Introducing DSpace 3.x

Hardy Pottinger, University of Missouri Library Systems
@hardypottinger
http://lso.umsystem.edu/~pottingerhj/
7/11/2013, OR-2013, Prince Edward Island, Canada

This work is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

Slides available at: http://tinyurl.com/ds3-introducing
Missouri, which coast is that?

- Neither, it’s in the middle.
1.7, 1.8... Wait a minute, what happened to DSpace 2.0?

- The version numbering scheme changed after 1.8
- Release numbers will now only consist of two numbers: \([\text{major}].[\text{minor}]\) (e.g. 3.0, 3.1, 4.0)
- MAJOR=new features
- MINOR=bug fixes (only)
- Details: [http://tinyurl.com/dspace-relno](http://tinyurl.com/dspace-relno)
A moment of silence for DSpace 2.0
Introducing Dspace 3.x, overview

- How did we get to Dspace 3.x?
- Notable Features of Dspace 3.0
- DSpace 3.1 bug fixes
- Dspace 3.2?
How did we get to DSpace 3.x?

Lots of hard work from these folks:

- DCAT (DSpace Community Advisory Team)
- DSpace Committers
- The Dspace 3.x Release Team (RT)
  - Sands Fish (MIT) – RT Leader
  - Hardy Pottinger (University of Missouri)
  - Robin Taylor (University of Edinburgh)
  - Ivan Masár (aka Helix84)
- Contributions from the community (you!)

http://tinyurl.com/dspace-dcat
Notable Features of DSpace 3.0

- New OAI-PMH interface
- Improvements to Solr-based statistics
- Batch import for bibliographic formats (Endnote, BibTex, etc.)
- Type-based submissions
- Enhanced browse and Solr provider (Discovery works for JSPUI)
- Advanced embargo support
- Item-level versioning
- ElasticSearch-based usage statistics
- A new mobile theme
- Easier environment property management with build.properties
Included automatically

- What you get “out of the box”
A new OAI-PMH interface

Contributed by Lyncode (committer: João Melo)

- Driver and Open-AIRE compatible
- Allows for multiple contexts (URL endpoints), each with a different configuration
- 12 default metadata export formats and easy way to write new ones using XSLT
- Runs on Solr for great performance
- Legacy mode over DSpace database supported
- Even faster thanks to caching
Improved Solr-based statistics

Contributed by @mire

- Workflow statistics
- Search Query statistics
- Solr version upgrade and performance optimization
## Search Statistics

### Top Search Terms

<table>
<thead>
<tr>
<th>Search Term</th>
<th>Searches</th>
<th>% of Total</th>
<th>Pageviews / Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>author_keyword: Deininger, Klaus</td>
<td>23</td>
<td>16.55%</td>
<td>0.00</td>
</tr>
<tr>
<td>author_keyword: Ali, Daniel Ayalew</td>
<td>22</td>
<td>15.83%</td>
<td>0.41</td>
</tr>
<tr>
<td>modeling</td>
<td>10</td>
<td>7.19%</td>
<td>0.10</td>
</tr>
<tr>
<td>subject_keyword: Energy</td>
<td>10</td>
<td>7.19%</td>
<td>0.00</td>
</tr>
<tr>
<td>subject_keyword: Environment</td>
<td>9</td>
<td>6.47%</td>
<td>0.00</td>
</tr>
<tr>
<td>topic_keyword: Health</td>
<td>9</td>
<td>6.47%</td>
<td>0.00</td>
</tr>
<tr>
<td>author_keyword: World Bank</td>
<td>8</td>
<td>5.78%</td>
<td>0.00</td>
</tr>
<tr>
<td>economic</td>
<td>8</td>
<td>5.78%</td>
<td>0.00</td>
</tr>
<tr>
<td>subject_keyword: Natural Resources</td>
<td>8</td>
<td>5.78%</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### Total

<table>
<thead>
<tr>
<th>Searches</th>
<th>% of Total</th>
<th>Pageviews / Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>139</td>
<td>100.00%</td>
<td>0.12</td>
</tr>
</tbody>
</table>
## Workflow Statistics

### Overall Statistics

<table>
<thead>
<tr>
<th>Step</th>
<th>Performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Accept/Reject/Edit Metadata Step Pool</td>
</tr>
<tr>
<td>2</td>
<td>Accept/Reject/Edit Metadata Step</td>
</tr>
<tr>
<td>3</td>
<td>Edit Metadata Step Pool</td>
</tr>
<tr>
<td>4</td>
<td>Accept/Reject Step Pool</td>
</tr>
<tr>
<td>5</td>
<td>Accept/Reject Step</td>
</tr>
<tr>
<td>6</td>
<td><strong>Score Review Pool</strong></td>
</tr>
<tr>
<td>7</td>
<td>Score Review</td>
</tr>
<tr>
<td>8</td>
<td>Single User Review Pool</td>
</tr>
<tr>
<td>9</td>
<td>Edit Metadata Step</td>
</tr>
<tr>
<td>10</td>
<td>Score Review Evaluation</td>
</tr>
</tbody>
</table>
Batch import for Bibliographic formats

contributed by Greek National Documentation Centre/EKT (our newest committer: Kostas Stamatis)

- Support for Endnote, BibTex, RIS, TSV, CSV
- Enhanced batch import routines
Controlled vocabulary support for XMLUI

contributed by @mire (committer: Kevin Van de Velde)

- Submission form vocabulary lookup
- Includes The Norwegian Science Index and the Swedish Research Subject Categories
Controlled vocabulary support for XMLUI
Google Analytics support for JSPUI

contributed by Denys Slipetskyy

- Support for statistics collection by entering the GA key into dspace.cfg
Salted passwords in simple password authentication

contributed by Mark H. Wood, with the support of IUPUI University Library

- Now stores salted hashes (which is a best practice for password storage)
- Old passwords will continue to work and will be automatically converted to salted hashes on next user login
Some configuration required

- Comes in the box, but you’ll have to fiddle with it
Discovery: faceted Search & Browse

- Enhancements for XMLUI (contributed by @mire, with the support of the World Bank)
  - Search snippets
  - Hit highlighting
  - Related items
  - Hiding restricted results
- Supported in JSPUI (contributed by CILEA, now CINECA)
Browse Enhancement & Solr provider

- (contributed by CILEA, now CINECA, and EKT)
- Browse system can rely on Solr index, and not the DB
- Can indicate number of items in a set (browse by author, subject) in both XMLUI and JSPUI (DS-1223)
- in the release notes, we expressed this as “Discovery works in the JSPUI”, it’s actually much more. Community development at it’s best, makes attribution difficult. So I will just say thanks, to:
  - Andrea Bolini (CINECA), Christina Paschou (EKT) and Kostas Stamatis (EKT)
Frequencies of Search Terms displayed in Discovery, in both JSPUI and XMLUI
Frequencies of Search Terms displayed in Discovery, in both JSPUI and XMLUI
Item-level versioning

contributed by @mire, with the support of MBLWHOI Library, WHOL, MBL CLI(HPS), ASU CBS and Dryad

- Create and preserve different item versions
- Enhanced Identifiers
Advanced Embargo

contributed by @mire with the support of the University of Michigan Libraries

- time-based restrictions on both bitstreams and metadata
- advanced-mode for additional user group restrictions
Embargo: simple form

Item submission

Access Settings

Private Item:
If selected, the item won't be searchable

Embargo Access until Specific Date:
Accepted format: mm/dd/yyyy, mm/yyyy, yyyy

Reason:

Complete
Embargo: advanced form

Policies List

<table>
<thead>
<tr>
<th>Name</th>
<th>Action</th>
<th>Group</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embargo Policy</td>
<td>READ</td>
<td>Anonymous</td>
<td>2013-01-01</td>
<td></td>
</tr>
</tbody>
</table>

Edit Bitstream Access

Name:

Groups:

Anonymous

Visible/Embargoed:

- Item will be visible once accepted into archive
- Embargo Access until Specific Date

Accepted format: mm/dd/yyyy, mm/yyyy, yyyy

Reason:


Embargo: upload with Embargo step

**Item submission**

**Upload File(s)**

**File:**
Please enter the full path of the file on your computer corresponding to your item. If you click "Browse...", a new window will allow you to select the file from your computer.

**File Description:**
Optionally, provide a brief description of the file, for example "Main article", or "Experiment data readings".

**Embargo Access until Specific Date:**
Accepted format: mm/dd/yyyy, mm/yyyy, yyyy

01/01/2013

**Reason:**
embargoed through 2012

Upload file & add another
Mobile theme for XMLUI

contributed by Elias Tzoc and James Russell with the support of Miami University

- technology preview
Mobile theme for XMLUI
Type-based submissions

contributed by Nestor Oviedo and SeDiCl

- show/hide metadata fields in the submission forms, based on the type of content submitted

```xml
<field>
  <dc-schema>dc</dc-schema>
  <dc-element>identifier</dc-element>
  <dc-qualifier>isbn</dc-qualifier>
  <label>ISBN</label>
  <type-bind>thesis,ebook</type-bind>
</field>
```
ElasticSearch-based usage statistics

contributed by Peter Dietz with the support of Ohio State University Libraries

- scalable ElasticSearch backend, runs on embedded node by default
- uses Google Chart API for graphs and maps
- export to CSV available
- displaying can be either public or restricted
Statistical Report for Sample Community
Showing Data (Last Five Years)

Number of File Downloads for Sample Community
Data Range: Last Five Years For more information.

Number of File Downloads: Sample Community
### Statistical Report for Sample Community

**Report Generator**

Used to generate reports with an arbitrary date range.

Showing Data ( All Data Available )

**Back to Summary Statistics for Sample Community**

**Top Downloads (all time)**

<table>
<thead>
<tr>
<th>Title</th>
<th>Creator</th>
<th>Publisher</th>
<th>Date</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Webpage</td>
<td>Cat, Lily</td>
<td></td>
<td>1982-11-11</td>
<td>18</td>
</tr>
<tr>
<td>Test Word Document</td>
<td>Smith, Samuel G</td>
<td></td>
<td>2012-06-26</td>
<td>14</td>
</tr>
<tr>
<td>Test PDF Document</td>
<td>Nelson, Sandra K</td>
<td></td>
<td>2002-06-26</td>
<td>12</td>
</tr>
<tr>
<td>Test Webpage</td>
<td>Cat, Lily</td>
<td></td>
<td>1982-11-11</td>
<td>9</td>
</tr>
<tr>
<td>aa</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Test PDF Document</td>
<td>Nelson, Sandra K</td>
<td></td>
<td>2002-06-26</td>
<td>8</td>
</tr>
<tr>
<td>Test PowerPoint Document</td>
<td>Doe, Jane I</td>
<td></td>
<td>1950-06-26</td>
<td>6</td>
</tr>
</tbody>
</table>
### Countries with most Downloads Sample Community
Data Range: Last Five Years

### Top Downloads for June 2013
Data Range: Last Five Years

<table>
<thead>
<tr>
<th>Title</th>
<th>Creator</th>
<th>Publisher</th>
<th>Date</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Webpage</td>
<td>Cat. Lily</td>
<td></td>
<td>1982-11-11</td>
<td>15</td>
</tr>
</tbody>
</table>
Improved LDAP authentication

contributed by Samuel Ottenhoff

- Hierarchical LDAP
- you can now map LDAP group membership to internal DSpace groups
Easier environment property management with build.properties

contributed by Steve Swinsburg, Australian National University, tweaked by Tim Donohue, DuraSpace

- A straightforward .properties file-based system for managing different settings for dev, staging and production environments
- No more fiddling with Maven .profiles or .m2/settings.xml
- Just drag one properties file into a new clone of Dspace, and you are in business!
- I won’t live demo this, but it was really tempting
- More details: http://tinyurl.com/ds3-build-properties
Bug fixes in DSpace 3.1

Full details: http://tinyurl.com/ds31-release-notes

- Oracle compatibility fix
  - big thanks to Qamar Abbas, Artur Konczak, Andrea Schweer and Ivan Masár
  - If your repository runs on Oracle, buy these people beer!
  - Oracle + DSpace users, please introduce yourself to me

- OAI-PMH interface fixes (João Melo)

- Fixed Discovery "More like this" feature (Kevin Van de Velde)

- Type-based submission ported to JSPUI (Ivan Masár)
DSpace 3.2?

- A minor release, bug fixes only
- What bugs?
  - OAI bug fixes
  - A security fix for JSPUI
- When? Shortly after OR13
- All fixes already exist, we just need to release them
- Security fix has been backported to DSpace 1.7 and 1.8
- Security fix will also be made available in patch form, to maximize your options to address the issue
Lessons learned

from institutions who are using 3.x in production repositories

- You will need to tweak your servlet container’s JAVA_OPTS
  [http://tinyurl.com/ds3-tomcat-tweaks](http://tinyurl.com/ds3-tomcat-tweaks)
- Be careful with debug logging (don’t leave it on by accident)
  see DS-1565 and DS-1564 for more info
- Expect to spend some time updating your themes, if you use
  XMLUI, use the new base themes, either dri2-xhtml-alt (aka
  Mirage) or dri2-xhtml, or…
  - Any theme can function as a “base theme” so best practice is to…
  - Base your theme on a standard theme and customize just the pieces
    you need to customize
More lessons learned

- Only deploy what you need (limit your work in progress)
- Here is what we are saving for a future iteration:
  - Discovery
  - ElasticSearch stats
  - Mobile theme
Image Credits

- “Map of the United States” by Jared Benedict
- “Dizzy” by Matt Brooks, from The Noun Project
- “Grave” by Alex Fuller from The Noun Project
- “Mesh Network” by Lance Weisser from The Noun Project
- “Box” by Dmitry Baranovskiy from The Noun Project
- “Wrench” by John Caserta from The Noun Project
- “Three stages in descent of dropping yo-yo” by Brews Ohare
- “Spray Can” by Luke Anthony Firth from The Noun Project
- “Post It” by Joe Harrison from The Noun Project
Thank you!

Hardy Pottinger, University of Missouri Library Systems
@hardypottinger
http://lso.umsystem.edu/~pottingerhj/