User Interfaces with Taxonomies

Heather Hedden

Principal, Hedden Information Management

www.hedden-information.com

About Heather Hedden

- Taxonomy consultant, Hedden Information Management
- Taxonomy consultant, Earley & Associates
- Instructor, Continuing Education, Simmons College Graduate School of Library & Information Science
- Formerly, taxonomist at Viziant Corporation and Thomson Learning (Gale)
- Author, *The Accidental Taxonomist* (Information Today, Inc., May 2010), www.accidental-taxonomist.com

Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Definitions

- Taxonomies (in general)/knowledge organization systems
- Controlled Vocabulary
- Synonym Ring
- Taxonomy (hierarchical)
- Thesaurus
- Ontology

Definitions

Taxonomy

1. Any of various forms of a set of designated terms for organizing information or knowledge
   - a “knowledge organization system”
2. A specific type of knowledge organization system, whereby the terms are structured into a hierarchy (or multiple hierarchies) or categories

Knowledge Organization System Definitions

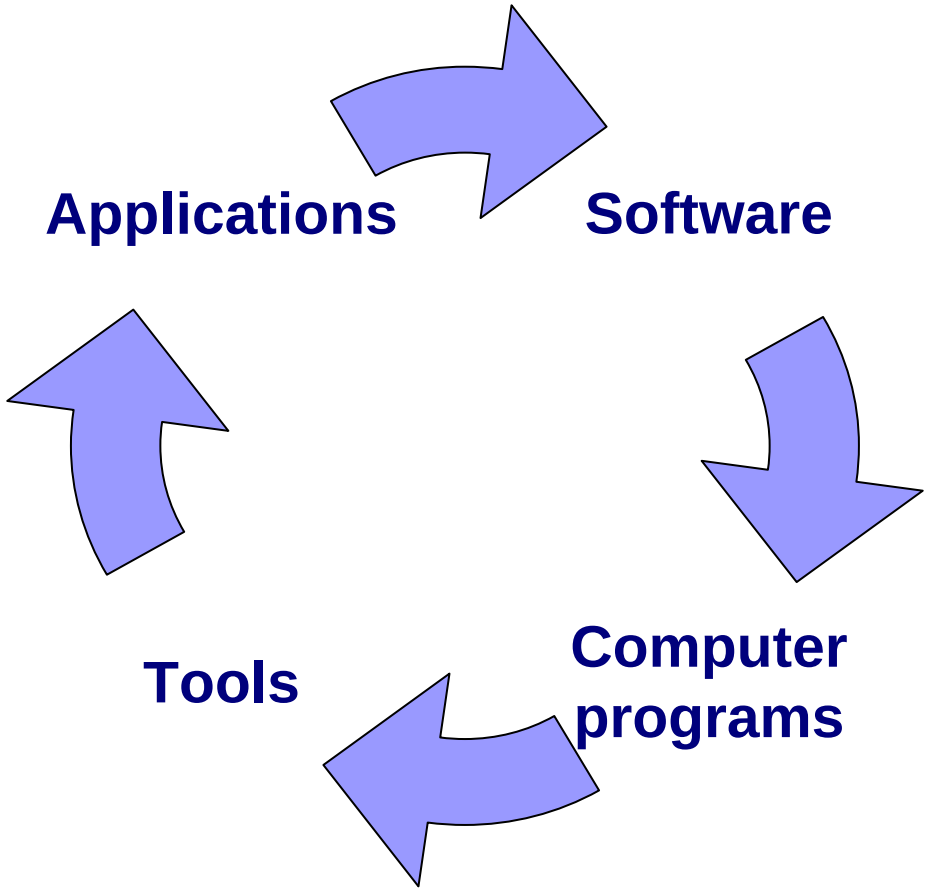
Controlled Vocabulary

- An authoritative, restricted list of terms (words or phrases) mainly used for indexing/tagging content to support information retrieval
- Controlled in who and when new terms can be added
- Usually includes equivalent variant terms (synonyms, etc.) for each “official” term
- May or may not have structured relationships between terms

Knowledge Organization System Definitions

Synonym ring

- A controlled vocabulary with synonyms or near-synonyms for each concept
- No designated “official” term: All terms are equal and point to each other, as in a ring.
- Also called syn-set



Knowledge Organization System Definitions

Taxonomy

- A controlled vocabulary with broader term/narrower term (hierarchical) relationships between terms
- All terms have hierarchical relationships, thus resulting in an overall hierarchical structure to the taxonomy (even if only 2 levels)
- Emphasizes categories and classification
- May or may not include equivalent synonyms/variants

Taxonomy example

Top Level Headings

- Business and industry
- Economics and finance
- Education and skills
- Employment, jobs and careers
- Environment
- Government, politics and public administration
- Health, well-being and care
- Housing
- Information and communication
- International affairs and defense
- Leisure and culture
- Life in the community
- People and organizations
- Public order, justice and rights
- Science, technology and innovation
- Transport and infrastructure

Leisure and culture

. Arts and entertainment venues
. . Museums and galleries
. Children's activities
. Culture and creativity
. . Architecture
. . Crafts
. . Heritage
. . Literature
. . Music
. . Performing arts
. . Visual arts
. Entertainment and events
. Gambling and lotteries
. Hobbies and interests
. Parks and gardens
. Sports and recreation
. . Team sports
. . . Baseball
. . . Basketball
. . . Football
. . Water sports
. . Winter sports
. Sports and recreation facilities
. Tourism
. . Passports and visas
. Young people's activities

Knowledge Organization System Definitions

Thesaurus

- A controlled vocabulary that has standard structured relationships between terms
  - Hierarchical: broader term/narrower term (BT/NT)
  - Associative: related terms (RT)
  - Equivalence: preferred term (“use for” or “used for”)/non-preferred term (use) (USE/UF)
- Also supports notes, such as scope notes (SN), for terms, as needed
- “Thesaurus” is most often used for controlled vocabularies used in human indexing for literature retrieval.
- Created according to ANSI/NISO Z39.19 (2005) or ISO 2788 (1986)

Thesaurus term entry example

materials acquisitions

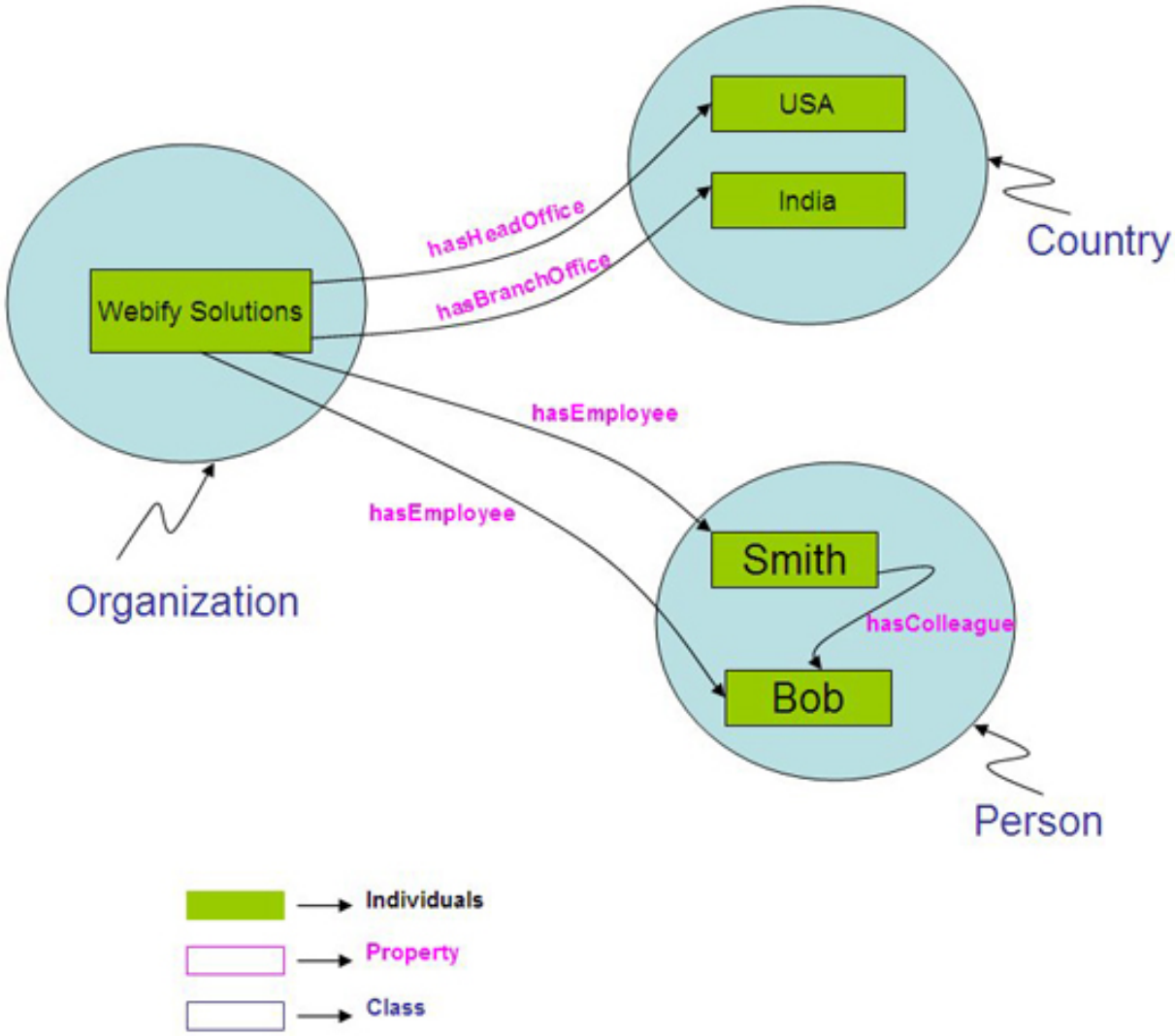
UF	acquisitions (of materials)
	library acquisitions
BT	collection development
NT	accessions
	approval plans
	gifts and exchanges
	materials claims
	materials orders
	subscriptions
RT	book vendors
	jobbers
	subscription agencies
	subscription cancellations

Knowledge Organization System Definitions

Ontology

- A complex type of thesaurus, in which terms have specified attributes and classes, and relationships are further specified
- Relationships contain meaning, are “semantic”
- Computer-readable according to W3C guidelines of the OWL Web Ontology Language Guide
- A form of “knowledge representation” that defines a domain of knowledge
- May be an end in itself, rather than a means to an end of information retrieval.

Ontology example



Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Taxonomy Purposes & Benefits

1. Controlled vocabulary:

   Brings together different wordings (synonyms) for the same concept

   - Helps people search for information by different names

2. Hierarchical taxonomy:

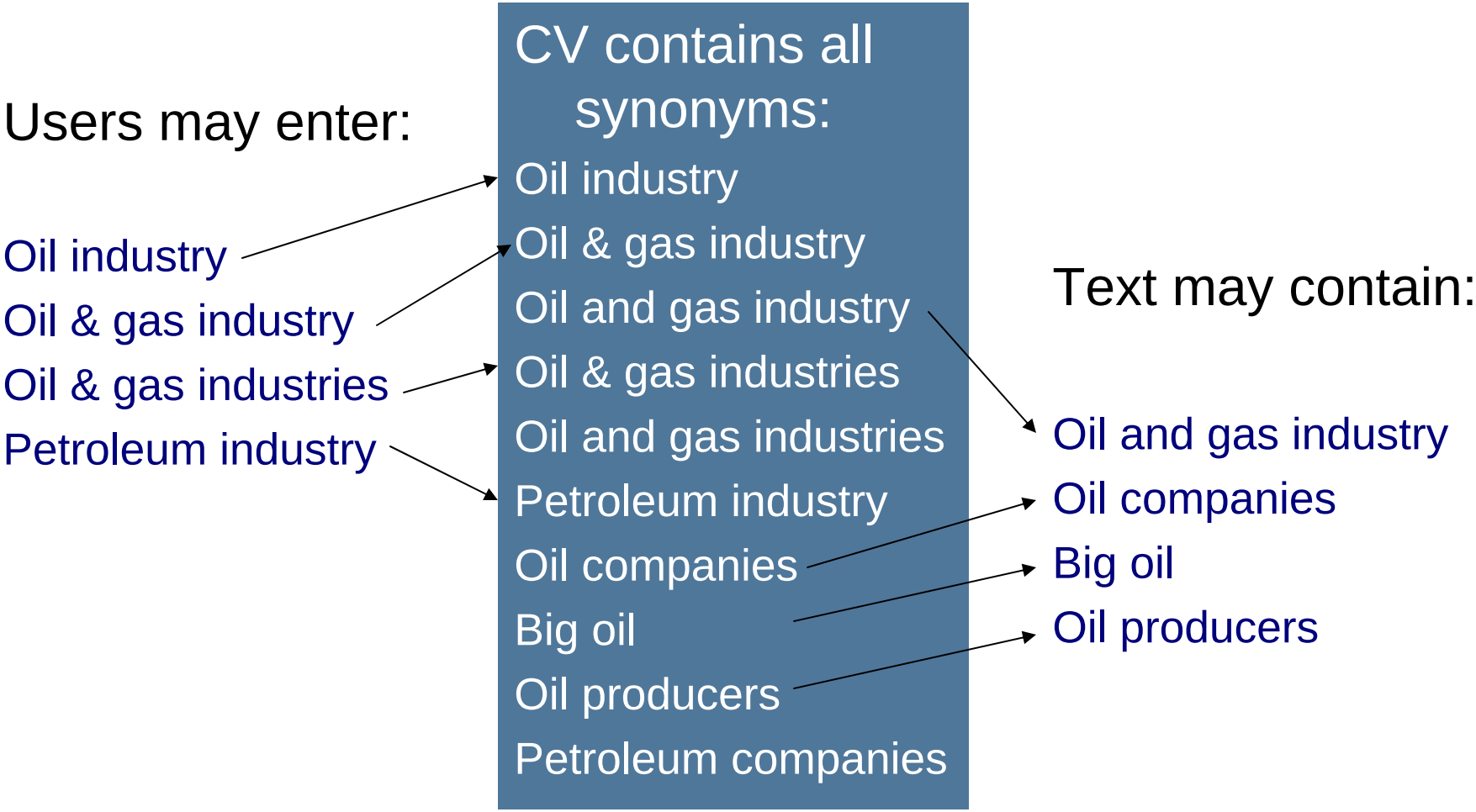
   Organizes information into a logical structure

   - Helps people browse or navigate for information

Taxonomy Purposes & Benefits

- There are multiple ways to describe the same thing.
- Without a controlled vocabulary keyword searches would miss relevant documents.
- A controlled vocabulary gathers synonyms, acronyms, variant spellings, etc.
  - Documents not missed due to use of different words (e.g. ***Automobiles***, instead of ***Cars***)
  - User does not need to guess the spelling of unusual or foreign names (e.g. ***Condoleezza Rice***)

Taxonomy Purposes & Benefits



Taxonomy Purposes & Benefits

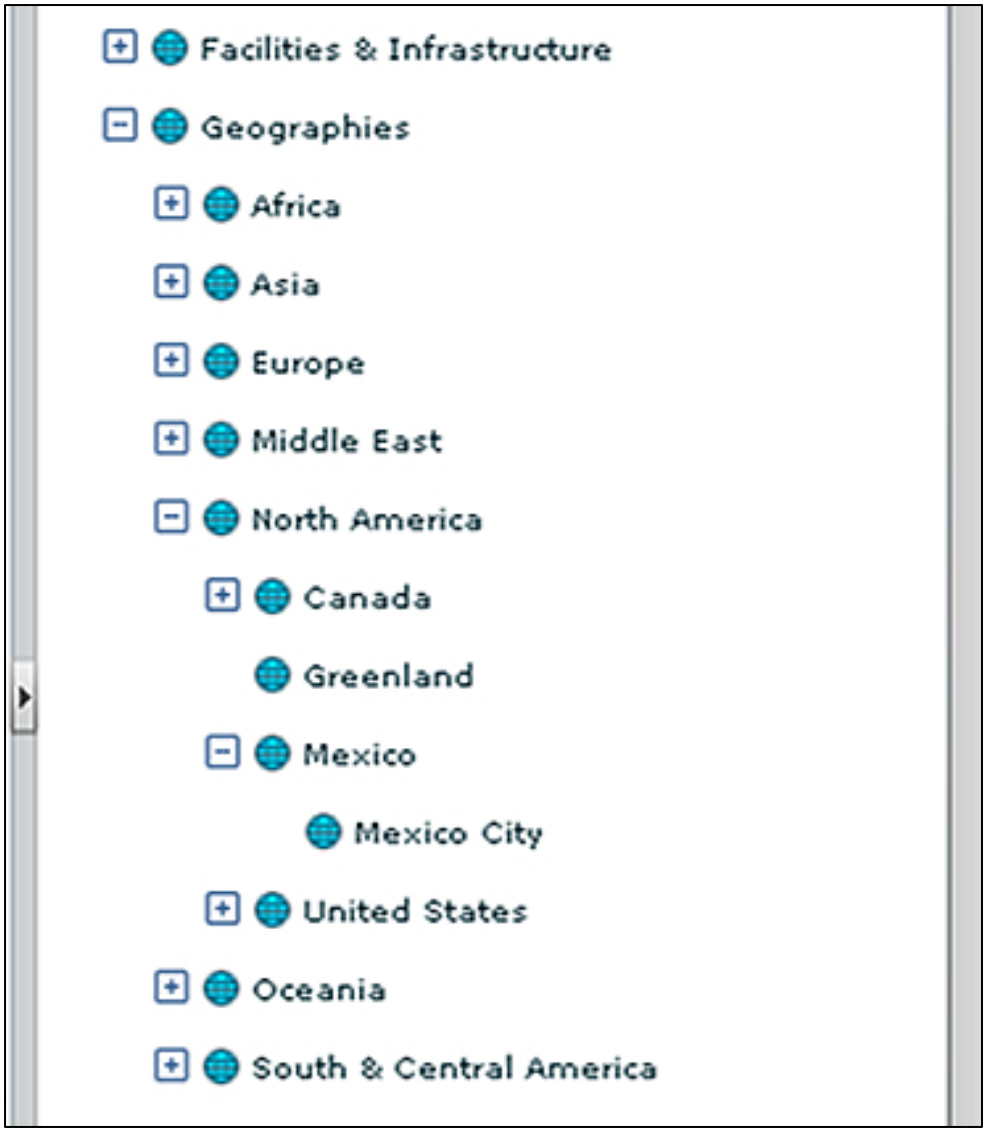
- A single term may have multiple meanings.
- Without a controlled vocabulary, too many irrelevant documents would be retrieved.
- A search restricted on the controlled vocabulary retrieves concepts not just words.
  - Excludes document with mere text-string matches (e.g. ***monitors*** for computers, not the verb "observes")

Taxonomy Purposes & Benefits

A hierarchical taxonomy provides guided search.

- Users can browse and locate narrower (more specific) subjects of interest.
- Users find out what is included and what is not.
-Users may find related subjects of interest.

Taxonomies reflect natural categories.



Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Taxonomy Types and Focus

Various types, all to serve “findability”

1. Organization and navigation support
2. Information retrieval support
3. Indexing/tagging support (and secondarily, retrieval)

Taxonomy Types and Focus

1. Organization and Navigation Support

- Categorical organization of a web site, intranet, portal structure, CMS areas
- Multi-level navigation menu labels
- Site maps (of topics, not just page titles)

Taxonomy Types and Focus

Examples of organizational taxonomies

- Digital Web Magazine topics
  http://www.digital-web.com/topics
- Information Architecture Institute
  http://iainstitute.org/en/site-map.php
- MyFlorida.com State of Florida site map
  http://www.myflorida.com/taxonomy

Taxonomy Types and Focus

2. Information Retrieval Support

- Displayed taxonomies
  - As browsabale hierarchies
  - As facets
- Non-displayed taxonomies
  - Synonym rings match user entered terms to taxonomy terms

Taxonomy Types and Focus

Examples of displayed hierarchical taxonomies:

- Verizon Superpages
  http://www.superpages.com/yellowpages
- Amazon.com book subject categories
  http://www.amazon.com/gp/homepage.html

Examples of faceted taxonomies:

- Microbial Life Educational Resources
  http://serc.carleton.edu/microbelife/resources
- Shoebuy.com - advanced search
  http://www.shoebuy.com/s.jsp/r_as

Taxonomy Types and Focus

3. Indexing Support

- Controlled vocabulary to support consistent indexing/tagging by multiple indexers who are indexing multiple different documents or records.
- Often in a thesaurus display to support indexer browsing and finding the best term quickly.

Agenda

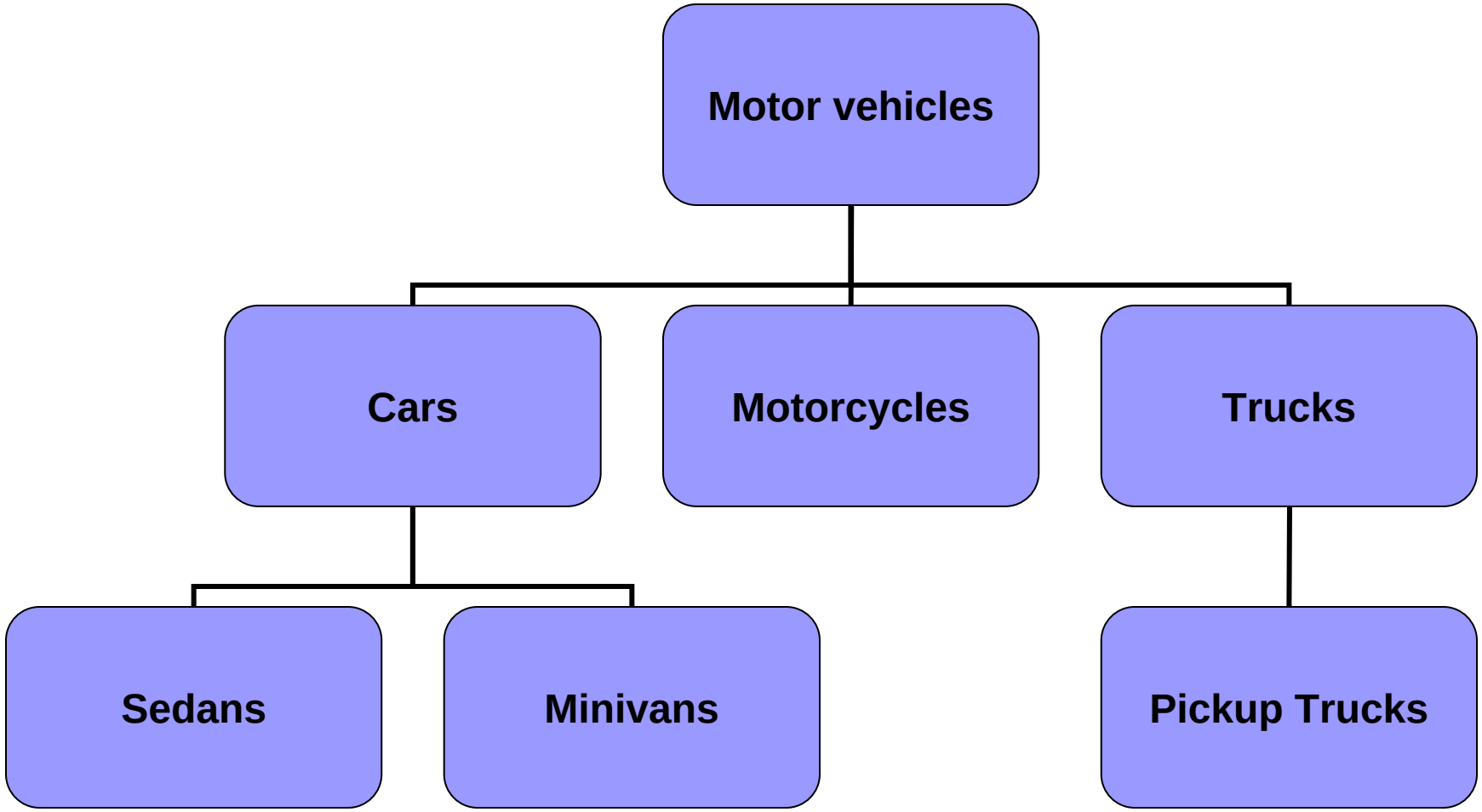
1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Hierarchies

Taxonomy Hierarchies:

- Each term has a broader (parent) and/or narrower (child) term that it's related to
- Terms have multiple narrower terms
- Terms usually have only one broader term, but more than one (polyhierarchy) might be permitted.
- Emphasizes categorization, classification, sorting
- Involves navigating from the top down
- Also known as "tree" structures

Hierarchies



Hierarchies

Hierarchical display options

- One level per web page
- Expandable tree
  - In same column
  - In second column
- Combination

Hierarchies

One level per web page

Yahoo directory

http://search.yahoo.com/dir

ThomasNet browse

http://www.thomasnet.com/browse

Hierarchies

One level per web page

Advantages:

- Large number of terms can display at each level
- Works well with polyhierarchies
- Technically easy to implement

Disadvantages:

- Display takes entire screen width, so search results don't display on same screen or are pushed lower down
- Users see only one level at a time
- Less appropriate for taxonomies with varied/inconsistent levels or levels containing just one or a few terms
- May lack method to indicate whether a level is final or contains more lower levels

Hierarchies

Expandable tree

USA Today

http://content.usatoday.com/community/tags/topic-index.aspx

Nicem

http://accessinn.com:8081/PerfectSearch/navtree/index.html

Hierarchies

Expandable tree

Advantages:

- Hierarchy is obvious to the user
- Supports visualization and interactive use of the taxonomy
- Presence of plus signs indicates presence of deeper levels
- Accommodates inconsistent numbers of terms per level
- Display takes up only part of screen

Disadvantages:

- Insufficient for displaying very large taxonomies or large numbers of terms at the same level
- Does not support polyierarchies as well due to space limitations, and could be confusing too
- More complex to develop and may take more time to load display

Hierarchies

Decision-making issues

- Deciding the number of levels and the number of terms per level
- Deciding the arrangement of the term hierarchy
- Named entities (proper nouns) within subject hierarchies or not

Hierarchies

Named entities separate out in hierarchies:

Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Facets

- For serving faceted classification, which allows the assignment of multiple classifications to an object
- A “dimension” of a query; a type of concept
- Intended for searching with multiple terms in combination (post-coordination), one from each facet
- Reflect the domain of content
- Similar, but not identical, to a single hierarchy
- May contain internal hierarchy or may not

Facets

For enterprise taxonomies:

Patrick Lambe,

Organising Knowledge

- *People and organizations*
- *Things and parts of things*
- *Activity cycles*
- *Locations*

For Web sites:

Rosenfeld and Morville,

Information Architecture

- *Topic*
- *Product*
- *Document type*
- *Audience*
- *Geography*
- *Price*

Facets

Shoebuy retail site

http://www.shoebuy.com/s.jsp/r_as

My recipes

http://search.myrecipes.com

Microbial Life Educational Resources

http://serc.carleton.edu/microbelife/resources

Facets

Advantages

- Supports more complex search queries by users
- Allows users to control the search refinement, narrowing or broadening in any manner or order

Disadvantages

- Only suitable for somewhat structured, unified type of content that all share the same multiple facets
- Does not support “advanced search” of multiple terms selected at once (“or”) from the same facet
- Requires investment of thorough indexing/tagging

Hierarchies & Facets Combined

ThomasNet businesss directory

http://www.thomasnet.com

Amazon.com

http://www.amazon.com

Buzzillions product reviews

http://www.buzzillions.com

Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Taxonomy Creation Best Practices

- Terms should reflect language of the users
- Terms should have consistent format and style
- Terms should not be ambiguous
- Taxonomy should have structural balance
- Broader term/narrower relationships must be either:
  1. generic/specific: “is a kind of”
  2. type/instance: “is an example of”
  3. whole/part: “is an integral part of”
- Taxonomy should have its own owner, policy, guideline, etc. – governance
- Use taxonomy/thesaurus management software
- Taxonomy should be periodical reviewed and updated

Taxonomies and Usability

- Taxonomy creation methods include:
  - Interviewing test users
  - Card-sorting exercises
  - Search logs, other web analytics

But also other methods of analyzing sample content; consulting other vocabularies, thesauri, glossaries; working with subject matter experts

Taxonomies and Usability

- Other Issues
  - Balancing number of terms per level and numbers of levels
  - Choosing the preferred term name

- Taxonomies serving usability
  - Synonyms/variants serve different ways users call the same concept
  - Terms in more than one hierarchy (polyhierarchy) serve different navigation perspectives

Agenda

1. Definitions
2. Taxonomy Purposes & Benefits
3. Taxonomy Types & Focus
4. Hierarchies
5. Facets
6. Taxonomy Creation Best Practices
7. Taxonomy Resources

Taxonomy Resources: Books

- ANSI/NISO Z39.19 (2005) *Guidelines for Construction, Format, and Management of Monolingual Controlled Vocabularies.* Bethesda, MD: NISO Press.

- Hedden, Heather. (2010). *The Accidental Taxonomist.* Medford, NJ: Information Today Inc. www.accidental-taxonomist.com

- Lambe, Patrick. (2007). *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Oxford, England: Chandos Publishing.

Taxonomy Resources: Organizations

- American Society for Indexing: Taxonomies and Controlled Vocabularies Special Interest Group
  http://www.taxonomies-sig.org

- Special Libraries Association (SLA): Taxonomy Division
  http://wiki.sla.org/display/SLATAX/Taxonomy%2BHome

- American Society of Information Science & Technology
  http://www.asis.org

Taxonomy Resources: Discussion Groups

- Taxonomy Community of Practice
  http://finance.groups.yahoo.com/group/TaxoCoP

- Taxonomies & Controlled Vocabularies SIG, ASI
  http://finance.groups.yahoo.com/group/taxonomies

Taxonomy Resources: Workshops, Seminars

- Taxonomy Community of Practice Webinar phone calls
  - $50 each. Occasionally free vendor-sponsored calls.
  - Usually first Wednesday of the month, 1:00-2:00 pm EST
  - www.earley.com/TaxoCoP.asp
- "Taxonomies and Controlled Vocabularies" Simmons College Graduate School of Library and Information Science Continuing Education Program
  - 6 weeks. $250. Next in October 2010.
  - www.simmons.edu/gslis/continuinged/workshops
- Taxonomy Boot Camp conference
  - Information Today Inc.
  - Next: November 15-16, 2010, Washington, DC
  - www.taxonomybootcamp.com

Taxonomy Resources: Web Sites

- **Taxonomy Community of Practice Wikispace**, http://taxocop.wikispaces.com
- **Taxonomy Guide, Faculty of Information Studies, University of Toronto** http://plc.fis.utoronto.ca/tgdemo/default.asp
- **Construction of Controlled Vocabularies: A Primer** http://www.slis.kent.edu/%7Emzeng/Z3919/index.htm
- **Thesaurus Construction tutorial by Tim Craven** http://publish.uwo.ca/~craven/677/thesaur/main00.htm
- **Willpower Information: Publications on thesaurus construction and use** http://www.willpowerinfo.co.uk/thesbibl.htm
- **Taxonomy Watch Blog by Linda Farmer**, http://taxonomy2watch.blogspot.com
- **Earley & Associates** http://www.earley.com
- **Taxonomy Strategies** http:// www.taxonomystrategies.com

Questions/Contact

Heather Hedden
Hedden Information Management
98 East Riding Drive
Carlisle, MA 01741
www.hedden-information.com
heather@hedden.net
978-371-0822
978-467-5195 (cell)