

# Better Living Through Taxonomies

## Got something to say?

Share your comments [[/articles/better\\_living\\_through\\_taxonomies/comments/](/articles/better_living_through_taxonomies/comments/)] on this topic with other web professionals

In: Articles [</types/articles/>]

By Heather Hedden [[/about/contributors/heather\\_hedden/](/about/contributors/heather_hedden/)]

Published on February 5, 2008

Large websites and intranets can benefit from improved methods of search and navigation. These include site maps, A-Z indexes, sophisticated search engines, and generally improved navigational design—and playing a potential role in all of these methods is well-planned taxonomy.

## What is a taxonomy?

For a small, simple taxonomy, you can create lists of terms in Excel spreadsheets for importing, but a more sophisticated taxonomy tool is needed if you want to support multiple non-preferred terms (synonyms) for each term, related-term relationships, term notes (definitions), more than one broader-term relationship for the same term (polyhierarchies), and attributes or categories for terms. A dedicated taxonomy tool will also check for duplicate terms, unlinked terms (orphans), and circular relationships, and adjust relationships if a term is re-named or deleted.

There are both different kinds of taxonomies and different definitions of taxonomies, so it is understandable that there exists uncertainty regarding what a *taxonomy* actually is. When we think of a taxonomy what often comes to mind is a tree-type hierarchy of terms for classifying things such as plants or animals. The word *taxonomy* originally meant the science of classifying things, but it has since become a popular term for any hierarchical classification or categorization system—thus, we no longer speak of *taxonomy* as a science, but rather *a taxonomy* (plural: taxonomies) as a kind of controlled vocabulary that has a hierarchy. The taxonomy terms are arranged or linked so that narrower/more specific/child terms fall under broader/more generic/parent terms. An example of a term hierarchy might be:

- Computers and internet
  - Software
    - Internet software
      - World Wide Web software
        - HTML editors

A controlled vocabulary, at a minimum, is “a restricted list of words or terms used for indexing or categorizing”. Most controlled vocabularies have the additional feature of cross-references pointing from a non-preferred term (synonym, variant, alternate spelling, etc.) to the preferred term. The objectives of a controlled vocabulary are to ensure consistency in indexing, tagging, or categorizing and to guide the user to the desired information.

Different kinds of taxonomies and controlled vocabularies are known by different names. If a controlled vocabulary is merely a list of concepts with their synonyms, and none of the synonyms is designated as the preferred term, it is called a *synonym ring*. A more complex controlled vocabulary, with hierarchical (broader-term/narrower-term), related-term, and used-from relationships (and often also term notes) is called a *thesaurus*. A more complex thesaurus with customized semantic relationships such as “located in”, “utilized for the purpose of”, “is a member of”, “is owned by”, etc., is known as an *ontology*. A taxonomy that consists of multiple smaller hierarchies, one for each type of term or characteristics (or facet), which can be searched in combination, is called a *faceted taxonomy*. (Typical facets might be People, Organizations, Topics, Products, Locations, and Activities.)

Recently the term taxonomy has also become popular as the term for any kind of controlled vocabulary or classification system, whether a simple standard glossary, a highly structured ontology, or anything in between. A taxonomy is thus either—or both—a **hierarchical classification scheme** and/or a **controlled vocabulary**. Both these characteristics of a taxonomy make it useful for improving the navigation and findability of a website.

## Hierarchies and information architecture

In print, pages of content are usually arranged in a linear sequential form; in digital format, pages of content are often arranged hierarchically. Obviously a linear format will not be followed in hyperspace, where users can click in any direction they want. But why a hierarchy? People “surf” the Web instead of reading everything in one place; the main ideas, overviews, and executive summary pages are the top level pages in a website, whereas more specific, detailed topics are at lower levels, requiring more clicks. This also occurs (to a limited degree) in print in daily newspapers, where the main headlines are on the front page with stories continued on later pages. The structural design (i.e. information architecture) of websites, therefore, has adopted a hierarchical format with main pages branching from the home page, and lower level pages branching from each of the main pages in an inverted tree-like structure. The structure is reflected in the navigation menu, often with second-level pages in pull-down sub-menus, and in the site map.

Member Services	IA Network	Learning IA	About Us
<ul style="list-style-type: none"> <li>^ Join IAI</li> <li>^ Member Benefits</li> <li>^ Job Board               <ul style="list-style-type: none"> <li>^ Post a Job</li> <li>^ View current jobs</li> </ul> </li> <li>^ Mentoring Program               <ul style="list-style-type: none"> <li>^ What is mentoring?</li> <li>^ What to expect from a mentor/protégé</li> <li>^ How to choose a mentor</li> <li>^ Sample mentoring agreement</li> </ul> </li> <li>^ Grant Program               <ul style="list-style-type: none"> <li>^ Progress Grant Details</li> <li>^ Grant Recipients</li> </ul> </li> <li>^ Event Sponsorship</li> <li>^ Email Lists               <ul style="list-style-type: none"> <li>^ IAI Newsletter</li> <li>^ IAI Members List</li> <li>^ Job Board List</li> <li>^ IAI Meta List</li> <li>^ Second Life Initiative List</li> </ul> </li> <li>^ Member Center</li> </ul>	<ul style="list-style-type: none"> <li>^ Networking Guide</li> <li>^ IAI Local Groups               <ul style="list-style-type: none"> <li>^ Local Group FAQ</li> <li>^ Start a Local Group</li> <li>^ Local Group Coordinator Survey</li> </ul> </li> <li>^ GeoLocate IAs</li> <li>^ Find a Network</li> <li>^ Find a Project               <ul style="list-style-type: none"> <li>^ Education</li> <li>^ IA Glossary</li> <li>^ Local Groups</li> <li>^ IA Library</li> <li>^ IA Tools</li> <li>^ Translation of IA</li> <li>^ Second Life</li> <li>^ Job Board</li> </ul> </li> <li>^ Discuss IA</li> <li>^ Contribute Documents</li> </ul>	<ul style="list-style-type: none"> <li>^ Research               <ul style="list-style-type: none"> <li>^ Salary Survey</li> <li>^ Industry Surveys</li> <li>^ IA White Papers</li> <li>^ IAI Grant Research</li> </ul> </li> <li>^ Education               <ul style="list-style-type: none"> <li>^ Careers</li> <li>^ Schools Teaching IA</li> <li>^ IA Curriculum</li> <li>^ Selected IA Books</li> <li>^ IA Glossary</li> <li>^ Elevator Pitches</li> <li>^ Testimonials</li> </ul> </li> <li>^ Library               <ul style="list-style-type: none"> <li>^ Introduction to IA</li> <li>^ Recently Added</li> <li>^ Search the Library</li> <li>^ About the Library</li> <li>^ Browse by:                   <ul style="list-style-type: none"> <li>^ Subject</li> <li>^ Resource Type</li> <li>^ Author</li> <li>^ Language</li> <li>^ All Resources</li> </ul> </li> </ul> </li> <li>^ Tools</li> </ul>	<ul style="list-style-type: none"> <li>^ Our Mission</li> <li>^ People               <ul style="list-style-type: none"> <li>^ Board of Directors</li> <li>^ Staff</li> <li>^ Advisory Board</li> <li>^ Volunteer Initiative Leaders</li> <li>^ Founders</li> </ul> </li> <li>^ Partners</li> <li>^ Global Initiatives</li> <li>^ Annual Report</li> </ul>
			<ul style="list-style-type: none"> <li>^ Utilities               <ul style="list-style-type: none"> <li>^ Log-In</li> <li>^ Join</li> <li>^ Contact Us</li> <li>^ Privacy Policy</li> <li>^ About this Site</li> </ul> </li> </ul>

Site map of the Information Architecture Institute web site exemplifying the three-level hierarchical structure

It goes without saying, then, that developing a good hierarchical structure is important for creating a well-designed and easy to navigate website. By understanding the fundamentals and best practices of taxonomy development, web designers and information architects can design better websites. This involves knowing whether concepts or topics are indeed of a broader-narrower (parent-child) relationship and not merely an associated relationship. A concept can be narrower to another concept only if it is a kind of, instance of, or part of the broader concept.

## Controlled vocabularies and findability

Another application of taxonomies in the digital medium is to aid findability and search. In addition to the hierarchical nature of the taxonomy, the naming of the terms or concepts must be carefully considered. Terms should be unambiguous, yet consistent in style. The category may also need to be concise to fit within menu labels. Terms used on menu labels should be used consistently in page titles and other areas of a site—this consistency aids in the navigation of a site.

When we speak of “search” we probably think of a search engine. For many of us, how a search engine works is some kind of mathematical mystery; there are algorithms for the frequency, location, and proximity of words and phrases, but often what it comes down to is having the words or phrase entered by the user match words or phrases in the text. If, however, a synonym or different wording is used, such as “car” instead of “automobiles”, relevant pages will be missed. A taxonomy, at least of the synonym ring type, can significantly improve search engine results. Instead of merely going straight from the user-entered search-string to the text, the search engine first searches for the word or phrase on a synonym look-up table, whose terms are in turn matched to the text of the documents. A synonym-ring controlled vocabulary is typically

not seen by the user.

A user-browsable taxonomy used with a search engine is another possibility. If browsed by the user, the hierarchical nature of a taxonomy is then also important.

Higher-end search engines may come with a taxonomy creation and maintenance feature, but—except for a very few customized enterprise search systems—a search engine rarely comes with a pre-built taxonomy. But even the simplest search engines, including freeware, provide the option to configure the search to target certain fields, including the keyword meta-tags. If you configure your search engine to search only the keywords tags and not the full text, and you ensure all pages are properly and consistently tagged with keywords, you can reap the benefits of a taxonomy. Note that the keywords need to be selected by the (human) taggers from a controlled vocabulary list to ensure consistency in application and retrieval.

A database is not necessary to make use of a taxonomy, but certain kinds of content are better categorized and searched through a database structure. A taxonomy that is based on facets and supports a simultaneous query on multiple facets in combination (such as activity, product, and location) requires that its content be stored in a database.

## **What kind of sites are best candidates for taxonomies?**

The design of any site can benefit from the basics of classification and categorization, although complex taxonomies are not needed for a small site. Large sites, though, such as intranets and web sites of large diversified corporations, government agencies, and universities can definitely benefit from a carefully planned taxonomy to aid both navigation and search. Sites that are collections of articles, images, or other data records can also be more easily searched with a taxonomy.

### Articles by Topic

Discover	Design	Build	Evaluate
<ul style="list-style-type: none"> <li>• <a href="#">Basics</a></li> <li>• <a href="#">Blogging</a></li> <li>• <a href="#">Business</a></li> <li>• <a href="#">Client Management</a></li> <li>• <a href="#">Content Management Systems (CMS)</a></li> <li>• <a href="#">Convergence</a></li> <li>• <a href="#">E-Commerce</a></li> <li>• <a href="#">E-Learning</a></li> <li>• <a href="#">E-Marketing</a></li> <li>• <a href="#">Globalization and Internationalization</a></li> <li>• <a href="#">History of Technology</a></li> <li>• <a href="#">Intranets</a></li> <li>• <a href="#">Knowledge Management</a></li> <li>• <a href="#">Legal</a></li> <li>• <a href="#">Planning</a></li> <li>• <a href="#">Technology</a></li> <li>• <a href="#">Web Guru</a></li> <li>• <a href="#">Web Teams</a></li> </ul>	<ul style="list-style-type: none"> <li>• <a href="#">Audio</a></li> <li>• <a href="#">Color</a></li> <li>• <a href="#">Databases</a></li> <li>• <a href="#">Graphic Design</a></li> <li>• <a href="#">Illustration</a></li> <li>• <a href="#">Information Architecture</a></li> <li>• <a href="#">Information Design</a></li> <li>• <a href="#">Interaction Design</a></li> <li>• <a href="#">Keep It Simple (KISS)</a></li> <li>• <a href="#">Liquid Web Design</a></li> <li>• <a href="#">Motion Graphics</a></li> <li>• <a href="#">Photography</a></li> <li>• <a href="#">Prototyping</a></li> <li>• <a href="#">Redesign</a></li> <li>• <a href="#">Search</a></li> <li>• <a href="#">Typography</a></li> <li>• <a href="#">User Experience</a></li> <li>• <a href="#">User-Centered Design (UCD)</a></li> <li>• <a href="#">Web Design</a></li> </ul>	<ul style="list-style-type: none"> <li>• <a href="#">API</a></li> <li>• <a href="#">Browsers</a></li> <li>• <a href="#">CSS</a></li> <li>• <a href="#">Content</a></li> <li>• <a href="#">DOM</a></li> <li>• <a href="#">Email</a></li> <li>• <a href="#">Flash</a></li> <li>• <a href="#">HTML</a></li> <li>• <a href="#">Mailing Lists</a></li> <li>• <a href="#">Navigation</a></li> <li>• <a href="#">PHP</a></li> <li>• <a href="#">Programming</a></li> <li>• <a href="#">RSS</a></li> <li>• <a href="#">SVG</a></li> <li>• <a href="#">Scripting</a></li> <li>• <a href="#">Web Standards</a></li> <li>• <a href="#">XHTML</a></li> <li>• <a href="#">XML</a></li> </ul>	<ul style="list-style-type: none"> <li>• <a href="#">Accessibility</a></li> <li>• <a href="#">Community</a></li> <li>• <a href="#">Critique</a></li> <li>• <a href="#">Human-Computer Interaction (HCI)</a></li> <li>• <a href="#">ROI</a></li> <li>• <a href="#">Search Engine Optimization (SEO)</a></li> <li>• <a href="#">Usability</a></li> <li>• <a href="#">Web Analytics</a></li> <li>• <a href="#">Web Maintenance</a></li> </ul>

Digital Web uses a 2-level taxonomy to categorize its articles by topics

Ecommerce sites, in which each product is categorized, are excellent candidates for faceted taxonomies. There can be facets for product type, purpose, customer-type, price-range, and other characteristics such as color or size. The user chooses a term/value from each facet and searches for results that meet the requirements of all chosen facets combined.

Consumer-oriented sites that do not include an ecommerce feature might utilize a single hierarchical taxonomy instead, such as in the case of the Verizon Superpages site’s category browse:

Arts & Entertainment >	Construction & Contractors >	Media & Communications >
Automotive >	Education >	Personal Care & Services >
Business & Professional Services >	Food & Dining >	Real Estate >
Clothing & Accessories >	Health & Medicine >	Shopping >
Community & Government >	Home & Garden >	Sports & Recreation >
Computers & Electronics >	Industry & Agriculture >	Travel & Transportation >
	Legal & Financial >	

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](#), ...

Verizon Superages has a single hierarchy of 19 top level categories on the first page, such as Arts & Entertainment, each of which of 20-30 second-level categories

## How are taxonomies created?

While many content management systems and search engines support the integration of a taxonomy, most do not include full capabilities to create and manage a taxonomy within the system. Rather, taxonomy construction must be created offline in a different tool and imported into the CMS or search engine.

For a small, simple taxonomy, you can create lists of terms in Excel spreadsheets for importing, but a more sophisticated taxonomy tool is needed if you want to support multiple non-preferred terms (synonyms) for each term, related-term relationships, term notes (definitions), more than one broader-term relationship for the same term (polyhierarchies), and attributes or categories for terms. A dedicated taxonomy tool will also check for duplicate terms, unlinked terms (orphans), and circular relationships, and adjust relationships if a term is re-named or deleted.

Affordable desktop taxonomy/thesaurus creation tools that can export complex taxonomies as XML or HTML include MultiTes [<http://www.multites.com/>], Webchoir TCS-10

[<http://www.webchoir.com/products/tcs10.html>] , and Term Tree 2000 [<http://www.termtree.com.au/html/termtree.html>] . Higher-end multi-user systems combine distributed taxonomy creation with indexing features; these include Synaptica [<http://solutions.dowjones.com/djcs/index.asp>] , WordMap [<http://www.wordmap.com/>] , SchemaLogic [<http://www.schemalogic.com/>] , and Data Harmony's Thesaurus Master [<http://www.dataharmony.com/products/tm.html>] . Finally, there are several enterprise search engines that feature auto-categorization in combination with taxonomies.

Use of the software alone does not necessarily result in a usable taxonomy, just as a good HTML editing tool does not guarantee a good website if the user is not skilled in the techniques of website design. A good taxonomy requires additional skills and knowledge in categorization practices. The best qualified people to build taxonomies are those with a background in library and information science, or experience in creating thesauri for indexed article retrieval. Often indexers of books or periodical articles transition well into taxonomy work. Other individuals could make good taxonomists if they have the right analytical skills, take a course or workshop on taxonomy creation, and read up on the field.

If you want to start a taxonomy project for your organization, first check with your corporate librarian, if you have one. If you have a large taxonomy project, especially one that involves more content than just what's on a website, there are several taxonomy consultancies that can help you. If you have a smaller taxonomy project that you can manage but not execute yourself, you can contract a freelance taxonomist.

## Summary

Whether your goal is to improve the organizational structure of your website or boost the effectiveness of the search engine on your site, a taxonomy of one kind or another can be of great help.

## Resources

- Taxonomy Community of Practice Wikispace [<http://taxocop.wikispaces.com/>]
- Taxonomies & Controlled Vocabularies SIG [<http://www.taxonomies-sig.org/>]
- National Information Standards Organization: ANSI/NISO Standard 39.19 – Guidelines for the Construction, Format and Management of Multilingual Controlled Vocabularies [[http://www.niso.org/standards/standard\\_detail.cfm?std\\_id=814](http://www.niso.org/standards/standard_detail.cfm?std_id=814)]
- Introductory Tutorial on Thesaurus Construction [<http://publish.uwo.ca/~craven/677/thesaur/main00.htm>]
- Construction of Controlled Vocabularies: A Primer [<http://www.slis.kent.edu/%7Emzeng/Z3919/index.htm>]
- Taxonomy Community of Practice Yahoo group [<http://finance.groups.yahoo.com/group/TaxoCoP/?yguid=24812470>]

## Got something to say?

Share your comments with other professionals [[/articles/better\\_living\\_through\\_taxonomies/comments/](/articles/better_living_through_taxonomies/comments/)]  
(2 comments)

- Add to Magnolia [[http://ma.gnolia.com/bookmarklet/add?url=http://www.digital-web.com/articles/better\\_living\\_through\\_taxonomies&description=Better+Living+Through+Taxonomies](http://ma.gnolia.com/bookmarklet/add?url=http://www.digital-web.com/articles/better_living_through_taxonomies&description=Better+Living+Through+Taxonomies)]
- Add to My Yahoo! [<http://myweb2.search.yahoo.com/myresults/bookmarklet?u=http://www.digital->

- web.com/articles/better\_living\_through\_taxonomies&t=Better+Living+Through+Taxonomies]*
- Add to Newsvine [*http://www.newsvine.com/\_tools/seed&save?u=http://www.digital-web.com/articles/better\_living\_through\_taxonomies&h=Better+Living+Through+Taxonomies]*
  - Digg this story [*http://digg.com/submit?phase=2&url=http%3A//www.digital-web.com/articles/better\_living\_through\_taxonomies&title=Better+Living+Through+Taxonomies]*
  - Add to Del.icio.us [*http://del.icio.us/post?url=http://www.digital-web.com/articles/better\_living\_through\_taxonomies&title=Better+Living+Through+Taxonomies]*

**Related Topics:** Navigation [*/topics/navigation]* , Information Design [*/topics/information\_design]* , Information Architecture [*/topics/information\_architecture]*



Heather Hedden [*/about/contributors/heather\_hedden/*] is an information taxonomist with Viziant Corporation [*http://www.viziantcorp.com*] in Boston, MA. She also does consulting through Hedden Information Management [*http://www.hedden-information.com*] and teaches continuing education workshops through Simmons College Graduate School of Library and Information Science [*http://www.simmons.edu/gslis/continuinged/workshops*] .

Copyright © 1994-2009 Digital Web Magazine. All Rights Reserved.