Utilities and Files

Utilities and Files

Administrators can create multiple utilities for processing different types of objects and performing different operations.

This section contains:

MIME Types

Administrators define MIME types that must be available to staff when they configure material flows. Administrators can configure MIME types by accessing the MIME Types code table as described in Working with Code Tables.

Stream Handlers (Deprecated)

Stream handlers are part of the Generic Representation task that creates derivative copies. Their table maps the name of the stream handler (such as `tiff_2_jpeg`) to the script that calls the utility/tool for migrating files. Any parameters are also stored in the table.

Creating a Stream Handler

The Rosetta system comes supplied with a number of predefined stream handlers (such as thumbnail generation and technical metadata extraction). In addition, Administrators can create a new generic representation or thumbnail stream handler.

Administrators configure the components that define task behavior as described in the table below:

Task Components

Component	Description	Defined In...
Stream handler utility	A program that performs an operation on an object	Stream Handler mapping table
Script	Code that wraps the stream handler utility and enables launching it with specified parameters	Any external text editor
Task rule	A rule that defines the stream handler utility to be used for the specific object	One of the following: • Generic representation task

Component	Description	Defined In...
		rules • Thumbnail rules

The process of stream handler creation consists of the following stages:

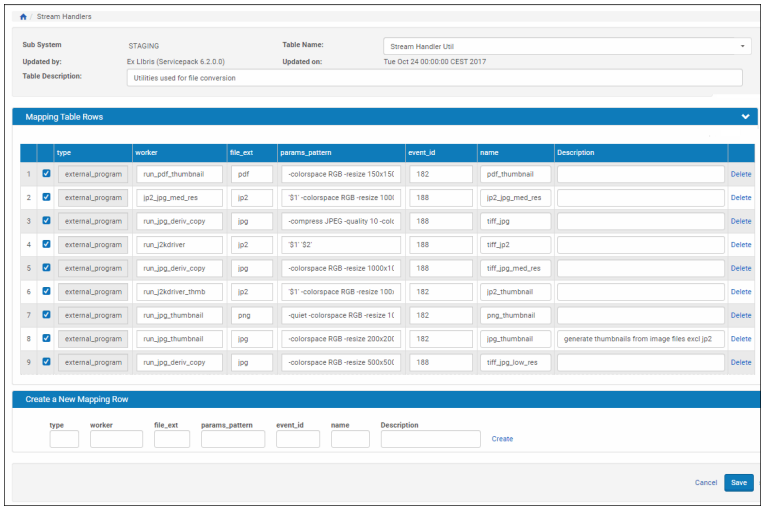
1. Creating a script that runs a stream handler with specified parameters
2. Creating an entry for the new stream handler in the Stream Handlers mapping table
3. Creating a rule that the Rosetta system uses to determine the stream handler to be used for processing a specific object

To create a stream handler:

1. Create a script that wraps the stream handler utility and enables launching it with specified parameters. The following example contains the code of the script that wraps the `pdftotext` utility.

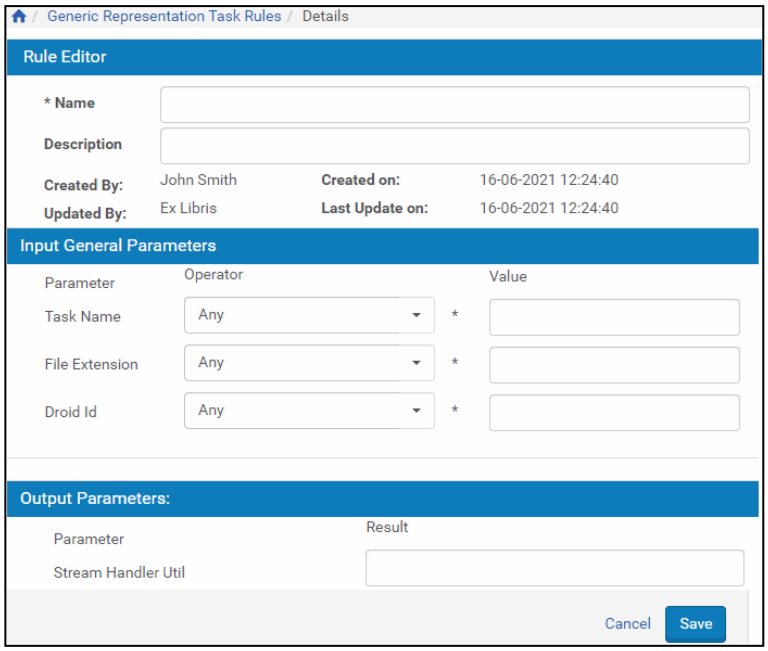
```
#!/bin/sh

for ARG do

T_ARG="\"$ARG\""

ARGS="$ARGS $ARG"

T_ARGS="$T_ARGS $T_ARG"

LAST_ARG="$ARG"

done

eval pdftotext $T_ARGS

exit $?
```

2. On the Administration page, click **Utilities and Files > Stream Handlers**. The Stream Handlers mapping table opens.



Stream Handlers Mapping Table

3. Enter information under the **Create a New Mapping Row** bar at the bottom of the page.
4. Click **Save**. The new entry is displayed in the mapping table.
5. On the Administration page, click **Utilities and Files > Generic Representation Task Rules**. The Generic Representation Task rules page opens.
6. Click **Add Rule**. The Task Rule Editor page opens.



Task Rule Editor Page

7. Complete the fields as required. For detailed information about parameter operators and values, see Operators Used in Rule Parameters.
8. Click **Save**. The new rule is saved to the Rosetta system.

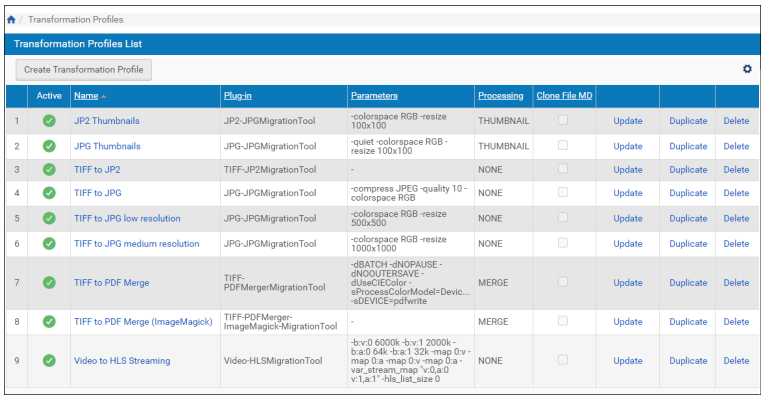
The task is saved to the Rosetta system. Administrators can now use this task in task chains.

Transformation Profiles

Transformation profiles are migration plugin-based alternatives to the Stream Handler framework. They can be used to generate derivative copies where one-to-many or many-to-one relations exist between the original file format and the derivative copy format (for example, one multi-page TIF file to several JPGs or several TIF files to one PDF).

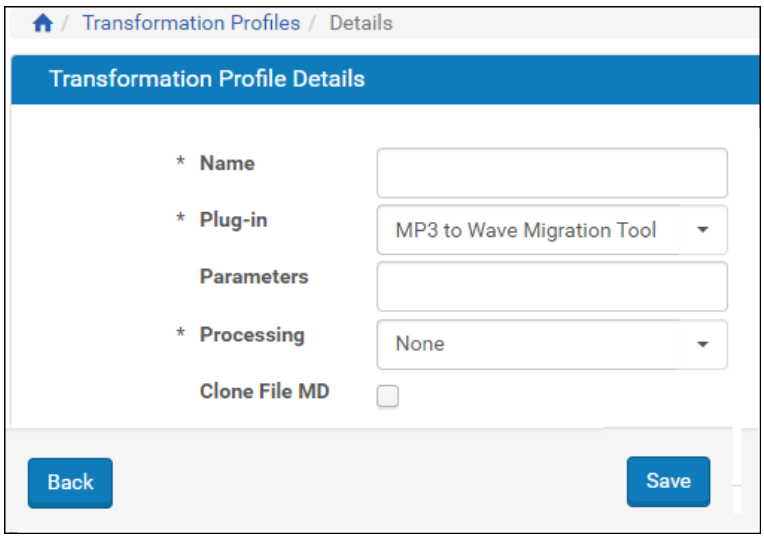
To create a Transformation Profile:

1. From the System Configuration page, open the Transformation Profiles page (**Repository > Utilities and Files > Transformation Profiles**):



List of Transformation Profiles

1. Click the **Create Transformation Profile** button.
   The Transformation Profile Details page opens.



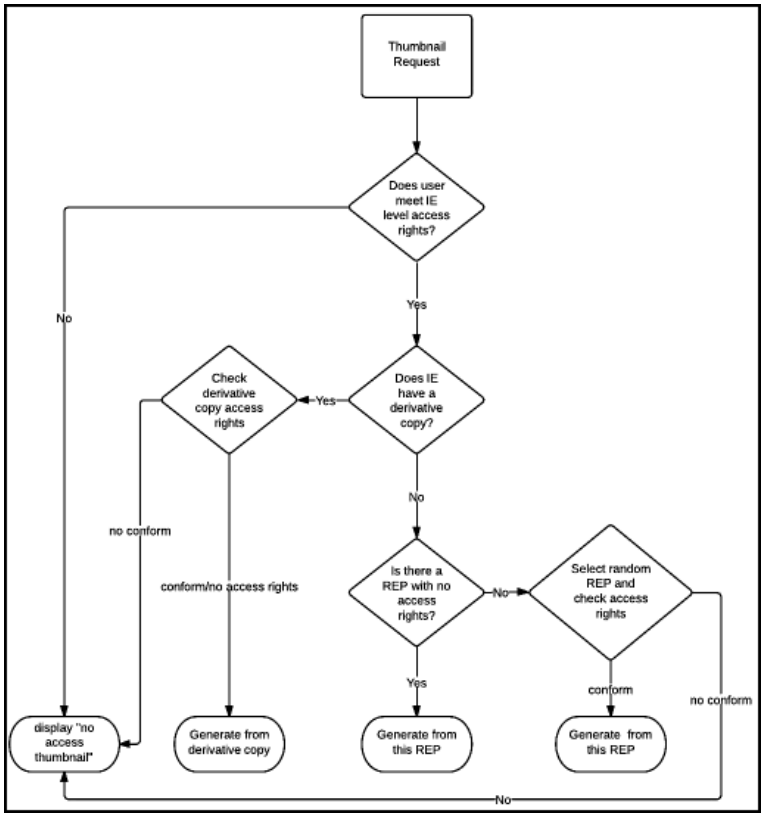
Transformation Profile Details Page

The following table describes the fields available on the Transformation Profile Details page.

Fields on the Transformation Profile Details Page

Parameter	Type	Description	Notes
Name	Text	Name of profile	
Plug-in	Combo/ drop-down	A list of installed migration plug-ins	
Parameters	Text	Arguments to pass to the plug-in	
Processing	Combo/ drop-down	Instructions to Rosetta on how to process the files.	• None - Each original file will be transferred into one derivative copy file • Merge - All files in the selected Representation will be merged into one derivative copy file • Split: A multiple-file derivative copy Representation will be created from the single-file original Representation

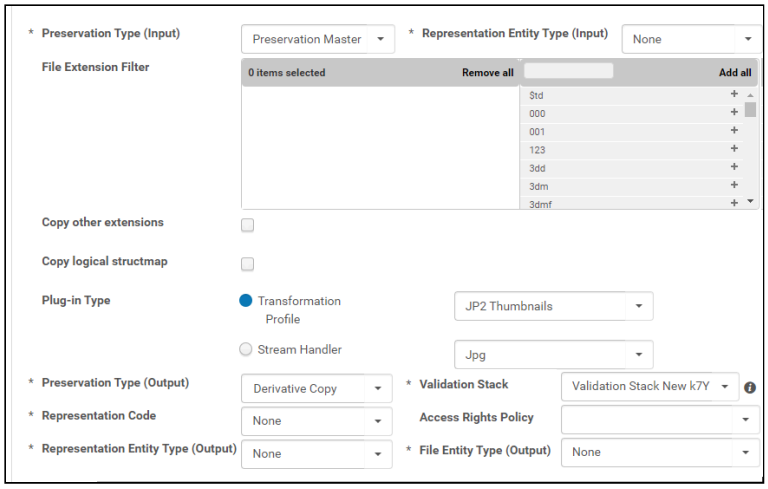
Parameter	Type	Description	Notes
			• Thumbnail: specifies this profile will be used for generating thumbnails and not derivative copies
Clone File MD	Checkbox	Copy file-level descriptive/source metadata from the original	



Note

- The responsibility for the actual splitting/merging of the files is the plug-ins.
- If no files match the profile configuration, no representation is created.

A transformation profile can be used in the Create Copy Representation task. See the following figure for an example.



Task Parameters for Transformation Profile

Generic Representation Creation Rules

Generic representation creation rules are used by the system to define which stream handler to use based on the original file format/extension and the task name.

The task name is configured as part of the Create Derivative Copy task chain.

For more information about configuring the task chain, see Managing Processes.

For information about creating a stream handler as part of a generic representation task, see Stream Handlers (Deprecated)

Thumbnail Creation Rules

The thumbnail image can now be added to an IE in a dedicated derivative copy representation. This dedicated derivative copy must be marked as such in order to be used as a thumbnail. This requires setting the **UsageType** to **THUMBNAIL** and the **RepresentationCode** to **IE**.

This is a standard representation, and it is expected to behave and be treated accordingly. For example, access rights should be allowed to be attached to the representation.

Note

The Rosetta thumbnail selection algorithm must be adapted to first attempt to display the thumbnail image in this representation before attempting to use any generated or generic thumbnail. If the user does not have access to the IE thumbnail representation, it should be ignored in the selection process.

Rosetta can also automatically create a thumbnail for each file when it's delivered. In order to create the correct thumbnail for each file format, the system uses these rules to match the file format with the appropriate utility that should be used to create the thumbnail.

Rosetta generates thumbnails from a given representation (REP) using the following algorithm:

1. Derivative copy
2. REP with no access rights
3. Any other REP

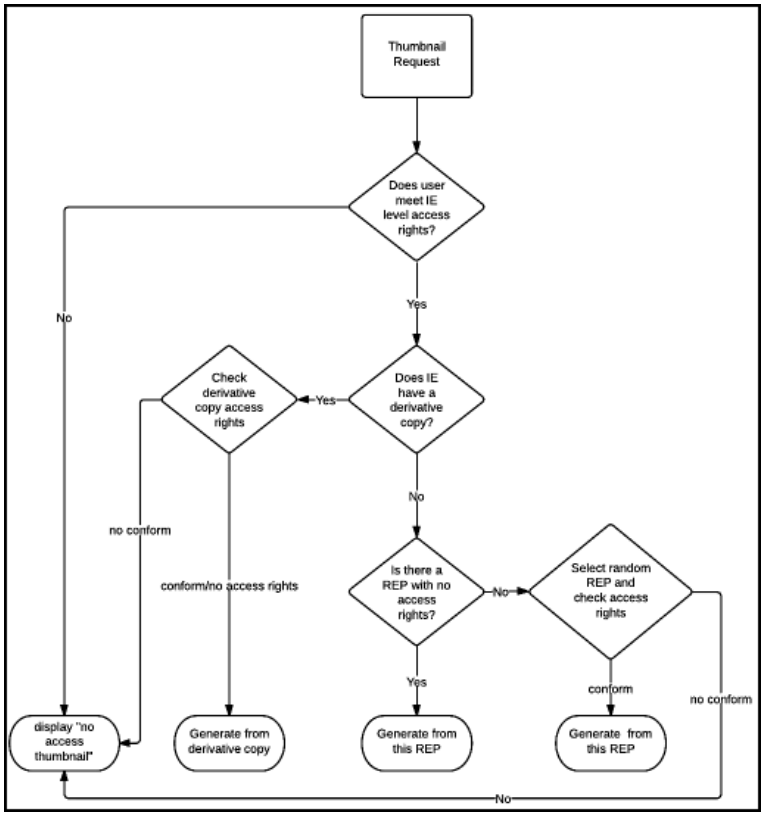
If access rights are not met for the selected REP, a "no access" thumbnail is displayed.

Note

Administrators can disable access rights checking for thumbnails by setting the check_thumbnail_access_rights general parameter to false. See Rosetta System Administration Guide, General Parameters - Delivery).

If there is no thumbnail creation rule for the selected REP's files, a generic thumbnail is displayed (see Generic Thumbnail Creation).

The following flowchart demonstrates the process of thumbnail request to thumbnail generation.



Thumbnail Creation Flow

The order of an IE's thumbnails will be according to the (first) logical structmap. If no logical structmap exists, order will be determined by the physical structmap.

By default, up to five thumbnails can be generated for an IE. This can be increased by the ie_thumbnail_limit general parameter (see Rosetta System Administration Guide, General Parameters - Delivery). The thumbnail view in user interfaces and viewers indicates the actual number of files per IE.

The recommendation remains to create thumbnails as part of the enrichment process or ongoing maintenance task chain to improve user experience for modules that display thumbnails."