Relativity®

ECA and Investigation Guide

December 8, 2021 | Version Server 2021

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

Table of Contents

1 Publishing data for ECA **3**

1.1 Selecting the ECA-optimized processing profile 3

1.2 Running the deduplication script 5

2 Promoting data with Integration Points **6**

2.1 Special considerations - job batches for large sync workflows 6

2.2 Integration point agent considerations 6

2.3 Working with promoted documents 7

2.3.1 Reusing coding decisions 8

2.4 Performance baselines 8

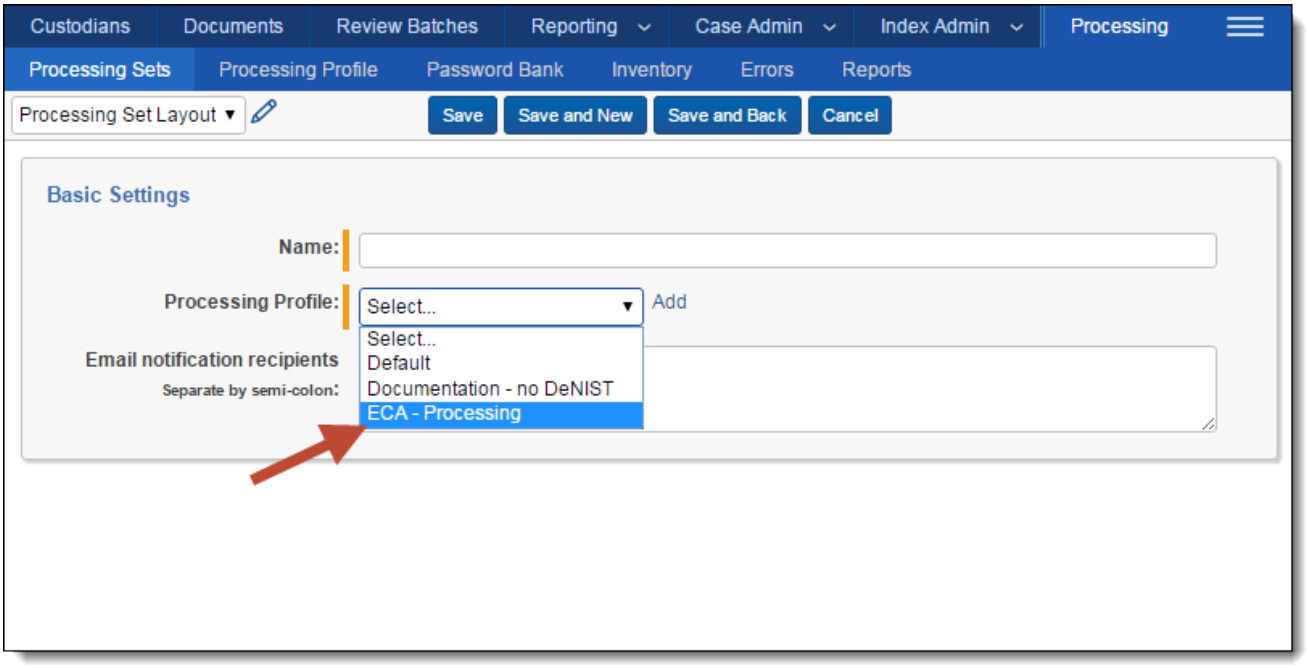
1 Publishing data for ECA

Before using the ECA Dashboard and Integration Points to promote culled data to a review workspace, you need to discover and publish the source data into your ECA workspace through Processing.

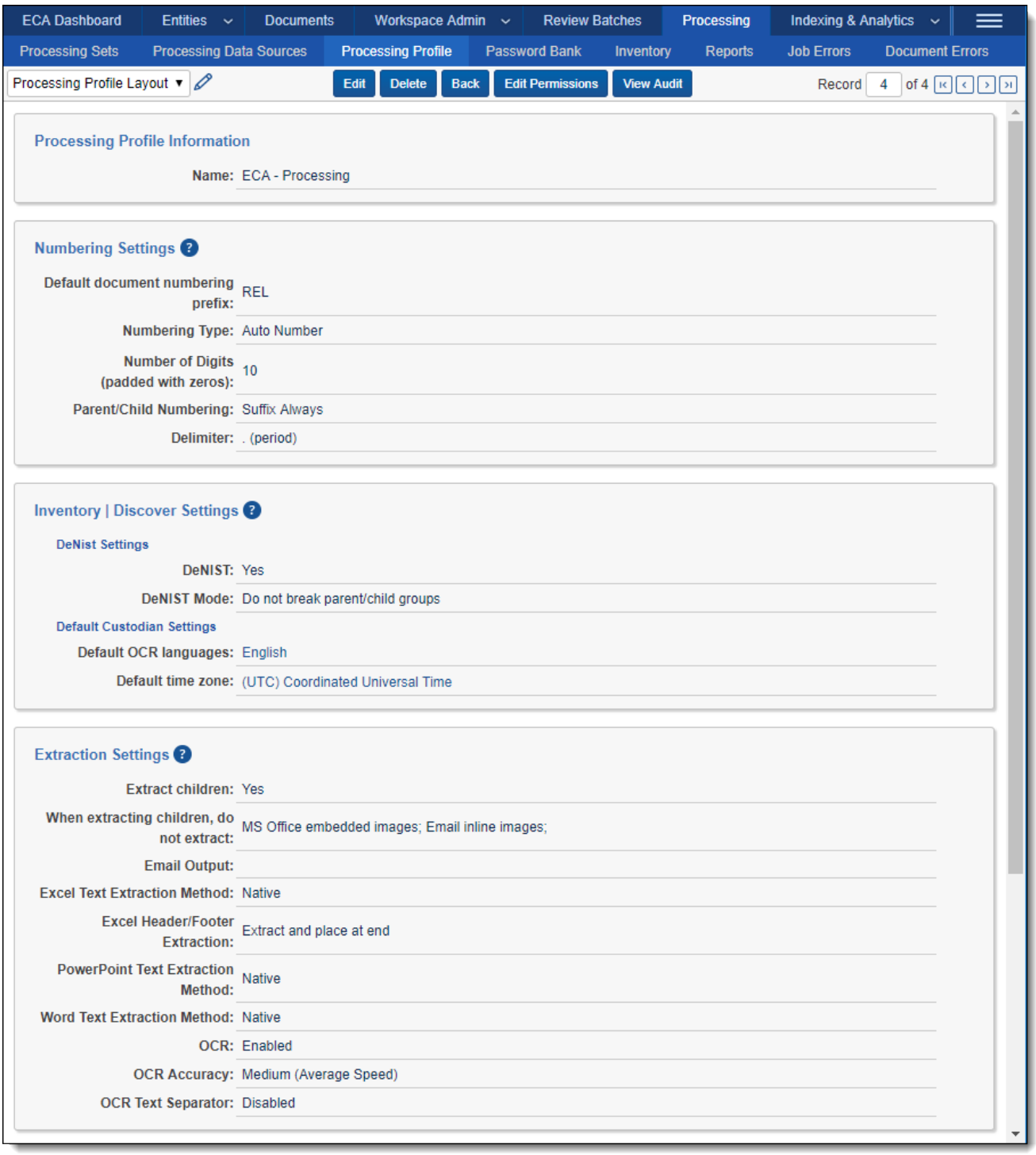
For details on discovering and publishing data outside of the ECA workflow, see the Processing User Guide.

1.1 Selecting the ECA-optimized processing profile

When processing data specifically for the ECA workflow, you need to use an ECA-optimized processing profile called **ECA - Processing**. This profile is automatically installed when you import the ECA & Investigation application. You can select this profile from the corresponding field on the Processing Set layout.



Among the settings that make the ECA - Processing profile optimized for this specific workflow is the Deduplication method field, which is set to None. This is because you perform deduplication through the Update Duplicate Status script that is installed automatically to your workspace when you install the ECA and Investigation application.



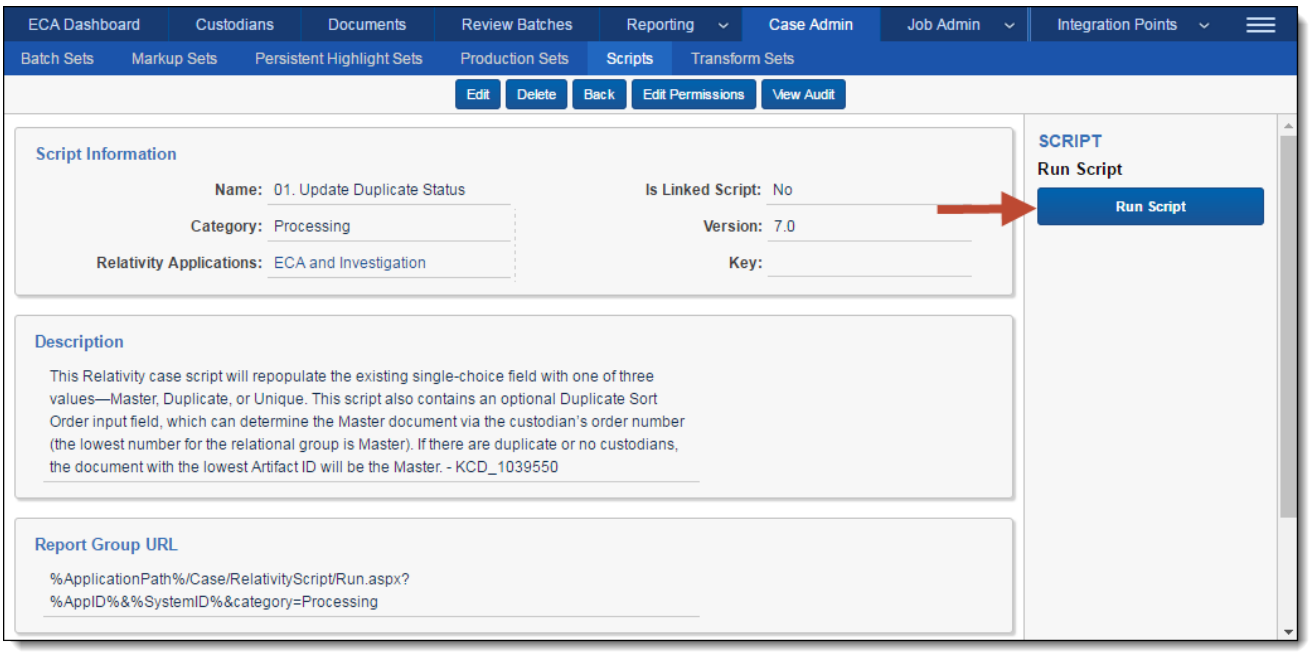
Note: If you intend to use both the RDC and Processing to bring data into the same workspace, note that if you select Custodial or Global as the deduplication method on your processing profile(s), the processing engine won't deduplicate against files brought in through the RDC. This is because the processing engine doesn't recognize RDC-imported data. In addition, you could see general performance degradation in these cases, as well as possible Bates numbering collisions.

1.2 Running the deduplication script

After you publish files to your ECA workspace, you need to run the Update Duplicate Status script to deduplicate those files.

Performing this workflow is required because you published data to your source workspace using a processing profile with a deduplication method of None.

Note: The All Custodians_Script field is a long text field and acts as a another piece of metadata for de-duplicated documents. You should select the new All Custodians_Script field when running the Update Duplicate Status script, as this will ensure that no de-duplicated documents make it into review.



For more information on the processing duplication solution, see Processing duplication workflow.

2 Promoting data with Integration Points

Once you've tagged documents for inclusion in and/or exclusion from the data you want to promote for review, you can access Integration Points to start the job or jobs that will send those documents to the review workspace.

This topic provides details on Integration Points as it is used in an ECA and Investigation workflow only. For information, see the Integration Points Guide.

Note: For information on how to develop with Integration Points, visit the Developers documentation site.

Note: For workspace-to-workspace integration points, Relativity supports transfers across the Document object only and not other RDOs.

2.1 Special considerations - job batches for large sync workflows

We recommend configuring integration point jobs into batches for large sync workflows. The following are recommendations when creating job batches:

- Document count should be no more than 500k documents in a single batch.
- When including extracted text as a mapped field in the integration point job, the sum of the extracted text in the batch should not exceed 25 GBs.
- The recommended number of fields to be mapped should be no greater than 100, but it is best to map as few long text fields as possible.
- When using the append/overlay or overlay only configuration, the batch size recommendations should be half the amount described above, specifically:
  - 250k document count
  - 12.5 GBs extracted text sum
- Choices names that contain restricted characters like hyphen (-), forward slash (/), or semicolon (;) are only supported in native push jobs. The restricted characters are not supported in integration points configured for images push, production set push, and retried jobs.

2.2 Integration point agent considerations

When you have one agent enabled, a scheduled integration points job is always queued and is always run. When you have more than one agent, however, the scheduled job may get queued or it may result in an error, after which that job is rescheduled. In other words, there is mixed behavior when you have enabled more than one agent.

The following table provides a breakdown of this behavior:

If	Then
1. You enable a single agent called Agent 1. 2. You schedule a job called IP1 to run at 10:00 AM daily. 3. You click Run or Retry Errors on IP1 at 9:59:59 AM.	1. Agent 1 picks up the Run or Retry Errors job for IP1 and completes it in one hour. 2. Agent 1 picks up the scheduled job for IP1, creates an error, removes the job history from IP1, and reschedules IP1 to run at the next interval.
1. You enable Agent 1 and Agent 2. 2. You schedule IP1 to run at 10:00 AM daily. 3. You click Run or Retry Errors on IP1 at 9:59:59 AM.	1. Agent 1 picks up the Run or Retry Errors job for IP1 and begins working on it. 2. Agent 2 picks up the scheduled job for IP1, creates an error, removes the job history from IP1, and reschedules IP1 to run at the next interval.
1. You enable Agent 1 and Agent 2. 2. You schedule IP1 to run at 10:00 AM daily. 3. You click Run or Retry Errors on IP1 at 9:59:59 AM. 4. All other agents are busy, including Agent 2.	1. Agent 1 picks up the Run or Retry Errors job for IP1 and begins working on it. 2. Agent 1 completes the Run or Retry Errors job for IP1. 3. Agent 1 picks up the scheduled job for IP1 and completes it.
1. You enable Agent 1 and Agent 2. 2. You click Run or Retry Errors on IP1. 3. You click Run or Retry Errors on IP2.	1. Agent 1 picks up the Run or Retry Errors job for IP1 and begins working on it. 2. Agent 2 picks up the Run or Retry Errors job for IP2 and begins working on it.

2.3 Working with promoted documents

To view the documents you promoted to the review workspace, perform the following steps:

1. Navigate to the destination workspace.
2. Select the Saved Search browser and select the search you created to bring back documents that were promoted from ECA.
3. Note the following ECA-relevant fields on the promoted-from ECA saved search view:
   - **Relativity Source Case** - the name of the workspace in which you tagged documents for inclusion and exclusion and from which you promoted your tagged documents to the review workspace.
   - **Relativity Source Job** - the name of the Integration Point that you used to promote tagged documents to the review workspace.

You can now review these documents and apply coding decisions for responsiveness and/or issues designation.

2.3.1 Reusing coding decisions

You can re-use the coding decisions you made on reviewed documents and promote them back into the ECA workspace through another Integration Point. For example, you could run another promote job to conduct a privilege overlay on documents in the source workspace.

To do this, perform the following steps:

1. Select the saved search you created to promote documents back to the ECA (source) workspace.
2. Navigate to the **Integration Points** tab.
3. Create a new integration point that specifies the following values, which differ from those you entered for the promote job you ran previously:
    a. **Destination Workspace** - select the original source workspace, specifically the workspace from which you previously promoted documents to the review workspace.
    b. **Saved Search** - select the saved search you created to promote documents back to the ECA (source) workspace.
    c. **Field Mappings** - map only **Control Number (Object Identifier)** and **Privilege Designation**.
    d. **Overwrite** - select **Overlay Only**.
4. Click **Run**.

2.4 Performance baselines

The following table presents performance baselines for using the Relativity provider in Integration Points to promote documents and metadata between workspaces.

To better identify Integration Points capabilities of exporting data between workspaces the testing sets of Integration Points jobs were divided into three bucket categories: small, medium and large.

The jobs which yielded the following metrics were performed on a basic test environment of Relativity. Due to differences in data, infrastructure, and configuration, these should not be used as a benchmark of what you expect to see in a production environment. The results may not scale linearly.

Small	Medium	Large
Size 10-50 MB Records max 200 Fields max 200 Append Only Duration 1-2 minutes	Size 51-1500 MB (natives) Extracted text - null Records max 5K Fields max 300 Append only Duration 2.30 min	Size 25GB (natives) Records max 500K Extracted text - null Fields max 100 Append only Duration 115min
	Size 51-1500 MB (extracted text) Records max 50K Fields max 300 Append only Duration 11 min	Size 200GB (natives) Records max 250K Extracted text - null Fields max 200 Append only

Small	Medium	Large
		Duration 8hrs
	Size 51-1500 MB (extracted text) Records max 50K Fields max 300 Overlay only Duration 4.5hrs	Size 12.5GB (natives) Records max 250K Extracted text - null Fields max 100 Overlay only Duration 4 hrs

Set Number	Copy Natives Files Set-ting	Overwrite Setting	Total Size	Natives	Metadata	Extracted Text	Records	Fields	Duration (seconds)
Small 1c	Physical Files	Append Only	50 MB	50 MB	3 MB	1 MB	100	190	58.33
Small 2b	Physical Files	Append Only	30 MB	30 MB	3 MB	1 MB	200	85	58.45
Small 3c	Physical Files	Append Only	10 MB	10 MB	3 MB	1 MB	100	190	61.87
Medium 1	Links Only	Append Only	600 MB	0	600 MB	0	30,000	200	319.81
Medium 2	No	Append Only	~1.5 GB	0	~5 MB	1.5 GB	50,000	2	15644.38
Medium 2a	No	Append Only	~1.5 GB	0	~5 MB	1.5 GB	50,000	2	663.66
Medium 3	Physical Files	Append Only	~1.55 GB	1.5 GB	~50 MB	0	5,000	100	124.13
Medium 4	Physical Files	Append Only	~1.65 GB	1.5 GB	~150 MB	0	5,000	300	133.48
Large 1	Links Only	Append Only	~9 GB	0	~9 GB	0	250,000	350	7569
Large 3	Physical Files	Append Only	~30 GB	25 GB	~5 GB	0	500,000	100	6909
Large 4	Physical Files	Append Only	~37.5 GB	12.5 GB	~2.5 GB	0	250,000	100	14831
Large 5	Physical Files	Append Only	~205 GB	200 GB	~5 GB	0	250,000	200	29256

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.