



Taxonomy for Delivering Targeted Technical Content

Content Wrangler Webinar, April 12, 2023

Heather Hedden
Semantic Web Company



About the Speaker

Heather Hedden

Product Communication Representative - Americas
and Knowledge Manager
Semantic Web Company

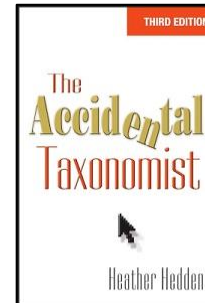


Over 25 years of experience in developing and managing taxonomies, metadata, and other knowledge organization systems for various organizations and applications.

Prior taxonomy consultant and staff taxonomist.

Instructor of self-paced online taxonomy courses.

Author of the book *The Accidental Taxonomist*
(3rd edition, 2022)



- ▶ A brief introduction to taxonomies
- ▶ Taxonomies delivering targeted content
 - ▷ Multiple, linked, specialized taxonomies
 - ▷ Selective taxonomy portions (concept schemes) for implementations
 - ▷ Recommended content based on added semantic relationships
- ▶ Personalization of taxonomies
 - ▷ Inclusion of synonyms for different users
 - ▷ Varied displays of the same taxonomy
 - ▷ Customization of faceted taxonomy displays

Intro to Taxonomies

The screenshot shows the 'Select Tags' interface in Adobe Experience Manager (AEM). It features a breadcrumb trail 'Tags / We.Retail / Apparel / Glasses' and a grid of tag categories. A modal window is open, displaying a hierarchical tree structure:

- Documents
- Library
- Categories
 - Category Root
 - Format Type
 - Information Type
 - Languages
 - Marketing Materials
 - Product Segment Type
 - Regions
 - Software Document Classification
 - Supporting Materials Type
 - Tags

- Tags

AEM

Taxonomies have become more common.

The screenshot shows the 'Select Tags' interface in Alfresco. It displays a hierarchical tree structure similar to AEM:

- Documents
- Library
- Categories
 - Category Root
 - Format Type
 - Information Type
 - Languages
 - Marketing Materials
 - Product Segment Type
 - Regions
 - Software Document Classification
 - Supporting Materials Type
 - Tags

- Tags

Alfresco

The screenshot shows the 'Term store' interface in SharePoint. It includes a search bar and a list of terms with expandable options:

- All About Cocktails_en
- Jobs and Skills Demo Appli...
 - Company Type
 - Job Roles
 - Location
- Skills
 - Languages
 - Operating Systems
- Programming
 - APL
 - ASP.NET
 - C

SharePoint

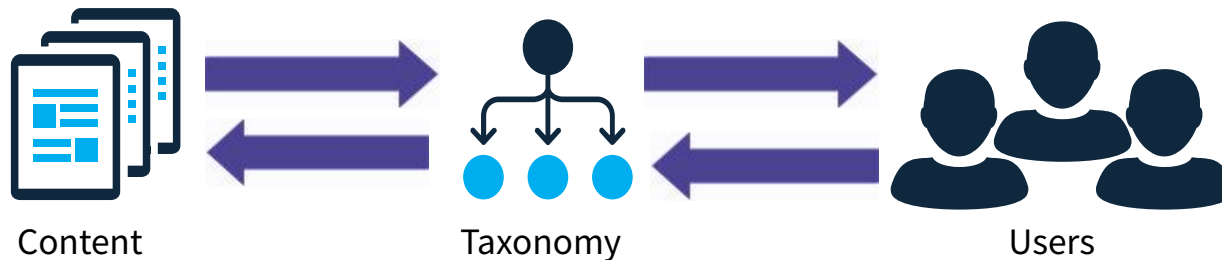
The screenshot shows the 'Term store' interface in Drupal. It features a table with a header 'OPERATIONS' and a list of terms, each with an 'edit' link:

OPERATIONS
edit
edit
edit
edit
edit
edit
edit
edit
edit
edit

Drupal

What and why taxonomies?

- ▶ A taxonomy is a controlled vocabulary organized into a hierarchical structure.
- ▶ Concepts/terms are used to tag/index/categorize content to make it easier to be found and retrieved by users
 - ▷ supporting better findability than search alone
- ▶ The taxonomy is an intermediary that links the user to the desired content.



- ▶ Terms are based on unambiguous concepts, rather than using text strings.
- ▶ Concepts are customized for the content. Term labels are customized for the users.

What is a taxonomy?

Controlled and *organized*

1. **Controlled:**

A kind of controlled vocabulary, metadata values, based on unambiguous concepts, not just words:

things, not strings

2. **Organized:**

Concepts are arranged in a structure of hierarchies, categories, or facets to organize them.

controlled

Preferred Label	
<input type="radio"/> Vorspeisen	de
<input type="radio"/> Appetizers	en
Alternative Labels	
<input type="radio"/> Hors d'oeuvres	en
<input type="radio"/> Starters	
Hidden Labels	
<input type="radio"/> Appetisers	en
<input type="radio"/> Horderves	
Scope Notes	
<input type="radio"/> Dishes usually served as appetizers	en
Definitions	
<input type="radio"/> A small dish of food or a drink taken before a meal or the main course of a meal to stimulate one's appetite	en

organized

- Recipes
 - Cooking methods (5)
 - Dishes (10)
 - Appetizers (3)
 - Bruschetta (0)
 - Dips (3)
 - Guacamole (0)
 - Hummus (0)
 - Spinach dip (0)
 - Quesadillas (0)
 - Breads and muffins (2)
 - Breakfast dishes (3)
 - Desserts (4)
 - Egg dishes (2)
 - Meat & poultry (4)
 - Pasta, rice, potatoes (3)
 - Salads (4)
 - Seafood (3)
 - Soups and stews (2)
 - Ingredients (5)
 - Occasions (3)

Benefits of taxonomies

- ▶ Providing consistent terms for tagging
- ▶ Bringing together synonyms for the same concept
- ▶ Organizing terms in to refinement filters
- ▶ Organizing terms into guiding hierarchies
- ▶ Supporting term definitions and other data (history, source, image, etc.)
- ▶ Supporting multilingual concepts, content, and users
- ▶ Potential to link to other taxonomies and thus other content/data

Intro to Taxonomies: Benefits

Organizing terms into search refinement filters

Department

- Research
- Sales
- Finance
- Marketing
- Exec Office
- SHOW MORE

Job Title

- Communications Director
- Developer
- Financial Controller
- Finance Assistant
- Head of IT Services
- SHOW MORE

Office Location

- London
- New York
- Shanghai

For people

File Type

- Word

Document Type

- Feature Overview
- Technical Specifications

Item Type

- Project Document

Project

- All
- CRM Support

Sprint

- Sprint 23
- Sprint 22

For documents

Result type

- Word

Author

- User1
- System Account
- AATISH AGARWAL
- SHOW MORE

Modified date

Community

- Technical
- Recreational
- Other Value
- Apply | Clear

Language

- English
- German

Content Manager

- TechNet Library (7)
- OfficeOnlineVNext (2)
- MSDN Code Gallery (1)
- MSDN Library (1)
- MSN Video (1)
- Other Value
- Apply | Clear

Distribution Channel

- Other Value
- Apply | Clear

Request status

- Published (12)
- Other Value
- Apply | Clear

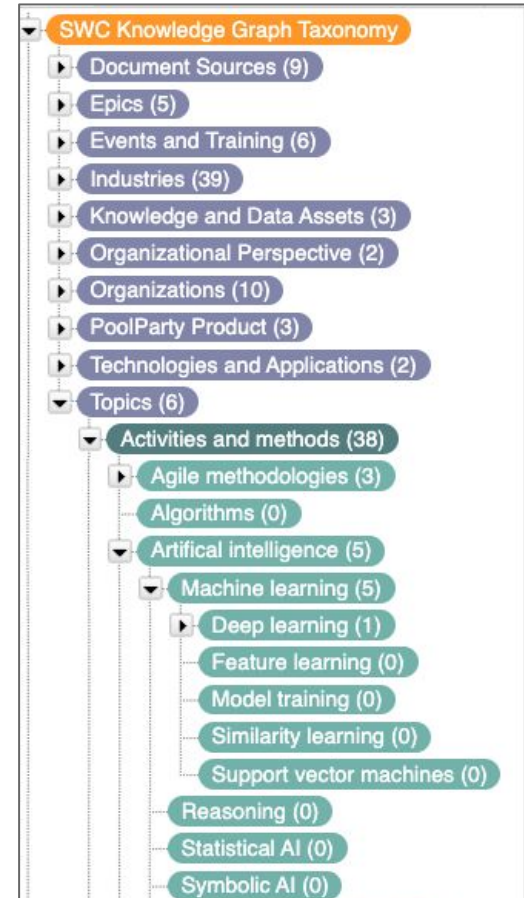
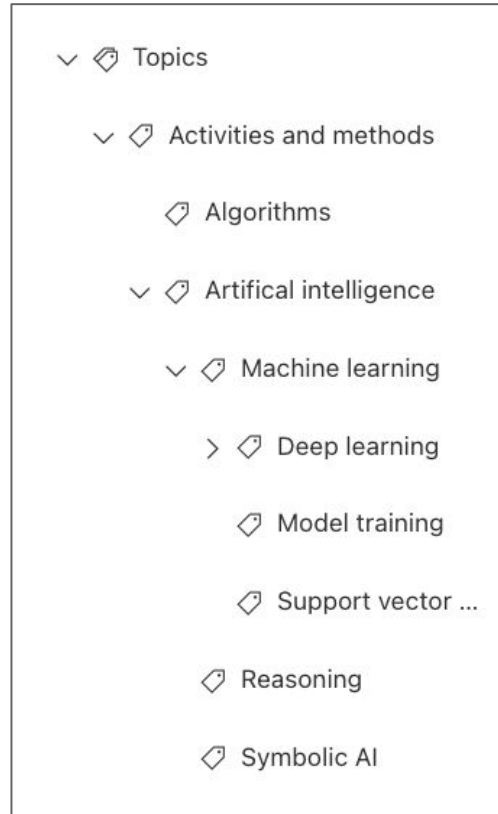
Content Type

- HVC library content item (8)
- OCMS article (2)
- Code Gallery content (1)
- Video content (1)
- Other Value
- Apply | Clear

Internal Writer

Intro to Taxonomies: Benefits

Organizing terms into hierarchies to guide users to specific topics



Tagging - assigning metadata to content

- ▶ Tagging with a taxonomy: “indexing,” “semantic tagging,” “controlled tagging”

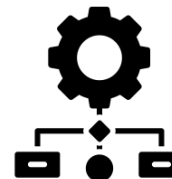
Benefits of manual tagging

- ▶ Greater precision/accuracy.
- ▶ Can respond to requests for new terms quickly.
- ▶ Better for complex content and subjects.
- ▶ Suitable also for non-text content (uncaptioned).



Benefits of auto-tagging (machine-learning or rules-based; and NLP)

- ▶ More comprehensive, complete tagging.
- ▶ More consistent tagging.
- ▶ Faster tagging for large volume of content.



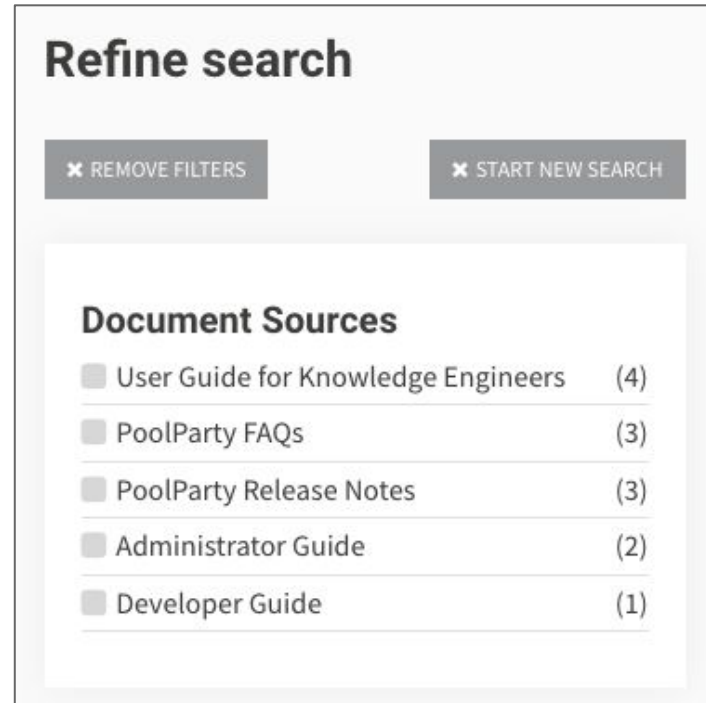
- ▶ **Targeted content**
 - ▷ Content created for a specific, niche audience
- ▶ **Personalization**
 - ▷ Content that is delivered to users based on their criteria and/or in a customized user experience
- ▶ **Taxonomies** serve both:
 - ▷ Targeted content:
 - by tagging content by content types, audience types
 - by having separate taxonomies or branches of taxonomies
 - ▷ Personalization:
 - by retrieving content based on user profiles, selections, filters, etc.

Targeted Taxonomy Options:

- ▶ Controlled vocabularies for content types and audience types
- ▶ Separate (but linked) taxonomies for separate audiences
- ▶ A single taxonomy with different branches (concept schemes) for different audiences
- ▶ A single taxonomy with designated “collections” applied to concepts for different purposes/audience/implementations
- ▶ A single taxonomy including branches (concept schemes) for audience metadata that can be linked to other taxonomy concepts by semantic relationships.

Targeting with Content Types and Audience Types

- ▶ Types tagged to all content
- ▶ Serving as a filter for users, even when it's not a fully faceted taxonomy



Targeting with Separate Linked Taxonomies

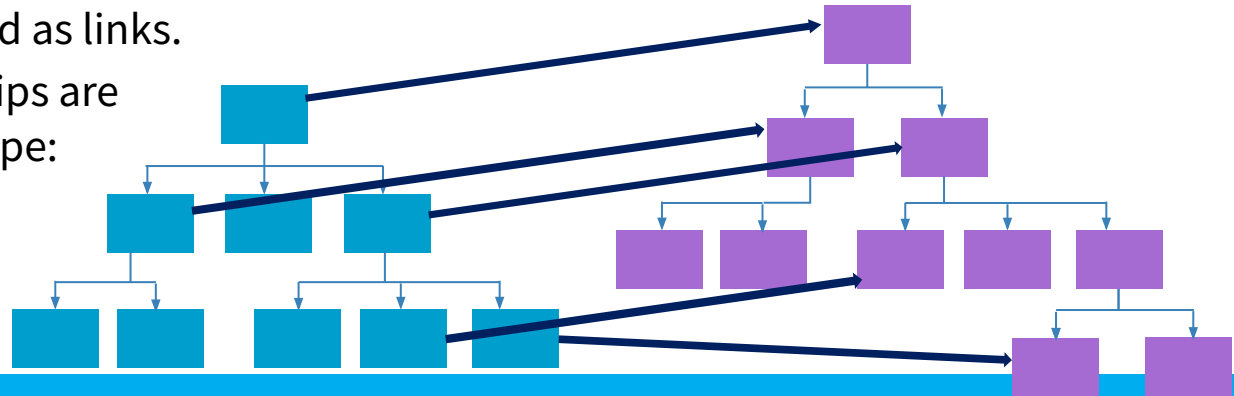
Why link taxonomies?

1. There is a reason to keep them distinct (distinct content, users, applications, management), *and...*
2. There is a need to query/access/manage all content or all taxonomies together.

How to link taxonomies

- ▶ Taxonomies are put into a single taxonomy system with linking capabilities.
- ▶ Matches between individual concepts are automatically identified and then manually approved as links.
- ▶ Linking relationships are of a distinct link type:

Exact match



Targeting with Distinct Taxonomy Branches

Taxonomy top-level branches selectively implemented

Refine search

✕ REMOVE FILTERS

✕ START NEW SEARCH

Document Sources

- PoolParty FAQs (10)
- User Guide for Knowledge Engineers (10)
- Administrator Guide (5)
- PoolParty Overview (4)
- PoolParty Release Notes (1)

PoolParty Product

- Thesaurus Management (30)
- Corpus Management (6)
- Extractor (5)
- PoolParty APIs (5)
- GraphSearch (3)
- Semantic Middleware Configurator (3)
- Blacklist (2)
- Concept scheme details (2)
- Custom schemes (2)
- Data Validator (2)

LOAD MORE

Topics

- Taxonomies (11)
- SKOS (8)

internal.semantic-web.at/GraphSearch/

KNOWLEDGE HUB

Knowledge Spaces

SWC Knowledge Hub

Facets

Document Sources (7)

Epics (5)

Events and Training (4)

Industries (40)

Knowledge and Data Assets (16)

Organizational Perspective (5)

Organizations (9)

PoolParty Product (3)

Skills (1)

Technologies and Applications (2)

Topics (8)

Thesaurus Management

EN

Selected facet values: Thesaurus Management (PoolParty Product) ✕ Clear all ✕

1213 results

[Web service method request subtree of concept or concept scheme](#)

<https://help.poolparty.biz/en/developer-guide/basic--advanced-server-apis/thesaurus--ontology-manager-api/thesaurus-services/concept-or-concept-scheme.html>

Web Service Method: Request Subtree of Concept or Concept Scheme Abstract Web Service Method: Request Description Returns a list of all narrower concepts (recursively) of a given concept or concept scheme. Note Request Child Concepts ...

[Show similar](#)

[Skos and poolparty](#)

<https://help.poolparty.biz/en/poolparty-overview/poolparty%2C-skos%2C-rdf-and-uris/skos-and-poolparty.html>

SKOS and PoolParty Abstract SKOS and PoolParty The following tables show all SKOS classes and properties Introduction to SKOS at W3C). Concept Scheme Class (SKOS) Class (PP) Notes ConceptScheme Concept Scheme statements about semantic relationships between those ...

[Show similar](#)

Targeting with Semantic Relationships

The screenshot shows a web application interface for managing semantic data. On the left is a hierarchical tree view of categories, including 'ESCO skill-transversal-groups (4)', 'Propel: Category (8)', 'Propel: Foundation (18)', 'Propel: Subcategory (63)', 'Propel: Topic (26)', 'SEMWEB Applications and Products (12)', 'SEMWEB Data resources (3)', and 'SEMWEB Roles and Skills (2)'. Under 'SEMWEB Roles and Skills (2)', there are sub-categories like 'Roles (19)' and 'Skills (42)'. The 'Skills (42)' category is expanded, showing a list of skills such as 'Algorithm Development (0)', 'Automatic Term Extraction (0)', 'Automatic Term Extraction Algorithm (0)', 'Big data (0)', 'C# (0)', 'cognitive computing (0)', 'Command and Control Information System (0)', 'Computer Vision (0)', 'constraint type (0)', 'Corpus (0)', 'Data Scientist Skills (8)', 'Delphi (0)', 'Distributed revision control system (0)', 'full-stack development (0)', 'GraphQL (0)', 'graph store (0)', 'HTML5 (0)', 'Image Processing (0)', and 'Java (0)'. The 'Java (0)' skill is selected, and its details are shown in the main panel. The details panel has tabs for 'Details', 'Notes', 'Documents', 'Linked Data', 'Triples', 'Visualization', 'Quality Management', and 'History'. The 'Details' tab is active, showing the skill's URI, its type 'Skill', and its relationships. Under 'isEssentialSkillFor', there are three roles: 'computer science lecturer', 'computer-aided design operator', and 'numerical tool and process control programmer'. Under 'isOptionalSkillFor', there are 15 roles: '3D modeller', 'application engineer', 'chief ICT security officer', 'chief information officer', 'chief technology officer', 'computer hardware engineer', 'computer hardware engineering technician', 'computer numerical control machine operator', 'Data Scientist', 'data warehouse designer', 'database designer', 'database developer', and 'digital games designer'. There are also buttons for '+ Add to Collection', '+ Add to Blacklist', and '+ Add to ExactMatch'.

Semantic relationships “is essential skill for” and “is optional for” to link skills to roles, so users of a particular role retrieve content tagged with the skill.

- ▶ **Synonyms/alternative labels for concepts**
 - ▷ Taxonomy concepts bring together synonyms or alternative labels, so when different users (developers, administrators, general users) call the same topic by different names, they can still retrieve the tagged content
- ▶ **The same taxonomy can be presented in different ways**
 - ▷ It can be both searched, by users who know what they are looking for, and browsed by users who want to explore the topics.
 - ▷ Facets may be adjusted by the users to display concepts in ranked order by the number content items tagged, alphabetical order, or with taxonomy hierarchies.

Defined: *Approximately* synonymous words or phrases to refer to a concept, for the *context* of the controlled vocabulary and the set of content.

Purposes:

- ▶ To capture different wordings of how different people might describe or look up the same concept or idea and used as alternative entries.
 - ▷ Differences between that of the content author and the end-users
 - ▷ Differences among different end-users
 - ▷ Differences between that of those tagging and the end-users
 - ▷ Differences among different people tagging
- ▶ To support auto-tagging by matching various ways the concept may be named in different texts.
- ▶ Also called: synonyms, non-preferred terms (NPTs), See/cross-references, variants

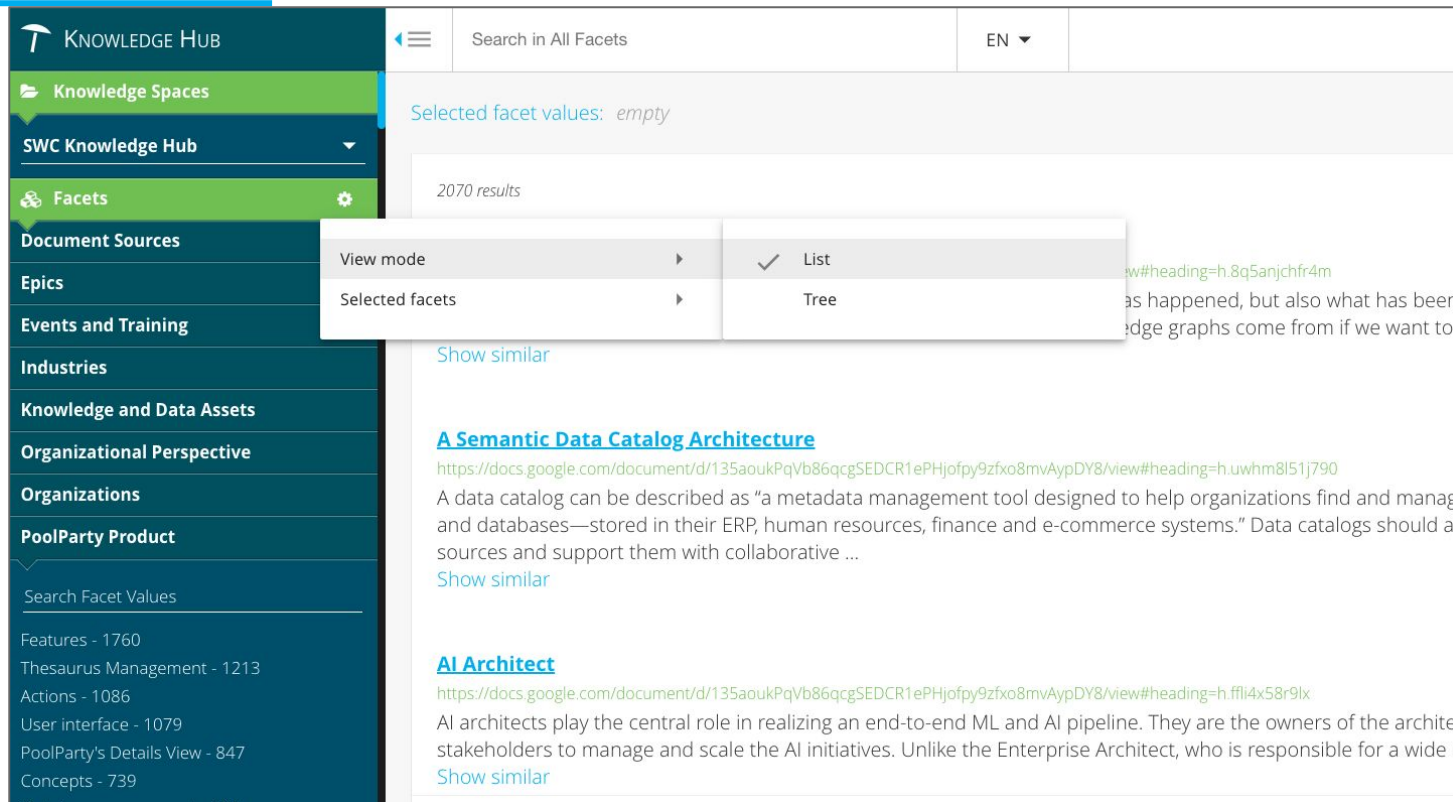
Types of alternative labels

- ▶ synonyms
- ▶ quasi-synonyms
- ▶ variant spellings
- ▶ lexical variants
- ▶ foreign language names
- ▶ acronyms/spelled out
- ▶ scientific/popular names
- ▶ antonyms (for characteristics)
- ▶ older/current names
- ▶ phrase variations (in alphabetical display)
- ▶ narrower terms that are not preferred terms

Types of alternative labels

- ▶ synonyms: **Cars / Automobiles**
- ▶ quasi-synonyms: **Politics / Government**
- ▶ variant spellings: **Taoism / Daoism; Email / E-mail**
- ▶ lexical variants: **Selling / Sales; Hair loss / Baldness**
- ▶ foreign language names: **Ivory Coast / Côte d'Ivoire**
- ▶ acronyms/spelled out: **GDP / Gross domestic product**
- ▶ scientific/popular names: **Neoplasms / Cancer**
- ▶ antonyms (for characteristics): **Flexibility / Rigidity**
- ▶ older/current names: **Near East USE Middle East**
- ▶ phrase variations (in alphabetical display): **Unions, labor USE Labor unions**
- ▶ narrower terms that are not preferred: **Genetic engineering USE Biotechnology**

Different Displays for the Same Taxonomy



The screenshot shows a knowledge hub interface with a search bar at the top containing the text "Search in All Facets" and a language dropdown set to "EN". Below the search bar, a facet menu is open, showing "Selected facet values: empty" and "2070 results". The facet menu has two sections: "View mode" with a right-pointing arrow and a checkmark next to "List", and "Selected facets" with a right-pointing arrow and "Tree" selected. The main content area displays search results for "Semantic Data Catalog Architecture" and "AI Architect", each with a URL, a title, and a brief description. The left sidebar contains a navigation menu with categories like "Knowledge Spaces", "SWC Knowledge Hub", "Facets", "Document Sources", "Epics", "Events and Training", "Industries", "Knowledge and Data Assets", "Organizational Perspective", "Organizations", and "PoolParty Product".

- ▶ The Accidental Taxonomist Blog
<http://accidental-taxonomist.blogspot.com>
- ▶ Hedden past presentations and past articles <http://www.hedden-information.com>
- ▶ Taxonomy Boot Camp London online
<http://www.taxonomybootcamp.com/london>
- ▶ Taxonomies and Controlled Vocabularies” self-paced online course from Hedden Information Management
<http://www.hedden-information.com/courses-workshops/taxonomy-course>
- ▶ Hedden, Heather. (2022) *The Accidental Taxonomist*, 3rd edition. Medford, NJ: Information Today Inc. <http://www.hedden-information.com/accidental-taxonomist/>

Questions/Contact

Heather Hedden

Data and Knowledge Engineer

Semantic Web Company Inc.

One Boston Place, Suite 2600

Boston, MA 02108

857-400-0183

heather.hedden@semantic-web.com

www.linkedin.com/in/hedden

<http://accidental-taxonomist.blogspot.com>

Semantic Web Company www.semantic-web.com

PoolParty software www.poolparty.biz

