

Taxonomies in Action: Tips for Taxonomy Hierarchies

Heather Hedden

9 October 2024

virtual



Techniques, tips & tools for taxonomists & knowledge managers

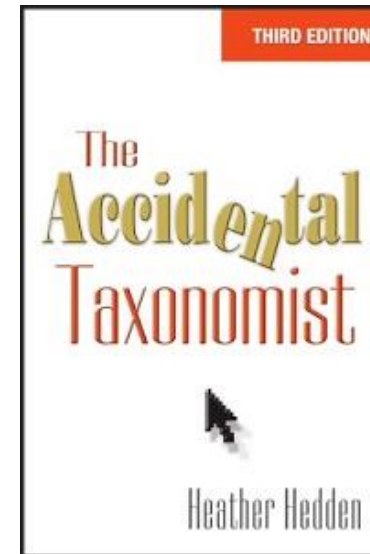
About Heather Hedden

- Independent taxonomy consultant, Hedden Information Management
- Instructor of online and corporate taxonomy courses and workshops
- Previously a taxonomy consultant in consulting firms, Enterprise Knowledge and PPC. Also, a contract consultant for others.
- Former taxonomy-related roles at Semantic Web Company, Gale/Cengage, Viziant, and First Wind.
- Author of *The Accidental Taxonomist*, 3rd ed. (2022, Information Today, Inc.)



**Hedden Information
Management**

Making information findable

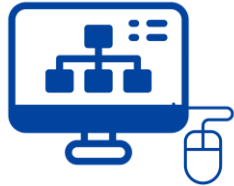


Outline

- Hierarchy purposes
- Hierarchical relationship standards
- Hierarchical relationship issues
 - Bending the rules for hierarchical relationships
 - Error of context-based narrower concepts
 - Named entities in hierarchies
 - Integrity of hierarchies (“is a”)
 - Single narrower concepts
- Polyhierarchy

Hierarchy Purposes

The hierarchy should serve a clear purpose:



- **Browsing:** Making it easier for the end user to quickly identify the concept they want for retrieving content.



- **Discovery:** Enabling end users to discover similar topics and content of interest, in the same part of the taxonomy.



- **Contextualizing tagging:** Providing context to concepts for indexers/taggers so that they apply the correct concept, or to support auto-tagging



- **Expanding retrieval:** Allowing a concept to retrieve not only what was tagged to it, but also what was tagged to each of its narrower concepts.

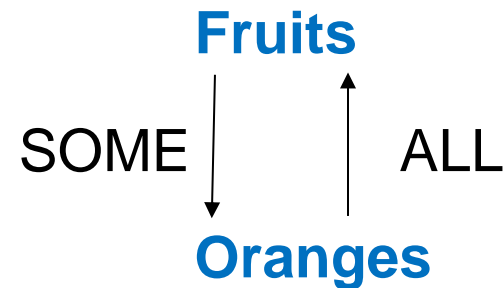
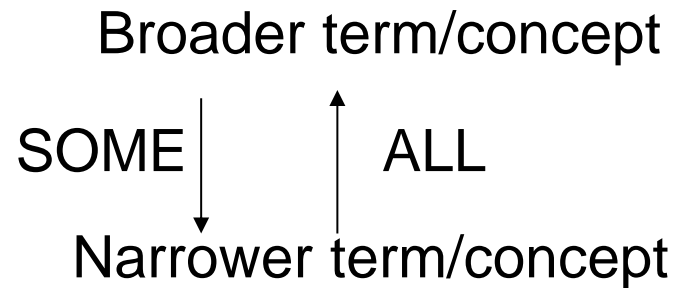


- **Instruction:** Educating users (such as students or employees) on the hierarchical structure of a subject area.

Hierarchical Relationship Standards

Hierarchical thesaurus relationship types by ANSI/NISO and ISO standards

Reciprocal (bi-directional) relationships, but asymmetrical



Fruits NT (*has narrower*) **Oranges**

Oranges BT (*has broader*) **Fruits**

Three types:

1. Generic – Specific: “is/are a kind of”
2. Generic – Instance: “is an instance of”
3. Whole – Part: “is/are within”

Hospitals has narrower: **Children's hospitals**

Hospitals has narrower: **Boston Medical Center**

Hospitals has narrower: **Emergency rooms**

Hierarchical Relationship Standards

In practice

1. **Generic – Specific**

Used the vast majority of the time. Also is “class – subclass” in ontologies.

2. **Generic – Instance**

Not used in faceted taxonomies (most taxonomies)

Named entities are in their own facets.

3. **Whole – Part**

Used in special cases: geography, disciplines, components, systems

Mixing types (hospital example) should also be avoided in displayed taxonomies (in contrast to thesauri).

Bending the Hierarchy Rules

May create hierarchical relationships that are *not* “a kind of” or “a part of” when all of the following apply:

- The taxonomy doesn’t use associative relationships.
- The broader concept is at or near the top of the hierarchy for grouping.
- The hierarchy functions for end-user browsing.
- The broader concept will *not* retrieve all content tagged with each of the narrower concepts, or will not even be tagged to content at all.

But best practice is to have consistency:

- Avoid “incorrect” hierarchical relationships sporadically throughout the taxonomy.
- Create them at the top level only, such as SKOS Top Concepts in a Concept Scheme.

Climate change

Climate change adaptation

Climate change mitigation

Climate technology

Extreme weather

Kyoto mechanism

Long-term temperature goals

Sea level rise

Bending the Hierarchy Rules

Do *not* create hierarchical relationships that are not “a kind of” or “a part of” when:

- It’s done sporadically throughout the taxonomy at any levels
- It’s done extensively for any grouping

Problem:

- Hierarchy is not logical for all users

Construction
Buildings
Building information management
Building management systems
Building materials
Asphalt
Concrete
Floors
Mortar
Tiles
Windows
Building services
Building automation
Elevators
Facilities management
Flexible structures
Modular construction

Error of Context-based Narrower Concepts

- A concept does not inherit meaning for its broader concept.
- Each concept is independent.

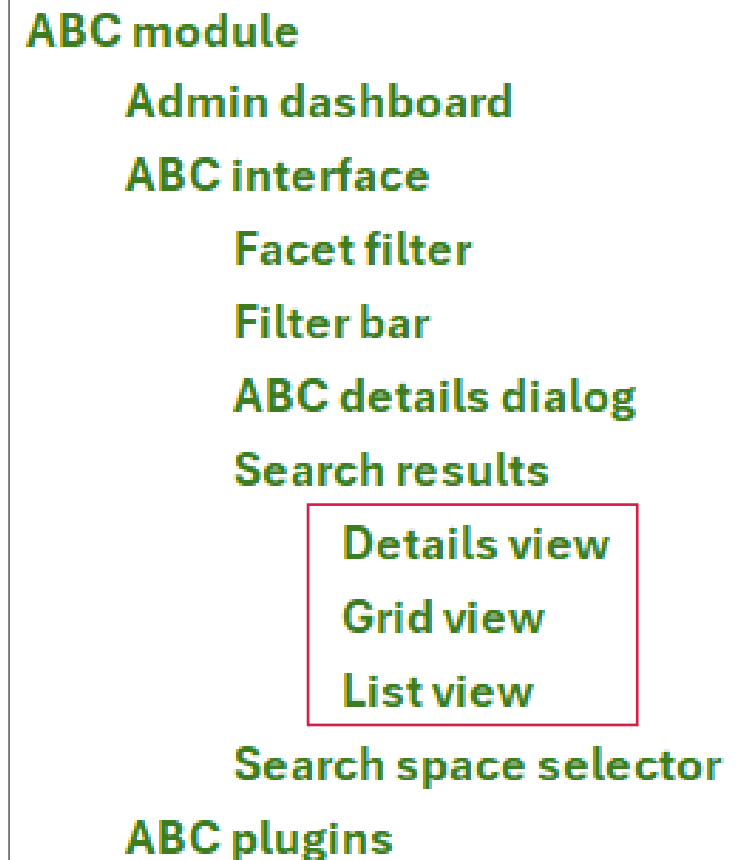


Problem:

- Narrower concept could be tagged for different meanings.

Fix the problem by:

- Refining the narrower concept with adjectives, e.g. **ABC search results details view**



Software product features

Named Entities in Hierarchies

Problems:

- Inconsistency
- Concepts may not appear as expected
- Does not serve natural facets
- Users may prefer to search **by type** or **by entity**, as separate aspects/facets.
- Users tend to browse for topics yet search for specific named entities.

Fix the problems by:

- Adding a concept scheme/facet for **Organization type** *Organizations facet*
- If categorizing organizations is useful merely for taxonomy management, set the organization types, appearing at top level only, as grouping categories, and not for tagging.



Integrity of Hierarchies

Problem:

- “is a” does not follow through.
- Broader concept retrieving what is tagged with narrower concepts would be incorrect.

Fix the problem by:

- Adding a concept scheme/facet for **Roles**, move out the job titles and roles, and leave only departments in the **Departments** or
- Rename the facet for what is for: **Roles**, and use the departments for grouping only, e.g. **Business administration roles** not for tagging, as “node labels”



Skills Taxonomy

Integrity of Hierarchies & Too Much Specificity

Problems:

Creating narrower concepts where the “is a kind of” logic no longer extends, and creating too many narrower concepts, where other concept remain broad.

Example: specific topics pertaining to **Oceanography**.

Fix the problems:

- Move specific concepts into another hierarchy.
- Delete and merge up the highly specific concepts, make them alternative labels



Integrity of Hierarchies

If an isolated concept does not fit the “is a” rule,
Perhaps it’s just a labelling problem,
and the “concept” is OK.

Example:

Problem:

Plant variety is not a kind of **Plant production**.

Fix the problem:

Change the label **Plant variety** to **Plant diversification**.



Single Narrower Concepts

A concept having a single narrower concept, instead of none or multiple

If full hierarchy is *not* displayed to the end-users, or it's a thesaurus, it's OK

Problem:

If the full hierarchy is displayed for drill-down browsing, each extra click should bring multiple narrower concepts, not just one. Poor user experience.

Fix the problem:

Merge the single narrower concept upward, and make it an alternative label.

The screenshot shows a hierarchical tree structure for the concept 'ENVIRONMENT'. The root node is 'ENVIRONMENT' (minus icon). It has two children: 'environmental policy' (minus icon) and 'environmental protection' (plus icon). 'environmental policy' has three children: 'waste management' (plus icon), 'pollution control measures' (minus icon), and 'environmental policy' (plus icon). 'pollution control measures' has four children: 'NT1 pollution control' (minus icon), 'NT1 prevention of pollution' (plus icon), 'NT1 polluter pays principle' (bullet), 'NT1 noise protection' (bullet), and 'NT1 decontamination' (bullet). 'NT1 pollution control' has one child: 'NT2 degree of pollution' (bullet). 'NT1 prevention of pollution' has three children: 'NT1 non-polluting vehicle' (bullet), 'NT1 noise protection' (bullet), and 'NT1 decontamination' (bullet). Each node includes a search icon and the word 'search'. The 'NT2 degree of pollution' node is highlighted with a red border.

Eurovoc Thesaurus excerpt

More Top Concepts than Narrower

Hierarchy that would have more top concept and fewer second level.

If full hierarchy is *not* displayed to the end-users, or it's a thesaurus, it's OK

Problem:

If the full hierarchy is displayed for drill-down browsing, poor user experience.

Fix the problem:

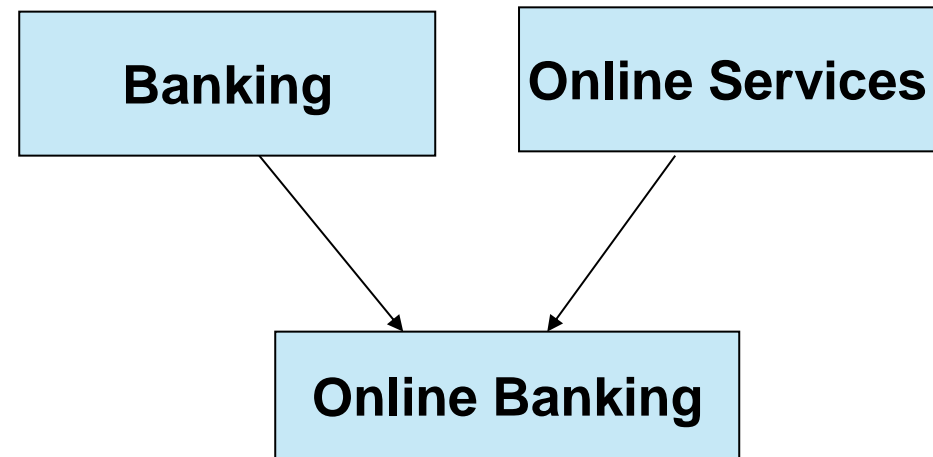
- If there are only 1-3 narrower concepts total, consider making it a flat list instead.
- If some have narrower concepts and others don't, use a symbol (+ or >) to indicate presence of narrow concepts

Industries		
	Accounting and business services	
	Aerospace and defense	
	Apparel and fashion	
	Aviation and airlines	
	Broadcasting and media	
	Chemicals	
	Construction and engineering	
	Consulting industry	
	Consumer electronics	
	Consumer goods	
	E-commerce and retailing	
	Education and training	
	Energy industry	
	Renewable energy	
	Entertainment	
	Computer games	
	Financial services	
	Banking	
	Food and beverages	
	Government	
	Healthcare	
	Higher education	
	Hospitality and tourism	
	Information services industry	
	Information technology and services	

Polyhierarchies

Sometimes a concept can have two or more broader concepts.

- Polyhierarchy is permitted if the hierarchical relationship is valid in both/all cases
- Remember “All-and-Some” test for each generic hierarchical relationship
- Systems may or may not support it.

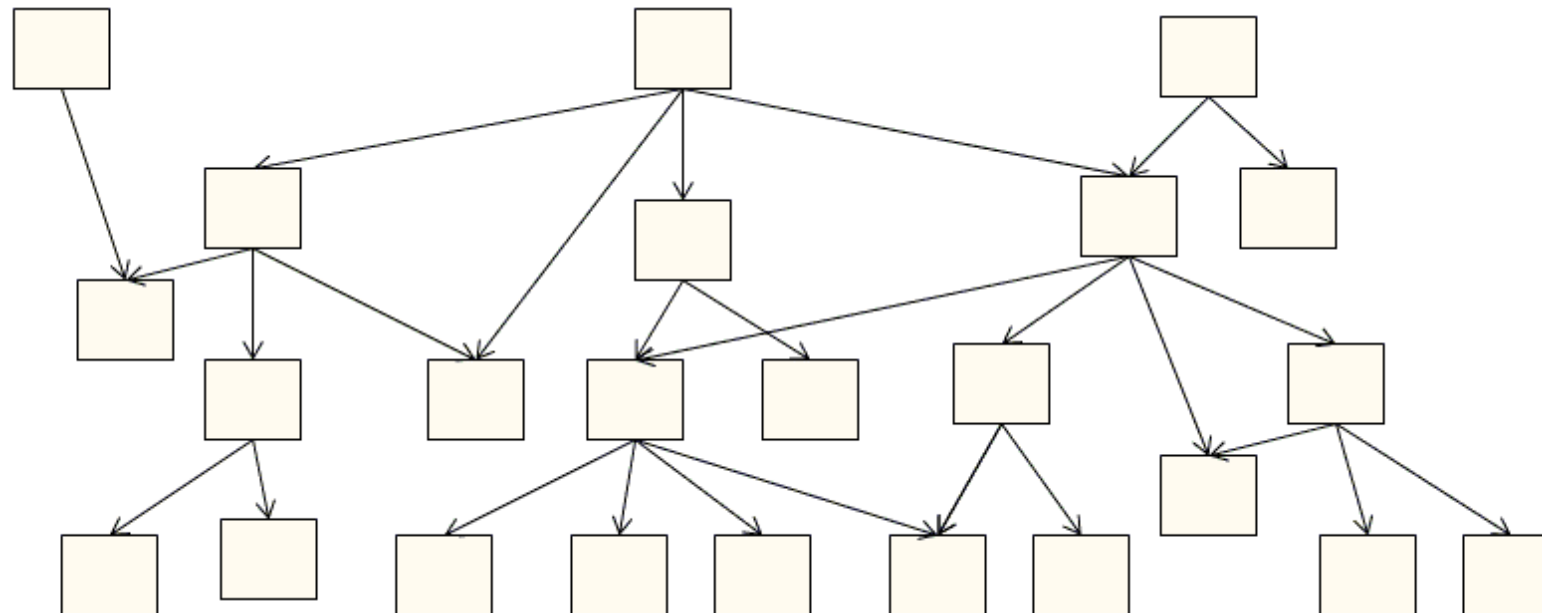


Polyhierarchies

Polyhierarchy Issues: excessive polyhierarchy

Problems:

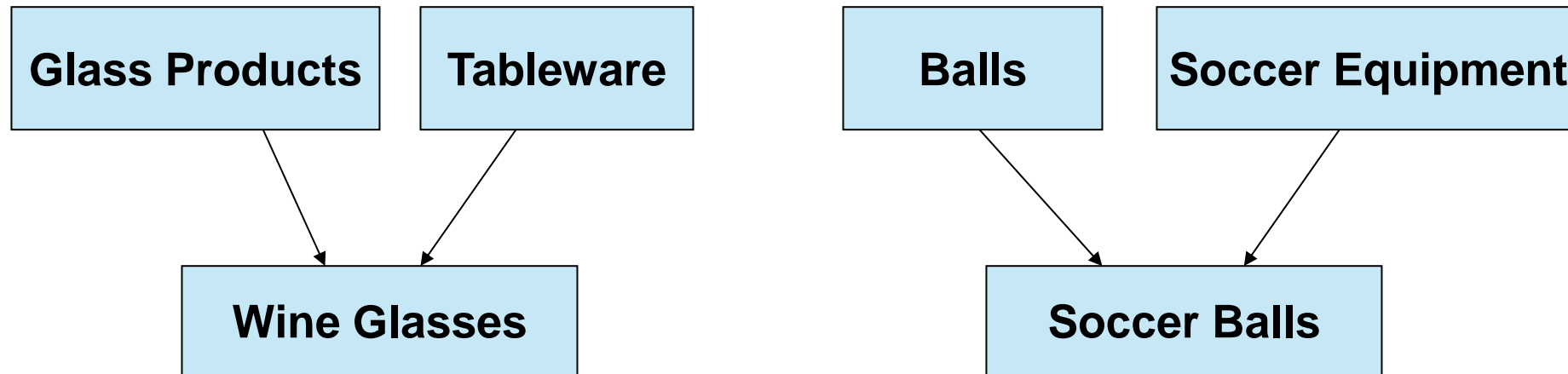
- Familiar tree structure is lost.
- Users cannot see the logical hierarchy.
- Users spend too much time clicking through categories.



Polyhierarchies

Polyhierarchy Issues: excessive polyhierarchy

Example: polyhierarchies for products based on different classifications could become very extensive.



Fix the problem:

If the hierarchies are all correct, then it's a faceted classification issue.

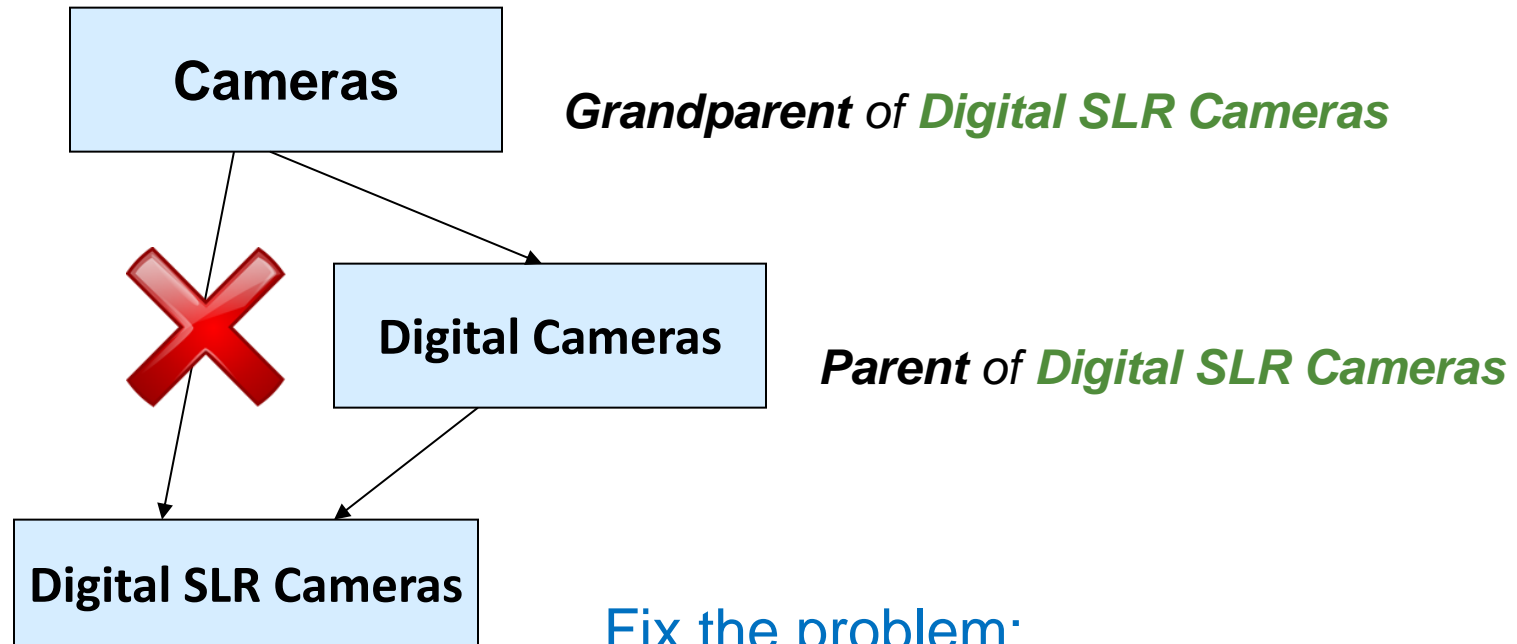
Consider creating a faceted taxonomy instead. A facet for type and facet for use.

Polyhierarchies

Polyhierarchy Issues: Polyhierarchy within the same hierarchy

Problem:

Polyhierarchy to both a “parent” and a “grandparent” of a concept.

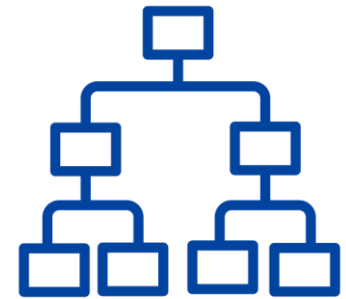


Fix the problem:

Delete hierarchical relationship to grandparent.

Conclusions

- If the hierarchy is primarily for the end-users
 - Consider the user experience in the hierarchy design: intuitive, logical, consistent
 - “Bending” the rules on hierarchy should be limited to the top grouping levels.
- If the hierarchy will be utilized in tagging and retrieval
 - Hierarchy standards (such as “is a”) must be closely followed.
- Keep facets “pure”
 - Don’t make named entities as narrower to generic types in facets.
- Very small controlled vocabularies (less than a screenful, such as 30 concepts) can be “flat” lists, without hierarchy.



Resources on Taxonomy Hierarchies

Posts from The Accidental Taxonomist Blog by Heather Hedden

- [“When a Taxonomy Should not be Hierarchical,”](#) June 20, 2020
- [“Navigation Schemes vs. Taxonomies,”](#) May 20, 2020
- [“Taxonomy Hierarchical Relationship Issues,”](#) August 30, 2018
- [“Taxonomy Hierarchy Levels,”](#) July 30, 2018
- [“Avoiding Mistakes in Taxonomy Hierarchical Relationships,”](#) February 19, 2017

"Topic Taxonomies are the Worst," by Bob Kasenchak, World IA Association / Taxonomy Talk presentation, June 21, 2023 ([Vimeo video recording](#))

Upcoming Taxonomy Presentations and Workshops

- "[Information Architecture vs. Artificial Intelligence in Taxonomies](#)"
Information Architecture for Artificial Intelligence, October 10, online
- "[Taxonomies to Ontologies: How, When and Why to Connect and/or Extend](#)"
HS Events' Semantic Data, October 23, New York, NY
- "[Building Taxonomies to Leverage Content](#)" half-day pre-conference workshop,
LavaCon content strategy conference, October 27, Portland, OR
- "[The Complete Guide To Sourcing Terms](#)"
Taxonomy Boot Camp, November 18, Washington, DC
- "[Taxonomy Design Best Practice for Knowledge Graphs](#)" 2-hour masterclass,
Connected Data London, December 11, London, UK

Questions/Contact

Heather Hedden

Taxonomy Consultant

Hedden Information Management

Carlisle, MA USA

www.hedden-information.com

accidental-taxonomist.blogspot.com

www.linkedin.com/in/hedden

heather@hedden.net



**Hedden Information
Management**

Making information findable

Hedden, Heather. (2022) *The Accidental Taxonomist, 3rd edition*.
Medford, NJ: Information Today Inc.

www.hedden-information.com/accidental-taxonomist

