

**Dagobert Soergel**

Department of Library and Information Studies  
Graduate School of Education  
University at Buffalo  
+1-716-645-1474 Fax +1- (716) 645-3775  
dsoergel@buffalo.edu www.dsoergel.com

# **Knowledge Organization Systems in Digital libraries**

## **Tutorial**

Note: The pdf does not include all pages that are in the paper tutorial booklet

**European Conference on Research and Advanced Technology  
for Digital Libraries (ECDL 2009)  
Corfu, Greece  
Sept. 27, 2009**

## Abstract

This introductory workshop is intended for anyone concerned with subject access to digital libraries. It provides a bridge by presenting methods of subject access as treated in an information studies program for those coming to digital libraries from other fields. It will elucidate through examples the conceptual and vocabulary problems users face when searching digital libraries. It will then show how a well-structured thesaurus or other KOS (Knowledge Organization System) can be used as the knowledge base for an interface that can assist users with search topic clarification (for example through browsing well-structured hierarchies and guided facet analysis) and with finding good search terms (through query term mapping and query term expansion — synonyms and hierarchic inclusion). It will touch on cross-database and cross-language searching as natural extensions of these functions. The workshop will cover the KOS structure needed to support these functions: Concept-term relationships for vocabulary control and synonym expansion, conceptual structure (semantic analysis, facets, and hierarchy) for topic clarification and hierarchic query term expansion. It will introduce a few sample thesauri and other KOS to illustrate these principles. Lastly the workshop will give a checklist for evaluating thesauri and other KOS.

## Course objectives

Participants should appreciate the complexity of subject access and understand the problems that a thesaurus or other KOS (Knowledge Organization System) can help solve.

Participants should understand the principles of KOS structure.

Participants should be able to apply KOS structure to solving subject access problems.

Participants should be able to identify and evaluate KOS suitable for a specific situation defined by a user community and by the collection of a digital library.

## Brief biography of the Instructor

Dagobert Soergel holds an MS equivalent in mathematics and physics (1964) and a PhD in political science (1970), both from the University of Freiburg, Germany. He is Professor of Information Studies, University of Maryland, where he teaches courses in information retrieval, thesaurus development, expert systems, and information technology, and an information systems consultant. He has been a visiting professor at the universities of Western Ontario, Chicago, and Konstanz, Germany. Among other books, he has authored *Organizing Information* (1985), which received the American Society of Information Science Best Book Award, *Indexing Languages and Thesauri. Construction and Maintenance* (1974) and numerous papers. He has designed several thesauri, most recently the Alcohol and Other Drug Thesaurus (for which he chairs the advisory committee) and is developing TermMaster, a thesaurus management software package. In 1997 he received the American Society of Information Science Award of Merit.

## Part 1. Structure and use. Outline

<b>9:00 - 10:00</b>	<b>KOS functions</b>	1
9:00 - 9:05	<b>Introduction. Challenges for digital libraries</b>	1
9:05 - 9:10	<b>Why thesauri: a first look with examples</b>	2
	User orientation in a concept space; avoid vocabulary confusion	
9:10 - 9:20	<b>What is a KOS? A first look with examples</b>	17
9:20 - 9:25	<b>KOS development principles</b>	25
9:25 - 10:00	<b>KOS functions</b>	27
	User-centered indexing	41
<b>10:00 - 10:40</b>	<b>KOS structure</b>	47
9:50 - 10:00	<b>Concept-term relationships</b>	48
10:00 - 10:40	<b>Conceptual structure: Semantic analysis and facets. Hierarchy</b>	53
<b>10:40 - 12:00</b>	<b>Implementation, evaluation, resources</b>	61
10:40 - 11:00	<b>Implementing KOS functions</b>	61
	Web-based display and search in the KOS	61
	Searching interaction using facets, Web search, multilingual	75
<b>11:00 - 11:30</b>	<b>Break</b>	
11:30 - 11:40	<b>Implementing KOS functions, continued</b>	
11:40 - 11:55	<b>Evaluation of thesauri. Yahoo classification as an example</b>	101
	Framework and process	115
11:55 - 12:00	<b>Resources</b>	123
<b>12:00 - 1:00</b>	<b>Examples of classifications and thesauri</b> after p.262, blue divider	
	<b>Alcohol and Other Drug Thesaurus (AOD Thesaurus)</b>	
	US National Institute on Alcohol Abuse and Alcoholism (NIAAA)	
	<b>Medical Subject Headings (MeSH) and Unified Medical Language System (UMLS)</b>	
	US National Library of Medicine (NLM)	
	<b>Art and Architecture Thesaurus (AAT).</b> Getty Foundation	
	<b>Dewey Decimal Classification.</b> US Libr. of Congress & OCLC/Forest Pr	
	<b>WordNet.</b> Princeton University, George Miller	
	<b>CYC Ontology</b>	

## Part 2. Design and development. Outline

<b>14:30 - 14:35</b>	<b>Introduction and overview</b>	160
	KOS development principles	162
<b>14:35 - 15:05</b>	<b>The process of KOS development</b>	164
13:35 - 13:40	<b>The overall process of KOS development</b>	165
13:40 - 13:55	<b>Sources of concepts, terms, relationships, definitions</b>	166
	<b>Methods of data collection</b>	
13:55 - 14:05	<b>Merging data from many sources</b>	171
<b>15:05 - 16:00</b>	<b>Developing the conceptual structure</b>	173
14:05 - 14:30	<b>Facet analysis 1: Education (starting with classes from DDC)</b>	174
14:30 - 14:40	<b>More facet examples: Job titles</b>	176
14:40 - 14:50	<b>Guidelines for meaningful arrangement</b>	179
14:50 - 15:00	<b>Rules for selection of preferred terms</b>	187
	<b>Rules for selection of concepts as descriptors</b>	
<b>16:00 - 16:30</b>	<b>Break</b>	
<b>16:30 - 17:15</b>	<b>Developing the conceptual structure, continued</b>	
16:30 - 17:15	<b>Facet exercise (in pairs)</b>	177
<b>17:15 - 18:15</b>	<b>The structure and processing of KOS data</b>	189
17:15 - 17:30	<b>Interoperability of thesauri/ontologies. Crosswalks</b>	189
17:30 - 17:45	<b>The structure of a KOS/ontology database</b>	193
17:45 - 17:55	<b>The many forms of Knowledge Organization Systems (KOS)</b>	203
	and their standards	
	SKOS (Simplified Knowledge Organization Systems) example	217
	OWL (Web Ontology Language) example	223
17:55 - 18:15	<b>KOS software and its evaluation</b>	209
	<b>Appendix. The KOS development process</b>	257

## **Challenges for digital libraries**

Improve retrieval effectiveness to handle the sheer mass of material

Provide unified access to materials in different media (esp. access to non-text materials)

Provide knowledge-based support for end users who access networked information without the benefit of an intermediary

Support creation and maintenance of personal or work-group information systems

Support information seeking as an integral part of problem solving, learning, and intellectual work

Support collaborative work:

Scholarly communication as computer-supported multi-party conversation

KOS (Thesauri, ontologies, taxonomies, ...) must support these functions

### **Support information seeking as an integral part of problem solving, learning, and intellectual work**

Help users to explore ideas in conjunction with exploring information

Support fine-grained retrieval and assimilation of information

Support processing of information along with or after retrieval

### **Support collaborative work**

Make users full participants in the continuing improvement of information systems through feedback and other contributions

Establish linkages between people

# **Why thesauri /KOS: A first look with examples**

## **Problems**

**Vocabulary confusion**

**User orientation in a concept space**

**Queries illustrating these problems**

## Queries:

### **Synonym expansion and Hierarchic expansion**

Query 1. Drug use by teenagers

Query 1.1 teenage\* AND drug\*

Query 1.2 Synonym expansion for teenage\*

(teenage\* OR teen OR teens OR youth\* OR adolescent\* OR kid\* OR "high school")  
AND drug\*

Query 1.3 In addition, synonym expansion and hierarchic expansion for drug\*

(teenage\* OR teen OR teens OR youth\* OR adolescent\* OR kid\* OR "high school")  
AND (drug\* OR substance\* OR alcohol OR nicotine OR smoking OR cigarette\* OR mari\*una OR cocaine OR crack OR heroin)

Query 1.4 Query more narrowly focused

(teenage\* OR teen OR teens OR youth\* OR adolescent\* OR kid\* OR "high school")  
AND (cocaine OR crack OR heroin)

**Query 1.1. teenage\* AND drug\* (AltaVista)**

- .  
About 30 documents match your query.

**1. CEIDA Druglinks - Info Centre - PARENTS TALKING TO TEENAGERS ABOUT DRUGS**

What do parents want from their teenagers? Basically, parents want: To know your kids are alright and not in danger. To know your kids think you're OK...

*[http://www.ceida.net.au/info\\_centre/drug~myths/what\\_do.html](http://www.ceida.net.au/info_centre/drug~myths/what_do.html) - size 3K - 21-May-97 - English*

**2. CEIDA Druglinks - Info Centre - PARENTS TALKING TO TEENAGERS ABOUT DRUGS**

Better Ways of Communicating. Different points of view Communication is the key to resolving problems, if they exist. Or to finding out if they exist....

*[http://www.ceida.net.au/info\\_centre/drug~myths/better.html](http://www.ceida.net.au/info_centre/drug~myths/better.html) - size 9K - 21-May-97 - English*

**3. Testimony of Donna E. Shalala, Secretary of HHS on Teenage Drug Use**

Testimony of Donna E. Shalala, Secretary of Health and Human Services on Teenage Drug Use. Testimony of. Donna E. Shalala. Secretary of Health and Human...

*<http://www.apa.org/ppo/shalala.html> - size 15K - 13-Sep-96 - English*

**4. Statement of Senator Richard C. Shelby on Teenage Drug Use**

Statement of Senator Richard C. Shelby on Teenage Drug Use. Statement of. U.S. Senator Richard C. Shelby. Before The. Senate Judiciary Committee Hearing..

*<http://www.apa.org/ppo/shelbyhtml> - size 3K - 13-Sep-96 - English*

**5. Testimony of John P. Walters on Teenage Drug Use**

Testimony of John P. Walters, President of The New Citizenship Project, on Teenage Drug Use. Testimony by. John P. Walters\* President of the New...

*<http://www.apa.org/ppo/walters.html> - size 28K - 13-Sep-96 - English*

**6. Drug Use Rises for Teenagers**

Parent News for November 1996. Of Interest. Drug Use Rises for Teenagers. by Anne S. Robertson. A recent report released by the Parents Resource ...

*<http://ericps.ed.uiuc.edu/npin/pnews/pnewn96/pnewn96f.html> - size 4K - 23-May-97 - English*

**7. CEIDA Druglinks - Info Centre - PARENTS TALKING TO TEENAGERS ABOUT DRUGS**



## Query 1.2. Synonym expansion of teenager

(teenage\* OR teen OR teens OR youth OR adolescent\* OR kid\* OR "high school")  
AND drug\*

About 249 documents match your query.

### 1. Adolescent Drug Abuse Treatment Outcome

Adolescent Drug Abuse Treatment Outcome. Executive Summary. This is a report on the evaluation of an inpatient adolescent drug abuse treatment program in..

<http://www.cbc.med.umn.edu/~andy/drugabuse/adoltx.htm> - size 3K - 28-Sep-96 - English

### 2. Poll finds parents overestimate communication with kids on drugs

03/03/97 - 07:26 PM ET - Click reload often for latest version. Poll finds parents overestimate communication with kids on drugs. NEW YORK - Most parents..

<http://cgi.usatoday.com/elect/eq/eq17&htm> - size 2K - 21-May-97 - English

### 3. Albany Youth Futures shows kids alternatives to drugs, alcohol/TITLE>

<http://www.indreg.com/9-11-96/FEATURES/feature5.htm> - size 5K - 13-Sep-96 - English

### 4. IPRC Version - Keeping Youth Drug-Free - Exercise #3

Re-posted by the Indiana Prevention Resource Center at Indiana University Indiana's RADAR Network State Center. Exercise 3. Building Social Skills. Offer..

<http://www.drugs.indiana.edu/pubs/radar/keeping/exer3.html> - size 2K - 28-Jun-96 - English

### 5. Online NewsHour: Teen Drug Use Doubling -- August 20, 1996

THEY'RE NOT SAYING "NO" AUGUST 20, 1996. TRANSCRIPT. Two new and deeply troubling reports have just been released showing that drug abuse among 12 to 17...

[http://web-cr01.pbs.org/newshour/bb/health/august96/teen\\_drug~abuse\\_8-20.html](http://web-cr01.pbs.org/newshour/bb/health/august96/teen_drug~abuse_8-20.html) - size 16K - 10-Sep-96 - English

### 6. Kmart: HOTNEWS/Kmart Kids Race Against Drugs Race Results

Kmart Kids Race Against Drugs. And the winner is... On Saturday. January 18. Jamie Barreiro of Port St. Lucie, FL, Joshua Brown of Willingboro, NJ and ...

<http://www.kmart.com/hotnews/hotnews.stm>. size 7K - 21-May-97 - English

### 11. OMH-RC Database Record: Drug Abuse Among Minority Youth: Methodological Issues

Office of Minority Health Resource Center Database Record. When available, information on where these materials may be obtained has been listed below...

<http://www.omhrc.gov/mhr2/docs/95D2315.htm> - size 3K - 1-May-97 - English

### Query 1.3. Plus synonym and hierarchic expansion of "drug\*"

(teenage\* OR teen OR teens OR youth\* OR adolescent\* OR kid\* OR "high school")

AND (drug\* OR substance\* OR alcohol OR nicotine OR smoking OR cigarette\*)

About 409 documents match your query.

#### 1. Smoking is NOT for kids!

We believe smoking is for adults only. We therefore require that you be at least 18 years of age in order to view this site. Click below to enter the...

<http://www.smokers.org/> - size 820 bytes - 20-Apr-97 - English

#### 2. Adolescent Drug Abuse Treatment Outcome

Adolescent Drug Abuse Treatment Outcome. Executive Summary. This is a report on the evaluation of an inpatient adolescent drug abuse treatment program in..

<http://www.cbc.med.umn.edu/~andy/drugabuse/adoltx.htm> - size 3K - 28-Sep-96 - English

#### 3. Poll finds parents overestimate communication with kids on drugs

03/03/97 - 07:26 PM ET - Click reload often for latest version. Poll finds parents overestimate communication with kids on drugs. NEW YORK - Most parents..

[http://cgi.usatodaycomelect/eq/eq\]7&htm](http://cgi.usatodaycomelect/eq/eq]7&htm) - size 2K - 21-May-97 - English

#### 4. Albany Youth Futures shows kids alternatives to drugs, alcohol/TITLE>

<http://www.indregcoml9-11-96/FEATURESfeature5.htm> - size 5K - 13-Sep-96 - English

#### 5. IPRC Version - Keeping Youth Drug-Free - Exercise #3

Re-posted by the Indiana Prevention Resource Center at Indiana University Indiana's RADAR Network State Center. Exercise 3. Building Social Skills. Offer..

<http://www.drugs.indiana.edu/pubs/radar/keeping/exer3.html> - size 2K - 28-Jun-96 - English

#### 6. Smoking still increasing among teens

Despite a chorus of ignorance one woman wanted to dance... To all of those people who say that national role models are a thing of the past, I want to...

[http://www.baschusgamma.org/bb\\_october/staff\\_view.html](http://www.baschusgamma.org/bb_october/staff_view.html) - size 5K - 11-Oct-96 - English

#### 7. Online NewsHour: Teen Drug Use Doubling -- August 20, 1996

THEY'RE NOT SAYING "NO" AUGUST 20, 1996. TRANSCRIPT. Two new and deeply troubling reports have just been released showing that drug abuse among 12 to 17...

[http://web-cr01.pbs.org/newshour/bb/health/august96/teen\\_drug\\_abuse\\_8-20.html](http://web-cr01.pbs.org/newshour/bb/health/august96/teen_drug_abuse_8-20.html) - size 16K - 10-Sep-96 - English

#### 8. KCEOC SUBSTANCE ABUSE/YOUTH PROGRAM

KCEOC SUBSTANCE ABUSE/YOUTH PROGRAM. Address: 1611 First Street. Phone Number: 336-5310. FAX Number: 336-5303. Contact Person: Robert Cubit. Target Group..

<http://www.bakersfield.org/ydc/secondary/kceoc.html> - size 2K - 15-Oct-96 - English

**9. Kmart: HOTNEWS/Kmart Kids Race Against Drugs Race Results**

Kmart Kids Race Against Drugs. And the winner is... On Saturday, January 18, Jamie Barreiro of Port St. Lucie, FL, Joshua Brown of Willingboro, NJ and...

<http://wwwkmart.coiri/hotnews/hotnews.stm> - size 7K - 21-May-97 - English

**10. Connecticut Kidslink - Substance Abusing Mothers and Their Children**

Inter-agency Committee on Substance Abusing Mothers and Their Children in Connecticut: A Summary of Problems and Solutions. Report Summary by Andy Dodge,..

<http://statlab.stat.yale.edu/cityroom/kidslink2/welffire/texts/9603-03.html> - size 9K - 7-Nov-96 -English

**11. OMH-RC Database Record: Drug Abuse Among Minority Youth: Methodological Issues**

Office of Minority Health Resource Center Database Record. When available, information on where these materials may be obtained has been listed below...

<http://wwwomhrc.gov/mhr2/docs/95D2315.htm> - size 3K - 1-May-97 - English

**12. Browne for President - Release - teenage smoking**

NEWS FROM THE BROWNE FOR PRESIDENT CAMPAIGN. FOR IMMEDIATE RELEASE August 23, 1996. Clinton's new "War On Teenage Smoking" is moral grandstanding, charges.

<http://www.harrybrowne96.org/release-teenage-smoking.html> - size 4K - 24-Aug-96 - English

**13. Teacher Talk, 3(3), Alcohol and Adolescents**

Alcohol and Adolescents: Prevention, Intervention, Treatment, Aftercare Volume 3, Issue 3 A Publication Just for Secondary Teachers. 1996 Indiana...

<http://education.indiana.edu/cas/tt/v3i3/v3i3toctext.html> - size 2K - 6-Jun-96 - English

**14. White House Conference on Youth Drug Use**

White House Conference on Youth Drug Use. (from March/April 1996 Marijuana Policy Report) In a further attempt to defuse criticism of being soft on drugs,.

<http://www.mpp.org/youthconfhtml> - size 2K - 21-May-97 - English

**15. Anti-Smoking Software Installed at Bronx High School of Science**

Anti-Smoking Software Installed at Bronx High School of Science. March 6, 1997: The Alumni Association of the prestigious Bronx High School of Science has.

<http://www.smokefreekids.com/relO2.htm> - size 2K - 21-May-97 - English

**16. Optum: Live Event! Talking to Kids about Alcohol and Drugs**

Optum: What is happening This Month at Optum? Check here and find out.

### Query 1.4. Drug component more specific

**( teenage\* OR teen OR teens OR youth OR adolescent\* OR kid\* OR "high school")**

**AND (cocaine OR crack OR heroin)**

2 documents match your query

#### **1. Teenage "Huffing" - Worse Than Cocaine**

Teenage "Huffing" - Worse Than Cocaine. May 22, 1996. MEEUWSEN: Imagine substances experts call deadlier than heroin or cocaine. Imagine that...

*<http://www.cbn.org/news/stories/huffinghtml> - size 7K - 29-Oct-96 - English*

#### **2. Teen is arrested with a kil of crack cocaine**

Teen is arrested with a kilo of crack cocaine. STROUDSBURG, Pa. (AP) - A 14-year-old New York City girl was busted during a bus trip through here last...

*<http://www.recordernews.com/1996/0703/natnews/teenare/teenare.html> - size 2K - 25-May-97 English*



## **Queries: Homonyms and polysemes**

Query 2. wordnet (homonym: 6 meanings)

Query 3. classification (polyseme)

Query 3.1. classification AND security

## Query 2. Wordnet (homonym: 6 meanings)

### 3. WordNet: A Lexical Database for English

Lexical Resources for Human Language Technology. Princeton University.  
DARPA/ITO

<http://www.ito.darpa.mil/Summaries95/B370--Princeton.html> - size 12K -  
12-Sep-96 -

### 4. VDI - Racial WordNet Networks

Racial Recorders. The WordNet Uses The TCP/IP "internet" Protocol, Allowing  
Easy Network Utilization. Search And Playback. Recorded Messages Via The  
Lan....

<http://www.fishnet.net/~ecs/racal3.htm> - size 539 bytes - 11-Oct-96 - English

### 6. WordNet lexical database

<http://www.grafnetix.com/thesaurus/QueryExpansionIntro/node1.html> - size 6K -

### 8. WORDNET, the new generation of digital communications recorders

Digital communications logging recorder.

<http://www.abds.net/dss/wordnet.htm> - size 2K - 30-Jan-97 - English

### 13. WordNet's Christian Links

Christian Web Sites. Below is your passport to a wider Christian on-line  
community. Some contain links to many other Christian sites

<http://www.wordnet.co.uk/links.html> - size 3K - 23-May-97 - English

### 18. The Wordnet Story

Wordnet Productions. Jesus, the Divine Word, casts his net, the Good News, to all  
through mass media. Wordnet is a Catholic television ministry dedicated..

<http://www.rlagroup.com/wordnet/wrdntstr.htm> - size 2K - 6-Feb-97 - English

### 30. Tesi di Laurea WordNet

Linguaggio Naturale. Proposta per Tesi di Laurea: WordNet. WordNet e` una base  
di conoscenza lessicale per l'inglese, disponibile gratuitamente su..

<http://ecate.itc.it:1024/cirave/wordnet.html> - size 2K - 30-Sep-96 - Italian

### 48. WORDNET Language Translation Service

WORDNET is a team of language experts specializing in foreign language  
translation, typesetting and printing. In recent years, we have helped a number of..

<http://www.wordnet.com/> - size 4K - 20-Jun-97

### 52. Consortium of the EuroWordNet project

All Rights reserved by Computer Centrum Letteren University of Amsterdam.  
Coordinator: builder of...

<http://www.let.uva.nl/~ewn/consortium-ewn.html> - size 3K - 22-Apr-97 - English

### Query 3. classification (polyseme)

#### Examples from AltaVista search

#### 1. GNWT Administrative Records Classification System BUILDINGS AND PROPERTIES

BUILDINGS AND PROPERTIES - DAMAGES 2063. Records relating to damages incurred by government buildings, facilities and structures. It includes...

<http://pingo.gov.nt.ca/Records/sections/2000/1995blg9.htm> - size 4K - 17-Oct-96 - English

#### 2. LC Classification: U - Military Science

U - Military Science. U. 1-900. Military Science (General). 21-22.3. War, Philosophy, Military Sociology. 27-43. History of Military Science. 164-167.5....

<http://www.library.yorku.ca/lc/u.html> - size 6K - 13-Nov-96 - English

#### 7. Table Tennis Classification Procedures

International Paralympic Committee. Sports Science | Medical | Sports | Secretariat | General. Table Tennis Classification Procedures. A. Purposes. 1) To..

<http://info.lboro.ac.uk/research/paad/ipc/table-tennis/class-proc.html> - size 7K - 2-Jul-96

-

#### 8. MPW Public Highways (Road Classification)

ROAD CLASSIFICATION. Law No. 13 of the year 1980 (UU 13/1980) concerning roads distinguishes the category of road into public and special roads. The...

<http://www.pu.go.id/publik/binama~1/html/eng/classifi.htm> - size 3K - 22-May-96 -

#### 9. Hurricane and Tropical Storm Classification

<http://www.hiwaay.net/cwbol/scale.html> - size 3K - 7-Jul-95 - English

#### 17. DEPARTMENT OF ENERGY FUNDAMENTAL CLASSIFICATION POLICY REVIEW

Secretary Hazel O'Leary has emphasized the importance of improved public accountability

<http://www.osti.gov/html/osti/opennet/fcprsum.html> - size 10K - 11-Feb-97 - English

#### 29. Subject guide to the classification

Subject guide to the Library of Congress classification. For subjects not listed here please consult the printed, red-bound Subject Index in the entrance..

<http://potter.cc.keele.ac.uk/depts/li/lctable.htm> - size 7K - 21-May-97 - English

#### 30. BRYOPHYTES: Hornwort Classification

Phylum ANTHOCEROTOPHYTA. DENDROCEROS. Gametophyte plant with horn-like sporophyte. copyright ©1996 Southern Illinois...

<http://www.science.siu.edu/bryophytes/anthocerotophyta.html> - size 940 bytes - 5-Apr-97

-



**31. Policy & Planning Support - Staff Level Classification**

Staff Classification & Level. All staff are assigned a classification on employment. This data element indicates the classification...

<http://www.plan.murdoch.edu.au/stats/descript/clssfctn.html-ssi> - size 4K - 21-May-97

**34. Classification Reform Approval**

March 5, 1996. FOR IMMEDIATE RELEASE. Release No. 14. POSTAL SERVICE APPROVES CLASSIFICATION REFORM RECOMMENDATIONS

<http://www.usps.gov/news/press/96/96014new.htm> - size 4K - 12-Apr-97 - English

**57. Universal Decimal Classification Index 5414**

NATURAL SCIENCES. MATHEMATICS. 54 CHEMISTRY. MINERALOGY. 541 GENERAL, THEORETICAL, AND PHYSICAL CHEMISTRY. 5414 CHEMICAL.

<http://www.chem.ualberta.ca/~plambeck/udc/u5414.htm> - size 827 bytes - 9-May-97 -

**61. Draft Public Guidelines to Department of Energy Classification of Information**

<http://www.osti.gov/html/osti/opennet/document/guidline/pubgf.html> - size 17K -

**71. The GNU C Library - Classification of Characters**

This section explains the library functions for classifying characters.

[http://www.ia.pw.edu.pl/Pl-iso/tex-info/libc/libc\\_55.html](http://www.ia.pw.edu.pl/Pl-iso/tex-info/libc/libc_55.html) - size 7K - 6-Apr-94 - English

**80. Dewey Decimal Classification System**

Dewey Decimal Classification System. Overview. 000 Generalities 100 Philosophy & psychology 200 Religion 300 Social sciences 400 Language 500 Natural...

<http://www.tnrplib.bc.ca/dewey.html> - size 38K - 7-Aug-96 - English

**88. Extended Computing Reviews Classification Scheme**

Extended Computing Reviews Classification Scheme. Computing Reviews Classification System. Copyright 1994, by the Association for Computing Machinery,...

<http://www.dpmms.cam.ac.uk/MR/CRclass.html> - size 37K - 1-Sep-95 - English

**89. 627.440 - Classification of costs.**

627.440 - Classification of costs. Standard Number: 627.440. Standard Title:

Classification of costs. SubPart Number: D. SubPart Title: Administrative...

[http://www.doleta.gov/regs/cfr/20cfr/toc\\_Part600-699/0627.0440.htm](http://www.doleta.gov/regs/cfr/20cfr/toc_Part600-699/0627.0440.htm) - size 12K -

**90. Pirelli Cumbria Rally 1996 Final Classification**

Pirelli Cumbria Rally 1996. Final Classification. POSITION OVERALL CLASS NUMBER CREW CLASS TOTAL 11 201 Richard Tuthill/Nick Kennedy Vauxhall Nova 1300...

[http://www.idiscover.co.uk/tcs21/1996/pirelli/c\\_class.html](http://www.idiscover.co.uk/tcs21/1996/pirelli/c_class.html) - size 2K - 5-May-96 - English

**117. Classification of Students**

Classification of Students. Students at Bemidji State University are classified as regular, special, or auditor. Regular: A regular student is one who is..

[http://bsuweb.bemidji.msus.edu/~catalog/catalog94\\_96/classify\\_stud.html](http://bsuweb.bemidji.msus.edu/~catalog/catalog94_96/classify_stud.html) - size 1K - 15-Mar-95 - English

## Query 3. classification

### Examples from Lycos search

#### 2) Classification of Signatures

<http://www.seas.gwu.edu/faculty/pbock/SignatureCla> [99%]

#### 5) Supervised Classification

Neural Network Classification of Multispectral Imagery Supervised Classificati .  
<http://www.ece.arizona.edu/~paola/SupervisedClass.> [99%]

#### 139) RESIDENCE CLASSIFICATION

Residence Classification Residence Classification Nonresident students seeking to become California residents for tuition/fee purposes must petition t.  
<http://www.reg.uci.edu/REGISTRAR/SOC/rc.html> [99%]

#### 152) PRODUCT CLASSIFICATION

EPA may classify a pesticide product for restricted use if its characteristics warrant special handling. Restricted use pestici.  
<http://hammock.ifas.ufl.edu/txt/fairs/26668> [99%]

#### 426) Dewey Decimal Classification Home Page

DDC 21 and Dewey for Windows now available! OCLC Forest Press is pleased to announce the publication of DDC 21, the latest edition of the Dewey Decima.  
<http://www.oclc.org/fp/> [99%]

#### 429) Dewey Decimal Classification Web Site

The Dewey Decimal Classification: Numbers You Can Count On catalog is now available. Use the online form to have.  
<http://www.oclc.org/oclc/fp/fptxthm.htm> [99%]

#### 634) Library of Congress Classification System Introduction

Introduction to the LC Classification System Some say Information is Power. Others say Information is the door to Knowledge. Libraries hold the key to.  
<http://snoopy.tbic.lib.fl.us/laudem/Introduction.h> [99%]

## Query 3.1. classification and security

### Examples from AltaVista search

Restricts results but also misses a lot.

#### 1. EXSYS: Specific Applications: Security Classification

Nuclear Weapons Security Classification. US Dept. of Energy. Nuclear...  
<http://www.exsysinfo.com/Appnotes/nuclear.html> - size 7K - 22-May-97 - English

#### 2. SLATE Application Note --Security Classification and Automatic Page Marking wi

Introduction. If your document contains classified information, you can identify the classification by.  
[http://www.slate.tdtech.com/app\\_notes/secclass-html.html](http://www.slate.tdtech.com/app_notes/secclass-html.html) - size 6K - 22-Feb-96 - English

#### 3. Computer Security Classification

The Classification. alert Advisories on various security vulnerabilities. dict Dictionaries and word lists. doc Security related documents.  
 access\_control.  
<http://www.cs.purdue.edu/coast/archive/Classification.html> - size 8K - 17-Mar-95 - English

#### 4. 355 Security Classification Control (R)

Top] -- MARC Field Guides Table of Contents -- 300 - Physical Description Fields. 355 Security Classification Control (R)Contains specifics pertaining to..  
<http://infoshare1.princeton.edu/katmandu/marc/355.html> - size 3K - 20-Jan-97 - English

#### 5. Security and Classification

By John Pike (johnpike@clark.net) The classification system is designed primarily to protect the confidentiality of certain...  
<http://www.tscm.com/classification.html> - size 17K - 28-Dec-96 - English  
<http://www.awpi.com/IntelWeb/US/misc/classification.html> - size 16K - 15-May-96 -

#### 6. National Security Classification Cost Estimates

A report to Congress from the Information Security Oversight Office  
<http://vwww.clark.net/fas/sgp/isoo/costs97.html> - size 9K - 10-May-97 - English

## What is a thesaurus / KOS? A first look

A **dictionary** is a listing of words and phrases giving information such as spelling, morphology and part of speech, senses, definitions, usage, origin, and equivalents in other languages (bi- or multilingual dictionary).

A **thesaurus** is a structure that manages the complexities of terminology and provides conceptual relationships, ideally through an embedded classification/ontology.

A thesaurus may specify descriptors authorized for indexing and searching. These descriptors form a **controlled vocabulary (authority list, index language)**.

A **monolingual thesaurus** has terms from one language, a **multilingual thesaurus** from two or more languages.

A **classification** is a structure that organizes concepts into a hierarchy, possibly in a scheme of facets.

The term **ontology** is often used for a shallow classification of basic categories or a classification used in linguistics, data element definition, or knowledge management or (increasingly) for any classification.

In AI-related contexts, an ontology is a classification with a rich set of semantic relationships that support reasoning.

AOD navigation page here

**EF****route of administration****EF2**— **by scope of drug action**

- EF2.2 . topical and local administration
- EF2.2.2 . . topical administration
- EF2.2.4 . . local drug administration
- EF2.4 . systemic administration

**EF4**— **by method or body site**

- EF4.2 . enteral administration
- EF4.2.2 . . oral enteral administration
- EF4.2.4 . . rectal enteral administration
- EF4.4 . mucosal administration
- EF4.4.2 . . transdermal administration
- EF4.4.4 . . inhalation, smoking, sniffing
- EF4.4.4.2 . . . smoking
- EF4.4.4.2.2 . . . . smoking w/out inhalation
- EF4.4.4.2.4 . . . . smoking with inhalation
- EF4.4.4.4 . . . nasal administration
- EF4.4.4.6 . . . pulmonary administration
- EF4.4.6 . . oral mucosal administration
- EF4.4.6.2 . . . buccal administration
- EF4.4.6.4 . . . sublingual administration
- EF4.4.8 . . rectal mucosal administration
- EF4.6 . parenteral administration
- EF4.6.2 . . intravenous injection
- EF4.6.2.2 . . . intravenous infusion
- EF4.6.4 . . intra-arterial injection
- EF4.6.6 . . intraperitoneal administration
- EF4.6.8 . . intracutaneous injection
- EF4.6.10 . . administration through skin implant
- EF4.6.12 . . subcutaneous injection
- EF4.6.14 . . intramuscular injection
- EF4.6.16 . . CNS injection
- EF4.6.16.2 . . . intrathecal injection
- EF4.8 . skin administration
  - (The full entry shows Narrower Term cross-references to the more specific methods involving the skin: EF4.4.2, EF4.6.8, EF4.6.10, and EF4.6.12)
- EF4.10 . oral administration
  - (NT to EF4.2.2, EF4.4.4.2, and EF4.4.6)
- EF4.10 . rectal administration
  - (NT to EF4.2.4 and EF4.4.8)

**EF6****drug administration by self vs. others**

- EF6.2 . self administration of drugs
- EF6.4 . drug administration by others

**Excerpt from a thesaurus hierarchy**

## EF route of administration

SN The way in which a substance reaches its site of action in the body. The substance may be administered for therapeutic or psychoactive effects - possibly as part of a human or animal experiment - by a third party or by the subjects themselves, or the subject may be exposed to the substance through the environment or in utero.

The major distinction between routes of administration is not the site where a substance is introduced or applied to the body, or even the way it is introduced or applied, but whether it takes effect merely in the local area where it is applied or whether it reaches its destination through systemic circulation. A further criterion is whether the drug reaches systemic circulation directly or whether it first passes through the liver, where it may be metabolized or excreted (first-pass effect in enteral administration). (Note: Drugs administered into the systemic circulation by any route, excluding intra-arterial injection, are subject to possible first-pass elimination in the lung prior to distribution to the rest of the body.)

Whether administration of a drug results in local or systemic action depends not only on the site and method of administration but also on the properties of the drug; sometimes the drug has both local and systemic action. This is particularly true for application to a mucous membrane, which may be intended for a local action but also may have - sometimes unwanted - systemic action. Furthermore, a drug may be absorbed at several sites (e.g., the mouth and the lung, the rectum and the intestine) in various proportions. To account at least partially for the very complex phenomena of the absorption of drugs into the body, the following classification uses two dimensions, or facets: By scope of drug action and by method or body site of administration. To index route of administration completely, use at least one descriptor from each facet.

- ST *medication route*
- ST *method of delivery of drugs or food*
- ST *mode of substance administration*
- ST *route of drug application*
- ST *route of drug entry*
- ST *route of exposure*
- BT +EE12 pharmacokinetics
- RT +AA2 AOD use
- RT +BS AOD substance by route of administration
- RT EE12.2e drug absorption
- RT +EE14.4.8 drug effect by location
- RT +HR drug therapy
- RT MD2.2.2.2 drug paraphernalia

## EF2 route of administration by scope of drug action

SN Use one of these descriptors in combination with a descriptor from **+EF4 route of administration by method or body site**.

### EF2.2 . topical and local administration

- SN The application of a substance to a localized area, chiefly for local effects at this site.
- NT HU4.2 local anesthesia
- RT GH10.2 chemical injury

#### EF2.2.2 . topical administration

- SN The application of a substance on the surface of the skin or on a mucous membrane (incl. the gastrointestinal membrane) so that the substance will take effect on the surface or on a localized layer under the surface. For example, for the administration of a decongestant spray, use **EF2.2.2 topical administration** combined with **EF4.4.4.4 nasal administration**.

ST *topical application*

#### EF2.2.4 . local drug administration

- SN The introduction of a substance into a localized area of the skin or other tissue, as through injection.
- NT EF4.6.4 intra-arterial injection
- NT EF4.6.8 intracutaneous injection
- NT +EF4.6.16 CNS injection

### EF2.4 . systemic administration

- SN The introduction of a substance into systemic circulation so that it is carried to the site of effect.
- NT +EF4.6.2e intravenous injection
- NT EF4.6.10 administration through skin implant
- NT HU4.4 general anesthesia
- RT +GH10.4 chemical poisoning

## Examples of full thesaurus entries



### Multilingual KOS problems

#### **simian**

monkey  
ape

#### **Affe**

*niederer Affe*  
Menschenaffe

#### **timepiece**

clock  
*wall clock*  
*standing clock*  
*tower clock*  
watch  
pocket watch  
wrist watch  
alarm clock

#### **Uhr**

*Wanduhr, Standuhr, Turmuhr*  
Wanduhr  
Standuhr  
Turmuhr  
*Taschenuhr, Armbanduhr*  
Taschenuhr  
Armbanduhr  
Wecker

#### **blanket, rug, carpet**

blanket  
*rug, carpet*  
rug (or carpet)  
*long, narrow rug*  
(wall-to-wall) carpet  
*hanging rug*

#### **Teppich**

Betteppich  
Bodenteppich  
*loser Bodenteppich*  
Läufer  
Teppichfußboden  
Wandteppich

*Italics* denotes terms created to express a concept not lexicalized in English or German, respectively.

Note that most English-German dictionaries would have you believe that the German equivalent for "monkey" is "Affe", but that equivalence holds only in some contexts.

Another difficulty arises when two terms mean almost the same thing but differ slightly in meaning or connotation, such as *alcoholism* in English and *alcoholisme* in French, or *vegetable* in English (which includes potatoes) and *Gemüse* in German, which does not. If the difference is big enough, one needs to introduce two separate concepts under a broader term; otherwise a scope note needs to clearly instruct indexers in all languages how the term is to be used so that the indexing stays, as far as possible, free from cultural bias or reflects multiple biases by assigning several descriptors.

## Examples of thesauri and other KOS

### **Alcohol and Other Drug Thesaurus (AOD Thesaurus)**

(US Nat. Inst. of Alcohol Abuse and Alcoholism)

<http://etoh.niaaa.nih.gov/AODVol1/Aodthome.htm>

### **Medical Subject Headings (MeSH) and Unified Medical Language System (UMLS)**

(US National Library of Medicine)

[www.nlm.nih.gov/mesh/meshhome.html](http://www.nlm.nih.gov/mesh/meshhome.html), [www.nlm.nih.gov/mesh/MBrowser.html](http://www.nlm.nih.gov/mesh/MBrowser.html),

[www.nlm.nih.gov/research/umls/umlsmain.html](http://www.nlm.nih.gov/research/umls/umlsmain.html), <http://umlsinfo.nlm.nih.gov>

### **Art and Architecture Thesaurus (AAT)**

(Getty Foundation)

<http://www.getty.edu/research/tools/vocabulary/aat/index.html>

### **Dewey Decimal Classification**

(US Library of Congress and OCLC/Forest Press)

[http://www.oclc.org/dewey/about/ddc\\_21\\_summaries.htm](http://www.oclc.org/dewey/about/ddc_21_summaries.htm)

### **WordNet (Princeton University, George Miller)**

[www.cogsci.princeton.edu/~wn/](http://www.cogsci.princeton.edu/~wn/),

[www.notredame.ac.jp/cgi-bin/wn](http://www.notredame.ac.jp/cgi-bin/wn) (Not reachable on July 6, 2002)

### **CYC Ontology (CYC Corporation)**

<http://www.cyc.com/cyc-2-1/cover.html>, <http://www.cyc.com/cyc-2-1/toc.html>

Example pages form part 2 of the tutorial materials. They will be examined briefly but are intended primarily for further study.

## More thesaurus / KOS examples

### A few sample pages included

**Yahoo**     **The Yahoo classification.** Web pages [www.yahoo.com](http://www.yahoo.com)

**Bloom**     Anderson, L. W., & Krathwohl, D. R.  
**A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.**  
 New York: Addison Wesley Longman; 2001.  
<http://www.unesco.org/webworld/ramp/html/r8810e/r8810e0e.htm>  
<http://websites.ntl.com/~james.atherton/learning/bloomtax.htm>,  
<http://sweep.riv.csu.edu.au/td/bloom.html>,  
<http://faculty.washington.edu/~krumme/guides/bloom.html>  
<http://rite.ed.qut.edu.au/oz-teachernet/training/bloom.html>  
<http://www.uwsp.edu/education/lwilson/curric/newtaxonomy.htm>

#### Original version, 1956

**Taxonomy of educational objectives. The classification of educational goals /** by a committee of college and university examiners ; Benjamin S. Bloom, editor.  
 Prelim. ed.  
 New York : Longmans, Green, 1954-

**SOC**     **Standard Occupational Classification 2000**  
 Bureau of Labor Statistics (BLS) + other agencies  
[http://stats.bls.gov/soc/soc\\_home.htm](http://stats.bls.gov/soc/soc_home.htm)  
 The SOC is augmented by the **Occupational Information Network (O\*NET)**, a database with additional occupational titles, definitions, and features of occupations.  
<http://www.doleta.gov/programs/onet>

**CSDGM**     **Content Standard for Digital Geospatial Metadata 1998**  
 Federal Geographic Data Committee (FGDC)  
<http://www.fgdc.gov/metadata/constan.html>

**ERIC**     **Education Resources Information Center Thesaurus.** 13th ed. Bibliographic retr.  
<http://searcheric.org/>

## Additional examples illustrating different functions

**HS Harmonized Commodity Description and Coding System.** World Customs Organization, Brussels. Info: <http://pacific.commerce.ubc.ca/trade/HS.html>

### **NAICS North American Industrial Classification System**

"common industry definitions for Canada, Mexico, and the US. Developed in cooperation with the US Economic Classification Policy Committee, Statistics Canada, and Mexico's Instituto Nacional de Estadística, Geografía e Informática to better compare economic and financial statistics and ensure that such statistics keep pace with the changing economy. NAICS will replace the countries' separate classification systems (in the US: Standard Industrial Classification, SIC) with one uniform system for classifying industries."

Info: [www.census.gov/epcd/www/naics.html](http://www.census.gov/epcd/www/naics.html), [www.naics.com](http://www.naics.com)

**ICD-10 The International Statistical Classification of Diseases and Related Health Problems, tenth revision.** Produced by the World Health Organization. Published in many languages. Info: [www.who.int/whosis/icd10/index.html](http://www.who.int/whosis/icd10/index.html), [www.cdc.gov/nchs/about/major/dvs/icd10des.htm](http://www.cdc.gov/nchs/about/major/dvs/icd10des.htm)

**CPT Physicians' Current Procedural Terminology. CPT 2003.** American Medical Association. November 2002  
(Info: <http://www.ama-assn.org/ama/pub/category/3113.html>, listing of codes <https://webstore.ama-assn.org/index.jhtml>)  
Health Care Finance Administration (HCFA) Common Procedure Coding System (HCPCS) for Medicare reimbursement for hospital outpatient services. It has three levels - CPT (level 1), HCPCS or National (level 2), and Local (level 3).  
In its data collection the Agency for Health Care Policy and Research (AHCPR) uses data standards that are based on those employed by the Census Bureau, the American Hospital Association, the Health Resources and Services Administration (Area Resource File), the National Center for Health Statistics, and codes for clinical diagnosis and procedures such as ICD-10 and CPT 1998. These standards facilitate data analysis and use by ensuring comparability, quality and interoperability. Further, uniform health care data advance medical and health care services research, the efficiency of the private sector health care delivery system, and quality improvement measurement.

- Further type of classification: **biological taxonomies**. Used in biology, agriculture, food science, and medicine. Several rivaling schemes for major areas (kingdoms) and many publications on specific areas.  
<http://www.itis.usda.gov/>  
<http://www.ucmp.berkeley.edu/help/taxaform.html>
- **Metadata schemas (such as CSDGM), data element dictionaries, object hierarchies in object-oriented programming**

## **KOS development principles**

- (1) Determine the right organizational scope or user group(s) to be served**
- (2) Determine the right function scope**
- (3) Ensure adoption by supporting many views through an inclusive, flexible KOS database**
- (4) Build on existing KOS. Reuse, reuse, reuse**
- (5) Use automation for efficiency**
- (6) Capitalize on collaborative creation and editing**

# KOS development principles

Good cost-effective KOS development stands and falls with the observation of six simple principles.

**(1) Determine the right organizational scope or user group(s)** to be served. A broad organizational scope results in higher return on investment but must be balanced against tailoring the KOS to specific requirements (but see principle (3)). The scope can range from a KOS for a single user who needs to organize a collection of documents, Web sites, files, email messages, notes, contacts, appointments, and tasks to a KOS that is used worldwide for organizing many databases in many languages, such as AGROVOC or the Medical Subject Headings. In between would be a KOS for an organization, an online community, or a specific database such as AERS, FDA's Adverse Events Reporting System database ([fda.gov/cder/aers/default.htm](http://fda.gov/cder/aers/default.htm)). A broad scope of application (many departments in an organization, several online communities with related purposes, databases in the same general subject domain) improves the return on investment into the KOS and fosters semantic interoperability. .

**(2) Determine the right function scope.** The more functions are served by a KOS the higher the return on investment. Therefore it is important to identify all possible applications that could profit from KOS support. This includes functions such as retrieval from any kind of database or through a Web search engine and structured output of results, project planning, insurance billing, natural language processing, reasoning in expert systems; see XXX for an extensive list. Understanding all these functions is crucial if one wants to maximize ROI. For each application state

- requirements and the spectrum of users' level of knowledge.
- required characteristics of the KOS

**(3) Ensure adoption by supporting many views through an inclusive, flexible KOS database.** Different applications need different views for intrinsic reasons that are rooted in the nature of the application (for example, to optimize reasoning over large knowledge bases) or for historical reasons (a given user group may not want to change what they are used to, or a large body of material can not feasibly be re-indexed). The solution is a flexible comprehensive KOS database in which different views can coexist or from which different views can be easily extracted; that way each application's or group's view of the domain can be accommodated, and that removes a key barrier to engaging several groups that might otherwise go their own ways (at great expense), even though all groups may agree on 80% and disagree on just 20% of concepts, terms, and relationships. There is no need to force everybody into the same mold; many views can coexist within the same KOS database, where they can be related to each other to the extent possible, capturing commonalities and explicating differences. Through the KOS database, different groups can learn from each other; this may lead to improvements of each group's view. In that sense, a KOS database

- records present usage by different groups (like a dictionary records usage);
- creates an overarching well-ordered structure for all the concepts from different views;
- through that structure, creates relationships between different views;
- through that structure, recommends improvements for each group's view..

**(4) Build on existing KOS. Reuse, reuse, reuse.** There is enormous intellectual capital in existing KOS; use it! Identify the KOS presently used, explicitly or implicitly, in the organizational and functional scope. Find other KOS that could meet some or all of the requirements, perhaps with some modification, or that would be useful sources for developing a new KOS (see Box 1. Searching for KOS).

**(5) Use automation for efficiency.** Automation can be used to massage data from existing KOS for the purpose at hand. It can also be used to extract terms, concepts, relationships, and global structure from texts and document collections. While automation cannot entirely replace intellectual effort, a KOS constructed using automatic processing alone is better than no KOS at all.

**(6) Capitalize on collaborative creation and editing** with central expert support and some measure of control. Possibly have a cadre of KOS editors throughout the organization or throughout the world.

**Functions of a KOS  
(thesaurus / classification / ontological knowledge base)**

in the context of digital libraries

Support learning and assimilating information.

Assist researchers and practitioners with problem clarification.

Support information retrieval.

Provide knowledge-based support for end-user searching.

Support meaningful information display.

Provide a tool for indexing.

Facilitate the combination of multiple databases or unified access to multiple databases.

Support document processing after retrieval.

**Support learning and assimilating information**

**Support learning** about any topic **by providing** the learner with a **coherent, age-appropriate conceptual framework**.

**Learning as information retrieval.** Conceptual framework for asking the right questions.

**Assist readers** in understanding text.

**Assist researchers and practitioners with problem clarification —**

provide the conceptual basis for the design of good research and implementation and for good query formulation. Includes help with

**exploring the conceptual context of a research or practical problem** — a study, policy, plan, or implementation project

and with

**structuring the problem.**

Examples of specific functions:

**Present the issues in a field or application area in a coherent framework.**

**Assist in problem-solving:** Assist in the exploration of the dimensions of a problem and aspects to be considered in its solution; provide a classification of approaches to solving a specific problem.

Provide classification and **consistent definition of variables for research / of evaluation criteria for practical problems**, thus enhancing the comparability of research and evaluation results and making research more cumulative.



## Support information retrieval

Provide **knowledge-based support for end-user searching**. Support

searching in multiple natural languages;

free-text searching;

searching multiple databases using different index languages.

**Elicitation of user needs** through a series of menus based on search tree, or through **guidance in the conceptual analysis of a search topic** (questions based on a facet structure, presentation of a segment of the concept hierarchy for each applicable facet).

**Browsing the classification structure** to identify useful concepts for a search at the level of specificity desired. Browsing a collection, as in a subject directory.

**Mapping from the user's query terms to descriptors** used in a database **or to the multiple natural language expressions** to be used for free-text searching.

**Inclusive** (hierarchically expanded) **searching**.

**Enhanced ranking algorithms** based on concept and term relationships.

**Searching multiple databases** by mapping the users query terms to the descriptors used in each of the databases, or mapping the descriptors from one database to another databases (switching); common search language.

## Support information retrieval, continued

**Support information display**, especially presentation of search results:

**Meaningful arrangement of units** (document records, paragraphs, property data on a given substance assembled from several databases), including knowledge-based clustering of records retrieved.

This supports **exploration of large retrieved sets** and, by extension, **exploration of the content of an entire collection** or subcollection.

**Meaningful arrangement of information within a record** (for example meaningful ordering of descriptors assigned).

### **Support information retrieval, continued**

#### **Provide a tool for indexing.**

##### **Vocabulary control.**

##### **User-centered** (request-oriented, problem-oriented) **indexing.**

Indexing **several databases** in a field with a **common index language** and sharing the results of indexing to reduce overall indexing effort.

##### **Mapping indexing descriptors from one system to another.**

### **Support information retrieval, continued**

#### **Facilitate the combination of multiple databases or unified access to multiple databases through**

**mapping the users query terms** to the descriptors used in each of the databases;

**mapping the query descriptors from one database to another** (switching);

providing a **common search language** from which to map to multiple databases;

providing a **common index language** for a number of databases in a field;

**mapping indexing descriptors from one database to another.**

## Support document processing after retrieval

For example

**Highlight descriptors responsible for retrieval**, using different colors for different facets.

**Highlight** terms belonging to a given category, for example, **personal names**, again using different colors for different categories.

**Prepare document summaries**, possibly in a different language, taking into account the query topic.

**Translate full documents.**

**Extract facts from text.** Compile and arrange facts extracted from several texts.

## The underlying function of a knowledge base on concepts and terminology

**Map out a concept space, relate concepts to terms, and provide definitions**, thus providing orientation and serving as a reference tool.

Provide a **semantic road map and common language** for an individual field and, perhaps more importantly, map the relationships among fields.

**Clarify concepts by putting them in the context of a classification / typology** and to provide a system of definitions.

**Relate concepts and terms across disciplines, languages, and cultures.**

# **KOS (thesaurus / ontology functions**

## **Reference list**

**Functions of a KOS**  
**thesaurus / classification / ontological knowledge base**  
**Overview**

Provide a **semantic road map** to individual fields and the relationships among fields. Map out a concept space, relate concepts to terms, and provide definitions, thus providing orientation and serving as a reference tool.

**Improve communication generally. Support learning and assimilating information.**

Support learning through conceptual frameworks. Conceptual framework to help the learner ask the right questions.

Support the development of instructional materials through conceptual frameworks.

Assist readers in understanding text by giving the meaning of terms.

Assist writers in producing understandable text by suggesting good terms.

Support foreign language learning.

**Provide the conceptual basis for the design of good research and implementation.**

Assist researchers and practitioners with problem clarification.

Consistent data collection, **compilation of statistics** (related to information analysis)

**Provide classification for action. Classification for social and political purposes**

a classification of diseases for diagnosis,  
of medical procedures for insurance billing,  
of commodities for customs.

**Support information retrieval and analysis. Organizing and keeping track of goods and services for commerce (esp. ecommerce) and inventory**

Provide a tool for searching, particularly knowledge-based support for end-user searching, including hierarchically expanded searching.

Provide a tool for indexing.

Facilitate the combination of or unified access to multiple databases

Support document processing after retrieval.

**Support meaningful, well-structured display of information.**

**Ontology for data element definition.** Data element dictionary.

**Conceptual basis for knowledge-based systems.**

**Do all this across multiple languages**

**Mono-, bi-, or multilingual dictionary for human use.**

**The underlying function of a knowledge base on concepts and terminology:**

**Provide a semantic road map to individual fields and the relationships among and across fields.**

**Map out a concept space, relate concepts to terms, and provide definitions,** thus providing orientation and serving as a reference tool.

Provide a **semantic road map and common language** for an individual field and, perhaps more importantly, map the relationships among fields.

**Clarify concepts by putting them in the context of a classification / typology** and to provide a system of definitions.

**Relate concepts and terms across disciplines, languages, and cultures.**

Many specific functions build on this foundation.

**Improve communication generally.  
Support learning and assimilating information**

**Support learning** about any topic by **providing** the learner/reader with a **coherent, age-appropriate conceptual framework**. Conceptual frameworks help the learner ask the right questions; **learning as information retrieval**.

**Assist readers** in understanding text; help them ascertain the proper meaning of a term and placing it in context.

**Support the development of instructional materials** by providing a conceptual framework to the instructional developer / writer and by suggesting didactically useful arrangements of topics.

**Assist writers** in producing understandable text by helping them to conceptualize the topic and suggesting from a semantic field the term that best conveys the intended meaning and connotation.

**Support foreign language learning**

## **Provide the conceptual basis for the design of good research and implementation.**

### **Assist researchers and practitioners with problem clarification**

Includes help with

**exploring the conceptual context of a research or practical problem** — a study, policy, plan, or implementation project

and with

**structuring the problem** and providing a conceptual framework for asking the right questions and devising good query formulations for retrieval.

Examples of specific functions:

#### **Present the issues in a field or application area in a coherent framework.**

**Assist in problem-solving:** Assist in the exploration of the dimensions of a problem and aspects to be considered in its solution; provide a classification of approaches to solving a specific problem (for example, a classification of approaches to drug abuse prevention as a help in designing drug abuse prevention projects).

Provide classification and **consistent definition of variables for research / of evaluation criteria for practical problems**, thus enhancing the comparability of research and evaluation results and making research more cumulative.

## **Support the compilation and use of statistics**

This is a very important function. The Census Bureau, the Bureau of Labor Statistics, and other statistical agencies are heavily involved in developing classifications and defining concepts.

### **Support data collection**

The concepts in a classification used for statistics not only make the collected data retrievable, they define the very nature of the data.

### **Support data aggregation**

For example, get the value of all *electronic goods* imported into the US in the year 2000, or the tonnage of *green leafy vegetables* produced in a given year in the US.

### **Support retrieval of specific numbers (also part of information retrieval)**

Support data tabulation and analysis (Need to have proper variables available)

### **Provide classification for action**

This list addresses the functions of formal classifications. In a broader perspective, classification is the basis for much of everyday action, where we put people, things, and events in certain categories and, based on these categories, predict the behavior of persons and things and the course and effects of events, determine our attitudes towards them, and plan action accordingly.

For example,

- a classification of diseases for diagnosis,
- a classification of medical procedures for insurance billing,
- a classification of medical outcomes to assist with treatment evaluation,
- a classification of commodities for customs,
- a classification of educational objectives for instructional development,
- a classification of occupations for matching job applicants with job openings and for pay scale;
- a classification of skills for employee task assignments.
- a classification of crimes for determining sentences
- a classification of types of expenses for tax purposes

### **Classification for social and political purposes.** Socially charged classification

For example

Establishing that a profession has its own knowledge base, thereby enhancing the recognition of the profession (for example, the Nursing Intervention Classification)

Establishing a persons condition or behavior as normal, or as a disease, or as a moral failing or otherwise deviant. Different groups may want the same condition or behavior classified in different ways to further their agenda

Examples:

Should homosexuality be classified as a disease?

Is alcoholism or other drug abuse a disease or a moral failing?

Is mental illness a disease on a par with physical illness, and thus covered by health insurance the same way?

Is some levy to be classified as a *tax* or as a *user fee*



**Support information retrieval 1:****A tool for searching, particularly knowledge-based support for end-user searching.** Support

searching in any kind of database — bibliographic, full-text and hypermedia, directory, numeric, etc.;

searching in any kind of medium — printed indexes, CD-ROM systems, online systems, and the Internet;

searching in multiple natural languages independent of the language used in each database;

free-text searching;

searching multiple databases using different index languages.

**Elicitation of user needs** through a series of menus based on a search tree, or through **guidance in the conceptual analysis of a search topic** (questions based on a facet structure, presentation of a segment of the concept hierarchy for each applicable facet).

**Browsing the classification structure** to identify useful concepts for a search at the level of specificity desired. (The user may not have command of the vocabulary needed.)  
Browsing a collection (as on the shelves or in a subject directory)

**Mapping from the user's query terms to descriptors** used in a database **or to the multiple natural language expressions** to be used for free-text searching.

**Inclusive** (hierarchically expanded) **searching**.

**Enhanced ranking algorithms** that use concept and term relationships.

**Searching multiple databases** by mapping the users query terms to the descriptors used in each of the databases, or mapping the descriptors from one database to another databases (switching); common search language.

**Support information retrieval 2: Provide a tool for indexing.**

**Vocabulary control.**

**User-centered** (request-oriented, problem-oriented) **indexing**.

Indexing **several databases** in a field with a **common index language** and sharing the results of indexing to reduce overall indexing effort.

**Mapping indexing descriptors from one system to another.**

### **Support information retrieval 3:**

**Facilitate the combination of multiple databases or unified access to multiple databases** through

**mapping the users query terms** to the descriptors used in each of the databases;

**mapping the query descriptors from one database to another** (switching);

providing a **common search language** from which to map to multiple databases;

providing a **common index language** for a number of databases in a field;

**mapping indexing descriptors from one database to another.**

### **Support information retrieval 4: Document processing after retrieval**

Sample functions that require knowledge-based support:

**Meaningful arrangement of search results** (see next box)

**Highlight descriptors responsible for retrieval**, using colors to show facets.

**Highlight** terms belonging to a given category, for example, **personal names**, again using different colors for different categories.

**Prepare document summaries**, possibly in a different language, taking into account the query topic.

**Translate full documents.**

**Extract substantive data from text.** Compile and arrange data extracted from several texts.

### **Support meaningful, well-structured display of information**

**Meaningful arrangement of units** (document records, paragraphs, property data on a given substance assembled from several databases), including knowledge-based clustering of records retrieved. This includes meaningful structure for **Web sites** and **subject directories**

This supports **exploration of large retrieved sets** and, by extension, **exploration of the content of an entire collection** or subcollection.

**Meaningful arrangement of information within a unit** (for example meaningful ordering of descriptors within a bibliographic record).

### **Organizing and keeping track of goods and services for commerce (esp. ecommerce) and inventory**

The functions detailed for information retrieval apply to this special case

Organize a store, an inventory, an online merchandise catalog, a yellow page directory so items can be found

Display the inventory in a meaningful arrangement so users can find things (as in a store)

Keep track of inventory

These functions apply both to business-to-consumer and to business-to-business commerce. Classification by function or purpose is especially important here.

### **Ontology for data element definition.**

Data element dictionary.

Consider data processing systems in a multinational corporation

### **Conceptual basis for knowledge-based systems.**

### **Do all this across multiple languages**

**Mono-, bi-, or multilingual dictionary for human use.**

Printed or machine-readable, such as dictionary on CD-ROM or a thesaurus used in conjunction with a word processor

**Dictionary/knowledge base for automated language processing**

Machine translation and natural language understanding (data extraction, automatic abstracting/indexing). (It should be noted that parsing natural language requires not only morphological information and information about the possible syntactic roles of a term but also a great deal of semantic information.)

Spell check dictionary

Knowledge base for grammar checking.

**Functions of an ontological knowledge base in software development**

Assist in the design and implementation of the **user interface, esp. choice of terms and icons.**

Terms and icons must be chosen with the sometimes conflicting goals of communicating to the intended user group and of adhering to standards.

Assist in the organization and formulation of **help messages and of documentation** and third-party software books.

Serve as the **lexicon for machine translation** of interfaces and software-related documents

**Assist the user in understanding interfaces and documentation, esp. in a foreign language.**

**Support retrieval of software** for the end user or for **software reuse.**

**Data element definition and standardization and organization of CASE tool databases.**

All this functionality must be provided in **multiple languages** (for example, **software localization** for end users, **CASE tool databases for multinational development teams**)

## **User-centered indexing / request-oriented indexing**

Construct a **classification/ontology**  
(embedded in a thesaurus)

**based on** actual and anticipated **user queries and interests.**

Thus provide a **conceptual framework**  
that organizes user interests and  
communicates them to indexers.

### **Index materials from users' perspective:**

Add need-based retrieval clues  
beyond those available in the document.  
Increase probability that a retrieval clue  
corresponding to a query topic is available.

### **Index language as checklist.**

Indexing = judging relevance against user concepts.  
Relevance rather than aboutness

### **Implementation:**

Knowledgeable indexers  
Expert system using syntactic & semantic analysis  
& inference.

## **User-centered indexing / request-oriented indexing.**

### **Sample concepts included in the index language due to user interest**

Systemic administration

Intergenerational social mobility

Biochemical basis of behavior

Longitudinal study

**User-centered / request-oriented indexing.  
Sample documents with descriptors**

**Document**

The drug was injected into the aorta

User concept: *Systemic administration*

**Document:**

The percentage of children of blue-collar workers going to college

User concept: *Intergenerational social mobility*

**Document:**

CSF studies on alcoholism and related behaviors

User concept: *Biochemical basis of behavior*

User concept: *longitudinal study*

(Longitudinal not mentioned in the document; determined through careful examination of the methods section.)

## Design of a classification scheme for fiction based on analysis of actual user-librarian communication

Annelise Mark Pejtersen

<i>Dimensions</i>	<i>Freq.</i>	<i>Sub-classes</i>	<i>Freq.</i>	<i>Examples</i>
1. subject matter	38	a. action and course of events (plot)	10	a. mystery novel, book with action
		b. psychological development / description	5	b. love story, book with psychology
		c. social relations	19	c. family chronicle, not with politics
2. frame	25	a. time	13	a. historical novel, books from 16th and 17th centuries
		b. place (geographical, social environment, professions)	12	b. travelogue, books from the countryside, books about working people
3. author's intention /attitude	37	a. emotional experience	34	a. humoristic, suspense, amusing
		b. cognition / information	3	b. philosophical, political, not too many problems
4. accessibility	34	a. readability	16	a. easy, not complicated, not heavy
		b. physical characteristics	18	b. typography, modern/old, series, size, volume
Other formulations	87	a. author's name / title	25	
		b. author's name / title as example	22	b. something like Emily Bronte
		c. good book	32	
		d. diverse	8	

From: **Theory and application of information research**. London: Mansell; 1980. ISBN 0-7201-1513-2. p. 149



## **Sample user concepts for indexing images**

Good scientific illustration

Useful for fundraising brochure

Appealing to children

Cover page quality

## User-centered /request-oriented / problem-oriented indexing

As summarized in the overheads, user-centered indexing involves analyzing actual and anticipated user queries and interests and constructing a framework, a hierarchically structured controlled vocabulary, that includes the concepts of interest to the users and thus communicates these interests to the indexers or an expert system that can infer user-relevant concepts from text. The indexers then become the "eyes and ears" of the users and index materials from the users' perspective. The indexer uses the structured list of user-relevant concepts as a checklist, applying her understanding of a document (or other object) to judge its relevance to any of these concepts. This process ensures that users will find the documents that they themselves would judge relevant upon examination.

Request-oriented indexing contrasts with document-oriented indexing, where the indexer simply expresses what the document is about or where simply the terms in the text are used. But, as the examples show, a document can be **relevant** for a concept without being **about** the concept: a document titled *The percentage of children of blue-collar workers going to college* is not necessarily about *intergenerational social mobility*, but a researcher interested in that topic would surely like to find it, so it is relevant.

Request-oriented indexing is essential for good performance in fiction retrieval and even more so in image retrieval. Image retrieval profit from descriptors that capture imponderables, such as the mood of an image or from descriptors indicating possible uses of an image (such as

This perspective on indexing has implications for cross-language retrieval: The conceptual framework must be communicated in every participating language to allow a meeting of minds to take place, regardless of the languages of the user and the indexer. This is particularly salient in the context of indexing : One needs to make sure that, as far as possible, the term used by the indexer in one language communicates the same mood as the term given to the user in another language for searching.

# **Thesaurus / KOS structure**

**Concept-term relationships**

**Conceptual structure**

**Semantic analysis and facets**

**Hierarchy**

## Concept-term relationships (Terminological structure)

### Controlling synonyms

<b><i>Term</i></b>	<b><i>Preferred synonym</i></b>
<b>Teenager</b>	<b>Adolescent</b>
<b>Teen</b>	<b>Adolescent</b>
<b>Youth (young person)</b>	<b>Adolescent</b>
<b>Pubescent</b>	<b>Adolescent</b>
<b>Black</b>	<b>African American</b>
<b>Afro-American</b>	<b>African American</b>
<b>Alcoholism</b>	<b>Alcohol dependence</b>
<b>Inheritance</b>	<b>Heredity</b>
<b>Ultrasonic cardiography</b>	<b>Echocardiography</b>

## **Soergel, p. 215, enlarged**

## Disambiguating homonyms

**administration 1 (management)**

**administration 2 (drugs)**

**Läufer 1 (Sportler)**

***English:* runner (athlete)**

**Läufer 2 (Teppich)**

***English:* long, narrow rug**

**Läufer 3 (Schach)**

***English:* bishop (chess)**

**discharge 1 (From hospital or program)**

***German:* Entlassung**

**discharge 2 (From organization or employment)**

***Preferred synonym:* Dismissal**

***German:* Entlassung**

**discharge 3 (Medical symptom)**

***German:* Absonderung, Ausfluss**

**discharge 4 (into a river)**

***German:* Ausfluss**

**discharge 5 (Electrical)**

***German:* Entladung (which also means unloading)**

## **Importance of terminological structure**

**The terminological structure is equally important in controlled vocabulary systems and in free-text searching.**

**In free-text searching,**

**synonym expansion of query terms is important for recall**

**homonym indicators can trigger a question to the user on the intended meaning of the query term.**





## **Conceptual structure**

**A well-developed conceptual structure**

*sine qua non* for user-centered indexing  
very useful for free-text retrieval as well.

**The two principles of conceptual structure**

**facet analysis**

**hierarchy**

**Facets and hierarchy interact**

## **Facets. Semantic factoring or feature analysis**

**Analyzing a concept into its defining components (elemental concepts or features).**

**Concept frame with facet slots**

### **liver cirrhosis**

<i>Pathologic process:</i>	<b>inflammation</b>
<i>Body system:</i>	<b>liver</b>
<i>Cause:</i>	<b>not specified</b>
<i>Substance/organism:</i>	<b>not specified</b>

### **alcoholic liver cirrhosis**

<i>Pathologic process:</i>	<b>inflammation</b>
<i>Body system:</i>	<b>liver</b>
<i>Cause:</i>	<b>chemically induced</b>
<i>Substance/organism:</i>	<b>alcohol</b>

### **hepatitis A**

<i>Pathologic process:</i>	<b>inflammation</b>
<i>Body system:</i>	<b>liver</b>
<i>Cause:</i>	<b>infection</b>
<i>Substance/organism:</i>	<b>hepatitis A virus</b>

## Facet principles

### **A facet groups concepts that fill the same role:**

- concepts that fall under the same aspect or feature in the definition of more complex concepts;
- concepts that can be answers to a given question.

In frame terminology:

The facets listed above are slots in a disease frame; a facet groups all concepts that can serve as fillers in one slot.

### **Using elemental concepts as building blocks for constructing compound concepts**

- drastically reduces the number of concepts in the thesaurus / KOS and thus leads to conceptual economy.
- facilitates the search for general concepts, such as searching for the concept *dependence*, which occurs in the context of medicine, psychology, and social relations.

Facets can be defined at high or low levels in the hierarchy, as illustrated in the next overhead.

## Top-level facets

organism

body part

chemical substances by function

chemical substances by structure

## Low-level facets

### **route of administration**

route of administration by scope of drug action  
(local/topical or systemic)

route of administration by body part

route of admin. by method of application  
(injection, rubbing on, etc.)

### **liver**

liver tissue (hepatocyte, Kupffer cell, etc.)

liver part (hepatic lobule, portal lobule, etc.)

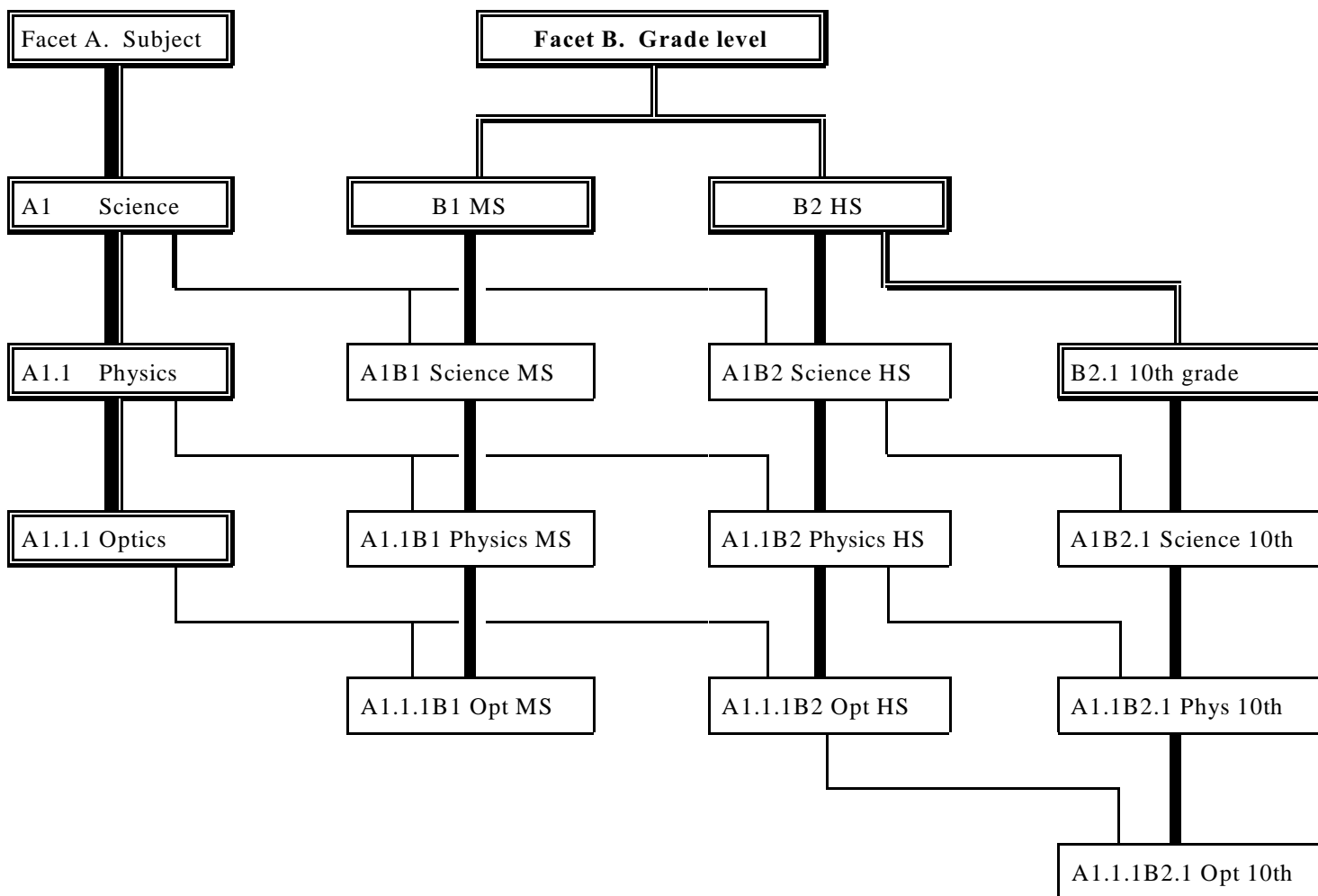
## Hierarchy

### groups at high risk of drug use

- . suicidal or physically or mentally disabled
- . persons from unstable or low-cohesion families
- . children of alcoholic or other drug-abusing parents
  - SN Adult or still under age
- . children of single teenage mothers
- . persons subjected to abuse or neglect (now or past)
  - . persons subjected to abuse/neglect by parents
    - . latchkey children
  - . persons subjected to abuse/neglect by spouse
- . single teenage mothers
- . school dropouts or those at risk of dropping out
- . unemployed or in danger of being unemployed
- . economically disadvantaged
- . homeless
  - . runaway youth
- . gateway drug users
- . persons engaged in violent or delinquent acts

**See also examples given previously in  
What is a thesaurus / KOS**

# Hierarchy from combining two facets



## Linear arrangement 1

### A Facet A. Subject

- . A1 Science
- . . A1B1 Science MS NT A1.1B1; BT B1
- . . A1B2 Science HS NT A1.1B2; BT B2
- . . . A1B2.1 Science 10th grade NT A1.1B2.1; BT B2.1
- . . A1.1 Physics
- . . . A1.1B1 Physics MS NT A1.1.1B1; BT A1B1
- . . . A1.1B2 Physics HS NT A1.1.1B2; BT A1B2
- . . . . A1.1B2.1 Phys 10th gr NT A1.1.1B2.1; BT A1B2.1
- . . . A1.1.1 Optics
- . . . . A1.1.1B1 Optics MS BT A1.1B1
- . . . . A1.1.1B2 Optics HS BT A1.1B2
- . . . . . A1.1.1B2.1 Optics 10th grade BT A1.1B2.1

### B Facet B. Grade level

- . B1 MS NT A1B1
- . B2 HS NT A1B2
- . . B2.1 10th grade NT A1B2.1

## Linear arrangement 2

### A Facet A. Subject

- . A1 Science NT B1A1, B2A1
- . . A1.1 Physics NT B1A1.1, B2A1.1
- . . . A1.1.1 Optics NT B1A1.1.1, B2A1.1.1

### B Facet B. Grade level

- . B1 MS
- . . B1 A1 MS Science BT A1
- . . . B1 A1.1 MS Physics BT A1.1
- . . . . B1 A1.1.1 MS Optics BT A1.1.1
- . B2 HS
- . . B2 A1 HS Science NT B2.1A1; BT A1
- . . . B2 A1.1 HS Physics NT B2.1A1.1; BT A1.1
- . . . . B2 A1.1.1 HS Optics NT B2.1A1.1.1; BT A1.1.1
- . . B2.1 10th grade
- . . . B2.1 A1 10th grade Science BT B2A1
- . . . . B2.1 A1.1 10th grade Physics BT B2A1.1
- . . . . . B2.1 A1.1.1 10th grade Optics BT B2A1.1.1



## **Uses of facet analysis and hierarchy**

Help to organize the concept space and establish concept relationships.

Discover concepts, esp. general concepts spanning several disciplines

Assist the user in analyzing and clarifying a search problem:  
elicit the facets involved  
present hierarchical structure within each facet

Facilitate the search for general concepts, such as inflammation or dependence (which occurs in the context of medicine, psychology, and social relations)

Hierarchical query term expansion

These functions are useful in both controlled vocabulary and free-text searching.



## Concept discovery

### through facet analysis and hierarchy building

Through facet analysis and hierarchy building, one often discovers concepts that are needed in searching or that enhance the logic of the concept hierarchy. Need to create terms for these concepts.

Consider

*train station, bus station, harbor, airport*

Common semantic component: **traffic station**

*gin, whiskey, cherry brandy, tequila, etc.*

common semantic component: **distinct distilled spirits**

(counterpart of the already lexicalized

neutral distilled spirits)

*transactional analysis, dream analysis, insight therapy, Gestalt therapy, reality therapy, cognitive therapy*

Umbrella concept for structuring the hierarchy and for retrieval: **analytic psychotherapy**

(methods that seek to assist patients in a personality reconstruction through insight into their inner selves)

*Payment in exchange for some consideration (see above)*

## Web-based thesaurus / KOS display and incorporation into search functions

**Vignette on thesaurus / KOS use in searching a digital library.** The director of a drug-free community coalition is faced with developing a prevention project and the funding for it. Signing on to the AOD Digital Library, she begins by **browsing the prevention section of the thesaurus / KOS hierarchy** to get a structured overview of various prevention approaches. From the **thesaurus / KOS scope notes**, some of these approaches seem particularly applicable to her community, so she follows the links to more in-depth explanations. She returns to the thesaurus / KOS and follows a link from *prevention through education* to a funding program announcement. She opens the guidelines for submitting proposals to this program and copies a proposal template into her private space (shown in another window) and fills in some text and copies some text (which is transferred with the proper source). From the program announcement, she follows a link to projects funded previously and further to project reports and evaluations. She comes across the **unfamiliar term *triangulation*** and clicks on it to see the **thesaurus / KOS entry, which gives an explanation and the hierarchical context**. In another document she highlights the phrase *prevention program evaluation* to initiate a search in the system and one external database. She copies three references with abstracts to her private space. (Later she will return to these, select one for detailed reading, and add more notes and quotes to her emerging proposal.) Returning to the program announcement, she follows a link to relevant research, selects some articles to read, and adds more material to her outline. One of the papers compares the effectiveness of several prevention curricula. She follows a link to the curriculum that came out on top and from there finds further reviews. She also finds some discussion of resources required. She needs some more data — namely, demographics of her community and funding sources for the required local match — so she initiates searches in two external databases, incorporating the results into her proposal. Now she completes the first draft, including the text itself and annotations that explain why a piece is included or why certain language is used. Before submitting the proposal, she emails two board members and a city staff member for comments, giving them access to her private space. The three people read the draft and add their annotations, including suggested wording. The director now revises the draft, requests the final document in PDF format, links to the agency's submission system, and sends off her proposal.

The next page shows a proposed digital library structure built around a thesaurus / KOS and the pages following show two steps in a search using the thesaurus / KOS.

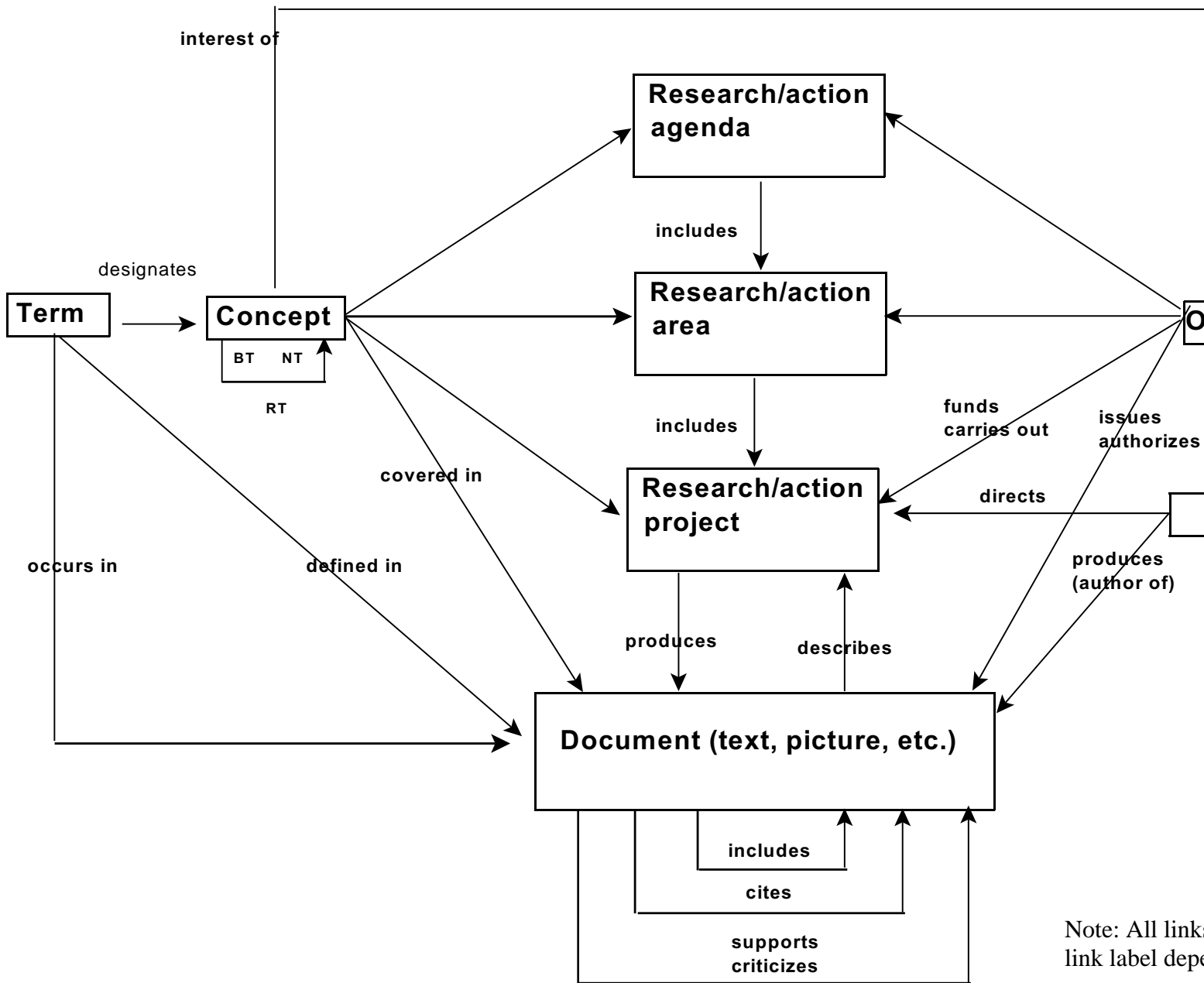
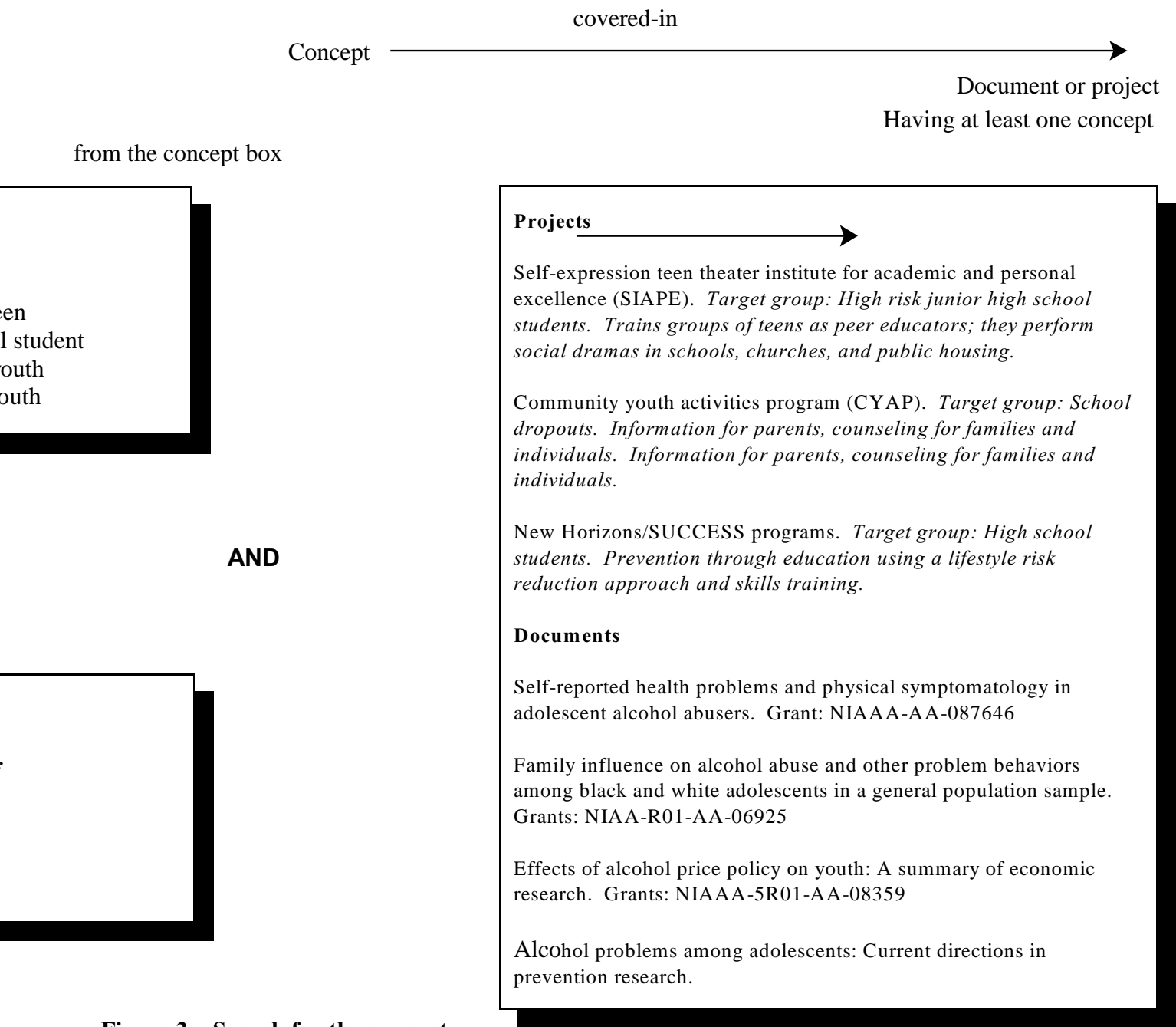
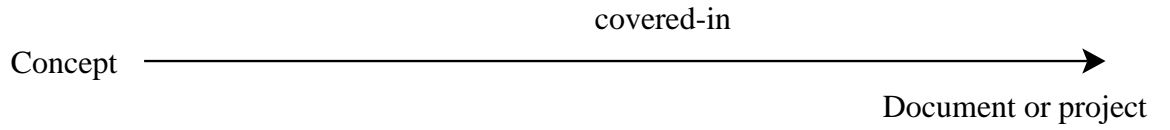


Figure 2. Portion from an information structure schema



**Figure 3a. Search for the concept “adolescent” (enhanced with further concepts from the AOD Thesaurus).**



Having at least one concept

from each concept box

en  
l student  
outh  
outh

AND

environment  
itions

**Projects**

Community youth activities program (CYAP). *Target group: School dropouts. Information for parents, counseling for families and individuals.*

**Documents**

Family influence on alcohol abuse and other problem behaviors among black and white adolescents in a general population sample.  
Grants: NIAA-R01-AA-06925

Family support decreases influence of deviant peers on Hispanic adolescents' substance abuse

**Figure 3b. Search results after ANDing a second concept related to *family*.**

## **Web-based thesaurus / KOS display Requirements**

**Browsing a hierarchy at different levels of detail**

**Hyperlinks for following relationships**

**Searching for compounds containing  
any combination of elemental concepts**

**Searching for a word or phrase (full complement of  
Boolean and adjacency operators).**

**Search in the combination of the descriptor field and  
the synonymous term field**

**For a controlled vocabulary search:  
Insert descriptor or descriptor + narrower terms  
into search form**

**For a free text search:  
Insert descriptor + synonyms or  
descriptor + synonyms  
+ narrower terms + their synonyms  
into search form**

The following pages have examples of a proposed interface that is very simple but functional.  
(Fancy graphics are often more a hindrance than a help.)

Thesaurus interface pages are in file dlthestut2.pdf

Fig. 2a from DL proposal



Fig 2b from DL proposal

# Searching with elemental concepts

## Example 1. AOD Thesaurus

**Search for:**

**central nervous system AND disorder**

**Result:**

**GH6.10.2 brain injury  
GX4 CNS disorder**

**Search for:**

**central nervous system**

**Result:**

**EF4.6.16 CNS injection  
EW8 CNS function  
GH6.10.2 brain injury  
GX4 CNS disorder  
XV4.4.4 CNS sensory pathway  
XZ central nervous system**

# **Searching with elemental concepts**

## **Example 2. LC Classification**

**Search for:**

**buildings, architecture AND acoustics**

**Result:**

**NA2800 Architectural acoustics**

**TH1725 Soundproof construction**

**Search for:**

**vehicles AND acoustics**

**Result:**

**TL681.S6 Airplanes. Soundproofing**

**VM367.S8 Submarines. Soundproofing**

## **Searching with elemental concepts. Ex. 3 DDC**

**Search for: Payment in exchange for some consideration**

**Result: general concepts containing this component with examples of more specific Dewey classes (many, but not all, in 330 Economics)**

### **Wage**

**331.21 Labor economics / Compensation**

**658.32 General management / Personnel management /  
Wage and salary administration**

### **Price/Cost**

**338.52 Production economics / Prices**

**339.42 Macroeconomics / Cost of living (Prices)**

**354.5285 Public administration / Admin. of agriculture /  
Agricultural price supports**

### **Interest**

**332.82 Financial economics / Interest**

**336.2426 Public finance / Income taxes / Interest  
income**

### **Rent**

**336.11 Public finance / Non-tax revenue / Revenues fr.  
rents**

## **Example for a word search in an online thesaurus / KOS**

**Search for**

**commercial AND organization**

**finds the following record**

**corporation**

**ST business organization**

**ST commercial enterprise**

**ST company**

**Search for the German words**

**Hirn AND Entzündung**

**finds the following record**

**meningitis**

***German* Hirn-haut-entzündung**



# **Searching interaction using facets:**

## **Facets for eliciting user needs**

User enters subject field of search.

System displays list of facets (limiting aspects).

User indicates first aspect for limiting the search

Subject field of search: **Education**

Indicate limiting aspects to be used:

- Level of education
- Ethnic origin of students
- Giftedness/handicap of students
- Curriculum subject
- Country of education
- Public/private education



## User selects *level of education* descriptor

### **Level of education**

- Preschool
- Kindergarten
- Elementary education
- Secondary education
- Higher education

# Searching interaction:

## Facets for eliciting user needs

System displays query formulated so far.

User indicates ***Ethnic origin of students***

as a limiting aspect

## Subject field of search

### **Education**

Indicate limiting aspects to be used

- Level of education                      **Elementary**
- Ethnic origin of students
- Giftedness/handicap  
of students
- Curriculum subject
- Country of education
- Public/private education

User selects *Ethnic origin of students* descriptor

## **Ethnic origin of students**

- Latin American / Spanish American
- Mexican American
- Puerto Rican
- African American
- Asian American
- Chinese American
- Japanese American
- . . . .
- Gypsy

# Searching interaction:

## Facets for eliciting user needs

System displays query formulated so far.

User indicates *Curriculum subject* as the next limiting aspect

## Subject field of search

### **Education**

Indicate limiting aspects to be used

- Level of education                    **Elementary**
- Ethnic origin of students           **African  
American**
- Giftedness/handicap  
of students
- Curriculum subject
- Country of education
- Public/private education

# **Searching interaction:**

## **Facets for eliciting user needs**

After a few more interactions, the system displays the completed query formulation

## Subject field of search

### **Education**

Indicate limiting aspects to be used

- Level of education      **Elementary**
- Ethnic origin of students      **African  
American**
- Giftedness/handicap  
of students
- Curriculum subject      **Reading**
- Country of education      **U.S.**
- Public/private education



## **Thesaurus / KOS-supported Web search engines**

These sites are still experimental; they come and go

### **Using synonym expansion**

Google (but limited by thesaurus / KOS)

### **Using homonym disambiguation**

[www.hotbot.com/](http://www.hotbot.com/) (used to do this at one time)

[www.oingo.com](http://www.oingo.com) (will change name to [appliedsemantics.com](http://appliedsemantics.com))

Finds Open Directory categories (Full Web search with  
homonym disambiguation is hard)

Acquired by Google, usage there not known

### **Using a large enriched thesaurus /ontology**

[www.seruba.com/](http://www.seruba.com/) (defunct)

### **Using facets**

[www.endeca.com](http://www.endeca.com) (used in [www.tesco.com/winestore/](http://www.tesco.com/winestore/))

[www.i411.com](http://www.i411.com) (used in [www.yellowpages.co.za](http://www.yellowpages.co.za))





















## Implementing thesaurus / KOS functions in retrieval with emphasis on cross-language retrieval

**Important principle: Let the system do the work.** Full exploitation of thesaurus / KOS power cannot rely on users learning how to use a thesaurus / KOS but rather requires a system that gives behind-the-scene knowledge-based assistance with the thesaurus / KOS serving as the knowledge base.

### Controlled vocabulary

With a controlled vocabulary there is a defined set of concepts used as descriptors in indexing and searching. The user can browse the thesaurus / KOS hierarchies to identify search concepts; or the user can start from a term or phrase and consult the thesaurus / KOS to find the proper descriptor(s) or let the system do the mapping behind the scene. In either case, the user need not worry about the various ways each search concept is expressed in natural language. In cross-language retrieval this simply means that the user should be able to use a term in his own language to find documents (or whatever the retrieval objects are) indexed by the corresponding descriptor (concept identifier). The user can consult a multilingual thesaurus / KOS that includes for each concept corresponding terms from several languages and has an index for each language. Or the mapping from a user term in any covered language to the corresponding descriptor could be automatic. As an example, consider a library catalog using the Library of Congress Subject Headings, for which French and Spanish translations are available. In the VTLS automated library system, each subject heading is identified by a number that is used in the document records. The authority file includes for each subject heading the preferred term and any synonyms; this information can be included in multiple languages. From any user term in English, French, or Spanish the system can map to the corresponding subject heading number through a free-text search on authority records to find any subject heading for which either the preferred term or any synonym contains the user's query word or phrase in any language.

Whenever the mapping from user terms to descriptors is done "behind the scenes", transparent to the user, the system should ask the user for clarification whenever the query word or phrase has multiple meanings and cannot be disambiguated automatically. Beyond that, showing the user the descriptor(s) the system came up with in their hierarchical context might improve the accuracy of the query formulation and thus retrieval. The success of this type of interaction depends on the quality of the hierarchy and the interface.

If voice input is available, one might even include the spoken form of terms in the thesaurus / KOS to enable voice input of query terms which would then be mapped to the appropriate descriptors.

A controlled vocabulary system must promote correct use of descriptors in indexing. Hierarchy and scope notes assist the indexer in understanding the meaning of a descriptor. Request-oriented indexing assures that important descriptors are not overlooked. In cross-language retrieval the thesaurus / KOS version in each language must make sure that the indexer in that

language fully understands the meaning of a descriptor that originated from another language; otherwise, the indexing of such a descriptor will not be consistent across the database.

Automated indexing with a controlled vocabulary, particularly if it is to take a request-oriented slant, can be accomplished with a knowledge base that (1) allows recognition of important words and phrases (in spoken form for speech indexing) and allows for homonym disambiguation and (2) gives mapping rules that lead from the (possibly weighted) set of words and phrases identified for a document to a set of descriptors that should be assigned.

Such mapping rules can take many forms. In their simplest form, they specify a direct mapping from text words or phrases to the appropriate descriptors for each word or phrase (and possibly even word or phrase combinations). To increase accuracy, the mapping can be made dependent on context (Hlava 97). A more complex mapping relies on association strengths between terms (words and phrases) and descriptors. Broadly speaking, the association strength between term T and descriptor D could be seen as the predictive probability that the document containing term T should be indexed with descriptor D. Such association strengths can be computed from a training set of indexed documents. This is the approach often taken in automated text categorization, where often, but not always, the goal is to index each document by only one descriptor (assign it to one of a set of non-overlapping categories). An advanced version of this approach is the use of "topic signatures", profiles consisting of a set of terms with weights; a document is assigned the topic if its terms match the topic signature (Lin 1997). In effect, a topic signature is a query which identifies documents relevant to the topic.

As the foregoing discussion illustrates, the knowledge base needed to support automated indexing is more complex than a Thesaurus / KOS for manual indexing. It must include more terms and term variants so that the words and phrases important for indexing can be recognized in the text, and it must include information needed for the disambiguation of homonyms (which often requires determining the part of speech of a text word).

For indexing and searching, a controlled-vocabulary cross-language retrieval system can be seen as a set of monolingual systems, each of which maps the terms from its language to a common system of concepts used in indexing and searching. For manual indexing and query formulation, this is accomplished through a multilingual thesaurus / KOS, which may in fact consist of multiple monolingual thesauri linked through common descriptor identifiers (such as Dewey Decimal class numbers). Automated indexing in cross-language text retrieval with texts in multiple languages means mapping from each language to the common conceptual structure represented in the controlled vocabulary. The knowledge base component dealing with identification of words and phrases for automated indexing can be developed independently for each language. Mapping rules that are entirely term-based can also be developed independently for each language. However, some mapping rules, for example rules based on context or topic profiles, may include conceptual elements that could be shared across languages.

There are a number of controlled-vocabulary cross-language retrieval systems based on manual indexing in use in bilingual or multilingual areas such as Switzerland, Belgium, Canada, and areas of the US with large Spanish-speaking populations; in international organizations, such as the European Community; and in international collaborative systems, such as AGRIS. These systems are based on the Universal Decimal Classification, which has been translated into many languages (library of the ETH, Zurich); on the Library of Congress Subject Headings (translated into French); on EUROVOC, an EC thesaurus in 9 languages; and AGROVOC, a thesaurus in

six plus languages created by translation from its original English-only version. There are a large number of thesauri that either have been developed as multilingual thesauri or have been translated into several languages.

### **Free-text searching**

High-recall (even moderate-recall) free-text searching requires query-term expansion as discussed above. Cross-language free-text searching, finding texts in one language that are relevant for a query formulated in another language without relying on controlled vocabulary indexing, is an extension of this principle: Each query term must be mapped to a set of search terms in the language of the texts, possibly attaching weights expressing the degree to which occurrence of a search term in a text would contribute to the relevance of the text to the query term. To assist with this task, a thesaurus / KOS must include the mapping information. If the thesaurus / KOS includes fine-grained definitions that deal with subtle differences of meaning, distance between such definitions can be used to derive term weights.

A major difficulty of this mapping is that a homonym used in the query gives rise to multiple translations, each corresponding to one of its meanings. The target terms may in turn be homonyms in their language and thus retrieve many irrelevant documents unless text terms are disambiguated. (This problem exists in synonym expansion in one language as well but is exacerbated in cross-language text retrieval.) When the mapping goes to a term that has multiple meanings, the specific meaning should be identified, possibly in interaction with the user. For best retrieval results the terms in the texts should also be disambiguated so that only documents that include the term in the right sense score

The issue of homonymy in retrieval is not as straightforward as it may seem at first glance (Sanderson 1994). First of all, quite a bit of disambiguation may occur “naturally”, in that a given term may assume only one of its meanings in the specific domain of the collection and therefore in the queries. Second, in a multi-component query, a document that includes a homonymous term from the first query component in a meaning other than that intended in the query is unlikely to also include a term from another query component; thus excluding irrelevant documents may not require disambiguation in either the query or the texts. On the other hand, with single-concept query to a general collection (such as the World Wide Web), disambiguation can be expected to have a beneficial effect on retrieval performance. Failing that, a system might be able to suggest to the user an additional query component that would separate out the documents that include the query term but in a different meaning. Note that information extraction is much more dependent on homonym disambiguation.

In any event, for best support of free-text retrieval a thesaurus / KOS should flag homonyms, give their senses, and include rules for disambiguation.

The greater difficulty of free-text cross-language retrieval stems in no small measure from the fact that one must work with actual usage, while in controlled-vocabulary retrieval one can, to some extent, dictate usage.

## **Thesauri for knowledge-based search support**

Whether searching is by controlled vocabulary or by free text, it is often helpful to the user to browse a well-structured and well-displayed hierarchy of concepts, preferably with the option of including definitions. A more sophisticated system may guide a user through a facet analysis of her topic. These aids provided by the system enable the user to form a better idea of her need and to locate the most suitable descriptors or free-text search terms. The guidance through facets and their hierarchical display must be available in the language of the user. These suggestions are based on the assumption that browsing a hierarchy is natural to most users and that users will appreciate the structure provided. This assumption rests on the belief that people try to make sense of the world and that guided facet analysis and browsing well-structured hierarchies help them do so. There is anecdotal evidence to support this assumption, but it needs to be investigated by building prototype systems and studying users' success (see, for example, Pollitt 1996).

This is one example of using a thesaurus / KOS as a knowledge base to make searching more successful. The assistance provided does not require that the user be an expert in thesauri / KOS. This is even more true for "behind-the-scenes" assistance. There is no need to teach users about following a cross-reference from a synonym to a descriptor if the system searches for the descriptor automatically. There is no need to tell the user to look under narrower terms also if the system can do a hierarchically expanded search. There is no need to tell the user about strategies of broadening the search if the system, in response to a user input that not enough was found, can suggest further descriptors to be searched based on cross-references in the thesaurus / KOS. Sophisticated retrieval software can make the use of thesauri / KOS in retrieval independent of the user's knowledge and thereby can get much more mileage out of the investment in thesauri / KOS.

## **Special issues in multilingual thesauri / KOS**

A multilingual thesaurus / KOS for indexing and searching with a controlled vocabulary can be seen as a set of monolingual thesauri / KOS that all map to a common system of concepts. With a controlled vocabulary, indexing is concept-based; cross-language retrieval is simply a matter of providing designations for these concepts in multiple languages so that queries can be written in multiple languages. However, as the example given above illustrates, conceptual systems represented in the vocabulary of different languages do not completely coincide.

The crux of the matter, then, is which concepts to include. Ideally, the thesaurus / KOS should include all concepts needed in searching by any user in any of the source languages. Language differences often also imply cultural and conceptual differences, more so in some fields than in others. We need to create a classification that includes all concepts suggested by any of the languages. At a minimum this includes all relevant concepts lexicalized in at least one of the source languages. Also, different languages often suggest different ways of classifying a domain; the system needs to be hospitable to all of these. The problem that has bedeviled many developers of multilingual thesauri / KOS is that a concept lexicalized in one language may not be lexicalized in another and that the terms that do exist often vary slightly in meaning, possibly giving rise to different relationships. Starting from the misguided notion that a thesaurus / KOS should include only concepts for which there is a term in the language and that term meanings cannot be adjusted for purposes of the thesaurus / KOS, they had difficulty making the system of concepts the same for all languages. But, as we have seen, even in a monolingual thesaurus / KOS the lexicographer often discovers concepts needed in searching or to enhance the logic of the concept hierarchy and then needs to create terms for these concepts. In multilingual thesauri / KOS this necessity arises more often, particularly when different languages differ in the hierarchical levels at which they lexicalize concepts.

The principle proposed here is to establish a common conceptual system, which may require an arduous and expensive, process of negotiation, and then arrange for the terms in all languages to fit, giving proper definitions, of course. In contrast, many multilingual thesauri / KOS have been produced by translating an established monolingual thesaurus / KOS, thus accepting the conceptual system of one language and limiting the number of synonyms (if any) in the other languages. EuroWordNet is based on a more comprehensive, but still limited approach: Rather than developing a conceptual structure based on an analysis of the vocabulary in all participating languages, EuroWordNet accepts the conceptual system of the English language WordNet. On the other hand, EuroWordNet does not simply translate WordNet but develops synonym sets independently in each language and then links them to the concepts (synsets) established in WordNet.

So far we have described a multilingual thesaurus / KOS with a common conceptual system, however constructed, where the terms of each language are linked to a concept but not to each other. Relationships between terms from different languages are established through their relationships to concepts. This simple model will do for most information retrieval. But unless the concepts are exceedingly fine-grained and include in their definition affective components of meaning and usage considerations, this model is too simplistic for natural language processing, especially translation. There one needs direct relationships between terms to enable the proper word choice in translation.

The problems discussed here and illustrated in the example above have major implications for cross-language free-text searching: Each query term should be mapped from the source language to its multiple equivalents in the target language; each of these equivalents may have other meanings in the target language, presenting potential problems for retrieval. The query term may not have a precise equivalent in the target language; one may need to map to broader or narrower terms, distorting the meaning of the original query.

# Key issues in multilingual thesauri / KOS

**Conceptual systems in different languages differ**

**What concepts are lexicalized differs from language to language**

**Translation of an English thesaurus / KOS into French does not make a French thesaurus / KOS**

**Develop common conceptual structure integrating perspectives from multiple languages. Harmonize concepts where possible, keep concepts where necessary, invent a term if a concept is not lexicalized in a language**

**Problems of structure: simplified versus real**

## Simplified

English term 1		French term 1
English term 2	Concept	French term 2
English term 3		French term 3

## Real

English term 1		French term 1
English term 2	Concept	French term 2
English term 3		French term 3



# **Evaluation of Thesauri**

## **Introductory example: Yahoo classification**

Yahoo home page

<p><b>Reference and General Interest</b></p> <p><b>Reference</b> Libraries, Dictionaries, Quotations ...</p> <p><b>Computers &amp; Internet</b> Internet, WWW, Software, Games ...</p>	<p><b>Subjects</b></p> <p><b>Science</b> Animals, Astronomy, Engineering ...</p> <p><b>Health</b> Medicine, Diseases, Drugs, Fitness ...</p> <p><b>Social Science</b> Archaeology, Economics, Languages ...</p> <p><b>Society &amp; Culture</b> People, Environment, Religion ...</p> <p><b>Government</b> Elections, Military, Law, Taxes ...</p> <p><b>Business &amp; Economy</b> B2B, Finance, Shopping, Jobs ...</p>
<p><b>News &amp; Media</b> Full Coverage, Newspapers, TV...</p> <p><b>Entertainment</b> Movies, Music, Humor, Cool Links ...</p> <p><b>Recreation &amp; Sports</b> Sports, Travel, Autos, Outdoors...</p>	<p><b>Education</b> College and University, K-12 ...</p> <p><b>Arts &amp; Humanities</b> Literature, Photography ...</p>
<p><b>Regional</b> Countries, Regions, US States ...</p>	

**Yahoo Classification. Home. Meaningful arrangement**

Yahoo education

## Yahoo classification. Education. Meaningful arrangement.

Home >

### Education

#### Categories

Browse by Region (170)

By Culture or Group (398)

By Subject (11)

#### Information sources

Bibliographies (4)

Web Directories (47)

News and Media (83)

Chats and Forums (40)

Conferences (52)

Journals (36)

Statistics (6)

#### Education by level

Early Childhood Education (90)

K-12 (53910)

Higher Education (16638)

Adult and Continuing Education (325)

#### Special students and subjects

Special Education (168)

Disabilities@

Literacy (12)

Bilingual (24)

Career and Vocational (236)

Correctional@

#### Educational methods

Theory and Methods (659)

Teaching (63)

Instructional Technology (334)

Distance Learning (476)

Standards and Testing (63)

Academic Competitions (79)

Graduation (53)

#### Political and economic aspects

Policy (52)

Reform (70)

Equity (27)

Financial Aid (395)

Employment (143)

#### Organizational aspects

Government Agencies (77)

Organizations (3008)

Companies@

Programs (322)

Several pages from Yahoo



















## Evaluation of Knowledge Organization Systems (KOS)

### Characteristics for describing and evaluating KOS

(classifications/ontologies/taxonomies/index languages/thesauri/glossaries/dictionaries)

(For some items, a section number from Soergel, Organizing information (starting with a digit), and/or Soergel, Indexing languages and thesauri (starting with a capital) is given.)

#### 1. Overall description and evaluation

##### 1.1 Purpose, for example (see full list given earlier)

Providing "conceptual infrastructure"

Mapping out the conceptual structure and providing a common language for a field

Providing classification/typology and concept definitions. Clarifying concepts by putting them into context. Thus providing orientation and serving as a reference tool for individual researchers and practitioners and thereby

Assisting with the exploration of the conceptual context of a research problem and in structuring the problem, thereby providing the conceptual basis for the design of good research, for the consistent definition of variables, and thus the cumulation of research results.

Providing the conceptual basis for the exploration of the various aspects of a program in program planning, in the identification of approaches and strategies, and in the development of evaluation criteria

Information retrieval (IR), database searching, expert systems, and other AI applications

One information system

Several IR systems, switching language. Support the coordination or combination of several databases in the same area to facilitate access to multiple databases.

Assisting readers in understanding text

Assisting writers with conceptualizing a topic and with finding the proper term

Translation

Language learning

In each case specify the intended audience

**If purpose is IR specify**

Information system(s) in which the vocabulary is to be used

Type of objects to be retrieved (text documents, images, people, etc.)

Use of the vocabulary

Vocabulary control in indexing and searching (controlled vocabulary)

Vocabulary control only for searching. Assist with clarifying a search topic and assembling all applicable concepts and terms, whether searching with a controlled vocabulary or free-text.

IR technique(s) (such as: printed index, computer search system). Support of inclusive (hierarchically expanded) searching

Automated vs. manual indexing or query formulation. Approach to indexing to be supported: Request-oriented vs. entity-oriented

Techniques for eliciting user needs (e.g., menu based on search tree; questions based on facet structure)

1.2 **Relationship to other KOS**, especially standard schemes

1.3 **Summary evaluation** of the vocabulary's adequacy for the stated purpose based on the more detailed analysis as outlined below.

Note:

In the following, links to two textbooks are given in [] as follows:

chapter or section numbers starting with a digit, such as 14.3, refer to

Soergel, *Organizing Information*

chapter or section numbers starting with a letter, such as F0.4.3, refer to

Soergel, *Indexing Languages and Thesauri. Construction and Maintenance*



## 2. Coverage of concepts and terms. Sources, quality of usage analysis.

2.1 Concepts: scope, breadth of coverage (See also 2.3.1 below)

2.2 Concepts: specificity, depth of coverage

Completeness of coverage at each level of specificity considering all concepts (descriptors and other preferred terms) and descriptors alone [F0.4.3]

Specificity must be adapted to the purpose. Assistance in the choice of terms or the comprehension of text requires many nuances. An IR system for propositions requires high specificity. A bibliographic IR systems may require only low specificity.

2.3 Sources from which concepts and terms are included (natural languages, classifications/thesauri / KOS).

Relationship to other vocabularies, especially standard schemes.

For each source:

2.3.1 Completeness of coverage; all vs. selected concepts; all vs. selected terms for each concept (this includes coverage of synonyms)

2.3.2 Quality of analysis of actual term usage in the source.

2.3.3 Recency

Specifically: Completeness of coverage of the terminology from a given language (English, French, German, Chinese, etc.; the language is the source)

2.4 Augmentation of sources through concepts created in concept analysis [15, C3]

2.4.1 Are all necessary facets included?

2.4.2 Formation of new concepts arising from semantic factoring and other methods of concept analysis. Specifically: Are the concepts applicable across disciplines? Are the concepts applicable across different societies and cultures? (See also 3.4 below)

2.5 Choice of terms

2.5.1 Form of terms - consistency, adherence to common usage.

2.5.2 Appropriateness of selection of preferred terms from among synonyms.

2.5.3 Choice of terms to designate descriptors [F0.4.2] Closeness to user terminology.

2.6 Nature of notation (if none, state that) [D4]

### 3. **Conceptual analysis and conceptual structure. Terminological analysis**

#### 3.1 Quality of conceptual structure [14, C1]

Types and degree of differentiation of conceptual relationships included:

3.1.1 Expression of concepts through elemental concepts (closely related to definition)

3.1.2 Hierarchical relationships (polyhierarchy)

3.1.3 Associative relationships

3.1.4 Rich set of semantic relationships (full-fledged ontology)

Completeness and correctness of conceptual relationships included.

#### 3.2 Quality of definitions, explications, scope notes (correctness, detail, clarity). [C3]

#### 3.3 Completeness of terminological relationships.

Does the thesaurus / KOS contain terms that are synonymous or quasi-synonymous without indicating the relationship?

### 4. **Use of precombination in the index language** (cuts across 2 and 3) [14, 15, C2]

#### 4.1 To what degree are descriptors precombined?

#### 4.2 To what extent are precombined descriptors enumerated and/or given in the alphabetical index? Built by the indexer? Updating characteristics.

Are precombined descriptors designated by an independent symbol or a string of symbols? Citation order free or fixed? To what extent do the components of a precombined descriptor determine its place in the arrangement? (Relates also to 5)

## 5. **Access and display. Format of presentation of the vocabulary**

Consider for each format access/retrieval by concepts versus access/retrieval by terms

Access can be provided through arrangement in a printed document or through a computerized search system.

### 5.1 Format of printed document

#### 5.1.1 Overall format [D1]

Thesaurus / KOS parts and information given in each, connections between them.

Is the overall format clear and helpful for finding the appropriate concepts and terms or notations in indexing and query formulation?

#### 5.1.2 Display of conceptual relationships

- through arrangement [15.5.2, C2, D3]
- through cross-references [D3.1.1,1]
- through descriptor-find index [15.5.1, D3.6]

How well does the display reflect the conceptual analysis (e.g., sequence of concepts on some hierarchical level) [D3.1.2]

#### 5.1.3 Display of terminological relationships. Format of alphabetical index [C5]

### 5.2 Access through computer systems. Retrieval of concepts and terms. Navigation. Format of on-line displays

#### 5.2.1 Overall format. Available windows and their relationships

5.2.2 Display of conceptual relationships, esp. hierarchy. Localized hierarchical chains vs. entire hierarchy. Overviews and total hierarchy. Expandable levels vs. expanded or expand-all option. Graphical displays, concept maps. Are cross-references active hyperlinks? Is there an online descriptor-find index.

5.2.3 Access by words and phrases. Is the thesaurus / KOS database searchable? How does the search work. What is searched? Just the term itself, synonyms, scope notes, all cross-references (not good!)?

### 5.3 Format of machine-readable form (if any). What standard is followed

### 5.4 Detail of keeping records of the origin of information included in the vocabulary.

## 6. **Updating**

## Outline for the analysis of KOS. DDC

### 1. Purpose

- 1.1 **Information system** or type of information system in which to be used  
*Bibliographic information system. Intended for public and school libraries.*
- 1.2 **Intended for** controlled vocabulary indexing  or query term expansion  [Ch. 12, Introduction]
- 1.3 **Type of file and search mechanism** for which originally designed  
Shelving  Card catalog  Online system  (Now promoted for Web subject directories)

### 2. Coverage and designation of concepts. Coverage and format of terms

- 2.1 **Concepts: Scope**, breadth of coverage. Recency of concepts  
*Universal — covers all of knowledge. But focus on Western culture, esp. US.*
- 2.2 **Concepts: Specificity**, depth of coverage. (Section 16.2.2). Coverage at each level of specificity.  
*Medium specificity. Would need closer analysis by subject area. Geography table quite specific.*
- 2.3 Are all needed **facets** included? Concepts formed in semantic factoring and facet analysis? (S.a. 3.1)  
*Some general concepts included in the general tables and the in-schedule tables. Many others not included by themselves but only as components in one or more precombined descriptors. Completeness of explicit and implied facets? Answer would require extensive analysis.*
- 2.4 **Terms: Completeness of coverage** (completeness of lead-in vocabulary). Recency of terms  
*Some lead-in synonyms included in the alphabetical index. How complete? Would need extensive analysis!*
- 2.5 **Form of terms**: Consistency, adherence to common usage. *Terms seem appropriate. Many classes cannot be expressed by a simple term but need a phrase devised by the editor.*
- 2.6 **Nature of notation** (if none, state that). [Section 15.5.2] *Decimal, highly expressive (with some exceptions).*

### 3. Terminological and conceptual analysis and conceptual structure.

- 3.1 **Quality of conceptual structure** (14): Facet analysis. Types and degree of differentiation of conceptual relationships included. For each type indicate the completeness of inclusion. (Fill in 3.1.1 - 3.1.3)
  - 3.1.1 Expression of concepts through elemental concepts (closely related to definition)  
*For enumerated compound concepts: Sometimes done implicitly in the relative index. For precombined descriptors constructed according to DDC rules: Done by the indexer.*
  - 3.1.2 Hierarchical relationships (polyhierarchy) (Shown by arrangement or Broader Term / Narrower Term X-ref)  
*Monohierarchical. A few additional BT/NT through cross-references. Many hierarchical relationships implied by the relative index (Example: The classes shown under Blind).*
  - 3.1.3 Associative relationships. (Implied by physical proximity in the arrangement or explicit Related Term X-ref)  
*Some explicit cross-references*
- 3.2 **Quality of definitions**, explications, scope notes (correctness, detail, clarity).  
*Many notes throughout the schedules and in the Manual. Mostly usage notes explaining the difference between classes or instructions on how to form new precombined descriptors. A few definitions*
- 3.3 Completeness of terminological relationships: Does the vocabulary contain terms that are synonymous or quasi-synonymous without indicating the relationship? *Not a problem in a classification like DDC.*

4. **Use of precombination in the index language** (concerns both 2 and 3) [14, 15, esp. 15.4]

4.1 To what degree are descriptors precombined?

*DDC can be used with a medium to high degree of precombination, depending on how many new precombined descriptors the indexer builds.*

4.2 To what extent are precombined descriptors enumerated and/or given in the alphabetical index?

*Medium degree of enumeration in the schedules, some addl. precombined descriptors in the index.*

To what extent can the indexer build additional precombined descriptors?

*To a large extent. Libraries differ in their use of this option.*

Are precombined descriptors designated by an independent symbol or a string of symbols? Combination order free or fixed? To what extent do the components of a precombined descriptor determine its place in the arrangement? (Relates also to 5) [Section 15.5.2]

*Enumerated precombined descriptors have their own independent symbol (which sometimes is constructed using notation components from tables). Combination order is fixed. The components completely determine the place of a precombined descriptor built by the indexer.*

5. **Access and display. Format of presentation of the vocabulary**

Consider for each format access/retrieval by concepts versus access/retrieval by terms.

Access can be provided through arrangement in a printed document or through a computer search system.

5.1 **Format of printed document** (Fill in 5.1.1- 5.1.3)

5.1.1 Overall format: Thesaurus / KOS parts and information given in each, connections between them. Is the overall format clear and helpful for finding the appropriate concepts and terms or notations in indexing and query formulation?

*Introduction (v.1), Tables (v.1), Schedules (v.2+3), Relative Index (v.4), Manual (v.5)  
Need to go back forth between schedules and manual, otherwise reasonably helpful.*

5.1.2 Display of conceptual relationships (Broader Term, Narrower Term, Related Term)

- through linear arrangement or graphical display [Section 15.5.2]

*In the tables and schedules.*

- through cross-references [Section 14.1]

*In the tables and schedules.*

- through descriptor-find index [Section 15.5.1]

*The relative index combines the functions of an alphabetical index and a descriptor find index of sorts.*

How well does the display reflect the conceptual analysis, e.g., sequence of concepts on the same hierarchical level (sequence of the children of a concept, that is, the concepts one level further down).

*Usually the sequence of classes makes good sense.*

5.1.3 Display of terminological relationships (Synonymous Term)

*Terminological relationships are displayed only in the relative index, which gives the lead-in term and points to the appropriate class number.*

5.2 **Access through computer system.** Navigation. Format of on-line displays

*This would be an analysis of Dewey for Windows. Not required here.*

## **Some points on procedure for evaluating a KOS**

### **What went into the development of a KOS**

Check sources used.

Check procedures used in KOS development.

### **Examine the KOS structure and content**

Use knowledge of KOS structure for analysis of structure and internal consistency.

Check against other KOS and against encyclopedias, dictionaries, or other authoritative sources.

In this examination, collect data on all the criteria in parallel by looking through the KOS, probing for examples and following leads as they arise. Interact with the KOS. Keep notes according to the outline of criteria. (Much like anthropological field work, where the investigator observes as events occur, keeping the variables of interest in mind but is always open to aspects not thought of beforehand. At some point, the notes are indexed and sorted by the variables of interest.)

### **Check how well the KOS supports the functions it was designed for**

If the KOS purpose is to support finding items (documents, software, products, people, jobs, etc., either indexed by the KOS or searched free-text), try the KOS to see how well search requests can be expressed and how well the aspects under which items should be found can be represented. Conduct indexing and retrieval experiments. Observe users interacting with the KOS online; see how it helps them in formulating a successful query. Study the effects of behind-the-scenes query expansion through the KOS. Such tests can be done by the evaluator - for example, to shed light on completeness - or by real indexers and real searchers, making sure that the indexers have sufficient domain knowledge and training in the use of the KOS. Indexing experiments where several indexers index the same documents can be very useful; disagreements may point out problems in the KOS.

If the purpose of the KOS is to support reasoning, run sample reasoning tasks and check the results. This may require tracing a reasoning process, paying particular attention to the steps where the KOS was used. If the purpose of the KOS is business process reengineering, talk to the people who do the task to see how the KOS helps them. If the purpose of the KOS is to collect and use statistics, check how well the KOS supports consistent data collection (depends on good definition of the concepts) and how well it supports finding data and processing them to derive indices or decision support.

## Resources

### Directories of thesauri / KOS and dictionaries on the Web.

Printouts from the following Web sites were included in the paper tutorial notebook:

<http://nkos.slis.kent.edu/>

Networked Knowledge Organization Systems (NKOS).

Has a workshop at ECDL and ECDL

<http://www.lub.lu.se/desire/desireIIindex.html#idx>

DESIRE I and II. The role of classification schemes in Internet resource description and discovery

[www.darmstadt.gmd.de/~lutes/thesoecd.html](http://www.darmstadt.gmd.de/~lutes/thesoecd.html)

Web Thesaurus Compendium (representative list with descriptions)

[www.onelook.com](http://www.onelook.com) OneLook Dictionaries. The Faster Finder

[www.yourdictionary.com/](http://www.yourdictionary.com/)

[www.strategic-road.com/pratique/dicofr.htm](http://www.strategic-road.com/pratique/dicofr.htm) Strategic Road Dictionaries

<http://linguistlist.org/sp/Dict.html>

[www.mikesart.net/giantglossarycom](http://www.mikesart.net/giantglossarycom) Terminology - Search

<http://www.asel.udel.edu/nli/nlp/lrd.html> The Language Representation Database Project

XXX

Enterprise search report:

<http://edge.networkworld.com/news/2008/070208-top-search.html>

**A short bibliography follows**

### Thesaurus / KOS software Web sites

<http://www.willpower.demon.co.uk/thessoft.htm>  
[http://sky.fit.qut.edu.au/~middletm//cont\\_voc.html](http://sky.fit.qut.edu.au/~middletm//cont_voc.html)  
[http://www.fbi.fh-koeln.de/fachbereich/labor/Bir/thesauri\\_new/indexen.htm](http://www.fbi.fh-koeln.de/fachbereich/labor/Bir/thesauri_new/indexen.htm)  
<http://bak-information.ub.tu-berlin.de/software/term.html> (covers a wider range of software, annotations in German)

### Concept mapping resources

<http://eaa-knowledge.com/ojni/ni/602/strategies.htm>  
<http://www.graphic.org/concept.html>  
<http://www.observatory.com/conceptmappingvs1.htm>

### Ontology resources. Ontology editor examples

**Protege** (client-based and Web-based versions)  
<http://protege.stanford.edu/>, click on Documentation

Biomedical ontologies: [www.obofoundry.org/](http://www.obofoundry.org/)  
<http://ontology.buffalo.edu/smith/>

On **description logic** see, for example <http://potato.cs.man.ac.uk/seanb/publications.php>

The URLs given on the standards page are also useful more generally, esp. on **RDF and OWL**

### Topic maps

[www.ontopia.net](http://www.ontopia.net).  
 For an example application, see [www.ontopia.net/omnigator/models/index.jsp](http://www.ontopia.net/omnigator/models/index.jsp) (try opera.xtm)

### Search terms for a Web search for thesauri / KOS

(ontolog\* OR classification\* OR Klassifikation\* OR taxonom\* OR thesaur\* OR dictionar\* OR  
 dictionnaire OR Woerterbuch OR glossar\* OR glossaire OR “word list” OR lexicon OR lexique  
 OR Lexik\* OR terminolog\* OR vocabulaire OR vocabulary OR “knowledge organization” OR  
 “knowledge structure” OR “authority list”)

Note: "classification" may pull a lot of irrelevant URLs

Possibly add OR concept OR mot-clé OR keyword OR “subject heading” OR definition

It is best to require these terms in the title. Otherwise there will be a lot of irrelevant material retrieved, especially by the term *classification*.





























































## Short bibliography

**Website:** <http://www.clis.umd.edu/faculty/soergel/dlthestut>

### Basic information retrieval and classification concepts

Soergel, Dagobert, 1985

**Organizing Information. Principles of data base and retrieval systems.**

Orlando: Academic Press, 1985. 450 p.

Vickery, Bryan C.

**Faceted classification.**

London: Aslib, 1970.

### Thesaurus textbooks

Soergel, Dagobert

**Construction and maintenance of indexing languages and thesauri**

New York: Wiley, 1974. 632 p.

Lancaster, F. Wilfrid

**Vocabulary control for information retrieval. 1.ed.**

Washington, D.C.: Information Resources Press, 1986. 233 p.

(2. ed. not as good)

Aitchison, Jean; Gilchrist, Alan; Bawden, David

**Thesaurus Construction and Use : A Practical Manual. 4. ed.**

London: Fitzroy Dearborn, 2000. 230 p.

Also watch for the Proceedings of the ASIS SIG/CR Classification Research Workshop, published as Advances in Classification Research by Information Today

**Standards** (use with caution) (see the section on Standards in Part 2 of the tutorial)

National Information Standards Organization

**Guidelines for the construction, format, and management of monolingual controlled vocabularies.**

Bethesda, MD: NISO Press; 2005. ANSI/NISO Z39.19-2005. ISBN 1-880124-65-3

<http://www.slis.kent.edu/~mzeng/Z3919/9Z3919.htm>

International Organization for Standardization.

**Documentation--guidelines for the establishment and development of monolingual thesauri. 2. ed.**

Geneva: International Organization for Standardization; 1986.

International Standard ISO 2788-1986(E). [www.collectionscanada.gc.ca/iso/tc46sc9/standard/2788e.htm](http://www.collectionscanada.gc.ca/iso/tc46sc9/standard/2788e.htm)

International Organization for Standardization.

**Documentation--guidelines for the establishment and development of multilingual thesauri.**

Geneva: International Organization for Standardization; 1985.

International Standard ISO 5964-1985(E). [www.collectionscanada.gc.ca/iso/tc46sc9/standard/5964e.htm](http://www.collectionscanada.gc.ca/iso/tc46sc9/standard/5964e.htm)

## **Machine-Readable Dictionaries and Computational Linguistics Research**

Walker, Don, ed. 1995; Zampolli, A., ed.; Calzolari, N., ed.. **Automating the Lexicon: Research and Practice in a Multilingual Environment.** Oxford University Press, 1995.

Cole, Ronald A., editor-in-chief 1996. **Survey of the State of the Art in Human Language Technology.** With Chapter 12 Language Resources and Section 12.4 Lexicons.

<http://www.cse.ogi.edu/CSLU/HLTsurvey/HLTsurvey.html>

Hutcheson, H.M. (1995) **Preparation of multilingual vocabularies.** *Standardizing and Harmonizing Terminology: Theory and Practice.* Philadelphia, PA: American Society for Testing and Materials. (1995): 102-114.

## **Other relevant publications by the tutorial instructor**

**A universal source Thesaurus as a classification generator.**

J. Amer. Soc. for Info. Sci. 1972.9; 23(5): 229-305.

**Indexing and retrieval performance: The logical evidence.**

J. Amer. Soc. for Info. Sci. 1994.9; 45(8): 589-599. (Invited paper)

Reprinted in From classification to "knowledge organization": Dorking revisited or "Past is prelude" / Edited by Alan Gilchrist. - The Hague: FID, 1997. - xiv, 186 p. - (FID pub. no. 714; FID Occasional paper 14). - ISBN 92 66 00 714 5

**The Art and Architecture Thesaurus (AAT): A critical appraisal.**

Visual Resources. 1995; 10(4): 369-400.

**Software support for Thesaurus construction and display.**

Proceedings of the 5th ASIS SIG/CR Classification Research Workshop. Held at the 57th ASIS Annual Meeting, Oct. 16-20, 1994, Alexandria, VA.

Silver Spring, MD: American Society for Information Science. Special Interest Group / Classification Research. 1994.10; 5: 157-184. (Advances in Classification Research. v. 5)

**Data structure and software support for integrated thesauri.**

Paper presented at the Research Seminar on Compatibility and Integration of Order Systems, Warsaw, Poland, September 13-15, 1995.

Published in *Compatibility and Integration of Order Systems: Research Seminar. Proceedings of the TIP/ISKO Meeting*. Issued by International Society for Knowledge Organization; Polish Library Association; Soc. for Professional Information. Warsaw: Wydaw. SBP; 1996. p. 47-57. (Included in the notebook).

**SemWeb. Proposal for an open, multifunctional, multilingual system for integrated access to knowledge base about concepts and terminology.**

Proceedings of the Fourth International ISKO Conference, 15-18 July 1996, Washington, DC. Frankfurt/Main: Indeks Verlag; 1996. (*Advances in Knowledge Organization*, v. 5). p. 165 - 173

**Multilingual thesauri in cross-language retrieval.** Paper presented at the AAI-97 Spring Symposium Series. *Cross-Language Text and Speech Retrieval*. Stanford, CA March 24-26, 1997. Published in the Symposium Technical Report

**The rise of ontologies or the reinvention of classification**

*Journal of the American Society for Information Science*. Sept. 1999; 50(12): 1119-1120.

**ASIST SIG/CR Classification Workshop 2000. Classification for user support and learning. Report.**

[A 3500 word substantive summary and an analysis into 8 themes in classification research.]  
*Bulletin of the American Society of Information Science and Technology*. April/May 2001; 27(4): 15-21 and *Knowledge Organization*. 2000; 27(3): 165-172.  
[www3.interscience.wiley.com/cgi-bin/fulltext/109860106/HTMLSTART](http://www3.interscience.wiley.com/cgi-bin/fulltext/109860106/HTMLSTART)

**The many uses of classification: Enriched thesauri as knowledge sources.**

Keynote address, ASIS SIG/CR Classification Research Workshop 12. 2001-11-4  
*Advances in Classification Research*—Vol. 12. *Proceedings of the 12th ASIS SIG/CR Classification Research Workshop*. Edited by Efthimis N. Efthimiadis. [dsoergel.com/cv/B78.pdf](http://dsoergel.com/cv/B78.pdf)

**Building a more meaningful Web: from traditional knowledge organization systems to new semantic tools.** Report on the 6<sup>th</sup> Networked Knowledge Organization Systems (NKOS) Workshop, JCDL 2003, May 31, 2003 in Houston, TX. *SIGIR Forum*. 2003 Fall; 37(2): 65-72.

Dagobert Soergel; Boris Lauser; Anita Liang; Frehiwot Fisseha; Johannes Keizer; Stephen Katz. **Reengineering thesauri for new applications. The AGROVOC example**  
*Journal of Digital Information*, Volume 4 Issue 4, 2004 March, Article No. 257,  
[jodi.tamu.edu/Articles/v04/i04/Soergel/](http://jodi.tamu.edu/Articles/v04/i04/Soergel/)

**Information organization**

David Bainbridge, ed. *Berkshire Encyclopedia of Human-Computer Interaction*. Great Barrington, MA: Berkshire; July 2004. ISBN: 0-9743091-2-5. p. 355-360

Nathalie Aussenac-Gilles, Dagobert Soergel. **Text analysis for ontology and terminology engineering.** *Applied Ontology*. 2005; 1(1): 35 - 46

## Bibliographies of thesauri

Gerstenkorn, A. 1985 ed.; Rolland, M. Th. ed.; et al.

**Thesaurus guide. Analytical directory of selected vocabularies for information retrieval.**

Amsterdam: Elsevier; 1985. 748p.

Basis for the Echo database of thesauri

**International Classification and Indexing Bibliography. Vol. I: Classification Systems and Thesauri 1950-1982.** ICIB 1. 160 pages, DIN A4, DM 48.80; ISBN 3-88672-300-3; FID-Publ.610. Frankfurt/M: Indeks Verlag; 1982. This comprehensive bibliography of all universal and special classification systems and thesauri which could be found in the literature as well as in libraries, listing some 2300 titles from the time 1950-1982,

Includes bibliography of editions in multiple languages of

Universal Decimal Classification (UDC)    Library of Congress Classification (LCC)

Dewey Decimal Classification (DDC)    Library of Congress Subject Headings (LCSH)

Chan, Lois Mai; Pollard, Richard. **Thesauri used in online databases: an analytical guide.**

New York: Greenwood; 1988. 268 p.

Brewer, Annie M.ed. **Dictionaries, Encyclopedias, & Other Word-Related Books.** 4th ed.

Detroit: Gale Research; 1988. 1333 p. ISBN 0810304406

Molho, Emanuel. **The dictionary catalogue.** Second edition., 178 pages; French & European Publications, Inc., New York; 1989. (Bibliography of mono-, bi-, and multilingual dictionaries)

Brühl, Bettina. **Thesauri und Klassifikationen : Naturwissenschaften - Technik -**

**Wirtschaft.** Potsdam : Verl. für Berlin-Brandenburg, 2005; 354 p. ISBN 3-935035-63-2

## Examples of multilingual thesauri

**Thesaurus EUROVOC:** Official journal of the European communities. Office for Official Publications of the European Communities; 1995.

Viet, J. and Georges van Slype. **EUDISED Multilingual Thesaurus for information processing in the field of education,** English version. 307 pages. Mouton Publishers, Berlin,, New York, Amsterdam; 1984.

**EUDISED R&D Bulletin,** volume 45, ISSN 0378-7192;. 127 pages.; K.G. Saur, Munich; 1993.

Food and Agriculture Organization of the United States. **AGROVOC multilingual agricultural Thesaurus.** Second edition, English version; 798 pages; APIMONDIA, Rome; 1992. (Not latest)

International Atomic Energy Agency. **INIS: Thesaurus.** 887 p. and **INIS multilingual dictionary.** 314 p. IAEA, Vienna; 1993, 1983 (not latest editions).

Organization for Economic Cooperation and Development. **Multilingual dictionary of fish and fish products.** Fourth edition., 352 pages; Fishing News Books, Cambridge; 1995. LCC Q164.7.M84.1995

Centre for Computer-Aided Egyptological Research. **Multilingual Egyptological Thesaurus**.  
<http://www.ccer.ggl.ruu.nl/thes/thsaur.html>. 1995.





# **Knowledge Organization Systems in Digital Libraries**

## **Tutorial**

### **Part 2**

#### **Design, evaluation, and development**



## Part 2. Design and development. Outline

<b>14:30 - 14:35</b>	<b>Introduction and overview</b>	160
	KOS development principles	162
<b>14:35 - 15:05</b>	<b>The process of KOS development</b>	164
13:35 - 13:40	<b>The overall process of KOS development</b>	165
13:40 - 13:55	<b>Sources of concepts, terms, relationships, definitions</b>	166
	<b>Methods of data collection</b>	
13:55 - 14:05	<b>Merging data from many sources</b>	171
<b>15:05 - 16:00</b>	<b>Developing the conceptual structure</b>	173
14:05 - 14:30	<b>Facet analysis 1: Education (starting with classes from DDC)</b>	174
14:30 - 14:40	<b>More facet examples: Job titles</b>	176
14:40 - 14:50	<b>Guidelines for meaningful arrangement</b>	179
14:50 - 15:00	<b>Rules for selection of preferred terms</b>	187
	<b>Rules for selection of concepts as descriptors</b>	
<b>16:00 - 16:30</b>	<b>Break</b>	
<b>16:30 - 17:15</b>	<b>Developing the conceptual structure, continued</b>	
16:30 - 17:15	<b>Facet exercise (in pairs)</b>	177
<b>17:15 - 18:15</b>	<b>The structure and processing of KOS data</b>	189
17:15 - 17:30	<b>Interoperability of thesauri/ontologies. Crosswalks</b>	189
17:30 - 17:45	<b>The structure of a KOS/ontology database</b>	193
17:45 - 17:55	<b>The many forms of Knowledge Organization Systems (KOS)</b>	203
	and their standards	
	SKOS (Simplified Knowledge Organization Systems) example	217
	OWL (Web Ontology Language) example	223
17:55 - 18:15	<b>KOS software and its evaluation</b>	209
	<b>Appendix. The KOS development process</b>	257



## **Introduction**

Developing a good Knowledge Organization System (KOS) – ontology, taxonomy, classification, thesaurus – for a given domain means to

- evolve the conceptual structure of the domain;
- discern and crystallize its key concepts and their relationships; and then
- represent this structure (1) externally in a form that readers can easily grasp and make their own and (2) internally in a form that computer programs can use.

This is an intellectually challenging task that requires considerable resources.

# **Introduction and overview**

## **Scope:**

**“Thesaurus / KOS” or just KOS is used for Knowledge Organization Systems (KOS) Includes Thesauri, classifications, ontologies, taxonomies, concept maps, dictionaries, etc.**

## **Main objective:**

**Participants should be able to crystallize the conceptual structure of a domain**

# **Outline**

**KOS development principles**

**The process of KOS development**

## **Developing the conceptual structure**

### **The structure and processing of KOS data**

## **KOS development principles**

- (1) Determine the right organizational scope or user group(s) to be served**
- (2) Determine the right function scope**
- (3) Ensure adoption by supporting many views through an inclusive, flexible KOS database**
- (4) Build on existing KOS. Reuse, reuse, reuse**
- (5) Use automation for efficiency**
- (6) Capitalize on collaborative creation and editing**

# KOS development principles

Good cost-effective KOS development stands and falls with the observation of six simple principles.

**(1) Determine the right organizational scope or user group(s)** to be served. A broad organizational scope results in higher return on investment but must be balanced against tailoring the KOS to specific requirements (but see principle (3)). The scope can range from a KOS for a single user who needs to organize a collection of documents, Web sites, files, email messages, notes, contacts, appointments, and tasks to a KOS that is used worldwide for organizing many databases in many languages, such as AGROVOC or the Medical Subject Headings. In between would be a KOS for an organization, an online community, or a specific database such as AERS, FDA's Adverse Events Reporting System database ([fda.gov/cder/aers/default.htm](http://fda.gov/cder/aers/default.htm)). A broad scope of application (many departments in an organization, several online communities with related purposes, databases in the same general subject domain) improves the return on investment into the KOS and fosters semantic interoperability. .

**(2) Determine the right function scope.** The more functions are served by a KOS the higher the return on investment. Therefore it is important to identify all possible applications that could profit from KOS support. This includes functions such as retrieval from any kind of database or through a Web search engine and structured output of results, project planning, insurance billing, natural language processing, reasoning in expert systems; see XXX for an extensive list. Understanding all these functions is crucial if one wants to maximize ROI. For each application state

- requirements and the spectrum of users' level of knowledge.
- required characteristics of the KOS

**(3) Ensure adoption by supporting many views through an inclusive, flexible KOS database.** Different applications need different views for intrinsic reasons that are rooted in the nature of the application (for example, to optimize reasoning over large knowledge bases) or for historical reasons (a given user group may not want to change what they are used to, or a large body of material can not feasibly be re-indexed). The solution is a flexible comprehensive KOS database in which different views can coexist or from which different views can be easily extracted; that way each application's or group's view of the domain can be accommodated, and that removes a key barrier to engaging several groups that might otherwise go their own ways (at great expense), even though all groups may agree on 80% and disagree on just 20% of concepts, terms, and relationships. There is no need to force everybody into the same mold; many views can coexist within the same KOS database, where they can be related to each other to the extent possible, capturing commonalities and explicating differences. Through the KOS database, different groups can learn from each other; this may lead to improvements of each group's view. In that sense, a KOS database

- records present usage by different groups (like a dictionary records usage);
- creates an overarching well-ordered structure for all the concepts from different views;
- through that structure, creates relationships between different views;
- through that structure, recommends improvements for each group's view..

**(4) Build on existing KOS. Reuse, reuse, reuse.** There is enormous intellectual capital in existing KOS; use it! Identify the KOS presently used, explicitly or implicitly, in the organizational and functional scope. Find other KOS that could meet some or all of the requirements, perhaps with some modification, or that would be useful sources for developing a new KOS (see Box 1. Searching for KOS).

**(5) Use automation for efficiency.** Automation can be used to massage data from existing KOS for the purpose at hand. It can also be used to extract terms, concepts, relationships, and global structure from texts and document collections. While automation cannot entirely replace intellectual effort, a KOS constructed using automatic processing alone is better than no KOS at all.

**(6) Capitalize on collaborative creation and editing** with central expert support and some measure of control. Possibly have a cadre of KOS editors throughout the organization or throughout the world.



# **The process of KOS development**

**The overall process of KOS development**

**Sources of concepts, terms, relationships,  
definitions**

**Methods of data collection**

**Merging data from many sources**

# The overall process of KOS development

Diagram from DS 1974 copied in here. Need orig

## Sources of concepts, terms, relationships, definitions

Reuse knowledge in existing Knowledge Organization Systems. Much intellectual capital was invested in their development

But: Adapt content and structure to user requirements and background.

Most important source: search requests and other statements of user requirements.

### Types of sources

- (1) **P rearranged sources** (terms are already arranged according to some principle)
- (2) **Open-ended sources** (terms are not ordered or terms must be inferred or derived)

Find machine-readable sources

Internal and external sources

## **Sources of concepts, terms, relationships, definitions**

- (1) **Prearranged sources** (terms are already arranged according to some principle)
  - (1.1) Descriptor lists, classification schemes, thesauri (universal classification schemes, such as LCC or UDC, and special classification schemes).
  - (1.2) Nomenclatures of single disciplines, esp. if approved by an international body.
  - (1.3) Treatises on the terminology of a subject field
  - (1.4) Encyclopedias, lexica, dictionaries, glossaries (universal or discipline-oriented; mono-, bi-, or multilingual).
  - (1.5) The tables of contents and indexes of conference proceedings, textbooks, handbooks, and course syllabi.
  - (1.6) Indexes of journals, abstracting journals, other publications, databases.
  - (1.7) Term-association lists produced by subjects in term association studies.
  - (1.8) Output from automatic classification programs based on term co-occurrence data or citations.

## **Sources of concepts, terms, relationships, definitions**

- (2) **Open-ended sources** (terms are not ordered or terms must be inferred or derived)
  - (2.1) Lists of search requests and interest profiles and other statements of user requirements obtained from search logs and user studies (individual interviews, focus groups).
    - (2.1a) Mooers' method: Focus group, present documents, ask "Why would this be of interest?"
  - (2.2) Descriptions of R&D projects and other activities to be supported.
  - (2.3) Free indexing of a sample of documents, each by several experts (to get synonyms).
  - (2.4) Titles, abstracts, full text, reviews of books, journal articles, conference papers, Web sites, internal documents, etc.
  - (2.5) For more information on individual terms: Web searches

# **Methods of data collection**

## **For prearranged sources**

**If machine-readable, include all information, can always delete later**

**If not machine-readable and highly relevant, scan or have transcribed**

**Otherwise go through and select**

**May need to reformat for input to thesaurus / KOS software; use Perl scripts or word processor macros**

# Methods of data collection

## For open-ended sources

Extract terms and **phrases** automatically, using a large general phrase dictionary, syntactic analysis, or a system such as <http://www.nzdl.org/Kea/>

Possibly use frequency data for further selection.

Extract term relationship automatically (often a feature of text mining programs).

Extract terms manually, being on the look-out for term relationships that can be inferred from text.

# **Merging data from many sources**

## **Merge terms**

Need to consolidate term variants

**Use broad-scope sources to get more information on terms collected**

## **Assemble synonym sets / concepts**

Use ST relationships from many sources

Source 1: elderly ST aged person,

Source 2: aged person ST senior citizen

## **Merge relationships**

Need to consider that often the same conceptual relationship is expressed in different terms

**“Afterburn” collection from specialized sources to fill gaps**



# **Developing the conceptual structure**

## **Facet analysis 1: Education**

### **More facet examples:**

**Yahoo Education (from Part 1)**

**Job titles**

### **Facet exercise (in pairs)**

### **Principles for meaningful arrangement**

**Rules for selection of concepts as  
descriptors. Rules for selection of terms**

# Facet analysis

**Education** (starting with classes from DDC)

## Conceptual analysis and synthesis

in three steps:

**Step 1. Semantic factor compound concepts, make a list of elemental concepts.**

**Step 2. Arrange elemental concepts into facets.**

**Arrange each facet in a well-structured hierarchy.**

**Step 3. If needed, fit compound concepts into the framework of the hierarchy.**

## Concept list for conceptual analysis and synthesis

(from Dewey Decimal Classification)

**Note:** A broader class is given in ( ), if necessary to specify the meaning of a term.

372.19	Curriculums of elementary schools
372.35043	Science in the elementary school curriculum
372.414	Methods of instruction for reading in elementary schools
372.72043	Arithmetic in the elementary school curriculum
373.19	Curriculums in secondary schools
373.243	Military schools (Secondary Education)
376	Education of women
376.63	Secondary education of women
378.19	Curriculum of colleges and universities
378.33	Fellowships (Higher Education)
371.911	Blind and partially sighted students
371.912	Deaf and hard-of-hearing students
371.95	Curriculums for gifted students

## **More facet examples**

**Job titles. Can you spot the facets?**

**Lawyer**

**Paralegal**

**Law office receptionist**

**Librarian**

**Library assistant**

**Library clerk**

**Physician**

**Physician's assistant**

**Doctor's office clerk**

**Ophthalmologist (eye doctor)**

**Ophthalmologic technician**

**Surgeon**

## **Facet exercise (in pairs?)**

### **Yahoo Health**

**Arrange the terms in front of you into meaningful groupings.**

**Use the blank strips to write a heading for each group.**

**Time: 30 minutes** (leaving 10 minutes for discussion)

Note: Back page has a list of the terms

**Alternative Medicine**

**Business to Business**

**Chats and Forums**

**Children's Health**

**Conferences**

**Death and Dying**

**Dentistry**

**Disabilities**

**Diseases and Conditions**

**Education**

**Emergency Services**

**Employment**

**Environmental Health**

**First Aid**

**Fitness**

**General Health**

**Health Administration**

**Health Care**

**Health Sciences**

**Hospitals and Medical Centers**

**Institutes**

**Law**

**Long Term Care**

**Medicine**

**Yahoo health**

**Strips to cut for term sorting  
exercise**

**Men's Health**

**Mental Health**

**Midwifery**

**News and Media**

**Nursing**

**Nutrition**

**Organizations**

**Pet Health**

**Pharmacy**

**Procedures and Therapies**

**Public Health and Safety**

**Reference**

**Reproductive Health**

**Senior Health**

**Sexuality**

**Shopping and Services**

**Teen Health**

**Traditional Medicine**

**Travel Health and Medicine**

**Web Directories**

**Weight Issues**

**Women's Health**

**Workplace**

•

-

## **Principles for meaningful arrangement**

Sequence and two-dimensional graphical arrangements (concept maps) can convey important information about concept relationships.

Collocate closely related concepts.

Often a principle of arrangement intrinsic to the subject matter suggests itself. The following examples and guidelines are intended to sharpen “informed intuition”.

# Meaningful arrangement

## Example 1

*<size: photograph formats>*

double whole plate

half plate

mammoth plate

ninth plate

quarter plate

sixteenth plate

sixth plate

whole plate

**Art and Architecture**

**Thesaurus**

**size: photograph formats**

sixteenth plate

ninth plate

sixth plate

quarter plate

half plate

whole plate

double whole plate

mammoth plate

**Suggested meaningful**

**sequence**

**Alphabetical vs. meaningful sequence on same**

**hierarchical level**



## Meaningful arrangement

### Example 2. **Body systems. Fuller version**

#### **XF body system or organ**

- XG . musculoskeletal system
- XH . skin system
- XJ . cardiovascular system
- XK . respiratory system
- XL . mouth, larynx, vocal organ
- XM . digestive system
- XN . urogenital system
- XP . . urinary system
- XQ . . reproductive system
- XR . blood, immune system
- XS . . blood
- XT . . immune system
- XU . endocrine system
- XV . sensory system
- XW . nervous system
- XX . . nervous system structures and components
- XY . . peripheral nervous system
- XZ . . central nervous system

# Meaningful arrangement

## Example 3. Art genres

Trying to find a meaningful arrangement for a list of concepts often reveals a facet structure.

See the example in the tutorial notebook.

## Graphical arrangement: Concept maps

See the examples in the tutorial notebook.

<art genres>

academic art  
amateur art  
apocalyptic art  
art brut  
children's art  
commercial art  
community art  
SN Includes art undertaken in conjunction with particular communities, often socially deprived, usually with the idea of producing an effect or inspiring response specifically within those communities, with no reference to widely established standards. For art intended to beautify or enrich public places, use **public art**.  
computer art  
court art  
crafts  
cybernetic art  
didactic art  
dissident art  
ethnic art  
fantastic art  
figurative art  
folk art  
funerary art  
naive art  
nonrepresentational art  
primitive art  
public art  
SN Use for art whose purpose is to beautify and enrich public places. For art undertaken in conjunction with particular communities, usually to produce an effect or inspire response specifically within those communities, use **community art**.  
rock art  
cave art  
serial art  
sofa art  
street art

a. Original alphabetical sequence

art genres

**art genres by content or other intrinsic characteristics**

figurative art  
fantastic art  
apocalyptic art  
nonrepresentational art  
cybernetic art  
serial art  
crafts

**art genres by standard**

academic art  
folk art  
dissident art

**art genres by type of artist or origin**

amateur art  
naive art  
art brut  
children's art  
computer art  
ethnic art  
primitive art

**art genres by audience, purpose, or display context**

sofa art  
court art  
public art  
SN Art whose purpose is to beautify and enrich public places.  
community art  
SN Public art undertaken in conjunction with particular communities, often socially deprived, usually with the idea of producing an effect or inspiring response specifically within those communities, with no reference to widely established standards.  
street art  
rock art  
cave art [prehistoric, esp. paleolithic]  
didactic art  
commercial art  
funerary art

b. Suggested meaningful sequence

Figure 3. Example from the Art and Architecture Thesaurus

## Concept map PHD

## Concept map instr design

# Meaningful arrangement

## Guidelines

### “Natural” principles

- (1) Chronological – e.g., historical events.
- (2) Evolutionary – arrange entities in the order they evolved, e.g., biological species, ideas.
- (3) Sequence of steps – e.g., production processes, research methods, sequence of logical steps
- (4) Increasing extension
- (5) Geographical – spatial proximity.

### More conceptual principles

- (6) Increasing complexity (integrative levels)
- (7a) From abstract to concrete or vice versa
- (7b) From general to specific
- (7c) From universal to local
- (8) Canonical – an order given by an authority, e.g., books of a holy scripture
- (9) Consistency of comparable subdivisions that appear in two or more different places
- (10) Importance for indexing and query formulation

## **Rules for selection of preferred terms from a group of synonyms**

**Include in the thesaurus / KOS any term that falls in scope.**

**A large lead-in vocabulary is good!**

**Then select preferred terms.**

**The preferred term should**

- be the best to reflect the meaning of the concept;
- be recognized in the user community;
- be unambiguous;
- be simple and short in spelling.

These criteria may conflict

Frequency data and occurrence in authentic sources can help in the selection.

# Rules for selection of concepts as descriptors

The following criteria are helpful:

- Usefulness for searching and other functions;
- Are there alternative solutions:  
use a combination of descriptors,  
use a broader descriptor,  
consolidate with another concept to form  
a broader concept;
- Logical structure: is the concept needed  
as a heading?
- Frequency in indexing.



# Interoperability of thesauri/ontologies.

## Crosswalks

### Primary question:

- take a query formulated in vocabulary A,
- map the descriptors to vocabulary B,
- how good is the search in B as compared to using a query formulated in vocabulary B directly?

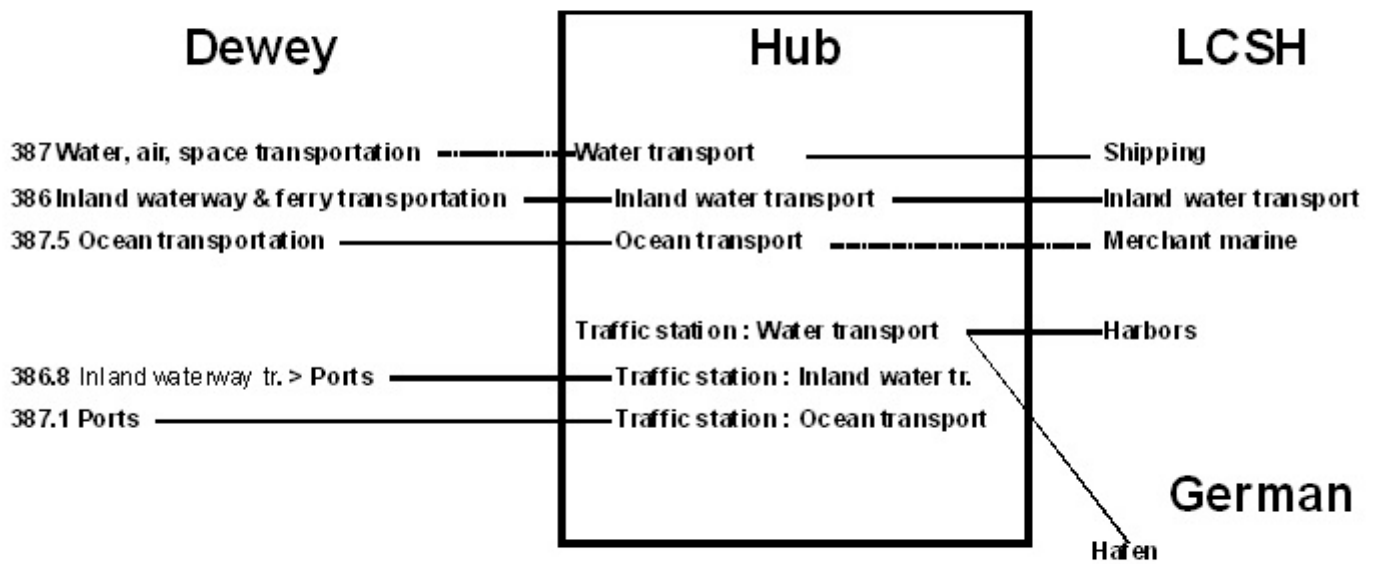
The answer determines searching compatibility.

**Searching compatibility is directional**, complex, and depends on the individual descriptors used.

<b>Vocabulary A</b>	<b>Vocabulary B</b>
Aircraft	Aircraft Airplane Helicopter
Military aircraft	Aircraft AND Military
Pest control Pesticides	Pest control (no narrower terms)

Insert index language page here

## Mapping through a Hub



Insert compat figures here

# The structure of a KOS database

Thesaurus / KOS data are relational.

**Relational database** is the most natural structure.

**Many types of relationships** – structure should not be restrictive.

(See sample list in notebook.)

Examples of Synonymous-Term-type relationships

ST	Synonymous Term
ET	Equivalent Term
SP	Spelling Variant
AB	Abbreviation
FT	Full Term

Structure should allow for a **relationship to be the object of another relationship**

(for example, a scope note explaining the relationship)

## Relationship strength

Appendix 2. Relationship types presently recognized by TermMaster

Note: This list is extensible by simply updating a table in the program and recompiling

<b>Sym bol</b>	<b>Meaning</b>	<b>Reci- procal</b>	<b>Group</b>	<b>Reference to</b>
FN	Full form Note (If full form of term > 61 char)	-	SN	Text
<b>SN</b>	<b>Scope Note</b>	-	SN	Text
QN	Qualifier Note	-	SN	Text
HN	History Note	-	SN	Text
<b>IN</b>	<b>Internal Note</b> Expands on the external scope note, esp. reasons for term inclusion, term placement, and other decisions.	-	SN	Text
<b>AN</b>	<b>Action Note</b> Notes on actions to be taken on the term, such as look up definition, ask Ms. X, etc.		SN	Text
SQ	Source (for additional subset record)	-	SN	
<b>SR</b>	<b>Detailed source</b>	-	SN	Text
SI	Synonym, Internal	SI	ST	Term
SH	From non-hyphenated to hyphenated	SG	ST	Term
SG	From hyphenated to non-hyphenated	SH	ST	Term
<b>SP</b>	<b>Spelling variant</b>	SP	ST	Term
SB	Spelling British	SA	ST	Term
SA	Spelling American	SB	ST	Term
<b>AB</b>	<b>Abbreviation</b>	FT	ST	Term
<b>FT</b>	<b>Full Term</b>	AB	ST	Term
<b>ST</b>	<b>Synonymous Term</b>	ST	ST	Term
<b>ET</b>	<b>Equivalent Term</b>	ET	ST	Term
TR	Translation	TR	ST	Term
NA	Narrower of Facet	FA	NT	Term
NX	Narrower term of a broad category used in preliminary sorting	BX	NT	Term
NF	Narrower term - compound containing factor	BF	NT	Term
NM	Narrower Term - compound containing Modifier	BM	NT	Term
NC	Reciprocal of BC	BC	NT	Term
<b>NT</b>	<b>Narrower Term</b>	BT	NT	Term
NG	Narrower term - Generic	BG	NT	Term
NTT	Narrower term - Token	BTT	NT	Term
NPT	Narrower term - Partitive	BP	NT	Term
FA	Facet	NA	BT	Term

BX	Broader term for preliminary. sorting	NX	BT	Term
BF	Broader term - Factor	NF	BT	Term
BM	Broader term - Modifier	NM	BT	Term
BC	Broader term that might have NT to be used in combination	NC	BT	Term
<b>BT</b>	<b>Broader Term</b>	NT	BT	Term
BG	Broader term - Generic	NG	BT	Term
BTT	Broader term - Type of token	NTT	BT	Term
BPT	Broader term - Partitive	NPT	BT	Term
RC	Related term for combination (pop-up menu showing terms to use)	RD	RT	Term
RD	Inverse of RC	RC	RT	Term
RG	One-directional related term	RH	RT	Term
RH	Inverse of RG	RG	RT	Term
RN	Related term in scope note, generated by the program	RO	RT	Term
RO	Inverse of RN	RN	RT	Term
<b>RT</b>	<b>Related Term</b>	RT	RT	Term
EX	Excludes	EF	EX	Term
EF	Excluded From	EX	EX	Term
UN	Unspecified relationship	UN	RT	Term
HT	Homonymous Term	HB	HT	Term
HF	Homonym From	HT	HT	Term
ME	Meaning Equivalent	MF	HT	Term
MF	Meaning equivalent From	ME	HT	Term
BW	Broader Word	NW	RT	Term
NW	Narrower Word	BW	RT	Term
AF	Affects	AY	AF	Term
AY	Affected by	AF	AF	Term
PC	Precursor	PB	AF	Term
PB	Produced by	PC	AF	Term
RW	reacts with	RW	AF	Term
IB		IB	ID	
//	From a relationship to a term. Internal symbol TH	/<		Relation
		TI		
/<	Inverse of //	//		Relation
	Internal symbol TI	TH		





# The structure of a KOS database

## Three levels

### Level 1: Link term variants to terms

AST FT aspartate aminotransferase

GOT FT glutamate oxaloacetate  
transaminase

(FT Full Term)

### Level 2: Link terms to concepts

aspartate aminotransferase

ST glutamate oxaloacetate  
transaminase

### Level 3: Relate concepts to concepts

aspartate aminotransferase

BT aminotransferases

**Levels 1 and 2 are often confounded.**

# The structure of a KOS database

## Two models

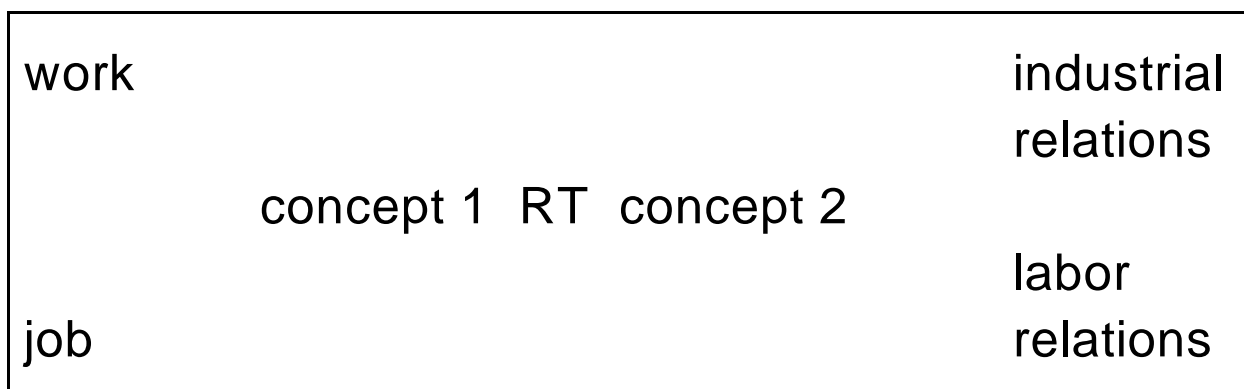
### Concept-based model

Terms are mapped to concepts. This mapping expresses Synonymous Term relationships.

Concept relationships are expressed using concept identifiers.

Elegant, but in a multi-thesaurus / KOS database requires universal commitment to the term-concept mapping.

UMLS uses this model



# The structure of a KOS database

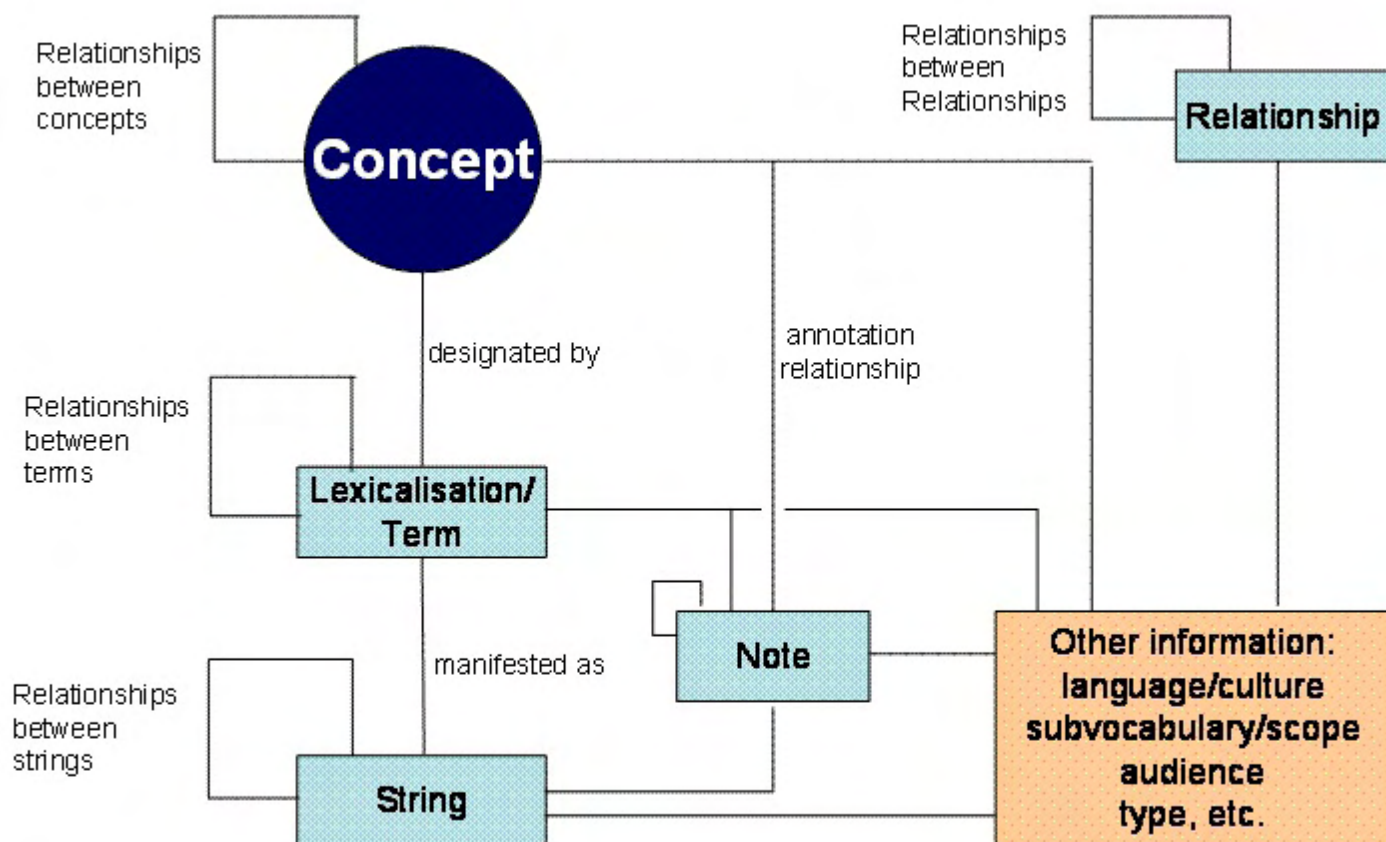
## Term-based model

All relationships are expressed as relationships between terms.

A concept relationship may be expressed in many ways, using different synonyms for each concept

Requires extensive processing to discover all concept relationships starting from a given concept.

job ST work	job RT industrial relations
industrial relations ST labor relations	work RT industrial relations
	job RT labor relations
	work RT labor relations



**Figure 1: Conceptual model for combining thesauri and ontologies**

page from JODI

page from JODI

# Standards 1



























## Standards 13



## SKOS (Simplified Knowledge Organization System) example. Sample records from AGROVOC

This example illustrates some of the core features of SKOS, a W3C standard  
[www.w3.org/2004/02/skos/](http://www.w3.org/2004/02/skos/).

Most records are very abbreviated.

The concepts of this example form the following polyhierarchy:

### Processed products (#c\_15472)

- . Processed animal products (#c\_29107)
  - . . Milk products (#c\_4830)
    - . . . Cultured milk (#c\_2023)
      - . . . . Acidophilus milk (#c\_8602)
    - . . . Milk protein (#c\_4831)
  - . Fermented products (#c\_15734)
    - . . Cultured milk (#c\_2023)
      - . . . Acidophilus milk (#c\_8602)
  - . Protein products (#c\_6256)
    - . . Animal protein (#c\_439)
      - . . . Milk protein (#c\_4831)

### Chemophysical properties (#c\_1521)

- . Acidity (#c\_8601)

## SKOS File

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
>

  <skos:ConceptScheme rdf:about="http://www.fao.org/aims/aos/agrovoc">
    <dc:title>AGROVOC</dc:title>
    <dc:description>FAO Multilingual Thesaurus.</dc:description>
    <dc:creator>
      <foaf:Organization>
        <foaf:name>FAO</foaf:name>
      </foaf:Organization>
    </dc:creator>
    <dc:rights>Free to all for non commercial use.</dc:rights>
    <dcterms:issued>Tue Apr 29 09:49:31 CEST 2008</dcterms:issued>
    <dcterms:modified>2008-04-14 10:17:36.0</dcterms:modified>
  </skos:ConceptScheme>
```

```

<!-- ===== -->
<!-- Complete AGROVOC concept record with all languages (but does not have a scope note) -->
<!-- ===== -->

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#c_3">
  <skos:prefLabel xml:lang="en">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="es">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="ar">ØçØ"Ø§</skos:prefLabel>
  <skos:prefLabel xml:lang="zh">è,,±è• ½é... </skos:prefLabel>
  <skos:prefLabel xml:lang="pt">Aba</skos:prefLabel>
  <skos:prefLabel xml:lang="th">à¹Ē,à,šà,µà¹Ē, </skos:prefLabel>
  <skos:prefLabel xml:lang="ja">ã,çãf-ã,·ã,ãf³é... </skos:prefLabel>
  <skos:prefLabel xml:lang="sk">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="de">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="hu">Aba</skos:prefLabel>
  <skos:prefLabel xml:lang="pl">Aba</skos:prefLabel>
  <skos:prefLabel xml:lang="fa">Øç.Ø"ÛĒ.Øç </skos:prefLabel>
  <skos:prefLabel xml:lang="it">ABA</skos:prefLabel>
  <skos:prefLabel xml:lang="hi">à¤• o à¤-à¥Ē à¤• o</skos:prefLabel>
  <skos:altLabel xml:lang="cs">kyselina abscisovÃĵ</skos:altLabel>
  <skos:altLabel xml:lang="de">ABSCISINSAEURE</skos:altLabel>
  <skos:altLabel xml:lang="en">Abscisic acid</skos:altLabel>
  <skos:altLabel xml:lang="es">Ã• cido abscÃsico</skos:altLabel>
  <skos:altLabel xml:lang="fa">ØçØ"Ø³ÛŠØ³ÛŠÛf Ø§Ø³ÛŠØ- </skos:altLabel>
  <skos:altLabel xml:lang="fr">Acide abscissique</skos:altLabel>
  <skos:altLabel xml:lang="hi">à¤• à¤-à¤,à¤;à¤,à¤;à¤• à¤... à¤@à¥• à¤²</skos:altLabel>
  <skos:altLabel xml:lang="hu">abszcizinsav</skos:altLabel>
  <skos:altLabel xml:lang="it">Acido abscissico</skos:altLabel>
  <skos:altLabel xml:lang="ja">ã,çãf-ã,·ã,ãf³é... </skos:altLabel>
  <skos:altLabel xml:lang="pl">Kwas abscysynowy</skos:altLabel>
  <skos:altLabel xml:lang="pt">Ã• cido abscÃsico</skos:altLabel>
  <skos:altLabel xml:lang="sk">kyselina abscisovÃĵ</skos:altLabel>
  <skos:altLabel xml:lang="th">à,• à,£à,"à¹• à,à,šà¹,à,à,à,à,à,, </skos:altLabel>
  <skos:altLabel xml:lang="zh">ABA</skos:altLabel>
  <skos:changeNote rdf:parseType="Resource">
    <rdf:value>The last modification for this concept was for the term in CS</rdf:value>
    <dc:creator>
      <foaf:Organization>
        <foaf:name>FAO/GILW</foaf:name>
      </foaf:Organization>
    </dc:creator>
    <dc:date>2008-03-05 00:00:00.0</dc:date>
    <dc:language>
      <dcterms:RFC1766>
        <rdf:value>EN</rdf:value>
        <rdfs:label>English</rdfs:label>
      </dcterms:RFC1766>
    </dc:language>
  </skos:changeNote>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_3397"/>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_32543"/>
</skos:Concept>

```

```
<!-- ===== -->
```

```
<!-- In the following records, change notes and prefLabel and altLabel for most languages are omitted
For most records, everything but the English prefLabel and the concept relationships is omitted-->
```

```
<!-- ===== -->
```

```
<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_439">
  <skos:prefLabel xml:lang="en">Animal protein</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Protéine animale</skos:prefLabel>
  <skos:prefLabel xml:lang="es">Proteínas de origen animal</skos:prefLabel>
  <skos:prefLabel xml:lang="de">TIERISCHES PROTEIN</skos:prefLabel>
  <skos:altLabel xml:lang="fr">Protéine d'origine animale</skos:altLabel>
  <skos:scopeNote xml:lang="de">Proteinprodukte tierischer Herkunft; fuer die chemischen Verbindungen
  PROTEIN (6259) benutzen</skos:scopeNote>
  <skos:scopeNote xml:lang="en">Protein products of animal origin; for chemical compounds use Proteins
  (6259)</skos:scopeNote>
  <skos:scopeNote xml:lang="es">Productos proteínicos de origen animal; para compuestos químicos use
  Proteínas (6259)</skos:scopeNote>
  <skos:scopeNote xml:lang="fr">Produits protéiques d'origine animale; pour les composés chimiques
  utiliser Protéine (6259)</skos:scopeNote>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_6256"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4831"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_3209"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_2931"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_28528"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_968"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4669"/>
</skos:Concept>

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_1521">
  <skos:prefLabel xml:lang="en">Chemicophysical properties</skos:prefLabel>
  <skos:scopeNote xml:lang="de">Fuer den Boden PHYSIKOCHEMISCHE BODENEIGENSCHAFT (7182)
  benutzen</skos:scopeNote>
  <skos:scopeNote xml:lang="en">For soil use Soil chemicophysical properties (7182)</skos:scopeNote>
  <skos:scopeNote xml:lang="es">Para el suelo use Propiedades físico - químicas suelo
  (7182)</skos:scopeNote>
  <skos:scopeNote xml:lang="fr">Pour le sol utiliser Propriétés physicochimique du sol
  (7182)</skos:scopeNote>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_36930"/>
  ...
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_8601"/>
</skos:Concept>

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_2023">
  <skos:prefLabel xml:lang="en">Cultured milk</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Lait fermenté</skos:prefLabel>
  <skos:prefLabel xml:lang="es">Leche fermentada</skos:prefLabel>
  <skos:prefLabel xml:lang="de">SAUERMILCHPRODUKT</skos:prefLabel>
  <skos:altLabel xml:lang="de">FERMENTIERTE MILCH</skos:altLabel>
  <skos:altLabel xml:lang="en">Fermented milk</skos:altLabel>
  <skos:altLabel xml:lang="pl">Mleko kwaszone</skos:altLabel>
  <skos:altLabel xml:lang="en">Lactic beverages</skos:altLabel>
  <skos:altLabel xml:lang="es">Bebidas lácticas</skos:altLabel>
  <skos:altLabel xml:lang="pl">Napoje mleczne fermentowane</skos:altLabel>
  <skos:altLabel xml:lang="de">SAUERMILCH</skos:altLabel>
  <skos:altLabel xml:lang="en">Sour milk</skos:altLabel>
  <skos:altLabel xml:lang="es">Leche agria</skos:altLabel>
  <skos:altLabel xml:lang="pl">Mleko kwaśne</skos:altLabel>
```

```

<skos:altLabel xml:lang="de">SKYR</skos:altLabel>
<skos:altLabel xml:lang="en">Skyr</skos:altLabel>
<skos:altLabel xml:lang="fr">Skyr</skos:altLabel>
<skos:altLabel xml:lang="pl">Skyr</skos:altLabel>
<skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4830"/>
<skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_15734"/>
<skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_23849"/>
<skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_15463"/>
<skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_8602"/>
<skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_28325"/>
<skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_32527"/>
</skos:Concept>

```

```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_4830">
  <skos:prefLabel xml:lang="en">Milk products</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_29107"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4831"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_7096"/>
  ...
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_1951"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_2023"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_2384"/>
  ...
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4146"/>
</skos:Concept>

```

```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_4831">
  <skos:prefLabel xml:lang="en">Milk protein</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_439"/>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_4830"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_28323"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_1357"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_28528"/>
</skos:Concept>

```

```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_6256">
  <skos:prefLabel xml:lang="en">Protein products</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_15742"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_8172"/>
  ...
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_439"/>
</skos:Concept>

```

```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_8601">
  <skos:prefLabel xml:lang="en">Acidity</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_1521"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_87"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_15734"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_5753"/>
</skos:Concept>

```

```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#_c_8602">
  <skos:prefLabel xml:lang="en">Acidophilus milk</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_2023"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#_c_23938"/>
</skos:Concept>

```



```

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#c_15734">
  <skos:prefLabel xml:lang="en">Fermented products</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_15742"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_248"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_2023"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_7060"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_27521"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_2855"/>
  <skos:related rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_8601"/>
</skos:Concept>

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#c_15742">
  <skos:prefLabel xml:lang="en">Processed products</skos:prefLabel>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_28228"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_28558"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_29107"/>
  ...
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_6256"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_7697"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_15734"/>
  ...
</skos:Concept>

<skos:Concept rdf:about="http://www.fao.org/aims/aos/agrovoc#c_29107">
  <skos:prefLabel xml:lang="en">Processed animal products</skos:prefLabel>
  <skos:broader rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_15742"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_35866"/>
  <skos:narrower rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_4830"/>
  ...
</skos:Concept>

<!-- ===== -->
<!-- Concept Schemes and Top Concepts: -->
<!-- ===== -->
<skos:ConceptScheme rdf:about="http://www.fao.org/aims/aos/asc#c_A">
  <rdfs:label>Agriculture</rdfs:label>
  <skos:hasTopConcept rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_199" />
  ...
  <skos:hasTopConcept rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_15742" />
  <skos:hasTopConcept rdf:resource="http://www.fao.org/aims/aos/agrovoc#c_1521" />
  ...
</skos:ConceptScheme>

</rdf:RDF>

```



## OWL (Web Ontology Language) example A simple Pizza ontology

This example illustrates some of the core features of OWL, a W3C standard.

[www.w3.org/TR/owl-ref/](http://www.w3.org/TR/owl-ref/)

It is fairly self-explanatory.

Tutorials

[www.co-ode.org/resources/tutorials/protege-owl-tutorial.php](http://www.co-ode.org/resources/tutorials/protege-owl-tutorial.php),

[www.cs.man.ac.uk/~horrocks/ISWC2003/Tutorial/](http://www.cs.man.ac.uk/~horrocks/ISWC2003/Tutorial/)

<!-- Preliminaries-->

```
<?xml version="1.0"?>
```

```
<!DOCTYPE rdf:RDF [
<!ENTITY owl "http://www.w3.org/2002/07/owl#" >
<!ENTITY owl1 "http://www.w3.org/2006/12/owl11#" >
<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
<!ENTITY owl11xml "http://www.w3.org/2006/12/owl111-xml#" >
<!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
<!ENTITY pizza "http://www.pizza.com/ontologies/pizza.owl#" >
<!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>
```

```
<rdf:RDF
```

```
  xmlns="http://www.pizza.com/ontologies/pizza.owl#"
  xml:base="http://www.pizza.com/ontologies/pizza.owl"
  xmlns:owl11="http://www.w3.org/2006/12/owl11#"
  xmlns:pizza="http://www.pizza.com/ontologies/pizza.owl#"
  xmlns:owl11xml="http://www.w3.org/2006/12/owl111-xml#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
  <owl:Ontology rdf:about="">
    <rdfs:comment>Protege tutorial for 775</rdfs:comment>
  </owl:Ontology>
```

<!-- Object Properties -->

```
<!--www.pizza.com/ontologies/pizza.owl#hasIngredient
```

```
Full URL of defined object. Given here as an example -->
```

```
<owl:ObjectProperty rdf:about="#hasIngredient">
  <rdf:type rdf:resource="#owl:TransitiveProperty"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#hasBase">
  <rdfs:subPropertyOf rdf:resource="#hasIngredient"/>
  <rdfs:domain rdf:resource="#Pizza"/>
  <rdfs:range rdf:resource="#PizzaBase"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#hasTopping">
  <rdfs:subPropertyOf rdf:resource="#hasIngredient"/>
  <rdfs:domain rdf:resource="#Pizza"/>
  <rdfs:range rdf:resource="#PizzaTopping"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#isIngredientOf">
  <rdf:type rdf:resource="&owl;TransitiveProperty"/>
  <owl:inverseOf rdf:resource="#hasIngredient"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#isBaseOf">
  <rdfs:subPropertyOf rdf:resource="#isIngredientOf"/>
  <owl:inverseOf rdf:resource="#hasBase"/>
  <rdfs:domain rdf:resource="#PizzaBase"/>
  <rdfs:range rdf:resource="#Pizza"/>
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#isToppingOf">
  <rdfs:subPropertyOf rdf:resource="#isIngredientOf"/>
  <owl:inverseOf rdf:resource="#hasTopping"/>
  <rdfs:domain rdf:resource="#PizzaTopping"/>
  <rdfs:range rdf:resource="#Pizza"/>
</owl:ObjectProperty>
```

### <!-- Classes -->

```
<owl:Class rdf:about="&owl;Thing"/>
```

```
<!-- Basic concepts: Pizza, PizzaBase, PizzaTopping -->
```

```
</owl:Class><owl:Class rdf:about="#Pizza">
  <rdfs:subClassOf rdf:resource="&owl;Thing"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasBase"/> <owl:someValuesFrom rdf:resource="#PizzaBase"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#PizzaTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <owl:disjointWith rdf:resource="#PizzaBase"/>
  <owl:disjointWith rdf:resource="#PizzaTopping"/>
</owl:Class>
```

```
<owl:Class rdf:about="#PizzaBase">
  <rdfs:subClassOf rdf:resource="&owl;Thing"/>
  <owl:disjointWith rdf:resource="#Pizza"/>
  <owl:disjointWith rdf:resource="#PizzaTopping"/>
</owl:Class>
```

```
<owl:Class rdf:about="#ThinAndCrispyBase">
  <rdfs:subClassOf rdf:resource="#PizzaBase"/>
  <owl:disjointWith rdf:resource="#DeepPanBase"/>
</owl:Class>

<owl:Class rdf:about="#DeepPanBase">
  <rdfs:subClassOf rdf:resource="#PizzaBase"/>
</owl:Class>

<owl:Class rdf:about="#PizzaTopping">
  <rdfs:subClassOf rdf:resource="&owl;Thing"/>
  <owl:disjointWith rdf:resource="#Pizza"/>
  <owl:disjointWith rdf:resource="#PizzaBase"/>

<!-- PizzaToppings arranged in a hierarchy through subClassOf -->
<!--Look at the first few, then skip down to Definition of specific types of Pizza -->

<owl:Class rdf:about="#CheeseTopping">
  <rdfs:subClassOf rdf:resource="#PizzaTopping"/>
  <owl:disjointWith rdf:resource="#SeafoodTopping"/>
</owl:Class>

  <owl:Class rdf:about="#MozzarellaTopping">
    <rdfs:subClassOf rdf:resource="#CheeseTopping"/>
  </owl:Class>

<owl:Class rdf:about="#ParmezanTopping">
  <rdfs:subClassOf rdf:resource="#CheeseTopping"/>
  <owl:disjointWith rdf:resource="#MozzarellaTopping"/>
</owl:Class>

<owl:Class rdf:about="#VegetableTopping">
  <rdfs:subClassOf rdf:resource="#PizzaTopping"/>
  <owl:disjointWith rdf:resource="#CheeseTopping"/>
  <owl:disjointWith rdf:resource="#SeafoodTopping"/>
</owl:Class>

<owl:Class rdf:about="#TomatoTopping">
  <rdfs:subClassOf rdf:resource="#VegetableTopping"/>
</owl:Class>

<owl:Class rdf:about="#PepperTopping">
  <rdfs:subClassOf rdf:resource="#VegetableTopping"/>
</owl:Class>

  <owl:Class rdf:about="#GreenPepperTopping">
    <rdfs:subClassOf rdf:resource="#PepperTopping"/>
  </owl:Class>

<owl:Class rdf:about="#RedPepperTopping">
  <rdfs:subClassOf rdf:resource="#PepperTopping"/>
  <owl:disjointWith rdf:resource="#GreenPepperTopping"/>
</owl:Class>
```

```
<owl:Class rdf:about="#JalapenoPepperTopping">
  <rdfs:subClassOf rdf:resource="#PepperTopping"/>
  <owl:disjointWith rdf:resource="#GreenPepperTopping"/>
  <owl:disjointWith rdf:resource="#RedPepperTopping"/>
</owl:Class>

</owl:Class><owl:Class rdf:about="#MushroomTopping">
  <rdfs:subClassOf rdf:resource="#VegetableTopping"/>
</owl:Class>

<owl:Class rdf:about="#OliveTopping">
  <rdfs:subClassOf rdf:resource="#VegetableTopping"/>
</owl:Class>

<owl:Class rdf:about="#OnionTopping">
  <rdfs:subClassOf rdf:resource="#VegetableTopping"/>
</owl:Class>

<owl:Class rdf:about="#MeatTopping">
  <rdfs:subClassOf rdf:resource="#PizzaTopping"/>
  <owl:disjointWith rdf:resource="#VegetableTopping"/>
  <owl:disjointWith rdf:resource="#CheeseTopping"/>
  <owl:disjointWith rdf:resource="#SeafoodTopping"/>
</owl:Class>

<owl:Class rdf:about="#HamTopping">
  <rdfs:subClassOf rdf:resource="#MeatTopping"/>
</owl:Class>

<owl:Class rdf:about="#PepperoniTopping">
  <rdfs:subClassOf rdf:resource="#MeatTopping"/>
  <owl:disjointWith rdf:resource="#HamTopping"/>
  <owl:disjointWith rdf:resource="#SalamiTopping"/>
  <owl:disjointWith rdf:resource="#SpicyBeefTopping"/>

</owl:Class><owl:Class rdf:about="#SalamiTopping">
  <rdfs:subClassOf rdf:resource="#MeatTopping"/>
  <owl:disjointWith rdf:resource="#HamTopping"/>
</owl:Class>

<owl:Class rdf:about="#SeafoodTopping">
  <rdfs:subClassOf rdf:resource="#PizzaTopping"/>
</owl:Class>

</owl:Class><owl:Class rdf:about="#AnchovyTopping">
  <rdfs:subClassOf rdf:resource="#SeafoodTopping"/>
  <owl:disjointWith rdf:resource="#PrawnTopping"/>
</owl:Class>
```

<!-- Definition of specific types of Pizza in terms of the toppings used -->

```

<owl:Class rdf:about="#NamedPizza">
  <rdfs:subClassOf rdf:resource="#Pizza"/>>
</owl:Class>

<owl:Class rdf:about="#MargheritaPizza">
  <rdfs:comment>A pizza that has only Mozzarella and Tomato toppings</rdfs:comment>
  <rdfs:subClassOf rdf:resource="#NamedPizza"/>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#MozzarellaTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#TomatoTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <owl:disjointWith rdf:resource="#AmericanaPizza"/>
  <owl:disjointWith rdf:resource="#AmericanaHotPizza"/>
  <owl:disjointWith rdf:resource="#SohoPizza"/>

</owl:Class>

<owl:Class rdf:about="#AmericanaPizza">
  <rdfs:subClassOf rdf:resource="#NamedPizza"/>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#MozzarellaTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#TomatoTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#PepperoniTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <owl:disjointWith rdf:resource="#AmericanaHotPizza"/>
  <owl:disjointWith rdf:resource="#SohoPizza"/>
</owl:Class>

<owl:Class rdf:about="#AmericanaHotPizza">
  <rdfs:subClassOf rdf:resource="#NamedPizza"/>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#MozzarellaTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#TomatoTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>

```

```

    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom
rdf:resource="#JalapenoPepperTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#PepperoniTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <owl:disjointWith rdf:resource="#SohoPizza"/>
</owl:Class>
<owl:Class rdf:about="#CheesyPizza">
  <rdfs:subClassOf rdf:resource="#NamedPizza"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#CheeseTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:about="#SohoPizza">
  <rdfs:subClassOf rdf:resource="#NamedPizza"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#MozzarellaTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#ParmezanTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#TomatoTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasTopping"/> <owl:someValuesFrom rdf:resource="#OliveTopping"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
</rdf:RDF>

```



**Appendix B. The Zthes Abstract Model in XML** ([www.loc.gov/z3950/agency/profiles/zthes-04.html](http://www.loc.gov/z3950/agency/profiles/zthes-04.html))**Appendix B.1. The Zthes DTD for XML**

This DTD was supplied by Thomas Place. It is put forward not as a "good" XML representation of thesaurus information (whatever that might be construed to mean) but as a pragmatically valuable alternative encoding of the Zthes abstract record. Real Zthes data sets have been exchanged in the form of XML documents conforming to this DTD.

```
<!-- Zthes DTD
  Based on Z39.50 Profile for Thesaurus Navigation, version 0.1 (20 Feb 1999)
  Version of DTD: 25 Feb 1999 -->

<!-- #PCDATA: parseable character data = text
  occurrence indicators (default: required, not repeatable):
  ?: zero or one occurrence (optional)
  *: zero or more occurrences (optional, repeatable)
  +: one or more occurrences (required, repeatable)
  |: choice, one or the other, but not both
-->

<!ENTITY % term "termId, termName, termQualifier?, termType?, termLanguage?">
<!ENTITY % admin "termCreateDate?, termCreatedBy?, termModifiedDate?, termModifiedBy?">

<!ELEMENT Zthes (% term; , termNote?, % admin; , relation*)>

<!ELEMENT relation (relationType, sourceDb?, % term;)>

<!ELEMENT termId      (#PCDATA)>
<!ELEMENT termName    (#PCDATA)>
<!ELEMENT termQualifier (#PCDATA)>
<!ELEMENT termType    (#PCDATA)>
<!ELEMENT termLanguage (#PCDATA)>
<!ELEMENT termNote    (#PCDATA)>
<!ELEMENT termCreateDate (#PCDATA)>
<!ELEMENT termCreatedBy (#PCDATA)>
<!ELEMENT termModifiedDate (#PCDATA)>
<!ELEMENT termModifiedBy (#PCDATA)>
<!ELEMENT relationType (#PCDATA)>
<!ELEMENT sourceDb    (#PCDATA)>
```

(This appendix should include a crosswalk with any pre-existing thesaurus DTDs if appropriate.)

## Appendix B.2. Sample Zthes-in-XML Document

This document was supplied by Thomas Place.

```

<?XML version="1.0" ?>
<!DOCTYPE Zthes SYSTEM "zthes.dtd">
<Zthes>
  <termId>102067</termId>
  <termName>video art</termName>
  <termType>PT</termType>
  <termNote>
    Use for works of art that employ video technology, especially videotapes. For the study and practice of the art
    of producing such works, use "video."
  </termNote>
  <relation>
    <relationType>UF</relationType>
    <termId>102067/001</termId>
    <termName>art, video</termName>
    <termType>ND</termType>
  </relation>
  <relation>
    <relationType>BT</relationType>
    <termId>185191</termId>
    <termName>[time-based works]</termName>
    <termType>NL</termType>
  </relation>
  <relation>
    <relationType>RT</relationType>
    <termId>54153</termId>
    <termName>video</termName>
    <termType>PT</termType>
  </relation>
  <relation>
    <relationType>RT</relationType>
    <termId>253827</termId>
    <termName>video artists</termName>
    <termType>PT</termType>
  </relation>
</Zthes>

```

Dagobert Soergel [dsoergel@umd.edu](mailto:dsoergel@umd.edu) [www.dsoergel.com](http://www.dsoergel.com)

### Elements of an XML thesaurus / KOS data specification

This proposed schema is parsimonious yet allows the recording of many types of data. It gives enough information to derive a full XML specification. It is more expressive than SKOS

This spec assumes that data from each source are grouped, so that source attribution is not needed for each element; otherwise the structure would be much more complex. This works for a communications format but not for an internal database format.

The term itself is indicated in a relationship of type TERM. This allows for terms in multiple languages for the same concept and simplifies the schema since elements in *term* would be the same as in *relationship target*.

Addition of the *scope* element was inspired by the Topic Map Standard (see [topicmap.com](http://topicmap.com))

The scheme needs a method for indicating a relationship set defined elsewhere and used within the source or for defining a relationship set for the source.

Default is minOccurs="1" maxOccurs="1"

Source (minOccurs="0" maxOccurs="unbounded")

Pointer to or definition of relationship set used

Unit: Concept or term or group of terms (minOccurs="0" maxOccurs="unbounded")

Unique identifier

Hierarchy position (minOccurs="0" maxOccurs="unbounded")

Hierarchical level

Class number / notation

Scope for which this concept/term holds (minOccurs="0" maxOccurs="unbounded")

Relationship (minOccurs="0" maxOccurs="unbounded")

Relationship type

Relationship target

/\* See below for structure. \*/

Relationship strength (minOccurs="0" maxOccurs="1")

Audience level /\* Of this relationship \*/ (minOccurs="0" maxOccurs="unbounded")

Perspective /\* Of this relationship \*/ (minOccurs="0" maxOccurs="unbounded")

Scope for which this relationship holds (minOccurs="0" maxOccurs="unbounded")

Relationship, added information (minOccurs="0" maxOccurs="unbounded")

/\* This could be a scope note explaining the relationship, an image illustrating the relationship, another term, etc. \*/

Type of added information /\* Relationship types might be reused here. \*/

Relationship target

Audience level /\* Of this piece of info. \*/ (minOccurs="0" maxOccurs="unbounded")

Perspective /\* Of this piece of information \*/ (minOccurs="0" maxOccurs="unbounded")

Where relationship target has this structure (unifying term, text, images, multimedia document)

Relationship target

Type

/\* Includes types of terms (descriptor, other preferred term, non-preferred term and types of texts and other documents, may be an elaborate hierarchy. \*/

Target value (a term or a document)

Term

Term variant (minOccurs="0" maxOccurs="unbounded")

Type of variant

/\* Such as Preferred Spelling, other SPelling, ABbreviation, Full Term. \*/

Term form (complete term or Stem plus suffix)

Complete term

Stem plus suffix

Stem

Suffix

Document

Language (zero to many, exactly one for terms)

Audience level /\* Of this relationship target \*/ (minOccurs="0" maxOccurs="unbounded")

Perspective /\* Of this relationship target \*/ (minOccurs="0" maxOccurs="unbounded")

Scope for which this/term holds (minOccurs="0" maxOccurs="unbounded")

# **Thesaurus / KOS software and its evaluation**

## **Different types of software**

- **Thesaurus management software specifically**
- **Concept mapping software**
- **Ontology editors**
- **Description-logic- based software**

# **KOS software selection criteria**

## **General criteria for evaluation of software**

Customizable

**Distributed, support for collaboration**

Track sources & changes. Version control

## **Special functions of KOS management**

### **A General system parameters**

Multiple KOS. Multiple languages

Relationship types supported

### **B Input and editing (batch and online)**

Preserve meaningful arrangement

### **C Output in various formats**

Nicely formatted hierarchical displays,  
concept maps. Web pages

Map detailed internal relationship types to  
less detailed external

### **D Processing of data**

Check or create reciprocal relationships

Create Description Logic (DL) formulas

Derive concept relationships from DL formulas

Create notations

A detailed list of criteria follows

## **Requirements for KOS Management Software.**

KOS = Knowledge Organization System. Also thesaurus management software

### **Criteria for Evaluation**

#### **Outline**

General criteria for description and evaluation of software

Special functions of thesaurus / KOS management

- A General system parameters
- B Input and editing (of input data files and online)
- C Output in various formats
- D Processing of data

#### **General criteria for description and evaluation of software**

Only a few points that are especially important in connection with thesaurus / KOS software are dealt with here.

**Database management system used.** Is it easy to produce tailor-made output. Performance.

#### **Deployment and collaboration support**

- . Can be deployed in a framework such as Web services or SOAP
- . Can be accessed by other programs that need to obtain information
- . **Supports collaborative creation and editing. distributed**

#### **Efficiency of storage**

**Version control.** Does the program keep track of all changes

#### **User interface**

- . Menus versus commands. Use of function keys, etc.
- . Use of windows
  - . Window positions fixed in program
  - . Window positions on the screen can be specified by user
- . Navigation possibilities (see editing)
- . Program asks for verification before actually recording a change in the KOS database.
- . Consistency of the user interface
- . Help



**Case sensitivity.** Are upper and lower case treated the same or different in sorting and retrieval? If the same, is this true for all characters or are there exceptions (for example, in Index 4.1 sorting is different for upper and lower case umlauts).

Note: Case is often important to distinguish words, e.g. turkey and Turkey. If case insensitive, need turkey (bird), Turkey (country)

**User influence on how the program works.**

- . The user can influence the program behavior through data input without changing the program itself.
- . . The program reads parameter from a file (possible from the line (s) at the beginning of an input file), that can be modified by the user.
- . . Program uses external files that can be changed by the user.
- . . Program accepts specifications written by the user (e.g. specification of a record structure through giving data fields) (example: database system).
- . . The user can change menus, error messages, help messages, etc.
- . The program itself can be modified according to user wishes
- . . Program change through the user himself or herself (source code available)
- . . Program modification only through the producer
- . . Effort needed for changing the program (this depends on the modularity of the program and the programming technique used. Example: in the program language C constants such as the maximum length of a term or the character used to mark a line as bold can be defined in a header file. To change these constants one needs only to change the header file and then compile the program anew, which could be done by a properly instructed non-programmer.)

**For the rest, see the pdf file**

## **Special functions of thesaurus / KOS management**

### **Note.**

For all parameters and functions of the program being evaluated the question arises how much the user can influence it. This criterion is always applicable and is explicitly mentioned only in special cases. For example, one should know whether the user can define term types, relationship types, etc. One should keep in mind, however, that many such values have a semantics which must be operated on by the program. For example, if the program has the ability to construct an overall hierarchal structure by binary NT relations, the parts of the program doing this function must use all NT-type relations, and only those. If the user defines a new relation that is a special case of NT then this can become complicated.

Whenever there are user choices, the system should provide default values so that the user who has no special requirements can use those defaults without further ado and need not concern herself with the choice of parameters and the methods for changing the parameters.

## General system parameters

### Types of KOS supported

Remark: The following types of vocabularies overlap considerably

- 1 Vocabularies used primarily for information retrieval
  - 1.1 . Classifications and thesauri
    - 1.1.1 . . Thesauri without a well-structured classification
    - 1.1.2 . . Well-structured classification
    - 1.1.3 . . Concept map
  - 1.2 . Topic map (relationally rich KOS)
  - 1.3 . Indexes for books or journals
  - 1.4 . Record filing scheme
  - 1.5 . Data dictionary (in systems analysis and software development)
- 2 Full-fledged ontology with precisely defined relationship types so that the information can be used for reasoning
- 3 Nomenclatures and taxonomy (chemistry, biology, etc.)
- 4 Dictionaries or lexica, general or special
  - 4.1 . Mono- or multi-lingual dictionaries
    - 4.1.1 . . Mono-lingual dictionaries
    - 4.1.2 . . Multi-lingual dictionaries
  - 4.2 . Glossaries
  - 4.3 . Lexica
  - 4.4 . Picture dictionary

### KOS database as a whole

- 1 Number of KOS in a KOS database
  - 1.1 . One KOS per database
    - 1.1.1 . . One of several KOS being worked on can be specified when calling the program (but each KOS is stored in its own database)
  - 1.2 . Multiple KOS integrated in one database
    - 1.2.1 . . Number of KOS that can be included
    - 1.2.2 . . Only KOS which are subsets of one unified KOS (micro-thesauri within one large KOS), or really different KOS?
    - 1.2.3 . . All KOS on an equal footing or one main KOS with connections to terms of other KOS
    - 1.2.4 . . Are there relationships between terms from different KOS? How are these relationships determined?

- 1.2.4.1 . . . Derived from the structure of the database
- 1.2.4.2 . . . Through reference to a "switching language"
- 1.2.4.3 . . . Through direct bilateral relationships between pairs of KOS
- 1.3 . Marking subsets in a single KOS (notations are the same across subsets)
- 2 Is there a starting database of terms and concepts that can be processed by the program?
- 3 Languages that can be processed: number of languages and list of languages. (This is relevant for functions that depend on the language such as normalization of plural forms to singular, decomposition of terms that include several roots - multi-word terms in English, composite words in German, spell checking, or use of a stop word list.)
  - 3.1 . All languages on a equal footing
  - 3.2 . One main language
- 4 Subjects that the program can work on: number and list (This is relevant for spell checking and possibly for operations that use certain structural properties of the terms in a special subject.)
- 5 Maximum number of terms
- 6 Stop word list
  - 6.1 . For data input (for example for the decomposition of terms that contain multiple roots, in English these are usually multi-word terms.)
  - 6.2 . Additional stop word list for KWIC or KWOC Index
  - 6.3 . Can the stop word list be changed by the user?
- 7 Does the program support hierarchical arrangement?
  - 7.1 . Maximum number of hierarchical levels
  - 7.2 . Does the program preserve sequencing on same level of the hierarchy (see below)

### **Other characteristics of the system as a whole**

- 1 Code lists for various types of data (term types, relationship types, languages, etc. that are used for checking input and/or for presentation of menus. Can the user change these lists?)

### **Data that can be given for each term and for relations between terms**

Note: This list is just a small subset of all the data that might be needed by varied applications.

- 1 Maximum term length (Recommended at least fifty, especially if there are many multi-word (or multi-root) terms and long names. Also important for input of source term lists that have long terms.)
- 1.1 . Maximum defined by the system
- 1.2 . Maximum can be defined by the user (within system limits) (This is needed if a KOS is produced for an ISAR system that has its own maximum term length.)
- 1.2.1 . . Is it possible to define a separate term length for each of multiple KOS integrated in a KOS database
- 2 Treatment of homonyms. How are the separate meanings of homonyms identified.
- 3 Language of the term. Maximal length of the language indication. Does the KOS use a standard list of language symbols (In a multilingual KOS databases indication of the language is necessary for the unique identification of a term.)
- 4 Sort form (if different from display form)
- 5 Part of speech for a term
- 6 The gender of a term
- 7 Other syntactic or morphological data
- 8 Language level (day-to-day language, discipline specific language, outdated, etc.)
- 9 Indication of whether this term may participate in relationships to other terms
- 10 Term types (See attached list for examples)
- 10.1 . Term types predefined in the system: number and list
- 10.2 . User definable term types: how many
- 10.3 . Can a separate list of term types be defined for each KOS included in an integrated KOS database?
- 11 Perspective, a value that can be used for selecting terms into lists (Index 4.1)

- 12 Marker, another value that can be used to select terms into lists (Index 4.1)
- 13 Notation
  - 13.1 . Coarse notation (for example, for identifying broad subject groupings or facets)
  - 13.2 . Detailed notation (can at the same time fulfill the functions of a coarse notation)
  - 13.3 . External notation
  - 13.4 . Internal notation (for example, a notation expressing the hierarchical structure to be used by a retrieval program for inclusive searching)
  - 13.5 . For each kind of notation: maximum length (can the maximum length be specified by the user?)
  - 13.6 . Can the user specify whether a descriptor can have several or only one notation (MeSH, for example, has for each descriptor as many notations as the descriptor has places in the parley hierarchy.)
  - 13.7 . How much influence does the user have on the form of the notation
  - 13.8 . Support for the generation of notations
- 14 The sequence of the terms on the same level of a hierarchy can be stored (This can be implemented through notation)
- 15 Relationship types (See attachment for examples.) (At a minimum, KOS software should support the relationship types specified in KOS standards.)
  - 15.1 . Relationship types predefined in the system: number and list
  - 15.2 . Relationship types that can be defined by the user: number (But see note at the beginning.)
    - 15.2.1 . . Can the user define/change the rules used by the system in processing relationship types?
  - 15.3 . Possibility of specifying many detailed relationship types in the database but map these to a few general relationship types in the user version
  - 15.4 . Can the relationship type names for the user version be freely chosen
- 16 Rules for relationship types (Rules serve for consistency checking but can also introduce unnecessary restrictions.) Examples for rules:
  - 16.1 . Synonym relationship always from descriptor to non-descriptor
  - 16.2 . Abbreviation relationship always from descriptor to non-descriptor
- 17 Data about relations
  - 17.1 . Strength of connection

- 17.2 . Aspect used in establishing the relation. For hierarchical relationships: The characteristic of subdivision (However, it is preferable to create an own heading for each characteristic of subdivision to group all the narrower terms that correspond to that characteristic.)
- 17.3 . Qualification through context (that is, the relation is valid only for a certain context, or in any case the connection strength is dependent on the context. Put differently, the relation is itself an object related to another object, such as a term.)
- 17.4 . Scope note for a relation. Explains why the relation was introduced.
- 17.5 . In what output formats should the relation appear (This does not refer to the relationship type, but to the specific relation between two terms)
- 18 Maximum number of relations of a given relationship type that can be given for a term (This may differ from one relationship type to another.) (For example, some systems allow only one BT; this is not good, since mono-hierarchy is too restrictive.)
- 19 Maximum number of relations for a term altogether
  - . Is it possible to establish two relations of different types for an ordered pair of terms (for example, NT as well as RT or ST as well as RT)? (There are cases where it makes sense to have two relations coexisting.)
- 19.1 . In a single KOS
- 19.2 . In the integrated KOS database
- 20 Scope note and other text information
  - 20.1 . How many types of text information (for example, is it possible to have internal notes)
  - 20.2 . How many notes of each type of term
  - 20.3 . Maximum text length
  - 20.4 . Can descriptors inside a scope note be marked and treated specially?
- 21 For terms, notations and/or relations
  - 21.1 . Status value (Such as *included in present edition*, *kept for later decision*, *deleted*. The *deleted* status is important so that decisions on the same term must not be made again when, for example, this term appears in a newly processed source; it is also needed in order to reconstruct the state of the KOS at the time of indexing a given document.)
  - 21.2 . Source indication
    - 21.2.1 . . Maximum length of the source indication

- 21.2.2 . . . Maximum number of sources of a term or relation
- 21.3 . Date indications (Dates for various events such as inclusion in the KOS database, inclusion in a given KOS, approval by an editor, deletion from the KOS, etc.)
- 21.4 . Frequency of use (in a system that indexes with weights: Frequency of use with weight 2, frequency of use with weight 1 or 2) (Keep in mind that one always must specify the frequency with a time span.)
- 21.5 . Indication of the editor/lexicographer and reviser
- 21.6 . Editing history (edited when and by whom, revised and approved when and by whom)
- 22 Data on the sources as such (Does the program allow for a directory of sources?)
- 23 Other kinds of data provided for in the program
- 24 Can the user define additional types of data? In what limits? (Since many data about a term can be given through relationships the possibility of defining additional relationship types is important.)
- 25 Data Structure

**Level of logic supported (description logic, first order predicate calculus, etc.)**



## Data input and editing

### Data input

- 1 Batch input
  - 1.1 . Batch input of other KOS
  - 1.2 . Batch input of KOS files that have been created with a word processor or otherwise. (In many cases this is the most efficient method of inputting data. This method also allows editors to work independently from the program wherever there is a computer.)
    - 1.2.1 . . Format(s) of such input files
  - 1.3 . Command structure that allows for scheduling the input of several files in sequence (this is important because the input of a file may take a long time. With such a command one can input several files over night without intervention.)
- 2 Online data input (see also online editing)
  - 2.1 . Input of individual terms and data about them
    - 2.1.1 . . Input of term and data about the term in one step
      - 2.1.1.1 . . . Online form for all data about a term. Details about this form (for example, are there fixed fields for relationship types or is the relationship type given through an explicit name, scope note as one continued text for a number of lines, scrolling if not all information fits on one screen)
      - 2.1.1.2 . . . Script: The system prompts for the various data for a term in a fixed sequence. Is the content and the sequence of these prompts defined by the system or definable by the user?
      - 2.1.1.3 . . . How does the system treat cross-terms that have not yet been entered as main terms
    - 2.1.2 . . Input of terms and relations in separate steps
    - 2.1.3 . . Are all data about a term shown on the screen once input is completed? Can they be modified at that point?
  - 2.2 . Input of whole lists, especially hierarchies, that have been composed on the screen under KOS program control (Functionally this is very similar to batch input of KOS files as discussed.)

## Editing

- . For the selection of the terms to be edited in an editing session and for the format of display of the data to be edited see the criteria under Output
- 1 General functions in editing (Some of these are also important for input.)
  - 1.1 . Effort for different types of changes
  - 1.2 . Consistency check for changes made (see consistency check under D)
  - 1.3 . Is the user asked to verify the change?
  - 1.4 . Can changes be made with "hierarchical force"? (E.g., deleting a broad term and all its narrower terms.)
  - 1.5 . Does the system give a message if the user enters a term or relation that was considered earlier and either rejected for inclusion or deleted after it was once included?
- 2 Types of changes. For each type: How much effort
  - 2.1 . Changes for terms
    - 2.1.1 . . Adding a term
      - 2.1.1.1 . . . Specifying of the position of the new term in the hierarchical sequence (the input of a BT relation alone is not sufficient if one wants to maintain a meaningful sequence of terms on the same level)
    - 2.1.2 . . Deleting a Term
      - 2.1.2.1 . . . Are all relations deleted as well (or at least not output any more? Possibility differentiated by output formats for editing and output formats for the user version)
      - 2.1.2.2 . . . Is there a consistency check after a term was deleted? Especially the effects on the hierarchy need to be checked. It is problematic to delete a descriptor that has narrower descriptors that ought to be kept. Some systems do not allow deletion of a term that is linked to other terms through relations; the editor must first delete these relations.
    - 2.1.3 . . Adding a term that was deleted earlier
      - 2.1.3.1 . . . Are the relations that were in the system while the term was still there also added automatically?
    - 2.1.4 . . Change in term type (especially from descriptor to non-descriptor and vice-versa)
    - 2.1.5 . . Replace one term through another
  - 2.2 . Notation changes

- 2.2.1 . . Are other affected notations automatically changed accordingly (important especially when a term is added at a given position)
- 2.3 . Changes in relationships
  - 2.3.1 . . Adding a relationship
  - 2.3.2 . . Deleting a relationship
- 2.4 . Global changes (for example, add EN to all terms in the KOS database if one wants to change from an English-only KOS database to a multilingual database)
- 3 Batch Editing
  - 3.1 . File of editing commands
  - 3.2 . The program produces a file for editing (as part of its output functions). This file can be edited and re-input (All data in the file for editing where given a temporary *deleted* status. For any data not contained in the edited file, that *deleted* status becomes permanent.)
    - 3.2.1 . . Format of the file for editing (for example, Generic Word Processor format or a format that can be used by an outline processor)
    - 3.2.2 . . See also online editing and output regarding the criteria for selection of terms and the display format
- 4 Online editing (most of the functions apply also to online input)
  - 4.1 . History functions
    - 4.1.1 . . Is navigation history kept? Can the user retrace steps?
    - 4.1.2 . . Complete transaction log for error recovery?
  - 4.2 . Manipulation of lists of terms that must undergo editing
    - 4.2.1 . . Editing lists can be stored and recalled
      - 4.2.1.1 . . . During one session
      - 4.2.1.2 . . . From one session to the next
      - 4.2.1.3 . . . Editing lists named by the user or by the system
    - 4.2.2 . . Navigation in the editing list
      - 4.2.2.1 . . . Screen by screen
      - 4.2.2.2 . . . Scrolling
    - 4.2.3 . . Deleting elements from an editing list
  - 4.3 . Navigation in a batch of forms
  - 4.4 . Switching between editing lists and editing batch of forms

- 4.5 . Editing data for an individual term
  - 4.5.1 . . Editing data about an individual term in a list
  - 4.5.1.1 . . . Which data are displayed (see C)
  - 4.5.1.2 . . . Which data can be edited (These editing changes can be changes to the database or they can be changes that influence further editing, such as marking a term as processed or moving a term to another list.)
  - 4.5.1.3 . . . Can new terms be input while working on an editing list?
  - 4.5.1.4 . . . Does system display available options (for example, when working on BT relationships, the system might display a list of the terms that would be legal and the user would select; see consistency checks)
  - 4.5.2 . . Editing data for a term on an online form (most systems would always allow input of new terms in this context by having the user request an empty form)
  - 4.5.2.1 . . . Screen format and editing options (for example, is it possible to do full screen editing as in a word processor using the general keys like arrows and delete, can text be copied from one place to another, from one form to another, can scope notes be edited as continuous text, mouse support.)
  - 4.5.2.2 . . . Function for exchanging descriptor with one of the synonyms
  - 4.5.3 . . Jumping to a cross-referenced term, editing it, and returning to the term previously worked on (possibly do this multiple steps)
  - 5 Editing entire structure, especially a section of a hierarchy, without detailed data for each term. This is functionally equivalent to editing and re-input of an editing file in hierarchical format as discussed above, but may be more convenient.
  - 5.1 . Functions offered for editing (for the editing of hierarchies the functions of outline processing are especially useful)
- Reports on inconsistencies** (For example, relationships to a nonexistent term) in a form that facilitates the input.
- 1 Batch
  - 2 Online
- Reports on changes, especially if there is a procedure for the edition and final approval.**

## Output

Note: Output can be for human use, either printed or online, by KOS users or for editing, or for use by another system. Furthermore, many of the functions/criteria discussed here apply also to the selection of a group of terms for online editing. This includes the selection and sequencing of terms to be edited online, the data displayed on the screen, and the extent to which the user can control these parameters.

### General criteria for all output functions

(One and the same KOS management program can have different values for different output formats.)

- 1 Domain of the output
  - 1.1 . An individual KOS (either the only KOS in the database or an individual KOS from an integrated database)
  - 1.2 . Terms that appear in multiple KOS
    - 1.2.1 . . User can specify a list of KOS
    - 1.2.2 . . Concordance
    - 1.2.3 . . Comparison print: a printout that shows how the terms occurring in one or more source KOS are dealt with in a target KOS, highlighting especially terms missing from the target KOS
- 2 Selection of terms from the domain (Many of these criteria are important especially for editing.)
  - 2.1 . Scope in a hierarchy (identified by beginning and ending notation or all terms under a broad term)
  - 2.2 . Selection by relationship to another term or object
  - 2.3 . Selection by facet
  - 2.4 . Selection by hierarchical level
  - 2.5 . Scope in alphabetical sequence (identified by beginning and ending term)
  - 2.6 . Selection by status
  - 2.7 . Selection by markers or perspective
  - . . Selection by absence from a given KOS. (This is important for editing: If a new source is added to the KOS database, check terms absent from the KOS being worked on to see whether they should be included.)
  - 2.8 . Selecting terms that are not yet revised and approved

- 2.9 . Select terms not included in the last printed or otherwise published version
- 2.10 . Selection by language
- . . Selection by string pattern contained (free text searching). How powerful are the possibilities for defining patterns (wild cards for characters, for strings, etc., phrase searching vs. just word searching, etc.)
- . .
- 2.11 . Selection by internal term number (record number)
- . . Selection by specific notation
- 2.12 . Selection by a boolean combination of the criteria
- 2.13 . Selection of a small list by marking terms in a big list
- 3 Sequencing of the selected terms for presentation (this is important to achieve a meaningful sequence for editing)
- 3.1 . Hierarchical sequence
- 3.1.1 . . Stored hierarchical sequence (usually implemented through notation)
- 3.1.1.1 . . . If the domain includes several KOS: Can the editor select one KOS as a guide that will determine the hierarchical sequence?
- 3.1.2 . . Hierarchical sequence generated by the program based on hierarchical relationships (This usually implies alphabetical sequence of the children under the same parent.)
- 3.2 . Alphabetical sequence
- 4 Determining the entry point for the list
- 5 Method for calling up a list (This may be different for the different selection criteria. For example: Index 4.1 the editor working on a term can position the cursor on the facet field and call up a list corresponding to the value; when the facet field for the term being worked on has the value "Person" then the list called up includes all terms from the facet "Person".)
- 6 Content and format of the output (for screen forms and for lists) (For each criterion: how much control does the user have?)
- 6.1 . Data for each term
- 6.1.1 . . Suppressing relationships that are shown through arrangement (especially suppression of hierarchical relationships that are shown through the sequence and indication of the hierarchical level)

- 6.2 . Differentiation of relationships types
- 6.3 . Symbols for relationship types
- 6.4 . Sequence of data and relationships for one term
- 6.5 . Sequence of the cross terms within the same relationship type
- 6.6 . Are cross terms shown with their notation
- 6.7 . Are cross terms that have narrower terms identified (for example, by a plus before or after the notation or before or after the term) (This is important because the searcher or indexer should check to see whether one of the narrower terms is more suitable than the cross term.)
- 6.7.1 . . If yes, is this indication fixed by the system or selectable by the user? Is the symbol chosen (in the example plus) user selectable?
- 6.8 . Orientation aids for the user (such as giving the first and last term on the page in an alphabetical list or the first and last notation on a page in a hierarchical list)
- 7 Number of languages presented in the output format
- 7.1 . Monolingual KOS
- 7.2 . Multilingual KOS
- 7.2.1 . . Parallel arrangement with a column for each language
- 8 Specification of the output format
- 8.1 . Only predefined formats (The evaluation of a KOS management program should include detailed descriptions and sample pages of these redefined formats.)
- 8.2 . Specification of the output format through the user
- 8.2.1 . . Specification online. Can the resulting specification be stored and recalled under a name?
- 8.2.2 . . Specification through a special specification file that can be produced with a word processor
- 8.2.3 . . How complex is this specification (this must be seen in relation to the number of formatting options offered)
- 8.2.4 . . How compact is the specification
- 8.2.5 . . Does the program come with predefined formats or specification files which the user can simply use as is or modified, which would be less work than creating these files from scratch. (Include in the evaluation detailed description and sample pages of these redefined specifications.)
- 9 Possibility to order several outputs at the same time (e.g., for overnight processing)

## Printed KOS for public use

Note: Many of the format specifications listed here apply also to online displays, particularly Web displays.

- 1 Printing methods supported: especially laser printer support (for example, through output of a file in the format of a word processing or desk top publishing program), Photo Type Setting Support. File with general markup language
  - 1.1 . Formatting into pages, especially considering proportional fonts and different font sizes
  - 1.2 . Formatting into columns
    - Note: Formatting into pages or columns important for producing orientation aids for the user
- 2 Can the output file be edited before printing?
- 3 Output formats
  - 3.1 . Hierarchal lists of terms
    - 3.1.1 . . Sequence of the hierarchy, see above
    - 3.1.2 . . Specificity of the hierarchical list
      - 3.1.2.1 . . . Hierarchical outline
      - 3.1.2.2 . . . Hierarchical list of all terms
    - 3.1.3 . . Degree of detail of the hierarchy
      - 3.1.3.1 . . . Quick hierarchical list
      - 3.1.3.2 . . . Annotated hierarchical list
    - 3.1.4 . . Method for showing the hierarchical level
      - 3.1.4.1 . . . Showing the hierarchical level through indentation
        - 3.1.4.1.1 . . . . Indentions with a special symbol (for example, a dot) for each level
        - 3.1.4.1.2 . . . . With additional explicit indication of the hierarchical level
        - 3.1.4.1.3 . . . . Indention, type size, and normal/bold as a function of the hierarchical level
        - 3.1.4.1.4 . . . . Maintaining the hierarchical context through repeating the hierarchical change at the beginning of each (left that is even)
      - 3.1.4.2 . . . Hierarchy without indention with explicit indication of the hierarchical level, esp. for two column printouts



- 3.2 . Graphical representation of conceptual relationships (concept maps, topic maps)
- 3.3 . Alphabetical lists of terms
- 3.4 . Alphabetical index
- 3.4.1 . . KWOC index
- 3.4.1.1 . . . KWOC index in which the access words are normalized to singular form

**Online search for navigation in the KOS using the Web or the program itself** (also important for editing)

- 1 Web files
- 1.1 . Generation of hyperlinks and anchors for jumping from an outline to a quick hierarchy to an annotated hierarchy and for following relationships
- 1.2 . Explorer-type expandable hierarchy
- 1.3 . Control over partitioning the KOS to get Web files of reasonable size
- 1.4 . Capability for showing coordinated windows on the Web

**Files for communicating KOS data to retrieval systems (such as DIALOG or search engines or intranet retrieval engines) or to other KOS management programs**

- 1 Files compliant with a given standard, for example ZThes
- 2 Files that can be input into a database system for searching the KOS. If the database is Web-enabled, this can be combined with KOS Web files.

**Change reports**

- 1 Report of changes since a given date
- 2 Report of changes since the last printed or otherwise published edition

**Statistical reports** (Number of descriptors and entry terms, number of descriptors in each major class, number of descriptors on each hierarchical level, number of each type of relationship)

## **Processing of data (consistency checks, etc.) through the program**

**In general: how much support does the program offer in the processing and generation of data** (for example, constructing a hierarchy from BT/NT relationships, generation of notations)? The other way around: To what extent is the program limited to managing the data input by the user.

### **Checking input data for formal correctness (in batch input or during online editing)**

- 1 Checking the term length for main terms and cross terms
- 2 Checking the relationship symbols, term type symbols, language symbols, etc.
- 3 Checking for illegal terms in a hierarchy (A jump by more than one level down is illegal.)
- 4 Checking completeness (for example, checking whether a notation is given for a term when one is required)
- 5 Spell check

### **Consistency checks (during batch input and online editing)**

- 1 General characteristics of consistency checks
  - 1.1 . Is the check mandatory or user selectable
  - 1.2 . Force of the check (maybe be different for different kinds of check)
    - 1.2.1 . . There is no way to input inconsistent data
    - 1.2.2 . . Merely a warning to the editor
- 2 Consistency checks for terms
  - 2.1 . Check for form of term
    - 2.1.1 . . Check whether the term agrees with the rules of form established for the KOS (for example, preference for singular, preference for nouns over adjectives or verbs)
    - 2.1.2 . . Singular/plural check (whichever is preferred in the KOS) or conversion
    - 2.1.3 . . Spell check
    - 2.1.4 . . Capitalization check or conversion (Some terms must always be capitalized; can this be enforced, for example by having these terms in the KOS database?)
  - 2.2 . Duplication check for terms

- 2.2.1 . . Does duplication check consider singular and plural as the same? (In an integrated database both can appear.)
- 2.2.2 . . Does the duplication check consider variant spellings the same?
- 2.2.3 . . Does the duplication allow the editor to take care of homonyms
- 2.2.4 . . Can the program handle identical strings of characters that denote separate terms in different languages?
- 3 Consistency check for relationships
- 3.1 . Duplication checks for relationships
- 3.2 . Check for reciprocal relationships and creation of reciprocal relationships where needed
- 3.3 . Check for several relationships between the same ordered pair of terms (If this is not allowed, it should be checked, but only within an individual KOS.)
- 3.4 . Check for a relation of a term to itself
- 3.5 . Check for terms that are not preferred terms used as cross term in a concept relationships. Alternatively, replacing a term through the appropriate preferred term when producing output
- 3.6 . Check for ST-type relationship from descriptor to descriptor
- 3.7 . Check the consistency of hierarchical relationships
- 3.7.1 . . Check for hierarchical relationships that jump a level, for example, A NT B, B NT C, A NT C
- 3.7.2 . . Check for hierarchy cycles, for example, A NT B, B NT C, C NT A (Such cycles could throw the program for a loop in the generation of a complete hierarchical structure from hierarchical relationships.)
- 3.8 . Check for incomplete relationships, for example, semantic factoring with only one semantic factor
- 3.8.1 . . Check for terms that do not participate in any relationship (orphan terms)

- 3.9 . More complex checks of the semantic consistency of a relationship (example for a rule: hierarchical relationships are allowed only between terms belonging to the same category (we do not say here whether this rule is good or bad). ST-type relationship only in some language if TRanslation relationship is used from one language to another (using ST-type relationships regardless of the languages involved might actually be better). Formal ontologies define properties of concepts such that only concepts that agree in these properties can be hierarchically related. For example, for a concept that represents a class of objects, each instance has identity, but for a concept that designates an amount this is not the case.

- 4 Check whether input data conform to the field or relationship rules (For example, some relationship may be valid only to places, so the cross term must be a place name.)

**Support hierarchical inheritance**

**Support reasoning over the knowledge base**

**Support in the editing of terms**

- 1 Normalization of terms to singular (while storing the original form)

**Generation of notations**

- 1 The editor can input some or all notations, but where notations are missing the program generates them through hierarchical extension.

- 2 Format of the notations generated

**Support for the processing of relationships, for example**

- 1 Support for the generation of relationships, for example

- 2 . Extracting single words from a multi-word term and presenting them as candidates for semantic factors

- 3 . Presenting candidates for semantic factors based on hierarchical inheritance from the broader terms

- 4 Generating hierarchical relationships from an input list in hierarchical format

. Conversely, generation of a hierarchical sequence from binary hierarchical Relationships

- 5 In an integrated KOS database: use of synonym relationships in detecting the identity of conceptual relationships, for example KOS 1: A BT B; KOS 2: A BT C; any KOS: B ST C; conclusion: the two hierarchical relationships are the same conceptual relationships.

## Appendix

### KOS development process, with emphasis on multilingual KOS

Building a KOS, especially a multilingual KOS, takes a lot of effort. Some term relationships can be derived by statistical analysis of term occurrence in corpora, but this will not result in the kind of well-structured conceptual system described above. Developing such a structure requires intellectual effort.

A common method for KOS development in a single language is to work bottom-up: One collects a list of terms (words and phrases), preferably from search requests, but also from documents, free-term indexing, and other KOS. These terms are then sorted into increasingly fine-grained groups, until a group contains only synonyms or terms that, for purposes of the KOS, can be considered synonyms. In this process at least some homonyms will be detected; they must be disambiguated into several senses, each expressed by its own (possibly newly coined) term having one meaning and being grouped accordingly. A group of synonyms can be considered to represent a concept; usually a preferred term to designate the concept is selected, but some other concept identifier can be used. A first rough hierarchy of concepts emerges from this process.

Now perform conceptual analysis, especially facet analysis at various levels, resulting in a well-structured faceted hierarchy. Next write definitions (scope notes) (often results in rethinking the hierarchy) and introduce relationships between concepts that complement the hierarchy.

The development of a multilingual KOS is, naturally, an even more complex undertaking; the basic approaches are summarized in the table below. The ideal way to develop a multilingual KOS is to start from a pool of terms in all covered languages and carry out the process without regard to the language of the terms. This will bring together terms from different languages that have the same meaning into one group. This process gives all languages an equal chance to contribute concepts and concept relationships. It also forces a careful analysis of the meaning of each term in each language to determine the degree of equivalence, making it possible to develop the fine-grained structure of definitions that has the potential of providing powerful support to free-text cross-language retrieval.

Of course, this process requires a lexicographer knowledgeable in the subject matter of the KOS and fluent in all covered languages, not a very practical requirement. A more practical variation that still maintains the spirit of this approach is to start with two languages and develop the conceptual structure — a bi-lingual lexicographer is needed in any event. Definitions should be written in both languages. One would then work on a pool of terms in a third language and fit it into the structure, creating new concepts as necessary. This is not at all the same as translating the KOS into the third language. This requires a lexicographer fluent in one of the starting languages and the third language. Add other languages the same way.

The result of such a process is a conceptual system that brings the conceptual structures embedded in the different languages under one roof, so to speak.

The most common approach to the development of a multilingual KOS is to translate an existing monolingual KOS into one or more languages. But this approach is problematic: The original language and its vocabulary determine the conceptual structure, and one merely looks for equivalent terms in the second language without covering its terminological richness. In some multilingual KOS, only one term in the target languages is provided, making the KOS unsuitable for query term expansion in free-text searching.

In between is an approach in which one starts with a monolingual KOS as the center and fits terms from one or more other languages into the structure of this central KOS without changing the concepts or the hierarchy. Euro-WordNet (Gillaranz 1997) takes an improved variation of this approach, working with the English WordNet as its central KOS. In EuroWordNet, separate and independent word nets are constructed in each language in parallel efforts, each identifying synonym sets in that language (A synset can be considered a concept). The synsets of each language are then mapped independently to WordNet synsets; no changes are made to WordNet. In addition to identity, this mapping allows for hyponym and hypernym relationships, thus indicating that the concept identified in the language being worked on is not included in WordNet, but giving at least the hierarchical location. EuroWordNet also uses a very weak variation of approach 5: The participants developed a “top ontology”, which presumably reflects and integrates perspectives from their individual cultures. In addition to being mapped to WordNet, the individual language synsets are also mapped to this top ontology.

### Building a multilingual KOS

**Requirements:** Must cover all concepts of interest to the users in the various languages, at a minimum all domain concepts lexicalized in any of the participating languages.

Must accommodate hierarchical structures suggested by different languages.

**Approaches** (by increasing complexity and quality)

(1) Start from monolingual KOS and translate. This approach does not capture concepts lexicalized only in another language and is biased to the conceptual structure underlying the starting language. May not produce all synonyms in the second language.

(2) Start from a monolingual KOS as the center. Collect terms from other languages and establish correspondences of these terms to the central KOS. Suffers from similar bias toward the starting language as (1), but may cover more synonyms in the other languages.

(3) Work with a central KOS as in (2), but after collecting terms from a second language first group them into synsets, that is, derive concepts each of which is represented by a set of terms, and then map each concept to the corresponding concept in the central KOS or indicate that the concept is new and give the nearest broader or narrower concept in the central KOS. Note that the central KOS remains unchanged.

(4) As (2), but add concepts not in the starting KOS. This mitigates bias, but the central KOS now becomes a moving target.

(5) Start from a pool of terms from all participating languages and organize them into a conceptual framework, establishing term correspondence in the process. This approach results in a true "conceptual interlingua" not biased to any one language, but offering a home to multiple conceptual perspectives. This approach requires most effort.

## **KOS development example**

### **Audience/Demographic Characteristics**

**Raw term list**

Terms collected from lists used in three NCADI databases, from the NCADI request form, and from *Breaking New Ground for Youth At Risk*, duplicates eliminated, in alphabetical order

A/D prevention professional	Legislator
A/D treatment professionals	Librarian/Information Specialist
Administrator/Manager	Media representatives
Adults (25-59 years)	Mental health professional
African Americans/Black	Native Americans (American Indians and Eskimos)
Asians and Pacific Islanders (Chinese, Japanese, Vietnamese, etc.)	Other
Attorney	Parents (specify age of child)
Biomedical researchers	Parole/Probation officer
Blacks	Patients
Caucasians	Police officer
Children subjected to abuse and neglect	Policy makers/Administrators
Children and youth who are economically disadvantaged	Preschool (age 4 and under)
Children (pre-adolescents)	Psychosocial researcher
Children of alcoholic or other drug-abusing parents	Recreation/Sports Personnel
Clergy	Reporter/writer
College students	Researcher
Community organization leader	School dropouts or those at risk of dropping out
Community service groups	School Administrator
Correction officer	Scientists and researchers
Criminal/juvenile justice	Single teenage mothers and their children
Disabled	Social service professional
EAP Practitioners	Sr. High Youth (16-18)
Educator/teacher/trainer (specify grade[s])	Student
Elderly (60 + years)	Unemployed youth or those in danger of being unemployed
Elementary youth (5-12)	Unknown/anonymous
Employees	Women
Employer	Young Adults (18-25 years) (19-25 years)
General public	Youth who use gateway drugs
General public, personal concern	Youth (adolescents)
General public, concern for family/friend	Youth who are suicidal or physically or mentally disabled
Grantee	Youth who are engaged in violent or delinquent acts
Handicapped/Disabled	
Health care providers (physicians, nurses, PAs, NAs, pharmacists)	
Health care professional	
High-risk families	
High-risk youth	
High-risk families/youth (including COAs and ACOAs)	
Hispanics/Latinos	
Homeless or runaway youth	
Homosexuals (males and females)	
IV drug users	
Jr. High Youth	
Judge	
Latchkey children	



**Terms collected arranged in broad groupings****Age**

Preschool (age 4 and under)  
 Elementary youth (5-12)  
     ST Children (pre-adolescents)  
 Youth (adolescents)  
     Jr. High Youth  
     Sr. High Youth (16-18)  
 Young Adults (18-25 years) (19-25 years)  
     College students  
 Adults (25-59 years)  
 Elderly (60 + years)  
 Student?

**Gender**

Women  
 Men

**Sexual preference**

Homosexuals (males and females)

**Racial/ethnic group**

African Americans/Black  
 Asians and Pacific Islanders  
 Caucasians  
 Blacks  
 Hispanics/Latinos  
 Native Americans (Amer. Indians, Eskimos)

**group by ability/handicap**

Disabled  
 Handicapped/Disabled

**Groups at high risk of drug use**

Children subjected to abuse and neglect  
 Children and youth economically disadvantaged  
 Children of alcoholic or other drug-abusing parents  
 High-risk families  
 High-risk families/youth (including COAs and ACOAs)  
 High-risk youth  
 Homeless or runaway youth  
 Latchkey children  
 School dropouts or those at risk of dropping out  
 Single teenage mothers and their children  
 Unemployed youth or those in danger of being unemployed  
 Youth who use gateway drugs  
 Suicidal or physically or mentally disabled youth  
 Youth who are engaged in violent or delinquent acts  
 IV drug users

**By profession or position**

A/D prevention professional  
 A/D treatment professionals  
 Administrator/Manager  
 Attorney  
 Clergy  
 Community organization leader  
 Community service groups  
 Correction officer  
 Criminal/juvenile justice  
 EAP Practitioners  
 Educator/teacher/trainer (specify grade[s])  
 Health care providers (physicians, nurses, Pas, NAS, pharm.)  
 Health care professional  
 Judge  
 Legislator  
 Librarian/Information Specialist  
 Media representatives  
 Mental health professional  
 Parole/Probation officer  
 Police officer  
 Policy makers/Administrators  
 Recreation/Sports Personnel  
 Reporter/writer  
 Researcher  
     Biomedical researcher  
     Psychosocial researcher  
  
 School Administrator  
 Scientists and researchers  
 Social service professional

**By employer/employee relationship**

Employees  
 Employer

**Other groupings**

Patients  
 Parents (specify age of child)  
 General public  
     General public, concern for family/friend  
     General public, personal Concern  
 Grantee  
 Other  
 Unknown/anonymous

### One area conceptually refined

#### Groups at high risk of drug use

Suicidal or physically or mentally disabled  
 Persons from unstable or low-cohesion families  
 Children of alcoholic or other drug-abusing parents  
     SN Grown up or still under age  
 Children of single teen-age mothers  
 Persons subjected to abuse or neglect  
     SN Now or in the past  
     Persons subjected to abuse and neglect by parents  
         Latchkey children  
     Persons subjected to abuse and neglect by their spouse  
 Single teenage mothers  
 School dropouts or those at risk of dropping out  
 Unemployed or in danger of being unemployed  
 Economically disadvantaged  
 Homeless  
     Runaway youth  
 Gateway drug users  
 Persons engaged in violent or delinquent acts

The concept *Youth at risk of drug use* or any of its subordinate concepts (as specified by group and age range) can be produced by combination with *Adolescent*

An observation on defining groups by combination: For any group defined by personal/demographic characteristics, there are several derivative groups, e.g.:

Parents of members of the group  
 Children of members of the group  
 Spouses of members of the group  
 Teachers of members of the group

The KOS needs to make provision for forming such combinations.

# **Examples of Thesauri and other Knowledge Organization Systems (KOS)**

The paper tutorial notebook includes copies of sample pages from print thesauri / KOS, Web pages on thesauri / KOS, and of search results in various Web thesauri / KOS. The pdf file does not include these pages, but it does include the URLs of applicable Web pages.



# Alcohol and Other Drug Thesaurus

## Alcohol and Other Drug Thesaurus: A Guide to Concepts and Terminology in Substance Abuse and Addiction.

3rd ed. Washington, DC: U.S. Department of Health and Human Services, 2000.

Volume 1: Introduction and Overview, 387 p.

Volume 2: Annotated Hierarchy, 848 p.

Volume 3: Alphabetical Index, 406 p.

Volume 4: Annotated Alphabetical List, 896 p.

		2nd ed.	3rd. ed.
Number of:	Descriptors	10,315	11,323
	Lead-in Terms	6,675	7,783
	Total Terms	16,990	19,106
	History Notes		2,900
	Scope Notes (2.ed. incl. HN)	2,351	2,085
	Total descriptor cross-references	35,108	<b>39,720</b>

Web: <http://etoh.niaaa.nih.gov/AODVol11/Aodthome.htm>

To order:

CSR Inc  
1400 Eye St, NW, Ste 200  
Washington, DC 20005  
tel. (202) 842-7600  
US\$ 100 plus shipping



# Art and Architecture Thesaurus

*Art and Architecture Thesaurus*. 2nd ed. Getty Art History Information Program.  
New York: Oxford University Press, 1994.

Volume 1 & 2: Introduction and Hierarchies, 455 p., 533 p.

Volume 3 - 5 : Aand - Zutu, 586 p., 586 p., 546 p.

US\$ 375

Number of:	Descriptors	24,500
------------	-------------	--------

	Guide Terms	2,750
--	-------------	-------

	Synonyms	20,000
--	----------	--------

	<b>Total Terms</b>	<b>47,000</b>
--	--------------------	---------------

	British Variants	2,000
--	------------------	-------

	Alternate Terms (singular/plural)	16,000
--	--------------------------------------	--------

	Permutations	27,000
--	--------------	--------

	<b>Terms and Variants</b>	<b>92,000</b>
--	---------------------------	---------------

## *Electronic Editions:*

See <http://www.getty.edu/research/tools/vocabulary/obtain.html>

## *On the Web*

<http://www.getty.edu/research/tools/vocabulary/aat/index.html>

Getty Vocabulary Program home page (copy included in this package)

<http://www.getty.edu/research/institute/vocabulary/introvocabs/>

## Review article

Dagobert Soergel. **The Arts and Architecture Thesaurus (AAT). A critical appraisal.**  
*Visual Resources* 1995; 10(4): 369-400.

A few sample pages from an expanded version of this article are included here.

## Dagobert Soergel

### The Arts and Architecture Thesaurus (AAT). A critical appraisal.

#### 1 Introduction: Thesauri in information retrieval

What is a thesaurus and what is its purpose? Describing the functions of a thesaurus in a nutshell will provide the background for a critical examination of the AAT. A thesaurus is a structured collection of concepts and terms for the purpose of improving the retrieval of information. A thesaurus should help the searcher to find good search terms, whether they be descriptors from a controlled vocabulary or the manifold terms needed for a comprehensive free-text search — all the various terms that are used in texts to express the search concept. Most thesauri establish a controlled vocabulary, a standardized terminology, in which each concept is represented by one term, a descriptor, that is used in indexing and can thus be used with confidence in searching; in such a system the thesaurus must support the indexer in identifying all descriptors that should be assigned to a document or other object in light of the questions that are likely to be asked. A good thesaurus provides, through its hierarchy augmented by associative relationships between concepts, a semantic road map for searchers and indexers and anybody else interested in an orderly grasp of a subject field.

A good thesaurus can be used for automatic search query expansion in two ways:

- (1) synonym expansion, adding all the synonyms for a search term needed for free-text searching. For example,

<b>color proofs</b>	add	color separations
<b>barrel vaults</b>	add	cradle vaults, tunnel vaults, wagon vaults, wagonhead vaults
<b>bluish gray</b>	add	aqua gray, baby blue, blue black, blue gray, centroid color 191, light Payne's gray, pewter, powder blue, slate

- (2) hierarchic expansion, adding all the narrower terms for a search term (also called inclusive searching). This is needed whether one searches with a controlled vocabulary or free-text, for example,

<b>humanities</b>	add	arts, linguistics, literature, philosophy, history, etc.
<b>gold</b>	add	electrum, chryselephantine sculpture
<b>barrel vaults</b>	add	annular vaults, half barrel vaults, rampant barrel vaults, spiral vaults
<b>saints</b>	add	<b>hagiography, hagiographies</b>



**B Associated concepts facet (1,018)**

BM Associated concepts (1018)

**D Physical attributes facet (890)**

DC Attributes and properties (353)

DE Conditions and effects (46)

DG Design elements (162)

DL Color (329)

**F Styles and periods facet (3,382)**

FL Styles and Periods (3,382)

**H Agents facet (1,093)**

HG People (958)

HN Organizations (135)

**K Activities facet (2,034)**

KD Disciplines (318)

KG Functions (287)

KM Events (177)

KQ Physical activities (87)

KT Processes and techniques (1,165)

**M Materials facet (2,869)**

MT Materials (2,869)

**P/V Objects facet (13,210)**

PC Object groupings and systems (202)

PE Object genres (154)

PJ Components (3,066)

**R Build Environment (1,943)**

RD Settlements and landscapes (241)

RG Built complexes and districts (287)

RK Single built works (1,185)

RM Open spaces and site elements (230)

**T Furnishings and equipment (5,592)**

TC Furnishings (1,363)

TE Costume (721)

TH Tools and equipment (1,463)

TK Weapons and ammunition (256)

TN Measuring devices (315)

TQ Containers (622)

TT Sound devices (607)

TV Recreational artifacts (183)

TX Transportation vehicles (462)

**V Visual and verbal communication (1,853)**

VC Visual works (574)

VK Exchange media (169)  
VW Information forms (1,110)

Numbers in parentheses give the number of descriptors to indicate emphasis.

Figure 1. **Top-level outline**

<b>Facet</b>	<b>Sample descriptors</b>
physical attributes	quarter plate, opacity, vivid red
styles and periods	Rococo
agents	painters (artists), photographers
activities and processes	gilding, gelatin silver process, color photography, carving, deterioration
materials	color film, wood
objects	chairs, negatives

Figure 4. **Facets and sample descriptors**

VC1	<visual works>
VC2	<visual works by form>
VC34	<visual works by function>
VC70	<visual works by location or context>
VC75	<visual works by medium or technique>
VC283	<b>photographs</b>
VC284	<photographs by form>
VC285	negatives
VC289	<negatives by color>
VC290	black-and-white negatives
VC291	color negatives
VC292	<negatives by process>
VC295	gelatin silver negatives
VC299	positives
VC310	photographic prints
VC312	later prints
VC315	<photographic prints by color>
VC316	black and white prints (photographs)
VC317	color prints (photographs)
VC318	<photographic prints by process>
VC322	chromogenic color print
VC346	<photographs by form: color>
VC347	black-and-white photographs
VC348	color photographs
VC349	<photographs by form: format>
VC357	slides (photographs)
VC358	black-and-white slides
VC359	color slides
VC360	<photographs by function>
VC363	news photographs
VC364	<photographs by technique>
VC365	<photographs by picture-taking technique>
VC366	aerial photographs
VC381	<photographs by processing or presentation technique>
VC390	manipulated photographs
VC391	composite photographs
VC400	<photographs by subject type>
VC406	studio portraits

Figure 5. Example for minor facets and precombined descriptors

## Photography

### D Physical Attributes Facet, DC Attributes and Properties

DC111 <*size: photograph formats*>  
 DC116 quarter plate

### D Physical Attributes Facet, DE Conditions and Effects

DE38 <*conditions and effects: photography*>  
 DE39 oxidative-reductive deterioration

### H Agents Facet, HG People

HG299 photographers

### K Activities Facet, KT Processes and Techniques

KT487 <*photography and photographic processes and techniques*>  
 KT503 photographic processes  
 KT526 gelatin silver process  
 KT567 <*photographic techniques*>  
 KT570 <*picture-taking techniques*>  
 KT571 chronophotography  
 KT598 <*photographic processing and presentation techniques*>  
 KT602 enlarging  
 KT616 reduction (photography)

### M Materials Facet, MT Materials

MT1416 paper  
 MT1463 <*paper by function*>  
 MT1481 photographic paper  
 MT2364 photographic materials  
 MT2367 photographic film

### P/V Objects Facet, TH Tools and Equipment

TH746 photographic equipment  
 TH747 <*cameras and camera accessories*>  
 TH788 <*photographic processing equipment*>  
 TH794 enlargers  
 [no reducers]

**P/V Objects Facet, VC Visual Works**

VC283    photographs  
VC284        <photographs by form>  
VC285        negatives  
VC292            <negatives by process>  
VC295            gelatin silver negatives  
VC364        <photographs by technique>  
VC364            <photographs by picture-taking technique>  
VC367            chronophotographs

Figure 6. **Facet arrangement dispersing concepts from same subject area.**  
a. **Hierarchy excerpts concerning the subject Photography**

<art genres>

academic art  
amateur art  
apocalyptic art  
art brut  
children's art  
commercial art  
community art  
SN Includes art undertaken in conjunction with particular communities, often socially deprived, usually with the idea of producing an effect or inspiring response specifically within those communities, with no reference to widely established standards. For art intended to beautify or enrich public places, use **public art**.  
computer art  
court art  
crafts  
cybernetic art  
didactic art  
dissident art  
ethnic art  
fantastic art  
figurative art  
folk art  
funerary art  
naive art  
nonrepresentational art  
primitive art  
public art  
SN Use for art whose purpose is to beautify and enrich public places. For art undertaken in conjunction with particular communities, usually to produce an effect or inspire response specifically within those communities, use **community art**.  
rock art  
cave art  
serial art  
sofa art  
street art

a. Original alphabetical sequence

art genres

**art genres by content or other intrinsic characteristics**

figurative art  
fantastic art  
apocalyptic art  
nonrepresentational art  
cybernetic art  
serial art  
crafts

**art genres by standard**

academic art  
folk art  
dissident art

**art genres by type of artist or origin**

amateur art  
naive art  
art brut  
children's art  
computer art  
ethnic art  
primitive art

**art genres by audience, purpose, or display context**

sofa art  
court art  
public art  
SN Art whose purpose is to beautify and enrich public places.  
community art  
SN Public art undertaken in conjunction with particular communities, often socially deprived, usually with the idea of producing an effect or inspiring response specifically within those communities, with no reference to widely established standards.  
street art  
rock art  
cave art [prehistoric, esp. paleolithic]  
didactic art  
commercial art  
funerary art

b. Suggested meaningful sequence

**Meaningful arrangement**  
**Example from the Art and Architecture Thesaurus**





# Medical Subject Headings

## **Medical Subject Headings, 2005 Supplement to Index Medicus (black and white MeSH)**

Alphabetical and categorized list of subject descriptors used to analyze the biomedical literature in NLM.

\$66.00. GPO S/N: 917-033-00000-5

General MeSH info: [www.nlm.nih.gov/mesh/meshhome.html](http://www.nlm.nih.gov/mesh/meshhome.html)

Ordering info: [www.nlm.nih.gov/mesh/pubs.html](http://www.nlm.nih.gov/mesh/pubs.html)

MeSH Files Available to Download: [www.nlm.nih.gov/mesh/filelist.html](http://www.nlm.nih.gov/mesh/filelist.html)

## **MeSH on the Web**

[www.nlm.nih.gov/mesh/MBrowser.html](http://www.nlm.nih.gov/mesh/MBrowser.html) (more powerful)

[www.ncbi.nlm.nih.gov/htbin-post/Entrez/meshbrowser](http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/meshbrowser) (simpler)

Used in searching the bibliographic database Medline through PubMed

[www.ncbi.nlm.nih.gov/PubMed/medline.html](http://www.ncbi.nlm.nih.gov/PubMed/medline.html)

PubMed uses MeSH and UMLS for synonym expansion and the MeSH hierarchy for hierarchic expansion

For UMLS also Internet Grateful Med: <http://igm.nlm.nih.gov/>

## **Unified Medical Language System (UMLS)**

<http://umlsinfo.nlm.nih.gov>

[www.nlm.nih.gov/pubs/factsheets/umlskss.html](http://www.nlm.nih.gov/pubs/factsheets/umlskss.html)

[www.nlm.nih.gov/pubs/factsheets/umlsmeta.html](http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html)



### Structure of the UMLS Metathesaurus

2000: 75 source vocabularies and 25 translations. Growth since 1998: x 1.5

Strings	Terms	Concepts
1,593,730 (1,718,083 tokens)	1,338,650	730,155

Substance Dependence Substance dependence substance dependence	Substance Dependence	Substance Dependence
Addiction, chemical addiction, chemical chemical addiction chemical addictions	Addiction, chemical	

adolescent	adolescent	adolescent
Teenager Teenagers teenager	Teenager	
teen teens	teen	
youth (young person) youths youth <1>	youth (young person)	

youth	youth	
-------	-------	--

youth <2> youth (stage of life)	youth <2>	youth <2>
------------------------------------	-----------	-----------

## UMLS semantic types

[https://umlsks.nlm.nih.gov/KSS/00/Specialist/Semantic\\_Net/semtype.list.html](https://umlsks.nlm.nih.gov/KSS/00/Specialist/Semantic_Net/semtype.list.html)

Last Modified: Monday, February 07, 2000, copied May 31, 2000

### Entity

#### Conceptual Entity

- Idea or Concept
  - Functional Concept
    - Body System
  - Temporal Concept
  - Qualitative Concept
  - Quantitative Concept
  - Spatial Concept
    - Body Location or Region
    - Body Space or Junction
    - Geographic Area
  - Molecular Sequence
    - Amino Acid Sequence
    - Carbohydrate Sequence
    - Nucleotide Sequence
- Finding
  - Laboratory or Test Result
  - Sign or Symptom
- Organism Attribute
  - Clinical Attribute
- Intellectual Product
  - Classification
  - Regulation or Law
- Language
- Occupation or Discipline
  - Biomedical Occupation or Discipline
- Organization
  - Health Care Related Organization
  - Professional Society
  - Self-help or Relief Organization
- Group Attribute
- Group
  - Age Group
  - Family Group
  - Professional or Occupational Group
  - Population Group
  - Patient or Disabled Group

#### Physical Object

- Anatomical Structure
  - Anatomical Abnormality
    - Acquired Abnormality
    - Congenital Abnormality
  - Embryonic Structure
  - Fully Formed Anatomical Structure
    - Body Part, Organ, or Organ
    - Component
  - Cell

- Cell Component
- Tissue
- Gene or Genome
- Manufactured Object
- Clinical Drug
- Medical Device
- Research Device
- Organism
  - Animal
    - Invertebrate
    - Vertebrate
      - Amphibian
      - Bird
      - Fish
      - Mammal
        - Human
      - Reptile
  - Archaeon
  - Bacterium
  - Fungus
  - Plant
    - Alga
  - Virus
  - Rickettsia or Chlamydia
- Substance
  - Body Substance
  - Chemical
    - Chemical Viewed Functionally
      - Biologically Active Substance
      - Neuroreactive Substance or Biogenic Amine
      - Hormone
      - Enzyme
      - Vitamin
      - Immunologic Factor
      - Receptor
    - Biomedical or Dental Material
    - Pharmacologic Substance

Antibiotic	<b>Event</b>
Indicator, Reagent, or Diagnostic	Activity
Aid	Behavior
Hazardous or Poisonous Substance	Social Behavior
Chemical Viewed Structurally	Individual Behavior
Organic Chemical	Daily or Recreational Activity
Amino Acid, Peptide, or Protein	Occupational Activity
Carbohydrate	Educational Activity
Lipid	Governmental or Regulatory Activity
Eicosanoid	Health Care Activity
Steroid	Diagnostic Procedure
Nucleic Acid, Nucleoside, or	Laboratory Procedure
Nucleotide	Therapeutic or Preventive
Organophosphorus Compound	Procedure
Inorganic Chemical	Research Activity
Element, Ion, or Isotope	Molecular Biology Research
Food	Technique
	Machine Activity
	Phenomenon or Process
	Human-caused Phenomenon or Process
	Environmental Effect of Humans
	Injury or Poisoning
	Natural Phenomenon or Process
	Biologic Function
	Physiologic Function
	Cell Function
	Molecular Function
	Genetic Function
	Organ or Tissue Function
	Organism Function
	Mental Process
	Pathologic Function
	Cell or Molecular Dysfunction
	Disease or Syndrome
	Mental or Behavioral
	Dysfunction
	Neoplastic Process
	Experimental Model of Disease

## UMLS semantic relations

[https://umlsks.nlm.nih.gov/KSS/00/Specialist/Semantic\\_Net/relation.list.html](https://umlsks.nlm.nih.gov/KSS/00/Specialist/Semantic_Net/relation.list.html)

Last Modified: Monday, February 07, 2000, copied May 31, 2000

### associated\_with

#### physically\_related\_to

- branch\_of
- connected\_to
- consists\_of
- contains
- ingredient\_of
- interconnects
- part\_of
- tributary\_of

#### spatially\_related\_to

- adjacent\_to
- location\_of
- surrounds
- traverses

#### functionally\_related\_to

- affects
  - manages
  - treats
  - disrupts
  - complicates
  - interacts\_with
  - prevents
- brings\_about
  - produces
  - causes
- performs
  - carries\_out
  - exhibits
  - practices
- occurs\_in
  - process\_of
- uses
- manifestation\_of
- indicates
- result\_of

#### temporally\_related\_to

- co-occurs\_with
- precedes

### conceptually\_related\_to

- analyzes
  - assesses\_effect\_of
- conceptual\_part\_of
- evaluation\_of
- degree\_of
- assesses\_effect\_of
- measurement\_of
- measures
- diagnoses
- property\_of
- derivative\_of
- developmental\_form\_of
- method\_of
- issue\_in

### isa

# Dewey Decimal Classification

*Dewey Decimal Classification and Relative Index.* 21st ed. Library of Congress.  
Albany, NY: OCLC Forest Press, 1996.

Volume 1: Introduction and Tables, 625 p.

Volume 2: Schedules 000 - 599, 1200 p.

Volume 3: Schedules 600 - 999, 1105 p.

Volume 4: Relative Index, 1207 p.

Cost: US\$ 325, UK 220

## *World Wide Web:*

Dewey Decimal Classification home page

[www.oclc.org/oclc/fp/](http://www.oclc.org/oclc/fp/)

[www.oclc.org/dewey/products/webdewey/about.htm](http://www.oclc.org/dewey/products/webdewey/about.htm)

Good display of top three levels

[www.oclc.org/dewey/about/ddc\\_21\\_summaries.htm](http://www.oclc.org/dewey/about/ddc_21_summaries.htm)

[www.tnrplib.bc.ca/dewey.html](http://www.tnrplib.bc.ca/dewey.html)

[www.anthus.com/CyberDewey/CyberDewey.html](http://www.anthus.com/CyberDewey/CyberDewey.html)

## Examples of Internet Resources Classified by Dewey

[www.oclc.org/dewey/worldwide/](http://www.oclc.org/dewey/worldwide/)

<http://link.bubl.ac.uk:80/linkbrowse>

<http://www.oclc.org/oclc/man/colloq/v-g>

## More complete listing to be on

[www.clis.umd.edu/faculty/soergel/dlthestut](http://www.clis.umd.edu/faculty/soergel/dlthestut)





# WordNet

**WordNet Lexical Database.** Version 3.0. Princeton University,  
Cognitive Science Laboratory, December 2006  
(sample pages are from an older version)

WordNet is an online lexical database that is organized  
semantically rather than alphabetically.

		synsets (concepts)	word senses (terms, homonyms disambiguated)
Number of:	nouns	117,798	146,312
(version 2.0)	verbs	11,529	25,047
	adjectives	21,479	30,002
	adverbs	4,481	5,580
	total	155,287	206,941

More WordNet statistics on the back

Web home page: <http://wordnet.princeton.edu/>

Especially useful:

<http://wordnet.princeton.edu/obtain>

<http://wordnet.princeton.edu/links>

<http://www.cogsci.princeton.edu/~geo/reader> (WNet as lexical  
aid: click on word in text)

<http://wordnet.princeton.edu/doc/man/> and the glossary link there

[www.illc.uva.nl/EuroWordNet/](http://www.illc.uva.nl/EuroWordNet/)

[www.illc.uva.nl/EuroWordNet/docs/](http://www.illc.uva.nl/EuroWordNet/docs/)

esp. D017, an overview of EuroWordNet

Best search: [www.notredame.ac.jp/cgi-bin/wn.cgi](http://www.notredame.ac.jp/cgi-bin/wn.cgi)

## WordNet Statistics from <http://wordnet.princeton.edu/man/wnstats.7W>

### Number of words, synsets, and senses

POS	Unique Strings	Synsets	Total Word-Sense Pairs
Noun	117,798	82,115	146,312
Verb	11,529	13,767	25,047
Adjective	21,479	18,156	30,002
Adverb	4,481	3,621	5,580
Totals	155,287	117,659	206,941

### Polysemy information

POS	Unique Strings	Mono- semous Words and Senses	Poly- semous Words	Poly- semous Senses	Avg no.of meanings, all words	Avg no.of meanings, polysemous words
Noun	117,798	101,863	15,935	44,449	1.24	2.79
Verb	11,529	6,277	5,252	18,770	2.17	3.57
Adjective	21,479	16,503	4,976	14,399	1.40	2.71
Adverb	4,481	3,748	733	1,832	1.25	2.50
Totals	155,287	128,391	26,896	79,450		

**WordNet Lexical Database**

<http://kaml1.csi.uottawa.ca:3000/wn.html>

**WordNet Lexical Database**

The WordNet database is an encoding of relations between synonym groups representing nouns, verbs, adjectives, and adverbs. Read the manual page for further information

**synonyms and immediate hypernyms**

Nouns Verbs Adjectives Adverbs

**antonyms**

Nouns Verbs Adjectives Adverbs

**familiarity and polysemy information**

Nouns Verbs Adjectives Adverbs

**recursive hypernyms**

Nouns Verbs

**immediate hyponyms**

Nouns Verbs

**hyponym tree**

Nouns Verbs

**coordinates**

Nouns Verbs

**meronyms**

Substance Part Member All Tree

**holonyms**

Part Member Substance All Tree

**entailment relations**

**verb sentence frames**

**Cause to relations**

**pertainyms**

Adjectives Adverbs

**adjective values / noun attributes**

Nouns Adjectives

**links between nouns and verbs that are morphologically related**

From dduchier@csi.uottawa.ca with some adaptation (Web site no longer active)

## D. Soergel Top level hierarchy of WordNet's main categories

Arranged building on the structure from the WordNet literature. Categories in [] added.

<b>nouns</b>	<b>verbs</b>	<b>adjectives</b>
<b>thing, entity</b> living thing, organism plant, flora animal, fauna person, human being and care non-living thing, object natural object body, corpus artifact substance food [other things or entities] group, collection	verbs of bodily function        contact verbs	
<b>[process, action, event]</b> process act, action, activity event, happening   natural phenomenon	<b>[process verbs]</b> verbs of change creation verbs motion verbs <b>[social interaction and            competition verbs]</b> verbs of social interaction competition verbs  consumption verbs  weather verbs	
<b>[time and place]</b> time place		
<b>[knowledge, communication,            feeling]</b> cognition, knowledge  communication feeling, emotion motive	<b>[knowledge, communication,            feeling verbs]</b> cognition verbs perception verbs verbs of communication emotion or psych verbs	
<b>[attributes and relations]</b> attribute, property state, condition shape quantity, amount possession relation	<b>[stative and possession verbs]</b> stative verbs verbs of possession	descriptive adjectives color adjectives   relational adjectives
		reference-modifying adjectives (e.g., <i>former</i> president)

# Cyc Ontology

## Cyc Ontology. OpenCyc

The old Cyc Ontology is a subset of the CYC system, a multi-conceptual knowledge base and inference engine. It is produced by

Cycorp, 3721 Executive Center Dr., Austin, TX 78731

Now replaced by OpenCyc which is not accessible directly through the Web but must be downloaded

Number of:	Concepts in the old Cyc Ontology guide (upper ontology) “the topmost few percent of the hierarchy in the Cyc® Knowledge Base.”	3000
	Concepts in the Cyc Knowledge Base	?

For info look at [www.cyc.com/](http://www.cyc.com/) and [www.opencyc.org/](http://www.opencyc.org/)

Some Web pages printed out here

Some samples from the old CYC ontology are attached.



## CYC ontology top level outline (43 classes)

From <http://www.cyc.com/cyc-2-1/toc.html> Updated 1997-8-12, accessed 2001-4-15

Reformatted

Fundamentals	Biology
Top Level	Chemistry
Time and Dates	Physiology
Types of Predicates	General Medicine
Spatial Relations	Materials
Quantities	Waves
Mathematics	
Contexts	Devices
Groups	Construction
	Financial
"Doing"	Food
Transformations	Clothing
Changes Of State	Weather
Transfer Of Possession	Geography
Movement	Transportation
Parts of Objects	Information
Composition of Substances	Perception
	Agreements
Agents	
Organizations	Linguistic Terms
Actors	Documentation
Roles	
Professions	
Emotion	
Propositional Attitudes	
Social	

Supporting Documentation

The Syntax of CycL

The CYC® Functional Interface

Glossary of Common CYC® Terms

**CYC Social Vocabulary Outline** (created by DS from full file)

Some groupings, indicated by blank lines, introduced by DS (this outline would profit from better organization)

controls : <Agent> <Individual>  
 SocialBeing  
 affiliatedWith : <Agent> <Agent>  
 acquaintedWith : <IndividualAgent> <IndividualAgent>

## Workplace

spectators : <Event> <Agent>  
 beneficiary : <Event> <Agent>

owns : <Agent> <SomethingExisting>  
 recipientOfService : <ServiceEvent> <Agent>  
 providerOfService : <ServiceEvent> <Agent>

socialParticipants : <SocialOccurrence> <Agent>  
 residesInDwelling : <Animal> <ShelterConstruction>  
 residesInRegion : <Animal> <GeographicalRegion>

## HumanOccupationConstructResident

languageSpoken : <IntelligentAgent> <NatLanguage>  
 fieldsOfFormalEducation : <Person> <FieldOfStudy>  
 fieldsOfCompetence : <Person> <FieldOfStudy>  
 fieldsOfActivity : <Person> <FieldOfStudy>

representsAgentToAgent : <Agent> <Agent> <Agent>  
 socialClass : <Person> <SocialClass-Lifestyle>  
 SocialClass-Lifestyle  
 competingAgents : <Competition> <Agent>  
 eventHonors : <SocialOccurrence> <Agent>  
 positiveVestedInterest : <Agent> <TemporalThing>  
 negativeVestedInterest : <Agent> <TemporalThing>

## AdultFemalePerson

HumanInfant  
 HumanChild  
 HumanAdult

SportsEvent  
 EntertainmentPerformance  
 EntertainmentEvent

spouse : <Person> <Person>  
 acquaintances : <Person> <Agent>  
     <AcquaintanceAttribute>  
 SimpleContactAcquaintance  
 AcquaintanceAttribute  
 friends : <Animal> <Animal>  
 boss : <Person> <Person>  
 cohabitingFamilyMembers : <Animal> <Animal>  
 cohabitants : <Animal> <Animal>  
 likesAsFriend : <SentientAnimal> <SentientAnimal>  
 loves : <SentientAnimal> <Agent>

maritalStatus : <Person> <MaritalStatusOfPeople>  
 MaritalStatusOfPeople

socialStatus : <Person> <SocialStatusAttributeType>  
 SocialStatusAttribute  
 SocialStatusAttributeType  
 SocialAttributeType  
 EducationLevelAttribute  
 schooling : <Person> <EducationalOrganization>  
 StudentStatusAttribute  
 educationLevel : <Person> <EducationLevelAttribute>  
 FieldOfStudy  
 ScientificFieldOfStudy  
 Religion

Title  
 CourtesyTitle  
 titleOfPerson-String : <Person> <CharacterString>  
 firstName : <Person> <HumanGivenNameString>  
 middleName : <Person> <HumanGivenNameString>  
 lastName : <Person> <HumanFamilyNameString>  
 ProperNameString  
 HumanNameString  
 HumanGivenNameString  
 HumanFamilyNameString  
 salutation : <Person> <CourtesyTitle>  
 nameOfAgent : <Agent> <ProperNameString>

ethnicity : <Person> <EthnicGroupType>  
 skinColor : <Person> <HumanSkinColor>  
 PersonalityAttribute  
 HumanCultureType  
 EthnicGroupType  
 Nationality



SocialOccurrence  
SociabilityBasedAction  
PublicEvent  
SocialGathering  
MeetingTakingPlace  
Transaction  
Party-Celebration  
SocialRitual  
Ritual  
  
RudeAction  
HostileSocialAction  
  
TransferringPossession  
GreetingSomeone  
MeetingSomeone  
VisitingSomeone  
  
Competition  
AthleticActivity  
Bartering  
MakingSomethingAvailable  
AppropriatingSomething  
ObtainingPermission  
  
CommercialActivity  
Advertising  
Negotiating  
BusinessRelationshipActivity  
  
PhysicallyAttackingAnAgent  
Battle  
WagingWar  
DisputeEvent  
Trial

## CYC Social Vocabulary. Annotated List

Excerpted from <http://www.cyc.com/cyc-2-1/vocab/social-vocab.html>. Updated 1997-10-14, accessed 2001-4-15

### **#\$controls : <#\$Agent> <#\$Individual>**

(#\$controls X Y) represents that assertion that agent X controls the object Y, in one of the following 2 senses: X can influence (prohibit, enable or constrain) the behavior of Y; or else X can at least influence (prohibit, enable or constrain) the behavior of other #Agents in/concerning Y. For example, Fred may control his horse directly, forcing it to do things, or not do them; and he also could control the horse indirectly, by deciding who else has access to and use of that horse. Control of one agent over another agent is rarely total, of course, so this predicate is most likely to apply to a Y which is a non-living possession, and/or to apply in a very narrow context. X's control over Y is usually either actual (de facto) control or legal (de jure) control. It is usually #cotemporal, meaning that some time slice of X controls the same temporal time slice of Y.

isa: #BinaryPredicate #CotemporalObjectsSlot

genlPreds: #PositiveVestedInterest #Cotemporal

some more specialized predicates: (1 additl more specialized public predicate, 4 unpubl. ones)

### **#\$SocialBeing**

The collection of beings whose existence is accepted by some social system. (Thus, the elements of #SocialBeing will vary with social contexts.) Social beings are entities able to perform social roles in the system that recognizes them. #SocialBeing includes elements of #Organization (e.g., the #QueensGuard) as well as the elements of #LegalAgent (in that system), so, for example, in modern industrial social systems, the elements of #LegalCorporation and #Person are instances of #SocialBeing.

isa: #ExistingObjectType

genls: #IntelligentAgent

some subsets: #JudicialAgent #MedicalCareProvider #Family-SocialEntity #LegalAgent

#Organization #Court-Judicial #MedicalCareProfessional #MedicalCareOrganization

#GeopoliticalEntity #SoleProprietorship #Partnership #LegalCorporation

#LegalGovernmentOrganization #Person #ManufacturingOrganization (plus 157 more public subsets, 1992 unpublished subsets)

### **#\$affiliatedWith : <#\$Agent> <#\$Agent>**

...

### **#\$acquaintedWith : <#\$IndividualAgent> <#\$IndividualAgent>**

(#\$acquaintedWith AGENT1 AGENT2) means the #IndividualAgent AGENT1 is acquainted with the #IndividualAgent AGENT2 (in the minimal sense that AGENT1 has come into physical or conversational contact with AGENT2, or that they have somehow knowingly communicated with each other). This typically means that each #IndividualAgent is aware of some facts about the other. In cases where one of the #IndividualAgents is sentient, this typically includes the ability of this agent to recognize the other by appearance, voice, scent, or some other physical attribute.

isa: #CotemporalObjectsSlot #BinaryPredicate #Predicate #ReflexiveBinaryPredicate

#\$SymmetricBinaryPredicate

genlPreds: #Cotemporal

some more specialized predicates: #boss #siblings #mate #cohabitants #likesAsFriend

#spouse #cohabitingFamilyMembers #loves #friends (plus 11 unpubl. more spec. pred.)

### **#\$Workplace**

The collection of places where people customarily work (not the employing organizations). #Workplace includes offices, restaurant buildings, construction sites, agricultural sites, the SpaceNeedle, etc. Some places may be #Workplaces only during a small part of their existence (a piece of residential property while the house is being built, perhaps); some may almost always be #Workplaces (grocery store buildings, office buildings, smithies, hospitals, etc.).

isa: #ContactLocationType

genls: #HumanlyOccupiedSpatialObject #PhysicalContactLocation

some subsets: (10 unpublished subsets)

### **#AdultFemalePerson**

The collection of all women; i.e., #Persons who are adult and female

isa: #ExistingObjectType

genls: #HumanAdult #FemalePerson

### **#HumanInfant**

The collection of #Persons in the infant stage of life. Functionally, this ends when the infant learns to walk (even just toddle) and/or talk (even a few words)... or, at latest, when the person's age greatly exceeds that at which most people develop those skills. Generally, this means that it spans the period from birth to about 12 - 18 months old. One of the subsets of this collection is #NewbornBaby.

isa: #ExistingObjectType #TemporalObjectType

genls: #HumanChild

some subsets: (3 unpublished subsets)

### **#HumanChild**

The collection of all #Persons in the childhood stage of life. Functionally, this ends when the child begins to take responsibility for themselves, work, have children of their own,... or, at latest, when the person's age greatly exceeds that at which most people reach those milestones. Generally, this means that it spans the period from birth to teenage years. This is highly dependent on context, of course; childhood in Shakespeare's culture ended around age 12.

isa: #ExistingObjectType #TemporalObjectType

genls: #JuvenileAnimal

some subsets: (1 more public subset, 8 unpublished subsets)

### **#HumanAdult**

The collection of human beings old enough to participate as independent, mature members of society. Since different societies have different age or maturity requirements for people to be considered adults, different axioms in various society-specific microtheories express these requirements. For most modern, Western, middle-class,... purposes, e.g., the current view is that anyone over 18 is an adult. In many cultures, adulthood occurs when one reaches puberty. Adulthood is #contiguousAfter childhood; that is, a #Person is a #HumanChild for a while, and then is a #HumanAdult.

isa: #ExistingObjectType #TemporalObjectType

genls: #AdultAnimal #Person

some subsets: #AdultFemalePerson (plus 16 unpublished subsets)



## Additional schemes

- Bloom**      **Taxonomy of educational objectives** 1956 (1 copy in the cataloging laboratory) (LB17.B55.1956), a summary at  
<http://www.unesco.org/webworld/ramp/html/r8810e/r8810e0e.htm>  
<http://websites.ntl.com/~james.atherton/learning/bloomtax.htm>,  
<http://sweep.riv.csu.edu.au/td/bloom.html>,  
<http://faculty.washington.edu/~krumme/guides/bloom.html>
- SOC**        **Standard Occupational Classification** 2000  
Bureau of Labor Statistics (BLS) + other agencies  
[http://stats.bls.gov/soc/soc\\_home.htm](http://stats.bls.gov/soc/soc_home.htm)  
The SOC is augmented by the **Occupational Information Network (O\*NET)**, a database with additional occupational titles, definitions, and features of occupations.  
<http://www.doleta.gov/programs/onet>
- CSDGM**     **Content Standard for Digital Geospatial Metadata** 1998  
Federal Geographic Data Committee (FGDC)  
<http://www.fgdc.gov/metadata/contstan.html>
- ERIC**        **Education Resources Information Center Thesaurus.** 13th ed.  
<http://searcheric.org/>

# Yahoo

**The Yahoo classification.** Web pages [www.yahoo.com](http://www.yahoo.com)

**Assignment 14****Assigned: Nov. 11****Due: Nov. 18/Dec. 2****Yahoo classification**

Time: 6 hours (first half - 3.5 hours; second half - 2.5 hours)

**Materials needed/available****0 The Yahoo classification itself as available on the Yahoo Web site****Materials attached in print.**

- 1 Figures 1a and b. **Yahoo Home (first summary): top level classes** (called *categories* in Yahoo) in the original Yahoo alphabetical arrangement and in a revised meaningful arrangement.
- 2 Figures 2a and b. **First-level breakdown of *Health*** in the original Yahoo alphabetical arrangement and in a revised meaningful arrangement.
- 3 Figure 3. **Second summary** (first two levels of the hierarchy).
- 4 Figure 4. **Excerpts from the Yahoo classification** designed to illustrate the structure of the classification, particularly its treatment of *Education* and *Transportation*.  
  
This list is by no means complete with respect to Yahoo classes that deal with *Education* or *Transportation*. It does illustrate patterns that are repeated throughout the classification. Some Yahoo classes do not have the number of associated Web pages; I sometimes used a simplified method for copying from the Yahoo displays that did not carry the number with it.
- 5 Figure 5. **Yahoo top level compared with LCC and Dewey top level.**
- 6 Figure 6. **Yahoo top level compared with the subdivisions under *State* and *City*.**
- 7 Figure 7. **Subdivisions of some Yahoo categories for comparison and analysis.**

**The Yahoo worksheet starts on the back of this page.****Plan to be online to Yahoo while doing the worksheet**, even though some of it needs just the attached materials. You need the Yahoo Web site for indexing and query formulation.

You can prepare your worksheet answers on a word processor, which gives you the opportunity to copy and paste examples from Yahoo.

**Deliverables**The filled-in worksheet (Examine the *Outline for the analysis of subject access vocabularies*, which is already filled in)

The filled-in document forms and query forms

## Yahoo Worksheet

Name: \_\_\_\_\_

### 60 min. 1. General layout of the classification. Formal structure

#### Familiarize yourself with the general layout of the classification.

- Look at the top level on the Yahoo home page and click down to *Education*.
- Look at Figures 1 and 2 and skim Fig. 3. the Yahoo classification second summary.
- Skim through Figure 4, the excerpt from the Yahoo classification; note the many places in which *education* and *transportation* concepts appear.

**Important principle:** In Yahoo, as in the Library of Congress Classification, the meaning of a category (class) is always defined by its total context. Thus in

Education

. K-12

. . Academic Competitions

. . . Debate

. . . . Clubs, Teams, and Societies

The last category means

*Clubs, Teams, and Societies* that engage in *Debates* staged as *Academic Competitions* for students in the *K-12* level of *Education*

If you click down to this level, this category will be shown with its **full caption:**

Home > Education > K-12 > Academic Competitions > Debate >  
**Clubs, Teams, and Societies**

### What is the degree of precombination in Yahoo?

#### Search for categories in Yahoo

A search for one or more words in Yahoo returns both categories whose caption contains the word(s) and individual sites. So Yahoo functions as a kind of **descriptor find index**. But, as a rule, retrieval of categories is based on the words in the full caption; while these words often reflect the conceptual components of the category, there are many cases where they do so only incompletely. See the facing page for an example. Try a search for *vehicles*; do you find *ships* or *aircraft*? To look at an exception, search for *driver education*.

Some good searches to try to explore the structure of the Yahoo classification

CDs, Records, and Tapes	Mexican Americans
Transportation	Bridge ( a homonym)
Teaching and Learning Aids	Humor

You can restrict your search to a category.



**Yahoo category search (descriptor find index): Words vs concepts**

<p><b>Search 1: law education Canada</b> (implied AND)</p> <p>Yahoo! Category Matches (1 - 4 of 4)</p> <p>Regional &gt; Countries &gt; Canada &gt; Provinces and Territories &gt; Saskatchewan &gt; Cities &gt; Saskatoon &gt; <b>Education</b> &gt; College and University &gt; University of Saskatchewan &gt; Departments and Programs &gt; College of <b>Law</b> [Note: <i>College</i> rather than <i>School</i>]</p> <p>Regional &gt; Countries &gt; Canada &gt; Provinces and Territories &gt; Ontario &gt; Counties and Regions &gt; Frontenac &gt; Cities &gt; Kingston &gt; <b>Education</b> &gt; College and University &gt; Queen's University &gt; Departments and Programs &gt; School of <b>Law</b></p> <p>Regional &gt; Countries &gt; Canada &gt; Provinces and Territories &gt; British Columbia &gt; Counties and Regions &gt; Capital &gt; Cities &gt; Victoria &gt; <b>Education</b> &gt; College and University &gt; University of Victoria &gt; Departments and Programs &gt; <b>Law</b> [Note: Simply a <i>department</i>]</p> <p>Regional &gt; Countries &gt; Canada &gt; Provinces and Territories &gt; Alberta &gt; Counties and Districts &gt; Edmonton &gt; Cities &gt; Edmonton &gt; <b>Education</b> &gt; College and University &gt; University of Alberta &gt; Departments and Programs &gt; <b>Law</b> [Note: Simply a <i>department</i>]</p>	<p><b>Search 2: law school Canada</b></p> <p>Yahoo! Category Matches (1 - 2 of 2)</p> <p>Regional &gt; Countries &gt; Canada &gt; Provinces and Territories &gt; Ontario &gt; Counties and Regions &gt; Frontenac &gt; Cities &gt; Kingston &gt; Education &gt; College and University &gt; Queen's University &gt; Departments and Programs &gt; <b>School of Law</b></p> <p>Regional &gt; Countries &gt; Canada &gt; Government &gt; <b>Law</b> &gt; <b>Law Schools</b></p>
--	---

In the Yahoo categories, all Search 1 retrievals happen to be also relevant for Search 2. (There could be one or more categories on *law education in Canada* generally, but there are none.). All Search 2 retrievals are by definition relevant for Search 1.

When running the same searches with *US*, Search 1 finds 62 categories as follows

- 44 have both *law* and *school* and are thus also found by Search 2
- 15 have *law* and *college* (These are relevant for Search 2 but not found by it)
- 3 have *law* and some other term, such as *center*

Search 2 finds 44 categories; they all happen to have *education*, and thus were found in Search 1 as well.

Searching for *law school* without restrictions would find categories not found by *law education* (such as Home > Government > Law > Law Schools; this category has under it categories that say *College of Law*) and vice versa (such as Home > Government > Law > Continuing Legal Education).

Note: Searching for *United States* finds no categories; you need to search for *US*.

### Explore the Yahoo “multi-tree”

As you know from Chapters 14 and 15, a compound concept fits in many places in a hierarchy. Put differently, in the Yahoo subject directory, a precombined category should be reachable thorough multiple paths down. How does Yahoo handle this problem? Probe the following example:

Click down to

Home > Education > K-12  
**Schools**

Click on Christian@

Try Home > Education > K-12 > By Region > Countries > France > Cities

Probe some on your own

Summarize your observations. What does the @ mean?

**A**

### Explore the nature of subordinate categories

In Example 1 on the facing page, why are the subordinate categories in the first group narrower than *Canada*, why the subordinate categories in the second group? What is the difference? (Hint: Remember *concepts narrower due to autonomous subdivision* and *concepts narrower due to combination*, p. 264 and 270 in *Organizing Information*)

In Example 2 (which is not quite as clear cut),

**mark with T** the subordinate categories that are transportation-specific  
**mark with G** the subordinate categories that are the broad concept of *transportation* I combined with some general concept not specific to the transportation domain.

Note your observations

**B**

**Example 1**

Home > Regional > Countries >

**Canada**

- . Cities (1697)
- . Provinces and Territories (89205)
  - . Alberta (7683)
  - . British Columbia (19882)
  - . Manitoba (3680)
  - ...
- . Arts and Humanities (2425)
- . Business and Economy (20557)
- . Computers and Internet (333)
- . Country Guides (21)
- . Education (1236)
- ...

**Example 2**

Home > Business and Economy >

**Transportation**

- . Auto-Free Transportation (23)
- . Aviation (513)
- . Buses (26)
- . Companies@
- . Employment (5)
- . Government Agencies (62)
- . Highways and Roads (127)
- . History (5)
- . Institutes (44)
- . Intelligent Transportation Systems (25)
- . Libraries (7)
- . Mass Transit (59)
- . Web Directories (3)

The remaining questions deal with content aspects of the Yahoo classification

30 min **2. Begin to develop a meaningful arrangement of the categories one level below *Education***

You can use the meaningful arrangement of the subcategories of *Health* (Figure 2b) as a general model. Just listing some facets with sample terms under them would be an acceptable answer

**C** Put your answer on a separate sheet

15 min **3. Compare Yahoo with Dewey and Library of Congress Classification**

Look over Figure 5 and get a sense of how different topics are treated in these classifications.

Why is *literature* given more prominence in DDC and LCC than in Yahoo? Can you find a general principle that would explain the differences in emphasis in DDC and LCC on the one hand and the Yahoo Classification on the other?

**D**

15 min **4. Compare Yahoo Home, *State* subdivision and *City* subdivision**

Look over Figure 6 and briefly describe the differences you see between Yahoo Home and the *State* subdivision and the *State* subdivision and the *City* subdivision

**E**

30 min **5. Examine some principles Yahoo uses when designing subdivisions**

Figure 7 gives a number of examples of category subdivisions. Where there are two groups, can you tell the difference between them?

Write your observations on any two of the examples or state a general principle.

**F**

**60 min 6. Overall facet analysis of the Yahoo classification**

Identify entity types / facets that occur throughout the Yahoo classification, preferably with some frequently occurring concepts under each. You can also mention concepts that occur as components in many places but that you cannot assign to a facet. Your listing would be the beginning of a faceted *core classification* for Yahoo (see Organizing Information, p. 299 and Section 15.6, p. 322-323).

It would be interesting to find out how many elemental concepts are in this core classification and how many precombined categories Yahoo has in its *extended classification*.

**G**

## Outline for the analysis of subject access vocabularies

For some items, a section number from Soergel, Organizing information is given in []

### 1. Purpose

#### 1.1 Information system or type of information system in which to be used

*Bibliographic information system: Organize Web pages*

#### 1.2 Intended for controlled vocabulary indexing or query term expansion [Ch. 12, Introduction]

#### 1.3 Type of file and search mechanism for which originally designed

Shelving  Card catalog  Online system  (*Web subject directories: Similar to shelving, but multiple locations for each class and multiple entries for each page*)

### 2. Coverage and designation of concepts. Coverage and format of terms

#### 2.1 Concepts: Scope, breadth of coverage. Recency of concepts

*Universal — covers all of knowledge. But focus on Western culture, esp. US, on topics in HTML Web pages.,*

#### 2.2 Concepts: Specificity, depth of coverage. (Section 16.2.2). Coverage at each level of specificity.

*Medium specificity. Would need closer analysis by subject area. Geographic names quite specific.*

#### 2.3 Are all needed facets included? Concepts formed in semantic factoring and facet analysis? (S.a. 3.1)

*Answer would require extensive analysis. Many general concepts are visible as one looks at subdivisions that repeat in many places, but no list of these. Bound to English language; if there is no English term for a general concept, that concept is unlikely to show up explicitly in the Yahoo Classification.*

#### 2.4 Nature of notation (if none, state that). [Section 15.5.2] *None.*

#### 2.5 Terms: Completeness of coverage (completeness of lead-in vocabulary). Recency of terms

*Only descriptors, no lead-in terms. Many terms are quite recent (very fast update).*

#### 2.6 Form of terms: Consistency, adherence to common usage. Terms seem appropriate. Most terms appear to be taken directly from generally used language, except for a few phrases like Arts and Humanities.

### 3. Terminological and conceptual analysis and conceptual structure.

#### 3.1 Quality of conceptual structure (14): Facet analysis. Types and degree of differentiation of conceptual relationships included. For each type indicate the completeness of inclusion. (Fill in 3.1.1 - 3.1.3)

##### 3.1.1 Expression of concepts through elemental concepts (closely related to definition)

*Category names are formed by stringing together terms that designate the category's conceptual components; to that extent, compound concepts are expressed through elemental concepts. However, compounds expressed through an accepted term in English, such as Ship, are not expressed through their components.*

##### 3.1.2 Hierarchical relationships (polyhierarchy) (Shown by arrangement or Broader Term / Narrower Term X-ref)

*Polyhierarchical; a category has a home place but may appear in many places in the hierarchy. From any place one get to the home place by a click, but the other places where the category appears are not indicated*

##### 3.1.3 Associative relationships. (Implied by physical proximity in the arrangement or explicit Related Term X-ref)

*None by the format. However, some of the hierarchical relationships should in fact be associative.*

#### 3.2 Quality of definitions, explications, scope notes (correctness, detail, clarity).

*No definitions.*

#### 3.3 Completeness of terminological relationships: Does the vocabulary contain terms that are synonymous or quasi-synonymous without indicating the relationship?

*Yahoo does not attempt to cover terminology beyond the category names.*

4. **Use of precombination in the index language** (concerns both 2 and 3) [14, 15, esp. 15.4]

4.1 To what degree are descriptors precombined?

*Yahoo categories are highly to very highly precombined.*

4.2 To what extent are precombined descriptors enumerated and/or given in the alphabetical index?

*Precombined descriptors are enumerated in the Yahoo directory. It is not known whether Yahoo indexers have schedules of just the elemental concepts to index from.*

To what extent can the indexer build additional precombined descriptors?

*Probably new categories built from existing components are added all the time either by indexers or by an editor based on indexer suggestions.*

Are precombined descriptors designated by an independent symbol or a string of symbols? Combination order free or fixed? To what extent do the components of a precombined descriptor determine its place in the arrangement? (Relates also to 5) (Section 15.5.2)

*By a string of symbols, the terms for the individual components. Exception: Words/terms like Ship that designate a compound concept. Combination order is free, with some apparent rules. The components completely determine the place of a precombined descriptor built by the indexer.*

5. **Access and display. Format of presentation of the vocabulary**

Consider for each format access/retrieval by concepts versus access/retrieval by terms.

Access can be provided through arrangement in a printed document or through a computer search system.

5.1 **Format of printed document** (Fill in 5.1.1- 5.1.3) *No print version. Comments here refer to online display.*

5.1.1 Overall format: Thesaurus parts and information given in each, connections between them. Is the overall format clear and helpful for finding the appropriate concepts and terms or notations in indexing and query formulation?

*A menu tree, walking down one level at a time.*

5.1.2 Display of conceptual relationships (Broader Term, Narrower Term, Related Term)

- through linear arrangement or graphical display [Section 15.5.2]

*In the subject directory "multi-tree", in which each descriptor (category) can appear in multiple places*

- through cross-references [Section 14.1]

*No cross-references, other than a category appearing in multiple places.*

- through descriptor-find index [Section 15.5.1]

*The search function provides a descriptor find index of sorts.*

How well does the display reflect the conceptual analysis, e.g., sequence of concepts on the same hierarchical level (sequence of the children of a concept, that is, the concepts one level further down).

*The sequence of categories at each level is usually strictly alphabetical, sometimes divided into groups based on meaningful criteria.*

5.1.3 Display of terminological relationships (Synonymous Term)

*Terminological relationships are not included..*

5.2 **Access through computer system.** Navigation. Format of on-line displays

*Described in 5.1.*



## Instructions for indexing documents with Yahoo

Give up to three categories per document. Yahoo does assign multiple categories to a document since a subject directory does not have the one-place-only constraints imposed by the physical arrangement of documents. Still only very important categories should be assigned (high-threshold, low-exhaustivity indexing). (I do not know the rules for number of categories that is given to Yahoo indexers.) Use the most specific category in each case. Give the categories in the format generally accepted on the Web (you can omit Home):

Home > Education > K-12 > Schools > Elementary Schools

Browse through the subject directory or use search to identify the categories under which a document should be found and under which it should therefore be indexed.

## Instructions for formulating queries with Yahoo

Try to list **all** categories where one should look for relevant documents; if there are more than 10 categories, just give a representative sample (enough to demonstrate that you know how to find all categories throughout the Yahoo classification).

Note: A broad category implies all the narrower categories; no need to list these narrower categories, they can be readily seen from the Yahoo multi-tree.

The query formulation is the OR combination of all the categories in your list. AND combinations of categories would run counter to the subject directory approach to searching; whether this approach makes best use of computer capabilities is a different question.

You can formulate each query in terms of elemental concepts and then use these to find categories (see Worksheet, Task 6).

<p><b>Arts &amp; Humanities</b> Literature, Photography ...</p> <p><b>Business &amp; Economy</b> B2B, Finance, Shopping, Jobs ...</p> <p><b>Computers &amp; Internet</b> Internet, WWW, Software, Games ...</p> <p><b>Education</b> College and University, K-12 ...</p> <p><b>Entertainment</b> Cool Links, Movies, Humor, Music ...</p> <p><b>Government</b> Elections, Military, Law, Taxes ...</p> <p><b>Health</b> Medicine, Diseases, Drugs, Fitness ...</p>	<p><b>News &amp; Media</b> Full Coverage, Newspapers, TV...</p> <p><b>Recreation &amp; Sports</b> Sports, Travel, Autos, Outdoors ...</p> <p><b>Reference</b> Libraries, Dictionaries, Quotations ...</p> <p><b>Regional</b> Countries, Regions, US States ...</p> <p><b>Science</b> Animals, Astronomy, Engineering ...</p> <p><b>Social Science</b> Archaeology, Economics, Languages ...</p> <p><b>Society &amp; Culture</b> People, Environment, Religion ...</p>
--	---

Figure 1a. **Yahoo classification. Home** (first summary)

<p><b>Reference and General Interest</b></p> <p><b>Reference</b> Libraries, Dictionaries, Quotations ...</p> <p><b>Computers &amp; Internet</b> Internet, WWW, Software, Games ...</p>	<p><b>Subjects</b></p> <p><b>Science</b> Animals, Astronomy, Engineering ...</p> <p><b>Health</b> Medicine, Diseases, Drugs, Fitness ...</p> <p><b>Social Science</b> Archaeology, Economics, Languages ...</p> <p><b>Society &amp; Culture</b> People, Environment, Religion ...</p> <p><b>Government</b> Elections, Military, Law, Taxes ...</p> <p><b>Business &amp; Economy</b> B2B, Finance, Shopping, Jobs ...</p>
<p><b>News &amp; Media</b> Full Coverage, Newspapers, TV...</p> <p><b>Entertainment</b> Movies, Music, Humor, Cool Links ...</p> <p><b>Recreation &amp; Sports</b> Sports, Travel, Autos, Outdoors...</p>	<p><b>Education</b> College and University, K-12 ...</p> <p><b>Arts &amp; Humanities</b> Literature, Photography ...</p>
<p><b>Regional</b> Countries, Regions, US States ...</p>	

Figure 1b. **Yahoo Classification. Home. Meaningful arrangement**

**Home > Health****Categories**

Alternative Medicine (480)	Men's Health (30)
Business to Business@	Mental Health (682)
Chats and Forums (52)	Midwifery (60)
Children's Health (153)	News and Media (201)
Conferences (19)	Nursing (431)
Death and Dying@	Nutrition (207)
Dentistry@	Organizations (21)
Disabilities@	Pet Health@
Diseases and Conditions (7392)	Pharmacy (1096)
Education (39)	Procedures and Therapies (292)
Emergency Services (236)	Public Health and Safety (740)
Employment (108)	Reference (93)
Environmental Health (194)	Reproductive Health (659)
First Aid (14)	Senior Health (79)
Fitness (166)	Sexuality@
General Health (81)	Shopping and Services@
Health Administration (65)	Teen Health (13)
Health Care (356)	Traditional Medicine (179)
Health Sciences (26)	Travel Health and Medicine (16)
Hospitals and Medical Centers (38)	Web Directories (50)
Institutes (34)	Weight Issues (77)
Law@	Women's Health (153)
Long Term Care (116)	Workplace (6)
Medicine (4955)	

Figure 2a. **Yahoo classification. Health.**

**Home > Health****Categories**

<p><b>Reference</b>  Reference (93)  Web Directories (50)  News and Media (201)  Chats and Forums (52)</p> <p><b>Health Sciences Fields</b>  Health Sciences (26)  Medicine (4955)  Dentistry@  Nursing (431)  Midwifery (60)  Pharmacy (1096)</p> <p>Traditional Medicine (179)  Alternative Medicine (480)</p> <p><b>Individual health condition</b>  Diseases and Conditions (7392)  Disabilities@  Fitness (166)  Nutrition (207)  Weight Issues (77)  Reproductive Health (659)  Sexuality@  Death and Dying@  Mental Health (682)</p> <p><b>Procedures and Therapies (292)</b></p>	<p><b>Health by place</b>  General Health (81)  Public Health and Safety (740)  Environmental Health (194)  Workplace (6)  Travel Health and Medicine (16)</p> <p><b>Health by population group</b>  Children's Health (153)  Teen Health (13)  Women's Health (153)  Men's Health (30)  Senior Health (79)  Pet Health@</p> <p><b>Health Care (356)</b>  Emergency Services (236)  First Aid (14)  Long Term Care (116)</p> <p><b>Health care organization</b>  Hospitals and Medical Centers (38)  Institutes (34)  Organizations (21)  Conferences (19)  Health Administration (65)</p> <p>Shopping and Services@  Business to Business@  Law@</p> <p>Education (39)  Employment (108)</p>
--	---

Figure 2b. Yahoo classification. Health. Meaningful arrangement.

### Figure 3. Yahoo classification second summary

#### Arts & Humanities

- SN Literature, Photography . . .
- . By Region (131)
- . Art History (742)
- . Artists (3474)
- . Arts Therapy@
- . Awards (16)
- . Booksellers@
- . Censorship (17)
- . Chats and Forums (45)
- . Companies@
- . Crafts (796)
- . Criticism and Theory (30)
- . Cultural Policy@
- . Cultures and Groups (483)
- . Design Arts (5177)
- . Education (631)
- . Employment (44)
- . Events (198)
- . Humanities (45830)
- . Institutes (38)
- . Museums, Galleries, & Cntrs (1016)
- . News and Media (260)
- . Organizations (360)
- . Performing Arts (6065)
- . Reference (35)
- . Thematic (456)
- . Visual Arts (12134)
- . Web Directories

#### Business & Economy

- SN B2B, Finance, Shopping, Jobs . . .
- . Business to Business (268905)
- . Shopping and Services (378101)
- . Business Libraries (23)
- . Business Schools@
- . Chats and Forums (24)
- . Classifieds (3632)
- . Consortia (42)
- . Consumer Advocacy and Information@
- . Conventions and Conferences (38)
- . Cooperatives (24)
- . Directories (347)
- . Economics@
- . Education (809)
- . Electronic Commerce (215)
- . Employment and Work (1742)
- . Ethics and Responsibility (46)
- . Finance and Investment (1804)
- . Global Economy (287)
- . History (20)
- . Intellectual Property@
- . Labor (725)
- . Law@
- . Magazines (130)
- . Management Science (203)
- . Marketing and Advertising (350)
- . News and Media@
- . Organizations (11880)
- . Quality Standards@
- . Real Estate (391)
- . Small Business Information (299)
- . Statistics and Indicators (5)
- . Taxes@
- . Television@
- . Trade (422)
- . Transportation (2094)

#### Computers & Internet

- SN Internet, WWW, Softw., Games . . .
- . Business to Business@
- . Shopping and Services@
- . Art@
- . Bibliographies (6)
- . Communications & Netw. (1128)
- . Computer Science@
- . Contests (26)
- . Conventions and Conferences@
- . Countries, Cultures, & Groups (38)
- . Cyberculture@
- . Data Formats (380)
- . Desktop Customization@
- . Desktop Publishing (53)
- . Dictionaries (32)
- . Employment@
- . Ethics (18)
- . Games@
- . Graphics (308)
- . Hardware (2292)
- . History (85)
- . Humor@
- . Industry Information@
- . Internet (5999)
- . Magazines@
- . Mobile Computing (64)
- . Multimedia (673)
- . Music@
- . News and Media (203)
- . Operating Systems@
- . Organizations (93)
- . People (119)
- . Personal Computers@
- . Product Reviews (2974)
- . Programming Languages (1515)
- . Science and Technology Policy@
- . Security and Encryption (589)
- . Semiconductors@
- . Software (6133)
- . Standards (45)
- . Supercomputing and Parallel Computing@
- . Technical Guides and Support (45)
- . Telecommunications@
- . Training@
- . User Groups@
- . Web Directories (14)
- . World Wide Web@
- . Year 2000 Problem (248)
- . Cnet
- . ZDNet

## Education

- SN College and University, K-12
- ...
- . Browse by Region (170)
- . By Culture or Group (404)
- . By Subject (12)
- . Academic Competitions (80)
- . Adult and Continuing Education (330)
- . Bibliographies (4)
- . Bilingual (23)
- . Career and Vocational (234)
- . Chats and Forums (40)
- . Companies@
- . Conferences (50)
- . Correctional@
- . Disabilities@
- . Distance Learning (491)
- . Early Childhood Education (92)
- . Employment (146)
- . Equity (27)
- . Financial Aid (396)
- . Government Agencies (78)
- . Graduation (52)
- . Higher Education (16594)
- . Instructional Technology (341)
- . Journals (38)
- . K-12 (54618)
- . Literacy (12)
- . News and Media (83)
- . Organizations (3094)
- . Policy (52)
- . Programs (335)
- . Reform (73)
- . Special Education (172)
- . Standards and Testing (59)
- . Statistics (6)
- . Teaching (63)
- . Theory and Methods (672)
- . Web Directories (47)

## Entertainment

- SN Cool Links, Movies, Humor, Music ...
- . Actors and Actresses (11355)
- . Amusement & Theme Parks (396)
- . Books and Literature@
- . Chats and Forums (92)
- . Comedy (794)
- . Comics and Animation (4979)
- . Consumer Electronics (588)
- . Contests, Surveys, and Polls (424)
- . Cool Links (1837)
- . Employment (366)
- . Entertainment and Media Production@
- . Events (214)
- . Food and Drink@
- . Gambling@
- . Games@
- . Genres (1363)
- . History (15)
- . Humor (4857)
- . Magic (296)
- . Movies and Film (19920)
- . Music (82025)
- . News and Media (394)
- . Organizations (11)
- . Performing Arts@
- . Radio@
- . Randomized Things (76)
- . Reviews (39)
- . Shopping and Services@
- . Television@
- . Trivia (109)
- . Virtual Cards (1019)
- . Web Directories (36)
- . Webisodes (94)
- . X of the Day, Week, etc. (171)

## Government

- SN Elections, Military, Law, Tax
- . Countries (147)
- . Chats and Forums (7)
- . Civic Participation (27)
- . Conventions and Conferences (17)
- . Documents (26)
- . Embassies and Consulates (99)
- . Ethics (14)
- . Intelligence (125)
- . International Organizations (531)
- . Law (2671)
- . Military (867)
- . National Symbols and Songs (50)
- . News and Media (11)
- . Politics (11446)
- . Public and Civil Service (8)
- . Research Labs (26)
- . Statistics (40)
- . Student Government@
- . Taxes (53)
- . US Government (11649)
- . Web Directories (14)

## Health

- SN Medicine, Diseases, Drugs, Fitness ...
- . Alternative Medicine (515)
- . Business to Business@
- . Chats and Forums (56)
- . Children's Health (186)
- . Conferences (18)
- . Death and Dying@
- . Dentistry@
- . Disabilities@
- . Diseases and Conditions (8505)
- . Education (62)
- . Emergency Services (536)
- . Employment (118)
- . Environmental Health (200)
- . First Aid (12)
- . Fitness (201)
- . General Health (90)
- . Health Administration (66)
- . Health Care (347)
- . Health Sciences (27)
- . Hospitals and Medical Centers (44)
- . Institutes (34)
- . Law@
- . Long Term Care (109)
- . Medicine (5071)
- . Men's Health (37)
- . Mental Health (753)
- . Midwifery (56)
- . News and Media (199)
- . Nursing (459)
- . Nutrition (218)
- . Organizations (21)
- . Pet Health@
- . Pharmacy (1231)
- . Procedures and Therapies (458)
- . Public Health and Safety (2207)
- . Reference (98)
- . Reproductive Health (713)
- . Senior Health (85)
- . Sexuality@
- . Shopping and Services@
- . Teen Health (23)
- . Traditional Medicine (197)
- . Travel Health and Medicine (24)
- . Web Directories (54)
- . Weight Issues (90)
- . Women's Health (174)
- . Workplace (67)

**News & Media**

- SN Full Coverage, Newspapers, TV . . .
- . By Region (21280)
- . Columns and Columnists (276)
- . Commercial Services@
- . Content Ratings@
- . Industry Information (821)
- . Internet Broadcasts (396)
- . Journals (33)
- . Magazines (3845)
- . Newspapers (8447)
- . Photojournalism@
- . Radio (9428)
- . Television (15866)
- . Web Directories (99)
- . Arts and Humanities@
- . Automotive@
- . Business (133)
- . College and University (1003)
- . Computers and Internet@
- . Crime@
- . Cultures and Groups (15)
- . Disabilities@
- . Education@
- . Entertainment@
- . Environment and Nature@
- . Good News (9)
- . Government@
- . Health@
- . History@
- . Home and Garden@
- . Humor (218)
- . Law@
- . Outdoors@
- . Personalized News (14)
- . Philanthropy@
- . Politics@
- . Real Estate@
- . Religion@
- . Science@
- . Sexuality@
- . Sports@
- . Technology (69)
- . Traffic and Road Conditions@
- . Transportation@
- . Travel@
- . Weather (1088)
- . Weird News (20)
- . World (71)
- . ABC News
- . BBC News
- . CNN
- . Fox News
- . MSNBC

**Recreation & Sports**

- SN Sports, Travel, Autos, Outdoors . . .
- . Amusement and Theme Parks@
- . Automotive (5657)
- . Aviation (810)
- . Booksellers@
- . Chats and Forums (7)
- . Cooking@
- . Dance@
- . Employment (7)
- . Events (10)
- . Fitness@
- . Gambling (315)
- . Games (17999)
- . Hobbies (3044)
- . Home and Garden (760)
- . Magazines (63)
- . Motorcycles@
- . Outdoors (9884)
- . Pets@
- . Sports (47097)
- . Television@
- . Toys (948)
- . Travel (112681)

**Reference**

- SN Libraries, Dictionaries, Quotations . . .
- . Acronyms and Abbreviations (25)
- . Almanacs (13)
- . Arts and Humanities@
- . Ask an Expert (556)
- . Bibliographies (7)
- . Booksellers@
- . Calendars (81)
- . Codes (24)
- . Dictionaries (153)
- . Directories (3)
- . Encyclopedia (23)
- . English Language Usage@
- . Environment and Nature@
- . Etiquette@
- . FAQs (18)
- . Finance and Investment@
- . Flags (26)
- . General (13)
- . Geographic Name Servers@
- . Health@
- . Journals@
- . Libraries (4803)
- . Maps@
- . Measurements and Units@
- . Music@
- . Parliamentary Procedure (13)
- . Patents@
- . Phone Numbers and Addresses (166)
- . Postal Information (21)
- . Quotations (235)
- . Research Papers@
- . Searching the Net@
- . Standards (75)
- . Statistics (27)
- . Thesauri (22)
- . Time@
- . Web Directories (16)
- . World Population Counts@

**Regional**

- SN Countries, Regions, US States . . .
- . U. S. States
- . Countries
- . Regions (8344)
- . Geography@
- . Web Directories (28)



**Science**

- SN Animals, Astronomy, Engineering. . . .
- . Acoustics (66)
- . Agriculture (2054)
- . Alternative (1047)
- . Amateur Science (18)
- . Animals, Insects, and Pets@
- . Anthropology and Archaeology@
- . Artificial Life (129)
- . Ask an Expert (21)
- . Astronomy (2519)
- . Aviation and Aeronautics (236)
- . Bibliographies (6)
- . Biology (16535)
- . Booksellers@
- . Chats and Forums (45)
- . Chemistry (1275)
- . Cognitive Science (94)
- . Complex Systems (23)
- . Computer Science (1516)
- . Dictionaries (27)
- . Earth Sciences (2831)
- . Ecology (746)
- . Education (549)
- . Employment (44)
- . Energy (554)
- . Engineering (4659)
- . Events (35)
- . Forensics (59)
- . Geography (3410)
- . Geology and Geophysics@
- . History (83)
- . Humor@
- . Hydrology@
- . Information Technology (71)
- . Institutes (58)
- . Journals (31)
- . Libraries (33)
- . Life Sciences (17)
- . Mathematics (1934)
- . Measurements and Units (219)
- . Medicine@
- . Meteorology@
- . Museums and Exhibits (146)
- . Nanotechnology (51)
- . News and Media (140)
- . Oceanography@
- . Organizations (160)
- . Paleontology@
- . People (52)
- . Physics (1647)
- . Psychology@
- . Religion and Science@
- . Research (161)
- . Science and Technology Policy
- . Science on Postage Stamps (8)
- . Space (1292)
- . Sports@
- . Web Directories (43)

**Social Science**

- SN Archaeology, Economics, Languages . . .
- . Anthropology and Archaeology (1145)
- . Area Studies (698)
- . Bibliographies (13)
- . Books@
- . Chats and Forums (12)
- . Communications (1700)
- . Conferences (17)
- . Critical Theory@
- . Disability Studies (8)
- . Economics (1113)
- . Education (25)
- . Employment (4)
- . Environmental Studies@
- . Ethnic Studies (173)
- . Futures Studies (25)
- . Gender Studies (21)
- . Genealogy@
- . Geography@
- . Gerontology (36)
- . History@
- . Humanities@
- . Institutes (83)
- . Journals (29)
- . Law@
- . Lesbian, Gay, & Bisexual Studies (62)
- . Libraries (13)
- . Library and Information Science@
- . Linguistics & Human Languages (2814)
- . Migration and Ethnic Relations (37)
- . Organizations (29)
- . Peace and Conflict Studies (129)
- . Political Science (1125)
- . Popular Culture Studies@
- . Psychology (1346)
- . Recreation and Leisure Studies (82)
- . Rural Development (36)
- . Science, Technology, and Society Studies (100)
- . Sexology (33)
- . Social Research (62)
- . Social Work (158)
- . Sociology (421)
- . Urban Studies (324)
- . Web Directories (13)
- . Women's Studies (178)

**Society & Culture**

- SN People, Environment, Religion . . .
- . Advice (108)
- . Bibliographies (11)
- . Chats and Forums (41)
- . Crime (4537)
- . Cultural Policy (11)
- . Cultures and Groups (13258)
- . Death and Dying (477)
- . Disabilities (1482)
- . Environment and Nature (6972)
- . Etiquette (36)
- . Events (35)
- . Families (878)
- . Fashion@
- . Firearms (155)
- . Food and Drink (5147)
- . Gender (35)
- . Holidays and Observances (2062)
- . Issues and Causes (3921)
- . Journals (4)
- . Magazines (219)
- . Museums and Exhibits (5366)
- . Mythology and Folklore (976)
- . People (46774)
- . Pets@
- . Relationships (447)
- . Religion and Spirituality (36093)
- . Reunions (370)
- . Sexuality (1470)
- . Social Organizations (440)
- . Web Directories (10)
- . Weddings (283)



## Figure 4

### Excerpts from the Yahoo Classification

These excerpts from the Yahoo classification are designed to illustrate the structure of the classification, particularly its treatment of *Education* and *Transportation*.

This list is by no means complete with respect to Yahoo classes that deal with *Education* or *Transportation*. Only the areas labeled “(complete)” list all the subdivisions or at least enough subdivisions to give a sense of the overall structure and content of the area. In other areas the subdivisions have been selected to *Education* or *Transportation* or otherwise make a point about the structure.

The list does illustrate patterns that are repeated throughout the classification. Some Yahoo classes do not have the number of associated Web pages. I sometimes used a simplified method for copying from the Yahoo displays that did not carry the number with it. Numbers were also deleted if they ran into a second line.

#### Symbols used

- Category falls conceptually under *Education* (this may or may not be shown in Yahoo)
- ▶ Category falls conceptually under *Transportation* (this may or may not be shown in Yahoo)

**SN** Scope Note. If SN is bolded, you should read it because it illustrates a general point.

At the top of each left page the hierarchical context is given by repeating the preceding levels of the hierarchy in italics, e.g. *Education*. Across a page spread, you can always find the hierarchical context.

The top level categories are **large and bold**, the categories on level 1 are **bold**. Sometimes important subdivisions further down are bolded to make the structure easier to see.

**Arts & Humanities**

SN Literature, Photography ...  
 . **Design Arts** (5187)  
 . . Architecture (1497)  
 . . . Buildings and Structures (474)  
 . . . . ▶Bridges@  
 . . . •Education (158)  
 . . •Education (73)  
 . •**Education**  
 . . Art History@  
 . . Art Schools (132)  
 . . College and University (252)  
 . . Companies@  
 . . Courses (12)  
 . . Design Arts@  
 . . Humanities@  
 . . K-12 (90)  
 . . . Curriculum Standards (28)  
 . . . Drama@  
 . . . Lesson Plans (17)  
 . . . Schools (35)  
 . . . Usenet (2)  
 . . Non-Degree Programs (47)  
 . . Organizations (32)  
 . . Performing Arts@  
 . . Teaching (6)  
 . . Workshops  
 . **Humanities**  
 . . History  
 . . . •Education  
 . . . . Art History@  
 . . . . College and University (217)  
 . . . . Courses (2)  
 . . . . K-12 (71)  
 . . . . . Courses (16)  
 . . . . . Curriculum Standards (5)  
 . . . . . Fairs and Competitions (3)  
 . . . . . Social Studies@  
 . . . . . Teacher Resources (26)  
 . . . . Organizations (4)  
 . . . U.S. History  
 . . . . By Time Period  
 . . . . . 19th Century  
 . . . . . . ▶Transcontinental Railroad  
 . . . Literature (16329)  
 . . . Authors (9971)  
 . . . . ▶Travel Writers (86)  
 . . . Genres (2194)  
 . . . Nonfiction (130)  
 . . . . ▶Travel Writing (16)  
 . **Performing Arts**  
 . . Dance  
 . . . •Education  
 . . . . College and Univ. Depts (24)  
 . . . . K-12 Curriculum Standards  
 . . . . Schools (138)  
 . . . . Summer Programs & Festivals

**Business & Economy**

SN B2B, Finance, Shopping, Jobs  
 . **Business to Business** (268905)  
 . . Construction (36758)  
 . . . General Contractors (448)  
 . . . . Commercial (344)  
 . . . . . ▶Railroads (15)  
 . . ▶Automotive (2053)  
 . . . •Training and Development (32)  
 . . . . Professional Driving Schools  
 . . •**Education** (complete)  
 . . . **By Subject** (909)  
 . . . . Adult Basic Education (8)  
 . . . . Art@  
 . . . . Business@  
 . . . . Computers & Technology (38)  
 . . . . Environment and Nature (11)  
 . . . . Health and Fitness@  
 . . . . Languages (242)  
 . . . . . Directories (1)  
 . . . . . Professional Development (2)  
 . . . . . Teaching & Learning Aids  
 . . . . . . Booksellers@  
 . . . . . . CDs, Records, & Tapes  
 . . . . . . Curriculum (3)  
 . . . . . . Publishers@  
 . . . . . . Software (35)  
 . . . . . . Specific Languages (163)  
 . . . . . . Amharic (1)  
 . . . . . . Arabic (4)  
 . . . . . . Bahasa Indonesia (1)  
 . . . . . . Chinese (11)  
 . . . . . . English 2nd Language  
 . . . . . . Farsi (1)  
 . . . . . . French (15)  
 . . . . . . German (4)  
 . . . . . . Greek (4)  
 . . . . . . Hawaiian (1)  
 . . . . . . Italian (2)  
 . . . . . . Japanese (13)  
 . . . . . . Korean (3)  
 . . . . . . Navajo (1)  
 . . . . . . Russian (7)  
 . . . . . . Scots English (1)  
 . . . . . . Sign Language (12)  
 . . . . . . Spanish (16)  
 . . . . . . Swedish (1)  
 . . . . . . Tagalog (2)  
 . . . . . . Thai (2)  
 . . . . . . Vietnamese (1)  
 . . . . . . Welsh (1)  
 . . . . . Videos (7)  
 . . . Life Skills (61)  
 . . . Mathematics (133)  
 . . . Media Studies (2)  
 . . . Music (11)  
 . . . Reading and Writing (176)  
 . . . Religion@

. . . . Science (166)  
 . . . . Social Studies (39)  
 . . . . Special Education (6)  
 . . . . Vocational Arts (12)  
 . . . Academic Competitions (8)  
 . . . Administration (198)  
 . . . Computers and Technology (81)  
 . . . Consulting (62)  
 . . . Directories (5)  
 . . . Financial Services (14)  
 . . . Fund Raising@  
 . . . International (9)  
 . . . Marketing and Recruiting (14)  
 . . . Presenters (13)  
 . . . Professional Development (55)  
 . . . School Management (42)  
 . . . Supplies and Equipment (83)  
 . . . **Teaching & Learning Aids**  
 (493)  
 . . . . By Subject (49)  
 . . . . . Adult Basic Education@  
 . . . . . . Adult Literacy@  
 . . . . . . . Curriculum (4)  
 . . . . . . . Software (3)  
 . . . . . . . Software (4)  
 . . . . . . . Adult Literacy@  
 . . . . . . Art (8)  
 . . . . . . Curriculum (3)  
 . . . . . . Software (2)  
 . . . . . . Videos (2)  
 . . . . . Business (6)  
 . . . . . Computers & Technology@  
 . . . . . . Curriculum (2)  
 . . . . . . Logo Programming  
 Language@  
 . . . . . . Software (4)  
 . . . . . . Videos (2)  
 . . . . . . Titles (1)  
 . . . . . Environment and Nature@  
 . . . . . Health and Fitness (12)  
 . . . . . Curriculum (6)  
 . . . . . . Sex Education (2)  
 . . . . . Languages@  
 . . . . . . Booksellers@  
 . . . . . . . Arabic (6)  
 . . . . . . . Chinese (5)  
 . . . . . . . French (5)  
 . . . . . . . German (5)  
 . . . . . . . Greek (4)  
 . . . . . . . Japanese (8)  
 . . . . . . . Titles (1)  
 . . . . . . . Polish (4)  
 . . . . . . . Portuguese (7)  
 . . . . . . . Publishers@  
 . . . . . . . Russian (7)  
 . . . . . . . Spanish (17)  
 . . . . . . . Antique, Rare, & Used  
 . . . . . . . Titles (4)

..... Tamil (3)	..... Videos (8)	..... Spelling@
..... Titles (7)	..... Titles (4)	..... Titles (7)
..... Dictionaries@	..... Spanish (15)	..... Speed Reading@
..... Japanese@	..... Booksellers@	..... Software (4)
..... Spanish@	..... Software (6)	..... Spelling (7)
..... CDs, Records, & Tapes	..... Swedish (1)	..... Games (2)
..... English as a second	..... Tagalog (2)	..... Software (3)
Language @	..... Thai (2)	..... Videos (4)
..... Curriculum (3)	..... Vietnamese (1)	..... Vocabulary (11)
..... Publishers@	..... Welsh (1)	..... Religion (10)
..... English as a second	..... Videos (7)	..... Christian Home
Language (5)	..... English as a second	Schooling@
..... Software (34)	Language@	..... Science@
..... Chinese@	..... Sign Language@	..... CDs, Records, & Tapes
..... English as a second	..... Titles (1)	..... Curriculum (12)
Language@	..... Life Skills@	..... Environment & Nature@
..... Japanese@	..... Character Education (11)	..... Games (1)
..... Resellers (5)	..... Curriculum (3)	..... Lab Equipment & Supplies
..... Sign Language@	..... Curriculum (13)	(23)
..... Spanish@	..... Character Education@	..... Manipulatives (6)
..... Titles (2)	..... Software (3)	..... Online Subscription
..... Specific Languages (157)	..... Videos (8)	Services (3)
..... Amharic (1)	..... Titles (2)	..... Posters and Charts (3)
..... Arabic (4)	..... Mathematics@	..... Software (43)
..... Bahasa Indonesia (1)	..... Books@	..... Chemistry (4)
..... Chinese (11)	..... Algebra (3)	..... Physics (6)
..... Booksellers@	..... Calculus (4)	..... Titles (4)
..... Software (7)	..... Chaos (3)	..... Videos (8)
..... English as a second	..... Statistics (1)	..... Social Studies@
Language (57)	..... Curriculum (5)	..... Curriculum (8)
..... CDs, Records, & Tapes	..... Games (2)	..... Online Subscription Serv.
(5)	..... Manipulatives (11)	..... Software (5)
..... Flash Cards (2)	..... Online Subscription	..... Videos (4)
..... Magazines@	Services (1)	..... Titles (1)
..... Online Subscription	..... Posters and Charts (1)	..... Special Education@
Services (2)	..... Software (77)	..... Publishers@
..... Publishers@	..... Testing & Assessment	..... Vocational Arts@
..... Software (17)	Software@	..... Software (4)
..... TOEFL Preparation@	..... Titles (2)	..... Academic Competitions (3)
..... Videos (4)	..... Titles (16)	..... Books@
..... Farsi (1)	..... Videos (6)	..... Retail@
..... French (14)	..... Titles (2)	..... Publishers@
..... Booksellers@	..... Media Studies@	..... Academic@
..... German (4)	..... Music@	..... .College & University
..... Greek (4)	..... Reading and Writing@	Presses@
..... Hawaiian (1)	..... Adult Literacy (9)	..... Organizations (1)
..... Italian (2)	..... Curriculum (4)	..... Textbooks@
..... Japanese (13)	..... Software (3)	..... Booksellers@
..... Booksellers@	..... Books@	..... Home Schooling
..... Titles (1)	..... Curriculum (24)	Materials (21)
..... Software (9)	..... Adult Literacy@	..... Language Education@
..... Korean (2)	..... Flash Cards (5)	
..... Navajo (1)	..... Games (4)	
..... Russian (6)	..... Manipulatives (3)	
..... Booksellers@	..... Software (49)	
..... Scots English (1)	..... Adult Literacy@	
..... Sign Language (12)	..... Handwriting (2)	
..... Software (2)	..... Speed Reading@	

**Business & Economy****Business to Business**

## . . . •Education

## . . . Teaching &amp; Learning Aids (493)

. . . . . Special Education Books

. . . . . Supplementary Materials

. . . . . Forensics &amp; Debate

. . . . . Teacher Aids (12)

. . . . . Textbooks (88)

. . . . . Higher Education (53)

. . . . . K-12 (17)

. . . . . Vocational Training

Books (4)

. . . . . Textbooks (43)

. . . . . College &amp; University

Bookstores@

. . . . . Textbooks@

. . . . . Science@

. . . . . Used (18)

. . . . . Titles (95)

. . . . . Administration (2)

. . . . . College &amp; University

. . . . . Admissions (10)

. . . . . Athletic Recruiting@

. . . . . Financial Aid (6)

. . . . . University Life (4)

. . . . . Distance Learning (3)

. . . . . Educating Your Child@

. . . . . Home Schooling (8)

. . . . . Reform (3)

. . . . . School Directories (4)

. . . . . Study Guides (7)

. . . . . Teaching (20)

. . . . . Reading (8)

. . . . . Test Preparation (12)

. . . . . Textbooks (8)

. . . . . Titles@

. . . . . Christian Home Schooling@

. . . . . Curriculum (38)

. . . . . By Subject (9)

. . . . . Christian Home Schooling@

. . . . . Preschool (9)

. . . . . School Reform (2)

. . . . . Textbooks@

. . . . . Learning Skills (21)

. . . . . Memory Improvement (10)

. . . . . Software (2)

. . . . . Study Skills (2)

. . . . . Books@

. . . . . Publishers@

. . . . . Resellers (23)

. . . . . Reviews (2)

. . . . . Supplementary Materials

. . . . . CDs, Records, &amp; Tapes (7)

. . . . . Languages@

. . . . . English 2nd Language@

. . . . . Science@

. . . . . Distance Learning (35)

. . . . . Online Subscription Serv.

. . . . . Email (2)

. . . . . Art@

. . . . . English 2nd Language@

. . . . . Mathematics@

. . . . . Science@

. . . . . Social Studies@

. . . . . Flash Cards (10)

. . . . . English as a second

Language@

. . . . . Reading and Writing@

. . . . . Software (5)

. . . . . Games (7)

. . . . . Mathematics@

. . . . . Reading and Writing@

. . . . . Science@

. . . . . Software@

. . . . . Manipulatives (5)

. . . . . Mathematics@

. . . . . Reading and Writing@

. . . . . Science@

. . . . . Montessori (10)

. . . . . Newspapers &amp; Magazines

. . . . . Posters and Charts (9)

. . . . . Mathematics@

. . . . . Science@

. . . . . Publishers@

. . . . . Software (149)

. . . . . By Subject (10)

. . . . . Adult Basic Education@

. . . . . Art@

. . . . . Computers &amp;

Technology@

. . . . . Languages@

. . . . . Life Skills@

. . . . . Mathematics@

. . . . . Reading and Writing@

. . . . . Science@

. . . . . Social Studies@

. . . . . Vocational Arts@

. . . . . Flash Cards@

. . . . . Games (1)

. . . . . Resellers (18)

. . . . . Shareware (2)

. . . . . Software Reviews@

. . . . . Teaching Tools (15)

. . . . . Admin. Functions@

. . . . . Classroom Organizers

. . . . . Consulting (1)

. . . . . Curriculum Planning

. . . . . Discipline Tracking

. . . . . Financial (5)

. . . . . Grading &amp; Evaluation

. . . . . Resellers (1)

. . . . . Titles (3)

. . . . . Home Schooling (5)

. . . . . Information

Management Systems

. . . . . School-to-Work

Progr. Management

. . . . . Student Information

. . . . . Inventory Management

. . . . . Registration (1)

. . . . . Resellers (3)

. . . . . Scheduling (8)

. . . . . Special Education

. . . . . Testing &amp; Assessment

. . . . . Classroom Computer

Control (4)

. . . . . Test Preparation@

. . . . . College Entrance@

. . . . . Titles (4)

. . . . . Mathematics@

. . . . . Reading and Writing@

. . . . . Videos (44)

. . . . . By Subject (13)

. . . . . Art@

. . . . . Computers &amp; Technol.@

. . . . . Languages@

. . . . . Life Skills@

. . . . . Mathematics@

. . . . . Reading and Writing@

. . . . . Science@

. . . . . Social Studies@

. . . . . Titles (5)

. . . . . Computers &amp; Tech..@

. . . . . Languages@

. . . . . Sign Language@

. . . . . Life Skills@

. . . . . Mathematics@

. . . . . Social Studies@

. . . . . Test Preparation@

. . . . . Titles (2)

. . . . . By Subject@

. . . . . Test Preparation@

. . . . . Teaching Methods (1)

. . . . . Testing and Assessment (27)

. . . **Emergency Services (803)**

. . . ▶Air Ambulance Services (34)

. . . Supplies and Equipment (159)

. . . ▶Vehicles (20)

. . . . . Ambulances (12)

. . . . . Fire Trucks@

. . . . . Rescue Watercraft@

. . . Labor

. . . . . Unions

. . . . . ▶Railroad@

. . . ▶**Transportation (3755)**

. . . . . Aerospace@

. . . . . Automotive@

. . . . . Aviation (1627)

. . . . . Aerial Advertising@

. . . . . Aerospace@

. . . . . Agricultural@

. . . . Air Cargo Services@	. . . . •Driving Schools (35)	. . . . Aviophobia@
. . . . Air Shows (8)	. . . . <b>Shopping and Services</b> (378101)	. . . . Classifieds@
. . . . Aircraft (900)	. . . . ▶Automotive (24751)	. . . . Companies@
. . . . Airlines@	. . . . •Driving Schools (524)	. . . . History (149)
. . . . . United Airlines	. . . . . By Region (491)	. . . . Military@
. . . . . •Education	. . . . . U.S. States (256)	. . . . Museums@
. . . . Airports (110)	. . . . . Massachusetts (5)	. . . . Pictures (46)
. . . . Aviation Weather@	. . . . . SN Boston Metro not here	. . . . Recreational Aviation@
. . . . Avionics@	. . . . . New York (42)	. . . . Safety (32)
. . . . Aviophobia@	. . . . . Cities (14)	. . . . Women
. . . . Books@	. . . . . New York@ <b>see ***</b>	. . . . Buses (26)
. . . . Classifieds@	. . . . Motorcycles (1054)	. . . . Canals (22)
. . . . Consulting (71)	. . . . •Driving Schools (5)	. . . . Commuting (30)
. . . . Directories (28)	. . . . Publishers (3285)	. . . . Companies@
. . . . In-Flight Entertainment (3)	. . . . •Education (277)	. . . . Employment (4)
. . . . Insurance@	. . . . Languages (35)	. . . . Events (2)
. . . . Navigation Systems@	. . . . •Language Education (17)	. . . . Freight (8)
. . . . Organizations (80)	. . . . . English as a second	. . . . Government Agencies (62)
. . . . Software (76)	. . . . . Language	. . . . Highways and Roads (127)
. . . . Supplies and Equipment (107)	. . . . • <b>Education</b>	. . . . Automated Highway Systems@
. . . . Trade Magazines (19)	. . . . Business Schools@	. . . . Bridges (52)
. . . . Training (189)	. . . . College and University (775)	. . . . Government Agencies (9)
. . . . Videos@	. . . . K-12 (26)	. . . . Organizations (13)
. . . . Buses (74)	. . . . Curriculum Standards (4)	. . . . Regional Information (38)
. . . . Consulting (73)	. . . . Organizations (22)	. . . . Road Maintenance@
. . . . Directories (10)	. . . . Organizations (6)	. . . . Roman Roads@
. . . . Fleet Services (92)	. . . . ▶ <b>Transportation</b> (2094) (complete)	. . . . Scenic Highways and Byways@
. . . . Intelligent Transportation Syst.	. . . . SN No subcategory Education	. . . . Traffic and Road Conditions@
. . . . Limousines and Shuttles (23)	. . . . Auto-Free Transportation (23)	. . . . Traff. Signs, Signals, & Lamps@
. . . . Maritime (307)	. . . . Bicycle Advocacy@	. . . . History (5)
. . . . Boat Transport (25)	. . . . Mass Transit@	. . . . Institutes (44)
. . . . Boating@	. . . . Organizations (8)	. . . . Intelligent Transportation Systems
. . . . Cargo Services@	. . . . Pedestrian Advocacy & Safety@	. . . . Libraries (7)
. . . . Construction@	. . . . Aviation (513)	. . . . Limousines and Shuttles@
. . . . Consulting (5)	. . . . Accidents (51)	. . . . Maritime (237)
. . . . Conventions &Trade Shows (6)	. . . . SN LCC has X-ref from Health	. . . . Boats@
. . . . Directories (10)	. . . . Grief Support@	. . . . Companies@
. . . . Insurance@	. . . . Specific Crashes (37)	. . . . Cruise Lines@
. . . . Navigation@	. . . . Aerospace Companies@	. . . . Employment (8)
. . . . Organizations (17)	. . . . Aviation@	. . . . Maritime History@
. . . . Parts and Accessories (39)	. . . . Consulting (45)	. . . . Merchant Marines (6)
. . . . Port Authorities@	. . . . Development (6)	. . . . Passenger Ferries (51)
. . . . Publications (4)	. . . . Electronic Warfare@	. . . . Port Authorities (128)
. . . . Shipbuilding (94)	. . . . Engines (16)	. . . . Ships (28)
. . . . Ships (33)	. . . . Equipment (214)	. . . . Submarines@
. . . . Software (7)	. . . . Industry Information (28)	. . . . Mass Transit (58)
. . . . Submarines (9)	. . . . Research and Design (32)	. . . . Museums (17)
. . . . Trade Magazines (2)	. . . . Simulations (15)	. . . . News and Media (247)
. . . . Movers@	. . . . Software (25)	. . . . Traffic & Road Conditions (234)
. . . . Organizations@	. . . . Spacecraft	. . . . Traffic Cams (15)
. . . . Road Maintenance (30)	. . . . Air Traffic Control (14)	. . . . Organizations (67)
. . . . Software (64)	. . . . Air Travel@	. . . . Port Authorities@
. . . . Trade Magazines (10)	. . . . Aircraft (119)	. . . . Safety (14)
. . . . Traffic Control (165)	. . . . Art@	. . . . Statistics (7)
. . . . Trains and Railroads (126)	. . . . Aviation and Aeronautics@	
. . . . Trolleys (8)	. . . . Aviation & Aerospace Medcin@	
. . . . Trucks	. . . . Aviation Weather@	
. . . . Trucking (594)	. . . . Aviators (55)	

**Business & Economy**. ▶ **Transportation**

- . . Streetcars, Trolleys, & Trams (45)
- . . Taxis (17)
- . . Traffic and Road Conditions@
- . . Trains and Railroads (390)
- . . . Companies@
- . . . High Speed Rail (11)
- . . . History (81)
- . . . Magazines (10)
- . . . Model Trains and Railroads@
- . . . Monorails (9)
- . . . Museums (64)
- . . . Organizations (27)
- . . . Railbiking@
- . . . Rails-to-Trails (27)
- . . . Railway Enthusiasts (55)
- . . . Railwayana (5)
- . . . Safety (17)
- . . . Steam Locomotives (8)
- . . . Subways (5)
- . . . Travel@
- . . . Web Directories (10)
- . . . Usenet
- . . Transportation Engineering@
- . . Trucking (46)
- . . Tunnels (32)
- . . Web Directories (3)
- . . Usenet (4)

**Computers & Internet**

SN Internet, WWW, Software,  
Games ...

- . **Internet** (5999)
- . . Devices Connected to the Internet (1316)
- . . . Web Cams (1233)
- . . . . Outdoor Cams (233)
- . . . . . ▶ Traffic Cams@
- . . •Online Teaching and Learning@
- . **Multimedia** (675)
- . . Audio (200)
- . . . Formats (133)
- . . . . MP3 (118)
- . . . . . Players@
- . . . . . . ▶Automotive (9)
- . **Software**
- . . Reviews
- . . . •Education
- . . . Scientific
- . . . . Math
- . . . . . •Educational

- Education** (complete, except for regional subdivisions)

SN College and University, K-12  
SN No subcategory Transportation  
either directly or indirectly

. **By Region** (170)

- . . Regions (6)
- . . . Africa@
- . . . Asia@
- . . . . Browse By Country (38)
- . . . . . Armenia@
- . . . . . Azerbaijan@
- . . . . . Bahrain@
- . . . . . Bangladesh@
- . . . . . Higher Education (10)
- . . . . . Organizations (3)
- . . . . . Primary and Secondary (3)
- . . . . . . Cities@
- . . . . . . Complete List@
- . . . . . Brunei@
- . . . . . Cambodia@
- . . . . . China@
- . . . . . India@
- . . . . . Indonesia@
- . . . . . Iran@ .
- . . . . . Government Agencies (1)
- . . . . . Higher Education (19)
- . . . . . Colleges & Universities
- . . . . . . Cities@
- . . . . . . Isfahan@
- . . . . . . Shiraz@
- . . . . . . Tehran@
- . . . . . . Complete List@
- . . . . . Teaching (1)
- . . . . . Iraq@
- . . . . . Israel@
- . . . . . Japan@
- . . . . . Jordan@
- . . . . . Kazakstan@
- . . . . . Korea, South@
- . . . . . Krygyzstan@
- . . . . . Kuwait@
- . . . . . Lebanon@
- . . . . . Macau@
- . . . . . Malaysia@
- . . . . . Mongolia@
- . . . . . Myanmar@
- . . . . . Nepal@
- . . . . . Oman@
- . . . . . Pakistan@
- . . . . . Philippines@
- . . . . . Qatar@
- . . . . . Russia@
- . . . . . Saudi Arabia@
- . . . . . Singapore@
- . . . . . Sri Lanka@
- . . . . . Taiwan@
- . . . . . Thailand@

- . . . . . Turkey@
- . . . . . United Arab Emirates@
- . . . . . Vietnam@
- . . . . . Yemen@
- . . . . . Companies@
- . . . . . Conferences (1)
- . . . . . Higher Education (1)
- . . . . . Organizations (3)
- . . . . . Primary and Secondary (3)
- . . . . . . Complete List@
- . . . . . . Guidance Counseling
- . . . . . . Student Resources
- . . . Europe@
- . . . Latin America@
- . . . Oceania@
- . . . Pacific Rim@
- . . Countries (113)
- . . . Andorra@
- . . . Argentina@
- . . . . Distance Learning (1)
- . . . . Higher Education (28)
- . . . . . Colleges and Universities(29)
- . . . . . . Cities@
- . . . . . . Buenos Aires@
- . . . . . . Cordoba@
- . . . . . . Lujan@
- . . . . . . Moron@
- . . . . . . Rosario@
- . . . . . . Salta@
- . . . . . . Provinces and Regions@ .
- . . . . . . Buenos Aires@
- . . . . . . Cordoba@
- . . . . . . Lujan@
- . . . . . . Moron@
- . . . . . . Rosario@
- . . . . . . Salta@
- . . . . . . Complete List@
- . . . . . Organizations (1)
- . . . . . Primary and Secondary (20)
- . . . . . . Cities@
- . . . . . . Buenos Aires@
- . . . . . . Cordoba@
- . . . . . . Provinces and Regions@
- . . . . . . Buenos Aires@
- . . . . . . Cordoba@
- . . . . . . Complete List@
- . . . . . . Schools (1)
- . . . . . . Programs (1)
- . . . . . Spanish Language Schools@
- . . . Armenia@
- . . . Australia@
- . . . Austria@
- . . . Bahrain@
- . . . Bangladesh@
- . . . Belarus@
- . . . Belgium@
- . . . Belize@
- . . . Bolivia@



. . . Bosnia and Herzegovina@	. . . Micronesia, Fed. States of@	. . . . Juneteenth@
. . . Brazil@	. . . Monaco@	. . . . Lynching@
. . . Brunei@	. . . Morocco@	. . . . Museums & Memorials (13)
. . . Bulgaria@	. . . Myanmar@	. . . . People@
. . . Canada@	. . . Nepal@	. . . . Science and Technology (4)
. . . Chile@	. . . Netherlands@	. . . . Slavery@
. . . China@	. . . New Zealand@	. . . . Sports@
. . . Colombia@	. . . Nicaragua@	. . . . Timelines (3)
. . . Congo, Democr. Republic of@	. . . Norway@	. . . . Web Directories (3)
. . . Costa Rica@	. . . Pakistan@	. . . . Institutes (31)
. . . Croatia@	. . . Panama@	. . . . Journals (2)
. . . Cuba@	. . . Papua New Guinea@	. . . . Libraries (2)
. . . Cyprus@	. . . Paraguay@	. . . . Literature@
. . . Czech Republic@	. . . Peru@	. . . . Theorists and Critics (7)
. . . Côte d'Ivoire@	. . . Philippines@	. . . Historically Black Colleges & Universities@
. . . Denmark@	. . . Poland@	. . . Lesbian, Gay, and Bisexual (207)
. . . Dominican Republic@	. . . Portugal@	. . . Migrant (9)
. . . Ecuador@	. . . Romania@	. . . Native American (44)
. . . Egypt@	. . . Russia@	. . . Religious@
. . . El Salvador@	. . . Saint Vincent and The Grenadines@	. . . Rural (11)
. . . Estonia@	. . . Saudi Arabia@	. . . Seniors (8)
. . . Federal Republic of Yugoslavia@	. . . Singapore@	. . . U.S. Hispanic and Latino (9)
. . . Fiji@	. . . Slovakia@	. . . Women (90)
. . . Finland@	. . . Slovenia@	. <b>By Subject</b> (11)
. . . France@	. . . South Africa@	. . . Art@
. . . Germany@	. . . Spain@	. . . Art History@
. . . Ghana@	. . . Sri Lanka@	. . . . Art Historians@
. . . Greece@	. . . Sudan@	. . . . College and University Departments (75)
. . . Grenada@	. . . Sweden@	. . . . Courses (3)
. . . Guatemala@	. . . Switzerland@	. . . Art Schools (126)
. . . Haiti@	. . . Taiwan@	. . . College and University (253)
. . . Honduras@	. . . Thailand@	. . . Companies@
. . . Hungary@	. . . Tonga@	. . . Courses (12)
. . . Iceland@	. . . Trinidad and Tobago@	. . . Design Arts@
. . . India@	. . . Tunisia@	. . . Humanities@
. . . Indonesia@	. . . Turkey@	. . . K-12 (88)
. . . Iran@	. . . Uganda@	. . . Non-Degree Programs (47)
. . . Iraq@	. . . Ukraine@	. . . Organizations (31)
. . . Ireland@	. . . United Arab Emirates@	. . . Performing Arts@
. . . Israel@	. . . United Kingdom@	. . . Teaching (6)
. . . Italy@	. . . Uruguay@	. . . Workshops (46)
. . . Jamaica@	. . . Venezuela@	. . . Business@
. . . Japan@	. . . Vietnam@	. . . Environment and Nature@
. . . Jordan@	. . . Zambia@	. . . Health@
. . . Kenya@	. . . Zimbabwe@	. . . Humanities@
. . . Korea, South@	. . . U.S. States (51)	. . . Languages@
. . . Kuwait@	. <b>By Culture or Group</b> (396)	. . . Mathematics@
. . . Latvia@	. . . African American (17)	. . . Music@
. . . Lebanon@	. . . African American Studies@	. . . News and Media@
. . . Liechtenstein@	. . . . History@	. . . Science@
. . . Lithuania@	. . . . Amistad@	. . . Social Science@.
. . . Luxembourg@	. . . . Black History Month (18)	
. . . Macedonia, Former Yugoslav @	. . . . Buffalo Soldiers (14)	
. . . Malaysia@	. . . . Civil Rights Movement@	
. . . Malta@	. . . . Civil War Units (4)	
. . . Marshall Islands@	. . . . Genealogy (15)	
. . . Mexico@	. . . . Harlem Renaissance@	



<ul style="list-style-type: none"> <li>.. Organizations (11)</li> <li>.. Research Centers (2)</li> <li>.. <b>Financial Aid</b> (386)</li> <li>.. College Aid Offices (163)</li> <li>.. Companies@</li> <li>.. Grants (57)</li> <li>.. K-12 School Funding@</li> <li>.. Loans (11)</li> <li>.. Organizations (13)</li> <li>.. Regional Resources (23)</li> <li>.. Savings &amp; Investment Planning@</li> <li>.. Scholarship Programs (88)</li> <li>.. <b>Government Agencies</b> (77)</li> <li>.. Canada@</li> <li>.. Equity@</li> <li>.. United States (53)</li> <li>.. <b>Graduation</b> (58)</li> <li>.. Clip Art and Graphics (6)</li> <li>.. Graduation Poems (3)</li> <li>.. Speeches (42)</li> <li>.. Virtual Cards@</li> <li>.. <b>Higher Education</b> (17643)</li> <li>.. Academic Competitions@</li> <li>.. College Bowl (13)</li> <li>.. Teams (11)</li> <li>.. Debate@</li> <li>.. Clubs, Teams, &amp; Societies (29)</li> <li>.. Teams (16)</li> <li>.. College Bowl@</li> <li>.. Debate@</li> <li>.. Forensics@</li> <li>.. Books@</li> <li>.. Admissions (10)</li> <li>.. Athletic Recruiting@</li> <li>.. Financial Aid (6)</li> <li>.. University Life (4)</li> <li>.. <b>College Entrance</b> (441)</li> <li>.. Admissions Offices (299)</li> <li>.. Books@</li> <li>.. Business Schools@</li> <li>.. Companies@</li> <li>.. Athletic Recruiting@</li> <li>.. Baseball@</li> <li>.. Books@</li> <li>.. Football@</li> <li>.. Soccer@</li> <li>.. Books@</li> <li>.. College Tour Operators (2)</li> <li>.. Counseling (29)</li> <li>.. Graduate School (4)</li> <li>.. International Students (9)</li> <li>.. Essays and Applications (25)</li> <li>.. International Students (5)</li> <li>.. Counseling@</li> <li>.. Online Applications (3)</li> <li>.. Software (2)</li> <li>.. Videos (3)</li> </ul>	<ul style="list-style-type: none"> <li>.. Educational Standards &amp; Testing@</li> <li>.. ACT (1)</li> <li>.. GED (2)</li> <li>.. GED Prep. Companies@</li> <li>.. GMAT (3)</li> <li>.. GRE (3)</li> <li>.. IELTS@</li> <li>.. By Region (52)</li> <li>.. Canadian Provinces (8)</li> <li>.. U.S. States (44)</li> <li>.. By Subject (13)</li> <li>.. K-12 Curriculum Standards@</li> <li>.. LSAT (4)</li> <li>.. MCAT (5)</li> <li>.. Companies@</li> <li>.. Occupational Standards@</li> <li>.. Australian Qualifications Framework@</li> <li>.. U.K. Natl Vocatl Qualific.@</li> <li>.. U.S. National Skill Standards</li> <li>.. SSAT (1)</li> <li>.. Test Preparation Companies@</li> <li>.. Books@</li> <li>.. Career Fields (79)</li> <li>.. Civil Service (5)</li> <li>.. Customs Broker (2)</li> <li>.. •Education (10)</li> <li>.. TEFL/TESL (9)</li> <li>.. Electrician (2)</li> <li>.. Engineering (2)</li> <li>.. Finance (29)</li> <li>.. CFA (11)</li> <li>.. CPA (8)</li> <li>.. EA (1)</li> <li>.. Insurance (2)</li> <li>.. NASD (3)</li> <li>.. Health (18)</li> <li>.. Medicine (11)</li> <li>.. Boards (6)</li> <li>.. USMLE (5)</li> <li>.. Mental Health (1)</li> <li>.. Nursing@</li> <li>.. Law (6)</li> <li>.. Bar Examination (6)</li> <li>.. Military (1)</li> <li>.. Social Work</li> <li>.. College Entrance (26)</li> <li>.. GED (5)</li> <li>.. Graduate School Entrance</li> <li>.. Online Subscription Serv. (7)</li> <li>.. Software (10)</li> <li>.. TOEFL (7)</li> <li>.. Videos (2)</li> <li>.. Testing Companies@</li> <li>.. Computer-Based Testing Services (7)</li> <li>.. Online (5)</li> </ul>	<ul style="list-style-type: none"> <li>.. Software@</li> <li>.. Math (3)</li> <li>.. Titles (2)</li> <li>.. Resellers (1)</li> <li>.. TOEFL@</li> <li>.. TOEIC@</li> <li>.. Voluntary National Testing (4)</li> <li>.. Financial Aid@</li> <li>.. College Aid Offices (163)</li> <li>.. Companies@</li> <li>.. Books@</li> <li>.. Counseling (4)</li> <li>.. Lenders (33)</li> <li>.. International (1)</li> <li>.. Loan Guarantors (4)</li> <li>.. Loan Servicers (6)</li> <li>.. Scholarship Search Serv. (16)</li> <li>.. Secondary Markets (5)</li> <li>.. Grants (57)</li> <li>.. Web Directories (4)</li> <li>.. K-12 School Funding@</li> <li>.. Organizations (7)</li> <li>.. Technology Funding@</li> <li>.. Technology Funding (16)</li> <li>.. Organizations (7)</li> <li>.. Programs (5)</li> <li>.. Loans (11)</li> <li>.. Lenders@</li> <li>.. International (1)</li> <li>.. Organizations (13)</li> <li>.. Regional Resources (23)</li> <li>.. Savings &amp; Investment Planning@</li> <li>.. Education Savings Plans@</li> <li>.. Scholarship Programs (88)</li> <li>.. Search Services@</li> <li>.. Web Directories (4)</li> <li>.. Search Services@</li> <li>.. Web Directories (4)</li> <li>.. Online Applications (83)</li> <li>.. Individual Schools (76)</li> <li>.. School Rankings (12)</li> <li>.. Graduate (7)</li> <li>.. Law School@</li> <li>.. Test Preparation Companies@</li> <li>.. Books@</li> <li>.. Career Fields (79)</li> <li>.. Civil Service (5)</li> <li>.. Customs Broker (2)</li> <li>.. •Education (10)</li> <li>.. TEFL/TESL (9)</li> <li>.. Electrician (2)</li> <li>.. Engineering (2)</li> </ul>
---	---	--

**Education****Higher Education**

- . . . *College Entrance*
- . . . . *Test Preparation Companies@*
- . . . . . *Career Fields (79)*
- . . . . . Finance (29)
  - . . . . . . CFA (11)
  - . . . . . . CPA (8)
  - . . . . . . EA (1)
  - . . . . . . Insurance (2)
  - . . . . . . NASD (3)
  - . . . . . . Health (18)
  - . . . . . . Medicine (11)
    - . . . . . . . Boards (6)
    - . . . . . . . USMLE (5)
    - . . . . . . . Mental Health (1)
    - . . . . . . . Nursing@
    - . . . . . . . Law (6)
    - . . . . . . . Bar Examination (6)
    - . . . . . . . Military (1)
    - . . . . . . . Social Work (2)
  - . . . . . . College Entrance (26)
  - . . . . . . AP (2)
  - . . . . . . SAT (13)
  - . . . . . . SAT II (1)
  - . . . . . . Software (3)
  - . . . . . . GED (5)
  - . . . . . . Graduate School Entrance (25)
    - . . . . . . . GMAT (4)
    - . . . . . . . GRE (3)
    - . . . . . . . LSAT (4)
    - . . . . . . . MCAT (4)
  - . . . . . . Online Subscription Serv. (7)
  - . . . . . . Software (10)
  - . . . . . . College Entrance@
  - . . . . . . TOEFL (7)
  - . . . . . . Videos (2)
  - . . . . . . Web Directories (4)
- . . . **Colleges and Universities (16861)**
  - . . . . By Region (15221)
    - . . . . . SN A long list as above
    - . . . . . Argentina@
    - . . . . . Armenia@
    - . . . . . . . . .
  - . . . . **Distance Learning@**
    - . . . . . Business@
    - . . . . . Graduate Programs@
    - . . . . . Community Colleges (20)
    - . . . . . Consortia (10)
    - . . . . . Credit for Prior Experience (6)
    - . . . . . Engineering@
      - . . . . . . Graduate Programs (9)
    - . . . . . Graduate Programs (24)
      - . . . . . . Business@
      - . . . . . . Engineering@
    - . . . . . Web Directories (2)
  - . . . **Graduate Education (74)**

- . . . . By Subject (49)
  - . . . . . Accounting and Auditing@
  - . . . . . American (United States) Studies@
  - . . . . . Archaeology@
  - . . . . . Arts@
  - . . . . . Asian Studies@
  - . . . . . Business Schools@
  - . . . . . Classics@
  - . . . . . Comparative Literature@
  - . . . . . Conflict Resolution@
  - . . . . . Cultural Studies@
  - . . . . . Economics@
  - . . . . . European Studies@
  - . . . . . Film@
  - . . . . . Finance@
  - . . . . . First Nations Studies (Canada)@
  - . . . . . Gender Studies@
  - . . . . . History@
  - . . . . . History & Philosophy of Science@
  - . . . . . Human Ecology@
  - . . . . . International Relations@
  - . . . . . Irish Studies@
  - . . . . . Japan Studies@
  - . . . . . Journalism@
  - . . . . . Latin American Studies@
  - . . . . . Law Schools@
  - . . . . . Library & Information Science@
  - . . . . . Literature@
  - . . . . . Marine Archaeology@
  - . . . . . Mathematics@
  - . . . . . Medical Schools@
  - . . . . . Mexican American Studies@
  - . . . . . Middle East Studies@
  - . . . . . Peace and Conflict Studies@
  - . . . . . Philosophy@
  - . . . . . Physics@
  - . . . . . Political Economy@
  - . . . . . Political Science@
  - . . . . . Psychology@
  - . . . . . Public Administration@
  - . . . . . Public Policy@
  - . . . . . Recreation and Leisure Studies@
  - . . . . . Rhetoric@
  - . . . . . Russian & East European Studies@
  - . . . . . Scandinavian Studies@
  - . . . . . Science, Technology, & Society Studies@
  - . . . . . Sexology@
  - . . . . . Slavic Studies@
  - . . . . . Social Work@
  - . . . . . Women's Studies@
- . . . . Distance Learning@

- . . . . Business@
- . . . . Engineering@
- . . . . Electronic Theses & Dissertations (ETDs)@
- . . . . Graduate Record Exam@
- . . . . Organizations (6)
- . . . . Rankings@
- . . . . Law School
- . . . . Web Directories (7)
- . . **Guidance (64)**
  - . . . . Academic Advising (2)
  - . . . . College Entrance@
  - . . . . Admissions Offices (299)
  - . . . . Books@
  - . . . . Business Schools@
  - . . . . Companies@
  - . . . . Educational Standards & Testing@
  - . . . . Financial Aid@
  - . . . . Online Applications (83)
  - . . . . School Rankings (12)
  - . . . . Test Preparation Companies@
  - . . . . Web Directories (4)
- . . . Medical School (34)
  - . . . . . MCAT@
    - . . . . . . Companies@
    - . . . . . . Medical Schools@
      - . . . . . . . MCAT@
      - . . . . . . . Medical Schools@
        - . . . . . . . . Admissions & Guidance@
          - . . . . . . . . . MCAT@
          - . . . . . . . . . Medical Schools@
            - . . . . . . . . . . Student Organizations@
              - . . . . . . . . . . . Web Directories (2)
              - . . . . . . . . . . . Caribbean Med. Schools (8)
              - . . . . . . . . . . . Dental Schools@
                - . . . . . . . . . . . . Orthodontic@
                  - . . . . . . . . . . . . Student Organizations@
                    - . . . . . . . . . . . . . Interns and Residents (3)
                    - . . . . . . . . . . . . . Optometry@
                      - . . . . . . . . . . . . . . Web Directories (2)
        - . . . . . Web Directories (2)
      - . . . Honors Programs (46)
      - . . **News and Media@**
        - . . . Individual Schools (419)
          - . . . . SN Long list of countries and US states
          - . . . . Australia@
          - . . . . Canada@
          - . . . . . . . . .
          - . . . . . Alabama (4)
          - . . . . . Arizona (3)
          - . . . . . . . . .
          - . . . . . Magazines (174)
          - . . . . . Individual Schools (139)
            - . . . . . . Alumni (40)
            - . . . . . . Humor@

- . . . . . Literary (52)
- . . . Newspapers (436)
- . . . . . Business Schools (8)
- . . . . . Medical Schools (2)
- . . . . . Newswires@
- . . . . . Web Directories (2)
- . . . Radio Stations@
- . . . . . Countries (13)
- . . . . . Australia@
- . . . . . Belgium@
- . . . . . Cities@
- . . . . . Provinces@
- . . . . . Complete List@
- . . . . . Brazil@
- . . . . . Canada@
- . . . . . Denmark@
- . . . . . Ireland@
- . . . . . New Zealand@
- . . . . . Singapore@
- . . . . . South Africa@
- . . . . . Sweden@
- . . . . . Switzerland@
- . . . . . United Kingdom@
- . . . . . Vietnam@
- . . . . Eastern United States (168)
- . . . . Internet Broadcasts@
- . . . . Western United States (110)
- . . Organizations (96)
- . . . Alumnae/I Associations@
- . . . Community College (9)
- . . . Graduate Education@
- . . . Transfer Student (3)
- . . Policy Research Centers@
- . . Seminaries@
- . . . Christian@
- . . . Jewish@
- . . Student Life (25)
- . . . Books@
- . . . Classifieds@
- . . . Organizations@
- . . Web Directories (14)
- . . . College Entrance@
- . . . Colleges and Universities@
- . . . Graduate Programs@
- . **Instructional Technology** (327)
- . . Conferences (23)
- . . . Online Teaching and Learning@
- . . . Past Events (2)
- . . Institutes (47)
- . . . College & University  
    Departments (34)
- . . Journals (9)
- . . Online Teaching & Learning (167)
- . . . Conferences (8)
- . . . Corporate Programs (10)
- . . . Courses (9)
- . . . Educational MOOs (7)
- . . . Magazines (3)

- . . . Organizations (14)
- . . . Papers (4)
- . . . . . Telementoring@
- . . . Projects (20)
- . . . . . Higher Education (4)
- . . . . . Regional (4)
- . . . . . Virtual Field Trips@
- . . . Research Institutes (3)
- . . . Teacher Resources (43)
- . . . . . Companies@
- . . . . . Email (2)
- . . . . . English as a 2nd Language@
- . . . . . Mathematics@
- . . . . . Science@
- . . . . . Social Studies@
- . . . Telementoring (12)
- . . . . . Papers (2)
- . . . . . Web Directories (1)
- . . . Web Directories (8)
- . . . Telementoring@
- . . WebQuests (19)
- . . Collections (8)
- . . Science@
- . . Social Studies@
- . . School Technology Funding@
- . . Organizations (7)
- . . Programs (5)
- . . Software Reviews@
- . . Titles (64)
- . . . College Preparatory (6)
- . . . Math@
- . . . Reading (23)
- . . . . . Reader Rabbit (7)
- . . Web Directories (6)
- . **Journals** (33)
- . . •Educational Theory@
- . . Instructional Technology@
- . . Music@
- . **K-12** (52293)
- . . **By Region** (27417)
- . . . Countries (7635)
- . . . Regions (9)
- . . . U.S. States (19773)
- . . **Academic Competitions@**
- . . . Debate@
- . . . . . Clubs, Teams, & Societies (25)
- . . . . . Lincoln-Douglas@
- . . . . . Institutes (2)
- . . . . . Policy Debate (7)
- . . . . . 1998-99 Debate Topic (3)
- . . . . . 1999-2000 Debate Topic (1)
- . . . History@
- . . Mathematics@
- . . . International (6)
- . . Science@
- . . . International (4)
- . . . . . Mathematics@
- . . . Mathematics@

- . . . . . International (6)
- . . . Projects and Ideas (20)
- . . . . . Science Project Books@
- . . . . . Web Directories (2)
- . . Teams (13)
- . . . Debate@
- . . . . . Lincoln-Douglas@
- . . . Forensics@
- . . Arts@
- . . Curriculum Standards (27)
- . . Dance@
- . . Drama@
- . . Lesson Plan (4)
- . . School Departments (10)
- . . Lesson Plans (17)
- . . Theater@
- . . Schools (35)
- . . . Departments (5)
- . . Usenet (2)
- . . Conferences (10)
- . . **Countries** (41)
- . . Australia@
- . . Belgium@
- . . Brazil@
- . . Canada@
- . . Chile@
- . . China@
- . . Columbia@
- . . Costa Rica@
- . . Cyprus@
- . . Denmark@
- . . Egypt@
- . . Estonia@
- . . Ethiopia@
- . . France@
- . . Germany@
- . . Ghana@
- . . Greece@
- . . Hong Kong@
- . . India@
- . . Ireland@
- . . Israel@
- . . Italy@
- . . Japan@
- . . Korea, South@
- . . Macau@
- . . Malaysia@
- . . Malta@
- . . Mexico@
- . . Netherlands@
- . . New Zealand@
- . . . . .

**Education****. K-12****. . Curriculum Standards (65)**

. . . . By Region (52)

. . . . . Canadian Provinces (8)

. . . . . U. S. States (44)

. . . . By Subject (13)

. . . . . Agriculture@

. . . . . Arts@

. . . . . Business@

. . . . . Dance@

. . . . . English Language Arts@

. . . . . Family and Consumer  
Sciences@

. . . . . Health@

. . . . . History@

. . . . . Languages@

. . . . . Library and Information  
Literacy@

. . . . . Mathematics@

. . . . . Music@

. . . . . Physical Education@

. . . . . Science@

. . . . . Social Studies@

. . Distance Learning@

. . . Home Schooling@

. . . . . Christian@

. . Environment and Nature@

. . . Programs (18)

. . Gifted Youth (65)

. . . Schools (14)

. . Guidance Counseling (17)

. . . K-12 Offices (7)

. . . Organizations (3)

**. . Home Schooling@**

. . . . . Christian@

. . . . . Companies@

. . . . . Publishers@

. . . . . Teaching &amp; Learning Aids

. . . . . . Curriculum (17)

. . . . . . Used (3)

. . . . . Distance Learning (11)

. . . . . Organizations (10)

. . . . . Usenet (2)

. . . . . Companies@

. . . . . Administrative Software@

. . . . . Books@

. . . . . Christian Home Schooling (36)

. . . . . Publishers@

. . . . . Teaching &amp; Learning Aids

SN This is its own category  
under Christian Home

Schooling

. . . . . . Curriculum (17)

. . . . . . Used (3)

. . . . . Publishers@

. . . . . Christian (9)

. . . . Teaching &amp; Learning Aids@

SN This is a reference to the  
generic Teaching & Learning  
Aids (see above), not  
restricted to K-12, much less  
to Home Schooling

. . . Conferences (2)

. . . Curriculum (1)

. . . Distance Learning (8)

. . . Islamic@

. . . Magazines (10)

. . . Opposing Views (1)

. . . Organizations (37)

. . . . Christian@

. . . Personal Experience (7)

. . . Unschooling (6)

. . . Web Directories (11)

. . . Usenet (2)

**. . Humanities@**

. . . Classics@

. . . Organizations (15)

. . . . Student (15)

. . . Teacher Resources (2)

. . . History@

. . . . Courses (15)

. . . . Curriculum Standards (5)

. . . . Fairs and Competitions (3)

. . . . Social Studies@

. . . . . Archaeology@

. . . . . Companies@

. . . . . Presenters (4)

. . . . . Professional Development

. . . . . Teaching &amp; Learning Aids

. . . . . Curriculum Standards (25)

. . . . . History@

. . . . . Lesson Plans (31)

. . . . . History@

. . . . . WebQuests (14)

. . . . . Organizations (13)

. . . . . Web Directories (6)

. . . . . Teacher Resources (25)

. . . . . Classroom Projects (5)

. . . . . Lesson Plans (13)

. . . Literature@

. . . . Courses (4)

. . . Reading@

. . . . Children's Reading Lists@

. . . . . Young Adult@

. . . . . Companies@

. . . . . Professional Development

. . . . . Teaching &amp; Learning Aids

. . . . . Lesson Plans (8)

. . . . . Literacy (37)

. . . . . Organizations (16)

. . . . . Organizations (5)

. . . . . Whole Language (3)

. . . . . Young Adult Reading Lists@

. . . . . Teacher Resources (40)

. . . . . Classroom Projects (2)

. . . . . Lesson Plans (27)

. . . Philosophy@

. . Institutes (7)

. . **Issues (31)**

. . . Bullying@

. . . Class Size@

. . . Religion in Public Schools (24)

. . . . School Prayer (16)

. . . School Choice@

. . . School Phobia@

. . . Shyness@

. . . School Violence@

. . . . Organizations (10)

. . . . School Shootings (80)

**. . Mathematics@**

. . . Academic Competitions (22)

. . . Courses (1)

. . . . Statistics@

. . . Curriculum Standards (38)

. . . Exercises@

. . . Algebra@

. . . Geometry@

. . . Online Equation Solvers (11)

. . . . Online Cryptarithmic &amp;

Alphabetic Puzzle Solvers@

. . . Pre-Algebra@

. . . Organizations (6)

. . . Programs (26)

. . . Magnet Schools@

. . . Summer (7)

. . . School Departments (8)

. . . Teaching (32)

. . . Lesson Plans (12)

. . Newspapers (142)

. . . Individual School Papers (130)

**. . Organizations (202)**

. . . Administrators@

. . . Alternative (7)

. . . Alumnae/I Associations@

. . . Reunions@

. . . . Alumnae/I Associations@

. . . Business@

. . . Charter School@

. . . Computer Clubs@

. . . Home Schooling@

. . . International Schools@

. . . Lesbian, Gay, and Bisexual@

. . . Parent@

. . . Reading@

. . . Literacy@

. . . School Associations (24)

. . . School Board Associations (13)

. . . School Funding@

. . . Social Studies@

. . . Violent Crime Prevention@

. . Parental Involvement@

. . Physical Education@

- . . Programs (124)
- . . Reading (66)
- . . Religion@
- . . School Funding (29)
- . . Schools (23726)
- . . Science@
- . . Social Science@
- . . Social Studies (111)
- . . Student Resources (30)
- . . Teaching (153)
- . . Teaching & Learning Aids@
- . . Web Directories (23)
- . . Usenet (22)
- . **Literacy** (12)
- . . Adult & Continuing Education@
- . . K-12@
- . . Media Literacy@
  - SN Following various branches of the hierarchy
  - . . . Culture Jammers@
  - . . . . Anarchism@
  - . . . . . Magazines (7)
  - . . . . . . United Kingdom@
  - . . . . . . United States@
  - . . . . . . Political Theory@
  - . . . . . . History (11)
  - . . . . . . Spanish Civil War@
  - . . . . . . Libraries (6)
  - . . . . . . Organizations@
  - . . . . . . People (28)
  - . . . . . . Situationists@
  - . . . . . . Bakunin, Mikhail Aleksandrovich (1814-1876)@
  - . . . . . . Bookchin, Murray (1921- ) (5)
  - . . . . . . Chomsky, Noam@
  - . . . . . . Durruti, Buenaventura (1896-1936)@
  - . . . . . . Godwin, William (1756-1836) (3)
  - . . . . . . Goldman, Emma (1869-1940)@
  - . . . . . . Kropotkin, Peter (1842 - 1921) (4)
  - . . . . . . Makhno, Nestor (1889-1934)@
  - . . . . . . Malatesta, Errico (1853-1932) (2)
  - . . . . . . Michel, Louise (1830-1905)@
  - . . . . . . Parsons, Lucy (1853-1942) (3)
  - . . . . . . Prominent Anarchists & Left-Libertarians -
  - . . . . . . Proudhon, Pierre-Joseph (1809-1865) (2)

- . . . . . . Stirner, Max (1806-1856)@
- . . . . . . Political Opinion@
- . . . . . . Web Directories (3)
- . . . . . . Publishers@
- . . . . . . Anti-Television@
- . . . . . . Cacophonists@
- . . . . . . Guerilla Art@
- . . . . . . Billboard Liberation (3)
- . . . . . . Culture Jammers@
- . . . . . . Graffiti@
- . . . . . . Anti-Graffiti (6)
- . . . . . . Companies@
- . . . . . . Exhibits (77)
- . . . . . . Magazines (11)
- . . . . . . Virtual Graffiti (4)
- . . . . . . Writers and Crews (38)
- . . . . . . Postering (6)
- . . . Youth Audience Issues (10)
- . . . Television Ratings@
- . . Organizations (9)
- . **News and Media** (84)
- . . Books@
- . . Retail@
- . . Publishers@
- . . Academic@
- . . . . . College & Univ. Presses@
- . . . . . Textbooks@
- . . . . . Booksellers@
- . . . . . Home Schooling Materials (21)
- . . . . . Christian (9)
- . . . . . Language Education@
- . . . . . English as a 2nd Language
- . . . . . Special Education Books (9)
- . . . . . Supplementary Materials (91)
- . . . . . Forensics and Debate (5)
- . . . . . Teacher Aids (12)
- . . . . . Textbooks (88)
- . . . . . Vocational Training Books
- . . . . . Textbooks (43)
- . . . . . Higher Education (53)
- . . . . . K-12 (17)
- . . . . . Titles (98)
- . . . Textbooks (8)
- . . . College & University Bookstores@
- . . . Science@
- . . . Used (19)
- . . . Titles@
- . . . Administration (2)
- . . . College and University (26)
- . . . Admissions (10)
- . . . Athletic Recruiting@
- . . . Financial Aid (6)
- . . . University Life (4)
- . . . Distance Learning (3)
- . . . Educating Your Child@

- . . . . . Home Schooling (8)
- . . . . . Reform (3)
- . . . . . School Directories (4)
- . . . . . Study Guides (7)
- . . . . . Teaching (21)
- . . . . . Reading (9)
- . . . . . Test Preparation (13)
- . . . College and University@
- . . . Individual Schools (410)
- . . . Magazines (175)
- . . . Individual Schools (139)
- . . . Newspapers (437)
- . . . Business Schools (8)
- . . . Medical Schools (2)
- . . . Newswires@
- . . . Web Directories (2)
- . . . Radio Stations@
- . . . Countries (12)
- . . . Eastern United States (169)
- . . . Internet Broadcasts@
- . . . Western United States (110)
- . . Journals@
- . . . •Educational Theory@
- . . . Instructional Technology@
- . . . Music@
- . . . Musicology@
- . . . . . Ethnomusicology@
- . . Magazines (50)
- . . College and University@
- . . . Individual Schools (139)
- . . . English as a Second Language@
- . . . Home Schooling@
- . . . Online Teaching and Learning@
- . . Newsletters (12)
- . . Newspapers (3)
- . . College and University@
- . . . Business Schools (8)
- . . . Medical Schools (2)
- . . . Newswires@
- . . . Web Directories (2)
- . . . K-12@
- . . . Individual School Papers (130)
- . . Policy@
- . . Television@
- . . . Distance Learning@
- . . . Educational Programs@
- . . . Television Curriculum (5)
- . . . High School Stations@
- . . . Shows@

## •Education

### . Organizations (2945)

. . . Alumnae/I Associations (459)  
 . . . Colleges and Universities (207)  
 . . . High Schools (239)  
 . . . . Reunions@  
 . . . . Reunions@  
 . . . . College and University (8)  
 . . . . High Schools (274)  
 . . Arts@  
 . . . College and University@  
 . . Business@  
 . . . Case Research (4)  
 . . . College and University@  
 . . . K-12@  
 . . Christian@  
 . . . Catholic@  
 . . . Home Schooling@  
 . . . Student (171)  
 . . . . Baptist@  
 . . . . Catholic@  
 . . . . Church of Christ@  
 . . . . Fraternities and Sororities (20)  
 . . . . Presbyterian@  
 . . . . Presbyterian Church in America@  
 . . . . United Methodist Church@  
 . . Early Childhood Education@  
 . . Equity@  
 . . Financial Aid@  
 . . **Higher Education@**  
 . . . Alumnae/I Associations@  
 . . . Community College (9)  
 . . . Graduate Education@  
 . . . Transfer Student (3)  
 . . **K-12@**  
 . . . Administrators@  
 . . . Alternative (7)  
 . . . Alumnae/I Associations@  
 . . . Business@  
 . . . Charter School@  
 . . . Computer Clubs@  
 . . . Home Schooling@  
 . . . . Christian@  
 . . . International Schools@  
 . . . Lesbian, Gay, and Bisexual@  
 . . . Parent@  
 . . . Reading@  
 . . . . Literacy@  
 . . . School Associations (24)  
 . . . School Board Associations (13)  
 . . . School Funding@  
 . . . . Technology Funding@  
 . . . Social Studies@  
 . . . Violent Crime Prevention@  
 . . Languages@  
 . . . English as a Second Language@

. . . . Student (1)  
 . . . . Teaching@  
 . . Lesbian, Gay and Bisexual@  
 . . Literacy@  
 . . Math@  
 . . Music@  
 . . . Teaching@  
 . . . . Kodaly Method@  
 . . Online Teaching and Learning@  
 . . Policy@  
 . . Professional (197)  
 . . Administrators (23)  
 . . Consultants (2)  
 . . Faculty (19)  
 . . Unions (77)  
 . . Reading (8)  
 . . Reform@  
 . . Rural Education@  
 . . Science@  
 . . . Engineering@  
 . . Student (2131)  
 . . Animal Rights@  
 . . Anthropology@  
 . . Architecture@  
 . . Community Service@  
 . . Computer Science@  
 . . Cultural (361)  
 . . . African (2)  
 . . . African-American@  
 . . . Arabic (5)  
 . . . Armenian (8)  
 . . . Asian American (23)  
 . . . Azerbaijani (1)  
 . . . Bangladesh (4)  
 . . . Bosnian (2)  
 . . . Brazilian (2)  
 . . . Cambodian (1)  
 . . . Caribbean (3)  
 . . . Chinese (23)  
 . . . Cypriot (2)  
 . . . Egyptian (2)  
 . . . Filipino (6)  
 . . . Finnish (1)  
 . . . Haitian (2)  
 . . . Hawaiian (3)  
 . . . Hellenic (27)  
 . . . Hispanic and Latino (32)  
 . . . . Law@  
 . . . Indian (21)  
 . . . Indonesian (14)  
 . . . Italian (3)  
 . . . Japanese (2)  
 . . . Korean (24)  
 . . . Lebanese (3)  
 . . . Malaysian (16)  
 . . . Mexican (3)  
 . . . Multicultural (7)  
 . . . Native American@

. . . . Pakistani (9)  
 . . . . Romanian (8)  
 . . . . Scandinavian (1)  
 . . . . Singaporean (29)  
 . . . . Sri Lankan (2)  
 . . . . Taiwanese (12)  
 . . . . Thai (21)  
 . . . . Turkish (12)  
 . . . . Vietnamese (22)  
 . . Economics@  
 . . . Accounting and Auditing@  
 . . Engineering@  
 . . . Materials Science@  
 . . . Mechanical Engineering@  
 . . Environmental@  
 . . European Union@  
 . . . Austria@  
 . . . Germany@  
 . . . Government (1)  
 . . . Netherlands@  
 . . Fraternities and Sororities (1443)  
 . . . Apparel@  
 . . . Christian@  
 . . . Collectibles@  
 . . . Directories (4)  
 . . . Hazing (2)  
 . . . Professional (128)  
 . . . . Music@  
 . . . Service (138)  
 . . . Social (1154)  
 . . . . Interfraternity Councils (14)  
 . . . Usenet (3)  
 . . Government (90)  
 . . Graduate (13)  
 . . High School (4)  
 . . Homelessness@  
 . . Honor Societies (74)  
 . . Law@  
 . . . Hispanic and Latino (7)  
 . . . Pre-Law Societies (2)  
 . . . Women@  
 . . Lesbian, Gay, and Bisexual@  
 . . . Alumni Associations (9)  
 . . . Campus Support Offices@  
 . . Medical@  
 . . . Interns and Residents (3)  
 . . Optometry@  
 . . Nursing@  
 . . Political (24)  
 . . . Democratic Party@  
 . . . Federalist Society@  
 . . . Libertarian Party@  
 . . . Reform Party@  
 . . . Republican Party@  
 . . . Young Americans for Freedom  
 . . Religious (5)  
 . . . Christian@  
 . . . Hindu@



- |   |   |   |
|---|---|---|
| <ul style="list-style-type: none"> <li>. . . . Islamic@</li> <li>. . . . Jewish@</li> <li>. . . . Sikh@</li> <li>. . . Residence Hall Associations (9)</li> <li>. . . Web Directories (2)</li> <li>. . Student Affairs (15)</li> <li>. . . Web Directories (1)</li> <li>. . Trade Associations (3)</li> <li>. . Vocational Education@</li> <li>. . . Private Industry Councils@</li> <li>. . . Training &amp; Enterprise Councils@</li> <li>. . Women@</li> <li>. . . Law@</li> <li>. <b>Policy</b> (52)</li> <li>. . College &amp; Univ. Dpts &amp; Programs</li> <li>. . News and Media (3)</li> <li>. . Organizations (6)</li> <li>. . Research Centers (23)</li> <li>. . . Higher Education (6)</li> <li>. <b>Programs</b> (309)</li> <li>. . Co-operative Programs (8)</li> <li>. . K-12@</li> <li>. . . Environment and Nature@</li> <li>. . . National Guard Youth Challenge Program@</li> <li>. . . . STARBASE (3)</li> <li>. . . Student Exchange@</li> <li>. . . . Study Abroad@</li> <li>. . . . . Companies@</li> <li>. . . . . Student Exchange@</li> <li>. . . . . Admissions@</li> <li>. . . . . Counseling@</li> <li>. . . . . Credential Evaluation (8)</li> <li>. . . . . Financial Aid@</li> <li>. . . Summer Programs (56)</li> <li>. . . . Learning Disabilities@</li> <li>. . . . Math@</li> <li>. . . . Science (22)</li> <li>. . . . . Math@</li> <li>. . Model Congress@</li> <li>. . Model United Nations@</li> <li>. . College and University (35)</li> <li>. . . Events (12)</li> <li>. . Events (2)</li> <li>. . . College and University@</li> <li>. . . High School@</li> <li>. . High School (33)</li> <li>. . . Events (16)</li> <li>. . National Writing Project (27)</li> <li>. . Student Exchange (78)</li> <li>. . Study Abroad (157)</li> <li>. . Summer Programs (18)</li> <li>. . K-12</li> <li>. <b>Reform</b> (67)</li> <li>. . Books@</li> <li>. . Class Size (6)</li> <li>. . Organizations (35)</li> </ul> | <ul style="list-style-type: none"> <li>. . School Choice (9)</li> <li>. <b>Special Education</b> (167)</li> <li>. . Companies@</li> <li>. . Administrative Software@</li> <li>. . Teaching &amp; Learning Aids (5)</li> <li>. . . Publishers@</li> <li>. . Conductive Education (19)</li> <li>. . Organizations (4)</li> <li>. . Schools and Institutes (14)</li> <li>. . Employment@</li> <li>. . Inclusion (7)</li> <li>. . Institutes (29)</li> <li>. . College and Univ. Departments</li> <li>. . Conductive Education</li> <li>. . Learning Disabilities@</li> <li>. . Attention Deficit Disorder@</li> <li>. . . Companies@</li> <li>. . . Conferences (1)</li> <li>. . . Organizations (8)</li> <li>. . . Web Directories (2)</li> <li>. . Companies@</li> <li>. . . Attention Deficit Disorder@</li> <li>. . . Dyslexia@</li> <li>. . Dyslexia@</li> <li>. . . Organizations (12)</li> <li>. . . Shopping and Services@</li> <li>. . Dysphagia@</li> <li>. . Dyspraxia@</li> <li>. . Gifted &amp; Learning Disabled (16)</li> <li>. . Institutes@</li> <li>. . College &amp; Univ. Departments</li> <li>. . Conductive Education@</li> <li>. . Organizations (20)</li> <li>. . Dyslexia@</li> <li>. . Schools (31)</li> <li>. . K-12@</li> <li>. . Summer Programs (3)</li> <li>. . U.S. Boarding Schools (9)</li> <li>. . Organizations (18)</li> <li>. . Conductive Education@</li> <li>. Schools (69)</li> <li>. . Adult &amp; Continuing Education@</li> <li>. . Colleges and Universities (3)</li> <li>. . Conductive Education@</li> <li>. . Early Education (16)</li> <li>. . Schools for the Blind@</li> <li>. . Schools for the Deaf@</li> <li>. . K-12 (46)</li> <li>. . Learning Disabilities</li> <li>. . Schools for the Blind@</li> <li>. . Schools for the Deaf@</li> <li>. . Schools for the Blind@</li> <li>. . Schools for the Deaf@</li> <li>. . Web Directories (3)</li> <li>. . Usenet</li> <li>. <b>Standards and Testing</b> (63)</li> <li>. . ACT (1)</li> <li>. . GED (2)</li> </ul> | <ul style="list-style-type: none"> <li>. . . GED Preparation Companies@</li> <li>. . GMAT (3)</li> <li>. . GRE (3)</li> <li>. . IELTS@</li> <li>. . K-12 Curriculum Standards@</li> <li>. . . By Region (55)</li> <li>. . . By Subject (14)</li> <li>. . . Agriculture@</li> <li>. . . Arts@</li> <li>. . . Business@</li> <li>. . . Dance@</li> <li>. . . English Language Arts@</li> <li>. . . Health@</li> <li>. . . History@</li> <li>. . . Languages@</li> <li>. . . Library &amp; Information Literacy@</li> <li>. . . Mathematics@</li> <li>. . . Music@</li> <li>. . . Physical Education@</li> <li>. . . Science@</li> <li>. . . Social Studies@</li> <li>. . LSAT (4)</li> <li>. . MCAT (5)</li> <li>. . Comapnies@</li> <li>. . Occupational Standards@</li> <li>. . Australian Qualific Framework@</li> <li>. . U.K. Natl Vocational Qualific.@</li> <li>. . U.S. National Skill Standards (5)</li> <li>. . SSAT (1)</li> <li>. . Test Preparation Companies@</li> <li>. . Testing Companies@</li> <li>. . Computer-Based Testing Serv.</li> <li>. . Software@</li> <li>. . TOEFL@</li> <li>. . TOEIC@</li> <li>. . Voluntary National Testing (4)</li> <li>. <b>Statistics</b> (6)</li> <li>. <b>Teaching</b> (91)</li> <li>. . Arts@</li> <li>. . <b>Curriculum</b> (9)</li> <li>. . Companies@</li> <li>. . . By Subject (9)</li> <li>. . . . Art@</li> <li>. . . . Computers and Technology@</li> <li>. . . . Health and Fitness@</li> <li>. . . . Languages@</li> <li>. . . . Life Skills@</li> <li>. . . . Mathematics@</li> <li>. . . . Reading and Writing@</li> <li>. . . . Science@</li> <li>. . . . Social Studies@</li> </ul> |
|---|---|---|

## •Education

- . **Teaching** (91)
  - . . . *Arts@*
  - . . . *Curriculum* (9)
  - . . . . *Companies@*
  - . . . . . Christian Home Schooling@
  - . . . . . Preschool (9)
  - . . . . . School Reform (2)
  - . . . . . Textbooks@
    - . . . . . . Retail
    - . . . . . . Used (3)
  - . . . . Home Schooling@
  - . . . . K-12 Lesson Plans@
    - . . . . . Arts@
    - . . . . . History@
    - . . . . . Literature@
    - . . . . . Math@
    - . . . . . News and Current Events (5)
    - . . . . . Reading@
    - . . . . . Science@
    - . . . . . Social Studies@
    - . . . . . Television Curriculum@
    - . . . . . Theater@
    - . . . . . WebQuests@
    - . . . . . Writing@
  - . . Early Childhood Education@
  - . . English@
  - . . English as a Second Language@
  - . . Evaluation (4)
  - . . Forums (3)
  - . . International Teaching (7)
    - . . . English as a Second Language@
    - . . . International Schools@
  - . . Internet@
  - . . Comapnies@
  - . . **K-12@**
    - . . . Lesson Plans (51)
      - . . . . . Arts@
        - . . . . . . Theater@
        - . . . . . . History@
        - . . . . . . Literature@
        - . . . . . . Math@
        - . . . . . . News and Current Events (5)
        - . . . . . . Reading@
        - . . . . . . Science@
        - . . . . . . Social Studies@
          - . . . . . . . History
          - . . . . . . . Webquests (15)
        - . . . . . . Television Curriculum@
        - . . . . . . Theater@
        - . . . . . . WebQuests@
          - . . . . . . . Collections (8)
          - . . . . . . . Science@
          - . . . . . . . Social Studies@
          - . . . . . . . Writing@
      - . . . School Library Resources@
      - . . . . Children's Literature@

- . . . . Young Adult Literature@
- . . . Substitute Teaching (6)
- . . . Teacher Certification (53)
  - . . . . U. S. States (50)
- . . . Web Directories (24)
  - . . . Math@
  - . . . Organizations (7)
  - . . . Professional Development Services (@)
  - . . Music@
    - . . . Organizations (40)
    - . . . Web Directories (3)
  - . . Online Teaching and Learning@
    - . . . . Conferences (8)
    - . . . . Corporate Programs (10)
    - . . . . Courses (10)
    - . . . . Educational MOOs (7)
    - . . . . Magazines (3)
    - . . . . Organizations (14)
    - . . . . Papers (4)
    - . . . . Projects (20)
    - . . . . Research Institutes (3)
    - . . . . Teacher Resources (45)
    - . . . . Telementoring (12)
    - . . . . Web Directories (8)
    - . . . . WebQuests (19)
  - . . Science@
    - . . . Science@
      - . . . . Biology@
        - . . . . . Lesson Plans (32)
        - . . . . . WebQuests (9)
      - . . . . Space@
    - . . Teacher Education (24)
      - . . . College & Univ. Departments@
      - . . . Distance Learning@
      - . . . Montessori Method@
      - . . . Professional Development Services@
      - . . . Teacher Certification@
  - . . **Theory and Methods** (644)
    - . . . Block Scheduling (6)
    - . . . Character Education (18)
      - . . . . Articles and Papers (4)
      - . . . . Bibliographies (3)
      - . . . . Companies@
      - . . . . Institutes (1)
    - . . . Charter Schools (15)
      - . . . . Organizations@
    - . . . Critical Thinking@
    - . . . Logic@
    - . . . Experiential Education (4)
    - . . . Feminist Pedagogy (6)

- . . Home Schooling (113)
- . . Institutes (240)
  - . . . College & Univ. Departments
  - . . . Early Childhood Education@
    - . . . . K-12@
  - . . . Journals (6)
  - . . . Libraries (24)
  - . . . Montessori Method (146)
  - . . . Schools
    - . . . . Teacher Education
  - . . . Multiple Intelligences@
  - . . . Service Learning (10)
  - . . . Theorists (12)
    - . . . . Dewey, John@
    - . . . . Freire, Paulo (8)
    - . . . . Montessori, Maria (2)
    - . . . . Steiner, Rudolf@
  - . . . Waldorf Method (50)
    - . . . . Schools@
  - . . . WebQuests@
- . **Web Directories** (45)

## Entertainment

SN Cool Links, Movies, Humor,

Music ...

### . Consumer Electronics

- . . Audio
  - . . . MP3 Players
    - . . . . ▶Automotive (9)

### . Humor (4865)

- . . ▶Cars (40)
- . . •Education (43)
- . . ▶Traffic Cones (2)

### . Music

- . . •Education
  - . . . By Instrument (3)
  - . . . Camps (29)
  - . . . Conferences (4)
  - . . . Courses and Lessons (8)
  - . . . Ear Training (5)
  - . . . Festivals (9)
  - . . . Jazz@
  - . . . Journals (2)
  - . . . K-12 Curriculum Standards (6)
  - . . . Kodály Method (18)
  - . . . Organizations (17)
  - . . . Schools (267)
  - . . . Shopping and Services@
  - . . . Suzuki Method (10)
  - . . . Teaching (50)
  - . . . Web Directories (6)
- . . Instruments
  - . . . Stringed Instruments
    - . . . . Guitar
      - . . . . . •Education

**Government**

- SN Elections, Military, Law, Taxes...

**Law**

- SN No narrower category

- Transportation

- Continuing Legal Education

- Companies@

- Court Reporting@

- Law Schools@..

**Military (867)**

- Weapons and Equipment (269)

- ▶Aircraft@

- ▶Armored Vehicles (17)

- ▶Ships (9)

**U.S. Government**

- Executive Branch

- Departments and Agencies

- Department of Education

- ▶Department of Transportation

- Military

- Training and Education (18)

- Navy

- ▶Fleets and Ships (247)

- ▶Submarines (83)

- Training and Education

- State Government

- Massachusetts@

- Department of Education

- ▶Department of Transportation

**Health**

- SN Medicine, Diseases, Drugs,  
Fitness ...

- Alternative Medicine (499)**

- Apitherapy (6)

- Applied Kinesiology (8)

- Aromatherapy (14)

- Companies@

- Organizations (3)

- Ayurveda@

- Companies@

- Herbal Supplements@

- Brand Names (8)

- Products@

- Biofeedback (8)

- Companies@

- Booksellers@

- Titles (33)

- Herbs (3)

- Massage (3)

- Yoga@

- Breathwork (9)

- Companies@

- Holotropic (3)

- Rebirthing (5)

- Vivation (1)

- Rebirthing (6)

- Companies@

- Business to Business

- Massage Therapy (132)

- Equipment and Supplies (33)

- Manufacturers (14)

- Training and Development (97)

- Massage Therapy@

- Buteyko@

- Companies@

- Chinese Medicine

- Acupuncture (24)

- Companies@

- Electro-acupuncture (6)

- Software@

- Supplies (5)

- Organizations (13)

- Chinese Herbs@

- Brand Names (26)

- Kombucha (3)

- Products@

- Medical Schools (28)

- Qigong (23)

- Companies@

- Falun Gong@

- Organizations (16)

- International Yan Xin Qigong  
Association (7)

- Chiropractic@

- Business to Business@

- Equipment and Supplies (13)

- Manufacturers (2)

- Internet Services (4)

- Software (6)

- Journals (5)

- Professional Organizations (33)

- Schools, Departments, &  
Programs (17)

- Shopping and Services@

- By Region (963)

- SN A long list of

- geographical. entities

**•Education**

- K-12 Curriculum Standards (24)

- Medicine@

- Midwifery@

- Nursing@

- Shopping and Services@

- Emergency Services (476)**

- Companies@

- ▶Air Ambulance Services (34)

- Supplies and Equipment

- ▶Vehicles

- Ambulances

**•Fitness**

- Aerobics (9)

- Jazzercise (3)

- Books@

- Yoga@

- Business to Business@

- Career Training (13)

- Yoga@

- Equipment Distributors (2)

- Equipment Manufacturers (30)

- Health Club Management  
Software (16)

- Events (6)

- Health Clubs@

- By Region (1060)

- Directories (2)

- Indoor Rowing@

- Institutes (8)

- Magazines (16)

- Bodybuilding@

- Yoga@

- Organizations (36)

- Professional@

- Physical Education (PE) (48)

- College & Univ. Departments

- K-12 (24)

- Curriculum Standards (15)

- Departments (1)

- Organizations (2)

- Organizations (2)

- Mental Health (755)**

- Bereavement (56)

- ▶Aircraft Accident Support (1)

**•Nutrition**

- Institutes (49)

- Schools, Departments, &  
Programs (31)

- Public Health and Safety (2206)**

- ▶Driving Safety@

- ▶Helmets (4)

- Injury Prevention (20)

- SN No X-ref to Accidents.

- ▶Travel Health and Medicine (24)**

- Deep Vein Thrombosis@

- Jet Lag@

- Motion Sickness@

- Shopping and Services@

- Tropical Diseases@

- Cholera@

- Dengue Fever@

- Ebola@

- Giardia@

- Hepatitis@

- Lassa Fever@

- Leprosy@

- Malaria@

- ...

**News and Media**

SN Full Coverage, Newspapers, TV

**. Industry Information**

- . . Journalism
- . . . •Education
- . . . . Career and Continuing (5)
- . . . . College and University (96)
- . . . . Organizations (4)
- . . . . Web Directories (1)
- . . •Media Education
- . . . Broadcasting@
- . . . Career and Continuing (3)
- . . . College and Univ. Departments
- . . . Journalism@
- . ▶Automotive@
- . •College and University (898)
- . Traffic and Road Conditions@
- . ▶Transportation@
- . ▶Travel@

**Recreation & Sports**

SN Sports, Travel, Autos, Outdoors

**. ▶Automotive**

- . . Alternative Fuel Vehicles (147)
- . . Audio (33)
- . . Auto-Free Transportation@
- . . Booksellers@
- . . British Cars (45)
- . . Bumper Stickers (4)
- . . Business to Business@
- . . Buyer's Guides (44)
- . . Car Art@
- . . Charitable Vehicle Donation (9)
- . . Chats and Forums (51)
- . . Classic Cars (145)
- . . Classifieds@
- . . Clubs and Organizations (177)
- . . Concept Cars (10)
- . . Driving (184)
  - . . . •Driving Schools@
- . . Dune Buggies (16)
- . . Electric Vehicles@
- . . Emissions (21)
- . . Employment (2)
- . . Engineering@
- . . Events and Shows (58)
- . . Famous Cars (30)
- . . Financing (8)
- . . Four Wheel Drive (111)
- . . Hearses (16)
- . . History (88)
- . . Humor@
- . . Industry Information@
- . . Kit Cars (22)
- . . Lemon Law@
- . . License Plates (39)
- . . Licensing and Registr. Agencies
- . . Lowriders (63)

- . . Maintenance (13)
- . . Makes and Models (2968)
- . . Model Cars@
- . . Motorcycles (1012)
  - . . . Pictures
- . . Museums@
- . . News and Media (113)
- . . Police Vehicles@
- . . Racing@
- . . Recreational Vehicles@
- . . Shopping and Services@
- . . Software (2)
- . . Special Needs@
- . . Sport Utility Vehicles (48)
- . . Station Wagons (28)
- . . Technicians (3)
- . . Theft@
- . . Travel@
- . . Trucks (49)
- . . Web Directories (16)
- . . Women (8)
- . . Wrecks (4)
- . . FAQs (3)
- . •Dance@
  - . . •Education
- . •Hobbies (3034)
  - . . Models (615)
    - . . . ▶Aircraft (216)
    - . . . ▶Armored Fighting Vehicles
    - . . . ▶Boats and Yachts (57)
    - . . . ▶Cars (73)
      - . . . . ▶Motorcycles (1)
      - . . . . ▶Radio-Controlled (13)
      - . . . . ▶Trains and Railroads (175)
  - . ▶Motorcycles@
- . •Sports
  - . . •College and University
  - . . •Physical Education@
  - . . Science
    - . . . •Education
  - . ▶Travel
    - . . Air Travel
      - . . . Airlines@
        - . . . . United Airlines
          - . . . . . •Education

**Reference**

- SN Libraries, Dictionaries, Quotations ...
- . . Libraries
  - . . . Library & Information Science
    - . . . . •Education
      - . . . . . College and University (55)
    - . . . •Education Libraries@
      - . . . . ▶Transportation Libraries@

**Regional**

SN Countries, Regions, US States

**. U.S. States**

- . . •Massachusetts
  - . . . Massachusetts Locations
    - . . . . Metropolitan Areas
      - . . . . . Boston Metro
        - . . . . . . Business and Shopping
          - . . . . . . . Shopping and Services
            - . . . . . . . . ▶Automotive
              - . . . . . . . . •Driving Schools
      - . . . . Counties and Regions
      - . . . . Cities
        - . . . . . **Boston**
          - . . . . . . Local Web Directory
            - . . . . . . . Business & Shopping
            - . . . . . . . Community
              - . . . . . . . •Education
                - . . . . . . . . Adult, Career, & Continuing (16)
                  - . . . . . . . . . Business to Business@
                  - . . . . . . . . . Child Care Centers & Preschools@
                  - . . . . . . . . . College & University
                  - . . . . . . . . . Government Agencies
                  - . . . . . . . . . K-12 (22)
                  - . . . . . . . . . Language Schools (11)
                  - . . . . . . . . . Organizations (25)
                  - . . . . . . . . . Shopping & Services@
                  - . . . . . . . . . Employment
                  - . . . . . . . . . Entertainment & Arts
                  - . . . . . . . . . Health
                  - . . . . . . . . . News & Media
                  - . . . . . . . . . Real Estate
                  - . . . . . . . . . Recreation & Sports
                    - . . . . . . . . . ▶Travel & Transportation
                      - . . . . . . . . . . Airports (4)
                      - . . . . . . . . . . Bicycle Advocacy@
                      - . . . . . . . . . . Car Rentals@
                      - . . . . . . . . . . Lesb., Gays, & Bisex.@
                      - . . . . . . . . . . Limos & Shuttles@
                      - . . . . . . . . . . Local Cruises@
                      - . . . . . . . . . . Local Guides (58)
                      - . . . . . . . . . . Lodging (71)
                      - . . . . . . . . . . Maps and Views (12)
                      - . . . . . . . . . . Mass Transit (8)
                      - . . . . . . . . . . Points of Interest (6)
                      - . . . . . . . . . . Restaurants@
                      - . . . . . . . . . . Taxi Services@
                      - . . . . . . . . . . Tour Operators@
                      - . . . . . . . . . . Traffic & Road Cond.
                      - . . . . . . . . . . Transportation Agencies
                      - . . . . . . . . . . Travel Agents@
                      - . . . . . . . . . . Travelogues (2)
                      - . . . . . . . . . . Marketplace
                      - . . . . . . . . . . Yellow Pages
      - . . . State Web Directory
        - . . . . Area Guides
        - . . . . Arts & Humanities

- ..... Business & Economy
  - ..... Community & Culture
  - ..... Computers & Internet
  - ..... •Education
    - ..... Business to Business@
    - ..... Career and Vocational (3)
    - ..... College and University (96)
    - ..... K-12 (11)
    - ..... Organizations (6)
    - ..... Shopping and Services@
    - ..... Employment
    - ..... Entertainment
    - ..... Government
    - ..... Health
    - ..... News & Media
    - ..... Real Estate
    - ..... Recreation & Sports
    - ..... Science
    - ..... Social Science
    - ..... ▶Travel & Transportation
      - ..... Cities@
      - ..... Counties and Regions@
      - ..... Complete List@
      - ..... Airports (3)
      - ..... Car Rentals@
      - ..... Destination Guides (13)
      - ..... Highways and Roads (5)
      - ..... Lodging (8)
      - ..... Maps and Views (4)
      - ..... Mass Transit (3)
      - ..... Transportation Agencies (2)
      - ..... Transportation Organizations
      - ..... Travel Safety (1)
  - .. **New York**
    - ..... Cities
      - ..... **New York**
        - ..... Business and Shopping
        - ..... Shopping and Services
          - ..... ▶Automotive
          - ..... •Driving Schools \*\*\*
          - ..... •Education
            - ..... Adult, Career, & Continuing
            - ..... Business to Business@
            - ..... Child Care Centers & Preschools@
            - ..... College & University (633)
            - ..... Events (2)
            - ..... Government Agencies (3)
            - ..... K-12 (115)
            - ..... Language Schools (18)
            - ..... English 2nd Language
              - ..... German (2)
              - ..... Italian (1)
              - ..... Spanish (2)
            - ..... Organizations (47)
            - ..... Primary and Secondary (1)
            - ..... Programs (7)
            - ..... Shopping and Services@
- ..... ▶Travel & Transportation
  - ..... **Washington, D.C.**
    - ..... •Education
      - ..... Adult, Career, & Continuing
      - ..... Business to Business@
      - ..... Child Care Centers & Preschools@
      - ..... College and University (427)
      - ..... Events (4)
      - ..... Government Agencies (1)
      - ..... K-12 (467)
      - ..... Language Schools (13)
      - ..... English as a 2nd Language
        - ..... French (2)
      - ..... Literacy (1)
      - ..... Organizations (98)
      - ..... Programs (3)
      - ..... Shopping and Services@
      - ..... Special Education (13)
    - ..... ▶Travel & Transportation
  - ..... **Countries**
    - ..... **Canada**
      - ..... •Education
        - ..... Academic Competitions (8)
        - ..... Adult and Continuing Education (22)
        - ..... Bilingual (2)
        - ..... By Culture or Group (41)
        - ..... By Subject (11)
        - ..... Career and Vocational (47)
        - ..... Companies@
        - ..... Conferences (5)
        - ..... Correctional@
        - ..... Distance Learning (23)
        - ..... Early Childhood Education (8)
        - ..... Employment (8)
        - ..... Financial Aid (17)
        - ..... Government Agencies (12)
        - ..... Higher Education (239)
        - ..... Instructional Technology (19)
        - ..... Literacy (6)
        - ..... News and Media (9)
        - ..... Organizations (142)
        - ..... Policy (2)
        - ..... Primary and Secondary (498)
        - ..... Cities@
          - ..... Metropolitan Areas@
          - ..... Provinces and Territories@
        - ..... Complete List@
          - ..... Academic Competitions@
          - ..... Alternative (8)
          - ..... Arts@
          - ..... Curriculum Standards (9)
          - ..... Distance Learning@
          - ..... Environment and Nature@
          - ..... Gifted Youth (1)
          - ..... Home Schooling@
          - ..... Mathematics@
- ..... Newspapers@
  - ..... Organizations (20)
  - ..... Parental Involvement@
  - ..... Physical Education@
  - ..... Programs (23)
  - ..... Reading (4)
  - ..... School Funding (3)
  - ..... Schools (373)
  - ..... Science@
  - ..... Social Science@
  - ..... Social Studies (7)
  - ..... Student Resources (10)
  - ..... Teaching (22)
  - ..... Teaching & Learning Aids@
  - ..... Web Directories (2)
    - ..... Programs (34)
    - ..... Reform (1)
    - ..... Special Education (11)
    - ..... Standards and Testing (8)
    - ..... Teaching (11)
    - ..... Theory and Methods (38)
    - ..... Web Directories (5)
  - ..... ▶Travel & Transportation

**Science**

- SN Animals, Astronomy, Engineering ...
- . **Engineering** (4691)
  - . . ▶Aerospace Engineering@
  - . . ▶Automotive Engineering (36)
  - . . Civil Engineering (475)
    - . . . ▶Transportation (31)
      - . . . . Companies@
      - . . . . Institutes (25)
        - . . . . ▶Tunnels@
    - . . •Education (304)
      - . . . Courses (4)
      - . . . Distance Learning (11)
      - . . . Institutes (247)
      - . . . Organizations (19)
      - . . . Vocational Schools (11)
  - . ▶Naval Engineering (3)
- . **Mathematics**
  - . . •Education
    - . . . Academic Competitions (2)
    - . . . College and University (366)
      - . . . . Companies@
      - . . . . Conferences (3)
      - . . . . Courses (4)
      - . . . . Exercises@
      - . . . . Gender Equity@
      - . . . . K-12 (149)
        - . . . . . Academic Competitions (22)
          - . . . . . Courses (1)
          - . . . . . Curriculum Standards (40)
            - . . . . . Exercises@
            - . . . . . Organizations (6)
            - . . . . . Programs (27)
            - . . . . . School Departments (8)
            - . . . . . Teaching (31)
              - . . . . . Organizations (28)
            - . . . . . Software@
            - . . . . . Teaching (32)
            - . . . . . Web Directories (3)

**Social Science**

- SN Archaeology, Economics, Languages ...
- . **Anthropology and Archaeology**
  - SN No narrower category
  - Transportation
- . **Linguistics & Human Languages** (2814)
  - . . Languages (2235)
    - . . . •Education (266)
      - . . . . Chinese@
      - . . . . English@
      - . . . . English as a 2nd Language@
      - . . . . French@
      - . . . . German@
      - . . . . Italian@
      - . . . . Japanese@
      - . . . . Russian@
      - . . . . Spanish@
      - . . . . College and University Departments (48)
        - . . . . . Commercial Products@
        - . . . . . Conferences (2)
        - . . . . . Courses (19)
          - . . . . . K-12 (28)
            - . . . . . . Chinese@
              - . . . . . . Curriculum Standards (20)
                - . . . . . . English as a 2nd Language@
                - . . . . . . French@
                - . . . . . . German@
                - . . . . . . Magnet Schools@
                - . . . . . . Spanish@
              - . . . . . . Language Labs (7)
              - . . . . . . Language Schools (125)
              - . . . . . . Organizations (8)
              - . . . . . . Web Directories (9)
  - . . Specific Languages
    - . . . Chinese
      - . . . . •Education
    - . . . French
      - . . . . •Education
    - . . . German
      - . . . . •Education

**Society and Culture**

- SN People, Environment, Religion ...
- . **Disabilities**
  - . . Assistive Technology (54)
    - . . . ▶Automotive (3)
      - . . . . Universal Design@
  - . . Recreation and Sports
    - . . . ▶Travel
      - . . . . ▶Transportation Resources
- . **Environment and Nature** (7070)
  - . . Pollution (204)
    - . . . Air (60)
      - . . . . ▶Auto Emissions@
- . **Food and Drink**
  - . . Cooking
    - . . . •Culinary Education
      - . . . . Baking (7)
        - . . . . . Vegetarian (4)
- . **Religion and Spirituality**
  - . . Faiths and Practices
    - . . . Christianity
      - . . . . Denominations and Sects
        - . . . . . Catholic
          - . . . . . . •Education
            - . . . . . . . Colleges and Universities
              - . . . . . . . K-12 (268)
                - . . . . . . . Organizations (14)
                - . . . . . . . Seminaries (13)
      - . . . Islam
        - . . . . Hajj
          - . . . . . Makkah@
            - . . . . . . ▶Travel and Transportation

**Figures 5 and 6**

**Yahoo top level comparisons**

| Dewey Decimal   | Library of Congress  | Yahoo  |
|---|--|--|
| 000 Computers, information, & general reference<br>000 Computers, Internet, & systems<br>010 Bibliography<br>020 Library and information sci.<br>070 News media | A General works<br><br>QA Math, incl. computer science<br>Z Bibliography and library science | Computers & Internet Reference<br><br>News & Media   |
| 100 Philosophy & psychology   | B-BJ Philosophy. Psychology.   | A & H > Humanities > Philosophy  |
| 200 Religion  | BL-BX Religion   | Society and Culture > Religion and Spirituality  |
| 300 Social sciences<br><br>330 Econ, 380 Commerce<br>320 Pol. sci., 350 Pub. admin<br>340 Law<br>370 Education  | H Social sciences<br><br>HB-HJ Economics<br>J Political science<br>K Law<br>L Education      | Social Science<br>Society & Culture<br>Business & Economy<br>Government<br>Government > Law<br>Education |
| 400 Language  | P-PM Language  | Social Science > Linguistics and Human Languages   |
| 500 Natural sciences & math.  | Q Science  | Science (parts of)   |
| 600 Technology (Applied sciences)<br>610 Medical sciences and medicine<br>630 Agriculture<br>Most of 600  | R Medicine<br>S Agriculture<br>T Technology  | Health<br>Science > Agriculture<br>Science (parts of)  |
| 700 Arts & recreation   | M Music and books on music<br>N Fine arts<br>GV Recreation, leisure                          | Arts & Humanities<br>Entertainment<br>Recreation & Sports  |
| 800 Literature & rhetoric   | PN-PZ Literature   | Arts & Humanities > Humanities > Literature  |
| 900 Geography & history   | C-F History, G Geography   | A & H > Humanities > History<br>Regional   |
|   | U Military science<br>V Naval science  | Government > Military  |

**Figure 5. Dewey, Library of Congress and Yahoo classification compared**

Note: Comparison is easier at lower levels of the hierarchy since different schemes may agree on specific narrower fields but group them differently at the top level of the hierarchy



| <b>Yahoo Home</b>                           | <b>Subdivision of states</b>                   | <b>Subdivision of cities</b>   |
|---|--|--|
| Regional                                    | <b>Massachusetts Locations</b>                 |  |
|   | Metropolitan Areas                             |  |
|   | Counties and Regions                           |  |
|   | Cities   |  |
| (Yellow Pages)<br>(Maps)                    |  | <b>Find Businesses</b><br>Yellow Pages, Maps Driving Dir               |
|   |  | <b>City Guides</b>   |
| (Classifieds)                               |  | <b>Classifieds</b>   |
|   | Real Estate                                    | Real Estate (category, general)<br>Local Real Estate (listings search) |
|   | Employment                                     | Employment (category, general)<br>Local Jobs (job listings search)     |
|   |  | <b>Online Community</b>  |
| Reference                                   |  |  |
| Computers & Internet                        | Computers & Internet                           |  |
| News & Media                                | News & Media                                   | News & Media   |
| Entertainment                               | Entertainment                                  | Entertainment & Arts   |
| Arts & Humanities                           | Arts & Humanities                              | (See Entertainment & Arts below)                                       |
| Education                                   | Education                                      | Education  |
| Recreation & Sports                         | Recreation & Sports<br>Travel & Transportation | Recreation & Sports<br>Travel & Transportation                         |
| Business & Economy<br>(Auctions) (Shopping) | Business & Economy                             | Business & Shopping  |
| Government                                  |  |  |
| Society & Culture                           | Community & Culture                            | Community  |
| Social Science                              | Social Science                                 |  |
| Health                                      | Health   | Health   |
| Science                                     | Science  |  |

**Figure 6. Yahoo Home compared with *State* subdivision and *City* subdivision**

My arrangement (compare Figure 1b). The items in ( ) in column 1 are not categories in Yahoo Home, but links under the search box; there are many other such links. Most of the subdivisions under states and cities are accessed through a drop-down box titled State/Local Web Sites.

**Figure 7. Subdivisions of selected Yahoo categories for comparison and analysis**  
**Arranged by ease of analysis.**

|   |                              |
|---|------------------------------|
| Home > Reference ><br><b>Libraries</b>      |                              |
| <b>Categories</b> (divided into two groups) |                              |
| Conferences (9)                             | Organizations (59)           |
| Countries (27)                              | Professional Resources (112) |
| History (6)                                 | Serials (6)                  |
| Librarians (29)                             | Web Directories (19)         |
| Library and Information Science (306)       |                              |
| Academic Libraries (451)                    | Military Libraries@          |
| Archives@                                   | Music Libraries@             |
| Arts Libraries@                             | National Libraries (33)      |
| Business Libraries@                         | Native American Libraries@   |
| Commercial Library Services@                | Performing Arts Libraries@   |
| Dance Libraries@                            | Philatelic Libraries@        |
| Digital Libraries (74)                      | Physics Libraries@           |
| Education Libraries@                        | Presidential Libraries@      |
| Environmental Libraries@                    | Prison Libraries@            |
| Government Documents@                       | Public Libraries (3494)      |
| Health Libraries@                           | Religious Libraries@         |
| Intellectual Property Libraries@            | School Libraries (35)        |
| Internet Filtering in Libraries@            | Science Libraries@           |
| Law Libraries@                              | Social Science Libraries@    |
| Lesbian, Gay and Bisexual@                  | Special Collections (41)     |
| Libraries for the Blind@                    | Sports Libraries@            |
| Literary Libraries@                         | Theological Libraries@       |
| Literature@                                 | Transportation Libraries@    |
| Map Libraries@                              | U.S. State Libraries (49)    |
| Masonic Libraries@                          |                              |

|  |   |
|--|---|
| Home >   |   |
| <b>News and Media</b>  |   |
| <b>Categories</b> (divided into two groups)  |   |
| By Region (21222)<br>Columns and Columnists (286)<br>Commercial Services@<br>Content Ratings@<br>Industry Information (824)<br>Internet Broadcasts (400)<br>Journals (33)  | Magazines (3835)<br>Newspapers (8419)<br>Photojournalism@<br>Radio (9418)<br>Television (15900)<br>Web Directories (100)  |
| Arts and Humanities@<br>Automotive@<br>Business (133)<br>College and University (844)<br>Computers and Internet@<br>Crime@<br>Cultures and Groups (15)<br>Disabilities@<br>Education@<br>Entertainment@<br>Environment and Nature@<br>Good News (9)<br>Government@<br>Health@<br>History@<br>Home and Garden@<br>Humor (219) | Law@<br>Outdoors@<br>Personalized News (14)<br>Philanthropy@<br>Politics@<br>Real Estate@<br>Religion@<br>Science@<br>Sexuality@<br>Sports@<br>Technology (69)<br>Traffic and Road Conditions@<br>Transportation@<br>Travel@<br>Weather (1087)<br>Weird News (20)<br>World (72) |

Home > Entertainment >

## **Humor**

**Categories** (divided into two groups)

|                             |                      |
|-----------------------------|----------------------|
| Archives (122)              | Lists (65)           |
| Chats and Forums (34)       | News and Media@      |
| Columns and Columnists (84) | Poetry@              |
| Comedy@                     | Quizzes and Tests@   |
| Comic Strips@               | Quotations@          |
| Companies@                  | Stories (21)         |
| Jokes (371)                 | Web Directories (20) |

|                                |                                   |
|--------------------------------|-----------------------------------|
| Advertising (41)               | Job Humor (90)                    |
| Advice (127)                   | Military (3)                      |
| Animals (176)                  | Movies and Film (62)              |
| Begging (40)                   | Murphy's Laws (6)                 |
| Bitterness (38)                | Music (93)                        |
| Bizarre (248)                  | Names (25)                        |
| Boredom (4)                    | Parenting (20)                    |
| Bubblewrap (4)                 | Parody (350)                      |
| Cars (40)                      | People (19)                       |
| Clean Humor (12)               | Philosophy (18)                   |
| Codes (8)                      | Politics (16)                     |
| Computers and Internet (631)   | Procrastination (13)              |
| Cultures and Groups (177)      | Rants (75)                        |
| Distorted Pictures (14)        | Religion (216)                    |
| Drugs and Addictions (16)      | Science (109)                     |
| Duct Tape (11)                 | Science Fiction and Fantasy (235) |
| Education (43)                 | Sex (130)                         |
| Fights (9)                     | Sports (11)                       |
| Food and Drink (139)           | Stupidity (70)                    |
| Furniture (5)                  | Tasteless (370)                   |
| Gender Wars (37)               | Traffic Cones (2)                 |
| Hair (27)                      | Useless Pages (123)               |
| Health and Medicine (22)       | Warning Labels (3)                |
| Holidays and Observances (109) | Wedding and Marriage (13)         |
| Horoscopes (19)                |                                   |

Home > Entertainment >

**Comics and Animation**

**Categories** (not divided into groups)

Animation (1821)

Artists@

Characters (163)

Chats and Forums (18)

Comic Books (1781)

Comic Strips (1023)

Conventions (22)

Cultures and Groups (21)

Editorial Cartoons (110)

Graphic Novels (15)

Magazines (6)

Organizations (22)

People (6)

Shopping and Services@

Small Press Comics (15)

Web Directories (8)

Home > Government >

## Law

**Categories** (divided into two groups)

Alternative Dispute Resolution (38)  
 Attorneys General@  
 Booksellers@  
 Cases (287)  
 Continuing Legal Education (24)  
 Countries (43)  
 District Attorneys@  
 Employment Resources (48)  
 Events (3)  
 Firms and Services@  
 History (27)  
 Journals (101)  
 Jury Duty (15)

Law Enforcement@  
 Law Schools (296)  
 Lawyer Jokes@  
 Legal Ethics (4)  
 Legal Research (98)  
 News and Media (49)  
 Organizations (340)  
 Self-Help (15)  
 Software Companies@  
 U.S. Judiciary and Supreme Court@  
 U.S. States (50)  
 Web Directories (40)

Administrative (9)  
 Business (65)  
 Constitutional (175)  
 Consumer (25)  
 Criminal Justice (110)  
 Disability (29)  
 Dog Breed Specific@  
 Elder (10)  
 Employment (24)  
 Entertainment (9)  
 Environmental (72)  
 Estate and Probate (10)  
 Federal (16)  
 Health (12)

Immigration and Naturalization (106)  
 Indigenous Peoples@  
 Intellectual Property (167)  
 International (48)  
 Lesbian, Gay and Bisexual@  
 Privacy (35)  
 Property (70)  
 Sexuality@  
 Tax (54)  
 Technology (104)  
 Trade@  
 Women's Resources@  
 Usenet (9)

|  |   |
|--|---|
| Home > Government ><br><b>Politics</b>   |   |
| <b>Categories</b> (not divided into groups, except for the fairly standard By Region)  |   |
| By Region (7647)<br>U.S. Politics@   |   |
| Activism Resources (47)<br>Chats and Forums (33)<br>Civic Participation@<br>Elections (3151)<br>General Information (3)<br>Humor@<br>News and Media (24)<br>Organizations (78) | Parties (134)<br>Political Consulting@<br>Political Issues (46)<br>Political Opinion (218)<br>Political Science@<br>Political Theory@<br>Regional Conflicts (33)<br>Web Directories (8) |





## Assignment 14. Yahoo classification

### Answer sheet

**What is the degree of precombination in Yahoo?** Very high

- A** Every Yahoo category (precombined descriptor) has a *home place* which is chosen from its many broader terms. But it appears also at other places where it logically belongs; at these other places the category is marked with @. Clicking on a category marked with @ takes you to a different place in the hierarchy. You could say that the @ represents a cross-reference.

The linkages are by no means complete. For example, under

- (1) Home > Education > K-12 > Schools > Elementary Schools

There is no reference to

- (2) Home > Society and Culture > Religion and Spirituality > Faiths and Practices > Christianity > Denominations and Sects > Anglican > Education > K-12 > Elementary Schools

However, sites are double-indexed. For example,

Christ Episcopal School

Is found under (2) and also under

Home > Education > K-12 > Schools > Elementary Schools > By Region > U.S. States > Maryland > Complete List

- B** Under **Canada** there are first cities and provinces (narrower by autonomous subdivision) and then precombined descriptors, such as

... > Canada > Arts and Humanities

which are narrower by combination

Similarly, under **Transportation** there are categories narrower by autonomous subdivision, such as *Auto-Free Transportation*, *Aviation*, and *Highways and Road*, and narrower categories formed by combination, such as ... > *Transportation* > *History*

- C** **Meaningful arrangement under Education** see attached.

- D** **DDC -LCC - Yahoo**

There are quite a few similarities, even at the top level. Differences in emphasis are based on *literary warrant* in the (book) literature (at the time DDC and LCC were devised) and the Web and/or difference the interests of users of the library vs. a Web search engine.

- E** **Compare Yahoo Home, State subdivision and City subdivision**

Some categories are unique to state and city subdivision: *Employment* and *Real Estate*. *Market Place*, *Yellow Pages*, *Maps* are unique to city subdivision. State subdivision replicates most of the categories under Home, but city subdivision combines some categories (less specific) and omits others altogether (less exhaustive).

## F Examine some principles Yahoo uses when designing subdivisions

In **Libraries**, you first have a group of categories that combine a general concept with *Libraries*. The categories in the next group are all formed by adding a subject component. The same principle is applied in **News and Media** and in **Humor**, but not in **Comics and Animation**, which has categories of both kinds mixed together. In **Law** the principle is applied again. In **Politics** there are no categories that are formed by combining with a subject, even though there could be plenty (*economic policy, energy policy, etc.*)

## G Some facets discernible in the Yahoo classification (in no particular order)

This list has just a few facets with a few sample concepts in each. A bit more than you were expected to do, but only the tip of the iceberg of a real analysis

|  |  |  |
|--|--|--|
| <p><b>Form of document</b></p> <ul style="list-style-type: none"> <li>. Web Directories</li> <li>. Chats and Forums</li> <li>. Bibliographies</li> <li>. Software</li> <li>. Videos</li> <li>. Flash cards</li> <li>. Games</li> </ul> <p><b>Subject</b></p> <ul style="list-style-type: none"> <li>. Business</li> <li>. Education</li> <li>. Health</li> <li>. Law</li> </ul> <p><b>Languages</b></p> <ul style="list-style-type: none"> <li>. Chinese</li> <li>. Japanese</li> <li>. Greek</li> </ul> | <p><b>Grade level</b></p> <ul style="list-style-type: none"> <li>. Early Childhood Ed.</li> <li>. K - 12</li> <li>. College and University</li> </ul> <p><b>Places</b></p> <p><b>Religions</b></p> <ul style="list-style-type: none"> <li>. Christian</li> <li>. Hindu</li> <li>. Islamic</li> </ul> | <p><b>Other elemental concepts repeated throughout</b><br/>(would have to develop facets for these)</p> <p>Software<br/>History<br/>Standards<br/>Tests<br/>Shopping ans Services<br/>Accident<br/>Companies<br/>Theory<br/>Method<br/>Employment<br/>Models</p> |
|--|--|--|



## Yahoo classification. Education

**Home** >

### Education

#### Categories

Browse by Region (170)  
 By Culture or Group (398)  
 By Subject (11)

|  |   |
|--|---|
| Academic Competitions (79)<br>Adult and Continuing Education (325)<br>Bibliographies (4)<br>Bilingual (24)<br>Career and Vocational (236)<br>Chats and Forums (40)<br>Companies@<br>Conferences (52)<br>Correctional@<br>Disabilities@<br>Distance Learning (476)<br>Early Childhood Education (90)<br>Employment (143)<br>Equity (27)<br>Financial Aid (395)<br>Government Agencies (77)<br>Graduation (53) | Higher Education (16638)<br>Instructional Technology (334)<br>Journals (36)<br>K-12 (53910)<br>Literacy (12)<br>News and Media (83)<br>Organizations (3008)<br>Policy (52)<br>Programs (322)<br>Reform (70)<br>Special Education (168)<br>Standards and Testing (63)<br>Statistics (6)<br>Teaching (63)<br>Theory and Methods (659)<br>Web Directories (47) |
|--|---|

## Yahoo classification. Education. Meaningful arrangement.

**Home** >

**Education**

### Categories

Browse by Region (170)  
 By Culture or Group (398)  
 By Subject (11)

|  |  |
|--|--|
| <p><b>Information sources</b><br/>           Bibliographies (4)<br/>           Web Directories (47)<br/>           News and Media (83)<br/>           Chats and Forums (40)<br/>           Conferences (52)<br/>           Journals (36)<br/>           Statistics (6)</p> <p><b>Education by level</b><br/>           Early Childhood Education (90)<br/>           K-12 (53910)<br/>           Higher Education (16638)<br/>           Adult and Continuing Education (325)</p> <p><b>Special students and subjects</b><br/>           Special Education (168)<br/>           Disabilities@<br/>           Literacy (12)<br/>           Bilingual (24)<br/>           Career and Vocational (236)<br/>           Correctional@</p> | <p><b>Educational methods</b><br/>           Theory and Methods (659)<br/>           Teaching (63)<br/>           Instructional Technology (334)<br/>           Distance Learning (476)<br/>           Standards and Testing (63)<br/>           Academic Competitions (79)<br/>           Graduation (53)</p> <p><b>Political and economic aspects</b><br/>           Policy (52)<br/>           Reform (70)<br/>           Equity (27)<br/>           Financial Aid (395)<br/>           Employment (143)</p> <p><b>Organizational aspects</b><br/>           Government Agencies (77)<br/>           Organizations (3008)<br/>           Companies@<br/>           Programs (322)</p> |
|--|--|

