DSpace Administration and Use

July 23rd, 2008
Baylor University

Scott Phillips
Research and Development
scott@library.tamu.edu

Jay Koenig
Metadata Librarian
jkoenig@tamu.edu
Schedule

9:00   - Introduction to DSpace
9:30   - Users, Groups, and Authorizations
10:00  - Hands-on: Creating Collections
10:30  - Item Submission and Workflow
11:00  - Hands-on: Submitting an Item

Lunch, 11:30 - 1:00

1:00   - Metadata Registries and Templates
1:30   - Administrating DSpace
2:00   - Configuration Options
2:30   - Batch Imports
Introduction to DSpace
What is DSpace?

✦ **Captures**
  ✦ Digital research material in any formats
  ✦ Directly from creators (faculty)
  ✦ Large-scale, stable, managed long-term storage

✦ **Describes**
  ✦ Descriptive, technical, rights metadata
  ✦ Persistent identifiers

✦ **Distributes**
  ✦ Via WWW, with necessary access control

✦ **Preserves**
  ✦ Bitstream guaranteed
The DSpace Community

- ~250 registered live sites
  - World-wide adoption
  - >1m digital assets and growing fast, largest sites several hundred thousand items

- Profile
  - Primarily research and higher education institutions
  - Cultural heritage organizations, state libraries/archives
  - Some commercial users and service providers
  - Active development community
How to get “connected”

✦ http://www.dspace.org/
 ✦ Documentation
 ✦ NewSpace Newsletter

✦ Mailing Lists
 ✦ dspace-tech
 ✦ dspace-dev

✦ http://wiki.dspace.org/
 ✦ Technical guides
 ✦ On-going projects
DSpace Data Model

- Community
- Collection
- Item
  - Bundle
    - Bitstream
  - Bundle
    - Bitstream
  - Metadata
    (Dublin Core)
What Metadata?

✦ **Descriptive Metadata**
  ✦ Qualified Dublin Core
  ✦ Limited expansion to other formats

✦ **Administrative Metadata**
  ✦ Internal access control

✦ **Structural Metadata**
  ✦ Bundles & bitstreams
Handles

✧ **Persistent identifier**
  ✧ Communities
  ✧ Collections
  ✧ Items

✧ **Format:**

  hdl:1721.123/4567
  http://hdl.handle.net/1721.123/4567

  ✧ prefix
  ✧ identifier
Communities and Collections

- **Communities**
  - Can contain sub-communities or collections
- **Collections**
  - Can contain items
- **Items**
  - Contains metadata and bitstreams (files)
Example

- Community
  - Sub-community
    - Sub-community
    - Collection
    - Items
  - Collection
    - Items
Example

- University (Community)
  - College (Sub-community)
    - Department (Sub-community)
      - Faculty Member (Collection)
        - Papers (Items)
    - Center (Collection)
      - Technical Reports (Items)
  - Historical Images (Collection)
    - Images (Items)
Users, Groups, and Authorizations
Users and Groups

✦ Users - individuals that have a role in the repository
✦ Groups - groups of users that share roles

✦ Special Groups
  ✦ “Administrator” - Group of repository administrators
  ✦ “Anonymous” - Anyone
User Authorizations

- **Bitstream**
  - READ - can open the file
  - WRITE - can alter the file

- **Bundle**
  - ADD/REMOVE - can add bitstreams to a bundle

- **Item**
  - READ - can view the item
  - WRITE - can modify the item
  - ADD/REMOVE - can add or remove bitstreams

- **Collection**
  - ADD/REMOVE - can add or remove items from the collection
  - DEFAULT_ITEM_READ - new items receive this READ attribute
  - DEFAULT_BITSTREAM_READ - new bitstreams receive this READ authorization
  - COLLECTION_ADMIN - can edit or withdraw items, or map items into the collection

- **Community**
  - ADD/REMOVE - can add or remove collections from the community
Workflow Steps

✦ No Workflow Steps
  ✦ Item is made available upon submission

✦ Workflow Step 1
  ✦ Administrator can accept or reject a submission

✦ Workflow Step 3
  ✦ Administrator can edit metadata before making item available

✦ Workflow Step 2
  ✦ Combination of 1 and 3
Workflow Steps

1. Unfinished Submission
2. Workflow Step 1
3. Workflow Step 2
4. Workflow Step 3
5. Available in DSpace

Steps:
- Submit
- Accept
- Reject
- Edit Metadata

Flow:
- Unfinished Submission → Workflow Step 1
- Workflow Step 1 → Workflow Step 2
- Workflow Step 2 → Workflow Step 3
- Workflow Step 3 → Available in DSpace
- Workflow Step 2 → Edit Metadata
- Workflow Step 3 → Edit Metadata

Conditions:
- Submit
- Accept
- Reject
- Edit Metadata
Hands-on: Creating Collections
Submission & Approval

**Submission** (web-based, single item)
- Metadata entry
- File upload
- License agreement

**Approval**
- Depends on the activated workflow step
- *Accept/Reject*
  - Accept puts the item into public view
  - Reject sends the workflow back to the submitter
- *Edit Metadata*
  - Once metadata is edited, item goes into public view
Item Submission and workflow
Hands-on: Submitting an Item
Break for Lunch

9:00   - Introduction to DSpace
9:30   - Users, Groups, and Authorizations
10:00  - Hands-on: Creating Collections
10:30  - Item Submission and Workflow
11:00  - Hands-on: Submitting an Item

Lunch, 11:30 - 1:00

1:00   - Metadata Registries and Templates
1:30   - Administrating DSpace
2:00   - Configuration Options
2:30   - Batch Imports
Metadata Registries and Templates
Metadata Registry and Templates

- Metadata Registry
  - Defines metadata fields
  - Add new fields
- Item template
  - Set default values for metadata fields
  - Affects all new submissions, does not change metadata for existing items
Item Mapper

- Items can appear in multiple collections
- Must be mapped from the destination collection
- Mapped items appear in the second collection but do not exist in the second collection
Administrating DSpace
Removing Items

✦ **Withdraw**
  - Removes item from view
  - Does not show up in search results
  - Recoverable

✦ **Permanently Delete**
  - Unrecoverable
  - Handle is not reused
  - Can only be done by a repository administrator
External Linking

✦ **Repository item**
  ✦ Handle
  ✦ http://repository.tamu.edu/handle/1969.1/6885

✦ **Bitstream**
  ✦ Handle + Bitstream Name
System-wide Alert

✦ Accessed through the Control Panel menu item
✦ Can be used to notify users of downtime or other maintenance
✦ Timer can be added to note expiration
✦ Other possible uses
Configuration Options
Configuration Locations

- **File:** dspace.cfg
  - General DSpace parameters
  - Catch all location

- **File:** xmlui.xconf
  - Where themes are installed
  - Interface plugins, known as aspects

- **Files:** input-forms.xml
  - Configure the submission questions
dspace.cfg: Email Options

# From address for mail
mail.from.address = dspace-noreply@myu.edu

# Currently limited to one recipient!
feedback.recipient = dspace-help@myu.edu

# General site administration (Webmaster) e-mail
mail.admin = dspace-help@myu.edu

# Recipient for server errors and alerts
# alert.recipient = email-address-here

# Recipient for new user registration emails
# registration.notify = email-address-here
dspace.cfg: Search & Index

search.index.1 = author:dc.contributor.*
search.index.2 = author:dc.creator.*
search.index.3 = title:dc.title.*
search.index.4 = keyword:dc.subject.*
search.index.5 = abstract:dc.description.abstract
search.index.6 = author:dc.description.statementofresponsibility
search.index.7 = series:dc.relation.ispartofseries
search.index.8 = abstract:dc.description.tableofcontents
search.index.9 = mime:dc.format.mimetype
search.index.10 = sponsor:dc.description.sponsorship
search.index.11 = identifier:dc.identifier.*
search.index.12 = language:dc.language.iso
dspace.cfg: Authentication

- **Password (default)**
  - User’s signup for an account with the repository

- **LDAP**
  - Access one university’s local account management system

- **Shibboleth**
  - Access multiple universities account management systems
dspace.cfg: Browse Indexes

- **Browse Metadata:**
  - Name
  - Metadata fields
  - Data type (title, text, date)

- **Browse Items**
  - Name
  - Sorting option

- **Sorting Options**
  - Name
  - Metadata field
  - Data type (title, text, date)
dspace.cfg: Browse Indexes

# Browse Configuration
webui.browse.index.1 = dateissued:item:dateissued
webui.browse.index.2 = author:metadata:dc.contributor.*:text
webui.browse.index.3 = title:item:title
webui.browse.index.4 = subject:metadata:dc.subject.*:text
#webui.browse.index.5 = dateaccessioned:item:dateaccessioned

# Sorting Options
webui.itemlist.sort-option.1 = title:dc.title:title
webui.itemlist.sort-option.2 = dateissued:dc.date.issued:date
webui.itemlist.sort-option.3 = dateaccessioned:dc.date.accessioned:date
xmlui.xconf: Themes & Aspects

- **Install Manakin themes**
  - For the whole repository
  - For specific communities
  - For specific collections
  - For a specific page

- **Install Manakin Aspects:**
  - For the whole repository
input-forms.xml: Metadata fields

✦ Define the questions asked during an item’s submission

✦ Create forms that can be attached to particular collections which:
  ✦ How many pages or steps are there to describe an item
  ✦ What metadata fields are presented on each page
Batch Import
Ingest Process

External SIP
Batch Importer
Web-based Submission
In Progress Submission
Workflow
Item Installer
Item in DSpace

DSpace Ingest Process
Batch Import

- Command line import
- Directory structure
  - “contents” file
  - “dublin_core.xml” file
  - “handle” file
  - bitstreams files
Simple Contents File

Format:

```
<filename>
```

or

```
<filename>  <tab> bundle:<bundle name>
```

Examples:

```
dissertation.pdf
mods.xml          bundle:METADATA
license.txt       bundle:LICENSE
```
Dublin Core Metadata

Example:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<dcollection>
  
  <dcvalue element="contributor" qualifier="author">John</dcvalue>
  
  <dcvalue element="language" qualifier="iso">en</dcvalue>
  
  <dcvalue element="subject" qualifier="none">Technology</dcvalue>
  
  <dcvalue element="title" qualifier="none">Sample Title</dcvalue>

</dublin_core>
```
Import Command

./dsrun org.dspace.app.itemimport.ItemImport

-a Add new items to DSpace
-c <coll> Which collection to add them to
-e <email> Existing user who is adding these items
-m <path> Create a log file for this import
-s <path> Location of the import files
-t Do not run, just test the import for validity
-h Print command line options and their description
Import Command (example)

Examples

```
./dsrun org.dspace.app.itemimport.ItemImport -a
   -c 123456789/5 -e scott@library.tamu.edu
   -m /path/to/file.map -s /path/to/import
```

- `-a` Add new items to DSpace
- `-c <coll>` Which collection to add them to
- `-e <email>` Existing user who is adding these items
- `-m <path>` Create a log file for this import
- `-s <path>` Location of the import files
Contact Information

Scott Phillips
Research and Development
scott@library.tamu.edu

Jay Koenig
Metadata Librarian
jkoenig@tamu.edu