Mission

- “Taxonomy ShareSpace -- access, deposit, share, save, and discuss taxonomy resources”
- There is no theoretical limit to the scope of subject matter coverage in ShareSpace.
Additional purposes

- To provide visitors with informative examples of taxonomy structure
- To show the variety and scope of taxonomies (and other terminologies)
- To reveal the possibilities that can be achieved with a taxonomy
- To implement and take advantage of JISC recommendations

JISC recommendations

- Published in Terminology Registry Scoping Study (TRSS): Final Report
  - By Koraljka Golub (UKOLN, University of Bath) and Douglas Tudhope (University of Glamorgan)
- JISC = the Joint Information Systems Committee of the Higher and Further Education Funding Councils (UK)
Outline of the talk

- JISC terminology recommendations (as implemented in Taxonomy ShareSpace)
- Tour of Taxonomy ShareSpace website
- Search and discovery of the metadata
- NICEM thesaurus for navigation and search
- Contributions of taxonomy metadata

What a terminology registry is

- TRSS: “A terminology registry (TR) lists, describes, identifies and points to sets of vocabularies available for use in information systems and services.”
- “The registry allows discovery of suitable schemes for information or, potentially, use, by exposing rich metadata about them for navigation and retrieval.”
Implemented recommendations

- Which JISC/TRSS recommendations are followed in Taxonomy ShareSpace
- How the metadata types are implemented on the website
  - Viewer interface
  - Contributor template

Group 1 – General information

- Vocabulary name
  - No-brainer – essential metadata
  - Also serves as a heading
- Vocabulary alternative name or acronym
  - Originally omitted; acronym included w name
  - Decided to include field to improve search
- Vocabulary type
  - Pick list in entry template for contributors
Group 1 (continued)

Vocabulary type: Thesaurus
- Authority file
- Classification System
- Gazetteer
- Subject Headings
- Controlled vocabulary
- Taxonomy
- Thesaurus
- Ontology
- Concept map

Group 1 (continued)

- Author or editor
  - Credits individuals
  - Includes institutional authors
- Current version/edition
- Date of current version
- Update frequency
Group 1 (concluded)

- Available format(s)
- Omitted – Available terminology services
- Vocabulary identifier (we used “Vocabulary URL”)
- Vocabulary sample URL
- Vocabulary description (essential)

Group 2 – Scope and Usage

- Language(s)
  - We sought out multiple language vocabularies to enrich the variety of the registry.
  - Pick list or free text? Problems w. short list
  - ISO language codes list - ISO 639-2 or -3
  - Numerous languages, some quite obscure (example: OmegaWiki)
Group 2 (continued)

- Major subjects covered
  - In cases for which major subjects were not included in the descriptive material on the terminology website, we used the top terms.

- Minor subjects covered
  - Included. However, the major subjects are often fully indicative of the subject matter scope.

Group 2 (continued)

- Purpose as given by author/publisher
  - We used the label “Purpose” for succinctness.
  - The purpose is sometimes discernible, but no explicitly stated by the author/publisher.

- Used by
  - Sometimes clearly stated in a separate paragraph on the terminology website
  - Often covered in the description or purpose
Group 2 (continued)

- Related vocabularies
- Overlap with related vocabularies
- Mappings to other vocabularies: which vocabularies, intellectual or automated

Group 2 (concluded)

- Optional Group 2 suggestions not used:
  - Description of collections where used
  - Usage case study
  - Use in application profiles
  - Rating
  - URL to vocabulary users’ discussion board
  - Change notification details
  - URL to tutorial for applying vocabulary
3 – Vocabulary characteristics

- Type of display
  - Check list
  - Hierarchical; Alphabetical; etc.

- Description of overall structure

- Type of terms
  - TRSS: “e.g. concept terms, geographic names

Group 3 (continued)

- Types of relationships
  - Check list (see next slide)
  - Hierarchical; associative; equivalence (etc.?)
  - Two major types of equivalence found
    - Preferred/Non-preferred
    - Equal synonyms (in synonym rings, termbanks)

- Total number of terms
- Total number of classes
### Group 3 (continued)

**Relationship types:**
- Equivalence (Synonym or alternate name)
- Equivalence (Primary/variant terms or Preferred/non-preferred)
- Equivalence (Language or dialect)
- Equivalence (Inter-language; may be between separate lists)
- Hierarchical (Broader/narrower terms)
- Hierarchical (Generic, isa) or Hierarchical (specific relationship)
- Hierarchical (Instance, isa)
- Hierarchical (Whole/part)
- Associative (Related terms)
- Associative (Symbols, animation, sound)

### Group 3 (concluded)

- **Number of preferred terms**
- **Number of non-preferred terms**
- **Depth of hierarchy**
  - On rare occasions, is stated on website
  - We sample branches (“at least n levels”)
- **Not used – Notes field (Optional)**
- **Not used Information given (Optional)**
Group 4 – Term and conditions

- Availability: free for all, free for registered users, costs
  - Initial focus was on terminology websites freely available for viewing

- Import/download instructions

- Purchase/subscription price

- Licensing options

Group 5 - Provider

- Vocabulary provider name
  - We generally listed the organization hosting the vocabulary website

- Vocabulary provider URL

- Vocabulary provider contact details
Current architecture

- Mainly an enriched directory
- Names and links
- Metadata
- Seed/starter entries to attract users and encourage outside contributions
- Is set up (in Drupal) for entries by contributors

Collaboration opportunities

- Users can create the following content:
  - Blog entries
  - Event announcements
  - Stories/articles
  - Terminology metadata
  - Eventually, downloaded terminologies
    - Preserve for endangered thesauri
Special case – student thesauri

- Thesauri usually on university sites
- Therefore in danger of extinction
- We offer to preserve thesauri/web pages

Interesting subjects
- Harry Potter thesaurus (by Dr. Stasa Milojevic)
- Sailspotting (Sails of tall ships)
- Thesaurus of Bellydancing (ThoBe) by the League of Bellydancing Librarians, Indexers, and Archivists (LoBeLIA)

Possible eventual architecture

- Storage of full contributed terminologies
- Use of SKOS wherever practicable
- Search capability across terminologies
- Automated indexing of descriptions
- Navigation tree based on NICEM thesaurus as “spine” for collection
Home page

Welcome to the Taxo ShareSpace Terminology Registry

The Taxonomy ShareSpace houses a collection of controlled vocabularies of all types and complexities. We invite you to both browse and contribute. Easy ways to be special special users, get ideas for building your own vocabulary, perhaps first one that can give you a quick start.

The information collected about each vocabulary follows a study (IRS) conducted by JISC, the Joint Information Systems Committee of the Higher and Further Education Funding Council. All of the recommendet fields included in the study’s first report, some of which are detailed in the study’s report, were not included in this project. See more about the Terminology Registry blogs at ShareSpace (2007) on this site. In addition, some of the information included was added in planning the site.

This is an interactive wiki site. To add information about a vocabulary, click on Create Content in the left navigation pane (you’ll need to register as a user first). You may need your name and password. There are only two required fields, but your location will become more useful if you complete all the applicable fields about your vocabulary.

Add a comment to an existing page - you’ve used the vocabulary, what did you add it to, how was it used in your document? Also, add articles, etc. Comments are welcome on specific Wiki pages as well. Click on Add Comment and write your thoughts. Even anonymous visitors can add comments, but they’ll be reviewed by a site admin before they’re made visible.

You may also update the Events section of the site. Taxonomy, Knowledge Systems, Information Architecture or Management. Web Data are all appropriate event themes. Click on Create Content and then on Events to add a new one (you’ll need to be a registered user).

Sample record

NAL Agricultural Thesaurus

Submitted by sk Nhà on Thu, 2003-10-15 09:19

"The Thesaurus and [an associated] Grüße are online vocabulary tools of agricultural terms in English and Spanish and are cooperatively produced by the National Agricultural Library, USDA, and the Inter-American Institute for Cooperative Agriculture through the Orton Memorial Library."

General information
Vocabulary alternative name or acronym:
NAL Thesaurus, Agricultural Thesaurus, NAL T
Vocabulary type: Thesaurus
Author or Editor: LentFinch, Thesaurus Coordinator
Current version date: Wed, 2008-12-31
Update frequency: Annually
Available formats: KIN, SKOS, PDF, MARC, and TXT
Vocabulary URL:
http://aclass.nal.usda.gov/ mirror site at http://aclass.carn.msu.edu/cgi-bin

Vocabulary Sample URL:

Was vocabulary created as a course project: No

Scopes and usage
Languages:
English, Spanish
Sample record (continued)

Major subjects covered:
- Agricultural subjects, biological sciences
- Minor subjects covered:
- Rural and agricultural sociology; physical sciences

Purpose:
"The NAL Agricultural Thesaurus (NALT) was originally prepared by staff of the National Agricultural Library to meet the needs of the United States Department of Agriculture (USDA), Agricultural Research Service (ARS). It is presently used for indexing and for improving retrieval of agricultural information."

Used By:
Currently, the thesaurus is the indexing vocabulary for the National Agricultural Library's bibliographic database of citations to agricultural resources, AGROVOC. The Food Safety Research Information Office (FSRIO) and Agricultural Network Information Center (AgNIC) also use the NAL Thesaurus (NALT) as the indexing vocabulary for their informatics systems. In addition, the NALT is used as an aid for locating information at the Agricultural Research Service, Economic Research Service and AgNIC web sites and databases.

Related vocabularies:
The NAL Thesaurus provides definitions for 2,574 of the thesaurus terms in English and 2,574 terms in Spanish.

Overlap with related vocabularies:
It can be assumed that NALT has considerable overlap with AGROVOC, another agriculture related thesaurus.

Mappings to other vocabularies:
NALT was aligned with AGROVOC by participants in the DARE 2007 food task. See http://www.fao.org/wdr-n/va/wha/992007/food.html

Vocabulary characteristics
Type of display: Hierarchy
Description of overall structure:
See http://aginis.nal.usda.gov/structure/sln#eqiv

Type of terms:
Agricultural and scientific concepts

Sample record (continued)

Relationship type:
- Equivalence (Primary/variant terms or Preferred/non-preferred)
- Hierarchical (Broader/narrower terms)
- Associative (Related terms)

Number of concepts:
17 subject hierarchies

Number of terms:
- English: 73,194
- Spanish: 69,118
- Number of preferred terms:
- English: 44,007
- Spanish: 44,007
- Number of non-preferred terms:
- English: 28,397
- Spanish: 24,281

Characteristics Comments:

Terms and Conditions
Availability:
Free to use
Comments, Availability:
Downloadable from the website in XML, SKOS, PDF, WARC and TXT formats. Other formats may be obtained by contacting the NAL Thesaurus (first by email, mail, phone, or fax; browseable online).
Sample record (concluded)

Licensing Options:
No license is required to obtain the data contained in the NAL Agricultural Thesaurus and Glossary. If the use is not personal, (1) the U.S. National Agricultural Library must be identified as the creator, maintainer, and provider of the data; (2) the version of the data must be clearly stated along with a description of the modifications.

Purchase/Subscription price:
$0.00
Import/download instructions:
http://explore.nal.usda.gov/download.shtml Downloads involve the acceptance of the NAL's Usage Agreement

Provider:
Vocabulary provider name:
National Agricultural Library, U.S. Department of Agriculture
Provider URL:
http://www.nal.usda.gov/
Provider contact details:
Vocabulary provider contact details: E-mail: techs@nal.usda.gov. Listing address:
National Agricultural Library, c/o Lon Finch, Thesaurus Coordinator, 8031 Baltimore Ave., Beltsville, MD 20705. Phone number: (301) 504-6883. Fax number: (301) 504-6212

Results from data base show all fields

Vocabulary characteristics
Type of display:
- Hierarchy
- Alphabetical list
Type of terms:
terminology peculiar to works of J.K. Rowling
Relationship types:
- Equivalence (Primary/variant terms or Preferred/non-preferred)
- Hierarchical (Broader/narrower terms)
- Associative (Related terms)
Number of classes:
23

Terms and Conditions
Availability:
Other arrangement (contact Provider)

Provider:
Vocabulary provider name:
Stana Miljevic
Provider URL:
Navigation by Taxonomy

Automatic Completion of Synonyms and Preferred Terms

Search by Taxonomy

<table>
<thead>
<tr>
<th>PSY</th>
<th>Psychiatric hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Psychic phenomena (nonpreferred)</td>
</tr>
<tr>
<td></td>
<td>Psychological disorders</td>
</tr>
<tr>
<td></td>
<td>Psychopharmacology (nonpreferred)</td>
</tr>
<tr>
<td></td>
<td>Psychotropic drugs (nonpreferred)</td>
</tr>
</tbody>
</table>

Abnormal psychology
Adolescent developmental psychology
Adult developmental psychology
Animal behavioral psychology
Animal psychology
Child developmental psychology
Clinical psychology
Cognitive psychology
Community psychology
Clinical psychopharmacology
Next steps

- Add navigation taxonomy and search
- Decide how to accommodate uploads
- Populate as KOS become known
- Allow others to add their KOS
- Spread the word on the registry
- Encourage others to join and participate
- Build the community

Thank you!

Marjorie M.K. Hlava
President
Access Innovations, Inc.
mhlava@accessinn.com
505-998-0800