



**Orbis Cascade Alliance
Content Creation &
Dissemination Program
Digital Collections Service**

Batch Editing Dspace Records

Produced by the Digital Collections Working Group of the Content Creation & Dissemination Team

Janet Hauck, Whitworth University (chair)

Anneliese Dehner, Metadata Applications Librarian, Alliance

Laura Zeigen, Oregon Health & Sciences University

Julia Simic, University of Oregon

Theodore Gerontakos, University of Washington (author of this document)

Jodi Allison-Bunnell, Alliance (ex officio)

Version 1.0, March 2017

Table of Contents

[Overview](#)

[Summary of steps in the process](#)

[Detailed steps](#)

[Step 0: Navigate to the community or collection you want to edit.](#)

[Step 1: Export metadata for a DSpace collection or community to your local computer.](#)

[Step 2. Edit the metadata on your local computer.](#)

[Step 3: Import the edited metadata into DSpace.](#)

[Additional resources](#)

Overview

This document proposes to describe the simplest method for batch metadata editing in DSpace.¹

This document describes DSpace 5.6 procedure.² Procedures for other DSpace versions³ may vary, but the procedures are all similar. See below, “Additional Resources,” for batch metadata editing procedures using other DSpace versions.

DSpace recommends editing batches of 1,000 records or less. For collections greater than 1,000 records, collections can be edited in subsets.⁴

¹ This method assumes use of the XMLUI, also known as the Manakin interface. DSpace batch metadata editing is also possible with both the JSPUI and command line interfaces, but these are not described here, as most DSpace instances use the XMLUI. For information on using the other interfaces, see the DSpace documentation cited in footnote 2 below.

² Official DSpace 5.x batch metadata editing documentation is at <https://wiki.duraspace.org/display/DSDOC5x/Batch+Metadata+Editing>

³ To determine (using your browser) what version of DSpace you are using: from any DSpace community or collection page, view the “page source.” Notice near the top of the page there are <meta> tags. One tag displays approximately as follows: <meta name=“Generator” content=“DSpace 5.6”>, containing the version number in a “content” attribute. Generally the display looks like the sample below, with the pertinent line highlighted:

```
<!DOCTYPE html>
<html class="no-js" lang="en">
  <head>
    <META http-equiv="Content-Type" content="text/html; charset=UTF-8">
    <meta content="text/html; charset=UTF-8" http-equiv="Content-Type">
    <meta content="IE=edge,chrome=1" http-equiv="X-UA-Compatible">
    <meta content="width=device-width,initial-scale=1" name="viewport">
    <link rel="shortcut icon" href="/researchworks/themes/Mirage2/images/favicon.ico">
    <link rel="apple-touch-icon" href="/researchworks/themes/Mirage2/images/apple-touch-icon.png">
    <meta name="Generator" content="DSpace 5.6">
    <link href="/researchworks/themes/Mirage2/styles/main.css" rel="stylesheet">
    <link type="application/rss+xml" rel="alternate" href="/researchworks/feed/rss_1.0/site">
  [etc...]
```

⁴ Although solving this problem is beyond the scope of this document, the DSpace documentation states that, when editing the CSV file exported from DSpace, rows that are removed will not affect the associated records; in other words, removing a row from the spreadsheet means you want to leave the associated record untouched. Collection subsets can be created therefore simply by deleting rows in the CSV file. See <https://wiki.duraspace.org/display/DSDOC5x/Batch+Metadata+Editing#BatchMetadataEditing-EditingtheCSV>, “Tips to Simplify the Editing Process.”

Summary of steps in the process

Step 1: Export metadata for a DSpace collection or community to your local computer.

Step 2. Edit the metadata on your local computer.

Step 3. Import the edited metadata into DSpace.

Detailed steps

Step 0: Navigate to the community or collection you want to edit.⁵

Step 1: Export metadata for a DSpace collection or community to your local computer.

Step 1.1: From the right sidebar, select the “Export Metadata” button from the “Context” section. This results in the option to save a .csv file.



⁵ DSpace groups items into “communities” and “collections.” Communities contain collections; collections contain items.

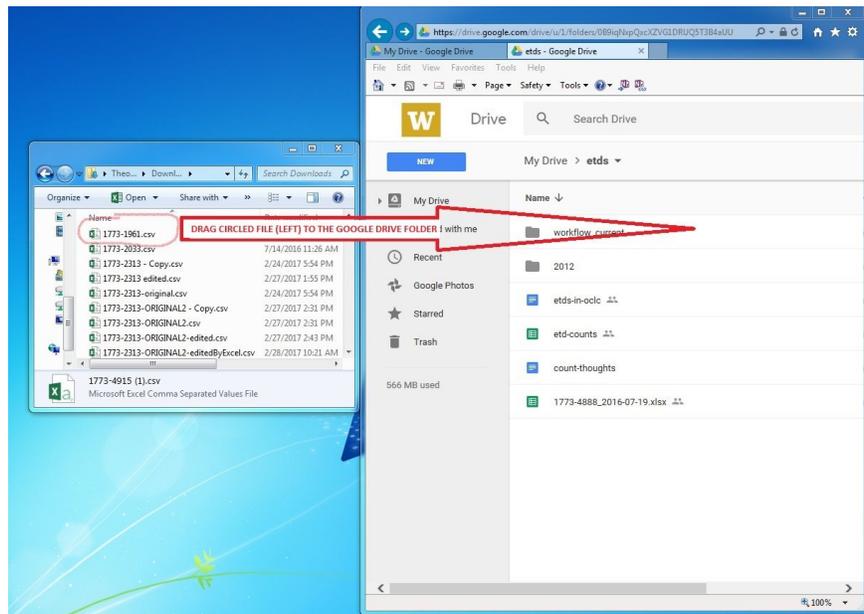
Step 1.2: Save⁶ the .csv file. Saving to the default downloads folder will work fine.

Step 1.3: Make a copy of the .csv file as a backup “just in case.” Name the copy whatever you like.

Step 2. Edit the metadata on your local computer.

Step 2.1: Open the file using spreadsheet software. We recommend Google Docs,⁷ as follows:

Step 2.1.1: Drag the .csv file from your computer’s Downloads folder to your Google Drive; you can drag it into “My drive” or into a folder of your choice. You can leave the filename as-is.



Step 2.1.2: Double click the newly displayed icon for the .csv file that loads onto your Google Drive. This opens the file in Google Sheets so that you can edit the metadata.

⁶ Save it to your Downloads folder, or wherever you prefer. Different browsers have different methods for this, but those methods are beyond the scope of this document.

⁷ Google Docs most reliably preserves the UTF-8 encoding of the DSpace-exported metadata. Preserving the UTF-8 encoding is important for the import back into DSpace, and Microsoft products often do not preserve UTF-8 encoding. This is the main reason we recommend Google Docs. In addition Google Docs reliably preserved RFC4180 CSV format and did not auto-change any data values; Microsoft Excel and Apache OpenOffice reportedly work just as well in this regard, but we did not test their performance. But if you want to use Microsoft Excel, follow the DSpace procedure at <https://wiki.duraspace.org/display/DSDOC5x/Batch+Metadata+Editing#BatchMetadataEditing-EditingtheCSV>.

Step 2.2: Perform the edits.⁸ This document groups the edits into three different types, as follows:

Step 2.2.1: Replacing already-existing metadata

- Just change it in the .csv file, as you would using a standard word processor.

Step 2.2.2: Deleting already-existing metadata

- To delete a full column: leave the column headings, delete all the values in the cells beneath the column heading.
- To delete contents of a cell in the .csv file: just delete the contents as you would using a standard word processor.
- Deleting a row (i.e. a full record) is an advanced option beyond the scope of this document.

Step 2.2.3. Creating new metadata

- Create a new column: first enter a column heading appropriate to your DSpace instance;⁹ once the column is created and headed, you can add data in the cells as needed.
- Create new data in an empty cell: just type your data as you would using a standard word processor.
- Creating a row (i.e. a full record) is an advanced option beyond the scope of this document.

Step 2.3: Save your changes. If using Google Docs, changes will be saved automatically. The filename can be whatever you like.

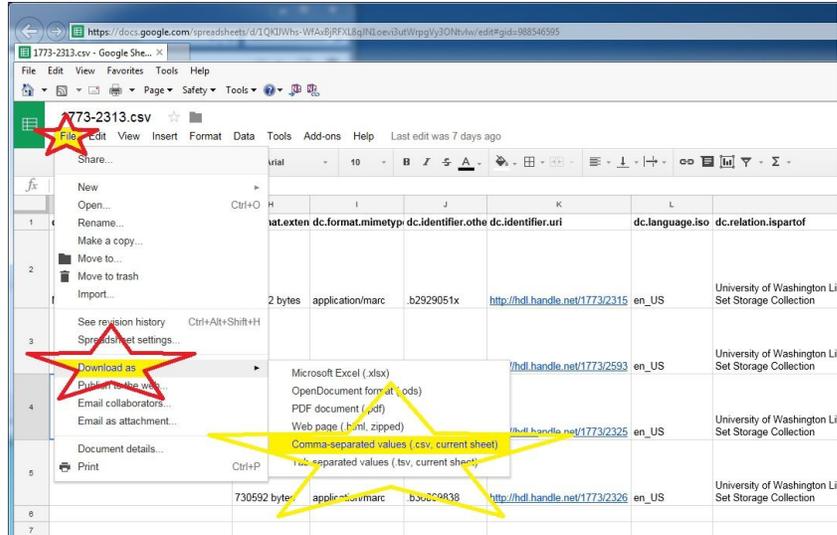
Step 3: Import the edited metadata into DSpace.

Step 3.1: If using Google Docs, download the edited .csv to your local computer:

Step 3.1.2: Select “Comma-separated values”: File menu > Download as > Comma-separated values (.csv, current sheet), as follows:

⁸ **Make sure the first column remains as-exported, headed “id” with all values untouched.** This is the data DSpace uses to make the changes.

⁹ The column heading needs to be a metadata field name that already exists in your DSpace instance. All DSpace instances have metadata “registries” that contain the metadata schemas used by your DSpace instance. To see these schemas, after logging in as an Administrative user, from the right sidebar select “Metadata” from the “Registries” group. The “Metadata registry” page displays, including a list of schemas. Select the schema from the list to display the available field names. Make sure you enter the full metadata field name as your column header, such as “dc.description.abstract.”



Step 3.1.2: Rename the file or accept the filename as downloaded; the filename is not significant. Downloading to your default Downloads folder is fine.

Step 3.2: At the right sidebar select the “Import Metadata” button from the “Administrative” section,¹⁰ as follows:

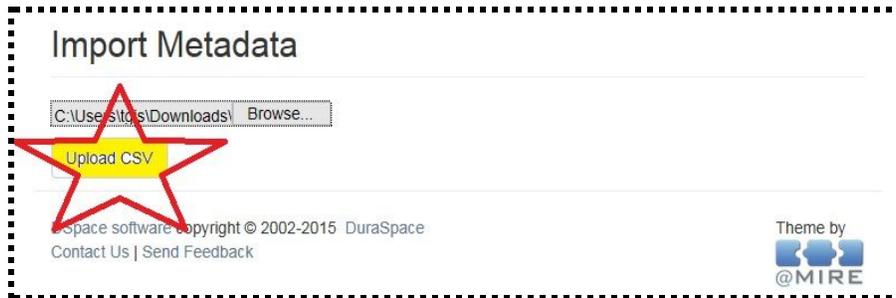


Step 3.2.1: Select “Browse” and navigate to the edited .csv file in the Downloads folder and select the file.

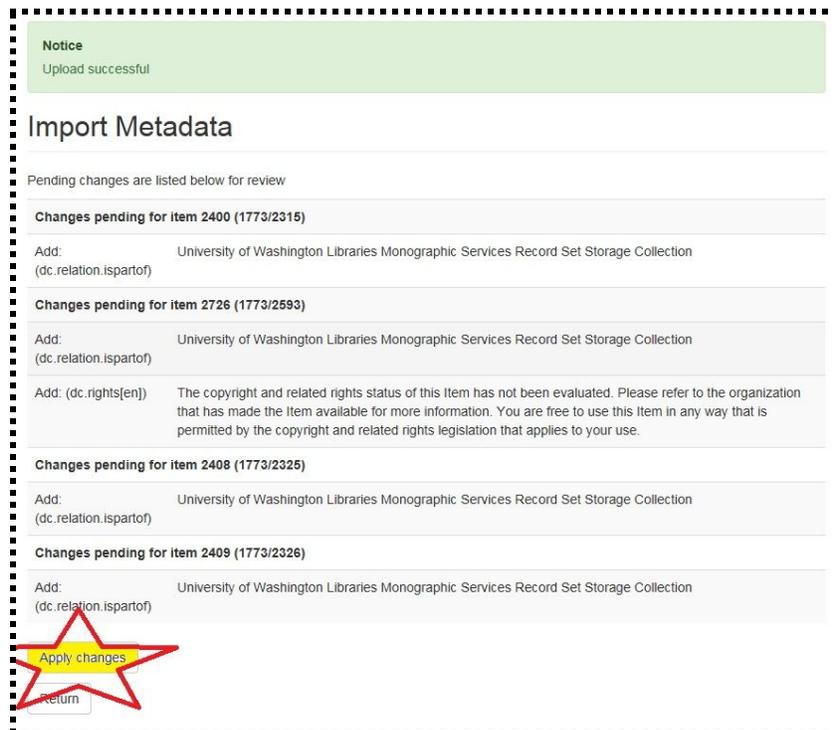
¹⁰ You must be logged into DSpace as an Administrative user to make this selection.



Step 3.2.2: From the resulting display select the “Upload CSV” button, as follows:



Step 3.2.3: “Pending changes are listed below for review” displays. If the pending changes look correct, select “Apply changes” at the bottom of the page (if the changes do not look correct, you can select “Return” to return to the previous screen; the changes will not be applied):



“Notice: Import successful” should then display with the list of changes made.

If the list of changes made looks correct, you are done.

Additional resources

Detailed batch editing instructions for DSpace versions 3 through 6:

- [DSpace 3.x](#)
- [DSpace 4.x](#)
- [DSpace 5.x](#)
- [DSpace 6.x](#)