# Intermediate DSpace: Metadata Imports & Exports

December 21, 2016

Laura Kane McElfresh, Texas A&M University at Galveston

# Intermediate DSpace: Metadata Imports & Exports

Today's training topics include...

Batch imports via Web UI

Exporting Items, Collections, & metadata

Batch edits & metadata cleanup

Moving/mapping Items en masse

Fun with spreadsheets!

2

# Batch Imports Via Web UI (User Interface)

# Why batch import?

Batch importer is good for uploading:

- Large number of Items all at once

- Large Bitstreams

- Anything where you can front-load your metadata and organization

  - Lets you bypass the 3 screens' worth of Web Submit UI...

  - ...but you don't get the 3 screens' worth of guidance.

# Batch importer: right sidebar, under "Content Administration"

## Import Batch Load (ZIP)

### Select a collection

**Collection:**

UT-RGV Training Community > Laura's Demo Collection ▾

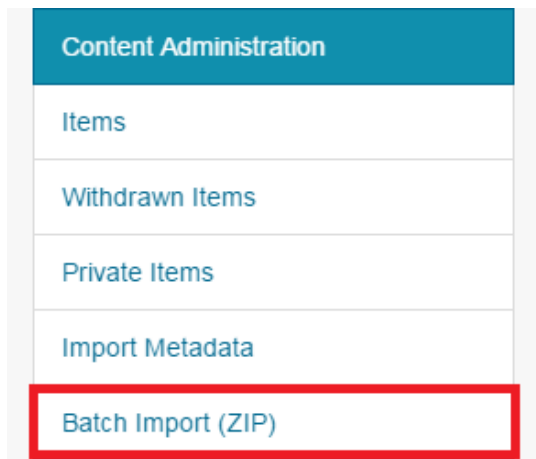Select the collection you wish to submit an item to.

Choose File | No file chosen

**Upload SimpleArchiveFormat ZIP**

DSpace software copyright © 2002-2015  DuraSpace
Contact | Support

Texas Digital Library

# Try it!

Batch upload via web UI



collection from last time and open up the Item "Intermediate DSpace".

2. Click on the sample_collection zip file and save it to your machine.

3. On the right sidebar in DSpace, scroll down to "Content Administration".

   ○ Click on "Batch Import (ZIP)".

4. Use the pulldown menu to select your own collection.

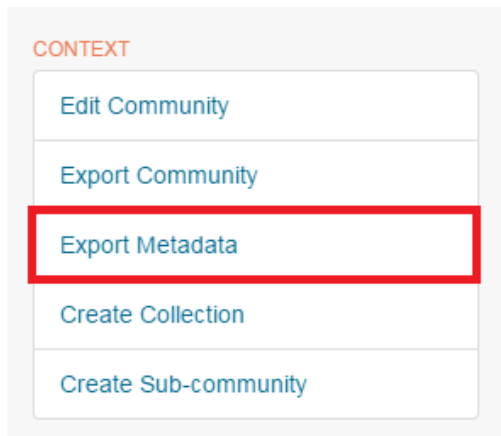   ○ Sorry, they are in some mystery order

5. Click "Choose File" and select the zip file you just downloaded.

6

# Exporting from DSpace

- Items (zip files)

- Communities & Collections (zip files)

- Metadata (comma-separated text, viewable as a spreadsheet)

# Try it!

## Exporting metadata



```
CONTEXT

Edit Community

Export Community

Export Metadata

Create Collection

Create Sub-community
```

Find "Export Metadata" in the sidebar on Community, Collection, or Item pages.

1. Go to your Collection

   - it should have the "Photographs" items you batch loaded

2. Click "Export Metadata"

3. Save the file, which will have a name like 123456789-XXXX.csv

   - The "XXXX" is the number of that collection, community, or item; 123456789 is the repository's number.

4. Double click to open the file -- it should come up in Excel already in columns. 8

# Result of a metadata export: UT-RGV Training Community

| id | collection | dc.acquisi | dc.contrib | dc.contrib | dc.date.is | dc.date.is | dc.descrip | dc.descrip | dc.descrip | dc.descrip | dc.descrip | dc.identif | dc.identif | dc.identif | dc.langua | dc.publish | dc.subject | dc.subject | dc.title.al | dc.title[] | dc.title[er | dc.type[] | dc.type[en_U |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 7030 | 123456789/5256||123456789/5; | | McElfresh, Laura::83; | 2009 | | | Chomp, rc | Graham ai | pictures of dogs play | http://hdl.handle.net/123456789/5263 | | | Dogs | | | | | Snow day! | | Image | |
| | 7032 | 123456789/5256 | | (Unknown)::ee09b1t | 2015 | | | | Pictures of puppies who look li | http://hdl.handle.net/123456789/5265 | | | Dogs | | | | | Not my dogs | | Image | | |
| | 7031 | 123456789/5256 | | Trask, Amy Kane::ba | ######## | | | | Emmie at Christmas in Tulsa | | http://hdl.handle.net/123456789/5264 | | | Cats | | | | | Emmie at Christmas | | Image | | |
| | 7029 | 123456789/5256 | | McElfresh, Laura::83; | 2003-12 | | | | Pictures of Graham, age approx | http://hdl.handle.net/123456789/5262 | | | Cats | | | | | Baby Graham | | Image | | |
| | 7028 | 123456789/5256 | | McElfresh, Laura::83; | 2004-01 | | Bed, crate | In which we see that Graham d | http://hdl.handle.net/123456789/5261 | | | Cats | | | | | Cammi and Graham | | Image | | |
| | 7042 | 123456789/5253 | | McElfresh, Laura::83; | ######## | Training n | Texas Digital Library | | | | | http://hdl en_US | | | DSpace | | | Intermediate DSpac | Learning Obje | |
| | 7024 | 123456789/5253 | Lindsey, Nerissa::a4; | ######## | | | | | | Slides and handout f | http://hdl.ne en_US | | | DSpace | | | Introduction to DSpa | Learning Obje | |
| | 7019 | 123456789/5253 | | McElfresh, Laura::83; | 2016 | | | | | | | http://hdl en_US | | | | | | Repository Structure | Learning Obje | |
| | 7040 | 123456789/5272 | S3E3, Joe::79c678c6-1 | ######## | | This is a bc | S3E3 University | | No particular descrip | http://hdl.handle.ne zh | | S3E3 Press | Genealogy||Computers | | | Technologies for Ani | Preprint | |
| | 7041 | 123456789/5272 | AB, CD::c265e8c1-7b | ######## | | This is an animation about nothing in part | I used RSTU softwar | http://hdl.handle.ne de | | HIJK Press | LMN||OPQ | | | EFG | | Animation | |
| | 7038 | 123456789/5272 | | S3E3::bcbc923f-67f2- | ######## | This is a book written by a robot. | | Originally published as a mono | http://hdl it | | S3E3 Press | Computer Science||Metadata | S3E3's guide to comp | Book | |
| | 7034 | 123456789/5272||123456789/5; | S3E3 Junior::9311b38 | ######## | This is a sc | S3E3 University | | Monograph chapter | | http://hdl other | S3E3 Press | Computer Genealogy | | | Computers for Gene | Book chapter | |
| | 7036 | 123456789/5270 | Ipsum, TV::55dd6577 | 2016 | | | | | | | http://hdl.handle.net/123456789/5274 | | | gibberish | | 1987 | | Working Pape | |
| | 7033 | 123456789/5273 | Smith, Sean William | ######## | | Image of a happy dog at a park located on | This pictu | From the c | http://hdl.handle.net/1234567 | Sean, Inc. | | Research : Happy Dog on a Fall I | Photograph of a Goo | Image | |
| | 7035 | 123456789/5271 | Monkee, Sock::7649( | ######## | | | | | | | http://hdl.handle.net/1234567 | university | | Research Subject Categories::N | D Space training test | Article | |

Column headers are Dublin Core fields (only the ones we used)

Notice use of double pipe || to repeat a field/include multiple values
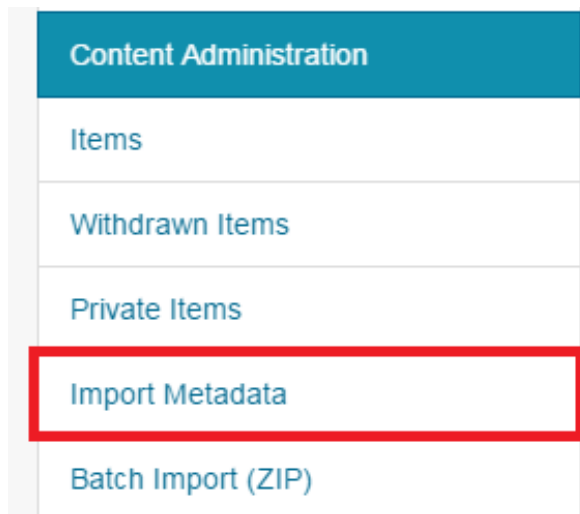
"id" is the item's number

"collection" can have multiple values: owning collection (first listed) &

# Metadata Cleanup & Batch Edits

- Edit metadata fields -- fix typos, standardize CaPiTaLiZaTiOn

- Add new metadata fields

- Remove unnecessary metadata fields (if needed)

- After editing the spreadsheet, upload it back into DSpace ("Import Metadata") and approve the changes.

# Try it!

## Batch metadata edits



1. Look at the spreadsheet you exported earlier. Make any changes you like.

   - Example: the "Photographs of..." Items all have author "O. Nonymous". Try changing it to something else.

2. Save your changes. Be sure to keep the spreadsheet in .csv format.

3. Under "Content Administration" on the right sidebar, click "Import Metadata" and select your .csv file.

4. Review your changes to make sure it's what you wanted.

5. BE SURE TO SCROLL DOWN & CLICK ON "APPLY CHANGES"

11

# Success! (Now scroll down and click "Apply Changes".)

**Notice**
Upload successful

## Import Metadata

Pending changes are listed below for review

### Changes pending for item 7030 (123456789/5263)

| | |
|---|---|
| Add: (dc.description[en_US]) | Graham and Cammi playing outside after a heavy snowfall in OKC |
| Remove: (dc.description[]) | Graham and Cammi playing outside after a heavy snowfall in OKC |
| Remove: (dc.description[en_US]) | pictures of dogs playing in snow |

### Changes pending for item 7032 (123456789/5265)

| | |
|---|---|
| Add: (dc.description[en_US]) | Pictures of puppies who look like mine might have when they were babies |
| Remove: (dc.description[]) | Pictures of puppies who look like mine might have when they were babies |

# ...or fix your metadata and try again.

**Notice**

Unknown metadata element in heading

dc.nonsense

## Import Metadata

No changes were detected

Return

# ...and did I mention, click on "Apply Changes"?

**Notice**
Upload successful

## Import Metadata

Pending changes are listed below for review

**Changes pending for item 7030 (123456789/5263)**

Remove: (dc.date.issued[])

2009

**THIS BUTTON HERE**

**SCROLL DOWN & CLICK IT**

Apply changes

**OTHERWISE DSPACE WON'T**
**SAVE YOUR WORK**

Return

# Moving & Mapping Items

- We move items by changing the Collection number.

- We map items by adding Collection number(s) for the mapped Collection(s), separating the Collection numbers with a double pipe.

  - Leave the owning Collection first in the listing.

15

# Try it!

Moving & Mapping Items

Try mapping some items into "Mapping Collection" within the UT-RGV Training Community.

Choose which items you'd like to map. Under "collection" in the metadata spreadsheet, append `||123456789/5284` to each entry.

Save your spreadsheet & import the metadata.

# You know the drill. ("Apply Changes".)

**Notice**
Upload successful

## Import Metadata

Pending changes are listed below for review

**Changes pending for item 7032 (123456789/5265)**

Map to collection        123456789/5284 (Mapping Collection)

Apply changes

Return

# Spreadsheet-Wrangling Techniques

Tips, tricks, questions

- Delete lines &/or columns you're not editing

    - Especially since DSpace does have a limit on length of import spreadsheets

    - Documentation says 1,000 lines; I thought it was more like 200

- `CTRL+H` (Find/Replace) is your friend!

- Sorting your spreadsheets can help you edit more efficiently

18

# Thank you!

Laura Kane McElfresh

Texas A&M University at Galveston

mcelfrel@tamug.edu