



This document contains instructions for the installation and setup of M-Files Discovery. The instructions are valid in M-Files January '25 Update (version 25.1) and later.

Contents

1. About M-Files Discovery	1
2. Prerequisites	2
3. Installing or Upgrading M-Files Discovery.....	2
4. Setting up M-Files Discovery.....	2
5. Dashboards	6
6. Solving Common Problems	7
7. Tips for Successful Use	8

1. About M-Files Discovery

With M-Files Discovery, you can find business critical information within the contents of large document archives. M-Files Discovery can automatically classify and categorize documents, find critical personally identifiable information (PII), and add metadata values based on document content.

M-Files Discovery can analyze the contents of an M-Files vault and the contents of a connected external repository. The analysis can be done to certain documents only (by document type or date), can be done in batches, can continually monitor for new content, or can do these all.

The types of analysis that M-Files Discovery can do:

1. Classify documents with M-Files Smart Classifier.¹
2. Find documents which contain specified text patterns.
3. Find documents which contain various forms of PII with M-Files Information Extractor.¹
4. Add metadata values based on document content with M-Files Information Extractor, M-Files Smart Metadata, and M-Files Smart Classifier.

When M-Files Discovery finds a document that meets the configured criteria, it can automatically set metadata properties, which of course can in turn, trigger workflows, set permissions, and so on.

¹ M-Files Smart Classifier and M-Files Information Extractor are M-Files intelligence services.

2. Prerequisites

Before you go forward, please make sure that these requirements are met:

- You have a license for M-Files Discovery.
- If you will use M-Files Discovery to classify documents, M-Files Smart Classifier is necessary. It is available from the [M-Files Solution Catalog](#).
- If you will use M-Files Discovery to find documents that contain PII, M-Files Information Extractor is necessary. It is available from the [M-Files Solution Catalog](#).
- If you will use M-Files Discovery to extract metadata from documents to save as properties on the object, you can configure M-Files Smart Classifier, M-Files Smart Metadata, or M-Files Information Extractor to provide these suggestions. These are available from the [M-Files Solution Catalog](#).

3. Installing or Upgrading M-Files Discovery

M-Files Discovery is an M-Files Vault Application. This means that you must install and update it separately from M-Files Server.

Note: When you upgrade, it is not necessary to first uninstall the application. The installation will replace the previous version.

When you install an older version, it is necessary to first uninstall the application. All configuration settings will be retained.

To install or upgrade M-Files Discovery:

1. Download the most recent version of M-Files Discovery from the [M-Files Solution Catalog](#).
2. Open M-Files Admin, expand a connection to M-Files Server, and go to **Document Vaults**.
3. Right-click on a vault and select **Applications**.
4. In the **Applications** dialog, click **Install**.
5. Browse to the folder where you downloaded M-Files Discovery and select the `M-Files_Discovery.mfappx` file.
When the installation is completed, *M-Files Discovery* will appear in the application list.
6. Install the M-Files Discovery license.
7. Click **Close**.
8. Restart the vault.

4. Setting up M-Files Discovery

This section is a quick start guide for the setup of each of the available scenarios. For a more detailed guide, refer to [M-Files Discovery - Administrator Guide](#).

Note: If your vault has M-Files Discovery release prior to 25.1 version, it is recommended to re-create all configured instances to ensure compatibility.

To configure M-Files Discovery in M-Files Admin:

1. For a vault that has M-Files Discovery installed, select **Configurations**.
2. Expand **Intelligence Services** and select **M-Files Discovery** from the list.
3. Continue with the instructions in the next sections.

4.1 Managing and Configuring Instances

M-Files Discovery can have multiple instances, each responsible for specific discovery tasks. Each instance is based on one of the supported scenarios: classifying documents, finding documents that contain specific text patterns, finding documents that contain PII, or extracting metadata values with other vault applications.

There is a top-level **Dashboard** that shows the list of instances. Each instance also has its own **Dashboard** (see section 5 for more details).

To create a new instance:

1. On the **Dashboard**, click **Add a new instance**.
The new instance is shown in the list of instances with the name "<New Configuration>".
2. Click **<New Configuration>**.
3. Select the **Configuration** tab to start the configuration for that instance.

There are three main areas of the configuration:

- General settings that include the selection of the type of discovery
- Settings specific to the target repository:
 - [Analyzing a vault](#)
 - [Analyzing an external repository](#)
- Settings specific to the type of discovery:
 - [Classifying documents](#)
 - [Finding documents that contain specified text patterns](#)
 - [Finding documents that contain PII](#)
 - [Adding metadata values based on document content](#)

You can remove an instance with the corresponding **Remove this instance** from the list of instances.

For advanced configuration options, refer to [M-Files Discovery - Administrator Guide](#).

4.2 Analyzing a Vault

To enable vault analysis:

1. In the **Target Repository** item, expand the **This Vault** section.
2. Set **Enabled** to **Yes**.

By default, M-Files Discovery will analyze *all* documents that have content. For details about filtering, refer to [M-Files Discovery - Administrator Guide](#).

4.3 Analyzing an External Repository

To enable external repository analysis:

1. In the **Target Repository** item, expand the **External Repositories** section.
2. Click **Add Connection**.
3. Expand the new entry and set **Enabled** to **Yes**.
4. In the **External Repository** drop-down menu, select an already configured and authenticated external repository.

By default, M-Files Discovery will analyze all documents that have content. For details about filtering, refer to [M-Files Discovery - Administrator Guide](#).

4.4 Classifying Documents

For this type of discovery, the M-Files Smart Classifier Intelligence Service is necessary. M-Files Smart Classifier must be installed, configured, and successfully trained before M-Files Discovery can start to classify documents.

An example of a basic configuration:

1. Expand **Discovery Settings**.
2. In **Classes**, select **All (default)**.
3. Set the **Target Property** to be a property definition for the list of **Classes**.
4. (Optional) Set the **Default Value** to a desired value from the list of **Classes**.
If no value is discovered, this value is used.
5. Save the configuration.

4.5 Finding Documents That Contain Specified Text Patterns

An example of a basic configuration:

1. In **Discovery Settings**, expand **Categories** and click **Add Category**.
2. Expand the new item and give it a descriptive name.
3. For the **Target Value List**, pick a value list that your target property chooses from.
It is important to set it to the same value list as **Target Property**.
4. Select a value in the **Target Category Value**.
5. Expand **Rules for Finding Phrases** and click **Add Phrase Rule**.
6. Expand the new item, give it a descriptive name, and then decide whether you want a regular expression phrase or a text phrase for the **Type of Rule** property.
7. Enter a rule based on your decision in the **Rule** property.
8. Leave the remaining settings in **Rules for Finding Phrases** and **Categories** as default.
9. For the **Target Property**, it is important to select a property that picks from the value list that you specified in the **Target Value List** section.
10. Save the configuration.

4.6 Finding Documents That Contain PII

For this type of discovery, the M-Files Information Extractor Intelligence Service is necessary. It is not necessary to enable or configure M-Files Information Extractor for M-Files Discovery to find documents that contain PII.

An example of a basic configuration:

1. Expand **Discovery Settings**.
2. For **Target Property (Indicator)**, select a Boolean property, and for **Target Property (Description)**, select a multiline text property.
3. Leave **Overwrite Existing Values** field to **No (default)**.
4. Save the configuration.

4.7 Adding Metadata Values Based on Document Content

You can use multiple vault applications for this type of discovery. You need to separately install, enable, and configure these applications.

These vault applications are currently supported:

- M-Files Information Extractor: Matcher
- M-Files Information Extractor: Text Analytics
- M-Files Smart Classifier
- M-Files Smart Metadata

Before you start to use this scenario, make sure that you update any of the vault applications that you want to use to their latest version.

An example of a basic configuration:

1. Expand **Discovery Settings**.
2. In Suggestion Strategy, select **Vault Application**.
3. Expand **Application Settings for Selected Scenario**.
4. Click **Add Vault Application**.
5. Expand the new node.
6. In the **Vault Application** list, select a vault application.

The vault application must be installed, enabled, and fully configured.

4.8 Starting the Analysis

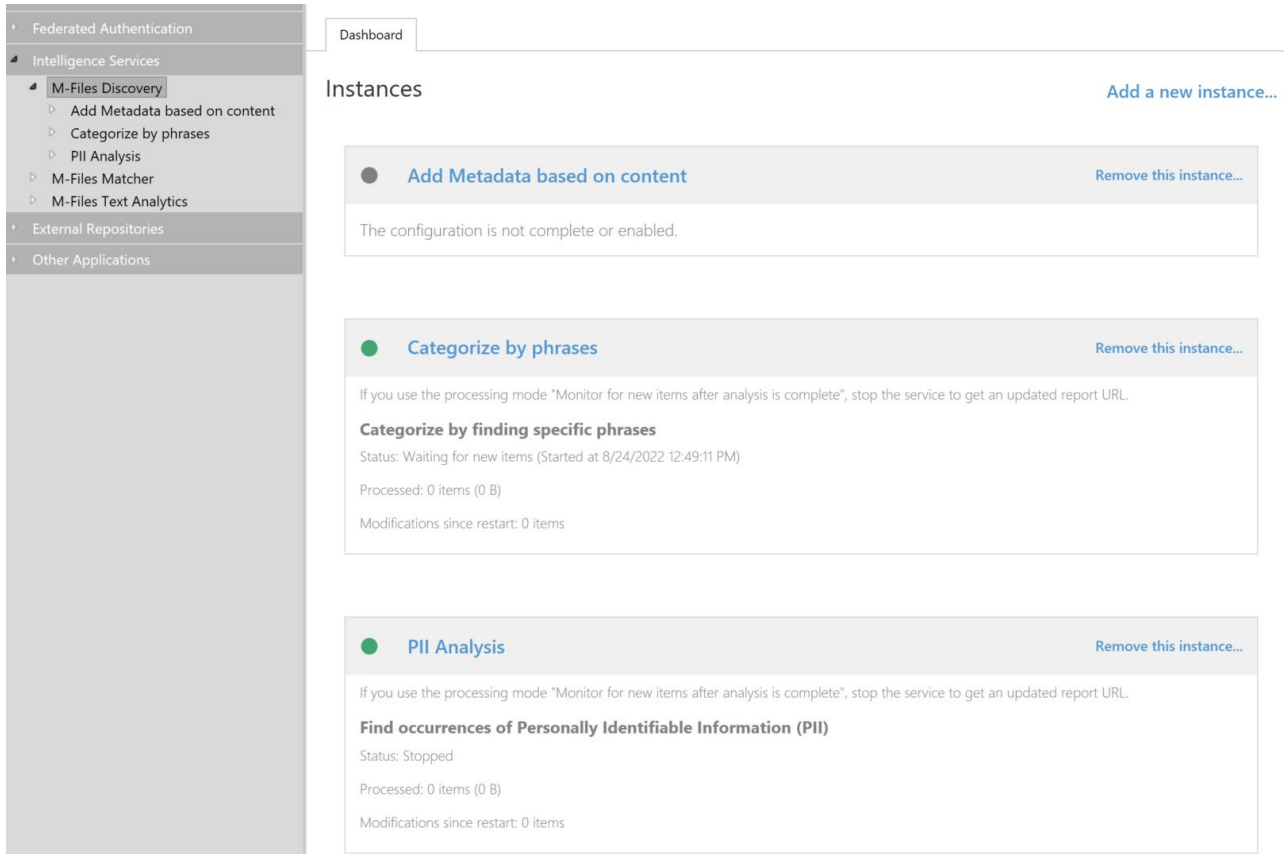
After you have configured all the instances, you can start the analysis. To start the analysis for any of the Discovery instances, click **Start** on the **Dashboard** of the instance.

Tip: Click **Evaluate** before starting the analysis to create an evaluation of the repository content that matches the filtering criteria.

5. Dashboards

The M-Files Discovery **Dashboard** shows the status and progress of any ongoing or completed analysis. The **Dashboard** has two levels: a top-level summary **Dashboard**, which gives an overview of all configured instances, and a detailed **Dashboard** for each instance which gives further information and management options for that instance.

On the summary **Dashboard**, each instance is listed with an icon that indicates its status. It is possible to add and remove instances from this dashboard.



The screenshot displays the M-Files Discovery Dashboard. On the left is a navigation sidebar with categories: Federated Authentication, Intelligence Services (expanded to show M-Files Discovery, Add Metadata based on content, Categorize by phrases, PII Analysis, M-Files Matcher, and M-Files Text Analytics), External Repositories, and Other Applications. The main content area is titled 'Dashboard' and 'Instances'. At the top right of the instances list is a link 'Add a new instance...'. Three instances are listed:

- Add Metadata based on content** (grey circle icon): 'Remove this instance...' link. Below the title, a message states: 'The configuration is not complete or enabled.'
- Categorize by phrases** (green circle icon): 'Remove this instance...' link. Below the title, a message states: 'If you use the processing mode "Monitor for new items after analysis is complete", stop the service to get an updated report URL.' The instance title is **Categorize by finding specific phrases**. Status: 'Waiting for new items (Started at 8/24/2022 12:49:11 PM)'. Processed: '0 items (0 B)'. Modifications since restart: '0 items'.
- PII Analysis** (green circle icon): 'Remove this instance...' link. Below the title, a message states: 'If you use the processing mode "Monitor for new items after analysis is complete", stop the service to get an updated report URL.' The instance title is **Find occurrences of Personally Identifiable Information (PII)**. Status: 'Stopped'. Processed: '0 items (0 B)'. Modifications since restart: '0 items'.

Image 1: Example of a summary dashboard.

The detailed dashboard shows more information about each instance. You can use it to create an evaluation of the content that matches the filter in this instance, and to start, stop, and restart the instance.

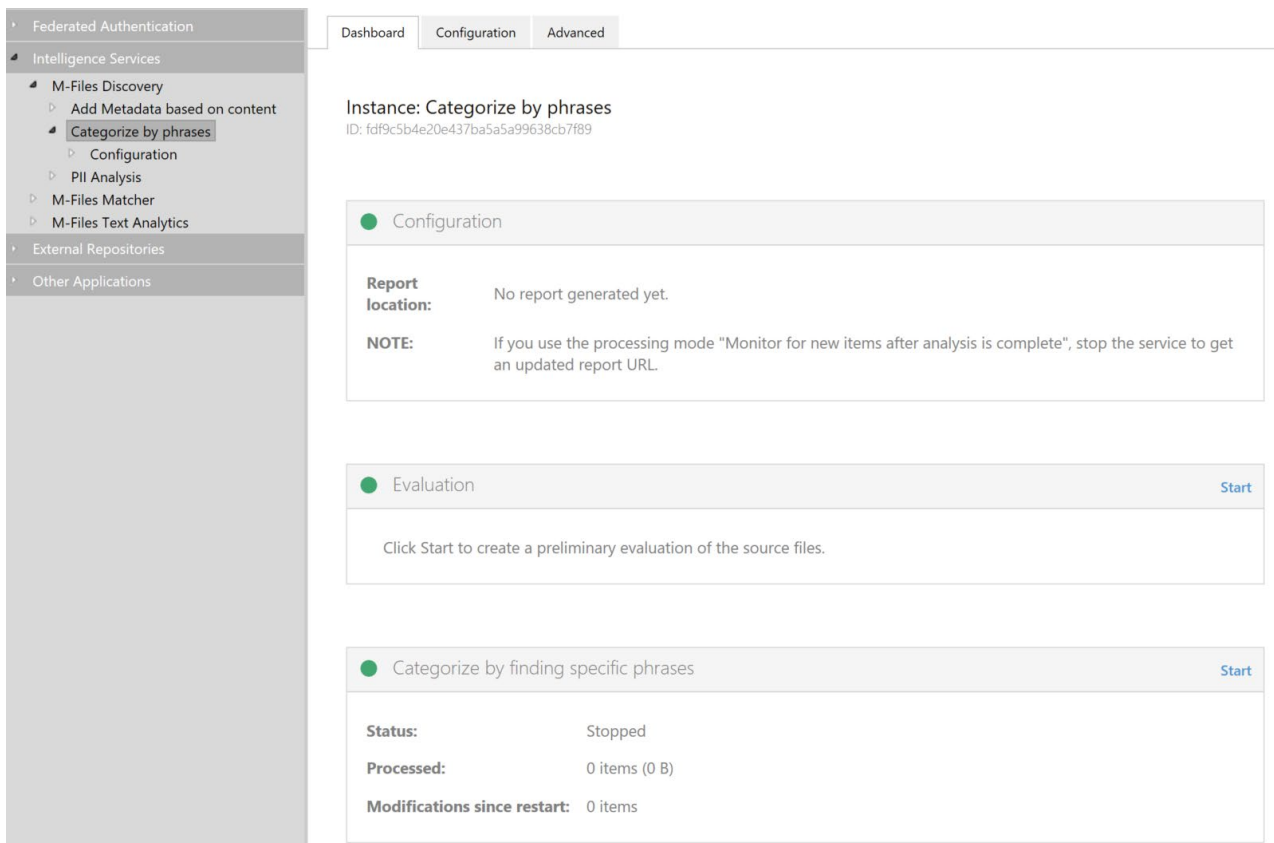


Image 2: Example of an instance dashboard.

When an instance has been configured, the **Start** option is shown on the **Dashboard** tab for the **Evaluation** and **Scenario** panels. Only one of these operations can be active at the same time. The **Stop** command pauses the operation and **Restart** starts the analysis from the beginning. After the operation is stopped, **Start** continues the analysis from the position where it was stopped.

The most typical scenario where you might want to restart the execution of an M-Files Discovery instance is after a configuration change. A restart will re-analyze everything again and apply the new results if necessary.

6. Solving Common Problems

In this section, you can find solutions to common problems.

PROBLEM	SOLUTION
No license.	To use M-Files Discovery, a valid M-Files Discovery license is required. Add a valid license.
M-Files Discovery is not listed under Intelligence Services.	The vault may not have been restarted after you installed the vault application. Restart the vault.
The M-Files Discovery Dashboard is not up to date.	Dashboard updates are triggered every 10 seconds. New information will only be shown after the update has been triggered.
Not analyzing an external repository.	Make sure that a valid user is configured and authenticated with the external repository.
Not classifying documents.	Make sure that M-Files Smart Classifier is configured, trained, and functions correctly. Only those classes listed on the M-Files Smart Classifier Dashboard will be discovered.
Not finding documents that contain specified text patterns.	This can happen if the target property does not pick from the same value list as each of the configured categories. Make sure that each category for this scenario uses the same value list and that the target property picks from this list.
Error occurs when attempting to classify documents or when attempting to find documents that contain PII.	Make sure that the latest versions of M-Files Smart Classifier and M-Files Information Extractor are installed.

7. Tips for Successful Use

In this section, you can find a few tips on how to prevent common problems.

Be careful when overwriting existing metadata

M-Files Discovery does not have an "undo" option. Any changes that it makes during its analysis cannot be easily undone – except restoring from a backup. Be very careful when overwriting values.

In some situations, as a trial run, it can make sense for M-Files Discovery to initially only update a new metadata property. When validated, this property can be removed, the configuration changed to overwrite the actual target property, and then the analysis restarted.

Start with small batches

M-Files Discovery is a powerful tool that can make sweeping changes to a vault. For this reason, it is recommended to first start with small batches and then evaluate whether the results are what you try to achieve.

As well, it may be helpful to have the same configuration for multiple instances where the only difference is **Filter Criteria**. With this, the filter options effectively create subsets of the analysis.

Minimize the number of documents to be processed

By using **Filter Criteria**, the efficiency of M-Files Discovery can be improved. If possible, it is recommended to focus on and analyze only candidate documents, not all documents whenever possible.

Be aware of the number of objects that may be promoted from an external repository

M-Files Discovery will automatically promote a document from an external repository to the vault when adding metadata values to that object. This of course requires storage within the vault. If you expect many documents to be promoted, make sure that you have enough allotted space in the vault.

Find and review items that were discovered but were skipped

If a document is, for example checked out, or for some other reason it is read-only, M-Files Discovery cannot update it to complete its processing for that document. In these situations, an entry is logged in the **Event Log** in M-Files Admin. It is important to check this log regularly.

Enable M-Files Discovery to initiate workflows and/or update permissions

Refer to the [M-Files Discovery - Administrator Guide](#) for initiating workflows and updating object permissions.

Be aware that M-Files Discovery will only modify objects that were last modified by a user

If a document was last modified by a server operation (for example M-Files Discovery instance adds a property), the added item is not re-processed by default. This prevents the object from getting into a modification loop. Refer to the [M-Files Discovery - Administrator Guide](#) for processing these objects.