



OpenText Core Data Discovery & Risk Insights

Application Version 25.2.0_25.2.1

Processing Agent Version 25.2.101

Release Notes

Document Release Date: April 2025
Software Release Date: April 2025

Legal notices

Copyright 2019-2025 Open Text

The only warranties for products and services of Open Text and its affiliates and licensors ("Open Text") 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. Open Text shall not be liable for technical or editorial errors or omissions contained herein. The information contained herein is subject to change without notice.

Contents

- OpenText Core Data Discovery & Risk Insights Release Notes 1
 - About OpenText Core Data Discovery & Risk Insights 1
 - About this release 1
- What's new 2
 - General 2
 - Processing Agent 4
 - Connect 5
 - Analyze 6
 - Manage 7
 - Administration 8
 - API updates 8
- Resolved issues for the application 10
- Resolved issues for the processing agent 12
- Known issues 13
- Documentation 14

OpenText Core Data Discovery & Risk Insights Release Notes

Application version: 25.2.0_25.2.1

Processing Agent version: 25.2.101

Publication date: April 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 25.2.0_25.2.1 and Processing Agent 25.2.101. 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 can 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 25.2.0_25.2.1.

- Added the ability to delete closed workspaces.
 - In the workspace list, a delete icon (X) has been added to the hover-over inline icons for closed workspaces. A **Delete** button has also been added to the detail panel for closed workspaces.
 - User must have the new **Delete any workspace** permission for Manage (application level). This permission is not automatically assigned to any roles.
 - The action is tracked in the Audit Log in Administration. Previous actions for the workspace are retained and viewable in the Audit Log.
- Added the ability to search for deleted documents in both Analyze and Manage. You can view limited details for deleted documents so long as the parent dataset has not been removed from the application.

From the action menu above the content list, click  > **View Deleted Documents**. You can export the list of deleted documents to CSV or further refine the list using the filter panel. For full details, see "Viewing deleted documents in Analyze" in the Help Center.

- When a grammar confidence score is selected as part of the criteria, the defined confidence score is included in the summary view in the following locations.
 - In the new and edit workspace and workspace template wizards, added to the data request summary for the workspace, to the criteria summary for a workspace, and to the overall workspace summary.
 - In the workspace and workspace template detail panels, added to the criteria information.
 - In the new and edit workbook and workbook template wizards, added to the criteria summary for a workbook and to the overall workbook summary,
 - In the workbook and workbook template detail panels, added to the criteria information.
 - In the tag detail panel, added to the criteria information.
 - In the search builders, added to the search criteria that displays on hover over for the VIEWING box above the filter panel.
- Added support for copy and paste of multiple values using the search builder in Analyze and Manage and selecting criteria for tags in Connect. This functionality is also available when

selecting criteria for custom datasets in workspaces and workspace templates, and for selecting criteria for query and dynamic query workbooks and workbook templates.

The following fields support entering multiple values and are available as appropriate for the context.

- CM Classification and CM Unique Identifier, values separated by a comma or on a new line
- Document ID, values separated by a space, a comma, or on a new line
- File Extensions, values separated by a space, a comma, or on a new line
- Job Run ID , values separated by a space, a comma, or on a new line
- Updated the label for "Processed Date" to "First Scanned Date". The label for this date did not represent the value that was being presented. The First Scanned Date defines the date that the document was initially scanned by the processing agent.
- Updated the label for "Last Processed Date" to "Last Scan Update Date". The label for this date did not represent the value that was being presented. The Last Scan Update Date defines the date that the document was last scanned by the processing agent. A file is re-scanned if it has been modified at the original source location or if a reprocessing action that requires going back to the original source location has been initiated (such as re-analysis when the content has not been stored).
- For workbook list page, your preferences for the column selection and resize are now saved for the browser until the cache is cleared.
- The following grammar updates are included in this release.
 - Added the following entities for Thailand.
 - Bank account and routing numbers
 - Driving license
 - Health ID (Thailand use the national UIN [unique identification number] in the health sector as UHI [Unique health identifier])
 - Email addresses
 - Names
 - National ID
 - Nationality
 - Passport number
 - SWIFT code
 - Telephone numbers
 - TIN

Processing Agent

The following features and enhancements apply specifically to the OpenText Core Data Discovery & Risk Insights 25.2.101 processing agent.

- [364814] Resolved an issue to where an agent activity for a SharePoint scan completed with an ingestion error stating that the list view threshold was exceeded. In this scenario, the SharePoint log files were not available because the SharePoint log retention period had expired. With this resolution, there is no error but a message is added to the SharePoint log in OpenText Core Data Discovery & Risk Insights about the log retention period and performing scans more often. All files are scanned as if the rescan is the initial scan
- Updated processing to exclude extraction from additional known non-textual files, such as operating system files.
- [349245] Added advanced configurations to the Agent Admin UI to define a hard blackout period for scanning. When you set a blackout period for an agent cluster or a dataset, no new job runs start during that time but jobs that have already started will continue until completed. If you use the new advanced configurations, no new jobs start during the blackout period and jobs that are in progress are stopped. The stopped in progress jobs resume where it left off when the blackout period stops. The advanced blackout configurations in the Agent Admin UI override any blackout periods defined for an agent cluster or dataset.

To configure the advanced blackout period for the processing agent

1. On the agent host machine, open **OpenText Core Data Discovery & Risk Insights > Agent Admin UI**.
2. In the left navigation, click **Advanced**.
3. In the category list, select **Common**.
4. Complete the following configuration options for the blackout period.

Option	Description
Blackout Period Duration (minutes)	Type the number of minutes that represents the duration of time for the blackout period. During this time no processing jobs run. All current jobs are paused and will resume after the blackout period duration has ended.
Blackout Period Start Time	Using cron format, define the time and days to start the blackout. The day and time is local to the agent host machine. Example: To start the blackout period every weekday at 9:00 AM, type <code>0 9 * * 1-5</code>

5. Click **Save**.

The defined advanced blackout period is saved and overrides any blackout period defines for the agent cluster to which this agent belongs and any datasets associated with the agent cluster.

Connect

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

- When connecting to Structured Data Manager, the URL format for JDBC sources using MySQL has changed from `jdbc:mysql://IP:PORT/schema-name` to `jdbc:mariadb://IP:PORT/schema-name`.

Example: `jdbc:mysql://198.51.100.0:3306/mysql` should now be `jdbc:mymariadb://198.51.100.0:3306/mysql`

- For Extended ECM datasets, you can now identify one or more Node IDs. A Node ID is the value of the "Item ID" property for the object in Extended ECM you want to process data from. The item ID can reference the workspace (default is Enterprise), a folder, or a container. You define the Node ID as part of the primary capture rules for the dataset.
- Added the ability to scan Exchange datasets for an individual mailbox. Previously you could specify an OpenText Core Data Discovery & Risk Insights user group. You could filter for an email address but it was not specific to a mailbox, just the reference to that email address (sender, recipient, mention).

At the source level, you can specify whether the data is to be **Based on a pre-defined user group** or **Based on an individual mailbox email address**. The option specified at the source level affects the options you can set at the dataset level.

- If based on user group, the dataset Primary Capture Rules options are as before and no changes. An information note in a blue banner has been added to the page, "Dataset scan is based on a pre-defined user group set at the source."
- If based on an individual mailbox, you can define a **Mailbox address** as part of the Primary Capture Rules options. An information note in a blue banner has been added to the page, "Dataset scan is based on an individual mailbox email address."
- Added support for sending documents to a Google Drive target.
 - When defining the access account for Google Drive targets and defining a shared drive, the account must have Editor permission if it is not the owner of the shared drive.
 - For Google Drive destinations, the relative path is case sensitive. If the case in the path you type is different than the case in the path on Google Drive, a new folder is created at the level of the case difference. For example, the relative path on Google Drive is `hr/recruiting` and you type `hr/Recruiting`. The end result is that items sent to this destination will be sent to a new sub-directory `/Recruiting` under the `/hr` directory alongside the `/recruiting` sub-directory.

Destination paths (or any portion of) associated with a given target must not overlap in any way.

For full details, see "Manage Targets", "Manage Destination", and "Google Drive connection" in the Help Center.

- Added the ability to define patterns that let you *exclude* specific information from the built-in grammars. Using Lua script patterns, you can define the grammar exclusions at either the grammar rule level or the grammar type level. If you add exclusions to the grammar type, the exclusions are inherited by all grammar rules under the grammar type.

For full details, see "Grammar exclusions" in the Help Center.

- [375550] For datasets based on custom adapters, you can now edit the custom adapter configuration fields even after data has been scanned.
- [379754] Removed the VAT Number grammar rule from applicable built-in grammar sets as VAT numbers are publicly available. Datasets that include these grammar sets will display a **Re-analyze** link next to "Grammar Sets" in the dataset detail panel. The VAT Number grammar rule remains available for inclusion in criteria selections.

The following grammar sets are affected.

- Australian Privacy Principles
- Brazil General Data Protection Law
- Canada Personal Information Protection and Electronic Documents Act
- EU General Data Protection Regulation
- EU General Data Protection Regulation (UK)
- New Zealand Privacy Act
- Turkey Personal Data Protection
- Better handling of distribution of work by cloud-to-cloud processing agents.
- Added a "spinner" to the custom grammar rule test function to let you know the testing is still in progress.

Analyze

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



- Updated dashboards.
 - Added more detailed information for the total number of items identified by grammars as being sensitive on the Unstructured Data > Grammar Analysis Trend dashboard. Previously only the total number of sensitive items was provided. You can now see the number of items identified as sensitive and the total number of items scanned, as well as the percentage of sensitive items compared to the total number.
 - When you click the link or view icon to view items related to a block or tree view in the Unstructured Data > Sensitive Data Heat Map, the content list of the data items now opens in a new tab.

Manage

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

- Added the ability to remove stored content for all documents within a workbook and to remove extracted grammar values for all documents within a workbook.
 - User must have the "Manage workbooks" permission at both the Manage and Workspace Security levels.
 - In the FILE ANALYSIS section of ACTIVITY tab in the workbook detail panel, a new action box has been added. The options are **Analyze files in workbook** (existing functionality) and **Remove analysis of files in workbook** (new functionality).

Select **Remove analysis in workbook**, click one or both of the following, and then click the **REMOVE** button.

- Click **Remove Stored Content**  to remove the content that has been stored in the index for each document in the workbook.
 - Click **Remove Extracted Grammars**  to remove extracted grammar values for all documents within the workbook. The association to an identified grammar rule (such as "Email addresses") is retained; only the grammar value (such as "jsmith@mailserver.com") is removed from the index.
 - When the analysis is in progress, the **REMOVE** button is replaced with a **CANCEL** button to allow you to cancel the request.
 - Remove and cancel analysis actions are included in the Audit Log in Administration and in the workspace Report and Audit in Manage.
- Added the option to create workspaces that refresh upon manual request. Prior to this release, workspace content is refreshed automatically. Refreshing a workspace retrieves documents that may have been ingested since the workspace was last scanned. Refreshing also updates query workbooks within the workspace.

When creating a new workspace or workspace template, you now have the option to select whether a refresh of the workspace content occurs automatically or if you have to manually refresh the workspace (default).

NOTE: This selection cannot be changed after the workspace is created.

- Select **Manual Refresh** to refresh the data items in the workspace and workbooks (if applicable) only when you initiate a refresh request.
 - Dynamic query workbooks cannot be created in workspaces with Manual Refresh.
 - A refresh button will display in the workspace detail panel on the right side of the workspace Overview page. The user must have the **Manage > Manage workspaces** and **Workspace Security > Edit workspace metadata** permissions to see the refresh

button.

- "Last Refreshed" date has been added to the workspace detail panel.
- Select **Automatic Refresh** to refresh the data items in the workspace on a continual basis as data matching the workspace criteria is identified.
 - Data requests (limiting criteria to a data subject or specific topic) cannot be defined for workspaces with Automatic Refresh.
 - Workspaces that exist prior to this release are set to automatic refresh.
- Improved handling and display of a large number of grammar rules matched and grammar rules masked in the content detail panel.
- In the workspace list, merged the indicator that the logged in user has a review pending into the current Status column.
- Updated the label for the deduplicate workbook option "Compare against official records" to "Compare against master records". This change makes the correlation between the option selected and the data point seen.

In addition, a document ID for one of the master records (if more than one) is now shown instead of the dataset ID.

Administration

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

- Updated the name of the "Analyze and collect data" permission for both the Manage and Workspace Security levels to "Collect data". The access for these permissions has not changed, just the label.

API updates

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

- With the removal of the internal concept of a workspace "datasource", the following API changes have been made.
 - The following APIs have been removed.
`POST /ca/v1/workspace/{workspaceId}/datasource`
`GET /ca/v1/workspace/{workspaceId}/datasource/{datasourceId}/metrics`
 - The following APIs now relate only to "Research" datasets.
`PUT /ca/v1/workspace/{workspaceId}/datasource/{datasourceId}`
`GET /ca/v1/workspace/{workspaceId}/datasource/{datasourceId}`
 - Instead of creating datasources and adding them to workspaces, the equivalent now is to

add datasets directly to the workspace using the `primaryDatasets` property of the `/workspace create` and `edit` methods.

- The following deprecated data subject methods have been removed.

`POST /ca/v1/workspace/{workspaceId}/data-subject`

`PUT /ca/v1/workspace/{workspaceId}/data-subject/{dataSubjectId}`

Similar to `datasources`, data subjects are added directly to the workspace using the `dataSubjects` property of the `/workspace create` and `edit` methods.

- All deprecated `/case` methods have been removed. Use the equivalent `/workspace` methods instead.
- When creating a workspace, the workspace will now default to the new Manual refresh type (recommended). If you wish to continue creating Automatic workspaces, set the `refreshType` property of the payload to `AUTO`. You may also explicitly set `refreshType` to `MANUAL` if you want to create a manual refresh workspace.
- Due to changes in deduplication to support multiple master documents, the response data of the following method has changed.

`POST /ca/v1/workspace/{workspaceId}/document/{documentId}/metadata`

- The `dedupeTopRankedDocumentId` and `dedupeTopRankedDocumentName` properties have been replaced by `dedupeTopRankedDocuments`.
- The type of the `dedupeTopRankedDocumentMetadata` property has changed to a map in order to support multiple master documents.

For more information about these properties, see the descriptions in Swagger UI.

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 a scan or rescan of an Exchange dataset resulted in a document count scan count of 0. This occurred when the Exchange group defined in Administration had been deleted. With this resolution, you will now see a warning that the group associated with the dataset no longer exists.
- Resolved an issue where the User field for send masked to target actions was not populated on the Processing page in Connect.
- Resolved an issue where the Data Volume Analysis dashboard in Analyze and the Manage Datasets page in Connect loaded slowly when the tenant includes a large number of datasets.
- [358526] Resolved an issue where an agent scan activity status failed to update in the Agent Activity list in Connect.
- [358526] Resolved an issue where the collection of agent logs failed from the agent detail panel on the Agent Clusters list in Connect.
- Resolved an issue where the notification banner about workbooks awaiting your review would disappear when the workbook list automatically refreshed and there are no workbooks needing your review that are currently visible on the screen (without scrolling). The "My Approvals" view toggle would also disappear in this same scenario.
- Resolved an issue where, in the SOURCES filter on the Manage Datasets page in Connect, sources with long names wrapped and overlapped the next source name. With this resolution, long source names are truncated in the SOURCES filter.
- Resolved an issue where the Date and Size columns were not sortable on the Manage Datasets page in Connect when viewing by datasets (instead of grouped by source).
- Resolved an issue where the tree view for the Unstructured Data > Sensitive Data Heat Map did not use the available vertical space.
- Resolved an issue where the swagger openAPI definition failed parsing.
- [368060] Resolved an issue where the single sign-on options on the login page were disabled after upgrading the application.
- Resolved an issue where the document counts on the General tab of the workbook detail panel in Manage were not clickable to go to the content list filtered for those documents.
- Resolved an issue where re-analysis of a workbook did not update the Last Processed Date for the items in the workbook. The label for this date did not represent the value that was being presented and has been updated to "Last Scan Update Date". The Last Scan Update Date defines the date that the document was last scanned by the processing agent. A file is re-scanned if it has been modified at the original source location or if a reprocessing action that requires going back to the original source location has been initiated (such as re-analysis when the content has not been stored).

- Resolved an issue where any custom feature that was enabled for a workspace was not showing as enabled when you edit the workspace.
- Resolved an issue where identified custom grammar values that include a percentage symbol (%) were not highlighted in the content view panel.
- [377629] Resolved an issue where adding a confidence score to an existing tag that includes the "Healthcare ID Number (Worldwide CJK)" grammar rule resulted in the tag not being saved.
- Resolved an issue where the OCR'd content of embedded images or the content of embedded files were not viewable in the content panel.
- Resolved an issue where multiple spaces between words or characters were not being properly displayed in the content view panel.
- Resolved an issue where special characters were not supported in the dataset file extension exclusion list. The following special characters are now allowed in the extension exclusion:
_ - & # + = ^ \$ ~ @
- Resolved an issue where a workspace created using the API without defining a user or group to a security role could not be opened.

Resolved issues for the processing agent

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

- Resolved an issue where a dataset and a workspace based on that dataset and all other identical criteria resulted in different document counts.
- Resolved an issue where documents deleted from the original Google Drive source location were not reported as "No longer at source" in the dataset or agent activity detail panels.
- Resolved an issue where 0 byte items failed to be sent to a Documentum target with an exception that the file content failed to upload.
- Resolved an issue where only items with a content type in Extended ECM of Document were being scanned for Extended ECM datasets. Now, items with a content type in Extended ECM of Emails are included in the scans of Extended ECM datasets.
- Resolved an issue where SharePoint Online (O365) could not connect using a configured web proxy.

Known issues

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

- None

Documentation

OpenText Core Data Discovery & Risk Insights includes 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](http://adobe.com) web site.