Uploading Files via Batch Ingest

(Note: the following borrows significantly from Maria Whitaker's documentation on creating a batch manifest file; see more at Batch Manifest)

Batch Ingest

Batch manifest template

Add items to Unit: AMP Pilot Unit

To add items as a batch, please use the AMP Batch Item template. Once items have been prepared using the template, click upload below.

Choose File | TemplateForDevs.csv   Upload

The Batch Ingest feature is how files are uploaded to AMP. To use the Batch Ingest, you will need to create a batch manifest in CSV format, and then upload it to AMP on the Batch Ingest page. All the files in the batch manifest first need to be uploaded to their respective collection's subfolder in the dropbox via an SFTP client; the Batch Ingest will fail if a file is not found in the expected dropbox subfolder. Detailed instructions for how to upload files to a dropbox using a tool like Cyberduck or WinSCP can be found here. It is highly recommended that users use Cyberduck, as it has built-in integration with Google Drive.

Batch manifests must conform to a specific format for the batch to be properly ingested. The following are the column names required for a batch manifest (note that some of these are optional):

Format

Collection name (required)

The collection must already exist in AMP, and collection names must match exactly. The name of the collection is used to determine:

- which collection dropbox holds the media files to be uploaded
- to which collection in AMP the information in that line refers.

Collection names cannot be changed by end users; however, they are able to be changed by AMPPD staff.

External Source (optional, strongly recommended if using External Item ID)

This field is optional. It is used to tell AMP the source system of the item. This information is added by AMP in the bag it provides with AMP-generated metadata for target systems to consume. If External Item ID is being used, it is strongly recommended to provide an item's External Source, as it allows items in the same collection from multiple sources to have the same External Item ID.

External Item ID (optional, recommended)

This field is optional, although recommended. When provided, it is used as the unique identifier of the item during the batch ingest process (if not, Item Title is the unique identifier). This allows items within the same collection to have the same Title (this is a relatively common occurrence, at least in some of the pilot collections).

Item Title (required)

The bibliographic title of the item. If the External Item ID is not provided, this field serves as the unique identifier for an item in a collection; in this case, one cannot have multiple items with the same title within a collection.

Item Description (optional)

When provided, this field will be displayed as the item description in AMP.

Primary File (required)

This is the file name of the media file that has been placed in the Dropbox for ingestion. The uniqueness of a primary file is the combination of Collection, Item/External ID, and Primary File Label. File names are unique within an item, but not necessarily across items (that is, two distinct items can have the same primary filename).

Warning: if a batch manifest includes two lines with the same value in the Primary File column for different items in the same collection, the validation step will let it pass, but the ingestion process will have a runtime error because in the dropbox there can only be one copy of a file with that name. To resolve this conflict, ingest the primary files with conflicting names using separate batches.

Primary File Label (required)

Users must provide a label (or title) for the file. That label will be used to uniquely identify the file within this item. In other words, one cannot have multiple primary files with the same label if they are associated with the same item within the same collection.

Primary File Description (optional)

This field may be left blank. When provided, it will be displayed as the file description in AMP.

Supplemental File fields

Multiple supplemental files can be specified per line. For each supplemental file, you need four fields:

- **Supplemental File Type** - The user must specify if they want to place the file at the Collection, Item, or Primary File level
- **Supplemental File** - the filename of the binary file (which needs to be found in the same Dropbox as the media files)
- **Supplemental File Label** - the user must provide a label (or title) for the supplemental file. That label will be used to uniquely identify the file in association with the item. One cannot have multiple supplementary files with the same name associated with the same item within the same collection.