Social Tagging and Dublin Core: A Preliminary Proposal for an Application Profile for DC Social Tagging.

Maria Elisabete Catarino[1]; Ana Alice Baptista[2]

[1] Information Systems Department, University of Minho.
Campus Azurém, Guimarães, Portugal. CAPES-MEC-Brazil.
e-mail:ecatarino@dsi.uminho.pt
[2] Information Systems Department, University of Minho.
Campus Azurém, Guimarães, Portugal. e-mail: analice@dsi.uminho.pt

Abstract

The Web 2.0 maximizes the Internet concept of encouraging its users to cooperate effectively for the offer of virtual services and content organization. Among the various potentialities of the Web 2.0, folksonomy appears as a result of the free assignment of tags to the Web's resources by their users/readers. Despite tags describe the Web's resources, generally they are not integrated in the metadata. In order for them to be intelligible by machines and therefore used in the Semantic Web context, they have to be automatically allocated to specific metadata elements. There are many metadata formats. The focus of this investigation will be the Dublin Core Metadata Terms (DCTerms) that is a widely used set of properties for the description of electronic resources. A subset of DCTerms, the Dublin Core Metadata Element Set (DCMES), has been adopted by the majority of Institutional Repositories' platforms as a way to promote interoperability. We propose a research that intends to identify elements of the metadata originated from folksonomies and propose an application profile for DC Social Tagging. That will allow tags to be conveniently processed by interoperability protocols, particularly the Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH). This paper will present the results of the pilot study developed in the beginning of the research as well as the metadata elements preliminarily defined.

Keywords: Social Tagging; Folksonomy; Metadata; Dublin Core.

1.   Introduction

Metadata may be defined as a group of elements for the description of resources [1]. There are many standards of metadata in the repository context; we can point out the Dublin Core Metadata Element Set (DCMES) or simply Dublin Core (DC) that is a metadata element set for the description of electronic resources. This standard is well diffused, used globally and on a broad scale due to some factors: a) it was created specifically for the description of electronic resources; b) it has an initiative which is responsible for its development, maintenance and spreading - the Dublin Core Metadata Initiative (DCMI); c) it is the metadata set used by default by the Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH).

The more active participation of the users in the construction and organization of Internet contents is the result of the evolution of the Web technologies. The so-called Web 2.0 is "the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an 'architecture of participation', and going beyond the page metaphor of Web 1.0 to deliver rich user experiences."[2].

One of the new possibilities of the Web 2.0 is the folksonomy that is "the result of personal free tagging of information and objects (anything with an URL) for one's own retrieval. The tagging is done in a social environment (shared and open to others). The act of tagging is done by the person consuming the information"[3].

The tags which make up a folksonomy would be key-words, categories or metadata [4]. Tags have several roles as a study from Golder and Huberman [5][6] points out: Identifying What (or Who) it is About, Identifying What it Is, Identifying Who Owns It, Refining Categories, Identifying Qualities or Characteristics, Self Reference and Task Organizing.

Another study, Kinds of Tags (KoT) [7], has the objective of verifying how the tags derived from folksonomies can be normalized aiming at their interoperability with metadata standards, specifically the DC. Their researchers observed that there are some tags that cannot be inserted in any of the already existing elements. Preliminary results indicate that the following new elements may have to be used: Action_Towards_Resource, To_Be_Used_In, Rate and Depth [8][9].

Generally digital repositories' metadata is input by authors or professionals that mediate deposit. In the Web 2.0 context, folksonomies arise, as a result of Web resource tagging by its own users. Tags are a complementary form of description which expresses the user's view of a given resource and, therefore, potentially important for its discovery and retrieval. The preliminary results of KoT indicate that the current DCTerms elements are not enough to hold user's descriptions by means of tags.

In the context shown, following up the analysis resulting from the KoT project, we propose an application profile for DC Social Tagging so as to enable that tags may be used in the context of the Semantic Web. This application profile will be a result of a research that aims at identifying metadata elements derived from folksonomies and compare them with DCTerms' properties.

2.   Investigation: Procedures

The procedures of this research project are divided in four stages. The first stage consists of an   analysis of all tags contained in the KoT project dataset. At this stage all tags assigned to the resources are analysed, grouped in what we call key-tags and then DC properties are assigned to them when possible. A Key-tag is a normalised tag that represents a group of similar tags. For instance, the key-tag `Controlled Vocabulary` stands for tags `controlledvocabulary`, `controlled vocabularies` or `vocabulars controlatis`.

Once that the meaning of tags is not always clear, it is necessary to dispel doubts by complementarily turning to lexical resources (dictionaries, encyclopaedias, Word Net, Wikipedia, etc), and analysing other tags of the same users. Contacting the users may be a last alternative to try to find out the meaning of a given tag. In this stage, a pilot study was developed in order to refine the proposed methodology and to verify whether the proposed variants for grouping and analysing tags are adequate.

The second stage aims at proposing complementary properties to the ones already existing in the DCMI Metadata Terms [10]. Key-tags that were not assigned to any DC property in stage one will now be subject to further analysis in order to infer new properties specific to Social Tagging applications. This analysis takes into account all DC standards and recommendations, including the DCAM model, the ISO Standard 15836-2003 and the NISO Standard Z39.85-2007.

The next stage comprises the adaptation of an already existing DC ontology. This will make use of Protégé, an ontology editor developed at Stanford University. The ontology will be encoded in OWL, a language endorsed by the W3C.

Finally, the fourth stage intends to submit the proposal to the DC-Social Tagging community for comments

and feedback via online questionnaires. After this phase, a first final version of a proposal to a DC Social Tagging profile will be submitted to the community.

This paper will present the results of the pilot study alongside with the preliminary results of the first research stage: tag analysis. The preliminary results of KoT indicate that an application profile for Social Tagging applications would benefit from the inclusion of new properties, other than those in DCTerms. Those terms will potentially accommodate tags that currently do not have a metadata holder. The results of this research will therefore allow to determine if the KoT preliminary findings are verified and to what extent.

3.   Pilot Study

The pilot study was carried out in order to improve the methodology proposed for the investigation project, since, as Yin [11] states, "The pilot study helps investigators to refine their data collection plans with respect to both the content of the data and the procedures to be followed".

The dataset used in this project is the same of the KoT project: it is composed of 50 records of resources which were tagged in two systems of social bookmarking: Connotea and Delicious.

Each record is composed by fields distributed in two groups of data: a) information related to the resource as a whole: URL, number of users, research date; and b) information related to the tags assigned to the resource: social bookmarking system, user, bookmarked date and the tags.

A relational database was set up with the DCMI Metadata Terms and the KoT data set that was imported from its original files. The following tables were created: Tags, Users, Documents, Key-tags and Metadata.

3.1   Tag Analysis

In the pilot study it was analysed data of the first five resources of the data set. This implied the analysis of a total of 311 tags with 1141 occurrences and assigned by 355 users. It was important to register not only the number of tags but also their total occurrence, since a *tag* could have different meanings to each one of the resources to which it was assigned. Therefore, in some cases, it was possible to analyse of the occurrence of a tag concerning an individual resource.

3.1.1 Grouping Tags in their different forms: Key-tags

Key Tag is the term that represents the various forms of a same Tag. In order to accomplish Tag grouping it was necessary to generate reports for each resource with the following information: Title (of the resource), User Nick and Tag, displaying information in the alphabetical order of the Tags to facilitate the visualization of the existing different Tag forms and definition of Key-tags.

In this stage it is necessary to use lexical resources (dictionaries, WordNet, Infopedia, etc) and other online services, such as online translators, in order to fully understand the meaning of tags. In some cases further research and analysis of other tags of a given user, or even a direct contact with this user by email may be necessary in order to understand the exact meaning of a tag.

An important concern regarding tag analysis is the fact that as tags are assigned by the resources' users, that inevitably leads to a lack of homogeneity in their form. Therefore, it was necessary to establish some rules in order to properly analyse tags, establish key-tags and relate DC properties with them.

The first rule to be observed concerns the alphabet. In this Project, only tags written in Latin alphabet were considered. Further studies should involve the analysis of tags written in different alphabets.

Another rule is directly related to language. The dataset comprises tags written in different languages. As English is the dominant one, it was the chosen language to represent Key-tags.

Depending on the Key-tags, certain criteria concerning the classification of words need to be established: simple or compound, singular or plural, based on a thesaurus structure in its syntactical relations. In these cases, the rules presented by Currás [12] were followed.

It was still necessary to create rules to deal with compound tags, as they contain more than one word. There are two kinds of compound tags: (1) the ones that are related to only one concept and therefore originate only one key-tag (e.g. `Digial Libraries`); and (2) the ones that are related to two or more concepts and therefore originate two or more key-tags (e.g. `Library and Librarians`).

In the first kind, compound tags are composed by a focus (or head) and a modifier [13]. The focus, i.e. the noun component which identifies the general class of concepts to which the term as a whole refers, and the modifier, i.e. one or more components which serve to specify the extension of the focus; in the example above: `Digital` (modifier) `Libraries` (focus). It is a compound term that comprises a main component or focus and a modifier that specifies it.

In the second kind, compound tags are related to two or more distinct Key-tags, as for example: `Library and Librarians`, which would be part of the group of two distinct Key-tags: `Library` and `Librarian`. Another example is `Cataloguing-Classification`, which would be assigned to the Key-tag `Cataloguing` and to the Key-tag `Classification`. In this second segment there isn't a relation of focus/difference between the components as their meanings are totally independent.

Following these pre-established rules, the 311 tags were grouped in their different forms, adding up to 212 Key-tags.

The first step of tag analysis comprises grouping tag variants: a) language; b) simple/compound; c) abbreviations and acronyms; d) singular/plural; e) capital letter/small letter. Then a Key-tag is assigned to each of these groups according to the rules presented above. Following, there are some examples of tags and their assigned key-tag:

- Tags: `_article, article, articles, artikel, article:sw.`
  Key-tag: `Article.`
- Tags: `biblioteca digital, biblioteques digitals, digital libraries, digital library, digital_libraries, digital_library, digitallibraries, digital-libraries, digitallibrary, dl.`
  Key-tag: `Digital Libraries.`

The above key-tags show a variation in :

- spelling: `_article / article; digital library / digital_library / digitallibrary and dl;`
- form (Singular/Plural): `article / articles; digital library / digital libraries;`
- language: `article` (EN) / `artikel` (DE); `Biblioteca digital` (PT) / `biblioteques digitals` (CA) and `Digital Library` (EN).

The examples above also show the two kinds of compound tags. Compound Tags focus/modifier like `biblioteca digital` and `digital library` are assigned to only Key-tag. Tags composed of two focus components like `article:sw` are assigned to two distinct Key-tags: `Article` and `Semantic Web`.

3.1.2 Tag Analysis in relation to DC

After Key Tag composition, an analysis was carried out in order to verify to which DC Properties these tags corresponded. What happens is that this analysis becomes more complex as the definitions of the DCMI Terms are intentionally very inclusive, so that the description of electronic documents with a small, however satisfactory number, of metadata is possible. This inclusiveness may cause some doubt when relating Key-tags to DC Properties. Another factor of complexity is that this is a qualitative study which is developed manually so that the analysis is the most detailed possible.

Due to these factors, it was necessary to define basic rules for the correspondence of Key-tags to the DC Properties.

In the occurrence of Simple tags there is a peculiarity to be noticed that relates to the way tags are inserted in the social bookmarking sites: the way tags are inserted can interfere with the system's indexation. When the user inserts tags in Delicious, the only separator is the space character and everything that is typed separated by spaces will be considered distinct tags. For example, if the compound term `Digital Library` is inserted containing only the space as separator, the system will consider two tags: `Digital` and `Library`. In order to be inserted as a compound tag it is necessary to use special characters such as underscore, dashes and colons. Some examples of such kind of compound tags are: `Digital_Library`, `Digital-library`, `Digital:Library`, `Digital.Library`.

In Connotea tags are also separated by a space or a comma. However, Connotea suggests to users to type compound tags between inverted commas. For example, if the user inserts `Information Science` without placing the words between inverted commas, the words will be considered two distinct tags; however, if they are typed between inverted commas ("`Information Science`") the system will generate only one compound tag. This simple, yet important issue, has a high implication on the system's indexation of the tags.

To exemplify what is said above there is an example of a Delicious user who, when assigning tags to the resource "The Semantic Web", written by Tim Berners-Lee, inserted the following tags: `the`, `semantic`, `web`, `article`, `by`, `tim`, `berners-lee`, without using the resources of word combination (`_;-` etc). The system generated seven simple tags. However, it is clear that these tags can be post-coordinated [14][15] to have a meaning such as Title, Creator and Subject.

Thus, as a first rule, in the cases when simple tags could clearly be post-coordinated, they were analysed as a compound term for the assignment of the DC Property. However, this analysis could only be carried out in relation to only one resource's user and never to a group, since it can mischaracterize the assignment of properties.

The second rule concerns tags that correspond to more than one DC Property. It is considered two different situations: simple and compound tags. The easiest case is the one of simple tags. If simple tags to which more than one property can be assigned occur, then all the properties are assigned to the tag. For example in the resource entitled DSpace, the properties "Title" and "Subject" are assigned to the Key-tag `dspace`.

As explained earlier, compound tags, however, can correspond to two or more key-tags. Thus the relationship with DC properties is made through the key-tags. These are treated as simple tags in the way they are related to DC properties. For example the tag `Web2.0:article`, corresponds to two Key-tags, `Web 2.0` and `Article`, each one of them corresponding to a different property: Subject and Type (respectively). There may also be cases of compound tags that represent two different values for the same property, as in Classification-Cataloguing, that was splitted into two Key-tags: `Classification` and `Cataloguing`, both SUBJECT.

Another rule is related to tags whose value corresponds to the property Title. Tags will be related to the element "Title" when they are composed by terms found in the main title of the resource. For example, Dspace, Library2.0. Another example is the case of the resource entitled "The Semantic Web", where the tags `The, Semantic, Web`, that were assigned by the same user, and thus, may be considered post-coordinated.

3.2   Definition of DC Properties

From the 311 tags analysed, 212 Key-tags were created. From this amount, 159 Key-tags (75%) of which corresponded to the following DC properties: Creator, Date, Description, Format, Is Part Of, Publisher, Subject, Title and Type.

From these, 90,5% correspond to Subject and Description. At this point it is worth to highlight that the tags that referred both to the main subject and to the other subjects related to the resource were allocated to Subject.

The other properties present the following percentages of allocation: Type - 5%; Creator, Is Part Of and Title 3,1% each, Date and Publisher 1,3% each and Format 0,6%.

The other 53 Key-tags (25%) could not be related to any DC property. New complementary properties were defined and their definition is still in process. The following properties that were identified in the pilot study will be described: Action, Category, Depth, Rate, User Name, Utility and Notes.

3.3. Proposed Properties

At this stage, potential new properties for the Key-tags to which it was impossible to assign any DC property were defined. The definition of these properties, at this stage of the research, is still preliminary, since it is based solely in the pilot study. The research on the full dataset will determine which properties will be included in the application profile, including any new that do not exist in DCTerms.

The preliminary new properties identified in the pilot study will be described below, and are the following: Action, Category, Depth, Rate, User Name, Utility and Notes. The following percentages for these properties proposed were observed: Action, Rate and Utility (15,1% each), Category (11,3%), Depth (9,4%), Notes (7,5%) and User Name (1,9%). There is still a 24,5% of Key-tags to which it was not possible to assign or propose any property as their meaning in relation to the resources and users was not possible to identify.

Below, each of these properties will be described, following the set of attributes used to specify the DCMI Metadata Terms [16]: Label, Definition, Comment and Example. Some additional information for better understanding these properties will also be included.

3.3.1. Action

There is a group of Key-tags that represent the action of the user in relation to the tagged resource. It is a type of Tag that can be easily identified since the action is expressed in the very term itself when tagging the resource. Eight Key-tags were identified: `Print, Read-Review, Read Later, Read This, Reading-List, To Do` and `To Read`.

Below, a descriptive table of the element to be proposed

Label	Action
Definition	Action of the user in relation to the resource.
Comment	Has the role of registering the action undertaken by the user to the resource
Example	As example the tags which represent the action `To Read`, attributed to 6 users, all from Delicious: `_toread`, `a_lire`, `toread`.

Table 1: Description of the property Action

3.3.2 Category

This property includes Tags whose function is to group the resources into categories, that is, to classify the resources. The classification is not determined by subjects or theme of the resource, since, in these cases, the key-tags could correspond to the Subject property.

This property is not easy to identify, since it is necessary to analyse the given tag in the context of the totality of tags that user has inserted, independently of the resource under analysis. In some cases it may become necessary to analyse the whole group of resources the user has tagged with the tag that is object of analysis.

Six Key-tags which could correspond to the Key Tag Category were identified: `Alternative Desktop`, `DC tagged`, `DMST`, `FW – Trends`, `Literature` and `Reference`. See descriptive table 2.

Label	Category
Definition	Terms that specify the category of a group of resources.
Comment	Applied to the tags which were attributed to group the resources in categories, but which aren't theme or subject categories, since for those Subject should be used.
Example	For instance, during the analysis of the Key-tag `DC Tagged` it was noticed that the corresponding resources had also other tags tags with the prefix `dc:` (e.g.: `dc:contributor`, `dc:creator`, `dc:Publisher`, `dc:language` or `dc:identifier`, among others). It was concluded that the tag `DC Tagged` could be being applied to group all the resources that were tagged by tags that were prefixed by `dc:`. Therefore it was considered a Category since it is not a classification of subjects or a description of the content of the resource.

Table 2: Description of the property Category

3.3.3 Depth

This type of tag confers the degree of intellectual depth to the tagged resource. As Word Net, Depth "degree of psychological or intellectual profundity" [17].

Label	Category
Definition	Terms that specify the category of a group of resources.
Comment	Applied to the tags which were attributed to group the resources in categories, but which aren't theme or subject categories, since for those Subject should be used.
Example	For instance, during the analysis of the Key-tag `DC Tagged` it was noticed that the corresponding resources had also other tags tags with the prefix `dc:` (e.g.: `dc:contributor`, `dc:creator`, `dc:Publisher`, `dc:language` or `dc:identifier`, among others). It was concluded that the tag `DC Tagged` could be being applied to group all the resources that were tagged by tags that were prefixed by `dc:`. Therefore it was considered a Category since it is not a classification of subjects or a description of the content of the resource.

Table 3: Description of the property Depth

The following Key-tags for this property were identified: `Diagrams, Introduction – Document, Overview, SemanticWeb – Overview, Semantic Web – Introduction,` that occurred only once.

3.3.4 Notes

This element may be proposed to represent the tags that are used as a note or reminder. As Wornet, "a brief written record" that has the objective of registering some observations concerning the resource, but that does not refer to its content and does not intend to be used as its classification or categorization [18].

Label	Notes
Definition	A note or annotation concerning a resource.
Comment	Used to make some type of comment or observation with the objective of reminding something, registering an observation, comment or explanation related to a tagged resource.
Example	For instance, there is a resource that received the tags `Hey` and `OR2007`. The first tag, `Hey`, refers to Tony Hey, a well-known researcher who made a debate on important issues that were related to the tagged resource. In this case the information was given by the user who attributed the tags himself. The second tag makes reference to the Open Repositories 2007, event where Tony Hey mentioned above made a Keynote speech. However, interestingly enough, the tagged resource does not have any direct relation neither with that event nor with Tony Hey, this information was confirmed by the user of the resource himself (creator).

Table 4: Description of the property Notes

A note should be understood as: an annotation to remind something; observation, comment or explanation inserted in a document to clarify a word or a certain part of the text [19].

From the five analysed resources, the following Key-tags considered as Notes were identified: `Hey, Ingenta, OR2007, PCB Journal Club.`

3.3.5 Rate

Rate, meaning pattern, category, class or quality is important to include tags that are evaluating the tagged resource. Thus, the user categorizes the resource according to its quality when using this type of tag.

Label	Rate
Definition	Categorizes the quality of the tagged resource
Comment	Used to register the evaluation of the user in relation to the quality of the tagged resource. Examples of this type of tag: good, great, important.
Example	A resource tagged with the tags `Good` and `Great` represent the qualification of the user according to the quality.

Table 5: Description of the property Rate

The following Key-tags were related to the property: `academic, critical, important, old, great, good` and `vision`. These are generally easily identified as Rate in each one of the terms. In other cases, the tags may be doubtful and it becomes necessary to analyse them in relation to the tags assigned by the user to the resource under analysis as well as to the whole collection of resources tagged by that user. For instance, the tag `Vision` could have several meanings, but, after an analysis to the collection of resources, it may be concluded that it is classifying the quality of the resource

Label	User Name
Definition	Name of the user of the resource.
Comment	Refers to tags which registered the Nick Name of the user of the resource.
Example	In the pilot study only one tag for this type of element was identified. The tag `Alttablib` was attributed by a user of Delicious to the resource 4 (Resource Description and Access (RDA)).

Table 6: Description of the resource User Name

3.3.6. User Name

The Tag User Name labels the resource with the name of a user. The analysed resource had the name of the user of the tagged resource.

Only one tag of this type was identified in the pilot study. Despite the preliminary results presented here, it is assumed that here may be other occurrences.

3.3.7 Utility

After an analysis of the tags and resources, it is proposed an element that would gather the tags that registered the utility of the resource for the user.

It represents a specific categorization of the tags, so that the user may recognize which resources are useful to him in relation to certain tasks and utilities.

In the pilot study the following tags were identified: `Class Paper`, `Research`, `Dissertation`, `Maass`, `Professional`, `Research`, `Search` and `Thesis`. It was not difficult to identify the majority as being Utility. However, three of them, `Class Paper`, `Maass` and `Professional`, required an analysis of other tags and resources from the same users.

`Class Paper` is a tag that is bundled in "1schoolwork" and was assigned to three resources. By analysing the group of resources and related tags, it supposedly refers to resources that would be or have been used for a certain activity.

`Maass` is a tag that was bundled in "Study". The term represents the name of a teacher, information found in the user's notes in two resources tagged with `Maass`: "Forschung von Prof. Maass an der Fakultat Digitale Medien an der HFU"; and "Unterlagen für Thema 'Folksonomies' für die Veranstaltung "Semantic Web" bei Prof. Maass".

`Professional` is a tag assigned by the user to separate those resources that are useful for work-related issues. This information was given by the user of the tag himself.

Label	Utility
Definition	Represents the purpose of use of the resource for the user.
Comment	Categorizes the resources according to utility, as for example: dissertation, thesis
Example	A group of resources useful for the development of a research could be tagged with the tag Research.

Table 7: Description of the property Utility

4.   Final Considerations

In the following cases it was not possible to make any correspondence with any property since it was impossible to understand the meaning of the tags in relation to their resources: resource 1 - Capstone;

resource 2: Suncat2; resource 4: Babel*, Exp, L and resource 5: Do it or Diet*, Inner Space*, Kynunan*, and W.

Nonetheless, these are the results of the pilot study and, therefore, they will, be presented to the DC community for evaluation and validation along with the result of the final research.

As result of this pilot study it is important to highlight that there is a meaningful part of tags, 25%, which could not be assigned the already existing DCTerms properties. This result strengthens what had already been concluded in the KoT project, where 37,3% of the analysed tags were not found to correspond to any of the DCTerms properties. Therefore, the adoption of new properties is justified so that the metadata deriving from folksonomies can be used by metadata interoperability protocols

5.   Acknowledgments

6.   Notes and References

[1]   DCMI. *Using Dublin Core: Dublin Core Qualifiers*. DCMI, 2005. Available at: http://dublincore.org/documents/usageguide/qualifiers.shtml , last accessed on August 30, 2007.

[2]   O'REILLY, T. *Web 2.0: Compact definition?* O'Reilly Radar Blog, 1 October 2005. Available at: http://radar.oreilly.com/archives/2005/10/web_20_compact_definition.html, last accessed on November 6, 2006.

[3]   WAL, Thomas Vander. *Folksonomy definition and wikipedia*. Available at: http://www.vanderwal.net/random/entrysel.php?blog=1750 , last accessed on November 22, 2006.

[4]   GUY, Marieke; TONKIN, Emma. Folksonomies: tidying up tags?. *D-Lib Magazine*, v.12, n.1, jan. 2006. Available at: http://wwww.dlib.org/dlib/january06/guy/ 01guy.html , last accessed on December 12, 2006.

[5]   GOLDER, Scott A.; HUBERMAN, Bernardo A. *The Structure of Collaborative Tagging systems*. Available at: http://arxiv.org/abs/cs.DL/0508082 , last accessed on November 14, 2006a.

[6]   _____. Usage patterns of collaborative tagging systems. *Journal of Information Science*, v.32, n.2, p.198-208, 2006b.

[7]   Preliminary data presented in DC-2007 and NKOS-2007.

[8]   BAPTISTA, Ana Alice et al. Kinds of Tags: progress report for the DC-Social tagging community. In: *DC-2007, International Cconference on Dublin Core and Metadata Applications: applications profiles: theory and practice*. 27-31 August, Singapure. Available at: http://hdl.handle.net/1822/6881, last accessed on September 4, 2007.

[9]   TONKIN, E. et al. Kinds of tags: a collaborative research study on tag usage and structure (Presentation). In: *European Networked Knowledge Organization Systems (NKOS)*, 6.; *EDCL Conference*, 11., Budapest, Hungary. Available at: http://www.us.bris.ac.uk/Publications/ Papers/2000724.pdf , last accessed on December 10, 2007.

[10]   DCMI Usage Board. *DCMI Metadata Terms*. 14 January 2008. Available at: http://dublincore.org/documents/2008/01/14/dcmi-terms, last accessed on January 21, 2008.

[11]   YIN, Robert K. *Case Study Research*: design and methods. Thousands Oaks, USA, 1989.

[12]   CURRÁS, Emília. *Ontologías, taxonomia y tesauros*: manual de construcción y uso. 3.ed. act. y ampl. Madri : Treas, 2005.

[13]   INTERNATIONAL STANDARDS ORGANIZATION. *ISO 2788: Documentation*: *Guidelines for the establishment and development of monolingual thesauri.* [S.L.] : ISO, 1986.

[14]   Post-coordination is the principle by which the relationship between concepts is established at the moment of outlining a search strategy [15].

[15]   MENEZES, E. M.; CUNHA, M. V.; HEEMANN, V. M. *Glossário de análise documentaria*. São Paulo : ABECIN, 2004. (Teoria e Crítica, 01).

[16]   DCMI Usage Board. *DCMI Metadata Terms*. 14 January 2008. Available at: http://dublincore.org/documents/2008/01/14/dcmi-terms, last accessed on January 21, 2008.

[17]   WORDNET. *A lexical database for the english language*. Princeton University, Cognitive Science Laboratory. Disponível em: <http://wordnet.princeton.edu/>. Acedido em 07 de Fevereiro de 2008.

[18]   WORDNET. ref 17.

[19]   INFOPEDIA. *Enciclopédias e dicionários*. Porto : Porto editora. Disponível em: <http://www.infopedia.pt>. Acedido em 7 de fevereiro de 2008.