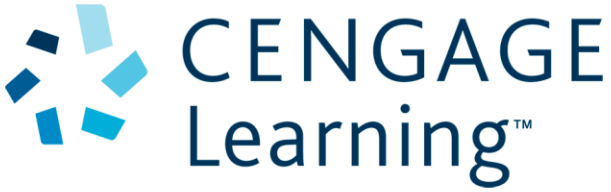
Taxonomies

Everything You Need to Know to Start a Taxonomy from Scratch

Heather Hedden
Senior Vocabulary Editor, Cengage Learning

SLA 2014, Vancouver, BC
June 8, 2014

Sponsored by the SLA Taxonomy Division



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

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

Initial Considerations

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

Initial Considerations

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.

Initial Considerations

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)

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

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

Definitions

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

Types

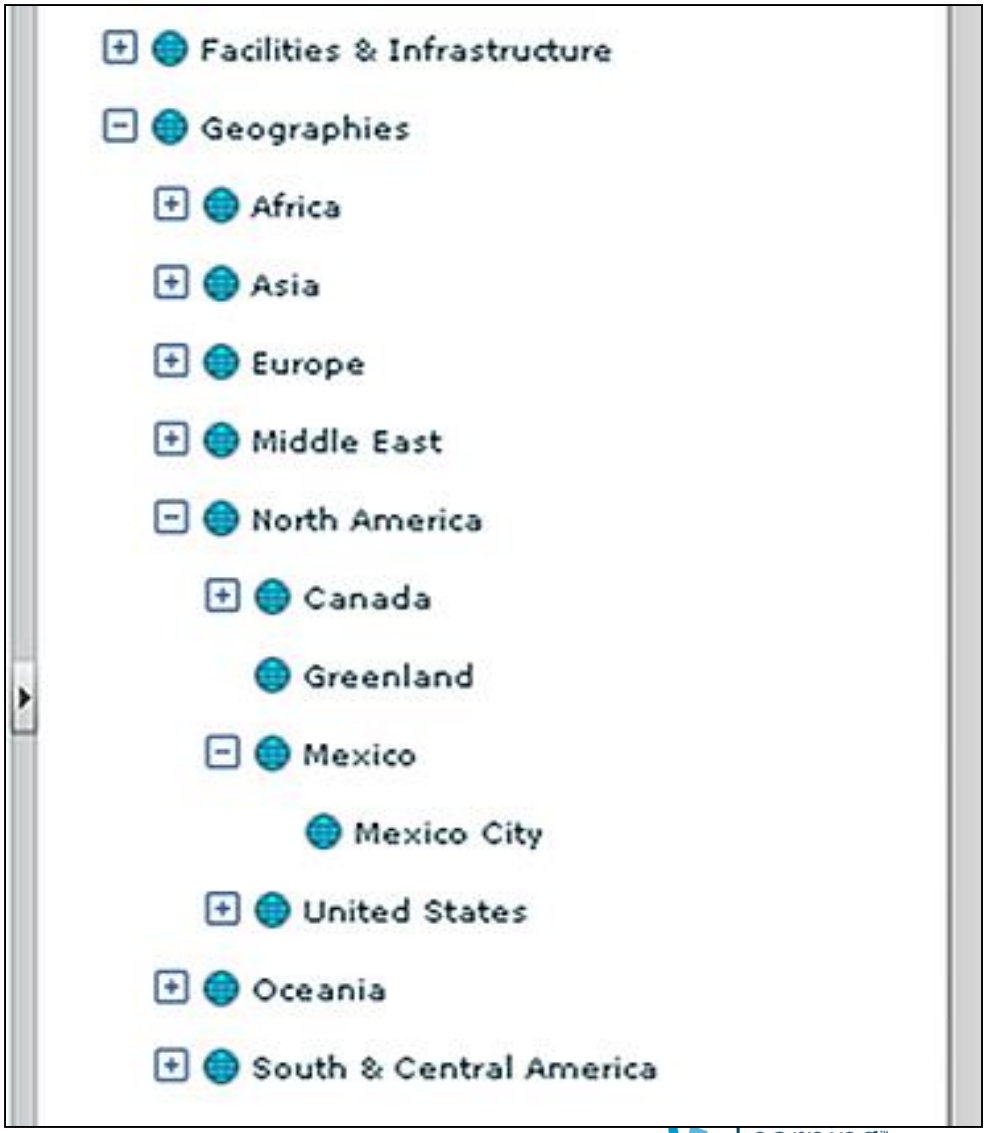
- Term List
- Synonym Ring
- Authority File
- Taxonomy (definition #1)
  - Hierarchical taxonomy
  - Faceted taxonomy
- Thesaurus
- Ontology

Types

Hierarchical Taxonomy –

Has broader term/narrower term relationships that include all terms to create a hierarchical structure

1. Example: Expandable, on same page



Types

Hierarchical Taxonomy

2. Example: Separate page for each level of hierarchy

Find business listings for Arts & Entertainment.

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, ...

Types

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

Types

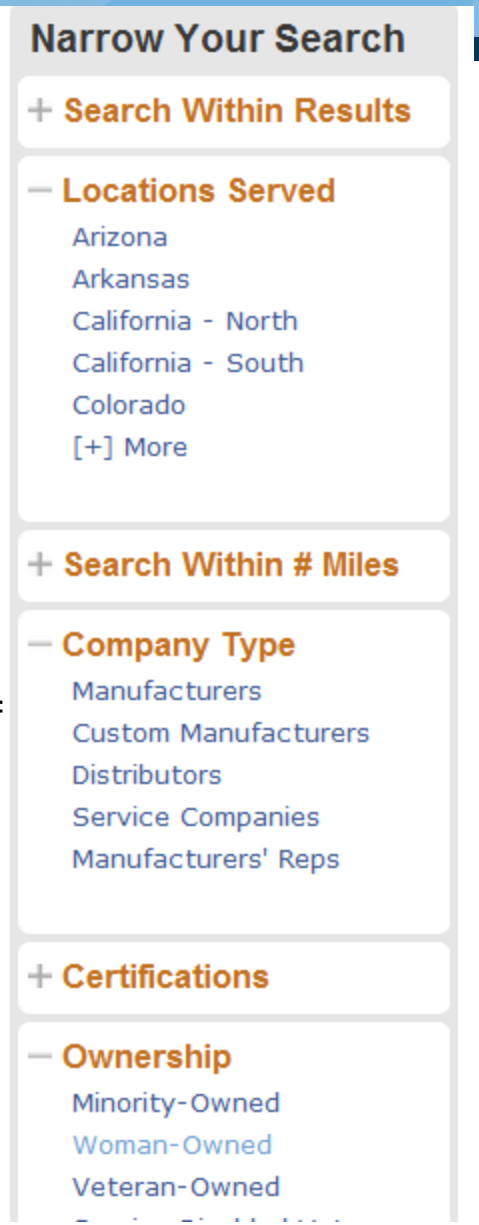
Facets

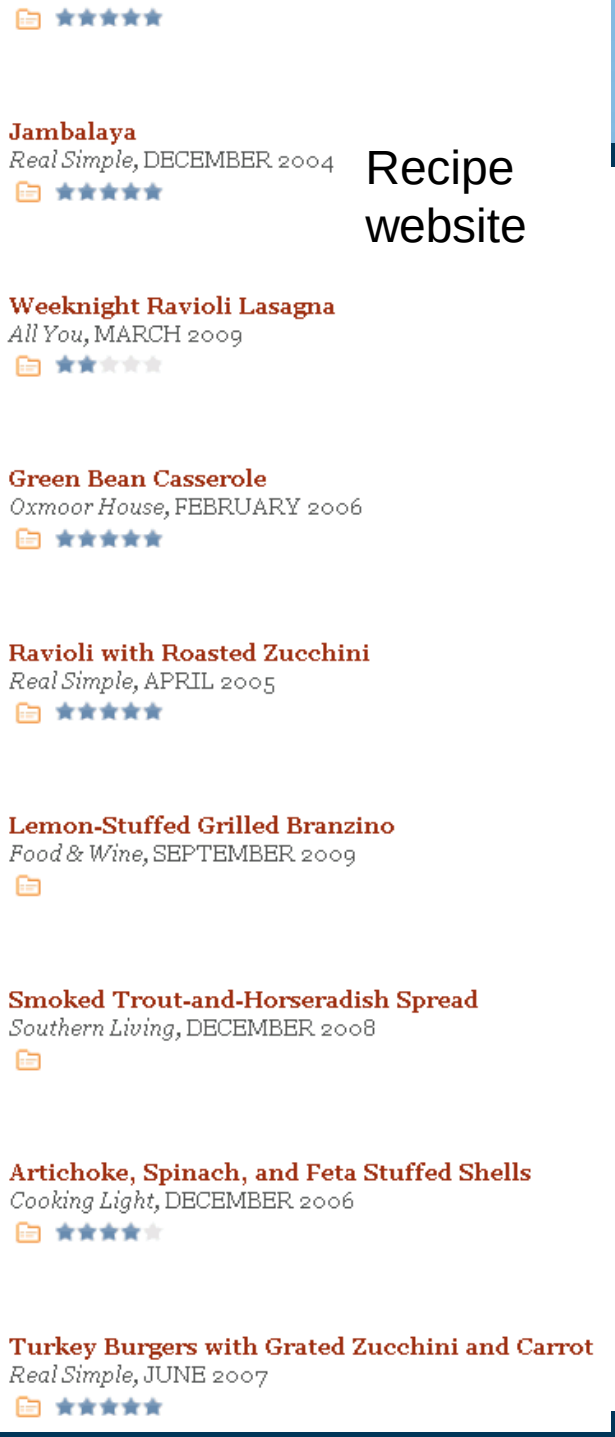
- For serving faceted classification, which allows the assignment of multiple classifications to an object
- A “dimension” of a query; a type of concept; domain of content
- Intended for searching with multiple terms in combination (post-coordination), one from each facet:
  to “limit by” filters

Types

Faceted taxonomy examples

Directory of companies





Recipe website

Types

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

materials acquisitions

UF	acquisitions (of materials)
	library acquisitions
BT	collection development
NT	accessions
	approval plans
	gifts and exchanges
	materials claims
	materials orders
	subscriptions
RT	book vendors
	jobbers
	subscription agencies
	subscription cancellations

Types: Summary

Pick List	Synonym Ring	Authority File	Taxonomy	Thesaurus	Ontology
Ambiguity control	Synonym control	Ambiguity control Synonym control (preferred & non-preferred terms)	Ambiguity control (Synonym control) Hierarchical Relationships	Ambiguity control Synonym control Hierarchical Relationships Associative Relationships	Ambiguity control (Synonym control) Semantic Relationships Classes

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

Taxonomy Scope, Breadth and Depth

Deciding the Taxonomy Scope:

- Content scope
- Subject area scope
- Term types to include
- Usage
- Breadth and depth

Taxonomy Scope, 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

Taxonomy Scope, Breadth and Depth

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

Taxonomy Scope, Breadth and Depth

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

Taxonomy Scope, Breadth and Depth

Scope of Term Types
(could also be considered as metadata or “facets”)

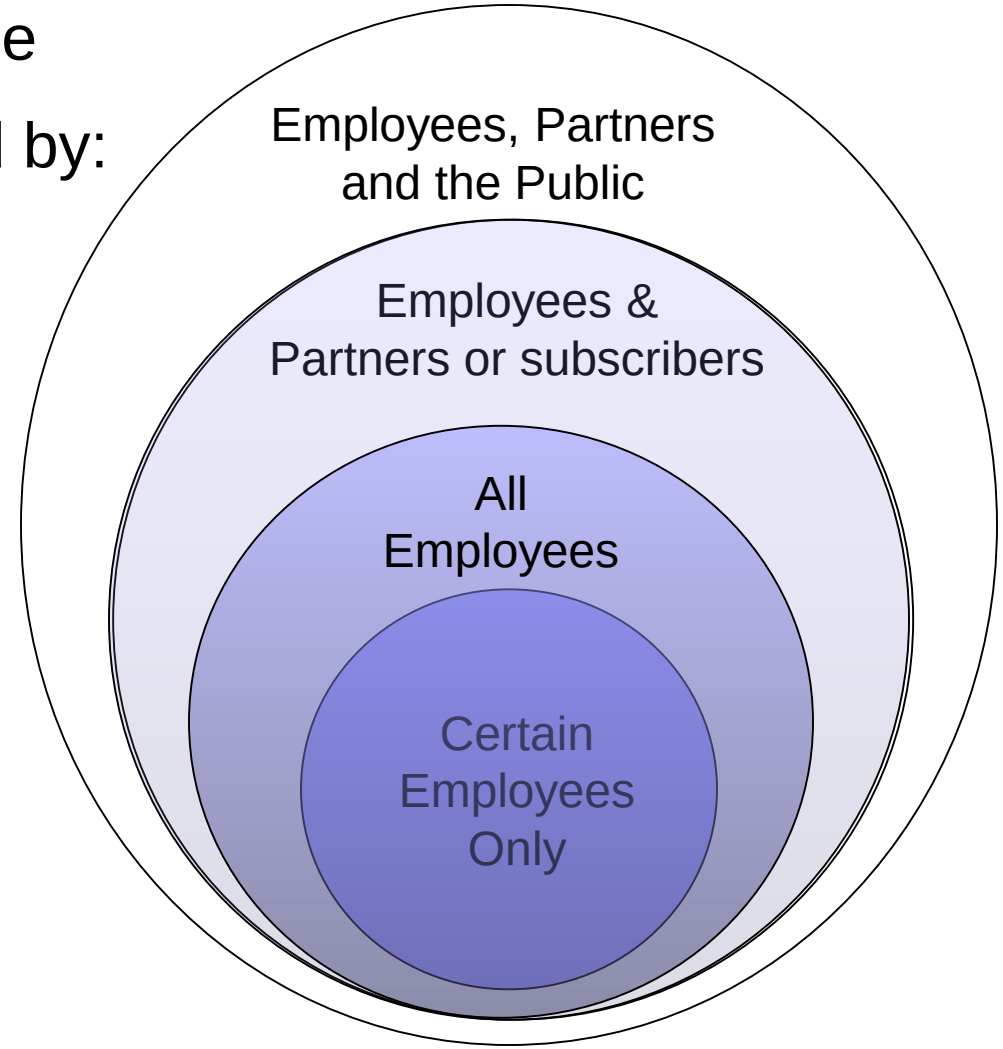
Term types to be included can be limited to some or all of:

- 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:



- Public Web Site
- Extranet or restricted access website
- Intranet
- Intranet restricted sections

Taxonomy Scope, Breadth and Depth

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

Drug trials	**Product testing** and **Drugs**
CRM Software	**Customer Relations Management** and **Software**

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

Taxonomy Terms and Relationships

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?

Taxonomy Terms and Relationships

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**

Taxonomy Terms and Relationships

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

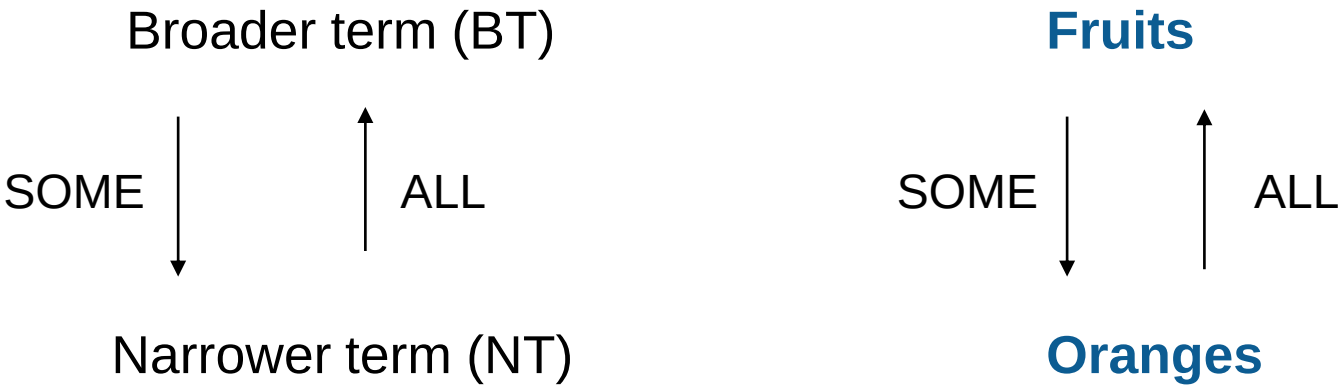
antonyms (for characteristics): Softness USE Hardness

narrower terms and instances that are not preferred terms: Hydroelectric power plants USE Power plants

Taxonomy Terms and Relationships

3. **Structuring the hierarchy**

Reciprocal (bi-directional) relationships, but asymmetrical



Fruits NT **Oranges** **Oranges** BT **Fruits**

Three types:

a) Generic - Specific

b) Common noun - Proper noun

c) Whole – Part

Taxonomy Terms and Relationships

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**

Taxonomy Terms and Relationships

a) **Hierarchical - Instance:**

Common noun

– Proper noun

Narrower term
"is a" broader term

Smartphones

NT **Samsung Galaxy**

Athletes

NT **Woods, Tiger**

Holidays

NT **Thanksgiving**

Taxonomy Terms and Relationships

c) **Hierarchical – Whole/Part:**

Concept or entity

– part

– subentity

Narrower term
"is a" broader term

Must be an integral part
that cannot taken out

United Nations

NT **UNICEF**

British Columbia

NT **Vancouver**

Digestive system

NT **Stomach**

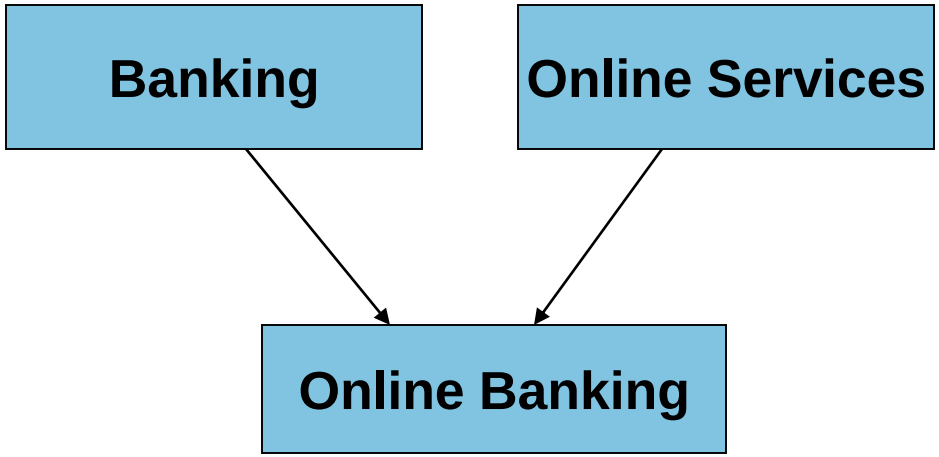
Engineering

NT **Electrical Engineering**

Taxonomy Terms and Relationships

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

Sources for Terms

People as sources

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

Sources for Terms: People

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

Sources for Terms: People

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

Sources for Terms: People

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

Sources for Terms: People

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

Sources for Terms: People

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.)

Sources for Terms: Content

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)

Sources for Terms: External

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

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

Gathering and Organizing Terms

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

Gathering and Organizing Terms

Term extraction for a taxonomy to cover a corporate website

Link (URL)/Location	Content Type	Terms	Audience
BB_BioForm.rtf	blank form	Bio form, employee bio form, employee skills f	HR personnel?
BB_Cerification_Application_Form.x	form template with instructi	Black Belt certification application instructions	certification applicant
black belt project descriptions.htm	table of projects, leaders, a	Black Belt projects	all employees
BULeaders.htm	table of business leaders a	business unit leaders, business unit contacts	all employees
Business Assessment Process.ppt	training/information	business assessment, DMAIC roadmap, value	Six Sigma participan
Control Charts.htm	navigation(?) To charts	control charts	Six Sigma participan
DMAIC_curriculum.xls	table of values	DMAIC curriculum	Six Sigma participan
DOEFORM.xls	blank form	DOE (Design of Experiments) approval form	Six Sigma participan
GB_faq.htm	FAQ	certification requirements, Green Belt program	Six Sigma participan
Information Sharing Meetings.doc	guidelines	meeting guidelines, information sharigng meeti	Six Sigma participan
Leadership Training.htm	presentation	training presentation, leadership training, prese	Six Sigma participan
lean tools.jsp.html	resource information	lean thinking, lean tools, waste elimination, valu	Six Sigma participan
LEAN.htm	porgram summary	lean activities, lean programs	
Link_Spring_2002.pdf	newlestter	The Link, Transformation, Full Supply Chain, Culture, Organizational	
MBB_Certification_Process.ppt	presentation	Master Black Belt Candidate Selection, Master Black Belt Certificatio	
mission.htm	mission statement	mission statement	
New Product & Service Introduction.	navigation	New Product & Service Introduction; Design for Six Sigma (DFSS)	
People.jsp.html	resource information	people, key players, key roles - master black belts, black belts, gree	
Problem Solving Techniques.htm	navigation to resource info	Problem Solving Techniques (Brainstorming, Pareto Analysis, Voting	
Program Life-Cycle Process.doc	policy	Program Life-Cycle Process, program management policy	
Program Life-Cycle Process	letter-style statement	Lean	
Reduce Variability - DMAIC.jsp.html	resource information	DMAIC, Reduce Variability - Define Phase, Measure Phase, Analyze	
Six Sigma Acronyms.htm	resource information, gloss	acronyms, glossary	
Six Sigma in Government.pdf	article	Six Sigma in federal government case studies	all employees
Software Downloads.jsp.html	software download	software downloads, simulation software - Minitab, Process Simulato	
Solution Trees.htm	resource information, gloss	solution trees, diagraming solutions	

Gathering and Organizing Terms

Draft taxonomy to cover a published collection of articles

	A	B	C	D	E	F
1	**Scheme**	**Concept**	**Concept**	**Concept**	**Concept**	**altLabels**
2		**Level 1**	**Level 2**	**Level 3**	**Level 4**	**(IT domain unambiguous)**
106	Data Centers	Data center power systems				
107	Data Centers	Data center power systems	Data center power requirements			
108	Data Centers	Data center power systems	Data center power requirements	Data center power reliability requirements		
109	Data Centers	Data center power systems	Data center power requirements	Data center power quality requirements		
110	Data Centers	Data center power systems	Data center power requirements	Data center power load requirements		
111	Data Centers	Data center power systems	Data center power requirements	Data center power load requirements	Data center critical loads	
112	Data Centers	Data center power systems	Data center power requirements	Data center power load requirements	Data center uninterruptible power supply requirements	
113	Data Centers	Data center power systems	Data center backup generators			Emergency power
114	Data Centers	Data center power systems	Data center backup generators	Data center backup generator capacity planning		
115	Data Centers	Data center power systems	Data center backup generators	Data center backup generator maintenance		
116	Data Centers	Data center power systems	Data center backup generators	Data center backup generator types		
117	Data Centers	Data center power systems	Data center backup generators	Data center backup generator types	Natural gas backup generators	
118	Data Centers	Data center power systems	Data center backup generators	Data center backup generator types	Diesel backup generators	
119	Data Centers	Data center power systems	Data center backup generators	Data center backup generator types	Bi-fuel backup generators	
120	Data Centers	Data center power systems	Data center backup generators	Data center backup generator types	Modular power systems	MPS
121	Data Centers	Data center power systems	Data center power conditioning			
122	Data Centers	Data center power systems	Data center power conditioning	Power quality problems		Dirty power;Power disturbances
123	Data Centers	Data center power systems	Data center power conditioning	Power quality problems	Brownouts	
124	Data Centers	Data center power systems	Data center power conditioning	Power quality problems	Voltage sags	
125	Data Centers	Data center power systems	Data center power conditioning	Power quality problems	Voltage surges	Power surges;Transients;Transient
126	Data Centers	Data center power systems	Data center power conditioning	Power quality problems	Voltage spikes	High voltage spikes
127	Data Centers	Data center power systems	Data center power conditioning	Power quality problems	Frequency variations	
128	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies		
129	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	In-line power conditioners	
130	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	UPS
131	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	
132	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	Monolithic UPS systems
133	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	
134	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	
135	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	
136	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	Rotary UPS
137	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	Battery rooms;Storage battery
138	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	
139	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	VFD;Standby UPS;SPS;Offline UPS
140	Data Centers	Data center power systems	Data center power conditioning	Power conditioning technologies	Uninterruptible power supplies	VI;Line interactive UPS

Gathering and Organizing Terms

Draft taxonomy to cover a published collection of articles

	A	B	C	D	E	F
1	**Level1**	**Level2**	**Level3**	**Level4**	**Level5**	**Synonyms**
2						
3	Agriculture and Rural Development					
4		Agribusiness				
5			Agribusiness Innovation			
6			Agribusiness Investment Climate			Agribusiness Regulation
7			Agribusiness Transport			
8			Agribusiness and Land			
9			Agricultural Markets and Risks			
10				Agricultural Risk Management		
11				Agricultural Trade		
12			Agricultural Value Chain			Agricultural Supply Chain, Agribusiness Supply Chain, Agribusiness Value Chain, Agribusiness Competitiveness
13		Agricultural Finance				Agrifinance, Agriculture Finance
14		Agricultural Policy				Agricultural Policy, Strategy and Institutions
15			Agricultural Laws and Regulations			
16				Irrigation Regulation		
17			Agricultural Statistics			
18		Agricultural Water Management				Agricultural Water, Water and Agriculture
19			Irrigation and Drainage			Irrigated Agriculture
20				Drainage		
21					Surface Drainage	
22					Subsurface Drainage	
23				Drainage Water Quality		
24				Irrigation Agronomy		Crop Water Requirements
25				Irrigation Efficiencies		Irrigation Efficiency, Efficient Irrigation
26				Irrigation Methods		
27					Drip Irrigation	
28					On-Demand Irrigation	
29					Sprinkler Irrigation	
30					Surface Irrigation	

Gathering and Organizing Terms

Subject-matter-expert proposed taxonomy and revisions

Key:
Red - Added
Blue - Moved
Purple - Changed wording

2	**Transport**		Mobility
3	Modes of Transport		Means of Transport, Transport
4	Air Transport		
5	Marine Transport	Change from Ports and Waterways	
6	Rail Transport	Changed from Railways	
7	Road Transport	Change from Roads and Highways	
8	Non-motorized Transport		
9	Multimodal Transport	moved from under Transport and Logistics services	
10	Transport Infrastructure	add this section from Tanase's proposal	
11	Airports		
12	Railways	change from Railway Infrastructure	Railroads, Rail Systems, Railw
13	Waterways		
14	Ports		
15	Roads and Highways		
16	Highways		
17	Urban Roads		
18	Rural Roads		
19	Transport and Logistics Services	also under Trade > Trade Facilitation and Logistics	
20	Freight Services	changed from Freight Transport	
21	Passenger Services	changed from Passenger Transport	
22	Transport Integrators		freight forwarders, travel agen
23	Transport Logistics Providers		Port Services, Transport Termi
24	Transport and Social Responsibility		
25	Gender and Transport		
26	Transport and HIV-Aids		
27	Universal Access in Transport		
28	Transport and Development		Transport Economics; Transpo
29	Transport and Poverty		
30	Transport Impact on Trade		
31	Transport and Economic Geography		
32	Transport and Urban Development		
33	Transport and the Environment		

Gathering and Organizing Terms

Comparing tables of contents of multiple books side-by-side

Taxonomy terms	Medical Terminology Systems A Body Systems Approach, 7 ed.,	Unlocking Medical Terminology, 2nd ed.,	Medical Terminology For Health Professions, 6th ed., Ann	Medical Terminology: A Living Language,5th ed., Bonnie F.
Human Body Structure	4. Body Structure	2. The Human Body in Heal	2. The Human Body in Health and	2. Body Organization
		3. Cells, Tissues, and Cancer	Vocabulary Related to the Human Body in Health and Disease	
		Anatomy and Physiology Overview		Body Organization at a Glance
	Introduction		Anatomic Reference Systems	Body Organization Illustrated
Body Structure Terms	Body Structure Terms			Terminology
Levels of Body Organization	Levels of Organization		Structures of the Body	Levels of Body Organization
Cells	Cell		Cells	Cells
Cell Membrane and Cytoplasm	Cell Membrane and Cytoplasm			
Cell Nucleus	Nucleus	Clinical Case Study: Diagnosis		
Tissues	Tissue		Tissues	Tissues
Organs	Organ			Organs and Systems
Glands			Glands	
Body Systems	System		Body Systems and Related Organs	
Organism Level	Organism			
				Body
Anatomical Position	Anatomical Position			
Planes of the Body	Planes of the Body			Body Planes
Body Cavities	Body Cavities			Body Cavities
Abdominopelvic Divisions	Abdominopelvic Divisions			
Abdominopelvic Quadrants	Quadrants			
Abdominopelvic Regions	Regions			Body Regions
Body Directional Terms	Directional Terms			Directional and Positional T
Spine and Spinal Divisions	Spine			
Genetics			Genetics	
Anatomical Terms Introduction				Anatomical Terms
Body Structure Word Elements	Medical Word Elements	Word Parts Focus		
Body Structure Term Roots and Combining Forms		Getting to the Root of it		

After Draft Taxonomy is Built

- Enter into taxonomy management system and tweak.
- 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.
- Develop taxonomy style and maintenance guidelines as part of a governance plan.
- Develop indexing guidelines.
- 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

Cengage Learning example

- 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

CENGAGE Learning

Cengage Learning example

- 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.

CENGAGE Learning™

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

CENGAGE Learning