Outline

1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
When to build a taxonomy from scratch

- Taxonomy is needed
- Taxonomy does not exist, or existing taxonomies/vocabularies are totally unsuitable
- Content collection is unique, so “off-the-shelf” taxonomies or thesauri will not suffice
Who will build the new taxonomy:

- Contracted consultancy
  -- large project, short time, externally led
- Taxonomist temp/freelancer/contractor(s)
  -- small project, internally managed
- Existing staff librarian or information architect
  -- developed over longer period of time
- Taxonomist new hire
  -- anticipating full-time maintenance after it’s built

➢ Additionally, subject matter experts may contribute.
➢ Consider: It needs to be designed and not just built.
Additional pre-building considerations

- Type of indexing/tagging of content
  - Manually by content creators
  - Manually by dedicated trained indexers
  - Automated

- Resources and constraints
  - Time
  - Money
  - Technology (content management, retrieval, and indexing user interfaces)
1. Initial Considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
Definitions

Controlled vocabulary (CV)
- The most general, broadest concept for all applications

- An authoritative, restricted list of terms (words or phrases) mainly used for indexing/tagging content to support retrieval
- Controlled in who and when new terms can be added
- Usually makes use of equivalent non-preferred terms (synonyms, etc.) to point to the correct, preferred terms
- May or may not have structured relationships between terms
Taxonomy

1. A controlled vocabulary with broader/narrower (parent/child) term relationships that include all terms to create a hierarchical structure
   - With focus for categorizing and organization concepts
   - May or may not have equivalent non-preferred terms (synonyms, etc.) to point to the correct, preferred terms

2. A controlled vocabulary used in corporate/enterprise applications
   - Used synonymously for any controlled vocabulary
- Term List
- Synonym Ring
- Authority File
- Taxonomy (definition #1)
  - Hierarchical taxonomy
  - Faceted taxonomy
- Thesaurus
- Ontology
Hierarchical Taxonomy – Has broader term/narrower term relationships that include all terms to create a hierarchical structure

1. Example: Expandable, on same page
Hierarchical Taxonomy

2. Example: Separate page for each level of hierarchy

- **Arcades & Amusements**
  - Amusement Equipment & Supplies, Video Arcades & Internet Gaming Centers, ...

- **Bars**

- **Carnivals, Fairs, & Festivals**
  - Fairgrounds, Music Festivals, ...

- **Children's & Family Entertainment**
  - Miniature Golf Courses, Theme Parks, Zoos, ...

- **Concert Tickets**

- **Cultural Attractions, Events, & Facilities**
  - Historical Places & Services, Libraries, Museums, ...

- **DVD Rentals**

- **Entertainment Clubs & Nightlife**
  - Comedy Clubs, ...

- **Entertainment Industry**
  - Entertainment Agencies & Bureaus, Motion Picture Producers & Studios, Music & Recording Industry, ...

- **Event Planning**
  - Convention & Meeting Planning Services, Party Planners, ...

- **Movie Theaters**

- **Movies, Videos, & DVDs**
  - Video Production Services, ...

- **Music**
  - Music Instruction, Music Stores, Musical Instruments Retail, ...

- **Night Clubs**

- **Performing Arts**
  - Entertainers, Live Theater, Orchestras, Symphonies, & Bands, ...

- **Sports Tickets**
  - Concert Tickets, ...

- **Tours & Charters**
  - Boat & Yacht Charters, Rental, & Leasing, Fishing Guides & Charters, ...

- **Video Game Rentals**

- **Visual Arts**
  - Art Supplies & Art Supply Stores, Artists & Art Studios, Professional Photographers, ...
Hierarchical taxonomy is suitable for:

- Content that is naturally categorizable: products, industries, government agencies, academic disciplines, scientific things, technologies
- Taxonomies of any size, but especially smaller taxonomies
- Browse navigations
- Certain kinds of auto-categorization, which puts documents into approximate categories
Facets

- For serving faceted classification, which allows the assignment of multiple classifications to an object
- A “dimension” of a query; a type of concept; domain of content
- Intended for searching with multiple terms in combination (post-coordination), one from each facet: to “limit by” filters
Faceted taxonomy examples
Example types of facets:

- For Products: *name*, series *number*, *category*, *size*, *color*, *price*
- For People: *name*, *job title*, *gender*, *birth year*, *location*, *dept.*
- For Reports: *title*, *author*, *subject*, *audience*, *document type*

Facets are suitable for:

- Structured data with discernable metadata fields or database records
- Homogeneous data with similar types of characteristics (e.g. products in an e-commerce site)
Thesaurus

Terms have:

- hierarchical relationships,
- associative relationships, and
- nonpreferred/preferred terms

Thesauri are suitable when:

- Terms are not easily categorized in hierarchies
- Content is manually indexed by trained indexers
- Users are subject-matter experts

**Thesaurus entry example**

<table>
<thead>
<tr>
<th>materials acquisitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF acquisitions (of materials)</td>
</tr>
<tr>
<td>library acquisitions</td>
</tr>
<tr>
<td>BT collection development</td>
</tr>
<tr>
<td>NT accessions</td>
</tr>
<tr>
<td>approval plans</td>
</tr>
<tr>
<td>gifts and exchanges</td>
</tr>
<tr>
<td>materials claims</td>
</tr>
<tr>
<td>materials orders</td>
</tr>
<tr>
<td>subscriptions</td>
</tr>
<tr>
<td>RT book vendors</td>
</tr>
<tr>
<td>jobbers</td>
</tr>
<tr>
<td>subscription agencies</td>
</tr>
<tr>
<td>subscription cancellations</td>
</tr>
</tbody>
</table>
### Types: Summary

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Less</th>
<th>Complexity</th>
<th>More</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pick List</td>
<td>Synonym Ring</td>
<td>Authority File</td>
<td>Taxonomy</td>
</tr>
<tr>
<td>Ambiguity control</td>
<td>Synonym control</td>
<td>Ambiguity control Synonym control (preferred &amp; non-preferred terms)</td>
<td>Ambiguity control Synonym control (Synonym control) Hierarchical Relationships</td>
</tr>
</tbody>
</table>
1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
Deciding the Taxonomy Scope:

- Content scope
- Subject area scope
- Term types to include
- Usage
- Breadth and depth
Content Scope: Content to be covered by the taxonomy can be limited to some or all of:

**Enterprise content types:**
- Manuals and policies
- Standards
- Product data
- Transaction records
- Reports & white papers
- Blog articles
- Marketing literature
- External publications
- Employee handbook content
- Meeting notes

**Library content types:**
- Books/monographs
- Multimedia
- Periodical articles
- Special collections
- Object/art collections
- Library internal reports
- Website content
Content Scope: Content to be covered by the taxonomy can be limited to some or all of document/file formats:

- Word documents
- PDF documents
- Presentations
- Spreadsheets
- Web pages
- Image files
- Videos, podcasts
- Database records
- Other published formats
Subject Area Scope

Example: Area to be covered by an enterprise taxonomy can be limited to some or all of:

- Products and services (names, descriptions, features, user issues)
- Technology and technical terminology (science, research & development oriented)
- People and their areas of expertise
- Human resources
- All information in an enterprise
Scope of Term Types
(could also be considered as metadata or “facets”)

Term types to be included can be limited to some or all of:

- Subjects/Topics
- Person names
- Organization/company names
- Product names
- Locations
- Activities/Actions
- Job titles
Taxonomy Scope, Breadth and Depth

Usage Scope
– to be used by:

- Employees, Partners and the Public
- Employees & Partners or subscribers
- All Employees
- Certain Employees Only
- Employees Only

- Public Web Site
- Extranet or restricted access website
- Intranet
- Intranet restricted sections
Breadth and depth

- How many terms (approximate)
- How many (if any) hierarchical levels deep
- How specific the terms get

A specific term vs. combining two less specific terms

<table>
<thead>
<tr>
<th>Drug trials</th>
<th>Product testing and Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRM Software</td>
<td>Customer Relations Management and Software</td>
</tr>
</tbody>
</table>
Outline

1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
Deciding whether a concept should be a term

- Is it within the subject-area scope of the CV?
- Is it important, likely to be looked up
- Is there enough information (or sufficient anticipated information)
- Do users want and expect it?
1. Choosing the preferred term
   - Doctors vs. Physicians
   - Movies vs. Motion pictures
   - Cars vs. Automobiles

2. Creating nonpreferred terms (for search or alphabetical browse, not for navigation)
   - Cars
     - UF Automobiles
     - UF Autos

3. Structuring the hierarchy or relationships (broader/narrower or related)
   - Motor vehicles  See also Drivers
     - Cars
       - - Sports cars
2. Creating Nonpreferred Terms aka synonyms, variants, USE/UF (used from), alternate labels, See references

Types include:

- **synonyms**: Automobiles USE Cars
- **near-synonyms**: Junior high schools USE Middle schools
- **variant spellings**: Defense USE Defence
- **lexical variants**: Hair loss USE Baldness
- **foreign language terms**: Luftwaffe USE German Air Force
- **acronyms/spelled out forms**: CDs USE Compact discs
- **scientific/technical names**: Neoplasms USE Cancer
- **antonyms (for characteristics)**: Softness USE Hardness
- **narrower terms and instances that are not preferred terms**: Hydroelectric power plants USE Power plants
3. Structuring the hierarchy

Reciprocal (bi-directional) relationships, but asymmetrical

Broader term (BT)

SOME \hspace{1cm} \text{ALL}

Narrower term (NT)

SOME \hspace{1cm} \text{ALL}

Fruits NT Oranges \hspace{1cm} Oranges BT Fruits

Three types:

a) Generic - Specific
b) Common noun - Proper noun
c) Whole – Part
a) Hierarchical - Generic/Specific:

Category or class
– members
– more specific types

Narrower term
“is a” or “are a kind of” broader term

Plants
NT Trees
Financial services
NT Investment services
Romance languages
NT Italian
a) Hierarchical - Instance:

Common noun
– Proper noun

Narrower term
“is a” broader term

Smartphones
NT Samsung Galaxy

Athletes
NT Woods, Tiger

Holidays
NT Thanksgiving
c) **Hierarchical – Whole/Part:**

Concept or entity
- part
- subentity

Narrower term
“is a” broader term

Must be an integral part that cannot be taken out

- United Nations
  - NT UNICEF
- British Columbia
  - NT Vancouver
- Digestive system
  - NT Stomach
- Engineering
  - NT Electrical Engineering
Polyhierarchies

Sometimes a term can have two or more broader terms.

*Systems may or may not support this.*
Outline

1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
Sources for Terms

1. People as sources
2. Content/material to be indexed as sources
3. External sources
   ➢ Sources for concepts
   ➢ Sources for preferred wordings and non-preferred terms
People as sources

- Owner/manager of the controlled vocabulary
- Subject matter experts
- Users
- Taxonomist
Owner/manager of the controlled vocabulary determines:

- Some or all of top-level terms or facets
- Some sample terms

➢ More for concepts than for preferred wording

More often for commercial, publicly used databases and search services and products

- Based on strategic/business need
- Database product/service design
- Perceived customer needs/ market research
Subject matter experts (SMEs) determine:

- Specific terms and their relationships (BT/NT, RT)
- For both concepts and preferred wording

Subject matter experts

- Used for highly specialized/technical CVs
- For internally used CVs, usually internal employees, perhaps borrowed from other departments (e.g. engineers)
- Can be tasked with creating proto-taxonomy portions
Users determine:

- What the needs are, scope
- What term concepts need to be included
- For both concepts and preferred wording

User input can be gathered:

- From internal users for “enterprise” taxonomy projects through interviews and workshops
- From external, commercial, subscriber, or public users, after some content/service is made available
Gathering information from internal users:

- User interviews (in-person or phone)
  - For concepts and scope
- User card-sorting exercises
  - For hierarchical relationships and concepts

From internal or external users:

- Query logs/search logs (search engine use)
  - For both concepts and preferred wording
- Surveys/questionnaires, feedback forms
  - For both concepts and preferred wording
Taxonomists determine:

- In hierarchical or faceted taxonomy, upper level terms (2-3 levels deep)
- More for concepts than for preferred wording

Based on:

- “General knowledge” of knowing what’s important
- Past similar taxonomy projects worked on
- Familiarity with standard classifications (SIC/NAICS industries, UNPSC products, academic disciplines, Dewey decimal system, etc.)
Content to be indexed:
Primary source of concepts for terms, other than top terms

From samples of the content to be indexed:
- Articles
- Documents
- Web pages
- Content management system assets

Look for concepts within:
- Article titles and sub-article headings
- Document tables of content
- Web site navigation menu labels, site maps, Web page titles
- Image captions
- Listed products, services, goals, people-types, issues
- Existing metadata (keywords, titles, short description)
External sources:
For preferred format, rather than for concepts

Library of Congress Subject Headings (LCSH) and Name Authorities: http://authorities.loc.gov

- Search engines, such as Google, for comparative hit counts
- Specialized glossaries and thesauri – Taxonomy Warehouse: http://www.taxonomywarehouse.com
- Other published thesauri – American Society for Indexing http://www.asindexing.org/site/thesonet.shtml

*Be careful, don’t copy the thesaurus!*

- Public web sites of similar organizations/companies
- Industry standards, regulatory agencies
1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
1. Draft preliminary top terms based on owner requests and consensus of key users (interviews, workshop)
2. Build out taxonomy from terms from content sources and (if available) subject matter expert proposals
3. Supplement with specific terms requested by users, search log terms, etc.
4. Revise top terms if needed
5. Add more nonpreferred terms, from external sources

Bulk of new taxonomy creation work is in step #2
Gathering and Organizing Terms

- Build out taxonomy from terms from content sources
- Despite taxonomy management software, use Excel.

Varying approaches:

- For taxonomy to cover a corporate website or intranet
  - Comprehensive coverage is important, but content is unknown
  - Conduct a content audit for list of “harvested” terms
  - Put terms into hierarchy as a second step

- For a taxonomy to covering defined (published) document collection
  - Terms can be gathered into hierarchy from start
  - Consider document sets, titles, headings in hierarchy
<table>
<thead>
<tr>
<th>Link (URL)/Location</th>
<th>Content Type</th>
<th>Terms</th>
<th>Audience</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB_BioForm.rtf</td>
<td>blank form</td>
<td>Bio form, employee bio form, employee skills for business</td>
<td>HR personnel</td>
</tr>
<tr>
<td>BB_Certification_Application_Form.xl</td>
<td>form template with instructions</td>
<td>Black Belt certification application instructions</td>
<td>certification applicants</td>
</tr>
<tr>
<td>black belt project descriptions.htm</td>
<td>table of projects, leaders, and teams</td>
<td>Black Belt projects</td>
<td>all employees</td>
</tr>
<tr>
<td>BULeaders.htm</td>
<td>table of business leaders and teams</td>
<td>business unit leaders, business unit contacts</td>
<td>all employees</td>
</tr>
<tr>
<td>Business Assessment Process.ppt</td>
<td>training/information</td>
<td>business assessment, DMAIC roadmap, value</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Control Charts.htm</td>
<td>navigation(?) To charts</td>
<td>control charts</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>DMAIC_curriculum.xls</td>
<td>table of values</td>
<td>DMAIC curriculum</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>DOEFORM.xls</td>
<td>blank form</td>
<td>DOE (Design of Experiments) approval form</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>GB_faq.htm</td>
<td>FAQ</td>
<td>certification requirements, Green Belt program</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Information Sharing Meetings.doc</td>
<td>guidelines</td>
<td>meeting guidelines, information sharing meeting</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Leadership Training.htm</td>
<td>presentation</td>
<td>training presentation, leadership training, presentation</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>lean tools.jsp.html</td>
<td>resource information</td>
<td>lean thinking, lean tools, waste elimination, value</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>LEAN.htm</td>
<td>program summary</td>
<td>lean activities, lean programs</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Link_Spring_2002.pdf</td>
<td>newsletter</td>
<td>The Link, Transformation, Full Supply Chain, Culture, Organizational</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>MBB_Certification_Process.ppt</td>
<td>presentation</td>
<td>Master Black Belt Candidate Selection, Master Black Belt Certification</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>mission.htm</td>
<td>mission statement</td>
<td>mission statement</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>New Product &amp; Service Introduction</td>
<td>navigation</td>
<td>New Product &amp; Service Introduction, Design for Six Sigma (DFSS)</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>People.jsp.html</td>
<td>resource information</td>
<td>people, key players, key roles - master black belts, black belts, green</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Problem Solving Techniques.htm</td>
<td>navigation to resource info</td>
<td>Problem Solving Techniques (Brainstorming, Pareto Analysis, Voting)</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Program Life-Cycle Process.doc</td>
<td>policy</td>
<td>Program Life-Cycle Process, program management policy</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Program Life-Cycle Process</td>
<td>letter-style statement</td>
<td>Lean</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Reduce Variability - DMAIC.jsp.html</td>
<td>resource information</td>
<td>DMAIC, Reduce Variability - Define Phase, Measure Phase, Analyze Phase</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Six Sigma Acronyms.htm</td>
<td>resource information, glossary</td>
<td>acronyms, glossary</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Six Sigma in Government.pdf</td>
<td>article</td>
<td>Six Sigma in federal government case studies</td>
<td>all employees</td>
</tr>
<tr>
<td>Software Downloads.jsp.html</td>
<td>software download</td>
<td>software downloads, simulation software - Minitab, Process Simulator</td>
<td>Six Sigma participants</td>
</tr>
<tr>
<td>Solution Trees.htm</td>
<td>resource information, glossary</td>
<td>solution trees, diagramming solutions</td>
<td>Six Sigma participants</td>
</tr>
</tbody>
</table>
Draft taxonomy to cover a published collection of articles

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>AllLabels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme</td>
<td>Concept</td>
<td>Concept</td>
<td>Concept</td>
<td>Concept</td>
<td>altLabels</td>
</tr>
<tr>
<td>Level 1</td>
<td>Level 2</td>
<td>Level 3</td>
<td>Level 4</td>
<td></td>
<td>(IT domain unambiguous)</td>
</tr>
<tr>
<td>106 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power requirements</td>
<td>Emergency power</td>
<td></td>
</tr>
<tr>
<td>107 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>108 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>109 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power reliability requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>110 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power requirements</td>
<td>Data center critical loads</td>
<td></td>
</tr>
<tr>
<td>111 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power load requirements</td>
<td>Data center uninterruptible power supply requirements</td>
<td></td>
</tr>
<tr>
<td>112 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center power load requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>113 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Emergency power</td>
<td></td>
</tr>
<tr>
<td>114 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Data center backup generator maintenance</td>
<td></td>
</tr>
<tr>
<td>116 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Natural gas backup generators</td>
<td></td>
</tr>
<tr>
<td>117 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Diesel backup generators</td>
<td></td>
</tr>
<tr>
<td>118 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Bi-fuel backup generators</td>
<td></td>
</tr>
<tr>
<td>119 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>Modular power systems</td>
<td></td>
</tr>
<tr>
<td>120 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td>MPS</td>
<td></td>
</tr>
<tr>
<td>121 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power systems</td>
<td>Data center backup generators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>122 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Dirty power; Power disturbances</td>
<td></td>
</tr>
<tr>
<td>123 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Brownouts</td>
<td></td>
</tr>
<tr>
<td>124 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Voltage sags</td>
<td></td>
</tr>
<tr>
<td>125 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Voltage surges</td>
<td></td>
</tr>
<tr>
<td>126 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Voltage spikes</td>
<td></td>
</tr>
<tr>
<td>127 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power quality problems</td>
<td>Frequency variations</td>
<td></td>
</tr>
<tr>
<td>128 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>129 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td>In-line power conditioners</td>
<td></td>
</tr>
<tr>
<td>130 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td>Uninterruptible power supplies</td>
<td></td>
</tr>
<tr>
<td>131 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td>UPS</td>
<td></td>
</tr>
<tr>
<td>132 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td>Monolithic UPS systems</td>
<td></td>
</tr>
<tr>
<td>133 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>134 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>135 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>136 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>137 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>138 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>139 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>140 Data Centers</td>
<td>Data center power systems</td>
<td>Data center power conditioning</td>
<td>Power conditioning technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------</td>
<td>---------------------------------</td>
<td>---------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Agriculture and Rural</td>
<td>Agriculture and Rural Development</td>
<td>Agribusiness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Agriculture and Rural</td>
<td></td>
<td>Agribusiness Innovation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Development</td>
<td></td>
<td>Agribusiness Innovation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Agriculture and Rural</td>
<td></td>
<td>Agribusiness Investment Climate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Development</td>
<td></td>
<td>Agricultural Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Agriculture and Rural</td>
<td></td>
<td>Agricultural Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Development</td>
<td></td>
<td>Agricultural Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Agriculture and Rural</td>
<td></td>
<td>Agricultural Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Development</td>
<td></td>
<td>Agricultural Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Value Chain</td>
<td>Agricultural Value Chain</td>
<td>Agricultural Value Chain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>Agricultural Value Chain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Agriculture Finance</td>
<td>Agricultural Finance</td>
<td>Agricultural Finance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Agricultural Policy</td>
<td>Agricultural Policy</td>
<td>Agricultural Policy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Agricultural Laws and</td>
<td>Agricultural Laws and</td>
<td>Agricultural Laws and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Regulations</td>
<td>Regulations</td>
<td>Regulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Agricultural Water</td>
<td>Agricultural Water</td>
<td>Agricultural Water Records</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Management</td>
<td>Agriculture Water Management</td>
<td>Agricultural Water Records</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Irrigation and Drainage</td>
<td>Irrigation and Drainage</td>
<td>Irrigation and Drainage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>Drainage</td>
<td>Drainage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>Subsurface Drainage</td>
<td>Subsurface Drainage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>Drainage Water Quality</td>
<td>Drainage Water Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>Irrigation Agronomy</td>
<td>Irrigation Agronomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>Irrigation Efficiencies</td>
<td>Irrigation Efficiencies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>Irrigation Methods</td>
<td>Irrigation Methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>Drip Irrigation</td>
<td>Drip Irrigation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>On-Demand Irrigation</td>
<td>On-Demand Irrigation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>Sprinkler Irrigation</td>
<td>Sprinkler Irrigation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>Surface Irrigation</td>
<td>Surface Irrigation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Gathering and Organizing Terms

Subject-matter-expert proposed taxonomy and revisions

<table>
<thead>
<tr>
<th>Transport</th>
<th>Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Transport</td>
<td>Mean of Transport, Transport</td>
</tr>
<tr>
<td>3 Modes of Transport</td>
<td>Change from Ports and Waterways</td>
</tr>
<tr>
<td>4 Air Transport</td>
<td>Changed from Railways</td>
</tr>
<tr>
<td>5 Marine Transport</td>
<td>Change from Roads and Highways</td>
</tr>
<tr>
<td>6 Rail Transport</td>
<td></td>
</tr>
<tr>
<td>7 Road Transport</td>
<td></td>
</tr>
<tr>
<td>8 Non-motorized Transport</td>
<td>moved from under Transport and Logistics</td>
</tr>
<tr>
<td>9 Multimodal Transport</td>
<td>services</td>
</tr>
<tr>
<td>10 Transport Infrastructure</td>
<td>add this section from Tanase’s proposal</td>
</tr>
<tr>
<td>11 Airports</td>
<td></td>
</tr>
<tr>
<td>12 Railways</td>
<td>railroads, rail systems, railways</td>
</tr>
<tr>
<td>13 Waterways</td>
<td></td>
</tr>
<tr>
<td>14 Ports</td>
<td></td>
</tr>
<tr>
<td>15 Roads and Highways</td>
<td></td>
</tr>
<tr>
<td>16 Highways</td>
<td></td>
</tr>
<tr>
<td>17 Urban Roads</td>
<td></td>
</tr>
<tr>
<td>18 Rural Roads</td>
<td></td>
</tr>
<tr>
<td>19 Transport and Logistics Services</td>
<td>also under Trade &gt; Trade Facilitation and</td>
</tr>
<tr>
<td>20 Freight Services</td>
<td>Logistics</td>
</tr>
<tr>
<td>21 Passenger Services</td>
<td>changed from Freight Transport</td>
</tr>
<tr>
<td>22 Transport Integrators</td>
<td>changed from Passenger Transport</td>
</tr>
<tr>
<td>23 Transport Logistics Providers</td>
<td></td>
</tr>
<tr>
<td>24 Transport and Social Responsibility</td>
<td>freight forwarders, travel agencies</td>
</tr>
<tr>
<td>25 Gender and Transport</td>
<td>port services, transport terminals</td>
</tr>
<tr>
<td>26 Transport and HIV-Aids</td>
<td></td>
</tr>
<tr>
<td>27 Universal Access in Transport</td>
<td></td>
</tr>
<tr>
<td>28 Transport and Development</td>
<td>Transport Economics; Transportation</td>
</tr>
<tr>
<td>29 Transport and Poverty</td>
<td></td>
</tr>
<tr>
<td>30 Transport Impact on Trade</td>
<td></td>
</tr>
<tr>
<td>31 Transport and Economic Geography</td>
<td></td>
</tr>
<tr>
<td>32 Transport and Urban Development</td>
<td></td>
</tr>
<tr>
<td>33 Transport and the Environment</td>
<td></td>
</tr>
</tbody>
</table>
### Comparing tables of contents of multiple books side-by-side

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Introduction</td>
<td>Introduction</td>
<td>Anatomic Reference Systems</td>
<td>Body Organization Illustrated</td>
</tr>
<tr>
<td>Body Structure Terms</td>
<td>Body Structure Terms</td>
<td>Body Structure Terms</td>
<td>Structures of the Body</td>
<td>Terminology</td>
</tr>
<tr>
<td>Levels of Body Organization</td>
<td>Levels of Organization</td>
<td>Levels of Body Organization</td>
<td>Body Systems and Related Organs</td>
<td></td>
</tr>
<tr>
<td>Cells</td>
<td>Cell</td>
<td>Cells</td>
<td>Cells</td>
<td></td>
</tr>
<tr>
<td>Cell Membrane and Cytoplasm</td>
<td>Cell Membrane and Cytoplasm</td>
<td>Clinical Case Study: Diagnosis</td>
<td>Tissues</td>
<td>Tissues</td>
</tr>
<tr>
<td>Cell Nucleus</td>
<td>Nucleus</td>
<td></td>
<td>Threads and Systems</td>
<td></td>
</tr>
<tr>
<td>Tissues</td>
<td>Tissue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organs</td>
<td>Organ</td>
<td></td>
<td>Glands</td>
<td>Organs and Systems</td>
</tr>
<tr>
<td>Glands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Systems</td>
<td>System</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism Level</td>
<td>Organism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anatomical Position</td>
<td>Anatomical Position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planes of the Body</td>
<td>Planes of the Body</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Cavities</td>
<td>Body Cavities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominopelvic Divisions</td>
<td>Abdominopelvic Divisions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominopelvic Quadrants</td>
<td>Quadrants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominopelvic Regions</td>
<td>Regions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Directional Terms</td>
<td>Directional Terms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine and Spinal Divisions</td>
<td>Spine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anatomical Terms Introduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Structure Word Elements</td>
<td>Medical Word Elements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Structure Term Roots and Combining Forms</td>
<td>Word Parts Focus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Gathering and Organizing Terms**
After Draft Taxonomy is Built

- Enter into taxonomy management system and refine.
- Do user testing with offline sample content, to test functionality in retrievability. Make adjustments as needed.
- Do testing with offline sample content to test suitability for use in indexing/tagging. Add more terms as needed.
- Have stakeholders and/or subject matter experts review it.
- Develop taxonomy style and maintenance guidelines as part of a governance plan.
- Develop indexing guidelines.

When done, deliver for implementation: CSV, XML, or RDF files.
Outline

1. Initial considerations
2. Definitions and determining the taxonomy type
3. Deciding on the taxonomy scope
4. Taxonomy terms and relationships
5. Taxonomy term sources
6. Gathering and organizing terms
7. Case example: Cengage Learning
Controlled vocabularies for research database products (Gale) not built from scratch

Taxonomies for learning (textbook content) products built from mostly scratch

– By a mix of internal and external (contract) taxonomists

– Relying on content-to-be-indexed as sources for terms
  - textbook tables of contents
  - back-of-the-book index terms
  - learning objectives

With some looking into external sources: college and university curricula
Book tables of contents source issues:

- Hierarchical but not according to taxonomy hierarchical relationship standards.
- Different textbooks for the same course take different approaches with varying main topics.
- Skills-based disciplines sometimes present divergent topics together in project-based instruction.

Back-of-the-book index source issues:

- Main entries might be good taxonomy terms, but subentries are not.
- Index’s typical inclusion of more than one term for the same concept can be confusing.
Questions?

Heather Hedden  
Senior Vocabulary Editor  
Cengage Learning  
20 Channel Center St., Boston, MA 02210  
www.cengage.com

(o) 617-757-8211  |  (m) 978-467-5195  
(e) Heather.Hedden@cengage.com