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



Page Not Found Bug Fix

Resource not found Go to Cadmus home org.apache.cocoon.ResourceNotFoundException: Page cannot be found Cocoon stacktrace [show] Java stacktrace [show] Java full stacktrace [show] The Manakin / DSpace digital repository software.

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. Contact site administrator || Show underlying error stack DSpace software copyright © 2002-2013 Duraspace Contact | Support

Old version

New version

exas Digital Library

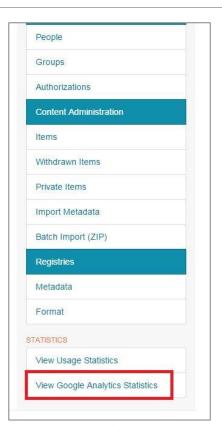
When a user selects a page that is no longer there, a 404 Error is returned. In previous versions of DSpace the page was a simple statement of "Resource not found" that looked like a more serious error. This bug fix now shows "Page cannot be found. Go to DSpace home. Please contact the site administrator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

LE YANG



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

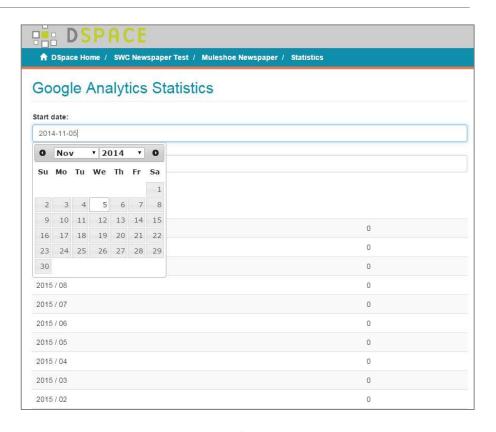
The feature pulls statistics of page views from Google Analytics, and display the usage statistics by month.





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.





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.

2015 / 11	0
2015 / 10	2
2015 / 09	0
2015 / 08	0
2015 / 07	0
2015 / 06	0
2015 / 05	0
2015 / 04	0
2015 / 03	0



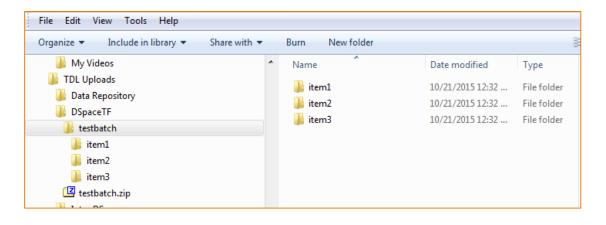
LAURA MCELFRESH AND ROBERT LUTTRELL

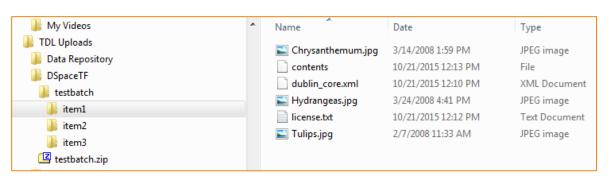


- Used to require command-line actions
- Now: much more user-friendly for those without programming experience
- Still needs file/directory arrangement as in the old command-line process
 - Simple Archive Format
 - http://www.tdl.org/wp-content/uploads/2009/04/DSpaceBatchImportFormat.pdf<-instructions for SAF packaging
 - Create and arrange your files, zip, and upload



Each Item in your batch upload has a folder containing the item's bitstreams, a metadata file, a license, and contents file.







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:

license.txt bundle:LICENSE

Chrysanthemum.jpg bundle:ORIGINAL description:Chrysanthemum

Hydrangeas.jpg bundle:ORIGINAL description:Hydrangeas

Tulips.jpg bundle:ORIGINAL description:Tulips



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:

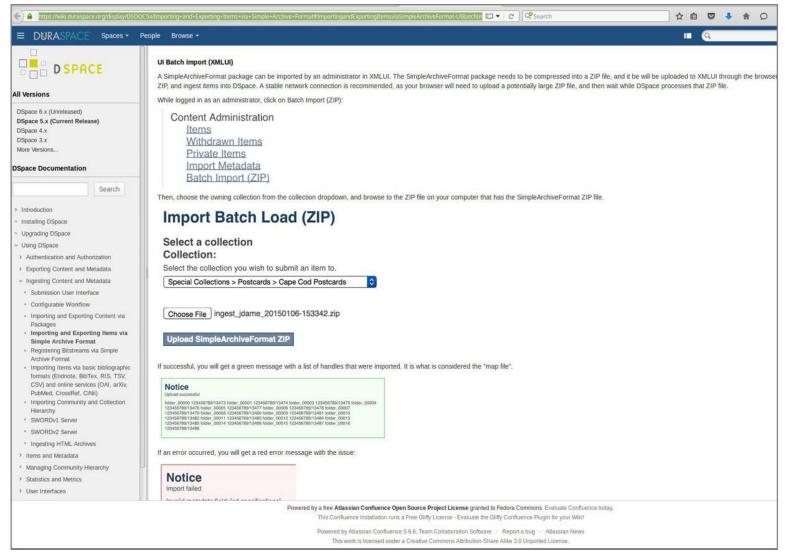
```
Must follow this format:
```

```
archive directory/
  item 000/
                         -- qualified Dublin Core metadata for metadata fields belonging to the dc schema
    dublin core.xml
    metadata [prefix].xml -- metadata in another schema, the prefix is the name of the schema as registered with the metadata registry
                      -- text file containing one line per filename
    contents
                      -- text file that contains the handles of the collections the item will belong two. Optional. Each handle in a row.
    collections
                  -- Collection in first line will be the owning collection
                      -- files to be added as bitstreams to the item
    file 1.doc
    file 2.pdf
  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







MathJax

LAURA MCELFRESH



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



Sources for MathJax markup how-to:

- http://meta.math.stackexchange.com/questions/5020/mathjax-basic-tutorial-and-quick-reference
- http://www.onemathematicalcat.org/MathJaxDocumentation/TeXSyntax.
- Have not (yet?) found a conversion engine online



Marked-up scientific/mathematical text will display in many fields! For example:

dc.author	Duck, \$ \lim_{x\to Donald} \$
dc.description.abstract	\$ x = \frac {-b \pm \sqrt {b^2 - 4ac}} {2a} \$
	<pre>\$ \require{cancel} \frac{(x+1) \cancel{(x+2)}}{3 \cancel{(x+2)}} \$</pre>
	This is how to do MathJax equations using the new DSpace feature documented here https://github.com/DSpace/DSpace/pull/680
	THE \$1 ^{1}B^{+}_{u} \leftarrow 1 ^{1}A^{-}_{g}\$ ABSORPTION OF JET-COOLED TRANS,TRANS-1,3,5,7- OCTATETRAENE
dc.title	Mathmagics Land: A Travelogue, Vol. \$ \sum_{n=1}^\infty \frac{1}{2^n} \$



...display as:





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

 Texas Digital Library

New "Lookup" button

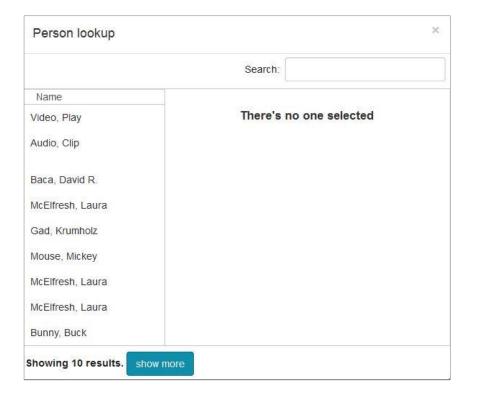
Authors can still be added by typing in the name and clicking "add."





Lookup screen

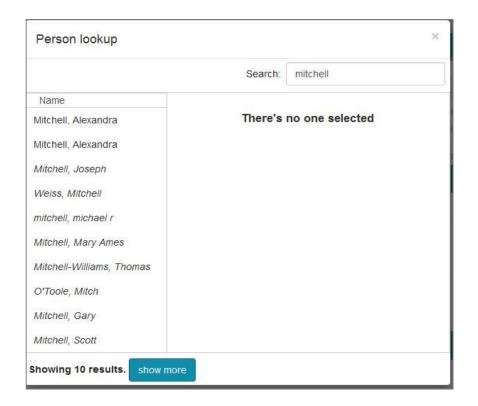
On the left are people already in your repository, even if they don't have an ORCID associated with their name.





Lookup screen

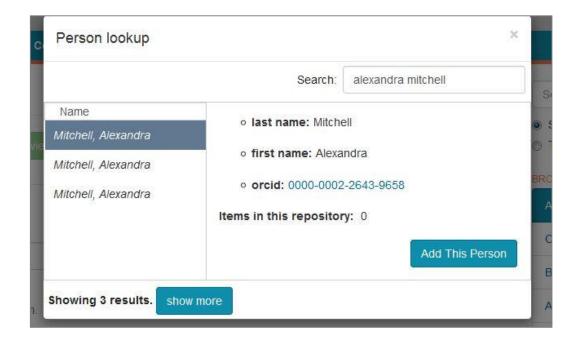
People with ORCID's associated with the name you type will be displayed in italics.





Lookup screen

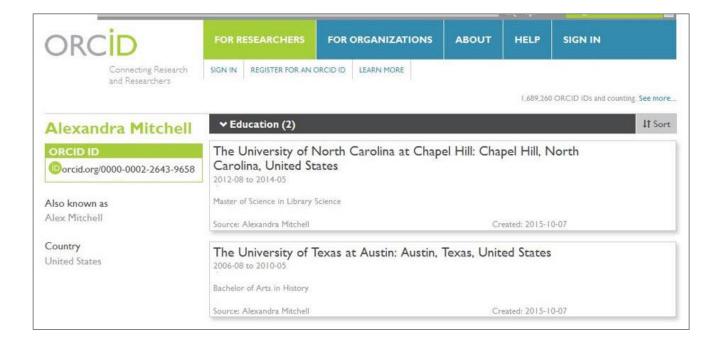
This is the screen displayed if you have multiple people with the same name in the ORCID database. Make sure that you are selecting the correct person.





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.





ALEXANDRA MITCHELL



- RoMEO is a database of publishers' policies "regarding the selfarchiving 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

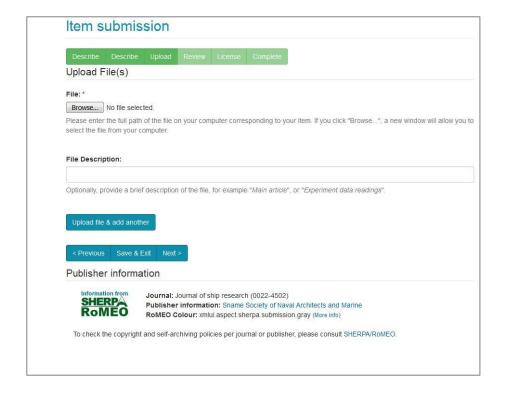


In order for the RoMEO lookup to work, you must enter the ISSN number of the journal that the article was published in.



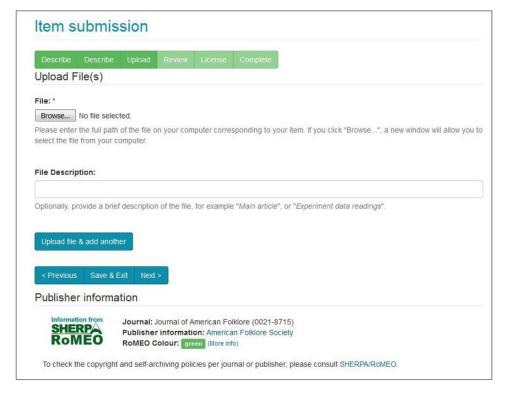


The system is color coded to tell you the level of permissions. If the journal isn't in the RoMEO system, you will get the "submission gray" message.





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



Video and Audio

NERISSA LINDSEY, ROBERT LUTTRELL AND LE YANG



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."



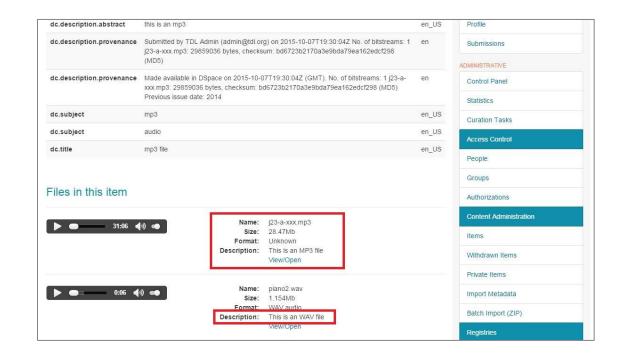


MP3 embedded player

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.





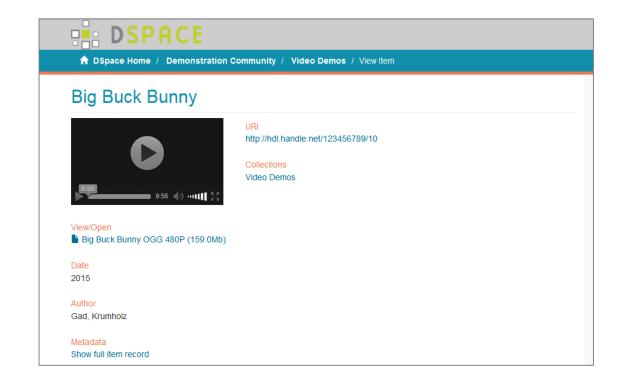
wav embedded player





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, fullscreen 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, .m4v, .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, .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
```

xas Digital Library

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







OAI-PMH and Openaire

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 imput the following URL into the browser, specifying your specific Dspace server information: https://YOURSERVER-oai/request?verb=Identify

Example:

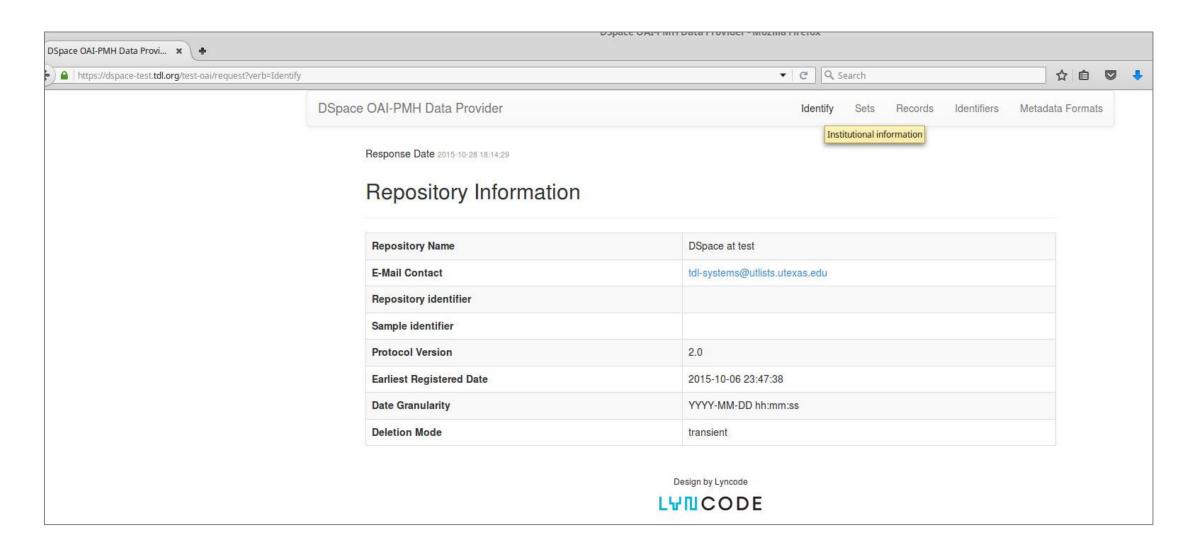
https://dspace-test.tdl.org/test-oai/request?verb=Identify

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 out test server has passed their verification server.







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



