Using Our Taxonom(ies)  
ASCE Case Study  

Charlotte McNaughton  
Director of Digital Product Development  
American Society for Civil Engineers  
Reston, Virginia  
26 September 2016
A CASE STUDY

- ASCE Library and Taxonomies
  - Charlotte McNaughton
  - Director of Digital Product Development
- Presented on her behalf by
  - Marjorie Hlava
  - President, Chief Scientist, Access Innovations Inc.
• 37 Journals: 73,000+ papers, 1983 to present
• 1 Magazine: 1000 articles + columns, 2005-2015
• 498 Proceedings: 54,000+ papers
• 333 E-books
• 82 Standards
Taxonomy Development

- Begun in 2009 finished 2011
- Content driven (our published content, not broad idea of discipline)
- Controlled vocabulary, 3500 terms
- Flat list (no hierarchy)
- Draft grew to 30,000 terms for Topics, Corporate, Geographic
- Dropped Corporate, condensed Geographic, edited Topics
- Review by Subject Matter Experts
More Development!

- Topics taxonomy: 2,170 Preferred Terms
- Geographic taxonomy: 778 Preferred Terms
- NEW!!! Content type taxonomy: XXX Preferred Terms
- Review never ends!
- Not just a matter of having a taxonomy…need to have a way to use it. Vendors and applications just now coming online.
- For specific uses, we found that we needed additional versions or variations of taxonomy
- Author Specialty taxonomy
- Author Country taxonomy
Primary implementation: ASCE Library

- Tagging of retro content in ASCE Library complete
- Two branches of taxonomy uploaded to Atypon Literatum Taxonomy Management Tool
- Tags for retro content: provided to Atypon
- Forward flow for tagging new content developed
- Redesign in progress…implementation still in development
- Tags will display on article-level pages; click to populate search?
- Type-ahead in Search using Preferred Terms?
- Taxonomy tree or other visualization for browsing?
New Taxonomy!

- Issue: create a filter for Content Types that would work with Search Results on our new ASCE Library platform
  - Need article types to match ASCE terminology
  - "Research article" should be "Technical Paper"
  - "Chapter" includes "Proceedings Paper": split!
  - Want to add Open Access filter
  - Want to add book-title level filters

<table>
<thead>
<tr>
<th>ARTICLE TYPE</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Article</td>
<td>35371</td>
</tr>
<tr>
<td>Chapter</td>
<td>33891</td>
</tr>
<tr>
<td>Brief Report</td>
<td>1133</td>
</tr>
</tbody>
</table>

Filter by:
- Book (3)
- Book - Proceedings (0)
- Book - Standard (0)
Development

• Select Preferred Terms (how they should appear in the filter)
• Map XML tags and doctypes from our XIS database
• Create new taxonomy in Thesaurus Master
• Tag content
• Each item in ASCE Library has one tag for Content Type
• Items that are Open Access get a second tag
• Prepare tag export file
• Provide taxonomy and tags to Atypon
Second Implementation: New Search Interface

CIVIL ENGINEERING DATABASE

AMERICAN SOCIETY OF CIVIL ENGINEERS
ASCE LIBRARY

search tips

all fields

author/affiliation

title

select serial title all titles

subject headings

publication year 

search reset
New Search Interface

- Subject headings = Taxonomy Preferred Terms, permuted!
New Search Interface

• Filter for content types

Subject headings: Snow loads
Publication year: __________ to __________

Search
Reset

Found 6 Records

2014  Snow-Related Roof Collapse during the Winter of 2010-2011: Implications for Building...
1996  Cold Regions Utilities Monograph
1981  Long Span Roof Structures
Third Implementation: Author Profiles

- Working with colwiz
- Snippet of javascript applied to templates in ASCE Library
- Connects to colwiz iPDF Reader; users can annotate, save, share…
- Additional features on Abstract pages: Awards
Author Specialty Taxonomy

Kebei Li
Professor, Key Laboratory of Civil Engineering Safety and Durability of China...
Damage, Humans, Concrete, Humidity, Swelling

55 Publications

71 Co-authors

55 Publications

79 Expertise

- Damage
- Humans
- Concrete
- Humidity
- Swelling
- Gene Expression
- Durability
- Electrical resistivity
- Moisture
Author Specialty Taxonomy Development

• Can Technical Topics tags be used to define Author expertise?
  • No!
• Too granular: Granularity works for articles, but not authors
• Develop from existing resources…
• Harvest terms from peer-review system used for matching Reviewers with articles
• Create a flat list of all terms; de-dupe
• Map Tech Topics tags to new flat list; run Frequency Report, delete low-usage terms
• Merge Tech Topics terms into new Author Specialty taxonomy; pulls in Rule Base
• Test run of selected fields against new taxonomy; examine 0 hits report
• Modify new taxonomy and re-run machine-aided indexing for final tags
• Preferred terms = 184
Fourth Implementation: Author Country Taxonomy

- Need: requests from Marketing for lists of Authors from various countries
- Obstacle: country information combined with other Affiliation data
- Obstacle: country name variation and inconsistency
- Develop from existing resources…
- Start with Geographic taxonomy, countries branch
- Run Affiliations against these terms, extract countries
- Success! First pass: 70% hits returned
- Analyze misses; revise rules for this task; re-run
- Preferred terms = 204
Evolution of ASCE Taxonomy

ASCE Taxonomy (includes technical, corporate and geographic terms)

ASCE Topics Taxonomy
ASCE Geographic Taxonomy

ASCE Technical Taxonomy
ASCE Geographic Taxonomy
ASCE Corporate Taxonomy

ASCE Technical Taxonomy
ASCE Corporate Taxonomy
ASCE Geographic Taxonomy
ASCE Author Specialty Taxonomy
Another project?
Institution name normalization

- Need: associate Authors with Institutions
- Obstacle: various and inconsistent data for Institution names
- Resource: GRID (Global Resource Identifier Database)
- https://grid.ac/
- Similar process as Author Country approach???
Another project?
Study or Structure Location with JournalMap
Study or Structure or Disaster Location with JournalMap

- Geotags needed for locations
- Return tags to bibliographic database
- Use locations to develop GEOthes
- What level of Topic tagging is appropriate?
- Geoparser + Manual review
- Model is location of a Study
- Civil engineers are interested in Projects and Structures and Disasters
Visualization Interface - Sunburst

<table>
<thead>
<tr>
<th>Hierarchy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>G</td>
</tr>
<tr>
<td>H</td>
</tr>
</tbody>
</table>

1. Hierarchy
2. Aerospace engineering
3. Aerodynamics
4. Aeroelasticity
5. Aerodynamic flutter
6. Gravity forces
7. Particle velocity
8. Aircraft and spacecraft
9. Aircraft wings
10. Flight
11. Launching and landing

[Image of a sunburst visualization interface]
Honeycomb Interface

- Water and water resources
- Engineering mechanics
- Engineering fundamentals
- Structural engineering
- Environmental engineering
- Geotechnical engineering
- Transportation engineering
- Materials engineering
- Practice and Profession
- Computing in civil engineering
- Infrastructure
- Aerospace engineering
- Sample Term
- Sample Term 2
More Drilling down through the Interface
More Drilling down through the Interface
What else???

- Visualizations using Tableau
- Recommender widgets
- Topic pages
- Develop other taxonomies (Events, Historical figures, e.g.)
- Leverage other taxonomies and databases (Structurae, e.g.)
- Other ideas from Text Mining???
- Limited only by time and resources!
Questions?
Thank you

Marjorie Hlava
Monica Bradford
Charlotte McNaughton