Turning the Tables on Print Textbooks: Opportunities & Challenges in Converting Print Tables of Contents to Digital Taxonomies

Ira Kleinberg
Heather Hedden

DAM New York 2012
May 11, 2012
About the Speakers

Ira Kleinberg

• Currently Senior Manager of Taxonomy Development at Pearson
• Previously a librarian at San Francisco Public Library, Information Architect at Netscape and indexer for companies such as Apple
• Interested in crowd-sourced taxonomies as an option for organizations without any and the point at which taxonomy management becomes master data management

Heather Hedden

• Taxonomy consultant with Project Performance Corporation
• Continuing education instructor with Simmons College Graduate School of Library and Information Science
• Previously an indexer and controlled vocabulary editor at Thomson/Gale
• Author of *The Accidental Taxonomist* (Information Today, Inc., 2010)
Agenda

• About the Project
• About Taxonomies
• Opportunities in Converting Print Tables of Contents to Digital Taxonomies
• Challenges in Converting Print Tables of Contents to Digital Taxonomies
• Taxonomy Resources
• Questions?
Agenda

- About the Project
- About Taxonomies
- Opportunities in Converting Print Tables of Contents to Digital Taxonomies
- Challenges in Converting Print Tables of Contents to Digital Taxonomies
- Taxonomy Resources
- Questions?
About the Project

• Move from producing one-version-fits-all textbooks to offering highly customizable learning solutions
• Transition from organizing around thousands of book tables of contents and associated text to organizing around taxonomies
• Enable staff, vendors and customers to find, access and repurpose millions of digital assets
About the Project

- Aggregate tables of contents and associated learning objectives
- Synthesize them into robust, discipline-level taxonomies useful for competitive analysis, strategic planning and digital asset management
Agenda

• About the Project
• **About Taxonomies**
• Opportunities in Converting Print Tables of Contents to Digital Taxonomies
• Challenges in Converting Print Tables of Contents to Digital Taxonomies
• Taxonomy Resources
• Questions?
About Taxonomies

A Taxonomy…

• An authoritative, restricted list of terms (words or phrases)
• Where each term is for a single unambiguous concept
• Governed by policies for who, when, and how new terms can be added
• Arranged in some (hierarchical) structure or groupings
• Supporting indexing/tagging/categorizing/metadata management of content to facilitate retrieval/findability
• Possibly including synonyms as cross-references
# About Taxonomies

<table>
<thead>
<tr>
<th>Humanities &amp; Social Sciences</th>
<th>Math &amp; Science</th>
<th>Professional &amp; Career</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropology</td>
<td>Anatomy &amp; Physiology</td>
<td>Business</td>
</tr>
<tr>
<td>Art</td>
<td>Biology &amp; Microbiology</td>
<td>Business Statistics</td>
</tr>
<tr>
<td>Communication, Film &amp; Theatre</td>
<td>Chemistry</td>
<td>Communication Sciences &amp; Disorders</td>
</tr>
<tr>
<td>English</td>
<td>Environmental Science</td>
<td>Computer Science</td>
</tr>
<tr>
<td>History</td>
<td>Geography &amp; Atmospheric Sciences</td>
<td>Counseling</td>
</tr>
<tr>
<td>Interdisciplinary Studies</td>
<td>Geology &amp; Oceanography</td>
<td>Criminal Justice</td>
</tr>
<tr>
<td>Music</td>
<td>Health, Kinesiology &amp; Nutrition</td>
<td>Culinary, Hospitality, Travel &amp; Tourism</td>
</tr>
<tr>
<td>Philosophy</td>
<td>Mathematics</td>
<td>Deaf Studies &amp; Deaf Education</td>
</tr>
<tr>
<td>Political Science</td>
<td>Physics &amp; Astronomy</td>
<td>Economics</td>
</tr>
<tr>
<td>Psychology</td>
<td>Statistics</td>
<td>Education</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td>EMS &amp; Fire Science (Brady)</td>
</tr>
<tr>
<td>Social Work / Family Therapy / Human Services</td>
<td></td>
<td>Engineering</td>
</tr>
<tr>
<td>Sociology</td>
<td></td>
<td>Fashion &amp; Interior Design</td>
</tr>
<tr>
<td>World Languages</td>
<td></td>
<td>Health Professions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information Technology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legal Studies &amp; Paralegal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nursing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Student Success &amp; Career Development</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trades &amp; Technology</td>
</tr>
</tbody>
</table>
About Taxonomies

Taxonomy Benefits

1. Managed as a controlled vocabulary…
   Brings together different wordings (synonyms) for the same concept
   - Helps users search for information by different names
   - Helps users retrieve matching concepts, not just words

2. Structured as hierarchies or facets…
   Organizes information into a logical structure
   - Helps users browse or navigate for information
Agenda

- About the Project
- About Taxonomies
- Opportunities in Converting Print Tables of Contents to Digital Taxonomies
- Challenges in Converting Print Tables of Contents to Digital Taxonomies
- Taxonomy Resources
- Questions?
Opportunities in Converting Tables of Contents to Taxonomies

Tables of Contents…

• Are freely available
• Are easily understood
• Already exist within one or more product hierarchies
• Already exist within one or more pedagogical hierarchies
Opportunities in Converting Tables of Contents to Taxonomies

Tables of Contents…

• Implicitly structure underlying assets
• Explicitly structure associated learning objectives
• Make gaps in coverage apparent
• Have been proven to help with information retrieval
Agenda

- About the Project
- About Taxonomies
- Opportunities in Converting Print Tables of Contents to Digital Taxonomies
- Challenges in Converting Print Tables of Contents to Digital Taxonomies
- Taxonomy Resources
- Questions?
Challenges in Converting Tables of Contents to Taxonomies

Issues with TOC Headings Becoming Taxonomy Terms

• Excessively lengthy/wordy headings
• Vague headings
• Headings that are questions or commands
• Headings for “introduction” to a topic
• Headings for “miscellaneous” or “other”
• Same heading with different meanings under different parent headings
Example Dealing with “Miscellaneous”

**TOC:**

Trusts
  - Trust A Type
  - Trust B Type
  - Trust C Type
  - Miscellaneous Other Trusts
    - Trust D Type
    - Trust E Type

**Taxonomy:**

Trusts
  - Trust A Type
  - Trust B Type
  - Trust C Type
  - Trust D Type
  - Trust E Type
Challenges in Converting Tables of Contents to Taxonomies

Issues with TOC Structure Becoming Hierarchical Taxonomy

• Hierarchies of headings (such as for balance) that do not reflect true taxonomy hierarchy of concepts
• Duplicated topics occurring in introductory chapters and then later in more detail
• Changes in order of subtopics from chronological or instructional order to default alphabetical
Challenges in Converting Tables of Contents to Taxonomies

Example TOC Hierarchy not True Taxonomy

TOC:

Property
  Types of Property
  Probate Versus Nonprobate Property
  Probate Property
  Nonprobate Property
  Real Property Versus Personal Property
  Real Property
  Personal Property

Taxonomy:

Property
  Probate Property
    Real Property
    Personal Property
  Nonprobate Property
Issues with Merging TOCs of Multiple Sources to Create the Discipline Taxonomy

• Differently worded headings covering the same content
• Different organization of headings within and across chapters
• Topics covered in Heading Level 1 or 2 in one title but in Heading Level 2 or 3 in another
• Headings or chapters that combine two topics taken up in two separate headings or chapters of another title
Challenges in Converting Tables of Contents to Taxonomies

Issues with Merging TOCs of Multiple Sources to Create the Discipline Taxonomy

• Differently worded headings covering the same content
• Different organization of headings within and across chapters
• Topics covered in Heading Level 1 or 2 in one title but in Heading Level 2 or 3 in another
• Headings or chapters that combine two topics taken up in two separate headings or chapters of another title
Example Dealing with a Topic at Different Levels

Child abuse is found differently in three different textbooks:
1. As its own chapter
2. Within the chapter on “Family Violence”
3. Within the chapter “Government Intrusion in Family Matters”

Consider:
• The amount of content
• The relative importance of the different sources
Agenda

• About the Project
• About Taxonomies
• Opportunities in Converting Print Tables of Contents to Digital Taxonomies
• Challenges in Converting Print Tables of Contents to Digital Taxonomies
• Taxonomy Resources
• Questions?
Taxonomy Resources


- Special Libraries Association (SLA), Taxonomy Division. [http://wiki.sla.org/display/SLATA](http://wiki.sla.org/display/SLATA)

- American Society for Indexing (ASI), Taxonomies & Controlled Vocabularies SIG. [http://www.taxonomies-sig.org](http://www.taxonomies-sig.org)

- “Taxonomies and Controlled Vocabularies,” 5-week online workshop, Simmons College Graduate School of Library and Information Science Continuing Education Program (next offered in September 2012). [http://alanis.simmons.edu/ceweb](http://alanis.simmons.edu/ceweb)

Agenda

• About the Project
• About Taxonomies
• Opportunities in Converting Print Tables of Contents to Digital Taxonomies
• Challenges in Converting Print Tables of Contents to Digital Taxonomies
• Taxonomy Resources
• Questions?
Questions?

Ira Kleinberg
http://www.linkedin.com/in/irakleinberg

Heather Hedden
http://www.linkedin.com/in/hedden