Margaret E. I. Kipp
College of Information and Computer Science
Long Island University
margaret.kipp@gmail.com

Searching with Tags: Do Tags Help Users Find Things?

Abstract
This study examines the question of whether tags can be useful in the process of information retrieval. Participants were asked to search a social bookmarking tool specialising in academic articles (CiteULike) and an online journal database (Pubmed) in order to determine if users found tags were useful in their search process. The actions of each participants were captured using screen capture software and they were asked to describe their search process. The preliminary study showed that users did indeed make use of tags in their search process, as a guide to searching and as hyperlinks to potentially useful articles. However, users also made use of controlled vocabularies in the journal database.

Background
Classification is practiced by all humans with varying purposes and agendas. (Bowker and Star 1999) In traditional library classifications, the classifier was the cataloguer or indexer, an individual trained in the rules of information organisation to assign important information about the physical media and the subject matter of the content. On the web, the classifier has typically been the creator of the item, or an automated system collecting basic word frequency information to determine approximate topics. There has been a growing move to classify materials manually using consensus classifications created on the web by large groups of users tagging material on social bookmarking sites. Users are encouraged to add descriptive terms or tags to each bookmark. Tagging is the process of assigning a label to an item.

While other groups have been involved in creating index terms (for example, keywords with submitted articles), these keywords generally have a small circulation and are not widely used. Small scale indexing is common but generally covers a narrow range of topics and is specific to the article. Collaborative tagging systems such as CiteULike (http://www.citeulike.org) or Connotea (http://www.connotea.org) allow users to participate in the classification of journal articles by encouraging them to assign useful labels to the articles they bookmark.

Related Studies
Previous research in classification suggests that there is a distinct difference between user created or naive classification systems and those created by professional indexers. Beghtol (2003) While both systems employ subject based terms, users tend to employ terms that remind them of current or past projects and tasks, terms which could have little meaning to those outside their circle of friends, but are very meaningful to the user. (Malone 1984; Kwasnik 1999; Jones, Phuwanartnurak, Gill and Bruce 2005)

Mathes proposes that librarians embrace user assigned tags as a third alternative to

traditional library classifications and author assigned keywords. (Mathes 2004) He and others also suggest that user classification systems would allow librarians to see what vocabulary users actually use to describe concepts and that this could then be incorporated into the system as entry vocabulary to the standard thesaurus subject headings. (Mathes 2004; Hammond, Hannay, Lund and Scott 2005)

Studies comparing the terminology used in tagging journal articles to indexer assigned controlled vocabulary terms suggests that many tags are subject related and could work well as index terms or entry vocabulary (Kipp 2006; Kipp and Campbell 2006; Hammond, Hannay, Lund and Scott 2005); however, the world of folksomonies includes relationships that would never appear in a traditional classification including time and task related tags, affective tags and the user name of the tagger. (Kipp 2007) These short term and highly specific tags suggest important differences between user classification systems and author or intermediary classification systems which must be considered.

Users searching online catalogues and databases often express admiration for the idea of controlled vocabularies and knowledge organisation system, but find the process of searching frustrating. (Fast and Campbell 2004) Controlled vocabulary indexing has proven costly and has not proven to be truly scalable on the web. Can the user created categories and classification schemes of tagging be used to enhance retrieval in these new environments? Much speculation has been advanced on the subject but so far few studies have been done. A few projects are currently examining the combined benefits of professional and naive classifications. (Trant 2006; Allen and Winkler 2007; Quintarelli, Resmini and Rosati 2006)

The following study explores the usefulness of tagging for enabling retrieval by performing an information retrieval study on a social bookmarking system and a more traditionally classified database to study the usefulness of tags in the support of information retrieval. All information retrieval studies using controlled vocabulary searches contain an implicit evaluation of the effectiveness of classification terms. In such an evaluation it is important to evaluate not only the retrieval effectiveness of the search term, but also how long it took the user to think of using this term in this context and whether or not the user thought the term was useful and accurate.

This study aims to explore those questions in a new context. Proponents of tagging and social bookmarking often suggest that tags could provide at worst an adjunct to traditional classification systems and at best a complete replacement for such systems. (Shirky 2005) A method for testing the usefulness of a classification system for enabling retrieval is to perform an information retrieval study on a social bookmarking system to study the usefulness of tags in the support of information retrieval.

One way to examine the potential uses of tags in the search process is to compare the search experience between social bookmarking tools and other methods of information

retrieval such as retrieval via controlled vocabulary or retrieval via free text search.

Research Questions
- Do tags appear to enhance retrieval? Do users feel that they have found what they are looking for?
- How do users find searching social bookmarking sites compared to searching more classically organised sites? Do users think that tags assigned by other users are more intuitive?
- Do tagging structures facilitate information retrieval? How does this compare to traditional structures of supporting information retrieval?

Methodology
A preliminary study was conducted using volunteer searchers. Participants are currently being recruited to continue the study. The searchers were asked to search Pubmed (an electronic journal database) and CiteULike (a social bookmarking site specialised for academics) for information on a specific assigned topic. The topic was provided as found below.

> "You are a reference librarian in a science library. A patron approaches the reference desk and asks for information about the application of knowledge management or information organisation techniques in the realm of health information. The patron is looking for 5 articles discussing health information management and is especially interested in case studies, but will accept more theoretical articles as well."

Screen capture software, a think aloud protocol and an exit interview were used to capture the impressions of the users when faced with traditional classification or user tags and their usefulness in the search process. While information concerning the usability of the systems themselves for searching may be of interest, data collection will be focused on a comparison of the terms entered by the participants.

Table 1: Timeline of Study Interview

Activity	Description	Length
Welcome	Initial greeting and welcome	2-3 minutes
Introduction to session	Discuss the session itself and the tasks they will be asked to perform.	5-7 minutes
First search task (CiteULike or Pubmed)	The first of two tasks consisting of: 1) user's generation of keywords for search, 2) collection of articles, 3) analysis of retrieved articles for relevance, and 4) assignment of relevance judgements to the articles, 5) assignment of new set of keywords for search	15 minutes
Second search task	same as first task	15 minutes

(Pubmed or CiteULike)		
Post search discussion	A semi-structured interview involving a discussion of the participant's results and their own thoughts as to the usefulness of the terms they used to search and the terms used to describe the documents they retrieved.	15 minutes
Conclusion	Final comments/ thank you for participating.	3-5 minutes

Each participant searched for information using both the traditional on-line database with assigned descriptors and a social bookmarking site. Participants were asked to perform the searches in the order specified so that their use of a social bookmarking site first versus an on-line database could be randomised.

Participants selected their own keywords for searches on both tools. Participants were asked to search until they had located 5 articles that appeared to match the query and assign relevance score to article based on an examination of available metadata. At the end of each search, participants were asked to make a list of what terms they would now use if asked to search for this information again. Participants did not have access to their initial set of search terms at this time to eliminate the learning effect.

Three sets of data were thus available for analysis: sets of initial and final keywords selected by the user, the recording of the search session and think aloud, and recorded exit interviews after the search session. Each set of data can be analysed to examine user impressions of the search process from the perspective of the keywords (tags or index terms respectively). Additionally, keywords and tags chosen by users will be compared and examined to see how or whether they are related.

Results
Preliminary results from the study show that users tended to prefer the search experience on the system used first, regardless of previous experience with either system. Further interviews are required to determine if this trend continues.

All users used multi word keywords initially, which is unsurprising as they are in training to be librarians. At the end of the search process, when users were asked to generate a new list of keywords they would now use for the search, a majority of the users separated their list of final keywords by tool, despite the fact that they were asked for only one list.

Table 2: Initial Keywords

Keyword	Frequency
knowledge management	3
information organisation	2
case studies/"case stud"	2

health information	2
health information management	1
consumer health	1
consumer health management	1

The four most commonly chosen terms were: knowledge management, information organisation, case studies/"case stud" and health information. Each of these terms is directly from the initial text of the information need.

Table 3: Final Keywords

CiteULike		Pubmed	
Keyword	**Frequency**	**Keyword**	**Frequency**
knowledge management/km	4	knowledge management	3
information management	3	health information	2
health-information/health information	2	"information management"[MeSH]	2
case studies/"case stud"	2	health care	1
health-informatics	1	information seeking	1

The most commonly used keyword, by far, was knowledge management. This term comes directly from the information need (described above) and is in keeping with previous information retrieval studies where users tended to select terms from the text for search. (Oppenheim, Morris and McKnight 2000) Information management (also commonly used), could be seen as a modification of knowledge management to fit the terminology of a different group of users who prefer the term information management. Another commonly chosen term was health information (from the information need). Both information management and health information were 2nd and 3rd most popular for CiteULike and Pubmed (or vice versa).While users considered their initial keyword sets to have been incomplete, they tended to choose the same or very similar terms as their suggestions for good search terms to use in order to produce better results. This suggests that their initial search terms were well chosen and matched closely those chosen by users tagging articles in CiteULike, but also came close enough to terms used in the Medical Subject headings used in Pubmed (or its entry vocabulary) or terms used by authors whose works are published in Pubmed for good results to be retrieved. Users selected from 3-5 keywords for both lists, but one user found the Pubmed descriptor "information management" to be the best keyword possible and suggested that users would be better to use this MeSH heading and browse the results.

Conclusions
The preliminary study showed that users did use the tags to aid in the search process,

selecting tags to see what articles would be returned. They also used the tags as a guide to suggest further search terms, suggesting that users do indeed pay attention to subject headings and metadata if it fits a pattern they recognise or makes sense in the context of their existing knowledge on the subject.

Users generally used the same number of keywords for both lists, though most insisted on dividing the final keyword list up by tool. Despite this, the 3 most commonly used terms were the same in each case and knowledge management was generally selected as a useful term for each tool. Interestingly, users tended to say they preferred searching the tool they used for the first search, regardless of prior experience. Further study is required to see if this phenomenon holds true.

It is expected that the continuing study will provide additional insight into user's choices of preliminary keywords for searching as well as participant insights into the process of searching via the tags or controlled vocabulary.

References
Allen, Laurie; Winkler, Michael. (2007). PennTags: Creating and Using an Academic Social Bookmarking Tool. In *Proceedings of the ACRL 13th National Conference*, Baltimore, MD, March 29-April 1, 2007.

Beghtol, Clare. (2003). Classification for Information Retrieval and Classification for Knowledge Discovery: Relationships between "Professional" and "Naive" Classifications. *Knowledge Organization* 30(2): 64-73.

Bowker, Geoffrey C.; Star, Susan Leigh. (1999). *Sorting Things Out: Classification and its Consequences*. Cambridge, MA: MIT Press.

Fast, Karl V.; Campbell, D. Grant. 2004. 'I still prefer Google': University student perceptions of searching OPACs and the Web. In *Proceedings of the 67th Annual Meeting of the American Society for Information Science and Technology*, Providence, Rhode Island, November 13-18, (2004) (Vol. 41, pp. 138-146).

Hammond, Tony; Hannay, Timo; Lund, Ben; Scott, Joanna. (2005). Social Bookmarking Tools (I): A General Review. *D-Lib Magazine* 11(4). http://www.dlib.org/dlib/april05/hammond/04hammond.html

Jones, William; Phuwanartnurak, Ammy Jiranida; Gill, Rajdeep; Bruce, Harry. (2005). Don't Take My Folders Away! Organizing Personal Information to Get Things Done. *CHI 2005*, April 2-7 2005, Portland, Oregon, USA.

Kipp, Margaret E.I. (2006). Complementary or Discrete Contexts in on-line Indexing: A Comparison of User, Creator and Intermediary Keywords. Canadian Journal of Information and Library Science. Preprint at: http://dlist.sir.arizona.edu/1533/

(forthcoming)

Kipp, Margaret E.I.; Campbell, D. Grant. (2006). Patterns and Inconsistencies in Collaborative Tagging Practices: An Examination of Tagging Practices. In *Proceedings Annual General Meeting of the American Society for Information Science and Technology*, Austin, TX, November 3-8, 2006. http://dlist.sir.arizona.edu/1704/

Kipp, Margaret E. I. (2007). @toread and Cool: Tagging for Time, Task and Emotion. *Proceedings of the 8th Information Architecture Summit*, Las Vegas, March 22-26. http://dlist.sir.arizona.edu/1947/

Kwasnik, Barbara. (1991). The Importance of Factors That Are Not Document Attributes in the Organisation of Personal Documents. *Journal of Documentation* 47(4): 389-98.

Malone, Thomas W. (1983). How Do People Organize Their Desks? Implications for the Design of Office Information Systems. *ACM Transactions on Office Information Systems* 1(1): 99-112.

Mathes, Adam. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata. http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

Oppenheim, Charles; Morris, Anne; McKnight, Cliff. (2000). The Evaluation of WWW Search Engines. *Journal of Documentation* 56(2): 190-211.

Quintarelli, E.; Resmini, A.; Rosati, L. (2006). FaceTag: Integrating Bottom-up and Top-down Classification in a Social Tagging System. *Proceedings of the 2nd European Information Architecture Conference*, Berlin, Germany, September 30-October 1.

Shirky, Clay. (2005). Ontology is Overrated: Categories, Links, and Tags. http://shirky.com/writings/ontology_overrated.html

Trant, Jennifer. (2006). Exploring the potential for social tagging and folksonomy in art museums: proof of concept. *New Review of Hypermedia and Multimedia* 12(1): 83 - 105. http://www.archimuse.com/papers/steve-nrhm-0605preprint.pdf