



Controlled and 'Uncontrolled' Vocabularies

Taylor and Joudrey
Chapter 10
Pages 349 - 368

Presentation by Chase Hanes
LIS 640 (Thursday Section)

PART ONE: CONTROLLED VOCABULARIES

Choose
your
WORDS
carefully

CONTROLLED VOCABULARIES

Controlled Vocabularies are created when all the various terms or phrases representing a concept are gathered together in a list (or database) of subject terms (Taylor and Joudrey 334)

Three types of controlled vocabularies:

- 1.) Subject Heading Lists
- 2.) Thesauri
- 3.) Ontologies



*All of these have merit, but they each present their own limitations.

SUBJECT HEADING LISTS

Libraries (rather than other domains) have created subject heading lists to offer consistent and predictable terminology.

Frequently Accessed Subject Heading Lists:

- Library of Congress Subject Headings (LCSH)
- Sears List of Subject Headings (Sears)
- Medical Subject Headings (MeSH)

LIBRARY OF CONGRESS SUBJECT HEADINGS (LCSH)

- Print version is produced once a year
- Electronic version is updated continuously
 - **updates can correspond to movements and trends in social justice platforms**
- Not restricted to any specific subject domain
- Intended to be used by all kinds of libraries in different settings and in different countries
 - **this means it is VERY inclusive!**

From [Library of Congress Subject Headings](#)

Details

Visualization


Suggest Terminology

Dogs

URI(s)

- > <http://id.loc.gov/authorities/subjects/sh85038796>
- > <info:lc/authorities/sh85038796>
- > <http://id.loc.gov/authorities/sh85038796#concept>

Instance Of

- > [MADS/RDF Topic](#)
- > [MADS/RDF Authority](#)
- > [SKOS Concept](#) 

Scheme Membership(s)

- > [Library of Congress Subject Headings](#)



Collection Membership(s)

- > [LCSH Collection - Authorized Headings](#)
- > [LCSH Collection - General Collection](#)
- > [LCSH Collection - May Subdivide Geographically](#)

Variants

- >  Canis canis
- >  Canis domesticus
- >  Canis familiaris
- >  Canis familiaris domesticus
- >  Canis lupus familiaris
- >  Dog
- >  Domestic dog

Broader Terms

- >  [Domestic animals](#)
- >  [Gray wolf](#)

Narrower Terms

- >  [Antis \(Dog\)](#)
- >  [Atticus \(Dog\)](#)
- >  [Balto \(Dog\)](#)
- >  [Bamse \(Dog\)](#)
- >  [Barney \(Dog\)](#)

SEARS LIST OF SUBJECT HEADINGS

- Intended for small collections
- Refrains from being very specific
- Used for people with general information needs
- Typically utilized in school and small (to medium-sized) public libraries
- Comprised of general terms
- Not conducive to seasoned researchers

Abbeys (May subdiv. geog.) 271; 726
SA names of individual abbeys [to be added as needed]
BT **Church architecture**
Monasteries
NT **Westminster Abbey**
RT **Cathedrals**

http://www.hwwilsoninprint.com/pdf/sears_pgs.pdf

[Eukaryota \[B01\]](#)

[Animals \[B01.050\]](#)

[Chordata \[B01.050.150\]](#)

[Vertebrates \[B01.050.150.900\]](#)

[Mammals \[B01.050.150.900.649\]](#)

[Carnivora \[B01.050.150.900.649.147\]](#)

[Canidae \[B01.050.150.900.649.147.153\]](#)

[Coyotes \[B01.050.150.900.649.147.153.150\]](#)

► [Dogs \[B01.050.150.900.649.147.153.200\]](#)

[Foxes \[B01.050.150.900.649.147.153.250\]](#)

[Jackals \[B01.050.150.900.649.147.153.400\]](#)

[Raccoon Dogs \[B01.050.150.900.649.147.153.700\]](#)

[Wolves \[B01.050.150.900.649.147.153.900\]](#)

MEDICAL SUBJECT HEADINGS (MeSH)

- Similar to a thesaurus in that it offers a strict hierarchical structure and focuses on specific subject areas
- Qualifies as a subject heading list because it precoordinates phrases

MeSH Heading	Dogs
Tree Number	B01.050.150.900.649.147.153.200
Annotation	check tag: no qualif; when IM, qualif permitted; TN 3: spelling of breeds
Scope Note	The domestic dog, <i>Canis familiaris</i> , comprising about 400 breeds, of the carnivore fam p1065)
Entry Term	<i>Canis familiaris</i>
Consider Also	consider also terms at CANINE
Allowable Qualifiers	AB AH BL CF CL EM GD GE IM IN ME MI PH PS PX SU UR VI
Date of Entry	19990101
Unique ID	D004285

- Updated annually
- Provides subject access points for every bibliographic record that the National Library of Medicine makes
- Provided free online

https://www.nlm.nih.gov/cgi/mesh/2016/MB_cgi

THESAURI

Thesauri, crafted in indexing communities, are very hierarchical and narrower in scope.

Frequently accessed thesauri:

- Art & Architecture Thesaurus (AAT)
- Thesaurus of ERIC Descriptors

ART & ARCHITECTURE THESAURUS (AAT)

- Assists in verbal access to all kinds of cultural heritage info
- Terms describe objects, textual materials, images, architecture
- Used in archives, libraries, museums, visual resources collections, and conservation spaces
- Arranged into seven facets that move from abstract to concrete
 - **this, along with FRBR, proves that catalogers MUST think conceptually and with high logic

Hierarchical Position:



Objects Facet

.... Visual and Verbal Communication (hierarchy name) (G)
..... Visual Works (hierarchy name) (G)
..... visual works (works) (G)
..... <visual works by material or technique> (G)
..... paintings (visual works) (G)
..... <paintings by material or technique> (G)
..... frescoes (paintings) (G)

http://www.getty.edu/vow/AATFullDisplay?find=fresco&logic=AND¬e=&english=N&prev_page=1&subjectid=300177433

THESAURUS OF ERIC DESCRIPTORS

- ERIC: Educational Resources Information Center
- Provides Access to education-related literature
- Indexes journal articles, descriptions/evaluations of programs, book reviews, research reports, curriculum and teaching guides, instructional materials, etc.
- Offers various means of display for nuanced searching
 - Alphabetical
 - Rotated
 - Hierarchical
 - Descriptor Group

ONTOLOGIES

- Ontologies are evolving to aid in the knowledge transfer between computers and humans.
- They extend from the computer science community and generate complex keyword searches.
- Taylor and Joudrey explain that an “ontology defines the nature of reality by identifying the concepts, entities, terms, and categories in a particular domain in order to model the relationships among them” (357).
- They aim to keep conceptual and semantic ambiguity at a minimum.
 - Meaning derived from language (spoken, written, etc.) is affected by context and the symbol systems our cultures attribute to words, phrases, and syntax. Ontologies seem to try to navigate towards the appropriate meanings of specific terms/phrases.

ONTOLOGIES (continued)

- People, databases, and applications use ontologies.
- Ontologies undergird and construct the Semantic Web in their incorporation of XML and RDF.
- Ontologies, according to Taylor and Joudrey, are “important if agents on the Web are to be able to search and/or merge information from diverse communities. This is because the same term may be used in different contexts with different meanings” (358).

PART TWO: NATURAL LANGUAGE OR 'UNCONTROLLED' APPROACHES TO SUBJECTS

Choose
your
DEEDS
carefully

NATURAL LANGUAGE OR 'UNCONTROLLED APPROACHES TO SUBJECTS

Types of uncontrolled approaches:

- 1.) Natural Language Processing
- 2.) Keyword Searching
- 3.) Tagging and Folksonomies

*All of these have merit, but they each
present their own limitations.



NATURAL LANGUAGE PROCESSING (NLP)

- Written language processing (rather than spoken word processing)
- Creates information retrieval systems that
 - “1. interpret users’ information needs as expressed in free text
 - 2. represent the complete range of meaning conveyed in documents
 - 3. Understand when there is a match between the user’s information need and...the documents that meet it” (Taylor and Joudrey 361)
- **LANGUAGE ISSUES, NUANCES, AND CONSTRUCTS MUST BE ADDRESSED
 - Ex: The door opened. Susan walked in → Did Susan open the door?
 - Ex: Where’s the water? → Are you thirsty, are you looking for the ocean?
- Ambiguities can sometimes be sorted out through the context of nearby words and sentences, but not always.

KEYWORDS

- “The success of keyword searching depends upon at least two assumptions:
 - 1.) that authors writing about the same concepts will use the same words in their writings, and
 - 2.) that searchers will be able to guess what words those authors used for the concept” (Taylor and Joudrey 363).
- Issues:
 - Multiple terms represent the same concept
 - Not all related (‘adjacent’) materials are received
 - Irrelevant materials can come up
- Complex keyword search functions consider synonymous adjectives to noun keywords, hypernymns, hyponyms, and other related terms to the initial keyword.
- The complex keyword functions can make up large lexical ontologies.

TAGGING AND FOLKSONOMIES



subject metadata for the
people, by the people!



This puppy
contributes tags
and you should
too!

TAGGING AND FOLKSONOMIES

- Tagging is a populist approach to subject description.
- Mass users apply their own keywords to web resources.
- Information organization and retrieval becomes **COLLABORATIVE!**
- Tags can sometimes be displayed in tag clouds.
- Everyday resource users add their own vocabulary/ idiosyncrasies to resources



This puppy
contributes tags
and you should
too!

TAGGING AND FOLKSONOMIES

Issues:

- Tagging lacks all of the benefits of controlled vocabularies.
- A large number of tags need to be assigned to records in order to add extra insight/ supplementary access points.

QUESTIONS?

Is there a certain model—within the controlled vocabulary sphere or within a natural language approach context—that you would prefer?

- Does it depend on the resources themselves and the demographics of your patron base?