**opentext™**

OpenText

# Core Data Discovery & Risk Insights

Application Version 25.1.0

Processing Agent Version 25.1.100

Release Notes

## Legal notices

# Contents

# OpenText Core Data Discovery & Risk Insights Release Notes

**Application version:** 25.1.0

**Processing Agent version:** 25.1.100

**Publication date:** January 2025

This document is an overview of the changes made to the OpenText Core Data Discovery & Risk Insights application and processing agent.

## Support matrix

For information about the installation requirements and compatibility with other products, see the OpenText Core Data Discovery & Risk Insights Support Matrix. The support matrix may be updated between releases so it is available only from the Support portal.

## About OpenText Core Data Discovery & Risk Insights

OpenText Core Data Discovery & Risk Insights lets you find, protect, and secure sensitive and high-value data within on-premises and cloud data platforms across your enterprise. Identify, collect, and organize content to ensure discovery of sensitive data. Configure how structured and unstructured sources and datasets are processed and categorized with Connect. Analyze your data under management with Analyze. Organize, review, and take action on documents and unstructured data items with Manage.

## About this release

This product release and release notes cover Application 25.1.0 and Processing Agent 25.1.100. These release notes reference changes included in this release only. For information about changes made in prior releases, see the release notes for the desired release.

While the prior release's Processing Agent remains compatible with this application's version for existing functionality, product enhancements and fixes may require an upgrade to the latest version of the agent. In order to facilitate transition to the new version, the existing processing agent an be used for a transitional period of four (4) weeks after upgrade of the application to this release. Beyond this grace period, out of date agents will be automatically disconnected and an upgrade will be mandatory to reconnect and process data.

# What's new

The following features and enhancements are included in this release.

## General

The following features and enhancements apply to more than one component in OpenText Core Data Discovery & Risk Insights 25.1.0.

- Enhanced support for grammar search. You can now set the desired minimum confidence score for the selected grammar rule to assist in reducing false positives. You can add the confidence score for a grammar when searching a content list in Analyze or Manage using the search builder, when creating dynamic tags that include grammar rules in Connect, and when creating a query or dynamic query workbook in Manage. You can also include the confidence score when creating an unstructured data workspace that includes topic-based data requests in the criteria or includes datasets associated with sources based on custom adapters.

  For each grammar rule added, you can set the confidence level. The presumed default confidence is 50% (50% confidence that the identified values are for the identified grammar type).

  If you wish to edit the confidence level, click the **CONFIDENCE** button above the grammar value that has been added and then click **Edit the minimum threshold**. Use the up or down arrow as necessary to define the desired minimum confidence level.
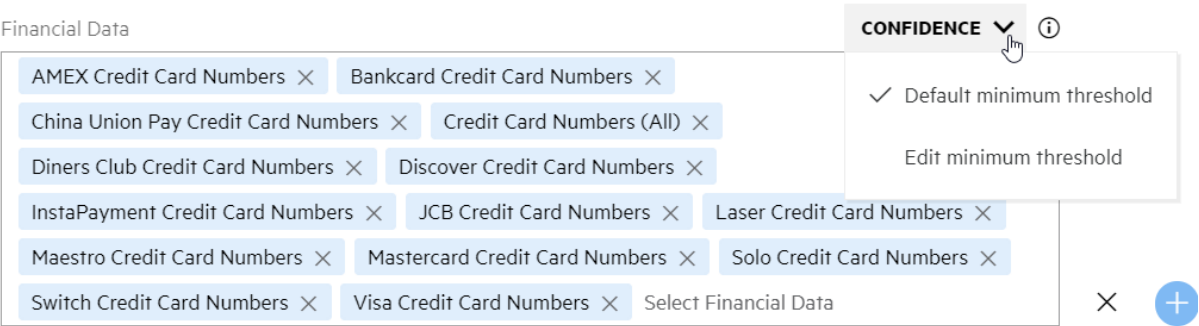


- Enhanced support for external metadata.

  ○ External metadata can now be processed for SharePoint, SharePoint Online, and Google datasets and is configured as part of the primary capture rules.

  ○ Added external metadata as a field that can be exported to CSV from a content list in Analyze and Manage.

  ○ Added external metadata as a field that can be added to a metadata field template in Manage. For more information about field templates, see "Field Templates" in the Help Center.

- Now supporting OpenText IDOL Media Server 24.4.x to perform optical character recognition.

  > **NOTE:** You will need to upgrade your implementation of Media Server to 24.4.x to support OCR in OpenText Core Data Discovery & Risk Insights. This is the only version of Media Server supported with this release

- Now capture additional metadata for data items. You can view this metadata in the content detail panel in both Analyze and Manage. You can also query for items by this metadata using the search builder and include this metadata when you export a content list to CSV.

  - *created by* - Defines the person that created the Office document. This is the *Author* property on the Office document.

  - *last accessed date* - Defines the date the item was last accessed on the original source. Captured on the initial scan or when items have been modified and then picked up during an additional scan. Applies to file system and SharePoint items.

    > **NOTE:** If the item was accessed but not modified since the initial scan, the last accessed date is not updated on a rescan.

    For file system items, if the last access date is not enabled in your company environment, this date is the last modified date. For SharePoint items, this date may not exist. In this scenario, the last accessed date will be blank.

  - *last modified by* - Defines the name of the person that last modified the Office document on the original source.

  - *owner* - Defines the owner of the items. Applies to file system and SharePoint items.

    SharePoint items can have multiple owners in SharePoint. OpenText Core Data Discovery & Risk Insights captures owners, based on availability, in the following order: SharePoint item owner, SharePoint site owner, SharePoint group owner, SharePoint site group owner. For OneDrive, the user is the owner.

## Processing Agent

The following features and enhancements apply specifically to the OpenText Core Data Discovery & Risk Insights 25.1.100 processing agent.

- Update the path to the common data store on the machine hosting the processing agent. The path is referenced in the Help Center in the procedure for installing the agent. This is related to a clean re-install of the agent on the agent host.

  - previous path: `C:\ProgramData\Fusion\Agent`

  - new path: `C:\ProgramData\OpenText\Core Data Discovery and Risk Insights\Agent`

# Connect

The following features and enhancements apply specifically to Connect, and are in addition to changes already mentioned in relation to general enhancements.

- Added primary capture rule criteria for Extended ECM datasets. You can now limit your dataset items by file extension.

- The Manage Datasets list and detail panel now auto-refresh every 10 seconds.

- Added a filter option, **SOURCES**, to the Manage Datasets page to let you display datasets for all sources or for a specific source.

- On the Data Volume, Manage Sources, and Manage Datasets pages, replaced the preferences icon (⚙) to view data size by either root or root and children with a dropdown. Click "Root Documents " or "Root + Child Documents" to change the view.

- Added data deletion tracking for Documentum sources. OpenText Core Data Discovery & Risk Insights tracks deletions of managed Documentum items using the Documentum audit trail. OpenText Core Data Discovery & Risk Insights uses the **dm_destroy** event from the Documentum audit trail to identify deleted Documentum items to be removed from the application. Without this information, items deleted from your Documentum environment cannot be removed from the application. Documentum items that have been added or modified are appropriately updated in OpenText Core Data Discovery & Risk Insights.

  For example, if your Documentum audit trail is configured to be purged after 90 days, verify that your Documentum datasets are updated at least every 89 days.

  > **CAUTION:** Failure to rescan Documentum datasets before the Documentum audit logs are purged will result in items being tracked incorrectly in OpenText Core Data Discovery & Risk Insights. Using the same example, if your Documentum audit trail is configured to be purged after 60 days and your Documentum datasets are updated every 90 days, you will lose 30 days of important information about deleted items—items deleted during this 30 day time frame will not be removed from the application.
  >
  > The loss of information cannot be reconciled in OpenText Core Data Discovery & Risk Insights; you would have to create a new dataset and start over.

- Prevent creation of a new dataset that is the same as a dataset that is currently being deleted. If identical source and path (directory or sub-directory of) criteria are defined for a new dataset, you will see an appropriate message in the dataset creation dialog.

# Analyze

The following features and enhancements apply specifically to Analyze, and are in addition to changes already mentioned in relation to general enhancements.

- Added the ability to perform file analysis on unstructured data items from the Research content list. Similar to performing file analysis on workbook items, you can choose to also store content,

perform OCR, and extract grammar values.

To perform file analysis from the content list, click ⊞ in the action ribbon above the content list and then click either **Analyze All** or **Analyze Selected**. In the Analyze Files dialog, you can select to also store content, perform OCR, and extract grammar values, or you can choose to just analyze files.

- When adding items to a new or existing workspace in Manage from the content list in Analyze, the workspace now opens in a new browser tab instead of the existing tab.

# Manage

The following features and enhancements apply specifically to Manage, and are in addition to changes already mentioned in relation to general enhancements.

- Simplified the process for deduplicating data to compare essential content (such as text, images and formatting) for Office documents and, for all other document types, compare documents based on the complete binary fingerprint including content and any embedded metadata.

  Previously, you had to create two deduplication workbooks, one selecting **Essential data** to compare essential data only for Office documents and a second one selecting **Full binary data** to compare the complete binary fingerprint for remaining non-Office data items.

  With this release, the **Essential data** deduplication option has been replaced with **Enhanced data** to compare essential content for Office documents and binary binary fingerprint for non-office documents in a single workbook. The **Full binary data** option is also available, as previously.

  **Enhanced data** has also been added to the available field output for exporting a list of data items to CSV in Analyze and Manage, and to the field templates in Manage.

- Added the ability to select the columns to display and to resize the columns in the Manage Workbooks list. Click and drag a column resize bar ( ⁝ ) to adjust the width of the columns as desired. Click the preferences icon (⚙) above the workbook list to select or deselect the columns to include in your view.

- Added an option to jump to recent workspaces from the initial workspace Overview page and the Manage Workspaces list page. In both scenarios, this is a menu accessible from the blue down caret to the right of the OVERVIEW and WORKSPACES page headers. The menu includes the workspaces needing your review (maximum 5, then scroll) followed by your recently viewed non-closed workspaces (maximum 5 and search option).

- When sending all items to a target or sending masked versions of items to a target from the context of a workbook, added the option to reflect the folder hierarchy of the original source location.

  > **NOTE:** This option does not apply to Content Manager.

  From the SEND TO TARGET or SEND MASKED dialog, do one of the following.

- Select **No folders** to send the files to the selected destination directory with no folder hierarchy.

- Select **Use source folder hierarchy** to send the files to the destination directory and replicate the original source folder structure. Set the number of folders **above** the file to maintain within the destination directory (10 maximum). The number of folder levels defined is beneath the destination directory.

> **Example**:
>
> - A file, `vision-overview.pdf` was captured from the following path on the original source host.
>
>   `\\example.host.com\HR\benefits\2025\January\prep`
>
>   You select a destination with a relative path of **reviewed** and select **No folders**.
>
>   The file is sent to the target. The relative path to the file is `\reviewed\vision-overview.pdf`.
>
> - A file, `vision-overview.pdf` was captured from the following path on the original source host.
>
>   `\\example.host.com\HR\benefits\2025\January\prep`
>
>   You select a destination with a relative path of **reviewed**, select **Use source folder hierarchy**, and set the folder hierarchy to **4** folder levels.
>
>   The file is sent to the target. The relative path to the file is `\reviewed\benefits\2025\January\prep\vision-overview.pdf`.

- Added deleted items counts in the following locations.
  - to the workspace Overview detail panel, add deleted item file size (next to delete count)
  - to the workspace report
  - to the workbook detail panel, add deleted item file size (next to delete count)

- Removed the option to export items in a workbook. You can send items or send masked items to a target.

- Added the ability to change a workspace state from the workspace detail panel on the Manage Worskpaces page. Previously, you could see the status in the detail panel, but you had to open the workspace to change the state from the individual workspace Overview page.

  From the workspace list page, click in the row for the desired workspace and open the detail panel (). At the top right of the detail panel, click the status button and select the desired status. The button showing the current status will be DRAFT, OPEN, or CLOSED.

- For new tenants as of this release, the default Trivial Data and Obsolete Data workbook templates are no longer restricted to gathering items from only file system, SharePoint, and Google Drive sources.

  Deletion tracking is not supported for Content Manager sources. If the workbook based on either of these templates includes Content Manager items and you attempt to delete those

items, the workbook activity will return errors for those documents, the same as it would if you tried to delete Content Manager data from a non-template workbook.

## Administration

The following features and enhancements apply specifically to Administration, and are in addition to changes already mentioned in relation to general enhancements.

- Added **Close and reopen workspaces** permission to the Manageapplication level permissions. The user must have the same permission at the Workspace Security level to close and reopen workspaces. Previously, a user had to have the **Manage workspaces** permission as the application level and allowed the user to create, edit, and close workspaces. In combination with the existing Workspace Security permission, this new permission allows a user to close or reopen workspaces without the ability to create or edit workspaces.

## API updates

The following updates have been made to the OpenText Core Data Discovery & Risk Insights APIs.

- To enhance API-based custom action, the `GET '/ca/v1/workspace/{workspaceId}/workbook/{workbookId}/documents'` method now supports returning deduplication information such as the standard `'metadata'` and `'csv export'` methods do.

  The method now accepts the following new query parameters.

  - `includeDedupeDetails - boolean` which, if true, will return the following additional deduplication properties, assuming this is a deduplication workbook.

    If there is a top-ranked document,

    - `dedupeTopRankedDocumentId`
    - `dedupeTopRankedDocumentName`

    Or, if there are master datasets,

    - `dedupeDataSources`

  - `dedupeTopRankedDocumentQueryFields` returns additional fields from the top-ranked document. These are in addition to the id and title which are returned by default in `dedupeTopRankedDocumentId` and `dedupeTopRankedDocumentName`.

# Resolved issues for the application

The following issues have been resolved in this release of the OpenText Core Data Discovery & Risk Insights application.

- Resolved an issue where the date and time did not display in the processing error dialog for container file deletion tasks.

- Resolved an issue where you could not fully scroll the Users and Groups lists in Administration at reduced browser resolutions.

- Resolved an issue where the User column on the Processing page in Connect was too narrow to display the content.

- [345272] Resolved an issue where, after collapsing an activity group with a large number of sub-activities, the scroll bar is no longer present. In this scenario, a "Load more results" link has been added.

- Resolved an issue to ensure container files that were processed for metadata only do not show the items contained within the container.

- [353455] Resolved an issue where duplicate Exchange items were ingested during a rescan a dataset.

- Resolved an issue where the data lines for "Overall items managed" and "Analyzed Items" were reversed in the Unstructured Data > Analyzed Sensitive Data dashboard in Analyze.

- Resolved an issue where the error details for a failed agent activity workunit did not display.

- [362842] Resolved an issue where performing a CSV export of documents would fail if the "Protection State" field was selected and any of the documents contained MIP labels.

# Resolved issues for the processing agent

The following issues were resolved in this release of the OpenText Core Data Discovery & Risk Insights processing agent.

- [339410] Resolved an issue where Extended ECM and Documentum items that contain special characters in the file name failed to be ingested.

- Resolved an issue where a scan of a SharePoint Online failed with an error that the connection request timed out. In this scenario, the connection to SharePoint Online was previously not retried.

- Resolved an issue where items could not be send to a Documentum destination that included a special character in the file path.

- Resolved an issue where data items with a content type of "application/octet-stream" were incorrectly ingested as an unknown file type.

- Resolved an issue where items with special characters in the file name could not be sent to an Extended ECM target.

# Known issues

The following issues are present in this release of OpenText Core Data Discovery & Risk Insights.

- None

# Documentation

OpenText Core Data Discovery & Risk Insightsincludes a single Help Center that is incorporated into each User Interface and is updated with each software release as appropriate.

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

**To navigate to the OpenText Core Data Discovery & Risk Insights documentation**

1.  From the primary menu on the Micro Focus home page, click **Support & Services** and then click **Support**, then **Documentation**.

2.  In the product selection list, begin typing **OpenText Core Data Discovery & Risk Insights**. As you type, matching product names display; click **OpenText Core Data Discovery & Risk Insights**.

    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 web site.