DSpace 5 Release Webinar

DSpace 5 Working Group
Texas Digital Library
December 7, 2015
DSpace 5 Working Group

Susan Elkins, *Angelo State University*

Robert Luttrell, *Texas Tech University*

Alexandra Mitchell, *Texas A&M Galveston*

Laura McElfresh, *Texas A&M Galveston*

Nerissa Lindsey, *Texas A&M International*

Le Yang, *Texas Tech University*

Ryan Steans, *Texas Digital Library*
Overview

• Page Not Found Bug Fix
• Google Analytics from DSpace Admin UI
• Batch Upload via DSpace UI
• MathJax
• ORCID
• SHERPA RoMEO
• Performance Improvement
• Video and Audio
• Additional Features Review
• Final Issues and Timeline
Page Not Found Bug Fix

SUSAN ELKINS
When a user selects a page that is no longer there, a 404 Error is returned. In previous versions of DSpace the page was a simple statement of “Resource not found” that looked like a more serious error. This bug fix now shows “Page cannot be found. Go to DSpace home. Please contact the site administrator if you wish to report this error. If possible, please provide details about what you were doing at the time this error occurred.”
Google Analytics from DSpace Admin UI

LE YANG
Google Analytics from DSpace Admin UI

In each digital item page, the feature “View Google Analytics Statistics” is available at the bottom of the right navigation menu.

The feature pulls statistics of page views from Google Analytics, and display the usage statistics by month.
Google Analytics from DSpace Admin UI

It allows administrators to select time range, but the usage statistics will still be grouped into month for display.

It provides succinct and direct views on usage statistics.
Google Analytics from DSpace Admin UI

The same page of Google Analytics Usage Statistics contains another section called “File Downloads.”

It provides insights of the “true” usage on digital items.

For example, a page gets 200 page views but only 20 file downloads, indicating only the 20 visits are “true” usage of this digital item.

<table>
<thead>
<tr>
<th>Date</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015/11</td>
<td>0</td>
</tr>
<tr>
<td>2015/10</td>
<td>2</td>
</tr>
<tr>
<td>2015/09</td>
<td>0</td>
</tr>
<tr>
<td>2015/08</td>
<td>0</td>
</tr>
<tr>
<td>2015/07</td>
<td>0</td>
</tr>
<tr>
<td>2015/06</td>
<td>0</td>
</tr>
<tr>
<td>2015/05</td>
<td>0</td>
</tr>
<tr>
<td>2015/04</td>
<td>0</td>
</tr>
<tr>
<td>2015/03</td>
<td>0</td>
</tr>
<tr>
<td>2015/02</td>
<td>0</td>
</tr>
</tbody>
</table>
Batch Upload via DSpace UI

LAURA MCELFRESH AND ROBERT LUTTRELL
Batch Upload via DSpace UI

- Used to require command-line actions
- Now: much more user-friendly for those without programming experience
- Still needs file/directory arrangement as in the old command-line process
  - Simple Archive Format
  - Create and arrange your files, zip, and upload
Batch Upload via DSpace UI

Each Item in your batch upload has a folder containing the item’s bitstreams, a metadata file, a license, and contents file.
Batch Upload via DSpace UI

The “contents” file is a tab-delimited text file.

• You can create it in Notepad, then delete .txt extension from the filename.

• If you don’t delete .txt, DSpace will not recognize the file!

• Example: Here’s the contents file for item1 in previous screenshots, a collection of pictures of flowers:

<table>
<thead>
<tr>
<th>license.txt</th>
<th>bundle:LICENSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrysanthemum.jpg</td>
<td>bundle:ORIGINAL description:Chrysanthemum</td>
</tr>
<tr>
<td>Hydrangeas.jpg</td>
<td>bundle:ORIGINAL description:Hydrangeas</td>
</tr>
<tr>
<td>Tulips.jpg</td>
<td>bundle:ORIGINAL description:Tulips</td>
</tr>
</tbody>
</table>
Batch Imports from the User Interface

- DSpace 5.x now has the ability to import items (photos, PDFs, etc.) AND metadata directly from the Graphical User Interface from a DSpace Administrator user from a zip file in the following format:

  Must follow this format:
  
  ```
  archive_directory/
  item_000/
      dublin_core.xml -- qualified Dublin Core metadata for metadata fields belonging to the dc schema
      metadata_[prefix].xml -- metadata in another schema, the prefix is the name of the schema as registered with the metadata registry
      contents -- text file containing one line per filename
      collections -- text file that contains the handles of the collections the item will belong two. Optional. Each handle in a row.
      file_1.doc -- Collection in first line will be the owning collection
      file_2.pdf -- files to be added as bitstreams to the item
      item_001/
          dublin_core.xml
          contents
          file_1.png
          ...
  ```

Please see the documentation for this at:
https://wiki.duraspace.org/display/DSDOC5x/Importing+and+Exporting+Items+via+Simple+Archive+Format
UI Batch Import (XSLUI)

A SimpleArchivePackage can be imported by an administrator in XSLUI. The SimpleArchivePackage needs to be compressed into a ZIP file, and it will be uploaded to XSLUI through the browser.

While logged in as an administrator, click on Batch Import (ZIP).

Content Administration
- Items
- Withdrawn Items
- Private Items
- Import Metadata
- Batch Import (ZIP)

Then, choose the owning collection from the collection dropdown, and browse to the ZIP file on your computer that has the SimpleArchivePackage ZIP file.

Import Batch Load (ZIP)

Select a collection
Collection:

Select the collection you wish to submit an item to.

Choose File ingest_plains_20150105-153342.zip

Upload SimpleArchivePackage ZIP

If successful, you will get a green message with a list of handles that were imported. It is also considered the "map file."

Notice

If an error occurred, you will get a red error message with the issue.
MathJax

LAURA MCELFRESH
MathJax: math and science typesetting

MathJax Javascript facilitates the display of mathematical and scientific formulas in the DSpace XMLUI.

• User must provide the MathJax coding; DSpace does not appear to perform conversions for you (from MathML, TeX, etc.)
• MathJax code is wrapped with dollar signs
MathJax: math and science typesetting

Sources for MathJax markup how-to:


• [http://www.onemathematicalcat.org/MathJaxDocumentation/TeXSyntax.htm](http://www.onemathematicalcat.org/MathJaxDocumentation/TeXSyntax.htm)

• Have not (yet?) found a conversion engine online
Marked-up scientific/mathematical text will display in many fields! For example:

$$\lim_{x \to Donald}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\frac{(x+1) \cancel{(x+2)}}{3 \cancel{(x+2)}}$$

This is how to do MathJax equations using the new DSpace feature documented here https://github.com/DSpace/DSpace/pull/680

THE $1^1B^+_{u} \rightarrow 1^1A^-_{g}$

ABSORPTION OF JET-COOLED TRANS,TRANS-1,3,5,7-OCTATETRAENE

Mathmagics Land: A Travelogue, Vol. $\sum_{n=1}^\infty \frac{1}{2^n}$
MathJax: math and science typesetting

...display as:

\[ \sum_{n=1}^{\infty} \frac{1}{2^n} \]
ORCID

ALEXANDRA MITCHELL
ORCID

- Unique ID that researchers/other users can sign up for
- DSpace 5 release comes with an ORCID lookup option for author and editor, which is supposed to create a dc field so you can look people up by ORCID
- You do not have to use it and can still add authors and editors in the usual way
- If you choose to use the ORCID, make sure that you have the correct person/id
- Clicking the ORCID number will bring up a pop up window so you can view that person’s information
New “Lookup” button

Authors can still be added by typing in the name and clicking “add.”
Lookup screen

On the left are people already in your repository, even if they don’t have an ORCID associated with their name.
Lookup screen

People with ORCID’s associated with the name you type will be displayed in italics.
Lookup screen

This is the screen displayed if you have multiple people with the same name in the ORCID database. Make sure that you are selecting the correct person.
ORCID Lookup

If you need to view a person’s information, you can click on the link in the ORCID pop-up window and view their profile information.
SHERPA/RoMEO

ALEXANDRA MITCHELL
SHERPA/RoMEO

- RoMEO is a database of publishers’ policies “regarding the self-archiving of journal articles on the web and in Open Access repositories”
- You must use the ISSN for the RoMEO policies to appear
- A color coded system is used to identify the publisher’s policies
- If you submit more than 500 requests per day from one IP address, you will need to register for a free account
In order for the RoMEO lookup to work, you must enter the ISSN number of the journal that the article was published in.
SHERPA/RoMEO

The system is color coded to tell you the level of permissions. If the journal isn’t in the RoMEO system, you will get the “submission gray” message.
SHERPA/RoMEO

This is what it looks like when the journal is found by RoMEO. You can click “more info” for details about the color coding system.
When you click “more info” for a journal that has been found by RoMEO, you can get more details about publishing policies.
Performance Improvement

NERISSA LINDSEY
Performance Improvements

Performance improvements for “Select Collection” dropdown in submission process.

• The “select collection” step in the DSpace workflow had a bug that caused it to perform badly when a large number of collections existed

• As per the documentation: https://jira.duraspace.org/browse/DS-682 this bug has been fixed
MP3 embedded player

This feature embeds HTML5 player for audio files in each digital item.

Variety of audio formats are compatible with the player.

Usually the first audio file is displayed at the view of “Simple Item Record.”
MP3 embedded player

At the view of “Full Item Record,” the audio files will be positioned in line at the bottom with some simple metadata.

Description with the bitstreams can be provided to identify each file.

Click “View/Open” to download the file.
wav embedded player
Video Player

• The programming team at TDL has developed a HTML5 video player for DSpace 5.0
• It has all the essential features for a video player: sound adjustment, full-screen capability, etc.
Video and Audio Player Allowed Formats

HTTP 206 + HTML5 audio/video
+HTTP 206 implementation for BitstreamReader -- needed for html5 video/audio players

+Adds html5 player to xmlui/Mirage2 for video files with certain file extensions:
  .3g2, .3gp, .asf, .avi, .m4v, .mov, .qt, .mp4, .m4p, .mpg, .mpeg, .m2v, .vob -- as <video> with <source src=xxx type=video/mp4>
  .ogg, .ogv -- as <video> with <source src=xxx type=video/ogg>
  .webm -- as <video> with <source src=xxx type=video/webm>
  .drc, .flv, .mkv, .mng, .mp2, .mpe, .mpv, .mxf, .nsv, .rm, .rmvb, .roq, .svi, .wmv, .yuv -- as <video> with <source src=xxx> with no type

+Adds html5 player to xmlui/Mirage2 for audio files with certain file extensions:
  .weba -- as <audio> with <source src=xxx type=audio/webm>
  .mp3 -- as <audio> with <source src=xxx type=audio/mpeg>
  .oga, .opus -- as <audio> with <source src=xxx type=audio/ogg>
  .wav, .wave -- as <audio> with <source src=xxx type=audio/wav>
  .act, .aif, .aiff, .aac, .amr, .ape, .au, .awb, .dct, .dss, .dvf, .flac, .gsm, .iklax, .ivs, .m4a, .mmf, .mpc, .msv, .ra, .raw, .sln, .tta, .vox, .wma, .wv -- as <audio> with <source src=xxx> with no type
Uploading Video

- The player supports all file types that HTML5 supports, but .avi, .mov, .mp4, .asf, and .wmv are recommended.
- It is also recommended that the uploaded videos be around 15 frames per second, at anywhere between 320x240 to 1080x720 resolution.
- Batch uploading videos is possible with DSpace 5, but depending on file sizes it could be a lengthy process (schedule jobs accordingly).
Minor Issues

• The video player has issues supporting all MIME types at the moment
• For videos with multiple parts, only the first part will load in the video player, subsequent parts will open in an installed video player on the computer
• The programming team is planning enhancements to the video player later this year
Additional Features Review

ROBERT LUTTRELL
PDF Thumbnail Generation

- PDF thumbnails are now generated from the first page of the first PDF file in the item upload. PDFs are generated in the nightly batch job when the DSpace media filter process is run. This process also generates the OCR bitstreams, thumbnails of regular images, and updates the search index.

- Thumbnails for primary files will now be “clickable”, generating a download of the primary file
PDF Thumbnail Generation
The OAI-PMH Interface may be used by other systems to harvest metadata records from your DSpace. There have been several bug fixes (including the QDC list records issue seen in DSpace 4.x) to the OAI data so they should now following the correct OAI standard.

Here is the link for the Dspace 5.x Documentation on OAI PMH Harvesting
https://wiki.duraspace.org/display/DSDOC5x/OAI

To View your OAI data from your Dspace server input the following URL into the browser, specifying your specific Dspace server information:
https://YOURSERVER.tdl.org/YOURSERVER-oai/request?verb=Identify

Example:
https://dspace-test.tdl.org/test-oai/request?verb=Identify

To see if you server is validated on the OAI standards goto:
http://validator.oaipmh.com/

Put in your OAI address for validation, and run the validation check ie:
https://dspace-test.tdl.org/test-oai/request

This version of Dspace has been updated for the European Openaire standard as well.. TDL has verified our test server has passed their verification server.
**Repository Information**

<table>
<thead>
<tr>
<th><strong>Repository Name</strong></th>
<th>DSpace at test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E-Mail Contact</strong></td>
<td><a href="mailto:ldl-systems@utlists.utexas.edu">ldl-systems@utlists.utexas.edu</a></td>
</tr>
<tr>
<td><strong>Repository identifier</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Sample identifier</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Protocol Version</strong></td>
<td>2.0</td>
</tr>
<tr>
<td><strong>Earliest Registered Date</strong></td>
<td>2015-10-06 23:47:38</td>
</tr>
<tr>
<td><strong>Date Granularity</strong></td>
<td>YYYY-MM-DD hh:mm:ss</td>
</tr>
<tr>
<td><strong>Deletion Mode</strong></td>
<td>transient</td>
</tr>
</tbody>
</table>
Final Issues and Timeline

RYAN STEANS
EZID for DOI’s

- Ongoing work to integrate EZID/DOI into DSpace for TDL members
- DSpace 5 will support EZID integration for implementation of DOI’s for each item
- TDL consortium members will now receive a discount on EZID memberships
- DOI’s will be tied to handles, but fully functional
Timelines and Upgrade procedure

• TDL DSpace 5 Working Group will be upgraded first
• No major changes to DSpace, just new features and bug fixes
• We will test “Upgrade in Place” with Working Group members week of December 7th
• Webinar redux in early January
• Upgrade remaining DSpaces in mid-January
Questions and Discussion