

links

- Versión española
- English version

- Home
 - Number 4, May 2006
 - Museums
 - Positioning
 - Usability
 - Health
 - Number 3, May 2005
 - Text retrieval
 - Visual elements
 - Ontology
 - Thesaurus
 - DigiDocMap
 - Taxonomies
 - XML
 - Open source
- Alert
- Laboratory
- About us

Taxonomies for categorisation and organisation in sites

Author: Miquel Centelles

Location in hierarchy: [Home](#) -> [Number 3, May 2005](#) -> Taxonomies

URL of this page: <http://www.hiptertext.net/english/pag1011.htm>

Citation: Miquel Centelles. *Taxonomies for categorisation and organisation in Web sites* [o "Hipertext.net", num. 3, 2005. <<http://www.hiptertext.net>> [Consulted: 12/02/2007]. IS:

Sumario

- ▼ 1. The concept of Taxonomy
- ▼ 2. Construction of taxonomy
 - ▼ 2.1. Processes for the construction of taxonomies
 - ▼ 2.2. Automation of the processes of construction of taxonomy
- ▼ 3. Resource categorisation
- ▼ 4. Application of taxonomy in the development of information systems
- ▼ 5. Bibliography
- ▼ 6. Notas

▲ 1. The concept of Taxonomy

At the moment when this article is published, a fact will have happened that should mark a before and after in the evolution of taxonomies and organisation systems: the appearance of the final draft of the revised ANSI/NISO Z39.19-1993 standard, *Guidelines for the construction, management of monolingual thesauri* [1]. This revision has been created between 2002 and 2004 by the Thesaurus Advisory Group (herein created in the National Information Standards Organization, for the introduction of a more user-friendly language in the standard, the update of the scope to the current environment of digital information and the extension of the scope to the wide range of production and content organisations.

We do not have the draft of the revised standard, but we do have access to its contents and the notes from the TAG meetings. From these we can see that one of the global modifications that have been proposed is the change of the standard title - *Guidelines for the construction, management of monolingual thesauri* - for *Construction, management of monolingual controlled vocabularies*. The controlled vocabularies include the four main types: the lists, the synonym rings, taxonomies and the thesauri. The revision of the standard ANSI/NISO proposes the "normalisation" definition of the four types, and establishes the essential elements for the construction and management of them. Specifically, in the "TAG Conference Call, June 30, 2003" (the provisional definitions below were included:

- List: "A set of words or phrases displayed in an organized series"
- Synonym rings: "A set of words or phrases that are considered as synonyms of a central word or phrase"

equivalent for the purposes of retrieval. Synonym rings are during input."

- Taxonomy: "An organized set of words or phrases used for information and primarily intended for browsing."
- Thesaurus: "A controlled vocabulary that indicates preferred variant terms, and term relationship. Usually considered to be a complex of controlled vocabularies." From the modifications proposed by the TAG, the final definition is the following: "A set of words with equivalent terms explicitly identified and with ambiguous phrases (e.g. homographs) made unique. This set of terms includes broader-narrower or other relationships."

In accordance with this definition, taxonomy does not require its components to be connected by a specific type of relationships; it simply requires its components to be organised. The defining characteristics are its prioritising browsing- and, therefore, the application environment-.

Nevertheless, in some documents relative to the process of revising the ANSI/NISO Z39.19 standard, the difference between the four types of controlled vocabularies is determined for the lesser or greater complexity presented. On one hand, the lists and synonym rings concern the equivalence relationship; on the other hand, the thesauri concern equivalence, hierarchy and associational relationships. In a central way, the taxonomies include equivalence and hierarchical relationships.

Waiting for the TAG works to provide a normative definition of the taxonomy, we should highlight that we currently do not have an accepted concept of said term.

Etymologically speaking, taxonomy comes from the Greek term *taxinomia*, meaning ordering, and "nomos", rule. Aristotle was one of the first to use the term around the year 300 before Christ, to name hierarchical schemes or prior classification of scientific objects. The botanist Carl Linnaeus (1707-1778) named with the term taxonomy the classification of the living organisms in hierarchical groups, ordered from the most generic to the most specific (kingdom, type, order, gender, and species). From this classic taxonomy developed as a subfield of biology, dedicated to the classification of organisms in accordance with their differences and similarities. In 1974, with Grove (2003, p. 2774), the principles providing a strict guide for the construction of taxonomies were the logical basis, the empirical observation of the hierarchical structure based on feature inheritance, the evolution, and the pragmatic use. The terminological sources of the general language include the meaning specifically oriented to the experimental environment, as proven by the article including the latest version of the *Diccionario de la lengua española* (2001) -Dictionary of the Spanish Language-:

"1. f. Science dealing with the principles, methods, and procedures for the classification. It is specifically applied, within Biology, to the ordering and systems, with the names of the groups of animals and

2. f. classification (? action and effect of classifying)."

In its basic concept, linked to the experimental sciences, taxonomy uses a mono-hierarchical criterion in the establishment of the classification.

that is: each one of the groups or types making it up can only have and only one, in the hierarchical structure.

At the beginning of the 90s, in the 20th Century, the concept of taxonomy was included in other fields of knowledge, such as Psychology, Social Sciences and Information Technology, to name almost all the access systems that attempt to establish coincidences between the terms used by the user, and that of the system. The first specialists developing web organisation systems were part of the knowledge management (KM) area, coming from fields close to information technology and information science (content management and information architecture); not being a tradition of the documental languages of the Information Sciences, they used the term taxonomy for the systems they developed. The term is currently used to name the content organisation systems in this context, although the theory and practice of the documental languages have been intensively applied in this context.

Before proposing a definition of the term of taxonomy in accordance with current development scopes, we have carried out a work of identification and confrontation of the semantic features with which they are defined. For this purpose, we have carried out an extensive search for definitions, development scopes and/or application of the term of taxonomy. Initially, we have not placed any limitation whatsoever on the origin of the definitions; we have only discarded those made from a classical context of the term. The result has been the localisation of 36 definitions of taxonomy between 2000 and 2005 in various types of sources [2].

The analysis of the definitions shows that these give important variables: the place occupied by taxonomy in the scope of the organisation systems (hereinafter KOS); the information content to which taxonomy is applied; the purposes sought by taxonomy; and the model with which the elements making up taxonomy interrelate.

From the documentation drafted by the NISO TAG, and in the light of the mainly accepted properties in the definitions formulated in the development and/or application scopes, the following definition is proposed:

Taxonomy is the type of controlled vocabulary where all the terms are connected by means of any structural model (hierarchical, tree, flat, etc.) and specially oriented to browsing, organisation systems and contents of the web sites.

It is necessary to specify three points in the contents of this definition:

- The terms (or categories) represent some aspect of the content or structure of the information resources, and not only the content.
- The structural models are not usually presented in a pure form, but a mixture of models is possible (and in the real world, usual) that a same taxonomy can be composed of different structures resulting from the mixture of models.
- The documents reflecting the discussions in TAG show an agreement regarding the applications and preferential uses of taxonomies. Some of the notes of the meetings of said TAG (for example, "TAG Conference Call, May 19, 2003" (2003)), reflect that the concept of taxonomy was initially oriented to browsing and that it was in prejudice to the recovery ("searching"); in the final version of the definition for taxonomy its application also includes this last model.

- The folksonomies or distributed classifications are excluded concept of taxonomy (Mathes, 2004).

Once the definition of taxonomy is established, we shall carry out on the taxonomy construction processes and the application of the categorisation of resources, and the development of information systems of the web sites. Both processes should be preceded by planning determining what characteristics the taxonomy should present and the analysis of the context-that will identify the priorities of the content, the organisation and presentation of the information on the web site, the audience -that will identify the needs and search behaviour and the information by the various user segments- and of the content to identify content patterns-.

▲ 2. Construction of taxonomy

▲ 2.1. Processes for the construction of taxonomy

The construction of corporate taxonomies involves the carrying out of the following processes:

1. Limitation of reality (entity, knowledge area, industrial sector, etc.) to be represented by the taxonomy.
2. Extraction of the group of terms or categories that represent said reality.

In order to carry out this process the establishment is necessary, in each place, of what the priority sources are and the ideal extraction mechanism for each one of them. There are three types: the personal sources, interviews with web users and specialists at the web domain; document sources, in the form of documents representative of the types of contents identified at the planning stage; and the taxonomies or knowledge representation already existing (from nomenclatures of the units and existing resources to the administration classification charts).

It is necessary to identify the extraction mechanisms for each type of source; thus, in the case of the personal sources, the interviews with users and the analysis of the search transaction registers are especially relevant.

The result of this process is a register of representative terms or categories.

3. Terminological control of the terms or categories.

This process involves the carrying out of two tasks. In the first place, terms making up a same concept are identified; in the event that there are two or more, it is necessary to specify which one is considered preferential and which are the less. Secondly, giving a correct and consistent shape to all the taxonomy elements is necessary, regardless of whether they are preferential or not.

The result of this process is the establishment of the equivalence between all the taxonomy terms.

4. Establishment of the scheme and organisation structure of the taxonomy categories.

The organisation scheme includes the criteria used to divide and categories. At the beginning, the criteria are limitless and their depends on the object that should be represented by the taxonomy of the most widely used criteria are the following: the subjects, t and/or disciplines; the people; the addressees; the process, ta functions; the types of documents; etc.

The structural model defines the type of relationship established b category groups derived from the organisation scheme. The gener: has been the application of the hierarchical model (based on th relationship) and the tree model (based on the "part of" relationsf fact, the international and national rules for thesauri designing that applied to the corporate taxonomies exalt these two structural mod model, the faceted, is a good alternative for the hypertext environn the breakdown of various perspectives from which a same concept be seen is key. In fact, this model is being used more and more fre certain types of web sites. Nevertheless, the documentation we h revision of standard ANSI/NISO Z39.19 does not seem to show the this alternative.

Traditionally, two techniques for the development of the structure o have been distinguished: the up to down technique and the d technique.

- The application of the up to down technique involves identification of a limited number of higher categories, and th of the rest of categories in successive levels of subordina reaching the most specific levels of categories. This technic oriented both to the application or a hierarchical structural mo tree model) as well as faceted. The possibility of exercising control on the main categories makes this technique applic construction of taxonomies that have, as exclusive or prioriti the development of browsing systems.
- The application of the down to up technique is based on identification of the most specific categories, which are successive levels of subordination up to reaching higher categories. Generally, this technique is mainly oriented to the of a hierarchical structural model (and/or tree model) althoug previous case, can facilitate the analysis for decision tak structural model most ideal to be applied. In any event, technique that is applied to the development of intervention representatives of real and potential users in the establishn structure of taxonomies (for example, the card sorting method

▲ 2.2. Automation of the processes of const of taxonomies

A critical factor in the construction of taxonomy is the degree of applied to the previously indicated processes. The degree of automa seen as a *continuum* : on the one hand the manual systems (or i are placed, and on the other, the automatic ones. The semi-automata are placed in a central point.

We should highlight that, currently, fully manual systems are rarely creation of taxonomies.

In the minimum level of automation, there are two types of solution: taxonomy templates, specialised in a certain industrial sector, that adapted to the specific conditions of a certain organisation [3] taxonomy edition tools. This second type of solution offers the administrator of the taxonomy a tank for term management, a friendly environment for establishing of relationships between terms, and various modes for presentation and viewing of results. Many of these applications already exist as thesauri administrators, and have not included excessive innovation in their new function in the context of taxonomies. Examples of these solutions can be the Multites 2005 (<http://www.multites.com>) or TermTree (<http://www.termtree.com.au>) products.

At the maximum level of automation, we find programmes that automatically extract a corpus of digital resources of a web site and extract categories in function of resources by means of the application of statistical analysis and/or linguistic processing. Generally, the process of construction of taxonomy and the categorisation of resources is the same; even in some cases, the taxonomy is directly editable as a browsing system. An extreme option of this modality is that giving rise to the so-called dynamic taxonomies: taxonomies resulting from the resources of a search in a search engine that usually resort to statistical analysis of frequencies than to linguistic processing. In these systems, the possibilities of establishing equivalence and relationships between the categories is very limited; the result is usually a taxonomy, closer to a clustering of resources than a classification. An example of these solutions is the Automatic Taxonomy Generator from IDOL Server (<http://www.autonomy.com/content/Products/IDOLServer>).

The completely automatic solutions have not offered, up to the present moment, satisfactory results on taxonomy construction. Consequently, automatic alternatives are being developed that, as Ultraseek Tool (<http://www.verity.com/products/ultraseek/index.html>), assist in the process of creation and maintenance of taxonomy at the same time that it provides a user interface for the revision and approval of categories. Said systems use an algorithm of statistical basis that analyses a resources corpus and establishes terms and relationships between terms to the administrator of the taxonomy, who can then accept them or reject them. All this in a friendly working environment.

▲ 3. Resource categorisation

Categorisation can be defined as the content representation process and/or structure of information resources by means of the assignment of terms from a documental language -categorisation by assignation- or by the extraction of terms of the own resources -categorisation by extraction-.

The most efficient categorisation model currently existing is that based on metadata. According to Méndez and Senso (2004), we can define metadata as:

" all that descriptive information on the context, quality, characteristics of a resource, data or object with the finality of facilitating its recovery, authentication, evaluation, preservation and/or interoperability ".

There are various models of metadata. The elements that differentiate the establishment of differences between these models are, basically, two:

- Which aspects of the resources represent (the elements).

- How those elements are represented (the syntax).

For example, Dublin Core, one of the most widely used models for the description of all the types of information resources, includes, in its simple format (simple level), fifteen elements [4]. The syntax of each element usually includes three components:

- Identification of an element. For example, in Dublin Core, the element is identified by means of the metalabel DC.Subject.
- One or more qualifiers that specify some specific attribute of the element. For example, a qualifier of the metalabel DC.Subject is DC.Subject.SCHEME, which identifies the name of the controlled vocabulary for the categorisation of the element.
- The value or values of the element assigned to the resource described. For example, the terms extracted from the language used for the categorisation of the element.

In a web page coded by means of HTML metalanguage, the syntax of an element would present the following aspect:

```
<META NAME="DC.Subject" SCHEME="TAGS" CONTENT="heritage; Cultural events; Exhibitions; Administration document management; Internet; Files; Information Management" >
```

In a categorisation model based on metadata, the taxonomy constitutes a controlled vocabulary that is very useful for value extraction. The elements will be assigned to the elements describing the information resources previously indicated, the application of taxonomies should not be limited to the elements expressing the contents of the resources, and more exactly to the matter, subject or discipline. The elements relative to the context and structure can also be expressed by means of categories extracted from the taxonomy.

The use of taxonomies in the information resource taxonomies offers general strong points of the controlled languages, as: the treatment of semantic and syntactic aspects of the language; the representation of concepts; the creation of a global vision of the domains objects; the exhaustiveness in the indexing; the solution of the problems involved by the multilingual contexts. From the management point of view, the use of taxonomies in the categorisation of resources offers two additional important benefits:

- On one hand, makes the construction and maintenance of a taxonomy and resource categorisation profitable, as the same can be re-used in the development of various search, browsing and personalisation applications.
- On the other hand, it allows maintaining the conceptual and semantic consistency in the representation of the elements of a taxonomy, which creates in the users an image of consistency in the who uses it and in the entity creating it and maintaining it.

The categorisation model applied by a certain organisation should respond to four essential questions: what information resources will be categorised? With what purpose? Who will categorise them? How will this be done?

The last two questions are closely related to the degree of automation in the assignment of values to the metadata. From this point of view, categorisation systems can be conceived as a *continuum*, on one end manual systems (or intellectual) are placed, and on the other, the automatic ones.

In the first case, an expert analyses the content, context and/or structure of a resource and assigns the appropriate categories to this from a predefined language (categorisation by assignment) or from the text of the resource (categorisation by extraction). The intellectual categorisation offers several strengths, a high level of exactness in the description of resource content, the possibility of including the contextual meaning in the description. Advantages: it facilitates the categorisation of non-textual documents (images, audio, etc.); the weak points are the limited scalability, the high cost of resources and the lack of consistency and exhaustiveness.

The automatic categorisation is based on algorithms that statistically analyse the document word sequence, identify word behaviour patterns and variables such as collocation, order, proximity, frequency, etc., and group documents that show similarities in said behaviour. The results are lists of resources that show similar behaviour patterns, labelled by means of terms or sequences extracted from the resources themselves that best represent their similarity.

A grouping system should be able to carry out the following tasks: analyse the resource word sequences; calculate the value representing the content of a document; and compare the values (sub) documents and determine their degree of similarity.

Currently, the algorithms designed for the analysis of frequencies use the following analysis methods, or a combination of various: statistical methods (Bayesian method Rocchio method, ...); vectorial methods (Neighbor method, Support Vector Machines...); and trees and decision

Examples of automatic categorisation can be the Automatic Categorisation module from the IDOL system (<http://www.autonomy.com/content/Products/IDOL>), based on the probabilistic method, and Lotus Discovery Server (<http://www.lotus.com/lotus/discovery>) based on the vectorial method [5].

The strong points of the automatic categorisation are the efficiency of processing, the high level of scalability and high level of consistency. Its biggest weak point is the low level of exactness that it usually requires making the very frequent use of these systems bases for decision by human categorisation experts.

The semi-automatic or hybrid categorisation systems combine human intelligence, which can identify the various levels of meaning existing in documents, and the efficiency of the automatism. Four types of semiautomatic systems of categorisation can be identified.

- Systems statistically analysing the resources and presenting human experts recommended terms for categorisation for them and approve them. An example of this type of system is Ultraseek Advanced (<http://www.verity.com/products/ultraseek/index.html>).
- Categorisation systems based on search rules. Allows linking to

of the taxonomy categories a search equation designed by sp means of advanced options (search rules). By means of an the system analyses the documents and determines which equation/s with more coincidence. Then, the document is assi category or categories which have said search rules linked. E this type of system can be K2 Enterp (http://www.verity.com/products/k2_enterprise/index.htm Ultraseek Content Classification (<http://www.verity.com/products/ultraseek/cce.html>), | Verity.

- Categorisation systems based on groups of training o documents. Allows linking each one of the categories of tax limited number of documents selected by specialists that are the most relevant. By means of an algorithm, the system a new documents that could be categorised and determines w example documents is most similar to. Then, the document to the category or categories of the most relevant ones. An this type of system can be Mohomine (<http://www.kofax.com/products/mohomine/classifier.asp> Mohomine.
- Categorisation systems based on the linguistic analysis. An this type of system can be Smart Discovery [7] from InXight.

The strong points of the semi-automatic categorisation systems ar balance between efficiency and exactness, the fact that the proces by human reasoning; and the capacity of accumulating and gene learning. Amongst the weak points, we should highlight the req knowledge, skills and efforts of management and maintenance.

In a questionnaire carried out by Delphi Research [8] , the mana large companies all over the world (60% North American) gave th answers to the question on the type of taxonomy implementa hybrid; 26%, automatic; 23%, manual; the rest, or other opt comment.

▲ 4. Application of taxonomy in the develop information search systems

As previously indicated, the differentiation of the taxonomy creation of resource categorisation by means of taxonomy categorie application of taxonomy offers multiple benefits. The objecti construction of this is the representation of a reality (an area of the scope of an organisation activity, etc.) in the most appropriate purpose and interests of the entity that could exploit said repr Additionally, it should be the expression of the image and corpora of the entity itself.

The applications of taxonomy in the web site context can be div focus on the information architecture scope, a same taxonomy car basic or auxiliary tool for the various browsing, organisation a search, labelling and personalisation systems. The re-use of a same for various information architecture tools offers various types of ben

- In the first place, it allows the profitability of the initial el creation of the taxonomy and of the subsequent maintenance

- Secondly, it facilitates the management of the function taxonomy applies: a modification of categories or in the relationships between categories of the taxonomy can uniformly and consistently be transferred to all the functionalities.
- Thirdly, it improves the use of the web site as a group as it reduces the requirement of cognitive, memory and learning load.
- Fourthly, it facilitates the interaction with the website and the creation of a consistent image of the organisation creating and applying a taxonomy.

There are various taxonomy presentation options.

- Integral presentation of the taxonomy, with all its category relationships interconnecting them (equivalence relationship, or faceted structural model, etc.).
- Partial presentation of the original taxonomy, to be able to filter contents from temporary or use criteria.
- Reduction of the taxonomy to the equivalence relationship, in that the taxonomy adopts the synonym ring shape.
- Reduction of the taxonomy to the hierarchical relationship, for category exploration system. In this case, this usually involves a decrease of the amplitude and depth levels to adjust the taxonomy to the recommendations derived from the cognitive, visual and capacity limitations of the standard user.
- Alternative presentations, as can be the alphabetic order of categories, or the tree, graphic and metaphoric presentations.

The selection of an option depends on various factors; the function which it is applied, the users to which it is addressed, etc. Generally, a combination between various presentations of a same functionality results.

One of the functionalities of the web sites where taxonomy plays a role is in the search for information. The systems that allow searching in the web environment can be classified into three main groups: searching and filtering.

The browsing search engines offer the users an organised set of categories where the information resources are included, and a search mechanism through said categories to find the relevant resource according to the information requirement. These browsing systems are especially useful in situations when the users are unable to specify the need for information at a high level (exploration search). The browsing system can be:

- The taxonomy original hierarchical or faceted structure, completely or partially reduced.
- One of the alternative presentations previously indicated: tree, graphic or metaphoric.
- The combination of two or more presentations in a way that the user can choose the most appropriate one.

select the most suitable for the information requirement condi

The information search systems offer the users the possibility of search equation from a word or word combination. These explorati are especially suitable for search situations where the users can information requirement with enough detail (search for a known taxonomy is included to the search system to help the user in the ic of relevant terms for the creation of the search equation, and also the result and presentation and search reformulation proce exploration and search systems imply interaction in real time betwe and the search mechanism.

The third modality, the filtering systems, offers the user the p create and declare an information need (user profile) and receive a reply when a certain period of time elapses, or when the system relevant resources for said need. In this case, taxonomy allows th selection of relevant terms for the specification of the profile.

▲ 5. Bibliography

| |
|--|
| Bennett, Paul. (2002). Introduction to text categorization. Consulted: 1-03-2005, http://www.softlab.ece.ntua.gr/facilities/public/AD/Text%20Categorization/Introduction%20to%20Text%20Categorization.ppt#256 , 1, Introduction to Text Categorization |
| Diccionario de la lengua española (2001). Consulted: 22-03-2005, http://buscon.rae.es/diccionario/drae.htm |
| Fast, Karl; Leise, Fred; Steckel, Mike (2003). "Controlled vocabularies: a glossothesaurus". In: Boxes & arrows, October 27, 2003. http://www.boxesandarrows.com/archives/controlled_vocabularies_a_glossothesaurus.php |
| Gilchrist, Alan; Kibby, Peter; Mahon, Barry. (2000). <i>Taxonomies for business: access connectivity in a wired world</i> . London: TFPL. ISBN: 1-870-889-83-5 |
| Grove, Andrew. "Taxonomy". (2003). In: <i>Encyclopedia of library and information science</i> 2nd ed., rev and enlarg. New York [etc.]: Marcel Dekker, p. 2770-2777 |
| IDOL Server. (2005). Consulted: 13-03-2005, http://www.autonomy.com/content/Products/IDOL |
| Information intelligence: content classification and the enterprise taxonomy practice (2004). Consulted: 25-01-2005, http://www.delphigroup.com/research/whitepapers/20040601-taxonomy-WP.pdf |
| K2 Enterprise. (2005). Consulted: 13-03-2005, http://www.verity.com/products/k2_enterprise/index.html |
| Lotus Discovery Server. (2004). Consulted: 1-sep-2004, http://www.lotus.com |
| Mathes, Adam. (2004). Folksonomies: cooperative classification and communication through shared metadata. Consulted: 26-01-2005, http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html |
| Méndez, Eva; Senso, José A. (2004). Introducción a los metadatos. Consulted: 14-03-2004, http://www.sedic.es/autoformacion/metadatos/introduccion.htm |
| Metainformación: Dublin Core. (2003). Consulted: 13-03-2005, http://www.rediris.es/metadata |

| |
|--|
| Mohomine Classifier. (2005). Consulted: 13-03-2005, http://www.kofax.com/products/mohomine/classifier.asp |
| Multites 2005. (2005). Consulted: 13-03-2005, http://www.multites.com |
| National Information Standards Organization. (2005). ANSI/NISO Z39.19-2003: gui for the construction, format, and management of monolingual thesauri. Consulted: 2005, http://www.niso.org/standards/standard_gather.cfm?pdfink = http://www.niso.org/standards/resources/Z39-19.pdf&std_id=518 . [Consulted: 9-(2005)] |
| Ruiz, Miguel E.; Srinivasan, Padmini. "Combining machine learning and hierarchical indexing structures for text categorization". In: ASIS/SIGCR Workshop on Classification Research (10è: Washington: 1999). Advances in classification research: proceeding. ASIS SIG/CR Classification Research Workshop, v. 10 (1999), p. 107-124 |
| Smart Discovery. (2005). Consulted: 13-03-2005, http://www.inxight.com/products/smartdiscovery |
| "TAG Conference Call, may 19, 2003" (2003). In: National Information Standards Organization. (2004). Developing the next generation of standards for controlled vocabularies and thesauri. Consulted: 23-04-2004. http://www.niso.org/committees/MTinfo.html |
| "TAG Conference Call, June 30, 2003" (2003). In: National Information Standards Organization. (2004). Developing the next generation of standards for controlled vocabularies and thesauri. Consulted: 23-04-2004. http://www.niso.org/committees/MTinfo.html |
| "TAG Notes November 1, 2004" (2004). In: National Information Standards Organization. (2004). Developing the next generation of standards for controlled vocabularies and thesauri. Consulted: 23-04-2004. http://www.niso.org/committees/MTinfo.html |
| Taxonomy strategies. Consulted: 25-01-2005, http://www.taxonomystrategies.com/index.htm |
| Taxonomy warehouse. Consulted: 22-02-2005, http://www.taxonomywarehouse.com |
| Term Tree. (2005). Consulted: 13-mar-2005, http://www.termtree.com.au |
| Ultraseek Advanced Classifier. (2005). Consulted: 22-02-2005, http://www.verity.com/products/ultraseek/index.html |
| Ultraseek Content Classification Engine (CCE). (2005). Consulted: 13-03-2005, http://www.verity.com/products/ultraseek/cce.html |
| Ultraseek Topic Advisor. (2005). Consulted: 22-02-2005, http://www.verity.com/products/ultraseek/index.html |
| Webopedia. Consulted: 28-01-2005, http://www.pcwebopedia.com/TERM/t/taxonor |

▲ 6. Notas

[1] In accordance with "TAG Notes November 1, 2004" (2004), the should be ready for January 2005. [\[volver\]](#)

[2] A copy of the references can be obtained by sending an e-mail this article's author (miguel.centelles@ub.edu). The reason for this req be included. [\[volver\]](#)

[3] An example of this option is Semio Taxonomy from Entr information from: <http://www.entrieva.com/entrieva/productHdr=scts> [Consultado: 13-mar-2005]- [\[volver\]](#)

[4] Information extracted from the Metainformation web site: [(2003), maintained by RedIRIS. [volver](#)]

[5] In accordance with the report Information intelligence classification and the enterprise taxonomy practice (2004 , p. 38), has a market share of 14% and Lotus Discovery Server of 7%. [volv](#)

[6] In accordance with the report Information intelligence classification and the enterprise taxonomy practice (2004 , p. 38) market share of 15%. [volver](#)]

[7] In accordance with the report Information intelligence classification and the enterprise taxonomy practice (2004 , p. 3 Discovery has a market share of 4%. [volver](#)]

[8] Information intelligence: content classification and the taxonomy practice (2004 , p. 26). [volver](#)]

versión para [imprimir](#)
versión mínima para [imprimir](#) o guardar