
Why Do Tagging Systems Work?

George W. Furnas (panelist)

School of Information
University of Michigan
Ann Arbor, MI, USA

Caterina Fake (panelist)

Yahoo! Inc.
Sunnyvale, CA, USA

Luis von Ahn (panelist)

Dept. of Computer Science,
Carnegie Mellon University
Pittsburgh, PA, USA

Joshua Schachter (panelist)

Founder, del.icio.us, Inc.
San Francisco, CA, USA

Scott Golder (panelist)

Information Dynamics Lab, HP
Laboratories
Palo Alto, CA, USA

Kevin Fox (panelist)

User Experience Designer
Google, Inc.
Mountain View, CA, USA

Marc Davis (contributor)

Yahoo! Research Berkeley and
UC Berkeley School of Information
Management Systems (SIMS)
Berkeley, CA, USA

Cameron Marlow (contributor)

Yahoo! Research Berkeley
Berkeley, CA, USA

Mor Naaman (moderator)

Yahoo! Research Berkeley
Berkeley, CA, USA

Abstract

The panel will explore the relevance of the emerging tagging systems (Flickr, Del.icio.us, RawSugar and more). Why do they seem to work? What kinds of incentives are required for users to participate? Will tagging survive and scale to mass adoption? What are the behavioral, economic, and social models that underlie each tagging system? What are the dynamics of those systems, and how are they derived from the specific application's design and affordances?

We will demand answers to these questions and others from some of the pioneering practitioners and academics in the field. Bring your wireless laptop to participate in a live tagging experiment! The experiment results will be shown and discussed at the end of the panel. To add to the fun, parts of the discussion will be motivated by short video segments.

Keywords

Tagging, social software, Flickr, del.icio.us, design guidelines

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Panel Overview

Web-based social tagging systems such as Del.icio.us [1] and Flickr [2] allow participants to annotate a

Copyright is held by the author/owner(s).

CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.

ACM 1-59593-298-4/06/0004.

particular resource, such as a web page or an image, with a freely chosen set of keywords (“tags”).

Tagging offers a number of benefits to users. Perhaps the most straightforward advantage comes at a personal level: the act of tagging a resource is similar to bookmarking a resource for oneself. A link to the resource is saved in the user’s account, and can be retrieved from any web-connected device by using any of the tags used to describe the resource. However, similar keyword-based systems have existed in popular web browsers, photo repository software, and other collection organization systems for many years. So what’s new?

Tagging systems have become increasingly popular after an element of social interaction was introduced. *Social Tagging Systems* connect the individual bookmarking activities of users into a network of tags and resources shared among multiple users.

Social tagging systems, then, allow users to share their tags for particular resources. The tag sharing allows multiple added benefits, in discovery as well as retrieval. Discovery is using the tag as an anchor in finding interesting resources. For example, the resources tagged “CHI Tagging Panel”, by anyone in the community, will represent a collective image of related material as seen by the community.

Tags are also useful for retrieval, possibly making it is easier to find resources. As a simple consideration, a shared pool of tagged resources enhances the metadata for all users in the system while distributing the workload amongst many contributors. Users can utilize the system by searching for resources by the

associated tags. When a resource has been tagged by many users, it is more likely that the tag used for search will match one of the previously assigned tags. Social tagging systems may thus offer a way to overcome the “Vocabulary Problem”, first articulated by George Furnas et al. in [3], showing different users use different terms to describe the same things (or actions).

The low participation barrier of tagging has also been a significant factor in the proliferation of social tagging. Tagging systems afford efficient free-text tag entry, usually limited to one or a few words for each tag, and a few tags for each resource [4]. In particular, the users are not obliged to use any controlled vocabulary or predefined ontology. Using a carefully crafted taxonomy or categorization requires an expert to design the category structure and training to those who assign categories to resources. Conversely, tagging is a much lighter task cognitively that requires no previous consideration or training, but may result in very “noisy” metadata as seen from a classic categorization point of view.

Because of their lack of predefined structure, tagging systems reflect the conceptual structure of the user community. Based on this observation, the popular tags in social tagging systems have recently been termed *folksonomy*, a folk taxonomy of important and emerging concepts within the user group. Tradeoffs between user-created folksonomies and expert-crafted ontologies have been discussed and debated at length ([5,6] and more), and may be best left for a different conference; we shall not focus on these as the central issue in our panel.

We want to look at some different questions about why tagging systems seem to work, especially looking at the affordances of various tagging system sociotechnical designs and user interfaces. This topic has not yet been explored or discussed at length. We will also focus on the incentive structure that is needed for tagging systems to work, and on how the design and the growth of tagging systems could affect their usefulness.

Panel Format

The panel will not be a series of presentations followed by a Q&A session. The panel will be a lively discussion between the panelists and crowd, motivated by a few short video presentations. Attendees could participate in a live tagging experiment, the results of which will be discussed towards the end of the panel. Some panel resources will be available for a limited time after the conference at <http://research.yahoo.com/chi2006>.

While the panel format will allow crowd participation, we will try to focus on a few fundamental issues around tagging systems. Amongst those issues are:

- Do users understand tagging? It seems as though tagging was adopted by the technology savvy, and not yet brought to the masses. Will the masses ever tag? Or better yet, will they understand the role of tags contributed by other users and find usefulness in them?
- What system of incentives should be in place for a tagging system to succeed? Does the success of the system depend only on participation in sheer numbers, or are there other factors such as group coherence and homogeneity that are required? What are the tradeoffs between different types of

social structures and incentives for various social tagging systems?

- How does the design of tagging systems affect user participation? Does the design influence the content of tags and whether or not the community convergences on a coherent folksonomy? How have different social tagging systems designs addressed the Vocabulary Problem?
- How do tagging systems grow without losing relevance? Do user behaviors and information patterns change as the system grows from early adopters to whole populations? What are the expected patterns of change over time and drift? What personal information dynamics are derived from a long-term use of tagging systems? Simply put, how can social tagging systems scale?

Contribution of Panelists

The panel participants bring with them a multitude of experiences and views. In a setup that tries to combine practitioners, theoreticians and those who combine the theory and practice, we have:

- Luis von Ahn. Luis is the creator of the ESP Game [7]. The ESP Game innovated by introducing tagging systems that are based on game incentives. Luis had been working on inventing techniques for utilizing “human cycles”, the computational abilities of humans. These human cycles can be harnessed to help solving problems that are still difficult for a computer, for example, the content of images. Such systems, naturally, require incentives for users to participate as well as clever strategies to elicit a beneficial response. The ESP Game provides a framework for tagging images that supports the two listed requirements.

- Caterina Fake. Caterina is a pioneer and an experienced practitioner in the world of tagging systems. As one of the co-founders of the Flickr photo-tagging system, Caterina has closely observed how social structure, design choices and interface affordances influenced a tagging system's behavior over time. Now at Yahoo!, Caterina helps shape the future of tagging systems and is a major contributor to the Web 2.0 community.
- George Furnas. A widely published and cited researcher in the CHI community. George has introduced the "Vocabulary Problem", discussed above, back in 1987. He is also known for his work on Latent Semantic Indexing. These two major innovations are hugely relevant in the discussion of tagging systems, their usefulness, their relevance and their sustainability.
- Scott Golder is the author of one of the first academic papers studying the semantics, the usage, and the information dynamics in a tagging system [4], focusing on Delicious [1]. He has also studied various forms of participation in online communities. Scott is a co-organizer of a tagging workshop at the 2006 WWW conference.
- Joshua Schachter is the founder of one of the web's first collaborative tagging system, Delicious [1]. Joshua has been a long contributor to the development of influential social software, having also been a cofounder of the proto-weblog Memepool [8] and creator of the geo-tagging standard GeoURL. Joshua is a tagging visionary that believes many difficult problems of information retrieval can be made tractable with a community of annotation.

- Kevin Fox has hands-on experience with the design and use of labels (you can call them tags if you wish) in Google's Gmail service. Kevin is also intrigued by the lexical trends that are derived by social tagging systems and their counterparts (e.g., HTML linking and trackbacks). Moreover, Kevin believes that previous publisher- and community-driven tag-like systems can help us predict and overcome problems in the current generation of tagging systems. He is interested in various tagging systems, current and legacy, starting from Library of Congress subjects to meta-tag keywords to tag folksonomies.

References

- [1] Delicious, <http://del.icio.us>
- [2] Flickr, <http://www.flickr.com>
- [3] Furnas, G. W., Landauer, T. K., Gomez, L. M., and Dumais, S. T. The vocabulary problem in human-system communication. *Commun. ACM* 30, 11 (1987), 964-971.
- [4] Golder, S., and Huberman, B. A. The Structure of Collaborative Tagging Systems. HP Labs technical report, 2005.
- [5] Shirkey, C. Ontology is Overrated: Categories, Links, and Tags. http://shirky.com/writings/ontology_overrated.html
- [6] Merholz, P. Clay Shirky's Viewpoints are Overrated. <http://www.peterme.com/archives/000558.html>
- [7] von Ahn, L. and Dabbish, L. 2004. Labeling images with a computer game. CHI 2004. ACM Press, 319-326.
- [8] Memepool, <http://memepool.com>