees that work for that specific department.
- **Use meaningful label names**, it is recommended not to use jargon, standards or acronyms as label names. Try to use names that resonate with the end user to improve adoption. Instead of using labels like PII, PCI, HIPAA, LBI, MBI and HBI consider names like Non-Business, Public, General, Confidential and Highly Confidential.

Continued on next page

Define a taxonomy and search criteria for each label

After developing a classification schema for your organization, the next step is to develop the taxonomy and search criteria for finding this data. For personal data, you've already completed this work by identifying sensitive information types and also by customizing or creating new sensitive information types for your environment.

The following table provides an example schema, taxonomy, and search criteria for an organization. The labels are ordered by sensitivity level from least sensitive to most sensitive to ensure that data that matches multiple label conditions is assigned the appropriate label.

Note: The configuration example is provided for illustration only and is not intended as deployment guidance or reference.

The important takeaway is to ensure that the work you invest to classify personal data for GDPR compliance fits together with the objectives for your entire organization.

Example schema, taxonomy, and search criteria

Label	Taxonomy	Method	Search syntax
Personal	Documents manually labelled personal by the end user.	Manual	Documents manually labelled personal by the end user.
Public	Documents containing the case insensitive phrase **Approved for Public Release ##/####** where # represents any digit.	KQL	Approved for Public Release*
		RegEx	(?i)(\bApproved for Public Release \d{2}\/\d{4}\b)
Customer data	Sensitive information types for EU citizen data. Custom sensitive information types for additional personally identifiable data.	Sensitive information types	
Human Resources — Employee Data	Documents that include the case sensitive employee id in the format **CONTOSO-9#####** where # represents any digit.	KQL	CONTOSO-9*
		RegEx	(\bCONTOSO-9\d{5}\b)
Human Resources — Salary Data	Documents that include the keyword (not case sensitive) **Contoso** AND either keyword (not case sensitive) **Salary** OR **Compensation**	KQL	Contoso AND (Salary OR Compensation)
		RegEx	(\bCONTOSO-9\d{5}\b)
Confidential	Documents containing the phrase (not case sensitive) **Contoso Confidential**.	KQL	Contoso Confidential
		RegEx	(?i)(\bContoso Confidential\b)
Highly confidential	Documents that include either pharase (case sensitive) **Contoso Secret** or **Secret-C####** where # represents any digit.	KQL	Contoso Secret OR Secret-C*
		RegEx	(?i)(\bContoso Secret\b)|(\bSecret-C\d{4}\b)

Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

Apply labels to personal data in Office 365

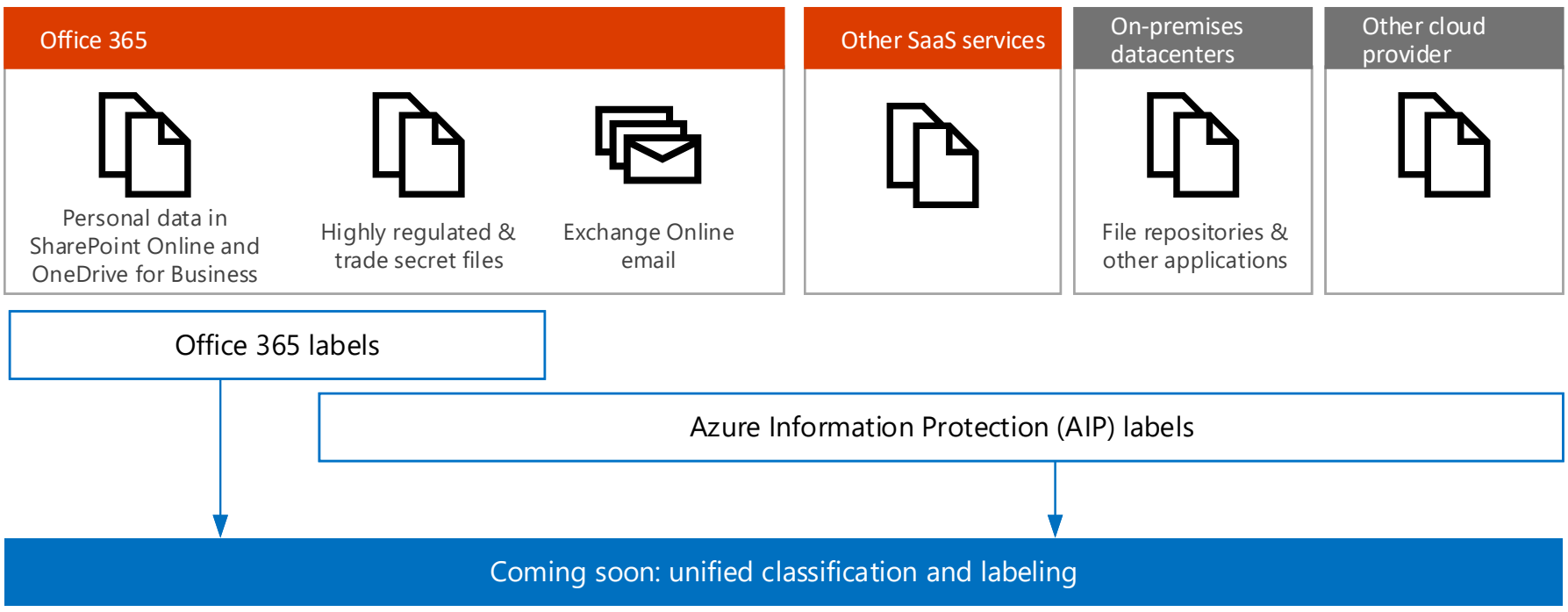
Use this topic if you are using Office labels as part of your GDPR protection plan. Today labels can be created in the Office 365 Security & Compliance Center and in Azure Information Protection. Over time these technologies will converge into a unified labeling and classification experience and you will be able to achieve even more.

If you are using labels for protection of personal data in Office 365, Microsoft recommends you start with Office labels. You can use Advanced Data Governance to automatically apply labels based on sensitive information types or other criteria. You can use Office labels with data loss prevention to apply protection. You can also use labels with eDiscovery and Content Search. You'll soon be able to use both labels and sensitive information types with Cloud App Security to monitor personal data that resides in other SaaS apps.

Azure Information Protection labels are currently recommended for applying labels to files on premises and in other cloud services and providers. These are also recommended for files in Office 365 that require Azure Rights Management (Azure RMS) encryption for data protection, such as trade secret files.

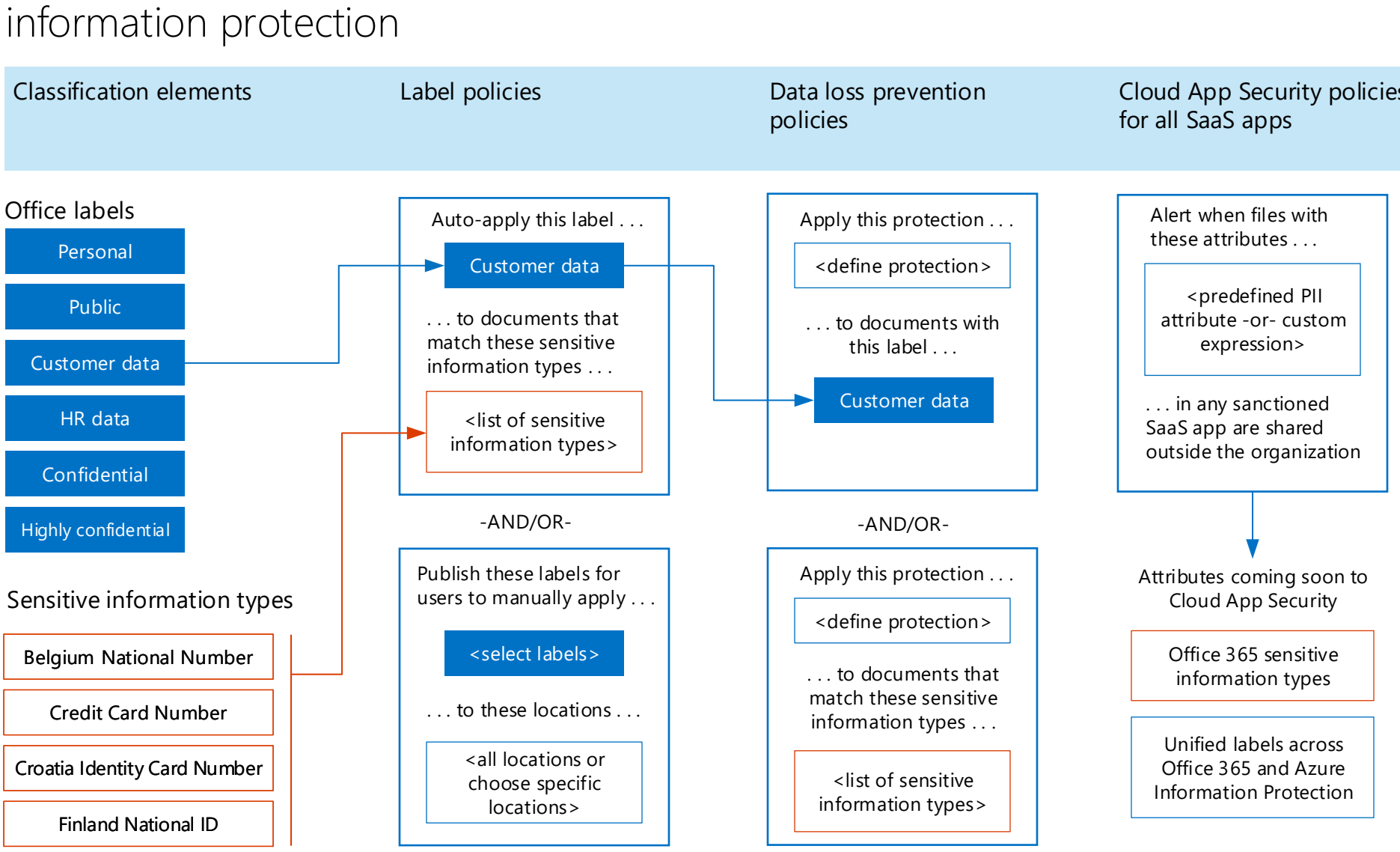
At this time, using Azure Information Protection to apply Azure RMS encryption is not recommended for files in Office 365 with data that is subject to the GDPR. Office 365 services currently cannot read into RMS-encrypted files. Therefore, the service can't find sensitive data in these files.

Azure Information Protection labels can be applied to mail in Exchange Online and these labels work with Office 365 data loss prevention. Coming soon with the unified classification and labeling experience you will be able to use the same labels for email and files, including automatically labeling and protecting email in transit.



Use Office labels and sensitive information types across Microsoft 365 for information protection

Classification elements	Label policies	Data loss prevention policies	Cloud App Security policies for all SaaS apps



Continued on next page

Prioritize auto-apply label policies

For personal data that is subject to GDPR, Microsoft recommends auto-applying labels by using the sensitive information types you curated for your environment. It is important that auto-apply label policies are well designed and tested to ensure the intended behavior occurs.

The order that auto-apply policies are created and whether users are also applying these labels affect the result. So it is important to carefully plan the roll-out. Here's what you need to know.

1 label at a time

You can only assign one label to a document.

Older auto-apply policies win

If there are multiple rules that assign an auto-apply label and content meets the conditions of multiple rules, the label for the oldest rule is assigned. For this reason, it is important to plan the label policies carefully before configuring them. If an organization requires a change to the priority of the label policies, they will need to delete and recreate them.

Manual user-applied labels trump auto-applied labels

Manual user applied labels trump auto-applied labels. Auto-apply policies cannot replace a label that is already applied by a user. Users can replace labels that are auto-applied.

Auto-assigned labels can be updated

Auto-assigned labels can be updated by either newer label policies or by updates to existing policies.

Be sure your plan for implementing labels includes:

- Prioritizing the order that auto-apply policies are created.
- Allowing enough time for labels to be automatically applied before rolling these out for users to manually apply. It can take up to seven days for the labels to be applied to all content that matches the conditions.

Example priority for creating the auto-apply label policies

Labels	Priority order to create auto-apply policies
Human Resources — Employee Data	1
Customer Data	2
Highly Confidential	3
Human Resources — Salary Data	4
Confidential	5
Public	6
Personal	No auto-apply policy

Create labels and auto-apply label policies

Create labels and policies in Security and Compliance Center

1 Give permissions to members of your compliance team

Members of your compliance team who will create labels need permissions to use the Security & Compliance Center. Go to Permissions in Security and Compliance Center and modify the members of the Compliance Administrator group.

Give users access to the Office 365 Security & Compliance Center

2 Create Office labels

Go to Classifications in Security and Compliance Center, choose Labels, and create the labels for your environment.

3 Create auto-apply policies for labels

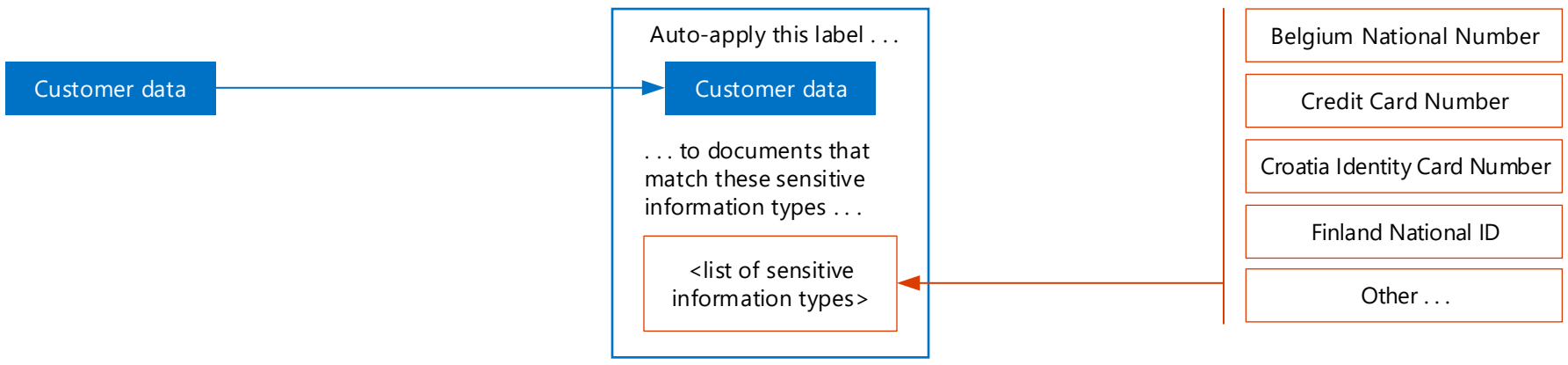
Go to Classification in Security and Compliance Center, choose Label policies, and create the policies for auto-applying labels. Be sure to create these policies in the prioritized order.

Labeling personal customer data for GDPR

Create a label for customer data

Create an auto-apply policy that assigns the label "Customer data" to any file that includes a sensitive information type

Add all of the sensitive information types you curated for your environment for GDPR



Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

Apply protection to sensitive data in Office 365

Protection of personal information in Office 365 includes using data loss prevention capabilities. With data loss prevention (DLP) policies in the Office 365 Security & Compliance Center, you can identify, monitor, and automatically protect sensitive information across Office 365.

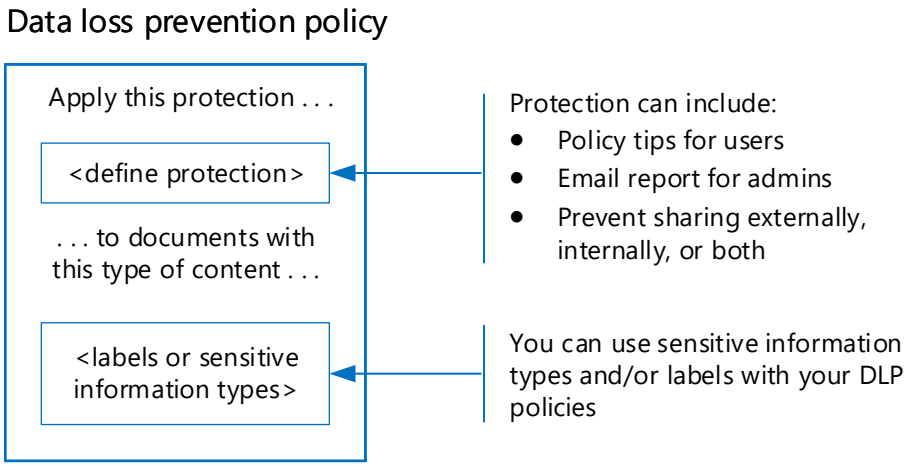
This topic describes how to use DLP to protect personal data. This topic also lists other protection capabilities that can be used to achieve GDPR compliance, including setting permissions in SharePoint libraries and using device access policies.

Apply protection using data loss prevention in Office 365

With DLP, you can:

- Identify sensitive information across many locations.
- Prevent accidental sharing of sensitive information.
- Help users learn how to stay compliant without interrupting their workflow.
- View DLP reports showing content that matches your organization's DLP policies.

For more information, see: Overview of data loss prevention policies



Using DLP for GDPR compliance

One of the primary uses of Office 365 DLP is to identify personal data related to EU data subjects in your Office 365 environment. Office 365 DLP can notify your compliance teams of where personal information is stored in SharePoint Online and OneDrive for Business, or when users send email containing personal information. DLP can also provide policy tips to your employees when working with personal information related to EU residents.

Educating and raising awareness to where EU resident data is stored in your environment and how your employees are permitted to handle it represents one level of information protection using Office 365 DLP. Often, employees who already have access to this type of information require this access to perform their day to day work. Enforcing DLP policies to help comply with GDPR may not require restricting access.

However, complying with GDPR typically involves a risk based assessment of the organization from both a legal and information security perspective, identification of what type and where personal information is stored, as well as if there is a legal justification to store and process that information. Based on this assessment, implementing policies to protect the organization and comply with GDPR might require removing access for employees to documents that contain personal information for EU data subjects. In cases where further protection is required, additional DLP protection can be configured.

The following table lists three configurations of increasing protection using DLP. The first configuration, awareness, can be used as a starting point and minimum level of protection for GDPR.

Example protection levels that can be configured with DLP policies and used for GDPR compliance

Protection level	DLP configuration for documents with personal information related to EU data subjects	Benefits and risks
Awareness	• Send email notifications to compliance teams when this data is found in documents in SharePoint Online and OneDrive for Business. • Customize and display Policy Tips to employees in SharePoint and OneDrive for Business when accessing documents containing this data. • Detect and report when this data is being shared.	• Raise awareness with compliance teams as well as employees regarding where this data is stored. • Educate employees on corporate policy for handling documents containing this data. • Does not prevent employees from sharing this data internally or externally. • You can review DLP reports for shared data and decide if you need to increase the protection.
Prevent external sharing	• Restrict access to documents that contain this data in SharePoint Online and OneDrive for Business when that content is shared with external users. • Prevent sending emails with documents that contain this data to external recipients. • Detect and report when this data is being shared.	• Prevents external sharing of this data while allowing for employees to work with this data internally. • You can review DLP reports for internally shared data and decide if you need to increase this protection.
Prevent internal and external sharing	• Restrict access to documents that contain this data in SharePoint Online and OneDrive for Business when that content is shared internally or externally. • Prevent sending emails which contain this data to both internal and external recipients.	• Prevents internal and external sharing of this data. • Employees might not be able to complete tasks that require working with this data. • You can review DLP reports for internally or externally shared data and decide if end user training is needed.

Note: As the levels of protection increase, the ability of users to access information will decrease in some cases, and could potentially impact their productivity or ability to complete day to day tasks.

Increasing protection levels by implementing policies that impact employees is typically accompanied by end user training, educating users on new security policies and procedures to help them continue to be productive in a more secure environment.

Continued on next page

Example DLP policy for GDPR — Awareness

Name: Awareness for personal data that is subject to GDPR	
Description: Display policy tips to employees, notify compliance teams when this data is found in documents in SharePoint Online and OneDrive for Business, detect and report when this data is being shared outside your organization.	
Control	**Settings**
Choose information to protect	Select a Custom policy template.
Locations	All locations in Office 365
Find content that contains	Click 'Edit' and add all the sensitive information types you curated for your environment.
Detect when this content is shared	Check this box and select 'with people outside my organization.'
Notify users when content matches the policy settings	Check this box ("Show policy tips to users and send them an email notification.") Click 'Customize the tip and email' and update these for your environment. See the default notifications in this article: Send email notifications and show policy tips for DLP policies.
Detect when a specific amount of sensitive info is being shared at one time	• 'Detect when content that's being shared contains: At least ___ instances of the same sensitive info type' — Set this to 1. • 'Send incident reports in email' — check this box. Click 'Choose what to include in the report and who receives it.' Be sure to add your compliance team. • 'Restrict who can access the content and override the policy' — clear this checkbox to receive notifications about sensitive information without preventing users from access that information.

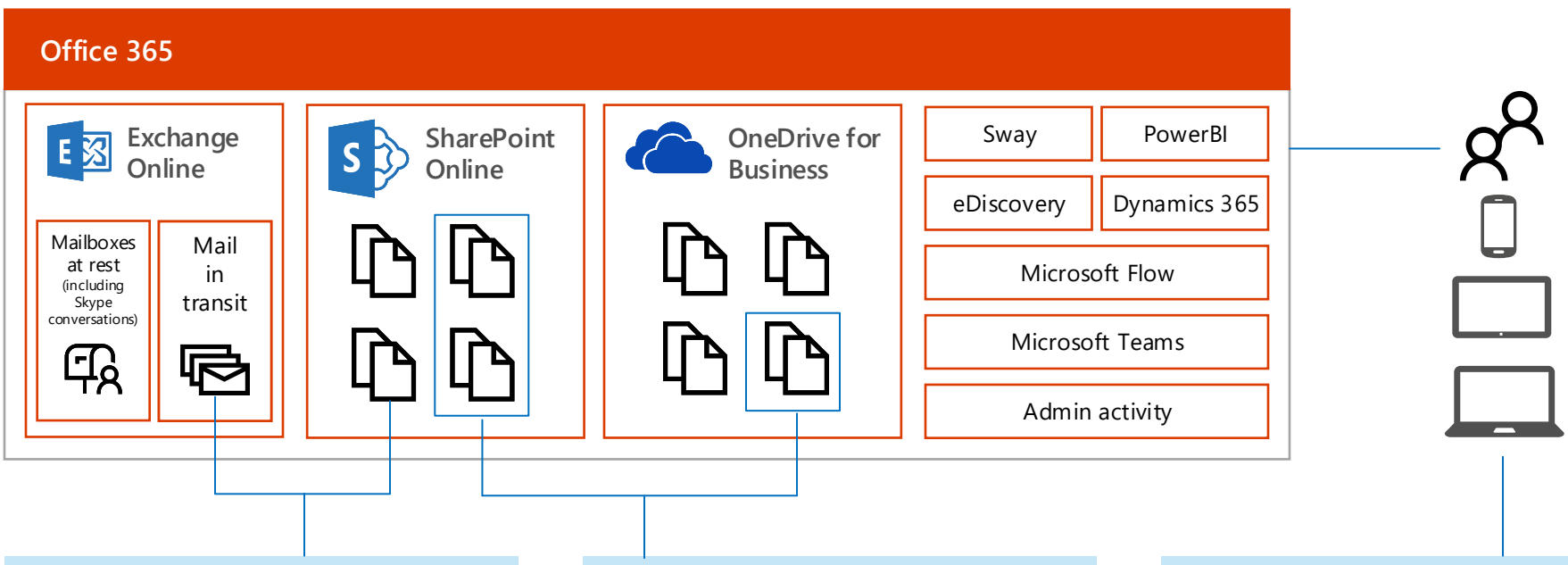
All locations includes:
- SharePoint Online
- OneDrive for Business accounts
- Exchange mailboxes

Because Content Search doesn't currently let you test sensitive information types with email, you might want to create separate policies for Exchange with a subset of sensitive information types in each policy and monitor the rollout of these policies.

Additional protection you can apply to protect personal data in Office 365

Sensitive information types, labels, and data loss protection policies help you identify documents containing specific data and apply protection. However, these protections depend on appropriate permissions being set for access to data, users with accounts that are not compromised, and devices that are healthy.

The following illustration details additional protection you can apply to protect access to personal data.



Document and email-level protection

Sensitive information types

Office labels

Data loss prevention policies

Office 365 Message Encryption for email

Site and library-level protection

Permissions for SharePoint Online and OneDrive for Business sites and libraries

External sharing policies for SharePoint Online and OneDrive for Business (site-level)

Site-level device access policies

Service access protection

Identity and device access protection in Enterprise Mobility + Security (EMS) suite

Privileged access management

Windows 10 security capabilities

Capabilities that are OK to use with GDPR

You can use the following capabilities in an environment configured for GDPR compliance. These capabilities are not necessary for GDPR compliance, but they can be used without adversely affecting your ability to discover, protect, monitor, and report on data related to GDPR compliance.

Customer Key — Allows customers to provide and retain control over the encryption keys that are used to encrypt data at rest within Office 365. Recommended only for customers with a regulatory need to manage their own encryption keys.

Customer Lockbox — Customer lockbox allows you to control how a Microsoft support engineer accesses your data, if needed, to fix a technical issue on a case by case basis. You can control whether to give the support engineer access to your data or not. An expiration time is provided with each request.

Continued on next page

Site and library-level protection

Permissions for SharePoint and OneDrive for Business libraries

Use permissions in SharePoint to provide or restrict user access to the site or its contents. Add individual users or Azure Active Directory groups to the default SharePoint groups. Or, create a custom group for finer-grain control.

More information:

[Understanding permission levels in SharePoint]

[Understanding SharePoint groups]

Full Control	Design	Edit	Contribute	Read	View Only
	Contribute + approve and customize	Contribute + add, edit and delete lists (not just list items)	View, add, update, delete list items and documents	View and download	View, no download

External sharing policies for SharePoint and OneDrive for Business libraries

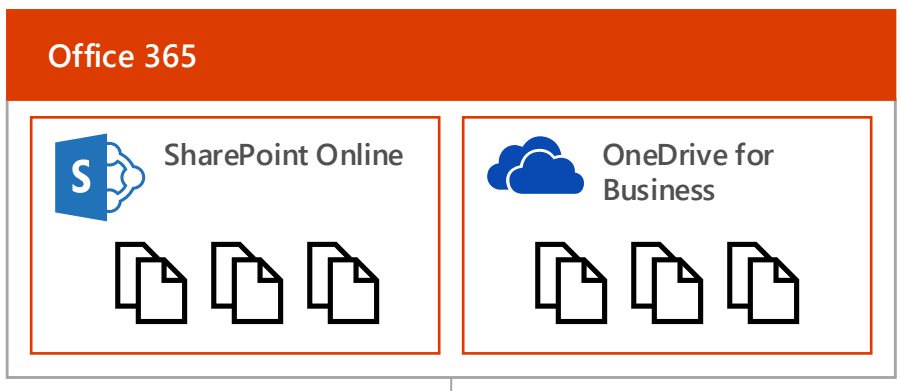
Many organizations allow external sharing to support collaboration. Find out how your tenant-wide settings are configured. Then review the external sharing settings for sites that contain personal data.

An external user is someone outside of your organization who is invited to access your SharePoint Online sites and documents but does not have a license for your SharePoint Online or Microsoft Office 365 subscription.

External sharing policies apply to both SharePoint Online and OneDrive for Business.

You must be a SharePoint Online admin to configure sharing policies.

You must be a Site Owner or have full control permissions to share a site or document with external users.



Type of sharing

- Don't allow sharing outside your organization
- Allow sharing to authenticated external users only (allow new or limit to existing)
- Allow sharing to external users with an anonymous access link
- Limit external sharing using domains (allow and deny list)
- Choose the default link type (anonymous, company shareable, or restricted)

These policies (in the blue box) can be set for individual site collections.

What external users can do

- Prevent external users from sharing files, folders, sites they don't own
- Require external users to accept sharing invitations with the same account the invitation was sent to

Notifications

Currently only available in OneDrive for Business. Notify owners when:

- Users invite additional external users to shared files
- External users accept invitations to access files
- An anonymous access link is created or changed

[Manage external sharing for your SharePoint Online environment]

[Share sites or documents with people outside your organization]

Site-level device access policies

SharePoint Online and OneDrive for Business let you configure device access policies at the site level. This lets you configure more protection for sites with sensitive data.

If you configure site-level device access policies, be sure to coordinate these with tenant-level policies and also with access policies that are configured in Azure Active Directory, Intune, and Intune App Management.

[SharePoint Online admin center: Control access from unmanaged devices]

Device access policies for SharePoint and OneDrive for Business require supporting policies in Azure Active Directory and Microsoft Intune depending on the scenario you are implementing. See the table below.

Device access scenarios and dependencies

Objective	Only allow access from specific IP address locations	Prevent users from downloading files to non-domain joined devices	Block access on non-domain joined devices	Prevent users from downloading files to non-compliant devices	Block access on non-compliant devices
SharePoint admin center	✓	✓	✓	✓	✓
Azure Active Directory		✓	✓	✓	✓
Microsoft Intune				✓	✓

Continued on next page

Service access protection for identities and devices

Microsoft recommends you configure protection for identities and devices that access the service. The work you put into protecting access to Office 365 services can also be used to protect access to other SaaS services, PaaS services, and even apps in other cloud providers.

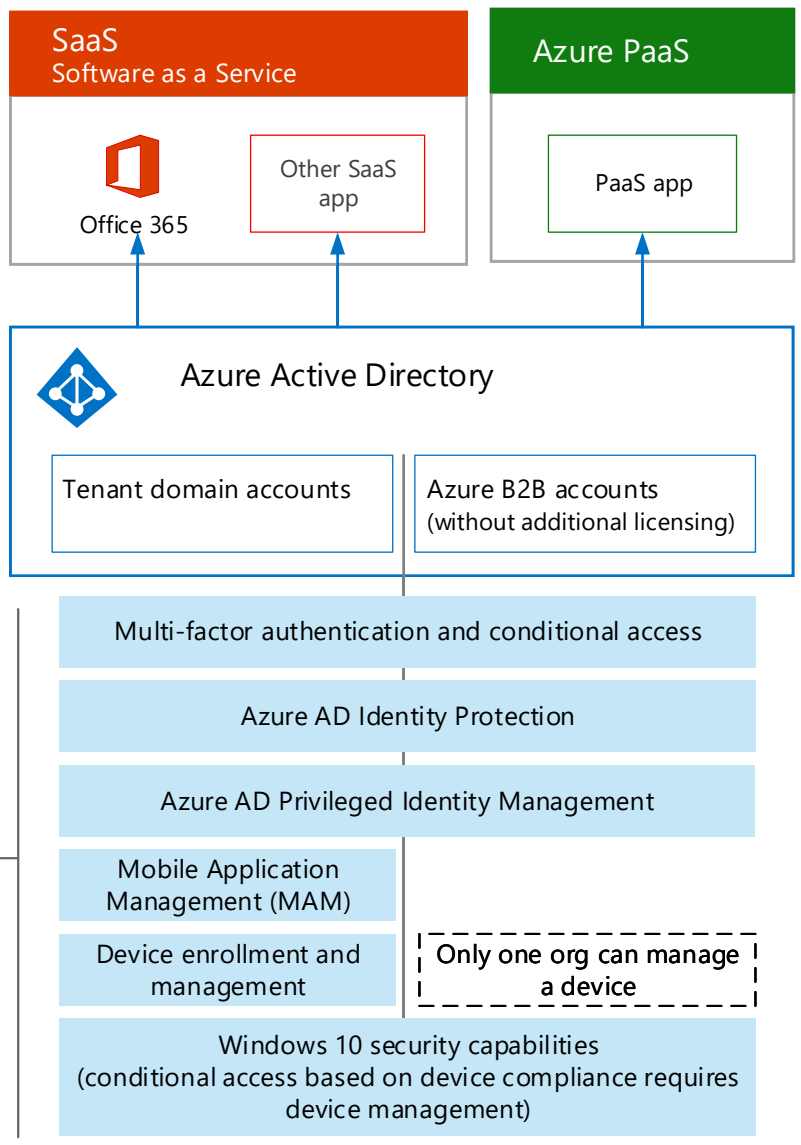
Access protection for identities and devices provides a baseline of protection to ensure that identities are not compromised, devices are safe, and organization data that is accessed on devices is isolated and protected.

For starting point recommendations and configuration guidance, see Microsoft security guidance for political campaigns, nonprofits, and other agile organizations.

For hybrid identity environments with AD FS, see Recommended security policies and configurations.

Cloud services

Azure Active Directory provides identity access to any cloud service, including non-Microsoft cloud providers such as Amazon Web Services.



Types of accounts

Tenant domain accounts — accounts you add to your tenant and manage directly.

B2B accounts — accounts for users outside your organization you invite to collaborate with. These can be other Office 365 accounts, other organization accounts, or consumer accounts (such as Gmail).

Capabilities

These capabilities protect identities and devices. This illustration shows which capabilities are available for each account type.

Capabilities in the B2B column are available without additional licensing. You can add licenses to B2B accounts to give these users additional capabilities, if needed, to protect access to personal data in your environment.

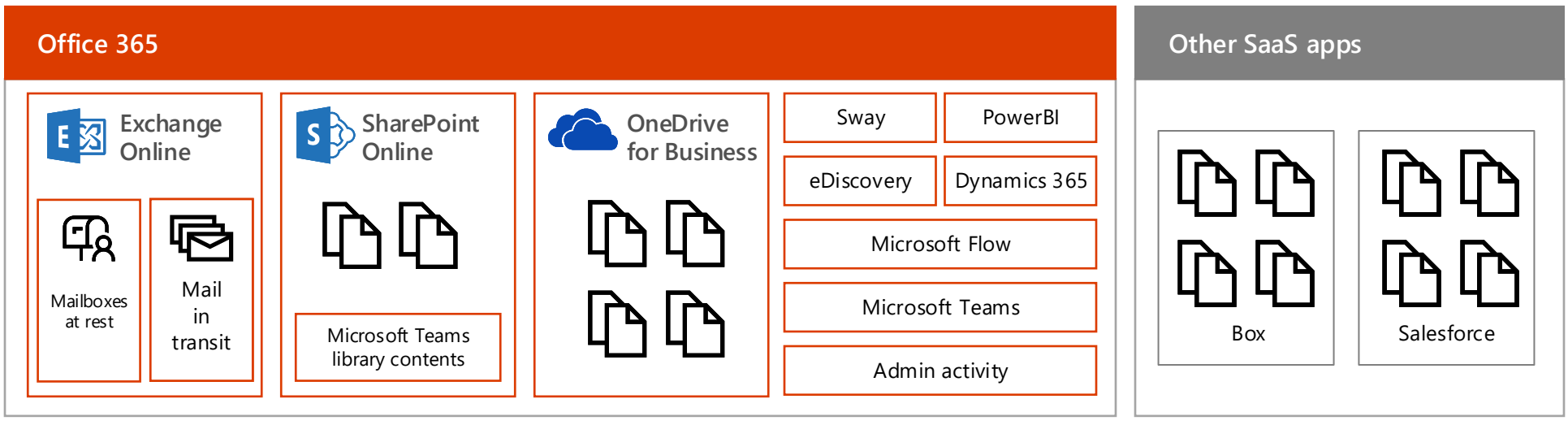
Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

Monitor for leaks of personal data

There are many tools that can be used to monitor the use and transport of personal data. This topic describes three tools that work well.

Tools recommended for monitoring personal data in Office 365 and other SaaS apps



Office 365 data loss prevention reports

After you create your data loss prevention (DLP) policies, you'll want to verify that they're working as you intended and helping you to stay compliant. With the DLP reports in Office 365, you can quickly view the number of DLP policy matches, overrides, or false positives; see whether they're trending up or down over time; filter the report in different ways; and view additional details by selecting a point on a line on the graph.

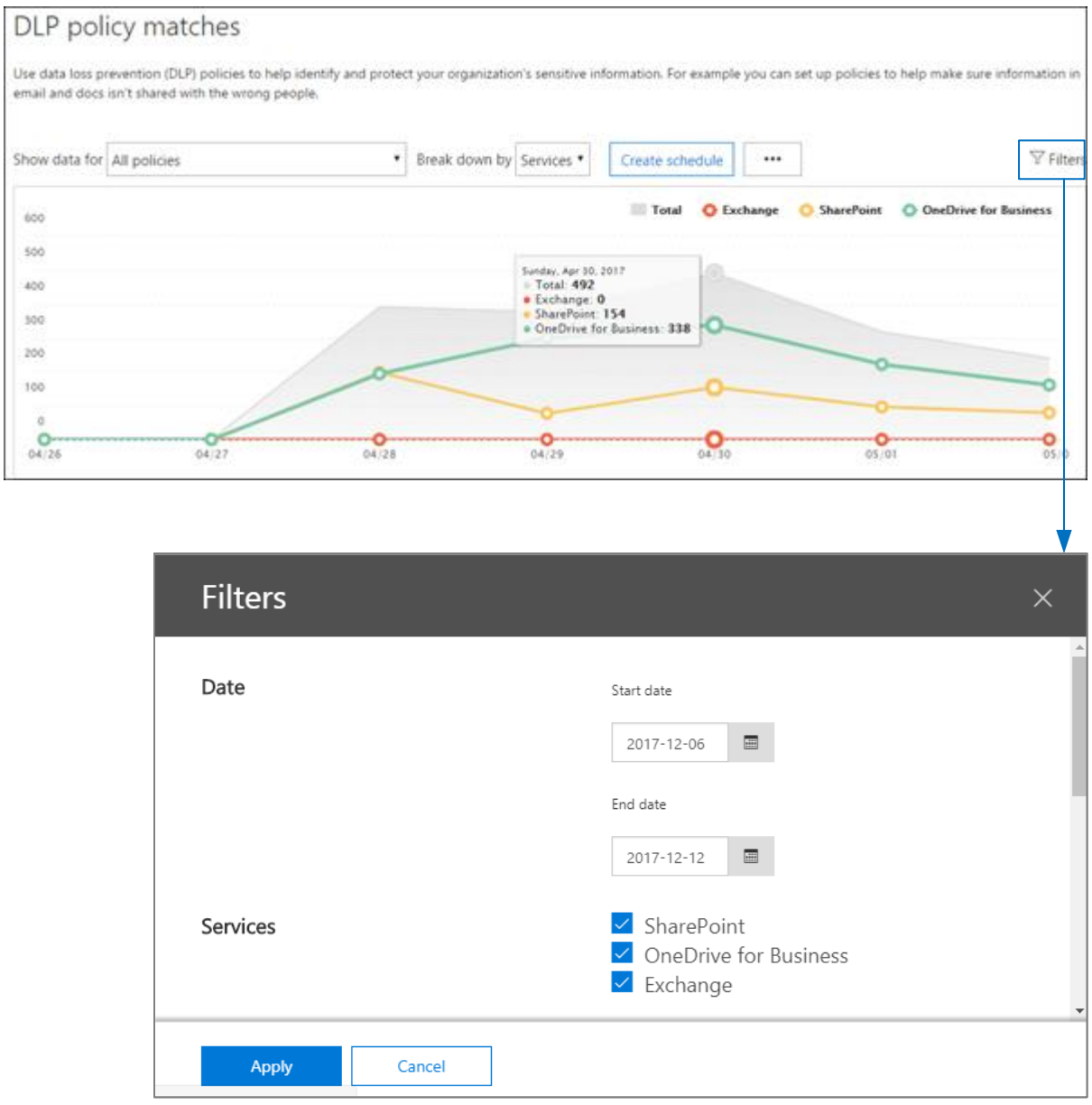
You can use the DLP reports to:

- Focus on specific time periods and understand the reasons for spikes and trends.
- Discover business processes that violate your organization's DLP policies.
- Understand any business impact of the DLP policies.
- View the justifications submitted by users when they resolve a policy tip by overriding the policy or reporting a false positive.
- Verify compliance with a specific DLP policy by showing any matches for that policy.
- View a list of files with sensitive data that matches your DLP policies in the details pane.

In addition, you can use the DLP reports to fine tune your DLP policies as you run them in test mode.

DLP reports are in Security and Compliance center. Navigate to **Reports** > **View reports**. Under **Data loss prevention (DLP)**, go to either **DLP policy and rule matches** or **DLP false positives and overrides**.

For more information, see View the reports for data loss prevention.



Continued on next page

Office 365 audit log and alert policies

The Office 365 audit log contains events from Exchange Online, SharePoint Online, OneDrive for Business, Azure Active Directory, Microsoft Teams, Power BI, Sway, and other Office 365 services.

The Office 365 Security and Compliance Center provides two ways to monitor and report against the Office 365 audit log:

- **Setup alert policies, view alerts, and monitor trends** — Use the new alert policy and alert dashboard tools in the Office 365 Security & Compliance Center. You can configure DLP rules to send alerts here, instead of sending email (requires E5).
- **Search the audit log directly** — Search for all events in a specified date rage. Or you can filter the results based on specific criteria, such as the user who performed the action, the action, or the target object.

Information security and compliance teams can use these tools to proactively review activities performed by both end users and administrators across Office 365 services. Automatic alerts can be configured to send email notifications when certain activities occur on specific site collections - for example when content is shared from sites known to contain GDPR related information. This allows those teams to follow up with users to ensure that corporate security policies are followed, or to provide additional training.

Information security teams can also search the audit log to investigate suspected data breaches and determine both root cause and the extent of the breach. This built in capability facilitates compliance with article 33 and 34 of the GDPR, which require notifications be provided to the GDPR supervisory authority and to the data subjects themselves of a data breach within a specific time period. Audit log entries are only retained for 90 days within the service - it is often recommended and many organizations required that these logs be retained for longer periods of time.

Solutions are available which subscribe to the Unified Audit Logs through the Microsoft Management Activity API and can both store log entries as needed, and provide advanced dashboards and alerts. One example is Microsoft Operations Management Suite (OMS).

More information about alert policies and searching the audit log:

Alert policies in the Office 365 Security & Compliance Center

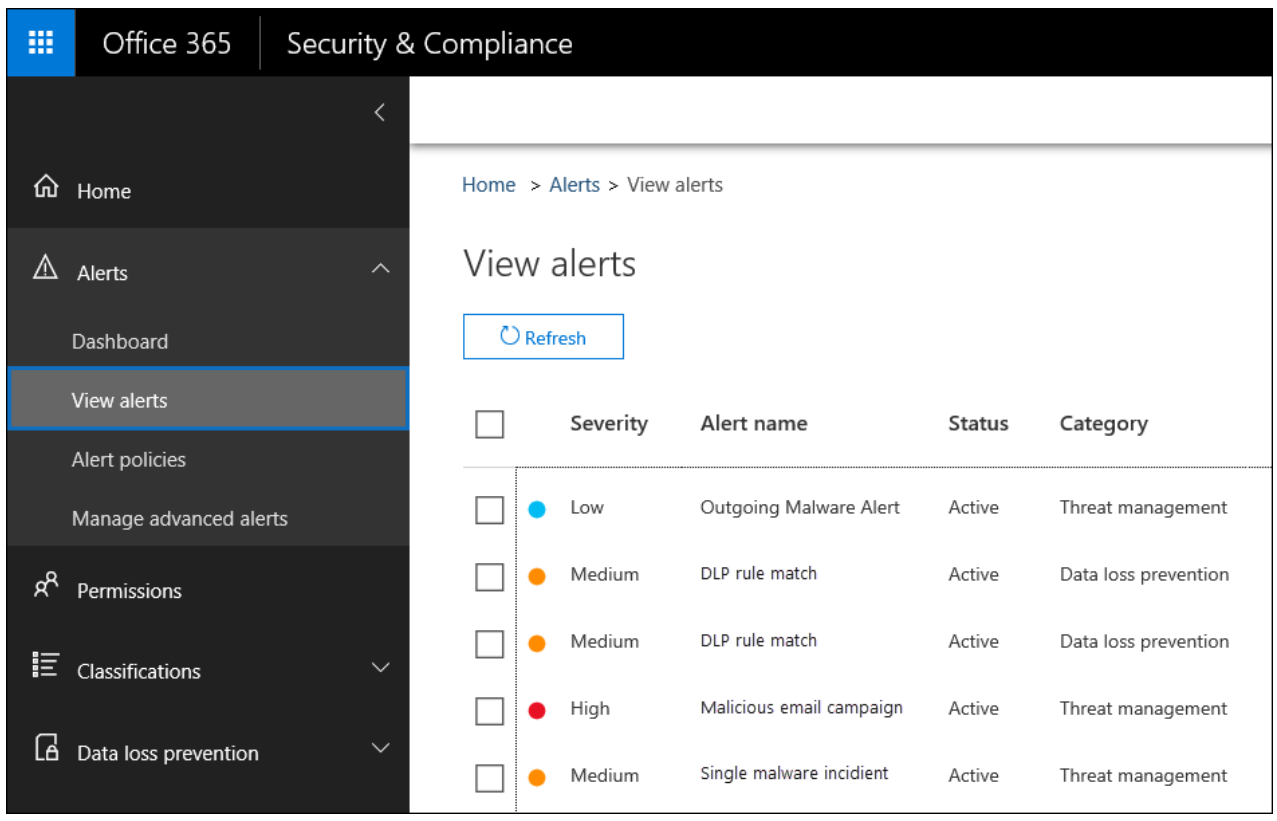
Search the audit log for user and admin activity in Office 365 (introduction)

Turn Office 365 audit log search on or off

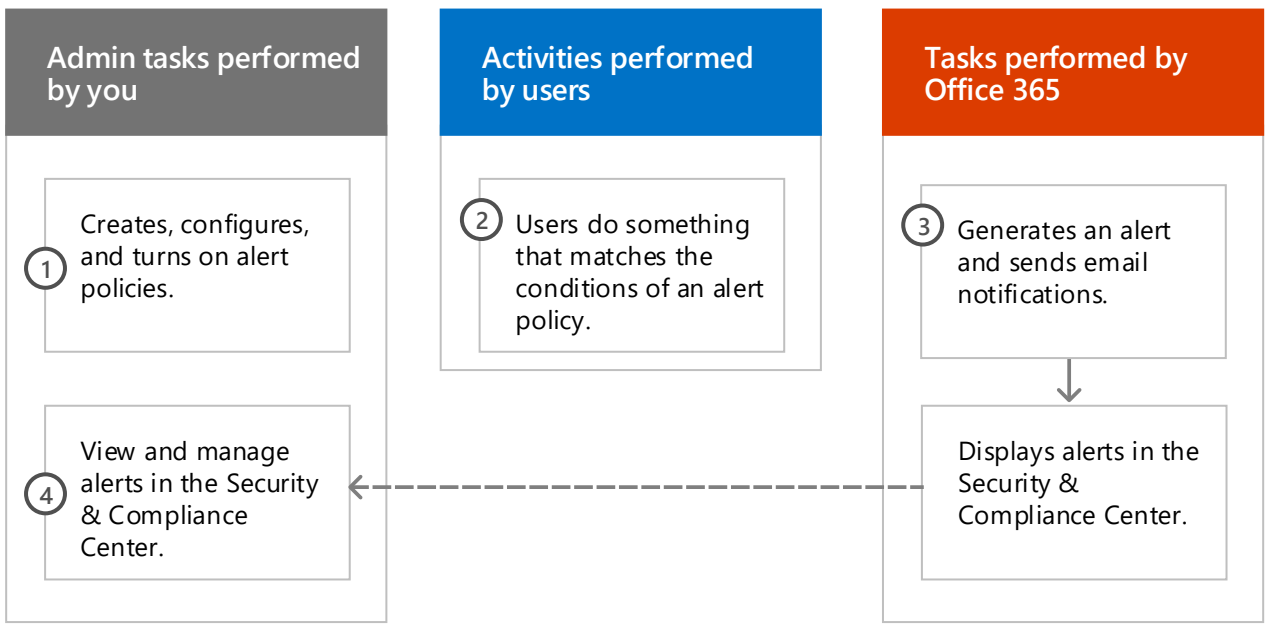
Search the audit log in the Office 365 Security & Compliance Center

Search-UnifiedAuditLog (cmdlet)

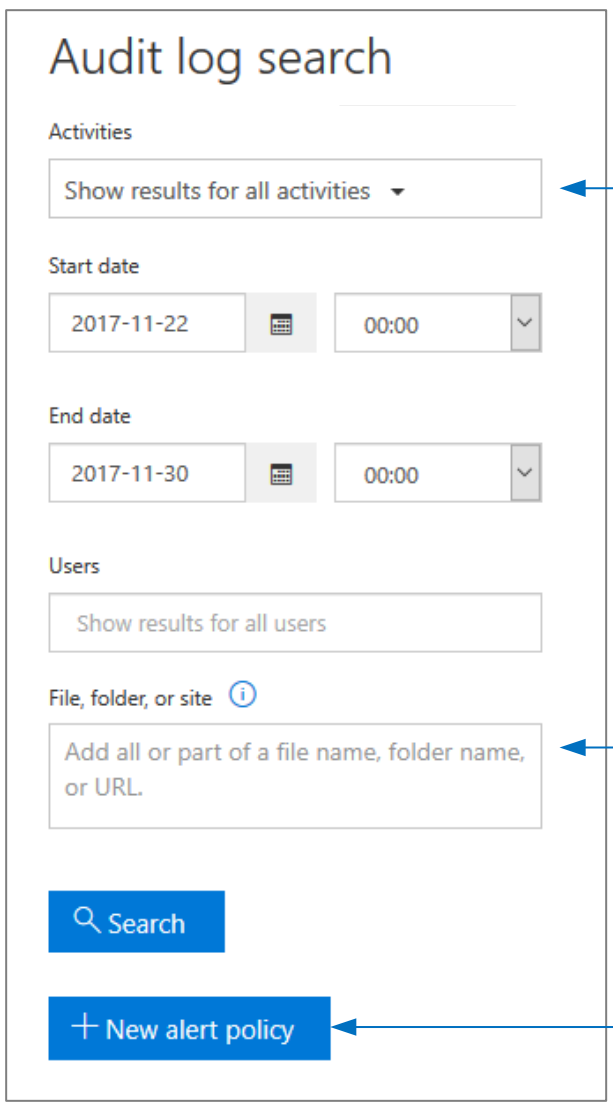
Detailed properties in the Office 365 audit log



How alert policies work



Searching the audit log



Example activities related to GDPR include:

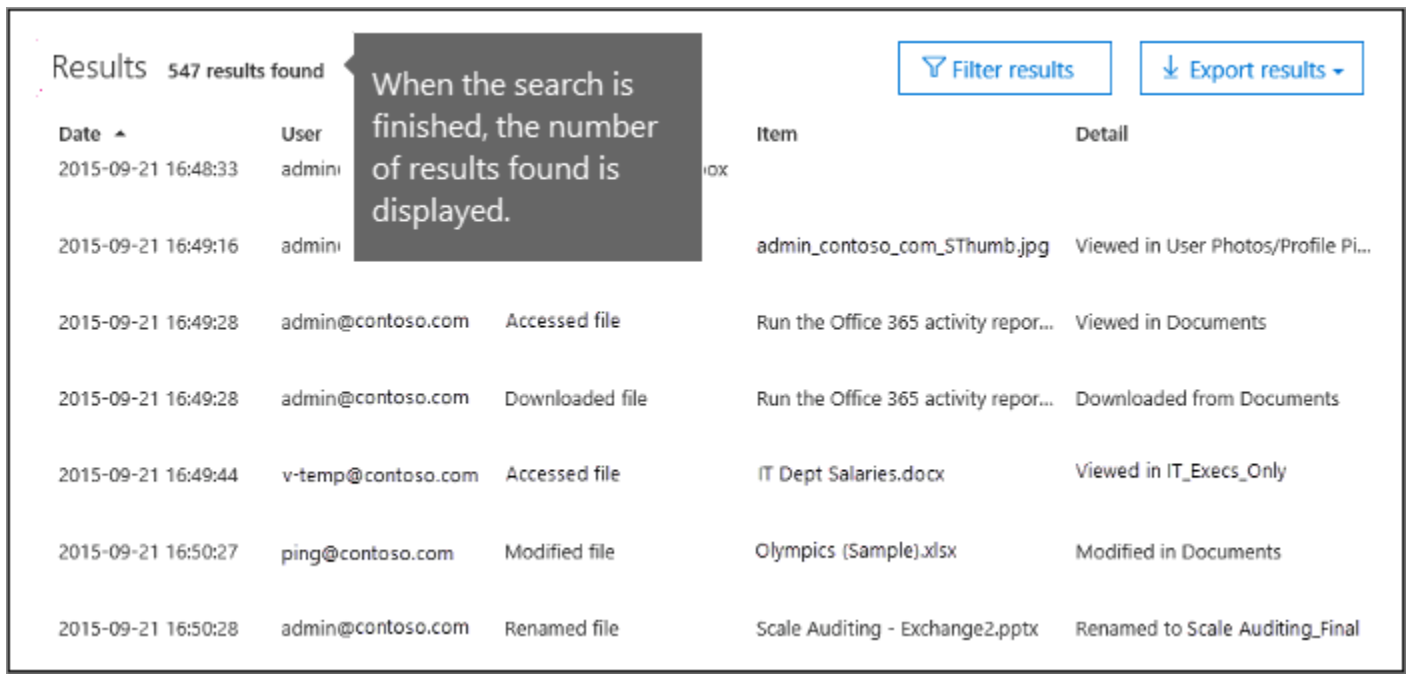
- Accessed file
- Copied file
- Downloaded file
- Accepted access request
- Accepted sharing invitation
- Broke permission level inheritance
- Created an anonymous link
- Created sharing invitations
- Allowed computer to sync files
- Downloaded files to a computer
- Changed a sharing policy (site admin)
- Many more . . . .

You can scope the search to sites with known personal data.

Create alert policies for continual monitoring of specific activities. These are different than alert policies in the new Alert dashboard. These alerts are sent to the specified recipients. Use PowerShell to view alerts that you've already created.

Continued on next page

Audit log search results



Microsoft Cloud App Security

Microsoft Cloud App Security helps you discover other SaaS apps in use across your networks and sensitive data that is sent to and from these apps.

Microsoft Cloud App Security is a comprehensive service providing deep visibility, granular controls and enhanced threat protection for your cloud apps. It identifies more than 15,000 cloud applications in your network-from all devices-and provides risk scoring and ongoing risk assessment and analytics. No agents required: information is collected from your firewalls and proxies to give you complete visibility and context for cloud usage and shadow IT.

To better understand your cloud environment, Cloud App Security investigate feature provides deep visibility into all activities, files and accounts for sanctioned and managed apps. You can gain detailed information on a file level and discover where data travels in the cloud apps.

Cloud App Security policies that can help with GDPR

Alert when files with these attributes . . .

<predefined PII attribute -or- custom expression>

. . . in these SaaS apps . . .

<choose or apply to all>

. . . are shared outside the organization

Block download of files with these attributes . . .

<predefined PII attribute -or- custom expression>

. . . in these SaaS apps . . .

<choose or apply to all>

. . . to any unmanaged device

Attributes coming soon to Cloud App Security

Office 365 sensitive information types

Unified labels across Office 365 and Azure Information Protection

Cloud App Security dashboard

If you haven't yet started to use Cloud App Security, begin by starting it up.

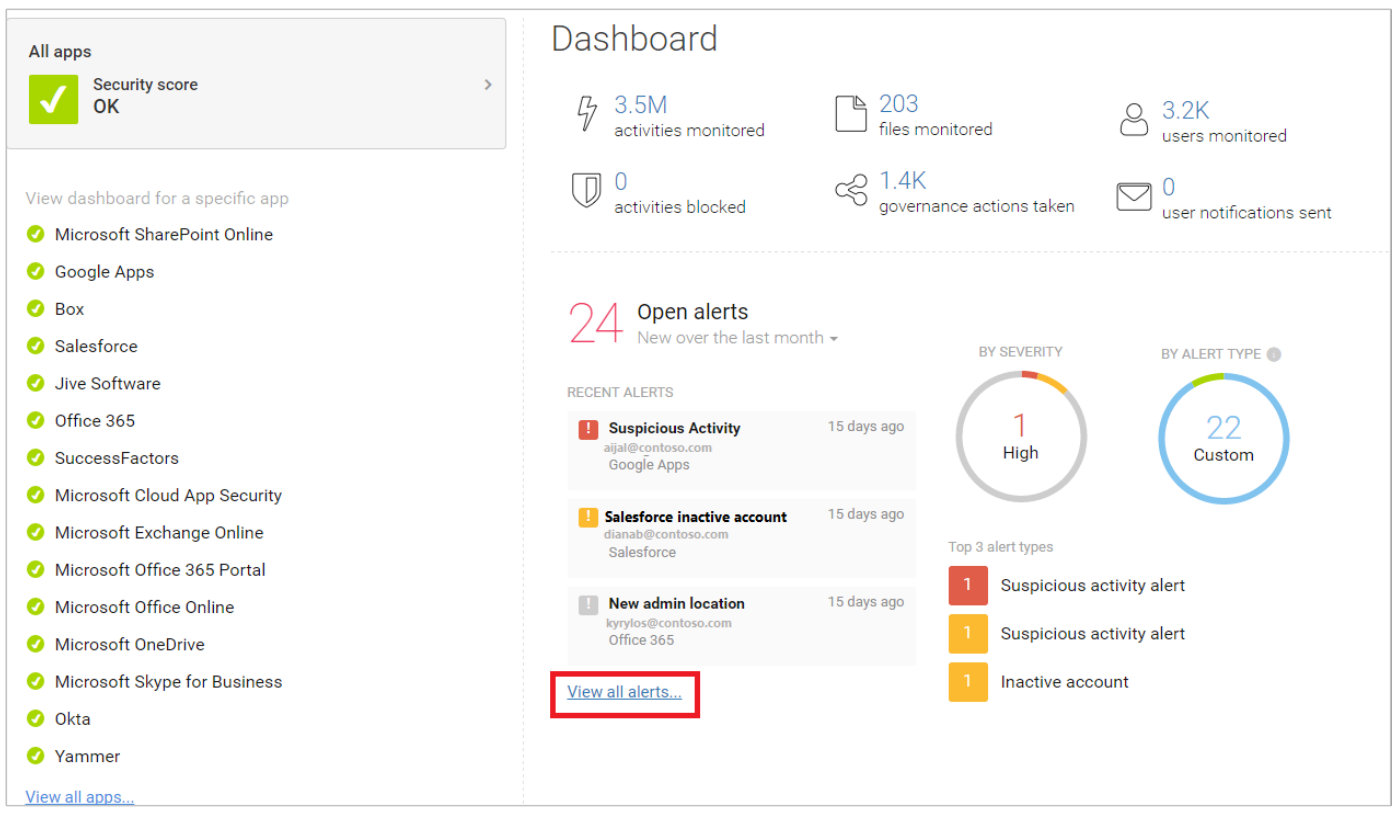
To access Cloud App Security: https://portal.cloudappsecurity.com.

Note: Be sure to enable '**Automatically scan files for Azure Information Protection classification labels**' (in General settings) when getting started with Cloud App Security or before you assign labels. After setup, Cloud App Security does not scan existing files again until they are modified.

Deploy Cloud App Security

More information about Microsoft Cloud App Security

Block downloads of sensitive information using the Microsoft Cloud App Security proxy



Continued on next page

Example file and activity policies to detect sharing of personal data

Detect sharing of files containing PII – Credit card number

Alert when a file containing a credit card number is shared from an approved cloud app.

Control	Settings
Policy type	File policy
Policy template	No template
Policy severity	High
Category	DLP
Filter settings	Access level = Public (Internet), Public, External App = <select apps> (use this setting if you want to limit monitoring to specific SaaS apps)
Apply to	All files, all owners
Content inspection	• Includes files that match a present expression: **All countries: Finance: Credit card number** • Don't require relevant context: unchecked (this will match keywords as well as regex) • Includes files with at least 1 match • Unmask the last 4 characters of the violation: checked
Alerts	• Create an alert for each matching file: checked • Daily alert limit: 1000 • Select an alert as email: checked • To: infosec@contoso.com
Governance	Microsoft OneDrive for Business • Make private: check Remove External Users • All other settings: unchecked Microsoft SharePoint Online • Make private: check Remove External Users • All other settings: unchecked

Similar policies:

- **Detect sharing of Files containing PII - Email Address**
- **Detect sharing of Files containing PII - Passport Number**

Note: this policy requires capabilities that are currently in private preview.

Detect Customer or HR Data in Box or OneDrive for Business

Alert when a file labeled as Customer Data or HR Data is uploaded to OneDrive for Business or Box.

Note: Box monitoring requires a connector be configured using the API Connector SDK.

Control	Settings
Policy type	Activity policy
Policy template	No template
Policy severity	High
Category	Sharing Control
Act on	Single activity
Filter settings	• Activity type = Upload File • App = Microsoft OneDrive for Business and Box • Classification Label (currently in private preview): Azure Information Protection = Customer Data, Human Resources—Salary Data, Human Resources—Employee Data
Alerts	• Create an alert: checked • Daily alert limit: 1000 • Select an alert as email: checked • To: infosec@contoso.com
Governance	All apps • Put user in quarantine: check • All other settings: unchecked Office 365 • Put user in quarantine: check • All other settings: unchecked

Similar policies:

- **Detect large downloads of Customer data or HR Data** — Alert when a large number of files containing customer data or HR data have been detected being downloaded by a single user within a short period of time.
- **Detect Sharing of Customer and HR Data** — Alert when files containing Customer or HR Data are shared.

Office 365 Information Protection for GDPR

Architecting information protection for sensitive information in Office 365

This topic is 7 of 7 in a series 1 2 3 4 5 6 7

Use Compliance Manager in the Service Trust Portal

The Compliance Manager in the Microsoft Service Trust Portal (STP) provides tools to track, implement, and manage the auditing controls to help your organization reach compliance with security or data protection industry standards when measured against Microsoft cloud services, such as Office 365 and Microsoft Azure. It helps the person who oversees the data protection strategy for your organization (sometimes called a Data Protection Officer) to manage the compliance and risk assessment process.

Compliance Manager in the Service Trust Portal (servicetrust.microsoft.com)

Use Compliance Manager in the Service Trust Portal

TechNet blog: Get to know the new Service Trust Portal

Blog: Compliance Manager Preview is now available

Blog: Manage Your Compliance from One Place – Announcing Compliance Manager

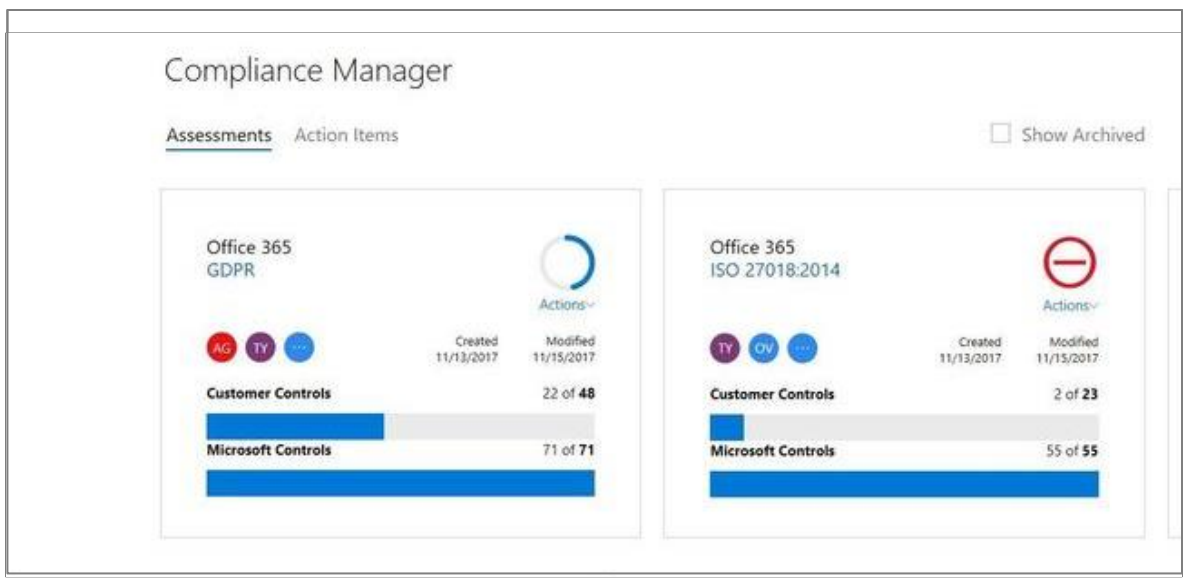
Compliance Manager:

- Combines the detailed information provided by Microsoft to auditors and regulators as part of various third-party audits of Microsoft's cloud services against various standards (such as International Organization for Standardization 27001:2013 and ISO 27018:2014) and information that Microsoft compiles internally for its compliance with regulations (such as the EU General Data Protection Regulation or GDPR) with your own self-assessment of your organization's compliance with these standards and regulations.
- Enables you to assign, track, and record compliance and assessment-related activities, which can help your organization cross team barriers to achieve your organization's compliance goals.
- Provides a secure repository for you to upload and manage evidence and other artifacts related to your compliance activities.
- Produces richly detailed reports in Microsoft Excel that document the compliance activities performed by Microsoft and by your organization, which can be provided to auditors, regulators, and other compliance stakeholders.

The core component of Compliance Manager is called an Assessment. An Assessment combines a Microsoft cloud service (such as Office 365) with a certification standard or data protection regulation (such as the GDPR). Assessments enable you to discern your organization's data protection and compliance posture against the selected industry standard for the selected Microsoft cloud service. Assessments are completed by the implementation of the controls that map to the standard being assessed.

Each assessment includes:

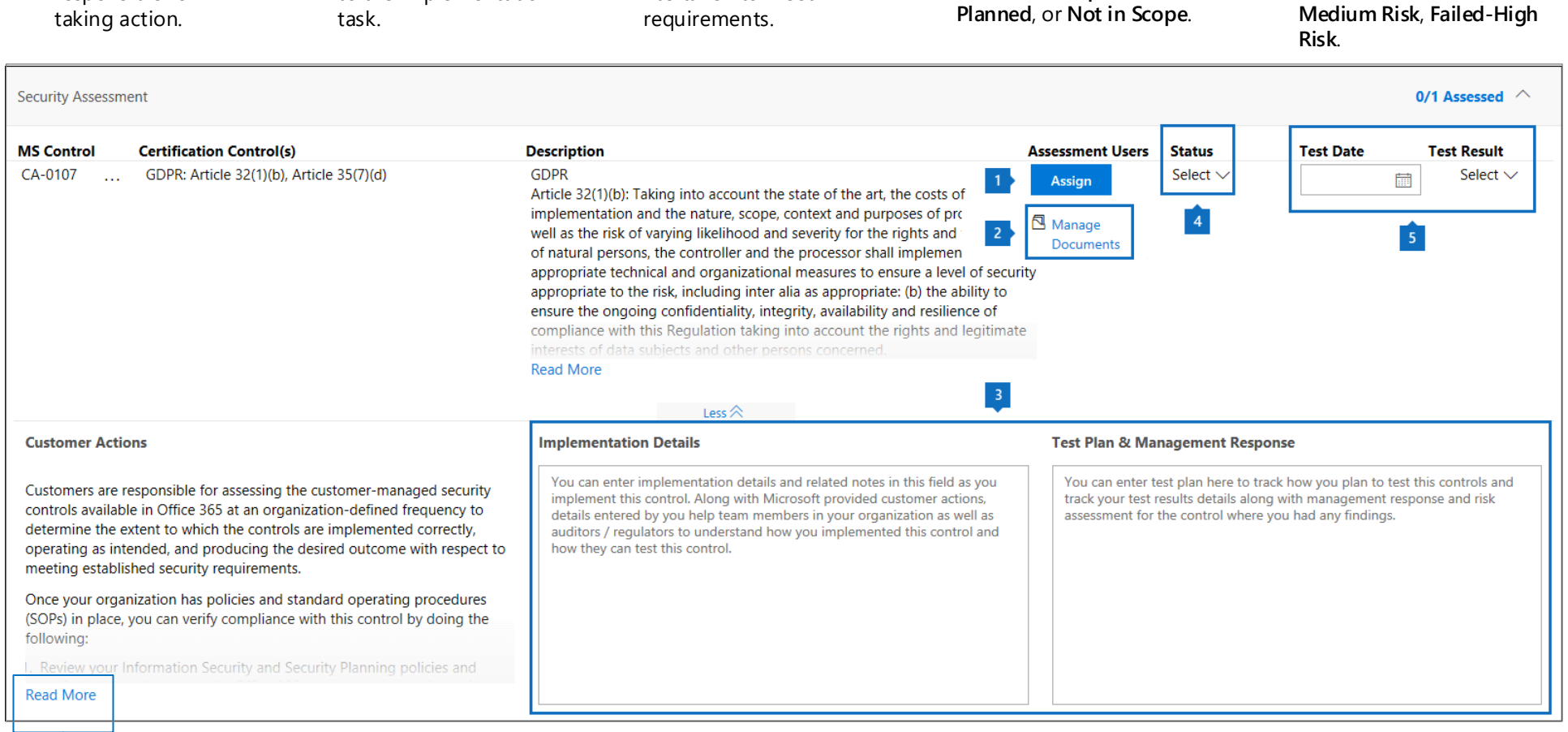
- **Microsoft managed controls** — For each Microsoft managed control, Compliance Manager provides details about how Microsoft implemented the control, along with how and when that control was tested and validated by an independent third-party auditor.
- **Customer managed controls** — Your organization is responsible for implementing these controls as part of your compliance process for a given standard or regulation.



Compliance Manager is a dashboard that provides a summary of your data protection and compliance stature and recommendations to improve data protection and compliance. This is a recommendation, it is up to you to evaluate its effectiveness in your regulatory environment prior to implementation. Recommendations from Compliance Manager should not be interpreted as a guarantee of compliance.

For each customer managed control . . .

1. Assign the item to the next person responsible for taking action.
2. Upload documents and other evidence related to the implementation task.
3. Add implementation steps your organization has taken to meet requirements.
4. Set status to **Not Implemented**, **Implemented**, **Alternative Implementation**, **Planned**, or **Not in Scope**.
5. Enter the test date and result: **Not Assessed**, **Passed**, **Failed-Low Risk**, **Failed-Medium Risk**, **Failed-High Risk**.



"Read More" to learn about Microsoft recommended actions for each article.

Microsoft