Relativity aiR

Collect - ChatGPT

November 13, 2026

For the most recent version of this document, visit our **documentation website**.

Table of Contents

1 Collect **4**

1.1 Data sources 4

1.2 Commercial and Government support 4

2 Installing Collect **4**

2.1 System requirements for Collect 4

2.2 Installing Collect 5

2.2.1 Installing Collect from the application library 5

2.3 Permissions to run Collect 6

2.3.1 Feature permissions 6

2.4 Uninstalling the application 6

3 Matters **6**

3.1 Creating a matter 6

3.2 Matter Details layout fields 7

3.3 Viewing or editing matter details 7

4 Custodian targets **8**

4.1 Custodians 8

4.2 Fields 8

4.3 Creating a custodian target 9

4.3.1 Generating targets in the wizard 10

4.4 Removing custodian targets 10

5 Data sources **10**

5.1 Creating a collect data source 10

5.1.1 Third-party vendor 11

5.2 Data source types 11

6 ChatGPT Enterprise data source **11**

6.1 Considerations 12

6.2 Prerequisites 12

6.3 Collection size 12

6.4 Collected metadata 13

6.5 Creating the data source 13

6.6 Settings fields 14

6.7 Configuring the data source 14

7 Collection **15**
7.1 Considerations 15
7.2 Best practices for time-sensitive collections 15
7.3 Creating a collection 15
7.4 Using the Collect wizard 16
7.4.1 Collection Details 16
7.4.2 Data source 17
7.4.3 Custodians 17
7.4.4 Non-custodial 18
7.4.5 Collection Summary 18
8 Viewing or editing Collection data **19**
8.1 Collection details 19
8.2 Viewing collected data 20
8.3 Collect console 21
8.3.1 Collection 21
8.4 Delete collections 22
9 Grouped collections **22**
9.1 Grouped data sources 23
9.2 Grouping messages 23
9.3 Custodian assignment considerations 23
9.4 Manual processing guidance 24
10 Migrate Collect workspaces **24**
10.1 Migrate Collect workspaces 24
10.2 Credentials do not migrate 24
10.3 After migration 24
10.3.1 Considerations 25
11 Identifying Collection data in Staging Explorer **25**
12 Reports **27**
12.1 Running reports 27
12.1.1 Collection Summary report 27
12.1.2 Collection Details report 27
12.1.3 Results report 27
12.1.4 Errors report 28
13 Status Summary **28**

13.1 Job status 28
13.2 Reviewing job statuses 28
14 Target Status **29**
15 Monitor **29**
16 Proprietary Rights **30**
17 Index **31**

1 Collect

Use this application for collecting your custodian's data through different sources. Start by setting up Collect as an application within your data source. For information on registering an app, see the source's documentation on their website. Once registered, start adding custodians, data sources, and targets to Collect. Once connected, start the collect job and begin collecting data from custodians.

Note: This document covers the Box source.

1.1 Data sources

In Collect, you can collect data from a variety of data sources. For more information, see Data sources on page 10.

- **ChatGPT Enterprise**—collect data from ChatGPT Enterprise. For more information, see ChatGPT Enterprise data source on page 11.

1.2 Commercial and Government support

There are no other limitations compared to Relativity aiR commercial instances aside from data source connector availability.

Use the following table to see available connectors in Relativity aiR Government instances compared to commercial instances.

Note: Inactive mailbox, and Inactive OneDrive for Government is currently only supported for customers in Microsoft GCC tenants.

2 Installing Collect

You can install Collect in a workspace by using the functionality available through the Application Deployment System (ADS). This system provides you with the option to install Collect by selecting it from the list of existing applications in the Application Library tab or by importing it from an external application file.

To install Collect, install Collect from the Application Library tab and, if required, enable access for the data source.

Confirm that you have the appropriate system admin permissions to install an application. For more information, see Workspace security on the Relativity aiR documentation site.

2.1 System requirements for Collect

Collect uses the ADS framework, so you install it as an application within a Relativity instance. Consequently, Collect has the same system requirements as Relativity aiR. For Relativity aiR's system requirements, see System Requirements on the Relativity aiR Documentation site.

2.2 Installing Collect

Collect is compatible with Relativity aiR. See Getting started in Relativity aiR on the Documentation site for requirements.

For a Collect-only installation, you do not need the following pre-requisities:

- Analytics server setup
- Database server for processing or native imaging
- Worker server for processing or native imaging
- Obtaining applications for native imaging and processing

Because Collect uses the ADS framework, you can install through the Relativity Application tab from the library. See Installing Collect from the application library below.

Note: You configure security permissions on Collect just as you would for any other Relativity application. For more information, see Workspace security on the Relativity aiR documentation site.

2.2.1 Installing Collect from the application library

If Collect is in the application library, you can install it to the current workspace. Confirm that you have the appropriate system admin permissions to install an application. For more information, see Workspace security on the Relativity aiR Documentation site.

Note: Analytics, Case Dynamics, Collect, Legal Hold, and Processing all share the Entity object. You may be prompted to complete additional steps to unlock and resolve conflicts of the listed applications in order to complete installation. For information, see Troubleshooting application installation errors on the Relativity aiR documentation site.

Use the following procedure to install Collect from the application library:

1. Navigate to the workspace where you want to install the application.
2. Navigate to the **Application Admin** tab.
3. Click **New Relativity Application** to display an application form.
4. Click the **Select from Application Library** radio button in the Application Type section.
5. Click the ellipses button in the **Choose from Application Library** field.
6. Select Collect on the Select Library Application dialog. This dialog only displays applications added to the Application Library. If Collect is not included in the list, see the Installing applications topic.
7. Click **Ok** to display the application in the **Choose from Application Library** field. The application form also displays the following fields:
    - **Version**—displays the version of the application that you are installing.
    - **User-friendly URL**—displays a user-friendly version of the application's URL. This field may be blank.
    - **Application Artifacts**—displays object types and other application components.
    - **Map Fields**—there are no fields available in Collect for mapping.
8. Click **Import** to install Collect into the workspace.
9. Review the import status of the application. Verify that the install was successful or resolve errors.

2.3 Permissions to run Collect

The following security permissions are required to run and complete the collection process.

2.3.1 Feature permissions

Feature Permissions provide an alternative to Relativity's security management by shifting the focus from Object Types and Tab Visibility to feature-based permissions.

Any feature-specific permissions information already in this and other topics is still applicable. The Feature Permissions interface enables administrators to manage permissions at the feature level, offering a more intuitive experience.

By viewing granular permissions associated with each feature, administrators can ensure comprehensive control, ultimately reducing complexity and minimizing errors. For details see Instance-level permissions and Workspace-level permissions.

2.3.1.1 Actions

With feature permissions, you can select different actions each group can take. With Collect you can select the following actions:

- **View Collections**—view existing collection jobs and their details.
- **Manage Collections**—create and edit collection jobs.
- **Manage Collection Sources**—create and edit collection data sources in Collect Admin.
- **Manager Collect Targets**—create and edit custodian targets.
- **Delete Collect**—delete collection jobs.
- **Secure Collect**—manage security settings for collection jobs.

2.4 Uninstalling the application

If you need to remove the Collect application from your workspace, contact Relativity Support.

3 Matters

In Collect, a matter represents a legal action or case requiring you to collect electronic data from sources such as Box. Matters can be used to group multiple related collection jobs.

You manage matters that are associated with a collection. You can create each of these items on their respective tabs, or you can create them when you add a new collection.

Note: Matters created from Home aren't available for use in Collect nor listed on the Matters tab in this application. Additionally, the matters created on this tab are only available for use in Collect.

3.1 Creating a matter

Use the following procedure to create a matter that you can associate with a collection:

1. Navigate to the Matters tab. Collect displays a list of the active matters currently available to this application.
2. Click **New Collection Matter**.
3. Complete the fields in the Matter Details layout. See Matter Details layout fields on the next page.
4. Click **Save**. Collect displays the matter details. See Viewing or editing matter details on the next page.

You can also create a matter when you add a new collection. Click the **Add** link next to the Matter field in the Collection layout. See Creating a collection on page 15.

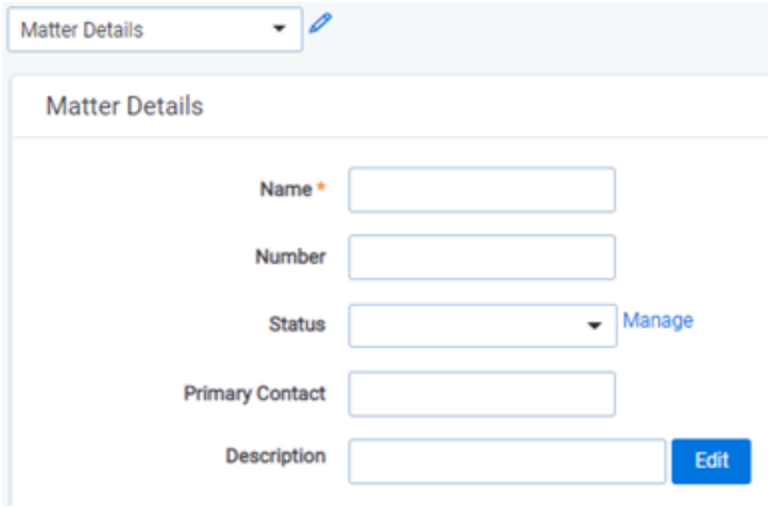
3.2 Matter Details layout fields

The Matter Details layout contains the following fields:

- **Name**—the name of the matter.
- **Number**—the number you assign to the matter for reporting purposes.
- **Status**—the status you assign to the matter for reporting purposes. Select an existing status from the drop-down menu or click **Add** to define a new one. Existing statuses include **Active** and **Closed**.

  Note: Assigning a status of Closed to a matter hides it from the Active Collect view on the Collect tab.

- **Primary Contact**—the name of an individual who handles communications related to the matter.
- **Description**—the description of the matter used for reporting purposes. Click **Edit** to display an HTML text editor where you can enter the description.



3.3 Viewing or editing matter details

Display the matter details by clicking the name of a matter on the Matters tab. Collect also displays these details immediately after you add a new matter. You can use the buttons at the top of the page to edit, delete, or perform other tasks with the matter.

The details page includes the following sections:

- **Matters Details**—lists the name, number, status, primary contact and description of the matter.
- **Collect**—lists all collections associated with the matter. You can also perform the following tasks in this section:
  - **Associate the matter with a new collection**—to create a new collection, click **New**. See Creating a collection on page 15.
  - **Remove a collection from Relativity**—click **Delete** to display a pop-up window. To view child objects and associated objects, click **Dependencies**.

- **Modify collection details**—click the **Edit** link for a collection. To modify the matter, click the **Edit** button at the top of the page.
- **Display the collection details**—click the name of the collection.

4 Custodian targets

In Collect, you manage custodians that have an associated data source. The custodian target is the combination of a custodian and data source. It is an endpoint from which Collect can connect to and collect from.

To collect from custodians, entities either have to already exist, be created, or be imported in Relativity. The custodian and their primary email address must also be associated to the data source. When the primary email address field for a custodian is present, Collect automatically generates the required targets for each data source while setting up a collection job.

4.1 Custodians

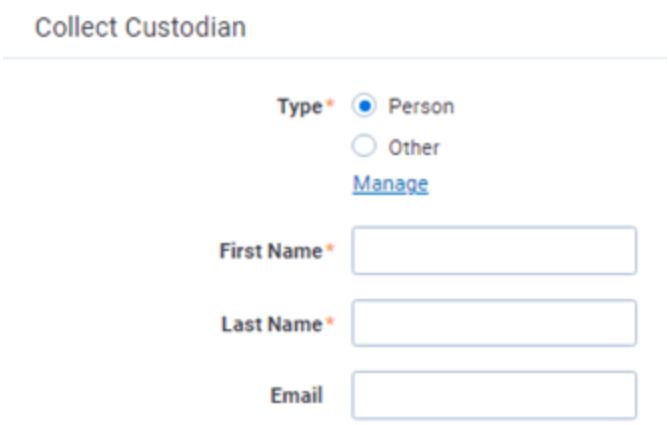
Once you create an entity in the Collect Custodian view and add it to a collect project, it becomes a custodian.

If a custodian does not exist, you can manually create an entity from the Entities tab. If a custodian does exist, you can create a custodian target. For more information, see Creating a custodian target.

You can add custodians to Collect at different times throughout the collection process. There are different ways to populate the entity list including using Integration Points, Import/Export, or manually. For more information, see Integration Points and Import/Export. To manually create a new custodian, follow the steps below:

1. Navigate to the **Entities** tab.
2. Click **New Entity** on the Custodians tab.
3. Select the **Collection History** layout from the drop-down menu and complete the fields. See Fields below.
4. Click **Save**.

Note: When creating a custodian from Legal Hold, the **Custodians - Legal Hold View** is selected by default. If Collection or Processing is also installed in the same workspace, you can view the **Custodians - Processing View** or **Custodians - Collection View**.



4.2 Fields

When creating an entity, the Collection Custodian layout provides the following fields and others. To successfully create a collect entity, at a minimum, complete the following fields:

- **Custodian Type** (Required)—select one of the following:
  - **Person**—select this option to enter first and last name of the individual acting as custodian of the data you wish to process.
  - **Other**—select this option if the custodian of the data to process is not an individual but is, for example, just a company name.
    - Selecting this option changes the Custodian layout to remove the required First Name and Last Name fields and instead presents a required Full Name field.
    - You can also select this option to enter an individual's full name without having that name include a comma once you export the data associated with it.
- **First Name** (Required)—the first name of the custodian. This field is only available if you've set the Custodian Type to Person.
- **Last Name** (Required)—the last name of the custodian.
  - This field is only available if you've set the Custodian Type to Person.
- **Full Name**—the full name of the custodian of the data you wish to process.
  - Required when you set the Custodian Type to Other, Organization, or a custom choice type.
  - When you enter the full name, it does not contain a comma when you export the data associated with it.
- **Email** (Required)—the email the custodian uses in the target.
  - The entity's email is required to connect the entity to a custodian target. The email is not required to save the entity.
  - This email address must match the email address within the connected data source.

4.3 Creating a custodian target

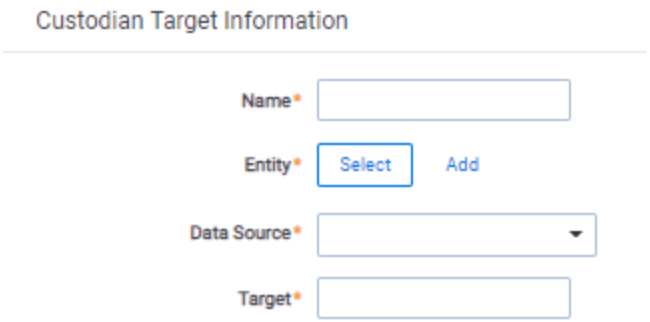
You can collect electronic data from custodians who are individuals, or entities, involved in a legal action or case. You may perform multiple collections from a single custodian.

On the Custodians tab, you can create and edit custodians as well as view their details, associate them with collections, and perform other tasks. Custodian Targets can also be automatically generated in the Collection Summary step of the Collect wizard. For more information, see Collection Summary on page 18.

Use the following procedure to create a custodian target that you can associate with a collection:

1. Enter the **Name** of the custodian target.
2. Click **Select** to select available custodians. If the custodian does not exist, click **Add** and complete additional steps.
3. Click the **Data Source** drop-down menu to select a data source. For more information Collection data sources, see Data source types on page 11.

4. Enter the custodian's email address in the **Target** field.



To avoid duplicate custodians, a custodian with multiple emails, you will need to link a second custodian target to the same entity. To make a second custodian target for the same entity, you will need to create another target.

In the target, select the same entity. Then, select the other data source and enter the target value. This way you can have multiple targets, different from the primary email address, for a single entity record.

4.3.1 Generating targets in the wizard

Custodian Targets can also be automatically generated in the Collection Summary step of the Collect wizard. For more information, see Collection Summary on page 18.

Click **Generate Targets** to check if targets exist for the custodians you've selected for collection. If the targets do not exist, Collect creates them based on the email address in the Entity details in each custodian.

If an invalid custodian target is created manually, the auto-generation of custodian targets does not remove the invalid custodian target. Remove it manually. Any errored targets display in the Collection Summary step or on the Status Summary tab.

To delete a custodian target from the Custodians Targets tab, click the data source's checkbox and use the Delete mass operation.

4.4 Removing custodian targets

To remove any custodian target, first remove the custodian from the collection or collections. You can use the Mass delete operation to delete collect custodians, also known as entities. For more information, see the Admin Guide.

Trying to delete a custodian target before removing a custodian results in an error.

5 Data sources

A data source allows you to define where and how you pull data from a communication channel. A data source stores the configuration necessary to retrieve data from a communication channel, process that data, and ingest it into Collect.

Set up workspace data sources before beginning collections. Data sources are stores of information from which you collect data. These data sources have parameters that you can set during the creation of a collection job.

Note: Data sources cannot be deleted once it's been run on a collection.

5.1 Creating a collect data source

The Collection Admin tab is where you create, edit, and remove data sources from your workspace. Setup only needs to be done once for each data source. You must create your data sources prior to setting up your custodian targets. For more information, see Custodian targets on page 8.

When creating data sources, you can select different types of data sources for obtaining files.

Use the following procedure to create a new Collect source instance:

1. Navigate to **Collection Admin**.
2. Click the **New Collection Source Instance** button.
3. Do the following:
    - **Name**—enter in a unique name for the data source.
    - **Type**—select the type of data source. For more information, see Data source types below.

      Note: Collect automatically collects any data that is preserved due to an in-place hold or litigation hold. Data on a hold is stored in a preservation library and separate folders. For more information, see Microsoft Retention Policies.

    - **Settings**—enter the data source-specific settings. For more information on specific fields and settings, locate your data source topic from Data source types below.
4. Click **Save**.

After clicking Save, Relativity verifies the parameters and connectivity to the data source instance. If successful, the data source is saved. If the connection fails, a message displays indicating that the connection failed. If verification fails, verify that the values are correct. The data source will save when it is corrected and is verified.

Once the data source is set up, you'll see the data source information on the Collection Admin page.

5.1.1 Third-party vendor

If you are a vendor setting up a data source on behalf of a client, the client does not need access to the Relativity workspace. You manage the workspace setup on their behalf. The client only needs to provide admin credentials for their platform to complete the data source setup. The client has two options:

- The platform admin shares their credentials with you directly.
- The client creates a dedicated admin account for you to use.

We recommend completing this step in a screenshare session with the client's admin to avoid sharing credentials outside of the client's organization.

5.2 Data source types

You can select any of the following data source types for step 3 of Creating a collect data source on the previous page:

- **ChatGPT**—select to collect data from ChatGPT. For more information on specific fields and settings, see ChatGPT Enterprise data source below.

6 ChatGPT Enterprise data source

This topic provides details on how to capture OpenAI's ChatGPT Enterprise with Collect.

Note: This documentation contains references to third-party software, or technologies. While efforts are made to keep third-party references updated, the images, documentation, or guidance in this topic may not accurately represent the current behavior or user interfaces of the third-party software. For more considerations regarding third-party software, such as copyright and ownership, see Terms of Use.

6.1 Considerations

Note the following considerations about this data source:

- This connector requires a ChatGPT Enterprise subscription.
- OpenAI's Compliance API only keeps data for 30 days. Collect can only retrieve ChatGPT data from the past 30 days.
    - OpenAI does not keep deleted data for more than 30 days, unless legally required otherwise.
    - Temporary chats, also called ephemeral chats, are only available for 30 days. These are conversations where OpenAI's Memory feature is disabled. The chat session is not saved to the user's history.
        - Temporary chats are accessible through the OpenAI Compliance API. They are included in Collect if your date filters are set correctly.
        - Temporary chats must fall within the 30-day retention window to be collected.
    - To avoid missing data, we recommend scheduling regular, ongoing collections before data is removed at the end of the 30-day retention window.
- Before your first collection, confirm that your organization enabled the Compliance API in its OpenAI account. This setting supports retrieval of both persistent and temporary conversations.
- For date criteria, ChatGPT Enterprise supports a start date. Collect retrieves all prompts and responses from the date you specify through the current date.
- If your organization has multiple ChatGPT Enterprise workspaces, you must configure a separate data source for each workspace in the Collect application. Name each data source to identify the workspace connects to.
- Relativity automatically converts collected conversation prompts and responses to Relativity's short message format (RSMF).
- ChatGPT metadata is collected. For more information, seeCollected metadata on the next page.

6.2 Prerequisites

Complete the following before setting up a ChatGPT data source in Relativity:

- Eligible Enterprise and Education workspace owners can create a workspace-scoped Admin key with the permissions needed by the integration.
    - Create Admin key on the *Admin Console* > *Credentials* > *Admin keys* page.
    - For more information, see OpenAI's documentation.
    - Existing authorized keys do not need replacing just because self-service is available.
    - Contact OpenAI Support only self-service is unavailable.
- Confirm that the Compliance API is enabled in your organization's OpenAI account to support retrieval of both persistent and temporary conversations.

Once you complete the prerequisites, you can begin creating the data source.

6.3 Collection size

ChatGPT Enterprise collections scale with conversation volume and duration.

Category	Details
Hard limits (enforced)	No explicit hard cap published
Planning thresholds (guidance)	High-usage enterprise tenants
When to split	Long date ranges
Notes	Conversion-based scaling

Focus on limiting time ranges in high-usage environments to control collection size.

6.4 Collected metadata

As of the current integration with OpenAI's Compliance API, the metadata fields are captured and included in collections:

The following metadata fields are captured from ChatGPT Enterprise conversations via the Compliance API and included in Collect:

Metadata	Collected
Conersation ID	Always
Conversation title	Always
User ID	Always
User name	When user is logged in.
User email	When user is logged in.
Timestamp of prompt	Always
Timestamp of response	Always
Prompt text	Always
Response text	Always
Edit history	When a prompt was edited before re-execution.
Source attributions	URLs of any publicly cited websites when browsing was enabled.
Attachments	Uploaded files submitted by users during the conversation. For example, PDFs, DOCX, images, CSV files. Generated files. For example, documents or images created by ChatGPT in response to prompts.

These fields are extracted for both persistent and ephemeral (temporary) conversations, as long as they fall within OpenAI's 30-day retention window and meet your search criteria in Collect.

6.5 Creating the data source

Use the following procedure to connect the ChatGPT Enterprise workspace data source to Collect.

In Relativity aiR:

1. Navigate to **Collection Admin** within Collect Admin of Set Up
2. Click the **New Collection Source Instance** button.
    - **Name**—Enter in a unique name for the data source.
    - **Type**—Select the **ChatGPT Enterprise** data source.
    - **Settings**—enter the required information in the Settings fields. For more information, see Settings fields below.
3. Click **Save**. The data source displays on the Collection Admin page.

6.6 Settings fields

To connect Relativity to a ChatGPT Enterprise workspace, you need to gather and enter the information for the following fields.

- **API Key**—enter the API Key that you created in OpenAI.
    - You can use the same API Key when configuring multiple ChatGPT Enterprise data sources from which you wish to collect.
    - For more information, see OpenAI's documentation.

    Note: Only contact OpenAI Support if you cannot create your own API Key.
- **Workspace ID**—enter the ChatGPT Enterprise WorkspaceID to which you want to connect.

6.7 Configuring the data source

Each data source used in Collect has different search criteria on the Collection Details step when creating a collection job. The criteria needs to be configured next.

Relativity collects ChatGPT Enterprise chat data in Relativity's short message format (RSMF).

Add search criteria to collect specific data. To configure the data sources, complete the following fields:

- **Select and unselected tabs**—choose the data sources to collect from by moving unselected data sources to the selected list.
- **Field**—choose the field to filter on within the data source.

    Note: You must set a Start Date to complete the Data Sources step.
- **Operator**—choose an operator. For example, equals or greater than.
- **Value**—enter a value to find in the selected field.

The following table lists the filter criteria supported for ChatGPT Enterprise collections.

All dates are in Coordinated Universal Time (UTC).

Criteria	Operators	Description	Example
Start Date	Greater Than or Equals	When you use the Start Date property in a query, the search returns prompts and responses that exist the day of and after the entered date.	When you search a Start Date of 1/1/2024, Relativity collects all prompts and responses from that date to today.

Note: Temporary, or ephemeral, chats will also be returned if they occurred within the 30-day retention window and meet your search criteria.

7 Collection

Before you begin collecting, you must create a collection job and associate it with a specific matter, custodians, and one or more data sources. Add the custodians, data sources, and other information using the Collect wizard. Once completed, start the collection using the Collect console. Finally, download a results report that details the items collected and a summary report of the entire collection job. For more information, see Reports on page 27.

7.1 Considerations

Before creating a collection for any data source, consider the factors that can affect collection performance.

- Collection time can increase when you use:
  - Large date ranges.
  - A high number of custodians.
  - Multiple collections running at the same time.
- Combining large scopes with high concurrency can slow down collections. This behavior is expected across all data sources.
- When you migrate a Collect workspace, data source credentials do not move with the workspace. You must re-enter credentials before running new Collect jobs. For more information, see Migrate.

7.2 Best practices for time-sensitive collections

To improve performance and reduce collection time:

- Use smaller date ranges whenever possible.
- Limit the number of custodians included in each collection.
- Reduce concurrency, or run collections sequentially when appropriate.
- Create multiple collection jobs when necessary.

Following these practices can help collections complete more quickly and predictably.

7.3 Creating a collection

Before you begin creating a collection, make sure to create a matter. For more information, see Matters on page 6.

Use the following procedure to create a collection:

1. On the Collect tab, click the **Collections** sub-tab. Collect displays a list of the collections currently added to this application.
2. Click **New Collection**.
3. Complete the steps in the Collect wizard. See Using the Collect wizard on the next page
4. On the Collection Details page, click **Run Collection** in the console. See Collect console on page 21.

7.4 Using the Collect wizard

The Collect wizard takes you through each step to create a collection. After completing the collection setup, run a collection from the Collection console.

Collection wizard security permissions

- **Custodian**—View

When a step is complete, click **Next** or the hyperlink under the next step shown. Click the **Previous** button to move to the previous step. Relativity auto-saves your information when moving between steps. If any required information is wrong or missing, an error message displays and you cannot move to the next step.

7.4.1 Collection Details

Complete the Collection Details step by entering information in the following fields:

- **Name**—the name of the collection. Enter a name using alphanumeric characters only. You cannot use special characters, such as periods, commas, and em dashes. Special characters will cause an error.
- **Collection Matter**—the name of the matter associated with this collection. Click **Edit** to select an existing matter or click **Add** to define a new one. See Creating a matter on page 6.
- **Job Number**—lists a number assigned to the job for reporting purposes.
- **Processing Source Location**—the file repository for collected data to be stored for future processing of documents or for storing collected data. All data sources that produce RSMF have a limit of 2 GB. For more information, see Processing documentation.
- **ZIP Collected Files**—toggle on to compress all collected data into ZIP64 formatted containers. Relativity can compress Microsoft 365, Slack, and X1 data into zip folders.
  - Selecting *Yes* on the Collect Files in ZIP field adds your collected data into containers and puts it into a password protected compressed folder. The compressed folders separated by the custodian target collected. Each custodian target collect has its own folder. These folders will split when reaching a set size. The compressed folders are then stored in the processing staging area by default.
  - You can secure these compressed folders with a password. You can enter a password that you, or another user, needs to enter to open the compressed folder. Relativity stores these passwords in the password bank. You can retrieve them there at a later time.
  - To export your collected data, use the Staging Explorer.
- **Zip Password**—enter a password that is required by anyone attempting to decompress the ZIP64 container files.
  - If you have Processing installed in the workspace, Collect automatically populates the Processing Password Bank with the password so it is available at the time the collection is processed.
  - Click the Show Password box to display the actual password in the field to ensure the characters are correct.
  - Zip is supported by Microsoft 365, Slack, and X1. Google is exported as pre-zipped.
- **Enable Auto-Processing**—toggle on to enable auto-processing. If enabled, select the workspace, profile, and document prefix. Relativity processes all data in a completed collection, or completed with errors collection, after a collection finishes.
  - **Workspace**—select a workspace within your instance to use when creating a processing job.
  - **Processing Profile**—select a processing profile available in the drop-down menu. The available profiles are from the selected workspace.

- **Document prefix** —select the document prefix option of *Use Entity Document Number Prefix* or *Use Processing Profile Document Number Prefix* to apply to each file in the processing set once it's published to a workspace.

- **Description**—enter a description of the collection used for reporting purposes.
- **Receive Progress Notifications**—toggle on to send or receive collection job status emails. The statuses include:
  - **Completed**—includes completed or completed with errors job status.
  - **Failed**—includes job status and reason for failure.
- **Notification Address**—enter the email address of person that wants to receive collection job statuses.
- **Data Source Type**—select one or more data sources to use in the collection. For more information, see Data source types on page 11.

7.4.2 Data source

Configure the data sources chosen in the Collection Details step. Each data source has different criteria to enter. See the list of data source types under Data sources on page 10 for information on each one.

You can select multiple data sources in the first step if you want to configure all or multiple sources in the step. Switch between each source to configure its criteria by using any of these methods:

- Click the name of the data source in the navigation menu.
- Click **Next** and **Previous** to move you through the data sources.
- Select individual data sources by clicking on the checkbox and then using the right arrows to select them.

7.4.2.1 Data source criteria

Add criteria to collect specific data. To configure the data sources, complete the following fields:

- **Select and unselected tabs**—choose the data sources to collect from by moving unselected data sources to the selected list.
- **Field**—choose the field to filter on within the data source.
- **Operator**—choose an operator such as equals, contains, greater than, or less than.
- **Value**—enter a value to find in the selected field.

After selecting field options, you must click **Add Criteria**. You can add multiple criteria to search data sources.

Note the following about criteria:

- Each criteria is then separated by an AND operator.
- Leave the data source criteria empty to collect all data from the sources.

7.4.3 Custodians

Complete the Custodians step by assigning custodians to the project. Follow the steps below to assign a custodian.

1. From the Unselected custodians table, use the column filters to locate custodians.
2. Select a checkbox next to a custodian.

   Notes: If you select multiple Slack custodians in the same collection, the channels they share will only collect once.

   - Collect supports up to 30 custodians, or entities, assigned in a collection job.
   - If you need more than 30 custodians, you need to create another collection job.
   - This limit only applies to the number of custodians.
   - There is no limit to the number of targets to be collected.
   - For example, you can select 30 custodians and three data sources for a total of 90 targets for the collection job.
3. Click the right arrow icon to add select custodians. Click double right arrow icon to add all custodians.
4. Click **Next**.

Note: There is a limit of 10,000 listed custodians with targets in the custodian picker.

7.4.4 Non-custodial

Select non-custodial data sources to complete this step. Non-custodial data means you select the specific locations or archives you want to collect from. Selecting custodians is not required.

To select non-custodial data:

1. Click one or more data sources in the **Select Sources** column.
2. With one of the data sources highlighted, select the check boxes next to the sites you want to collect.
3. (Optional) Toggle on the **Show selected only** option to display only the locations selected for collection. The selected locations display in a concise list for review before you proceed to the Summary step.
4. Click **Next**.

7.4.5 Collection Summary

Complete the creation of the collection by reviewing all steps, custodians, data sources, and targets, before finalizing.

If custodian targets were not created before you started the project, click **Generate Targets**.

Click **Generate Targets** to check if targets exist for the custodians you have selected for collection. If the targets do not exist, Collect automatically creates them based on the email address contained in the Entity record for each custodian.

7.4.5.1 Targets

In the Targets section, a number appears next to each custodian's name. The number listed is the number of custodian targets found in the associated data source.

Number	Meaning
0	Collect did not find any custodian targets with that email address in that data source.
1	Collect found a single custodian target associated with the email address within the data source.
2+	Collect found multiple custodian targets with that email address within the data source.

The number is also color-coded to indicate how the target was generated:

- **No highlight**—Collect already found and generated the custodian target.
- **Green**—Collect autogenerated the custodian target.
- **Red**—Collect could not autogenerate the custodian target.
  - You can still manually generate it.
  - For more information, see Creating a custodian target on page 9.

Click **View Collection Details** to finish setup. The Collection Details page appears. From the Collection Details page, you can preview and run the collection from the Collection Summary on the previous page.

8 Viewing or editing Collection data

You can view and edit collection details. You can also use the Collect console to start and stop collections and view reports.

8.1 Collection details

You can display the collection details by clicking the name of a collection on the Collections tab. Collection also displays these details immediately after you add a new collection.

On the Collection Details page, use the buttons at the top of the page to edit, delete, go back, edit permissions, or perform other collection tasks. Editing a collection takes you to the first step in the wizard. For more information, see Using the Collect wizard on page 16.

Note: Once a collection has started, the collection details are read-only and cannot be changed.

- **Collection Details**—displays the information that you entered or selected when you created the collection:
  - **Name**—lists the name given to the collection.
  - **Collection Matter**—the matter used in the collection.
  - **Job Number**—the number assigned to the job for reporting purposes.
  - **Description**—the description of the collection used for reporting purposes.
  - **Processing Source Location**—the file repository that collected data is stored for future processing of documents or for storing collected data.
  - **Auto Processing Workspace**—the processing workspace if you toggled the field on.
  - **Auto Processing Profile**—the processing profile if you toggled the field on.
  - **Auto Processing Document Numbering Prefix Option**—the prefix option if you selected one.
  - **Job Status**—status of the collection. Statuses are New, Not Started, Started, Completed, Error, and Completed with Errors.
  - **Error Message**—the message if collection job did not complete due to errors.
  - **Zip Collected Files**—if you selected to zip collected files this box is checked.
  - **Zip Password**—the password that everyone needs to use to decompress the ZIP64 container files.
  - **Receive Email Notifications**—if you toggled on to send or receive collection job status emails this checkbox is checked.
  - **Notification Address**—the email address of the person that will receive collection job statuses.

- **Collection Console**—displays buttons that you can use to perform the tasks listed. See Collect console on the next page.
- **Data Sources**—lists all the collection activities associated with this collection.
  - **Custodian**—the custodian associated with the collection.
  - **Source Instance**—the name of the data source associated with the collection.
  - **Data Type Name**—the name of the target associated with the collection.
  - **Status**—displays one of the following statuses:
    - **Not Started**—the collection has not been started.
    - **Started**—the collection is in progress.
    - **Completed**—the collection is done.
    - **Completed with Errors**—the collection of the target completed and one or more targets had errors. For more information, see the Errors report on page 28.
    - **Failed**—the collection failed. For more information, see Reports on page 27.
  - **Error**—lists the error message if the status is Failed.
  - **Collected Items**—the number of files collected from the target without error. If nothing is collected, a 0 is listed.
  - **Collected Item Total**—the number of files that there are to collect without errors. If nothing is collected, a 0 is listed.
  - **Target**—the custodian target associated with the data source.
  - **Result Link**—a Comma Separated Values file download listing all individual items collected from the target. It contains all the associated metadata for each collected item as well. If no results, the file is empty. For more information, see Viewing or editing Collection data on the previous page.
  - **Error Link**—a Comma Separated Values file download listing any individual items that couldn't be collected because of errors during the collection. The report provides as much metadata as it can along with as much error information as we can get from the source to help identify what caused the error. If no errors, the file is empty. For more information on errors, see the Errors report on page 28.
- **Previews**—the status and estimated number of items and size of collection. This is available after starting a preview. For more information, see Viewing or editing Collection data on the previous page.
- **Custodian Details**—the status of the custodian and the source instance. This card also includes the filter criteria, items, size, and errors.

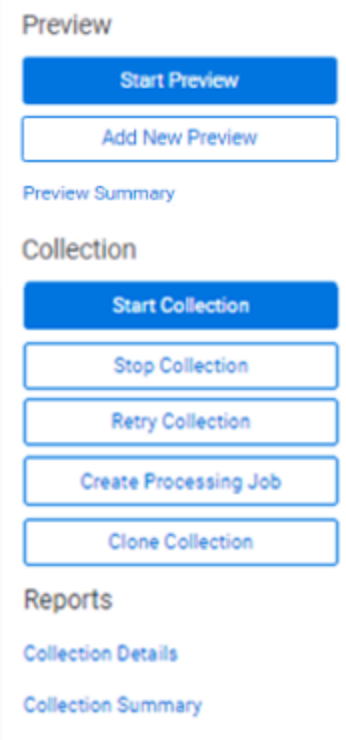
8.2 Viewing collected data

When Relativity collects the data, Relativity accepts the path names and file names that the source provides. On occasion, the collection source modifies the path name or file name.

There is a difference between viewing grouped and individual collections data. To view grouped collections data, for sources like Bloomberg Chat, Google Chat, Slack, and Teams, click the **Get Results** link to view your collection data. Although the Data Sources table will show N/A in the Collected Items and Collected Item Total columns, the collected data is still available in the download. These columns show N/A because more than one set of custodian data is included in the collection.

8.3 Collect console

After completing a setup, run the collection with the collection console. Verify connection, start or stop a collection, and view a collection report from the Collection console on the Collection Details page.



8.3.1 Collection

Once you decide to move forward with your collection, you will use this section of the console to start, stop, and adjust your collection job.

8.3.1.1 Start Collection

Click the **Start Collection** button the begin the collect project.

8.3.1.2 Stop Collection

Click **Stop Collection** to end the collection project that is currently running. Once you click this button, a warning pop-up message appears to confirm that you want to stop. Once you stop a collection, the collection cannot be restarted.

8.3.1.3 Retry Collection

The Retry Collection button is only available when a collect job does not complete because of errors. Click **Retry Collection** to start another collect job that only retries the targets that have failed. You cannot retry targets that completed successfully.

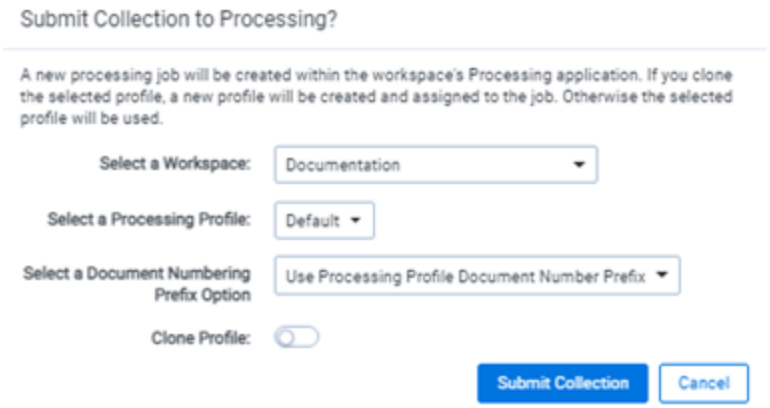
8.3.1.4 Create Processing Job

Click **Create Processing Job** to create a processing set from the collected documents from the data sources. Locate these documents in the Processing Source Location set in Collection Details.

When you click the Create Processing Job button, a pop-up window displays with two fields:

- **Select a workspace**—select a workspace within your instance to select a Processing profile to use when creating a processing job.
- **Select a Processing Profile**—select a processing profile available in the drop-down menu. The available profiles are from the selected workspace.

- **Select a Document Numbering Prefix Option**—select *Use Processing Profile Document Number Prefix* or *Use Processing Profile Document Number Prefix* to apply to each file in the processing set once it is published to a workspace.
- **Clone Profile**—toggle off to use the selected processing profile. Toggle on to clone the selected processing profile and use the clone profile in conjunction with the created processing set.



After selecting a processing profile and deciding on cloning the profile, click the **Submit Collection** button. Once you submit the collection, Relativity creates a processing set with the same name as the collection job. The processing set includes all data that was collected in the collection job.

8.3.1.5 Clone Collection

Click **Clone Collection** to duplicate the open collect job. You can clone any collect job that has a status other than New.

After clicking the Clone Collection button, a pop-up modal displays the progress and completion of the clone.

Once the job is cloned, it is placed in the Not Started status and you can find it in the Collections list under the same name with "Cloned - YYYY-MM-DD HH.MM.SS" amended to the end. All collection totals for the cloned job associated targets are reset to zero. The Collection Detail Custodian Target fields are reset to zero. The cloned collect job also generates the new targets.

8.3.1.6 Reports

Collect includes comprehensive reporting capabilities that you can use to view information about your collections. You can generate these reports in the collection console within a collection project. Click on the name of a report to download. When generating a report, Relativity downloads different files through your browser. For more information on specific reports, see Reports on page 27.

8.4 Delete collections

A collection will not delete if they have an existing custodian tied to it.

When you delete a collect job, Relativity deletes all of the related custodian targets for the job. The custodian targets table remain the same. The targets within Custodian Targets are not related to any specific collect jobs. Relativity creates job-specific copies of the custodian targets for the job when it's created.

9 Grouped collections

A grouped collection means that all conversations for the selected custodians are combined into a single set of RSMF output files.

This approach is used to:

- Minimize collection size
- Prevent duplicate conversations

If each custodian were collected separately, any shared conversations, such as direct messages, group chats, or shared channel content, would be collected multiple times. This would:

- Increase the overall data volume
- Expand storage requirements
- Require additional downstream deduplication effort during processing and review

9.1 Grouped data sources

Within Collect, there are short message data sources that Relativity collects in a grouped format when more than one custodian is selected.

Grouped collection data sources:

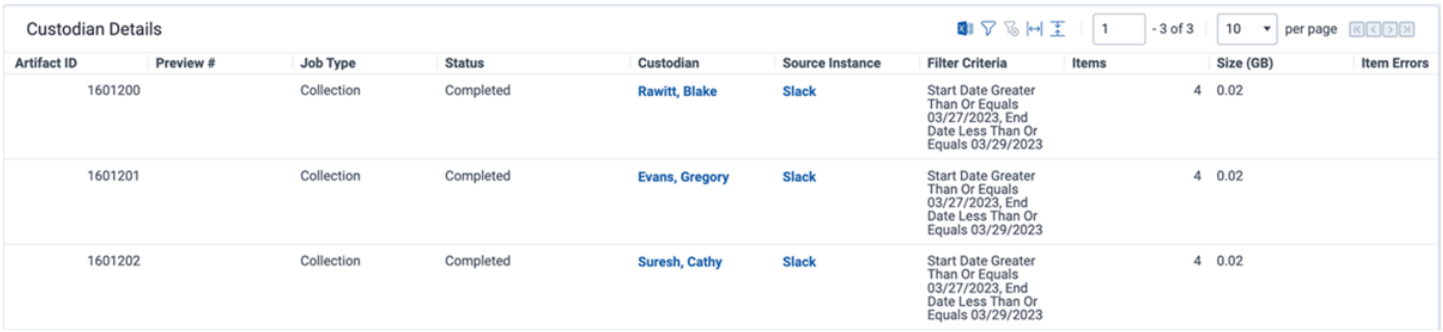
- Microsoft 365 Teams
- Slack
- Google Chat

9.2 Grouping messages

Grouping ensures that shared content is collected once. This helps reduce duplicates and keeps the dataset consistent.

When viewing grouped collection results on the Collect job details page, the reported artifacts collected and collection size show the same value for all involved custodians. This reflects that there is one collection of RSMFs containing all chats for the collected custodians.

The reported *Items* collected is the number of RSMF files that were generated during the collection. We currently do not support reporting the number of individual messages attributable to each custodian independently. The reported *Size (GB)* represents the aggregate size of all generated RSMF files.



Custodian Details

1 - 3 of 3 | 10 per page

Artifact ID	Preview #	Job Type	Status	Custodian	Source Instance	Filter Criteria	Items	Size (GB)	Item Errors
1601200		Collection	Completed	Rawitt, Blake	Slack	Start Date Greater Than Or Equals 03/27/2023, End Date Less Than Or Equals 03/29/2023	4	0.02	
1601201		Collection	Completed	Evans, Gregory	Slack	Start Date Greater Than Or Equals 03/27/2023, End Date Less Than Or Equals 03/29/2023	4	0.02	
1601202		Collection	Completed	Suresh, Cathy	Slack	Start Date Greater Than Or Equals 03/27/2023, End Date Less Than Or Equals 03/29/2023	4	0.02	

9.3 Custodian assignment considerations

Grouped collections create a limitation during processing because the Processing capability supports only one custodian per processing set data source.

Relativity assigns an Entity representing the data source as the custodian for the processing set when using *auto-processing* or the *Create Processing Job (on-demand)* feature in Collect.

This prevents the system from applying a single custodian to the entire RSMF dataset, which would incorrectly suggest all messages belong to one person.

9.4 Manual processing guidance

When manually creating a processing set, you encounter the same limitation of needing to assign a single custodian to the grouped RSMF data.

Recommended approach:

- Create or use an Entity of type *Other*
- Name the Entity to clearly represent the data set

Example: *Slack – Acme Matter – 2026-02-10*

This approach accurately reflects that the data set represents multiple custodians and maintains consistency with Collect's automated processing behavior.

10 Migrate Collect workspaces

When you migrate a Collect workspace, using *Archive, Restore, and Migrate (ARM)* or *Integration Points*, Relativity moves workspace data but does not move data source credentials.

After the migration, you must re-enter credentials before you can run Collect jobs.

Note: This behavior applies to all Collect data sources.

10.1 Migrate Collect workspaces

You can migrate Collect workspaces using ARM or Integration Points.

During a migration:

- Relativity restores the Collect workspace and its objects.
- Relativity does not restore data source credentials.
- Data sources appear in the restored workspace but are not connected.

Relativity does not provide a supported way to move or copy Collect credentials. This protects sensitive authentication information.

10.2 Credentials do not migrate

To protect from compromising sensitive authentication information, Relativity does not support moving or copying Collect credentials from migration.

10.3 After migration

Until you re-enter credentials, Collect cannot access the data sources.

After a migration:

- Data sources appear in the restored workspace.
    - The credentials will not work.
    - The connections are invalidated.

- You must re-enter the credentials for each data source and revalidate each data source.
- New collection jobs cannot be run until the data sources are connected again.

Note: You do not need to remake the credentials for the data sources. If the existing credentials are still valid you need to re-input.

10.3.1 Considerations

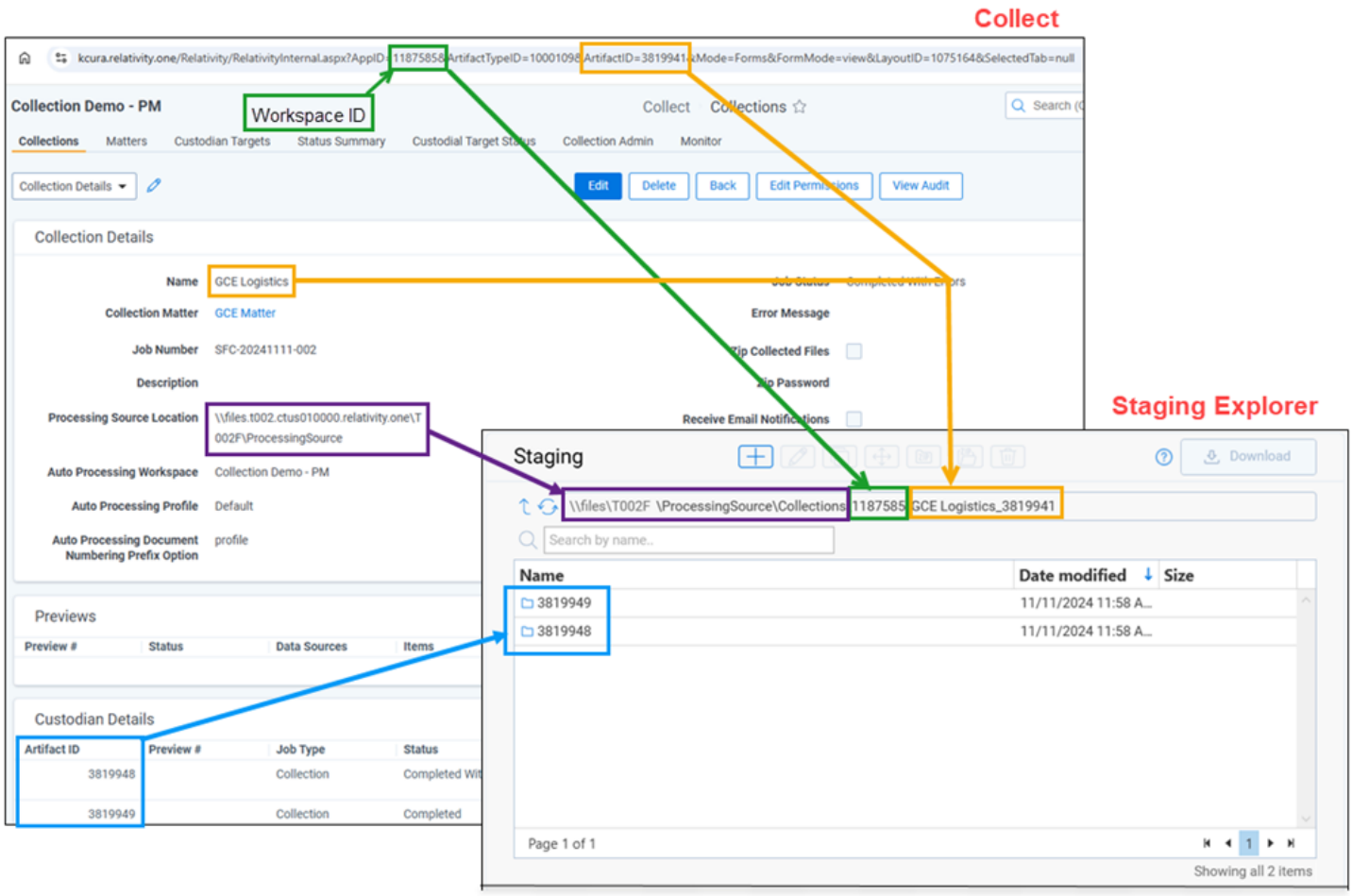
Consider the following when using Microsoft or Google:

- **Microsoft 365**—the app created for Microsoft 365 data sources has a secret value which is only exposed on initial creation.
  - This secret value must be stored and re-entered correctly for the data source to work.
  - If the Secret Value is forgotten, please generate a new secret value.
- **Google**—for Google data sources, a refresh token is generated on the first Google data source that was set-up and used for subsequent Google data sources.
  - This refresh token value must be stored and re-entered in the data sources in order for the data sources to work.
  - If the token is invalidated or forgotten, please see Troubleshooting for instruction on generating a new refresh token.

11 Identifying Collection data in Staging Explorer

This section helps to identify the components of the folder information within the Staging pane of the Staging Explorer as it relates to the Collections setup fields.

Use the sample screen below as a general guide for each data source type.



Here is a breakdown of the folder components within the Staging pane of the Staging Explorer:

\\files\<T####>\ProcessingSource\Collections\<WorkspaceID>\<CollectionName>_<CollectionArtifactID>\<DataSourceArtifactID>

- The Processing Source Location from the Collections page becomes the file location for the data in Staging Explorer. It contains a sub-folder for Collections to house the Collection data. Using the above example, it would be:

  \\files\T002F\ProcessingSource\Collections

- The Workspace ID folder in the Staging pane corresponds to the Workspace ID number listed after the "AppID=" in the URL. For example, 1187585 would be the Workspace ID from this URL: kcur-a.relativity.one/Relativity/RelativityInternal.aspx?AppID=**1187585**...

  \\files\T002F\ProcessingSource\Collections\1187585

- The Name of the Collection from the Collections page and the Collection Artifact ID become the folder name within the Workspace ID folder. Using the above example, it would be:

  \\files\T002F\ProcessingSource\Collections\1187585\GCE Logistics_3819941

- The Artifact ID of the data source used becomes the sub-folder. Using the above example, they would be:

\\files\T002F\ProcessingSource\Collections\1187585\GCE Logistics_3819941\3819948
\\files\T002F\ProcessingSource\Collections\1187585\GCE Logistics_3819941\3819949

In the case of grouped collections, such as Teams and Slack, there will be a folder within the collection folder corresponding to the data source type. Using Teams as an example, the file structure would be:

\\files\T002F\ProcessingSource\Collections\1187585\GCE Logistics_3819941\Teams

12 Reports

Collect offers comprehensive reporting capabilities that you can use to view information about collections. You can set options to generate these reports based on matter and collection as well as other combinations.

12.1 Running reports

You can generate these reports in the collection console within a collection project. Click on the name of a report to download. Locate the Collection Summary and Collection Details report in the collection console. Locate the Results and Error report in the Data Sources table on the Collection Details page.

When generating a report, Relativity downloads different files through your browser.

Note: All report timestamps are in UTC.

12.1.1 Collection Summary report

The Collection Summary report includes the target, target status, number of items collected, and the collection size. The report is grouped by custodians. Grouping by custodians makes it easy to sort the targets for each custodian with subtotals for each custodian. Grand totals are at the end of the report. The report downloads as a PDF file.

This report also includes filters that were used at any point in the collection. For example, if a modification date that is greater than or equal to 1/1/2010 is added at the criteria level, then that filter is listed in the summary report table. If no filter criteria was specified for the job, the report lists a "No Filter Criteria Applied" message.

Group-type data sources are also included in the list of data sources. This section is at the beginning of the report. For example, if you collect from Outlook calendars, Outlook mailboxes, and Slack, all collection summaries are included in the report.

The short message grouped collections lists the custodians involved in the collect, along with the data sources. The top of the report includes the custodian list. This section's title is "Short Message Grouped Collections." The report table includes the data source, target status, the number of items collected, and the collection size in gigabytes.

12.1.2 Collection Details report

The Collection Details report includes two files: the first file includes the successfully collected results of all items and their metadata. The second file includes the errored collection data. The error file includes as many of the items and as much of the items' metadata as it can. The report downloads as a CSV file.

This report, both files, is also stored in the assigned Entra ID file share. It is included in the collection output.

12.1.3 Results report

The Results report link for each target downloads as a CSV or XML file that contains a list of all individual items collected. These items include emails, files, or other data. It contains all of the associated metadata for each item. If no items were collected the file is empty.

12.1.4 Errors report

A Comma Separated Values file download of the errors that occurred during the collection from the target. If no errors occurred, the file is empty.

If the application is reporting errors with requests, creating objects, or parsing, check for correct permissions, check for healthy connections, and check if the fileshare is working. If the setup is correct, start diagnosing errors.

12.1.4.1 Errors.csv

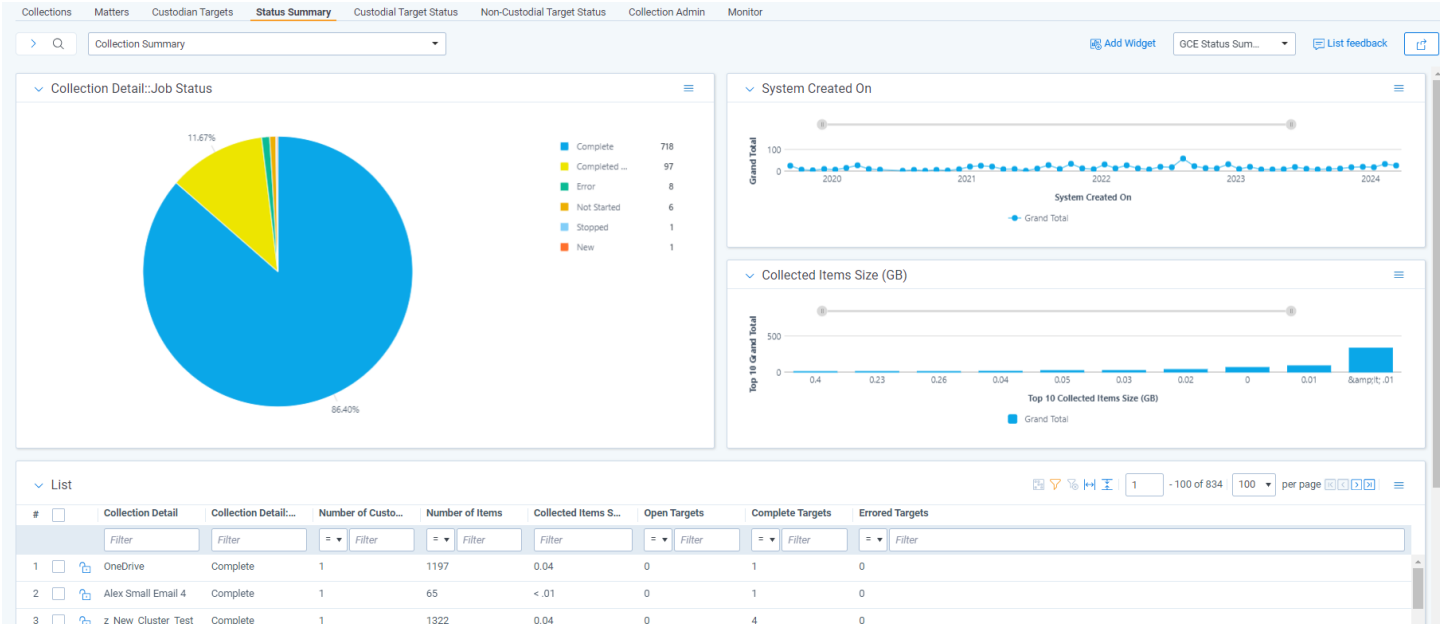
The report file lists one error per line. Each item is listed with an error ID alongside the message of the error that caused the item failure. Per-item errors only occur in the download phase of the collection; if an error occurs before (for example, if data is unavailable during our check) or after (for example, the worker cannot write the results.csv file to the fileshare) then there will be no record in the errors.csv report.

13 Status Summary

In Collect, you manage multiple collect jobs and you need to track all of them.

13.1 Job status

The job status dashboard to see the statuses of collection jobs. You can drill into each job from this dashboard. Able to look into the targets by custodian, data sources, or status to find out more about your collections. Focus on the collected components with this dashboard.



13.2 Reviewing job statuses

The job status dashboard is available after generating targets. To learn how to generate targets for a collection, see Collection Summary on page 18. Once you generate targets, the dashboard organizes by collection jobs.

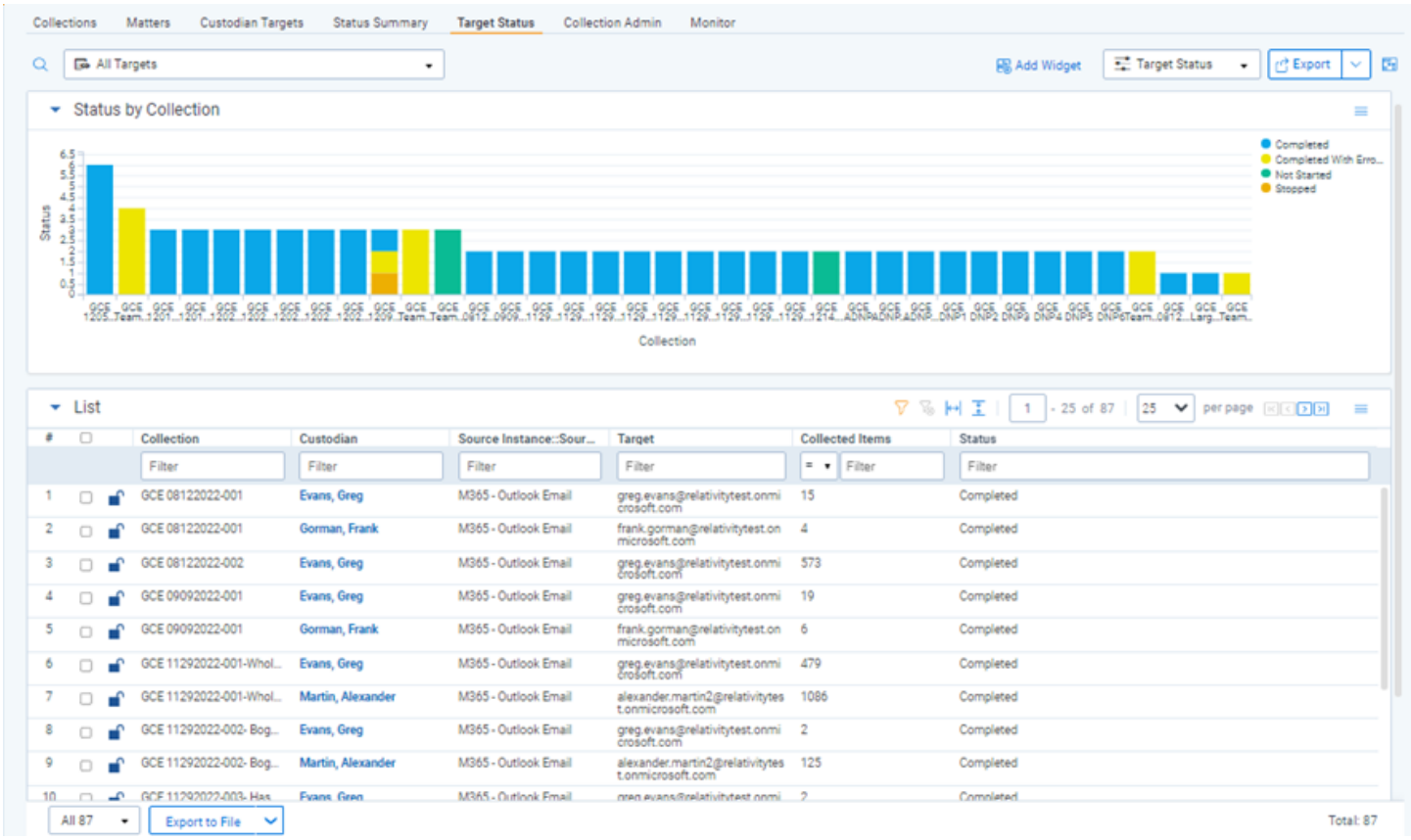
Status:

- **Not started**—the collection is set up, but hasn't been started.
- **Started**—the collection started and has not completed.
- **Completed**—the collection of the target completed without any errors.

- **Completed with Errors**—the collection of the target completed and had individual items that couldn't be collected. For more information, see the Errors report on the previous page.
- **Error**—the collection did not run successfully and couldn't collect from the target.

14 Target Status

The Target Status tab is a dashboard to see the statuses of collections. You can drill into each target from this dashboard. Able to look into the targets by custodian, data sources, or status to find out more about your collections. Focus on the collected components with this dashboard.



15 Monitor

Monitor pending, running, and completed collect jobs in the Monitor tab. The Monitor page only tracks collect jobs from the last 24-hour time range.

- **Queued**—this column lists the created collect jobs that have not started.
- **Running**—this column lists collect jobs in progress and with their current progress displayed in a status bar.
- **Completed**—this column lists collect jobs completed successfully, completed with errors, the amount of data collected, and the elapsed time.

16 Proprietary Rights

This documentation ("**Documentation**") and the software to which it relates ("**Software**") belongs to Relativity ODA LLC and/or Relativity's third party software vendors. Relativity grants written license agreements which contain restrictions. All parties accessing the Documentation or Software must: respect proprietary rights of Relativity and third parties; comply with your organization's license agreement, including but not limited to license restrictions on use, copying, modifications, reverse engineering, and derivative products; and refrain from any misuse or misappropriation of this Documentation or Software in whole or in part. The Software and Documentation is protected by the **Copyright Act of 1976**, as amended, and the Software code is protected by the **Illinois Trade Secrets Act**. Violations can involve substantial civil liabilities, exemplary damages, and criminal penalties, including fines and possible imprisonment.

©2026. Relativity ODA LLC. All rights reserved. Relativity® is a registered trademark of Relativity ODA LLC.

17 Index

C

Collect 15, 27

Collect wizard 16