Uploading Content via Batch Ingest (deprecated)

⚠ This page is deprecated as of Release 6.3. For current information about Batch Ingest, see Uploading Content using Batch Ingest.

Introduction

While content can be uploaded to Avalon either individually or several at a time using an Avalon dropbox, the batch ingest feature creates many items (and their metadata) at one time outside of the Avalon user interface. This user guide will explain how to form an Ingest Package and upload it into Avalon using the Batch Ingest.

Manifest Template

- manifest_example.xlsx

Definitions

- **Batch Ingest:** the process of creating multiple media items by uploading an Ingest Package to an Avalon dropbox
- **Ingest Package:** the combination of content and metadata, consisting of one Manifest File and one or more Content Files
- **Manifest File:** a spreadsheet containing the metadata for the items to be created as well as the names of the Content Files that make up each item
- **Content File:** a single media file that is part of an item; one or more Content Files can make up an item in the Avalon Media System
- **Item:** a single media object accessible through one Avalon page; consists of one or more media files and metadata describing the media file(s)
- **Collection:** a grouping of items in Avalon for administrative and discovery purposes. Items belong to one and only one collection

Preliminaries

When a new collection is created, Avalon creates a sub-directory with the name of that collection (substituting underscores for any blanks), beneath the Avalon dropbox directory. A Batch Ingest is started by uploading an Ingest Package to the collection sub-directory in the dropbox. To connect to the Avalon dropbox, see Uploading Content to an Avalon Dropbox.

The manifest file is a spreadsheet (xls, xlsx, csv, or ods) containing the metadata for the items to be created, as well as the names of the content files that make up each item; download manifest_example.xlsx for a blank manifest template. The manifest file contains many fields for metadata to be attached to an item when the batch ingest begins. Only a few of these fields are required: "Title," "Date Issued," and "File." For a description of the other fields, as well instructions for adding structure or captions, see Batch Ingest Package Format.

Creating a Basic Manifest File

1. Open the Manifest File that will contain the metadata for the Content Files.
2. Column A Row 1 contains a reference name for the batch; this field is only for reference and should be renamed to something convenient or descriptive.
3. Column B Row 1 contains the email address or username of the submitter, to be used for notifications and exceptions. The submitter's username or email address must be listed as a manager, editor, or depositor for the collection to which the batch is submitted.
4. Each of the columns on Row 2 specifies the potential metadata fields for each Item. If a field for an item is multi-valued (e.g. has more than one topical subjects or languages, or the Item is made up of more than one Content File), the column can be duplicated with the same header.
5. Row 3 and and all subsequent rows each represent a single Item to be created in Avalon. The minimum required fields for each item include:
   a. **Title:**
      i. Not repeatable: each item may only have one title.
      ii. The title field represents the name of a single item, which may consist of one or more Content Files.
      iii. If title is not available or missing, create a title that describes something about the content of the item. This is necessary for identifying items in search results.
   b. **Date Issued:**
      i. Not repeatable: each item may only have one date issued.
      ii. Date should be the main publication date associated with the item.
      iii. Simple format should follow "yyyy-mm-dd", see EDTF specifications for other options, including a date range or an unknown date.
   c. **File:**
      i. The name of a Content File, which consists of two parts:
         1. The path to the file, relative to the Manifest File
         2. The name of the file itself, including the file extension (e.g. video1.mp4, songB.aac)
         3. Both path name and file name are case-sensitive; incorrect cases will result in a processing error.
         4. Example: A Manifest File and a folder titled "video_content" (with .mp4 Content Files) are uploaded to the dropbox collection "MyVideos". The Manifest File and the folder are now located in the same directory; therefore, each file name would follow the format of *video_content/videoname.mp4*.
      ii. Repeatable: if an item consists of several Content Files, the column "File" should be duplicated for each individual Content File constituting the item.

1. Each "File" column should be immediately followed by a "Label" column.
2. "Label" represents the name or short description of each Content File (e.g. Part 1, Introduction, essay topic, musical movement, etc.)

Example Manifest File showing reference name, username, Title, and Date fields (please note there are some intentional errors in this example):

	A	B	C	D	E	F	G	H	I	J	K	L
1	Short Films	johndoe										
2	Bibliographic ID La	Bibliographic ID	Other Identifier Type	Other Identifier	**Title**	Creator	Contributor	Genre	Publisher	Date Created	**Date Issued**	Abstract
3					Jellyfish Swimming							
4					Reindeer in Alaska						2017-04-18	
5					Bubbles in Central Park						2017-04-19	
6					Snowy Branch						2017-04-20	
7											2017-04-21	
8					Wilderness Lookout						2017-04-22	
9					Goldfish in Wine Glass						2017-04-23	
10					Ink Drops						2017-04-24	
11												

Continued (file name and path name, file labels):

AB	AC	AD	AE	AF	AG
File	Label	File	Label	File	Label
short_films/jellyfish.mp4					
other_short_films/reindeer.mp4					
other_short_films/bubbles.mp4					
other_short_films/branch.mp4					
short_films/iridescence.mp4					
other_short_films/wilderness.mp4					
other_short_films/goldfish.mp4					
short_films/ink_1.mov	Part 1	short_films/ink_2.mov	Part 2	short_films/ink_3.mov	Part 3

Uploading an Ingest Package to the Dropbox

A Batch Ingest will begin automatically once the Manifest File and the Content Files are uploaded to the appropriate collection in the Avalon dropbox. If you are not yet connected to the Avalon dropbox, follow the instructions at Uploading Content to an Avalon Dropbox. Once you are connected, uploading the files is a simple drag-and-drop process. The image below shows the files to be transferred in the left pane, and the files after they have been transferred in the right pane.
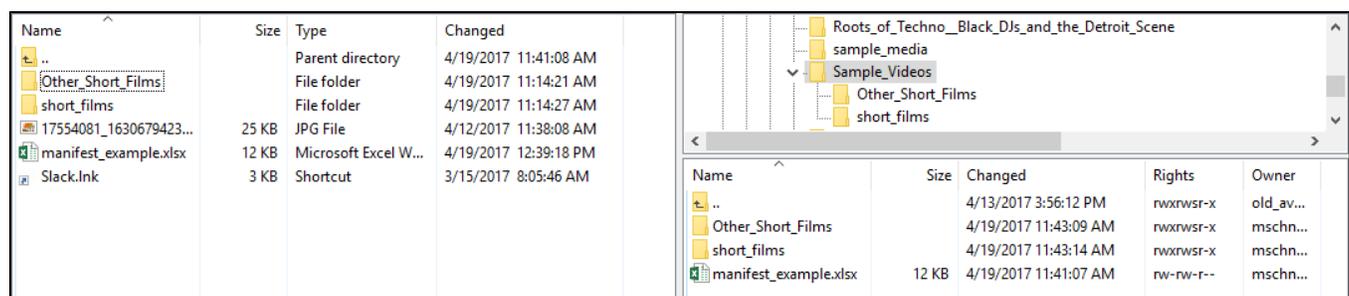
Troubleshooting a Failed Ingest

Once Avalon detects the presence of an unprocessed Manifest File, it will first verify that the metadata columns are recognizable, that the required columns are present and have values in them, and that the package is *complete* (i.e., all content files specified in the manifest are present and not open by any other processes) before attempting to ingest. If these conditions are met, Avalon will generate and send an email to the user listed on the Manifest File. A successful ingest will state:

"Your metadata package was validated and no errors were found. Your batch is now processing.

Manifest file: /N/beryllium/avalon/dropbox/Sample_Videos/manifest_example.xlsx"

At this point, no further action is necessary. The content files will automatically be created as items in Avalon with the metadata from the Manifest file.

However, an unsuccessful ingest will generate a message to the user with the details of the attempted process. Please note that even if only a single record contains an error, the entire batch ingest will fail and no items will be created.

Example failure message:

Initial processing of your batch ingest failed. Details are below. Please see the Batch Ingest Package Format for further help.

Manifest file: /N/beryllium/avalon/dropbox/Sample_Videos/manifest_example.xlsx

Row 2: Jellyfish Swimming
Date Issued - field is required.
Row 3: Reindeer in Alaska
Content - File not found: /N/beryllium/avalon/dropbox/Sample_Videos/other_short_films/reindeer.mp4, No files listed
Row 4: Bubbles in Central Park
Content - File not found: /N/beryllium/avalon/dropbox/Sample_Videos/other_short_films/bubbles.mp4, No files listed
Row 5: Snowy Branch
Content - File not found: /N/beryllium/avalon/dropbox/Sample_Videos/other_short_films/branch.mp4, No files listed
Row 6: Untitled
Title - field is required.
Row 7: Wilderness Lookout
Content - File not found: /N/beryllium/avalon/dropbox/Sample_Videos/other_short_films/wilderness.mp4, No files listed
Row 8: Goldfish in Wine Glass
Content - File not found: /N/beryllium/avalon/dropbox/Sample_Videos/other_short_films/goldfish.mp4, No files listed

This example, which was generated from the failed batch ingest of the Ingest Package shown in the above images, demonstrates some of the most common issues preventing a successful ingest. A few of the files were missing fields (either Date Issued, Row 2, or Title, Row 6), but the rest of the files were indicated as "No files listed." This can either indicate that a Content File was not included in the transferred files, or that there was an error in the "File" field. In this particular case, the folder was named "Other_Short_Films" while the path names were written as "other_short_films". Path and file names are case-sensitive, and therefore the batch ingest failed.

Other issues may involve incorrect formatting in other metadata fields; refer to the Batch Ingest Package Format for specific details about each field.