📖 **OhioLINK**   **DRMC**   [   ]   [Search this site]

Navigation

Collections in queue

Meetings

Membership

OhioLINK Digital Resources Management Committee Wiki

Requested Features

Issues, Problems, & Progress

Task Forces and Subcommittees

Metadata

DRC Liaison Assistance

Best Practices

Event Promotion and Coordination

Sitemap

121
days since
next DRMC Meeting

Bulk Submission

Creating Your Own Excel Metadata Spreadsheet

You can download the "Batch Importer 2010" file attached at the bottom of this page to fill in your metadata.
Before you begin, you might find this video demo helpful (the URL is http://www.screencast.com/t/nVlnYXWYb).

Preparing collections for bulk submission to the DRC is a very straightforward process.
There are two main parts: Organizing the files you want to submit, and creating a spreadsheet of your metadata.

Tutorial -- Organizing Your Files

For this example I will be using 5 images from the Ohio Agricultural Experiment Station Forestry Image Collection. (If you'd like to follow along, the attached file, 'Example_files_batch_import.zip' contains everything you'll need.)
Create a folder to hold all of your items. I created one called 'forestry' and placed it directly on my 'C:\' drive, but you can place it anywhere you're familiar with.
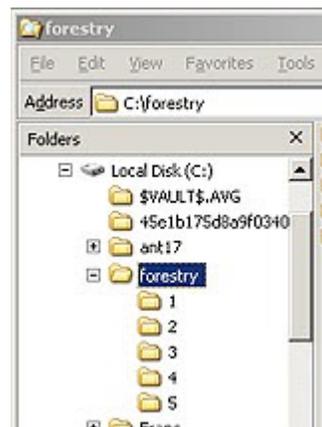
Top level folder

You'll also notice I've made subfolders for each of the items in the collection. I've simply named them 1, 2, 3, 4, and 5, but you can use any naming convention you wish.

Each subfolder contains all of the files that make up a single item. Don't let this confuse you: most collections are made up of just one file and its metadata. The forestry example has only a single JPEG image for each item, and we will be

adding the metadata record in the next step.
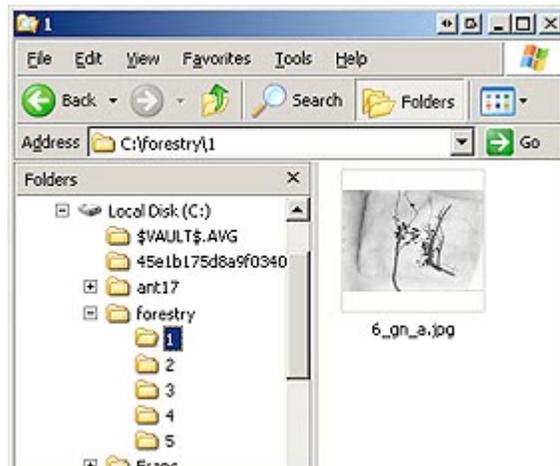
Forestry subfolder 1 only has a single JPG bitstream

We use the term 'bitstream' to describe the files that make up an item: single bitstream item vs multi-bitstream item

Oftentimes, a multi-bitstream item contains an archival master TIFF scanned at high resolution, and a smaller presentation-scale JPG file, or perhaps an Adobe PDF, if the TIFF is a scan of a text document.
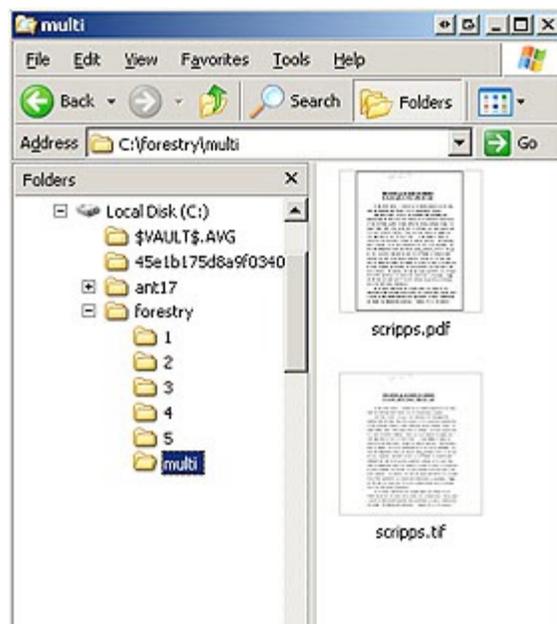
Multi-bitstream: an original TIFF and a PDF file

A collection can contain a mix of single- and multi-bitstream items.
Once you've created subfolders with all of your items, it's time to make the matching metadata records.

Tutorial -- Creating a Metadata Spreadsheet

Each of your subfolders needs to contain two text files, a 'dublin_core.xml' metadata record and a 'contents' manifest. The XML file contains the Title, Author, Description information, and the 'contents' file lists every bitstream being submitted with each record. Luckily, the Excel spreadsheet automatically creates these files for you.

Launch the file Example_files_batch_importer.xls and make sure to 'allow (or enable) macros' and follow these easy steps:

Type your records one per row in the spreadsheet. If needed, add more metadata fields by typing new column names at the top (contributor:author, subject:lcsh, etc.).

| A1		fx contributor:author			
	A	B	C	D	
1	contributor:author	coverage:spatial	date:creat	identifier:o	de
2	Ohio Agricultural Experiment Station.	Ohio	1904-6	0006_gn_a	Ca
3	Ohio Agricultural Experiment Station.	Ohio	1910-8	0016_gn_a	Ca
4	Ohio Agricultural Experiment Station.	Ohio	1908-12	0001_gn_a	Ca
5	Ohio Agricultural Experiment Station.	Scioto County (Ohio)	1911-8	0010a_fn_a	Fo
6	Ohio Agricultural Experiment Station.	Ohio	2/26/1931	0009_fn_a	Ju

The last column, named ResourceLocation lists the name of the subfolder for each item

P	Q
subject	ResourceLocation
Forests and forestry -- United States	1
Forests and forestry -- United States	2
Forests and forestry -- United States	3
Forests and forestry -- United States	4
Forests and forestry -- United States	5

Lastly, point the spreadsheet to the directory that holds all the files and press Run The Export Macro.
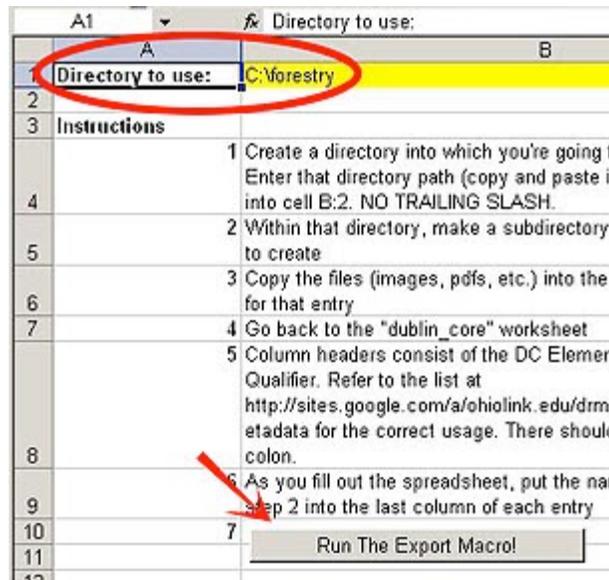
You should now see the dublin_core.xml and contents manifest files in each subfolder. Burn your collection to DVD and mail it to the OhioLINK office c/o DRC Team.
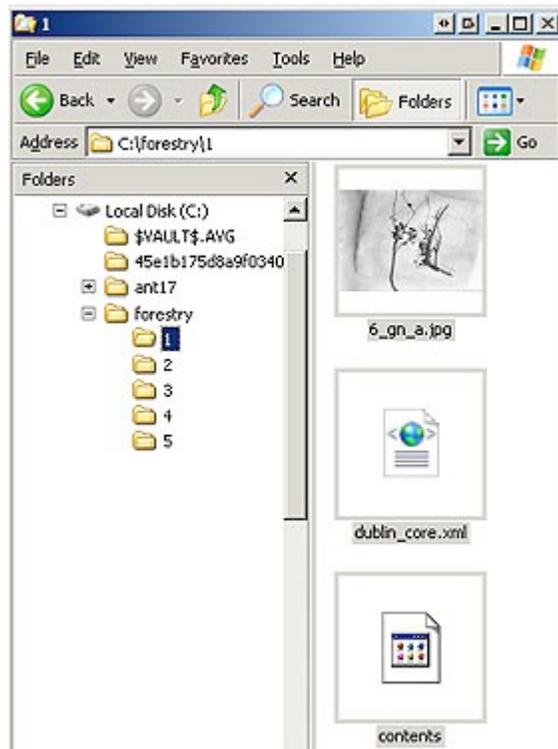
Things to remember:

1. Make sure you enable macros or allow macros when you first open up the spreadsheet. It will not run if you forget this. If you have forgotten and already filled in the rows, simply save the workbook, close it, and re-open, remembering to enable macros.
2. This is a two sheet workbook. The first (default) sheet is the dublin_core sheet; once you are ready to run the macro and generate the XML, look at the bottom of the sheet for the "Settings" tab. Switch to that worksheet and follow the instructions.
3. When you are adding columns (contributor:author, subject:lcsh, etc.), make sure that the punctuation between the element and the qualifier is a colon, not a period (the DSpace software uses a period between for elements and qualifiers; don't do that here!)
4. The first column of data must not contain any blank cells or the process will run but error out on the blank row.
5. After you run the macro, the easiest way to check to make sure that all files were generated correctly is to open the subfolder containing the item from the very last row (In the forestry example, folder 5). Check to make sure that the original files are there as well as the two files generated by the macro: "contents" and "dublin_core.xml". If those files are not there in the final folder, the process generated an error somewhere.

Attachments (3)

Example_files_batch_import-2010.zip - on Nov 17, 2010 2:42 PM by Gwen Evans (version 1)
202k Download

batch_importer-2010.xls - on Nov 17, 2010 2:41 PM by Gwen Evans (version 1)
49k Download

cudac_batch_importer.xls - on Aug 12, 2009 10:54 AM by John Davison (version 1)
62k Download

Sign in   Recent Site Activity   Terms   Report Abuse   Print page | **Powered by Google Sites**