Bibliography for KMPro, New England Chapter,
Differentiating Taxonomies & Ontologies

Agrawal, Rakesh
Multilevel taxonomy based on features derived from training documents classification using fisher values as discrimination values**,** inventors: Agrawal, Rakesh; Chakrabarti, Soumen; Dom, Byron Edward; R 31p. 05/15/2001 United States
Link to patent
A system, process, and article of manufacture for organizing a large text database into a hierarchy of topics and for maintaining this organization as documents are added and deleted and as the topic hierarchy changes. Given sample documents belonging to various nodes in the topic hierarchy, the tokens (terms, phrases, dates, or other usable feature in the document) that are most useful at each internal decision node for the purpose of routing new documents to the children of that node area automatically detected. Using feature terms, statistical models are constructed for each topic node. The models are used in an estimation technique to assign topic paths to new unlabeled documents. The hierarchical technique, in which feature terms can be very different at different nodes, leads to an efficient context-sensitive classification technique. The hierarchical technique can handle millions of documents and tens of thousands of topics. A resulting taxonomy and path enhanced retrieval system (TAPER) is used to generate context-dependent document indexing terms. The topic paths are used, in addition to keywords, for better focused searching and browsing of the text database.

Delphi group
The Delphi proving ground for taxonomy and information architecture 4p.+ Delphi Group 03/01/2004
http://www.delphigroup.com/research/surveys/complete/2004/2004-taxonomy-survey-done.htm
Taxonomy survey questions and results of about 280 responses. Includes questions about use ofsearch and taxonomy technology, as well as its perceived value to the researcher.

Delphi Group
Taxonomy & Content; Classification Market Milestone Report 01/01/2002
http://www.stratify.com/infocenter/download/DelphiResearchReport.pdf
Detailed analysis of why taxonomy utilization is beneficial in managing a large corpus of information; description of the various methodologies for automating indexing and taxonomy production (Rules-based, Bayesian, Linguistic and Semantic, Support Vector Machine, Pattern Matching and Other Statistical Algorithms, Neural Networks).

Defines how and why various search engines operate with and without taxonomy support. Analytical presentations on eleven competing products: Autonomy, Convera, Entopia, Mohomine, Quiver, Semio, Stratify, Textology, TopicalNet, Verity,Wherewithal.

Farber, Dan
Top strategic technologies for 2005 ZDNet 04/22/4004
http://techupdate.zdnet.com/techupdate/stories/main/top_10_technology_trends.html
Areas Gartner sees as important in the next year: Instant messaging, Wider use of WLans, Taxonomies, IP Telephony, Software as Services, Real-time infrastructure, Utility computing, Grid, Network Security Convergence, RFID

Frauenfelder, Mark
A Smarter Web 11/01/2001
http://www.ontoprise.de/documents/A_Smarter_Web.pdf
"These days, DARPA is contributing tens of millions of dollars to the Web consortium's Semantic Web project and has developed a semantic language for the U.S. Department of Defense called DARPA Agent Markup Language that allows users to add metadata to Web documents and relate it to ontologies." This effort is in support of Tim Berners-Lee's latest vision of the semantic Web and the effort to metatag all documents to insure ontological integrity. The article implies ongoing human intervention to apply the metadata.

Grimes, Seth
The Word on Text Mining; Text analytics provide concept discovery, automated classification, and innovative displays for volumes of unstructured documents 3p. Intelligent Enterprise 12/10/2003
http://www.intelligententerprise.com/031210/619decision1_1.shtml
Views text mining as a two stage process of categorization and classification. Autonomy focuses on pattern matching and statistical analysis while Convera believes that high-end (developed with human intelligence) taxonomies and ontologies will serve a better result. Definition: "...an ontology provides meaning for a knowledge domain, while a taxonomy organizes that knowledge."

Knox, Rita E.
What Taxonomies Do for the Enterprise, by Rita E. Knox and Debra Logan 2p. Gartner Group 09/10/2003
http://www.cw360ms.com/research/gartner/117204.pdf
"In the past six months, Gartner has fielded a flurry of questions about taxonomies, a key component of managing corporate intellectual assets. In this Spotlight, we cover many of the most important issues." Cites other Gartner articles on taxonomy and ontology development and use.

Lehman, John
Taxonomies = > Categories relating digital content to information needs 18 slides 11/01/2001
http://www.infotoday.com/il2001/presentations/lehman.pps
Makes the case for human built taxonomies as opposed to automatic generation.

McCloskey, Paul
Knowledge Management Building Blocks, Tech Briefing 4p. FCW.COM 04/14/2003
http://www.fcw.com/fcw/articles/2003/0414/spec-building-04-11-03.asp
Cites several search and classification products with comments on the use of system integrators and consultants for creating classification schemes and building taxonomies.

Moulton, Lynda W.
Why do You Need a Taxonomy Anyway? And How to Get Started LWM Technology Services 06/01/2003
http://www.lwmtechnology.com/publish/print_ezine/nlp0603.htm
Terminology, determining need, how to get started, tools for building, implementing and maintaining. Links to useful backup information.

Plumley, Deborah

Process-Based Knowledge Mapping 3p. Destination KM 03/03/2002
http://www.destinationkm.com/articles/default.asp?ArticleID=1041
Procedural knowledge maps show knowledge (and the sources of knowledge) mapped to a business process. This could be any process for a business or organization -- for example, a process for a R&D function/organization, or a selling process, etc. Major use: map is for planning and implementation of knowledge management efforts. Conceptual knowledge maps, which Caldwell calls a "taxonomy," a method of hierarchically organizing and classifying content. In knowledge management, a taxonomy is used for content management within a Web site or some other repository. Explores process mapping: mapping, analysis, application in detail.

Competency knowledge maps document the skills, positions, and even career path of an individual -- to create a competency profile. Among other uses, competency maps can be converted into a 'yellow pages' directory, which enables employees to find needed expertise (skills, techniques, and/or job tasks). Another type of commonly used mapping is called Social Network Mapping or Social Network Analysis. Social network analysis shows networks of knowledge and patterns of interaction among group members, organizations, and other social entities. Used for analysis of information sharing within a social context.

Polikoff, Irene

Ontology Tool Support; Ontology Development Lifecycle and Tools 14p. TopQuadrant 04/10/2003 http://www.topquadrant.com/FPweb/documents/TQ1202_Ontology%20Tool%20Survey.pdf
List of Ontology tool support and status of development organized by type (Creating, Populating, Deploying, Maintaining & evolving), list of tool interoperability.

Quint, Barbara

Taxonomy Developments in the Health Field 3p. Information Today 07/21/2003
http://www.infotoday.com/newsbreaks/nb030721-2.shtml
The National Library of Medicine (NLM) has announced an agreement with the College of American Pathologists (CAP; http://www.cap.org) to add CAP.s SNOMED CT (Systematized Nomenclature of Medicine . Clinical Terms) to NLM.s Unified Medical Language System (UMLS). With 344,000 concepts, SNOMED CT is the world.s most complete clinical taxonomy. "Underlying top quality databases one usually finds top quality taxonomies.

"On the commercial side, Factiva now offers a pharmaceuticals and healthcare taxonomy with over 800 industry-focused terms to help companies build data from both internal and external sources. Focused on business intelligence, the Factiva taxonomy generally comes packaged with support and guidance from Factiva.s consulting service staff." For more information on UMLS, go to: http://www.nlm.nih.gov/research/umls/

Rao, Madanmohan

A Decade of KM; a Report on "Real-World Best Practices" from American Productivity and Quality Center's 8th KM Conference 4p. Destination KM 06/11/2003
http://www.destinationkm.com/articles/default.asp?ArticleID=1065
Cites key lessons in KM: - Connect KM with bottom line - Integrate KM effectively into business strategy - Recognize the importance of taxonomy and governance - Foster organizational citizenship and democracy of knowledge - Integrate KM with Six Sigma and Lean - Extend KM across organizational boundaries - Document Successes

Describes measurements used at Intel & Halliburton; KM as part of business strategy at Siemans and GE; taxonomy development at Johnson Controls.

Russom, Philip
An Eye for the Needle 6p. Intelligent Enterprise 01/14/2002
http://www.intelligententerprise.com/020114/502feat2_1.shtml
"Knowledge workers want content management applications to impose order on document chaos...Accurately representing knowledge workers' domain expertise in a corporate portal's taxonomy is one of the greatest challenges to the development of portal-based content management applications."

"The taxonomy - the structure for categorizing text content by topic - is the piece of the content management application that knowledge workers depend on most and, therefore, the piece they use for measuring its success."

Russom, Philip
Managing Spaghetti Content 4p. Intelligent Enterprise 05/28/2002
http://www.iemagazine.com/020528/509decision1_1.jhtml?_requestid=171277
Every content management application demands a well-ordered taxonomy. The challenge is to maintain taxonomy quality as content evolves over time. Cites the need to hire librarians.

TopQuadrant
Proceedings from the One-Day Conference: Semantic Technologies for E-Government, September 8th, 2003, White House Conference Center, Washington, DC, sponsored by TopQuadrant, jointly with CIO Council's 01/01/2003
http://www.topquadrant.com/conferences/tq_proceedings.htm
Agenda includes:
Keynote: "The Semantic Web" - Eric Miller, Activity Lead for W3C's Semantic Web Initiative.
"Data Independence and the Roadmap to the Semantic Web", Michael Daconta, Chief Scientist, McDonald Bradley
"Positioning Semantic Technologies: The Emerging Vendor Landscape", Irene Polikoff, Executive Partner, TopQuadrant.
Introducing the Semantic Application Gallery and Vendors, Dr. Dean Allemang, TopQuadrant,

Exhibitors: Ecosystems, Coolheads Consulting, Linkspace, McDonald Bradley, Modulant, Ontoprise, Semagix, Software AG, TopQuadrant, Unicorn Solutions.

Ralph Hodgson, CEO TopQuadrant: "Potential of Semantic Technologies for eGovernment"
Interactive Panels discuss "Applying Semantic Technologies in Government" (Panel: George Strawn, NSF, William Sonntag, Chief of Staff at EPA, Lillian W. Gassie, Senior Systems Librarian of Homeland Security Digital Library, Con Kenney, Chief Enterprise Architect at FAA, Luis G. Kun, Ph.D., National Defense University/ DOD, Tom West, IC Metadata Working Group at DIA)

Turocy, Pat
No More Information Overload; Companies must consider how they classify data so employees can find it fast**,** by Pat Turocy, Jeff Phillips, and Bob Anders of Doculabs
2p. Information Week 12/16/2002 http://www.informationweek.com/story/IWK20021212S0007/
Cites the need for taxonomy development tools as part of all content management systems. Lists autocategorization and taxonomy tools available.

Warner, Amy
A taxonomy primer 5p. Lexonomy 01/01/2002
http://www.lexonomy.com/publications/aTaxonomyPrimer.html
Good overview of how taxonomy, thesauri, and classification systems relate, and how they are organized. References to other useful information included.