Classification is practised 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.

While other groups have been involved in creating index terms (for example, journal article authors who are asked to provide keywords with their 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.

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 et al 2005; Morville 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 et al 2005); however, the world of folksonomies includes relationships that would never appear in a library classification or thesaurus including time and task related tags, affective tags and the user name of the tagger. (Kipp 2007; Kipp and Campbell 2006; Kipp 2006) These short term and highly specific tags suggest important differences between user classification systems and author or intermediary classification systems which must be considered.

Although users searching online catalogues and databases often express admiration for the idea of controlled vocabularies and knowledge organisation system, they may find it difficult to accommodate their vocabulary to the thesaurus and often find the process of searching frustrating. (Fast and Campbell 2004) Additionally, controlled vocabulary indexing has proven costly and has not proven to be truly scalable when dealing with digital information, especially information on the web. Can the user created categories and classification schemes of tagging be used to enhance findability in these new environments? Much speculation has been advanced on the subject but so far no empirical studies have been done. (Shirky 2005)

The following study explores the usefulness of a classification system for enabling retrieval by performing an information retrieval study on a social bookmarking system and a more traditional online database to study the usefulness of tags in the support of information retrieval. A key component in all modern information retrieval studies is the relevance judgements of users of the retrieval system. (Tang and Sun 2003; Oppenheim et al. 2000) Relevance judgements will be used to provide additional context for participant comments and choices of articles and search terms.
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.

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.

- Do tags appear to enhance findability? 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?

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 a paragraph describing an information need. 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.

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.
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. 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.

Users used between 3-5 keywords initially and suggested 4-5 keywords for CiteULike use and 1-4 for Pubmed. Users did use the tags to aid in the search process, selecting tags to see what articles would be returned.

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


http://www.dlib.org/dlib/april05/hammond/04hammond.html


http://dlist.sir.arizona.edu/1947/


Oppenheim, Charles; Morris, Anne; McKnight, Cliff. 2000. The Evaluation of WWW Search