

# The Role of Tags in Information Retrieval Interaction

**Yong-Mi Kim**

School of Information, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112

Email: [kimym@umich.edu](mailto:kimym@umich.edu)

## Introduction

According to a recent Pew Internet report (Rainie, 2007), 28% of Americans have tagged content online. Tags are descriptive terms people attach to online content, either their own or other people's; tagging is the practice of attaching tags. Tagging has been rapidly adopted on the Web, particularly by sites based on user-contributed content, such as blogs and photo sharing sites. According to these sites, tags make it easier to find tagged items later, make tagged items more findable by others, and also help organize collections of items. Surprisingly, there is little empirical support for these claimed benefits of tags. Research on tagging has focused on analysis of the tags themselves (Golder & Huberman, 2006; Marlow et al, 2006; Kipp, 2007) or the motivations and behavior of taggers (Ames & Naaman, 2007; Wash & Rader, 2007). Users reported failing to find the information they wanted using tags (Wash & Rader, 2007), and recall and precision of searches using tags lagged behind search engines (Morrison, 2007).

What is still needed is a more nuanced understanding of the role tags play in the interaction or communication of the user with the IR system, in this case the site implementing tagging. Despite the proliferation of tags and tagging on the Web, we do not yet have a clear understanding of how to integrate tags into current models of information seeking and retrieval. Recall and precision measures tell us little about a user's IR interactions at the cognitive level, such as in making relevance judgments. Examining IR interaction can elucidate the relationship between how the IR interface handles tags and the quality of the search experience for the user. This study seeks to address these questions by examining the role of tags across different phases and activities of IR interaction.

## Tags in Web Information Retrieval Interaction

The Web IR interaction process is framed as starting once the user is on the Web site in which he or she intends to enter their query. This process can take place on any site that provides search functionality in the form of a search box. This process includes some or all of the following user activities: query formulation, examination of the search results, examination of specific documents selected from the search results, and query reformulation. Documents in this case may be images, video, or audio, in addition to text. Interaction with the results provided by the IR system is seen as a two-stage process, involving examination of the search results or document surrogates, and examination of the documents.

Tags are often characterized as keywords or labels to help find content later. Thus a tagging site, or Web site providing tagging functionality has available a collection of possible search terms generated by its users. This collection of tags can be used to recommend search terms to support query formulation and reformulation. Tags also convey information – in their study of tags on del.icio.us, Golder and Huberman (2006) found that tags provided information about a bookmarked item, such as its topic, what kind of item it is (e.g., article, blog, book) and the identity of the owner or creator of the bookmarked item. Tags thus provide information for making “predictive judgments” (Rieh, 2002), if they are used to help a user decide whether to

open a Web page or not, whether for collections of bookmarks in del.icio.us or displayed search results in other systems. In tagging systems that display the tags along with their associated information object, tags can also be used to make “evaluative judgments” of relevance (Rieh, 2002).

Currently Web sites vary in how they present tags in their interfaces. For example, search results pages present tags in one of the following three ways: 1) no tags are shown; 2) a list of related tags is displayed separately from search result items, with no tags shown for individual search result items; 3) for each search result item, tags are shown if available. Sites also differ on how tags are displayed on document pages. On Flickr, a page for a specific photograph displays its associated tags in light gray and often requires scrolling down to view, making them easy to miss. Blogs often display tags in a much smaller font than the main text, again making them easy to miss.

How tags are integrated into search is also not transparent to the user. Some sites offer a “tag search” functionality separate from the usual keyword search. Depending on the site, the result of a tag search can be a list of tags matching the search term, or a list of items tagged with the specified search term. On sites that do not provide a separate tag search, search results include items that included the search term, whether in the tags associated with the item or in other parts of the item such as its title or text. On sites with a separate tag search functionality, a user has to make the additional decision of which type of search to carry out, while on sites with no separate tag search, the user avoids this decision at the expense of less control over the search.

### **Preliminary Results and Future Work**

An exploratory pilot study was carried out to determine the most fruitful way to investigate the role of tags in Web IR interaction. Users of two or more Web sites with tagging functionality – del.icio.us, Flickr, Last.fm, LiveJournal, or YouTube – were interviewed on their use of these sites. In the interviews, users were asked to interact with two of the sites as they would normally, while answering questions from the researcher. Interactions typically took the form of searches for known items. Preliminary findings indicate that interviewees did not display consistent usage of tags across sites, in that use of tags varied across sites for the same person. The differing presentation of tags at the interface level also influenced their use. For example, on YouTube tags were visible in the search results, but not visible in the actual video page, while tags were visible for all phases of interaction on del.icio.us. Interviewees saw tags on the search results page of YouTube as a source of possible search terms but did not mention them at all regarding relevance or query reformulation when viewing the actual video page. These inconsistencies indicated that use of existing Web sites as-is would not be suitable for examining the role of tags across the IR interaction process.

The next step is a user study in a laboratory setting, where users will be asked to use tags in their search tasks. This will require the implementation of an experimental system allowing data collection across the entire IR interaction process. A contribution of this study is better understanding of IR interaction on the Web, especially on the utility of tags. In addition, findings from this study can inform IR system designers on when and how to make use of tags to improve the IR interaction experience for users.

### **References**

Ames, M. & Naaman, M. (2007). Why we tag: Motivations for annotations in mobile and online media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 971-980). New York: ACM Press.

- Golder, S. A. & Huberman, B. A.. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198-208.
- Kipp, M. E. I. (2007). @toread and Cool: Tagging for time, task and emotion. In *Proceedings 8<sup>th</sup> Information Architecture Summit*, Las Vegas, Nevada.
- Marlow, C. et al. (2006). Position paper: Tagging, Taxonomy, Flickr, Article, ToRead. In *Proceedings of the 17th Conference on Hypertext and Hypermedia* (pp. 31-40). New York: ACM Press.
- Morrison, P. J. (2007). Tagging and searching: search retrieval effectiveness of folksonomies on the World Wide Web. Unpublished master's thesis, Kent State University, Kent, Ohio.
- Rainie, L. (2007). 28% of online Americans have used the internet to tag content. Retrieved February 17, 2007, from [http://www.pewinternet.org/pdfs/PIP\\_Tagging.pdf](http://www.pewinternet.org/pdfs/PIP_Tagging.pdf)
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology*, 53(2), 145-161.
- Wash, R. & Rader, E. (2007). Public bookmarks and private benefits: An analysis of incentives in social computing. In *Proceedings of the American Society for Information Science and Technology*, Milwaukee, Wisconsin.