Collaborative Tagging as a User-Generated Tool for the Organization of Online Healthcare Resources

Valérie DURIEUX
Université Libre de Bruxelles
Avenue Franklin D. Roosevelt 50
1050 Brussels, Belgium
vdurieux@ulb.ac.be

Introduction

With the exponential increase of resources available on the Internet, it has become crucial to sort this mass of documents in order to retrieve the information relevant to a particular need. For their information retrieval needs, most Web users rely on search engines, such as Google, performing a full text indexing (Doyle & Hammond, 2006). Opposite to this automated indexation there is the one accomplished by professionals using controlled vocabularies allowing to reach a high precision level in terms of search results. Nevertheless the amount of online resources exceeds the professionals' capacities to index.

For some time now, a new practice called collaborative tagging has been considered as an alternative to professional indexing for online resources. Collaborative tagging describes the process of ordinary users adding metadata in the form of tags to online content (i.e. websites bookmarks, photographs, computer games or music) in order to store and organize it (Golder & Huberman, 2006). The assigned tags become immediately available for others to see and use as a means of information retrieval. The emergent list of freely assigned tags is commonly referred to as a 'folksonomy' (short for "folk taxonomy"), meaning a user-generated taxonomy (Vander Wal, 2005).

Research question

The doctoral research explores the potential of the collaborative tagging process to equal or even surpass professional indexing. More specifically, the study analyses the possibility to use collaborative tagging to improve information retrieval on the Web by organising its content. Focusing on healthcare resources, the main research question is therefore:

To what extent could collaborative tagging process serve healthcare professionals in their information seeking?

Methods

Part 1: The tags

In order to answer this question, several studies comparing tags assigned by Web users with descriptors provided by librarians for the same online resources have been conducted.

A first study examined online resources and their associated tags assigned by Delicious users compared with MeSH descriptors provided by librarians from CISMeF[1] (Durieux & Kerdelhué, 2009). For the comparison process, the seven-point scale proposed by Kipp has been adopted (Kipp, 2005). Each tag of the data set was manually compared to descriptors assigned to the concerned article and placed in one of the following categories: (1) Same – the tag and one of the descriptors are the same or almost the same (e.g. plurals, spelling variations and multi word terms); (2) Synonym – the tag and one of the descriptors are synonyms according to the MeSH thesaurus; (3) Broader Term – the tag is a broader term of one of the descriptors according to the MeSH thesaurus; (4) Narrower Term – the tag is a narrower term of one of the descriptors according to the MeSH thesaurus; (5) Related Term – the tag is a related term of one of the descriptors according to the MeSH thesaurus (e.g. the "See also" relation); (6) Related – there is a relationship (e.g. conceptual, etc.) between the tag and a descriptor but it is not formally expressed in the MeSH thesaurus; (7) Not Related – the tag has no apparent relationship to one of the descriptors. This study showed that Delicious users tend to assign tags that are identical to descriptors or more general. It was also found that numerous tags provide additional access points to the tagged resources compared with descriptors. The majority of tags are then relevant and useful for information retrieval. Nevertheless a quarter of the assigned descriptors were not represented at all by any of the users' tags. This study thus demonstrated that users' tags complement librarians' descriptors and even compete with them but can in no way be substitute for them.

To strengthen the findings of this previous study, metadata assigned to another type of resources by three different groups was examined. MeSH descriptors from PubMed librarians were compared to tags assigned by users of Delicious and CiteULike to a sample of scholarly articles in the Healthcare field. The study showed that CiteULike users tend to assign tags that are very similar to the descriptors from PubMed whereas Delicious users tend to provide additional access points to the articles. The term comparison demonstrated the majority of tags in both CiteULike and Delicious are relevant and useful for the information retrieval process. Moreover, numerous tags provide the tagged articles with additional access points compared with descriptors. Nevertheless, there were a high proportion of useless tags and a low added value provided by the tags in comparison with the descriptors. The study concluded that more articles have to be tagged and more tags have to be assigned per article to actually compete with professionals' descriptors. These objectives can be met by encouraging users to tag resources, especially in the field of Life Sciences where collaborative systems are not widely adopted.

Part 2: The end-users

To evaluate the potential of collaborative tagging, MeSH descriptors assigned by professionals as been considered as the reference. However, as demonstrated by Trant (2006), the professionals' descriptors do not completely fulfill end-users needs, which is an important criterion for metadata quality. Future research will thus focus on end-users. It will aim to measure the potential of collaborative tagging to meet the users' needs in their information retrieval process.

The methodological choices for this second part of my research are still in progress. First, the concept of "end-user" has to be precisely defined. People searching for Healthcare resources are of two types: professionals (e.g. practitioners, nurses) and non professionals (i.e. patients).

---

[1] a French expert gateway of online resources in the healthcare field.

As an assistant in a Medical University Library (attached to a teaching hospital), it seems logical to focus on professionals. Second, the most effective methods to collect the data about their information needs and seeking behaviour have to be determined. Online surveys seem to be an appropriate way to gather general information on "professional end-users". In depth interviews would follow. Third, depending on the results of the surveys and interviews, a collaborative tagging system could be created from scratch or an existing one adopted by the "professionals" working at the University Hospital.

References

Doyle T. and Hammond JL. Net cred: evaluating the internet as a research source. *Reference Services Review*, 34(1):56-70, 2006.

Durieux V. and Kerdelhué G. Looking for Health Information on the Internet: Can Social Bookmarking Systems Replace Expert Gateways?. *Workshop EAHIL* (Dublin, Ireland, June 2-5, 2009). Retrieved August 15, 2010, from:
http://www.eahil2009.ie/images/Eahil/Images/papers/fri8/durieux_paper.pdf

Golder S.A. and Huberman J.L. Usage Patterns of Collaborative tagging systems. *Journal of Information Science*, 32(2):198-208, Apr 2006.

Kipp M.E. Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator and Intermediary Keywords. *Canadian Journal of Information and Library Science*, 29(4):419-436, 2005.

Trant J. Exploring the potential for social tagging and folksonomy in art museum: proof of concept. *The New Review of Hypermedia and Multimedia*. 12(1):83-105, 2006.

Vander Wal T. Explaining and Showing Broad and Narrow Folksonomies. *vanderwal.net*, 2005. Retrieved August 10, 2010, from:
http://www.vanderwal.net/random/entrysel.php?blog=1635.