

File Analysis Suite Data Discovery

Software Version 3.4

Release Notes



Document Release Date: January 2021
Software Release Date: January 2021

Legal notices

Copyright notice

© Copyright 2019-2021 Micro Focus or one of its affiliates.

The only warranties for products and services of Micro Focus and its affiliates and licensors ("Micro Focus") are as may be set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Micro Focus shall not be liable for technical or editorial errors or omissions contained herein. The information contained herein is subject to change without notice.

Contents

Data Discovery Release Notes	1
About Data Discovery	1
What's new	2
New in 3.4	2
New in 3.3	5
New in 3.2	10
Documentation	12

Data Discovery Release Notes

Software version: 3.4

Publication date: January 2021

This document is an overview of the changes made to Data Discovery.

Support matrix

For information about the installation requirements and compatibility with other products, see the Data Discovery Support Matrix. The support matrix may be updated between releases so it is available only from the [Support portal](#).

About Data Discovery

Micro Focus Data Discovery gives you multi-repository management and review across your enterprise data. Identify, collect, and organize content to ensure discovery of sensitive data. Configure how repositories are processed and categorized with Connect. Analyze your data under management with Analyze. Organize, review, and export data with Manage.

What's new

This section describes the new features and enhancements to Data Discovery.

New in 3.4

The following features and enhancements are included in this release.



General



The following features and enhancements apply to more than one component in Data Discovery.

- Added the ability to secure data in-place. Create one or more protection systems to connect to your existing file protection applications and apply rules to sensitive data identified by Data Discovery. Protect all documents in a workbook using a defined protection system.

Initial file protection supports Voltage SmartCipher 4.2.1.10 and newer, letting you leverage your existing SmartCipher implementation. SmartCipher supports file protection for DOC, DOCX, PPT, PPTX, XLS, XLSX, and PDF files.
 - Added ability to filter protection status using the filter panel in Analyze and Manage.

Added ability to include or exclude documents that have been protected using the search builder. This ability is part of a search in the content view in Analyze and Manage.
 - Added "encryption" as a reported action on the Agent Activity page in Connect.
 - Added encrypted/protected field as an available field when exporting to CSV.
 - Added new "Protect documents" permission to Manage and Workspace Security for appropriate roles.
- Added the ability to sample data based on a percentage of data across all data or a subset of data from a search, a workspace, or a workbook.
 - Added a Random Sampling option to the task workbook in Manage. When you select random sampling as the task, you select the data sources to sample and then set the percentage of data for each data source to sample. The resulting documents are gathered into a workbook for you to review or index content for data risk assessment using estimated sensitive analysis.
 - Added an "Estimated Sensitive Analysis using Sampling" workspace template in Manage. This template include a default workbook category, Sampled, that is created when a workspace is created based on this template
- Added and updated dashboards.
 - In Analyze, renamed the Sensitive Data dashboard to Analyzed Sensitive Data dashboard to more accurately reflect the data represented.

- In Analyze, added an Estimated Sensitive Data dashboard. This new dashboard provides insight into how much sensitive data may be present in your Data Discovery environment based on the sensitive data already found in your analyzed data.
- In Manage, added file protection data to the workspace summary panel on the workspace Overview page.
- In Analyze and Manage, updated the "hold" icon to allow for clearer meaning alongside the new "protect" icon. Both icons are used for the in-line information icons in the document list views and for the action icons on the Activity tab of the workbook detail panel. The hold icon is now  and the protect icon is .

In the workbook detail panel, the release hold icon () displays for workbooks that are already on hold; the release protection icon () displays for workbooks that are already protected.

- Updated the function of viewing duplicate items. From the metadata tab of the document view panel in Analyze and Manage, the View all Duplicates link refreshes the current document list to show you the duplicates—instead of opening in a new browser tab.
- Updated the options for viewing the document lists in Analyze and Manage. You can now view the document list only, the document list and either content or the detail panel, or the document list, content, and detail panel. The last view you choose is saved as your preference until you change it again.
- Updated the recommended screen resolution for viewing the UIs. The new recommended minimum resolution is 1366 x 768.
- During implementation of the Data Discovery environment, added an optional configuration to enable field standardization during processing with custom connectors. Field standardization structures the metadata so that documents have a consistent structure and consistent field names. Use field standardization so documents indexed through different connectors use the same fields to store the same type of information. Field standardization only modifies fields that are specified in the connector's dictionary. If a connector does not have any entries in the dictionary, field standardization has no effect.

Processing Agent

The following features and enhancements apply specifically to the processing agent.

- Updated authentication for Exchange Online repositories to use OAuth authentication instead of basic authentication. When creating an Exchange Online repository in Connect, you are prompted to define the Application ID, Directory ID, and Client Secret as registered in the Exchange admin center during implementation of Data Discovery.

This change supports the retirement of Basic Auth for Exchange Web Services (EWS) for Exchange Online/Office 365.

Connect

The following features and enhancements apply specifically to Connect.

- When selecting entities while creating and editing repositories and repository templates, you can now select multiple languages in the Languages filter. This lets you filter for and select language-specific entities that are not tied to regions or countries selected on the previous page of the

create and edit wizards.

- Updated support for entities.
 - Added global (not country-specific) device entities. These entities include information that can identify electronic devices, such as ICCID, IMEI, IMEISV, IMSI, IP addresses, MAC addresses, MEID, and MSISDN.
 - Added a National ID entity for - Argentina, Bahrain, Colombia, China, Dominican Republic, Egypt, Hong Kong, India, Indonesia, Malaysia, Mexico, Pakistan, Russia, Singapore, Thailand, and United Arab Emirates.
 - Added Date of Birth language support for Afrikaans, Northern Sotho, Romansh, Swati, Tsonga, Venda, Xhosa, and Zulu.
 - Added entities specific to Brazil, Japan, South Africa, Switzerland, and Taiwan.
 - Updated the geographic regions to reflect current county association.
 - Added two geographic regions, *APAC (Including CJK)* and *Worldwide (including CKJ)*. These new regions add American Samoa, China, Hong Kong, Japan, Singapore, Taiwan, and Thailand to *APAC* and *Worldwide*.
- Added the ability to define a custom entity using terms. When creating a custom entity, you can now choose to base the entity on terms instead of based on a pattern. You can define individual terms or select from an existing term list.

Term Lists can be accessed under Tags or Grammars in the navigation panel. The term lists available under each are identical. Your access to each is determined by the permission you have to tags and grammars.

- Added support for File System as a target.
- To the Destinations page, added path and origin information and target sorting.
- To the Agent Activity page, added a column for the user who initiated the task.
- To the Processing page, added column sorting. You can now sort the processing information by all data points except the User.

Analyze

The following features and enhancements apply specifically to Analyze.

- Added the ability to add all or only selected documents in the document list to an existing workspace in Manage. The user must have permission to add and create data sources (Manage data sources) and workspaces (Manage workspaces) in Manage.

Manage

The following features and enhancements apply specifically to Manage.

- Added the ability to create multiple data sources, one per repository, at once. When creating a repository, you now have to select whether you want to create a Single Data Source that includes one or more repositories that you select or to create multiple Data Source per Repository. This new option lets you create data sources for each repository, with the same criteria, without having to create each data source individually.

Creating multiple data sources per repository allows for more streamlined detection of duplicates when creating a deduplication task workbook.

- Added the ability to reopen a workspace. A REOPEN button has been added to the summary panel on the workspace Overview page for closed workspaces.
- Updated the data represented in the Category cards on the Workbooks page. The data now shows the percentage of documents in the workspace associated with the category instead of the percentage shared with other categories.
- Added auto-refresh for the workbook status in the list of workbooks in a workspace and in the workbook detail panel. This lets you see the status while in progress.
- Added the workbook description to the workbook detail panel.

Administration



- Removed the Inherited Roles tab from the Edit Users function.

New in 3.3

The following features and enhancements are included in this release.

General

The following features and enhancements apply to more than one component in Data Discovery.

- Added features and functionality to support data optimization.
 - Added a new processing method to process only the document metadata. From the Activity tab of the workbook detail panel, you can process metadata-only documents to index the content or to fully collect the documents.
 - Added a new workspace template that includes new categories and new workbook templates. Use the new Data Cleanup workspace to gather redundant, obsolete, and trivial files (such as very old, system or temporary files) to determine files that can be deleted from your environment.
- In Analyze and Manage, updated the "enriched" and "content indexed" in-line information icons in the document list views to better articulate their meaning. The enriched icon is now ; the content icon is now .
- In Analyze and Manage, added information about any issues encountered during document processing.
 - When viewing documents, you will see an error or warning icon in the document view panel to the far right of the document name or email subject line, as applicable. Hover over the icon to see the error or warning message. Open the document detail panel to the metadata tab to view the full details of the error or warning. If a document has both errors and warnings, you will see the error icon and details for all errors and warnings are available on the metadata tab.
 - You can also search for documents that generated errors or warnings during processing. The document list filter panel now includes a section for Warnings to select one or more specific warnings to filter for, and a section for Errors to select whether to filter for documents with or

without errors. The search builder also now includes warnings and errors in the available criteria.

- Added the ability to search for documents based on the document ID. Using the search builder, you can now specify one or more document IDs to include in your query when creating dynamic workbooks and workbook templates, when specifying data source criteria for custom repositories, and when searching within a document list.
- In Analyze and Manage, added a new document list view. You can now view documents in the current detail view or the new grid view. The view you select determines the initial information that is available and any additional panels available. The activity ribbon and filter panel are present for both views.
 - The *detail view* includes the document list, the document view panel, and the document detail panel. In this view, you can see the content of the selected document if available.
 - The *grid view* includes the document list, but displays different information than the document list in the detail view. The grid view include In this view, you cannot see the content of the selected document and the document detail panel is not available.
- Updated the navigation panel for all UIs to a retractable, but still visible, view. Previously, you could hide the navigation panel, but you then had to reveal it in order to navigate through the UI. With this update, you can retract the navigation panel to show only the icons for the top level menu items. When you click one of the icons in the retracted navigation panel, the panel expands to full width. If you click a menu item that has sub-items, the navigation panel and menu items expand and the page does not change until you click the sub-item; if you click a menu item that does not have sub-items, the navigation panel expands and you are taken to the appropriate page.
- In Analyze and Manage, changed references to and metrics for "unique content" to "shared content". In Analyze, this affects metrics on the Sensitive Data dashboard. In Manage, this affects metrics in the workbook category panel, the workbook detail panel, and the category charts in the workspace report.
- Updated the date pickers for all areas that use the query-type selection.
 - In Analyze and Manage, the search builder date-related criteria let you define dates based on a date range (from x to y), before a specific date, or after a specific date.

These date-related options are also available in the tag criteria selection in Connect and in the dynamic workbook criteria selection in Manage.
 - In Manage, the workbook template date-related criteria let you define dates based on a date range (from x to y), before a specific date, after a specific date, or older than a specific number of days, weeks, months, or years.
- Added collection status, enrichment status, and whether an item has been sent to a target to the available information to include when exporting a document list to a CSV. If an item has been sent to a target, the target and destination names are included in the export.
- Added new date fields to all advanced search query builders and when exporting items from a document list in Analyze and Manage to a CSV file. The following date fields have been added.

Date field	Description
First Processed Date	Defines the date that the document was initially processed by Data Discovery. This date may differ from the Archive Date if the document was not collected during initial processing.
Last Processed Date	Defines the date that the document was last processed by Data Discovery. This date may differ from the First Processed Date or Archive Date if the document has been reprocessed in any way. NOTE: This was previously referred to as just <i>Processed Date</i> .
Modify Date	Defines the last modified (saved) date for the document itself.
Record Date	Defines the last modified date for the document itself. <ul style="list-style-type: none">For file items, this would be the last modified date, and the same as the "Modify Date" field.For email items, this would be the sent date.

- For charts and graphs in all applications, the processed date now refers to the *first processed date*.
- Across all applications, updated all values rounded down to "0" to "<1" for accuracy.

Processing Agent

The following features and enhancements apply specifically to the processing agent.

- Removed support for installing the processing agent on Windows Server 2012 R2.

Connect

The following features and enhancements apply specifically to Connect.

- When creating a repository or repository template, added the ability to filter grammars and entities by geographic region to help you select the appropriate grammars based on established geographic regions. Select a geographic region to pre-select the included countries and then the applicable entities. You have the option to select or deselect countries and entities as desired.
- Removed the region-specific default repository templates.
- Added information about document processing. View the progress and status of background processing tasks on the new Processing page.

Some analysis actions, such as creating or editing tags, create a background task. Previously these tasks could only be monitored by Micro Focus.
- Removed the option to collect documents when creating a repository. Documents can be collected at the workbook level and at the document level.
- Added collection status, enrichment status, and whether an item has been sent to a target to the available criteria for tags.
- For SharePoint Online repositories, added option to use federated authentication.

- Added sort and search functionality to the list of Term Lists. By default, the list of term lists is sorted alphabetically, ascending.
- Charts and data presented in the Data Volume dashboard are now relative to the user's timezone.

Analyze

The following features and enhancements apply specifically to Analyze.

- Added collection status, enrichment status, and whether an item has been sent to a target to the filter panel and the search builder on the document list page of Research.
- On the Data Analysis dashboard, replaced the trailing period date slider with a more flexible selector. You can now define the desired number of previous days, weeks, months, or years to view data for.
- Added an option to create a workspace from just the selected documents in the document list. Previously, you could only create a workspace from all documents in the document list.

Manage

The following features and enhancements apply specifically to Manage.

- Added new workbook types, *query* and *task*, and updated the label for the dynamic workbook, now labeled *dynamic query workbook*. The following workbook type are available in this release.
 - A *static* workbook represents a set of items that you have manually gathered. You can create an empty workbook and then add items to it, or create the workbook and add items to it at the same time from a document list. You can continue to add items to this workbook as needed.
 - (New) A *query* workbook represents a set of items based on defined search criteria. The search criteria you used to gather the items is saved as part of the workbook, but cannot be edited. The search query is performed once at the time the workbook is created. Similar to a static workbook, you can continue to manually add items to this workbook as needed.
 - A *dynamic query* workbook represents a set of items based on defined search criteria gathered on an ongoing basis. The search query is performed at the time the workbook is created and as new items are processed or existing items are reprocessed; matching items are added to the workbook. You can edit the search criteria as desired and, as a result, the actual items within a dynamic workbook can fluctuate.
 - (New) A *task* workbook represents a set of items gathered as a result of a specific action, such as duplicate identification (deduplication). The defined action is taken on items associated with selected data sources and performed once at the time the workbook is created. View the task history on the Activity tab of the workbook detail panel. Similar to a static workbook, you can continue to manually add items to this workbook as needed.
- Added a new workspace template to identify redundant, obsolete, and trivial data. This template, Data Cleanup, includes categories (Redundant, Obsolete, and Trivial) and query workbooks (Obsolete Data and Trivial Data) that are created when the template is used to create a workspace.
 - The Trivial Data workbook created within the data cleanup workbook includes the Trivial category and specifies to gather processed temporary and system files into the new workbook. By default the Trivial Data workbook gathers these files from File System and SharePoint repositories only.

You can edit the Trivial Data workbook template, and workbooks based on the template, to change the workbook criteria, including the data source types and the defined file types.

- The Obsolete Data workbook created within the data cleanup workbook includes the Obsolete category and specifies to gather processed items that are older than 10 years. By default, the Obsolete workbook gathers these files from File System and SharePoint repositories only.

You can edit the Obsolete workbook template, and workbooks based on the template, to change the workbook criteria, including the data source types and the desired age of matching items.

- At the workbook level, added the option to process metadata-only documents to index the document content. A button, "Index Content", has been added to the Activity tab of the workbook detail panel. This action applies to all documents within the workbook.
- Added collection status, enrichment status, and whether an item has been sent to a target to the filter panel on the Content page.
- Added collection status, enrichment status, and whether an item has been sent to a target to the search builder. This applies to the search builder on the Content page as well as the search builders that are part of the workbook criteria, workbook template criteria, and the data sources criteria for data sources created automatically when a workspace is created from a search results list from within Analyze.
- Improved tagging processing to allow concurrent tagging operations and bulk application of tags applied to a document family.
- Added enforcement of restricting the \ / : * ? " < > | . special characters from the name of export locations. This applies to exporting content from a workspace. The enforcement aligns with the restricted characters for folder names on the supported operating systems of the export destination.
- The bar charts on the workspace overview page are now clickable. You are taken to the Content tab, showing documents related to the chart clicked.

The bar charts in the detail panel on the Data Source, Data Subject, Hold, and Workbook pages are also now clickable to the Content tab showing the appropriate documents based on what you clicked.

Administration

The following features and enhancements apply specifically to Administration

- Added a Usage report to review information about your total data under management for a selected time frame. To view the usage report, you must have the *Connect Manage repositories and targets* permission. View overall total counts and file size of managed data, total entities scanned for and found, as well as total counts and file size for each repository defined in Data Discovery.
- Updated the *Collect data* permission for Manage and Workspace Security to "Index and collect data". With the addition of being able to process only document metadata, this permission lets the assigned user reprocess those documents, at the workbook level, to process and index the document content.

New in 3.2

The following features and enhancements are included in this release.

General

- Standardize filter selection across all list pages in all interfaces.

Processing Agent

- Streamlined Agent installation and Agent Administration application. Database configurations have been combined to reduce possible errors when defining connections.
- Improved reporting on the agent data processing. From the System Status page of the Agent Administration application, you can review the status of each step of the data processing flow. If the process resulted in information messages, warnings, or errors, the number of messages displays in the appropriate column for the processing step. Click the number in the column to go to the Event Console and review the details of the selected information message, warning, or error.
- Added the agent Gateway URL to the agent download page for easy reference during installation of the agent.

Connect

- New region-specific entities to allow for identifying sensitive data for specific geographic locations.

Region-specific entities have been added for the following countries.

- Australia
- Canada
- New Zealand
- Turkey
- United States of America

Region-specific entities have been added for the following geographic regions.

- Americas (AMER) - US, Canada
- Asia-Pacific (APAC) - Australia, New Zealand
- European Economic Area (EEA) - Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, United Kingdom, Greece, Croatia, Hungary, Ireland, Iceland, Italy, Liechtenstein, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia
- Europe, Middle East, Africa (EMEA) - Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, United Kingdom, Greece, Croatia, Hungary, Ireland, Iceland, Italy, Liechtenstein, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, Turkey

- New repository template to allow for creation of repositories based on common criteria. Default region-specific repository templates have region-specific entities pre-selected.
- Repository selection on the Data Volume dashboard now applies to all data presented.
- Improved entity selection when creating a tag.
- Increased maximum value for weighted labels to allow for broader sensitivity range. New maximum value is 10,000.
- On the Grammars page, added a button to add an entity to the selected grammar.
- Added the ability to create a reporting group and weighted label when creating a tag.

Analyze

- From within the document view panel, links to external content are only clickable using the keystroke Ctrl+click. The destination opens in a new browser tab.
- Renamed the "privacy score" to "risk score" for clarity and enhanced the computation method based on an increased maximum value for weighted labels.

Manage

- New workbook template to allow for creation of workbooks based on common criteria.
- Updated workspace template to allow for selection of workbook template and data source template. Any workbooks or data sources defined in the workspace template are created when a workspace based on the template is created.
- New hold statuses, Release Pending and Release Failed, to provide more complete overview of the status of the hold release action.
- From the Map view of a workspace overview, added the ability to go to the documents represented by the blocks in the map. Hover over a block in the subordinate rows to view the details; click the VIEW CONTENTS link in the pop up to view the items related to that block.
- Renamed the "privacy score" to "risk score" for clarity and enhanced the computation method based on an increased maximum value for weighted labels.

When calculating the risk score, each sensitivity level may be configured with a weight from 0 to 10000. Each entity discovered in a document is scored based on that weight and the total score is computed. The total of the weighted entities is normalized to a total score between 0 (not sensitive) to 100 (extremely sensitive).

- Updated the workspace report to include an Overview section with additional charts.
- From within the document view panel, links to external content are only clickable using the keystroke Ctrl+click. The destination opens in a new browser tab.

Administration

- New Administration role and permission to provide access to Data Discovery APIs. The *API Developer* role, and permission of the same name, provide access to the APIs through the included Swagger interface.

Documentation

Data Discovery includes Help Centers that are incorporated into each User Interface and are updated with each software release as appropriate.

To view the Data Discovery documentation outside of the product, visit Support & Services on the Micro Focus web site, www.microfocus.com.

To navigate to the Data Discovery documentation

1. From the Micro Focus home page click **Support & Services** in the primary menu along the top of the page and then click **Support**.
2. In the Browse Resources section, click **Documentation**.
3. In the product selection list, begin typing **File Analysis Suite**. As you type, products matching what you type display; click **File Analysis Suite Data Discovery (FAS DD)**.

Select the desired release version.

You must have Adobe® Reader installed to view files in PDF format (*.pdf). To download Adobe Reader, go to the [Adobe](http://adobe.com) web site.