

---

## How to create and load a CSV file into Rosetta

- **Product:** Rosetta
  - **Relevant for Installation Type:** Local;
- 

### Procedure:

1. Go to Deposits > Deposit arrangements > CSV Templates and click "Add CSV Template" to create a CSV template:

a. Select metadata fields from the right column and double-click or drag them to the left column and click "Save" to save the template.

See "Additional Information" for example csv with basic metadata fields.

2. Go to Deposits > Deposit arrangements > Content Structures

a. Select the "CSV Loader Converter" under Add Content Structure and click "Add" to create a new content structure for the CSV file.

b. Give the new content structure a name, link it to the appropriate CSV template (drop-down), and choose the appropriate Generate CSV Option (drop-down).

c. Download the newly-created CSV template by clicking on the "Download CSV Template" button and save locally.

d. Click "Save" to save the newly-create content structure.

3. Go to Deposits > Deposit arrangements > List of Submission Formats and click "Add Submission Format"

It's possible to create either a "Detailed CSV" which requires submission of a csv and zip files or an "NFS Acquiring" which only requires submission of the csv, as long as the "File Original Path" includes a resolvable URL to the source file.

If using Detailed CSV use the default "Detailed CSV" submission format which is not editable (as of Rosetta v4.2.1).

### If using NFS Acquiring:

a. Under "Submission Format Details" add the following:

Name: <mandatory>

NFS Path: (e.g. /mnt/operational\_shared/sipTmpDir/test\_load)

Allow Navigation: Yes

Min. Number of Files: 1

4. Go to Producers > Deposit Arrangements > Material Flow List and click "Add Material Flow"

a. Update the following mandatory fields:

Material Flow Definition (manual or automated)

-Name: <add name>

-Internal: No

Technical Definitions

-Select content structure: <add content structure defined in step 2. above>

-Select submission format: <add submission format defined in step 3. above>

Descriptive Definitions

-Select Metadata form: <add a metadata form appropriate for CSV ingest>

b. Click "Save" to save the Material Flow

5. Deposit with the material flow and review the log and BIRT reports for results.

---

## Additional Information

[CSV Workflow Recording \(6 minutes\)](#)

**Example CSV metadata Fields to include ([see also attached file here](#)):**

SIP: Title (DC)

Collection: Publish Collection [Title (DC) and Is Part Of (DCTERMS) are added automatically]

IE: Title (DC), Creator (DC), Type (DC), Identifier (DC), Identifier - URI (DC), Description (DC), Date (DC), Subject (DC), IE Entity Type

REP: Preservation Type, Usage Type

NOTE: Rosetta only allows for one PRESERVATION\_MASTER preservation type per IE.

Preservation Type options defined in (Administrative module): General > Code Tables> Subsystem: Preservation > Details  
PRESERVATION\_MASTER

MODIFIED\_MASTER  
DERIVATIVE\_COPY

Additional values can be added if needed.

File: FILE - Identifier (DC) [File Original Name and File Original Path are added automatically]

### Repeating DC metadata fields

DC fields in the CSV file can be repeated as many times as needed.

For example, a CSV file can include several IE level Subject (DC) columns.

Each column will become separate dc:subject tags in Rosetta METS.

Please note the following limitation: multiple columns with the same DC field on certain level can be only added by manually editing the CSV file downloaded from the CSV template page.

Currently the UI in Deposit > CSV template does enable to add each field only once.

### How to add File Labels for viewer display

During csv ingest the file names in the "File Original Name" column are added as the file labels by default (e.g. "123.jpg" - complete file name).

If it is preferred to use instead the DC titles as the file labels, in addition to the "File Original Name" column, add two additional columns:

1. File - Title (DC) with the "123.jpg" (complete file name as would be located in the path provided in the "File Original Path")
2. File Label with the file title "Map of Indonesia" or whatever file label that should display in the viewer's table of contents on the left.

### Important notes about the "Detailed CSV" submission format when used with a CSV file:

1. Make sure the file names listed in the 'File Name' column of the CSV match the names of the actual files.

2. Unless the zip contains internal folders:

- For Rosetta 7.0 and below - 'File original path' is: / (forward-slash)
- For Rosetta 7.1 and newer - 'File original path' is empty.

3. When the zip contains internal folders - 'File original path' is the path of files within the zip.

Make sure 'File original path' do not start with '/' (forward-slash).

Examples:

- a file directly under zip's root: *image.jpg*
- a file under few subfolders under zip's root: *folder1/folder2/image.jpg*

### Important notes about the "NFS Acquiring" submission format when used with a CSV file:

In v4.2.1 the CSV automated deposit now supports absolute and HTTP paths.

To add a remote path, add the url, starting with "http://" to the 'File Original Path' column including the file name.

Make sure that the 'File Original Name' column is *empty*.

To add an absolute path add the url to the 'File Original Path' and the file name to 'File Original Name', make sure it starts with '/' ."

## Scientific notations and Foreign language characters:

1. If automating CSV creation, the script should assure that non-byte order mark (BOM) UTF-8 encoded CSV is created.
2. For manual CSV creation, use LibreOffice or Notepad++ instead of Excel to save CSV files in non-byte order mark (BOM) UTF-8.

## Collection Hierarchy Building Structure:

<u>Object Type</u>	<u>Title (DC)</u>	<u>Is Part Of (DCTERMS)</u>	<u>Publish Collection</u>
Collection	collection1		TRUE
Collection	collection2	collection1	TRUE
Collection	collection3	collection1/collection2	TRUE
Collection	collection4	collection1/collection2/collection3	TRUE

Will create the following collection hierarchy in Rosetta's Collection Management: collection1/collection2/collection3/collection4/<IEs>

**NOTE:** Rosetta Collection Management only displays dc:title in the Title column of the Content tab (not dcterms:title).

Refer to the "Depositing SIPs in CSV Structure" section of the Rosetta's Producer Guide:

[http://knowledge.exlibrisgroup.com/@api/deki/files/39701/Rosetta\\_Producers\\_Guide.pdf](http://knowledge.exlibrisgroup.com/@api/deki/files/39701/Rosetta_Producers_Guide.pdf)

Refer to the "CSV Content Structure" and "CSV Templates" sections of the Rosetta Staff User's Guide:

[https://knowledge.exlibrisgroup.com/@api/deki/files/39696/Rosetta\\_Staff\\_User's\\_Guide.pdf](https://knowledge.exlibrisgroup.com/@api/deki/files/39696/Rosetta_Staff_User's_Guide.pdf)

For explanation with screen shots see attached file

Please see the relevant documentation for additional information and limitations

## CSV Loading Pre-checks

Below is a list of preliminary checks that can be performed before ingesting a CSV.

Following these pre-checks will reduce the number of potential errors encountered during loading.

1. Confirm that all of the files listed in the "File Original Name" column in the csv correspond to the ones listed in /content/streams/ directory.
2. Confirm that all of the file names listed in the "File Original Name" column in the csv match the actual file names (e.g. case sensitivity)?
3. Confirm that the "File Original Path" column accurately reflects where the files reside (e.g. path matched was is defined in the submission format).
4. Confirm that there are no extra spaces after the column names and/or column content (the submission to fail and get

routed to the TA workbench).

5. Confirm that there are no typos in the SIP folder name that would NOT match the path in the csv's "File Original Path" column (e.g. sampleJam instead of sampleJan).

6. Confirm that there is a "/" after the "streams" folder path in the "File Original Path" column in the csv (e.g. /operational\_shared/sipTmpDir/test/content/streams/).

7. Confirm that producer agent to be used for csv ingest is linked to the producer. If they are not linked, the producer/producer agent will not display in the Submission Job.

#### **Update to Rosetta version 5.4**

Rosetta version 5.4 supports also the Source MD ingest with CSV deposit.

It is available by adding a source MD object in the "{IE/REP/FILE} Source metadata content" column.

---

## Attachment

[How to load CSV file](#) (for Rosetta 7.0 and below | for more information see 'Important notes about the "Detailed CSV" submission format when used with a CSV file')

[CSV Example](#)

**Category:** Deposit

---

- **Article last edited:** 27-Jul-2021