Standards (or lack thereof) in Institutional Repositories

Sarah L. Shreeves
IDEALS Coordinator
University of Illinois at Urbana-Champaign

DLF Fall Forum 2007
issues in the ir environment

- Very diverse types of material from diverse disciplines - difficult to impose one controlled vocabulary for subjects
- Reliance on faculty deposit and description – tension between ease of deposit and full description
- Reliance on full text indexing – metadata not a high priority
- Limited resources for technical development and resource description
IDEALS News

What is IDEALS? IDEALS collects, disseminates, and provides persistent and reliable access to the research and scholarship of faculty, staff, and students at the University of Illinois at Urbana-Champaign. Learn more.

Faculty, staff, and graduate students can deposit their research and scholarship - unpublished and, in many cases, published - directly into IDEALS. IDEALS includes research papers, technical reports, published papers, pre-prints, conference papers and presentations, datasets, and other digital research material. Learn more.

Departments can use IDEALS to distribute their working papers, technical reports, or other research material. Contact Sarah Shreeves, IDEALS Coordinator, for more information.

Recently Added Documents

2007-Nov-5 - Agriculture's Drive And Dine Menu

2007-Nov-5 - Pork Industry Looks For Third Profitable Year

2007-Nov-5 - Can Pork Producers Handle $3.00 Corn?

Top Downloads this Month

DSpace How-To Guide: Tips and tricks for managing common DSpace chores (Now serving DSpace 1.4.2 and MongoDB) [total: 156]
some basics about our environment

- Use DSpace 1.4

- Allow departments to make decisions about workflows in their ‘communities’ [within the collection policy, of course]

- Have an open community for faculty, staff, and grad students so no direct contact with faculty or department is necessary
our metadata workflow

- We don’t have a metadata workflow

- We have very few required fields – title, year of ‘publication’, type of content (as in dcmi type), one subject keyword

- We don’t approve metadata, we don’t check metadata item by item
we do some work with metadata

- Bulk import
  - IDEALS maps metadata from native formats (usually highly customized) to the QDC in use
  - Determine what to include and what not to include
  - Normalize and add fields where possible

- Deposit items for faculty
we do use some standards

- Qualified Dublin Core (DSpace out of the box schema with some very slight additions and corrections)

- OAI PMH data provider with simple Dublin Core

- Enforce a few encoding schemes for date, language, type
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>contributor</td>
<td>illustrator</td>
</tr>
<tr>
<td>contributor</td>
<td>other</td>
</tr>
<tr>
<td>contributor</td>
<td>student</td>
</tr>
<tr>
<td>coverage</td>
<td>spatial</td>
</tr>
<tr>
<td>coverage</td>
<td>temporal</td>
</tr>
<tr>
<td>creator</td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>accessioned</td>
</tr>
<tr>
<td>date</td>
<td>available</td>
</tr>
</tbody>
</table>

**Spatial characteristics of content.** Used when mapping from MARC / MODS. Repeatable. Not available from user deposit interface. MARC Mapping: 650$a; 651$a; 752$a$b$c$d$f

**Temporal characteristics of content.** Used when mapping from MARC / MODS. Repeatable. Not available from user deposit interface. MARC mapping: 650$y; 651$y

**An entity primarily responsible for making the resource.** Repeatable. Depositor supplied. Uncontrolled. MARC Mapping: 100, 110, 111

**Date IDEALS takes possession of item.** Required. Not repeatable. System supplied. ISO 8601 encoding scheme in use.

**Date or date range item became available to the public within IDEALS.** Required. Not repeatable. System supplied. ISO 8601 encoding scheme in use.

**[NOT USED BY IDEALS] - Date of copyright.**
we don’t use others

• Metadata in use is not ‘standard’

• No standard controlled vocabularies in use for:
  • Subject
  • Genre

• No authority control
  • Names
With the availability of relatively easy-to-use tools for online video creation and distribution, people are increasingly producing videos not just for artistic expression, but also as a form of communication. This article examines two types of online video conversations, one within a specific online subculture, and one a series of responses to one particularly well-known video. Videos lend themselves to the expression of ambiguity, contradictory meanings, taboo topics, and emotions. Within these two video conversations, topics include discussions of masculinity, identity, and fame.

http://hdl.handle.net/2142/705
Title: 3D scene modeling for Distributed Video Coding
Authors: Maitre, Matthieu
Guillemot, Christine
Morin, Luce
Keywords: Video coding
Stereo vision
Issue Date: 2006-Oct
Publisher: IEEE
Citation Information: International Conference on Image Processing (ICIP)
Abstract / Summary: The compression efficiency of Distributed Video-Coding (DVC) suffers from the necessity of transmitting a large number of key-frames which are intra-coded. This paper describes a new 3D model-based DVC approach which reduces the keyframe frequency. The decoder first recovers a 3D model from the key-frames. It then predicts the intermediate frames by projecting it onto 2D image planes and applying image-based rendering techniques. This paper also introduces a new quasi-DVC method relying on a limited point tracking at the encoder. It greatly improves the prediction PSNR, while only slightly increasing the encoder complexity. It also allows the encoder to adaptively select the key-frames based on the video motion-content.
URI: http://hdl.handle.net/2142/159
Type of Resource: text
Genre of Resource: conference paper
Publication Status: published or submitted for publication
Peer-Reviewed: is peer reviewed
Communities have developed their own ‘controlled’ vocabularies
<table>
<thead>
<tr>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
</tr>
<tr>
<td>Laboratories -- High schools -- Illinois -- Chicago -- Safety measures</td>
</tr>
<tr>
<td>Laboratories -- High schools -- Pollution prevention</td>
</tr>
<tr>
<td>Laboratories -- High schools -- Source reduction (Waste management)</td>
</tr>
<tr>
<td>Lake Calumet</td>
</tr>
<tr>
<td>Lake Depue (Ill.) -- Environmental aspects</td>
</tr>
<tr>
<td>Lake sediments</td>
</tr>
<tr>
<td>Lake sediments -- Toxicity -- Illinois -- Peoria Lake</td>
</tr>
<tr>
<td>Lake sediments -- Analysis -- Illinois -- Cook County</td>
</tr>
<tr>
<td>Land Reform</td>
</tr>
<tr>
<td>Landscape Architecture</td>
</tr>
<tr>
<td>Language</td>
</tr>
<tr>
<td>Latin America</td>
</tr>
<tr>
<td>Latino children's books</td>
</tr>
<tr>
<td>Latino/a</td>
</tr>
<tr>
<td>Law</td>
</tr>
<tr>
<td>law</td>
</tr>
</tbody>
</table>

LCSH (from working paper series)
Ethnography of the University Initiative
problems

• Lack of consistency in:
  • Completeness of description
  • Fields used
  • Vocabulary used

• Interoperability suffers (not very shareable metadata)

• Not recording relationships between items very well

• Metadata in use sometimes not sufficient

• Beginning to get more non-textual items
what can we do to address these problems?*

- Developing (slowly) an AJAX based application that allows a community to build a common set of subject terms as deposits occur
  - At the point of deposit, can select from terms used by others in the community or create new one
  - Also could be used for author field
- Reusable templates of recurring metadata for different collections

*These are things that we’re actually working on or on our development list
our wish list

- Ability to select from creation of a collection appropriate metadata schema
- Ability to easily use different controlled vocabularies
- An accepted standard for author identifiers
- Ability to map relationships between items
contact Information

Sarah L. Shreeves
sshreeve@uiuc.edu
217-244-3877