Tagging with del.icio.us: Social or Selfish?

Emilee Rader
University of Michigan
School of Information
ejrader@umich.edu

Rick Wash
University of Michigan
School of Information
rwash@umich.edu

1. INTRODUCTION

del.icio.us is a website for "social bookmarking" where users can store and access their bookmarks online, along with descriptive keywords or "tags." When a user of del.icio.us logs in to their account and adds a bookmark, she may also tag that bookmark with any 10 or fewer single words that she feels are somehow related to that web page. Both the tags and the bookmarks are then publicly available to other users; searching by a tag returns all of the bookmarks stored in the system ever tagged with that word. Because the tags are public, that is, they can be viewed and searched by anyone, it is possible that users' choices regarding what tags to apply could be influenced by the tagging practices of others. A consensus might subsequently emerge for which tags should be used in a given context [6]. However, it has long been accepted that people use language imprecisely, and meaning is negotiated on-the-fly during conversation [2]. This imprecision is evident not only in communication, but also when people are asked to create keywords for recipes and names for common editing operations [4], and when user-generated index terms are compared with Library of Congress subject headings [3]. In fact, the probability that two people will generate the same label for the same object, called the "vocabulary problem", is widely held to be less than 20% [1,4].

When a user wants to take advantage of the collective properties of social bookmarking by browsing tags for the purpose of discovering new information, the "vocabulary problem" becomes apparent. Users who tag selfishly, without thinking about the public audience for their tags, are unlikely to choose the same tags to represent the same topics or concepts as other users. In del.icio.us this diversity might enhance the findability of specific pages. Because the system stores every single common and obscure tag ever used to refer to a particular bookmark, an individual looking for that content is more likely to search using a tag that someone else has already used. Unfortunately, it is precisely this diversity that decreases search precision and makes community convergence on a recognized and learnable tag vocabulary unlikely. When a given tag is applied to bookmarks in an inconsistent manner by many users, more variability exists in the content returned when a user searches with that tag. The desired bookmark may be returned, but there would be too much other "noise" in the results for it to be noticed. Users of del.icio.us interested in discovery would presumably be better off if there were tag convergence; however, due to the vocabulary problem this convergence seems unlikely to emerge.

CSCW'06, November 4-8, 2006, Banff, Alberta, Canada.

A question remains about whether users of del.icio.us practice social or selfish tagging. Marlow et al. [6] pointed out that while some people use tags for the purpose of organizing their own bookmarks, others intentionally choose to contribute to the value of "conceptual clusters", or call attention to the pages they bookmark, by adhering to conventions. Golder & Huberman [5] reported that over time the relative frequencies of tags applied to a web page stabilize into a pattern such that the most commonly used tags remain so and do not fall out of favor. They speculated that this could be because users imitate each other, or the user community is similar enough to naturally tag things the same way, or the stable content of the web page itself acts as a limit on tags people might choose. If it is true that in general users of del.icio.us actively practice "social tagging" (i.e. they choose to strive for tag convergence), then it is reasonable to assume that an analysis of tag frequencies for individual users and web pages would not show a similar pattern to the original "vocabulary problem" work published by Furnas et al. [4]. Therefore, an analysis of bookmark, user and tag data for 349 web pages downloaded via del.icio.us was conducted to discover whether the "vocabulary problem" is present in the way users select tags for web pages. Results indicate that there is very little inter-user agreement, suggesting that most users consciously or inadvertently tag selfishly.

2. DATASET

We collected a sample of 500 web pages that were bookmarked on del.icio.us, and listed on the "popular" and "recent" pages on several days between August and November 2005. We downloaded the "URL pages" for each, which list every user who bookmarked the web page and all the tags they applied. All "URL pages" were then re-collected at one time, in late December 2005, to obtain a consistent snapshot in time, and parsed to pull out the relevant data. Tags in del.icio.us are not case sensitive, but the system is sensitive to misspellings, tenses, and plurals. So for example, we treated the words "book" and "books" as unique tags, but "book" and "BOOK" as equivalent. Then, we eliminated the "extreme" web pages from the sample, retaining those bookmarked by 10 to 500 users (20^{th} to 80^{th} percentiles), and with 10 to 200 tags (20^{th} to 95^{th} percentiles). Both of these variables exhibit long-tail distributions consistent with Zipf's law [1,4] and we believe the 349 web pages that remain represent "typical" usage patterns for that time period.

We replicated analyses from Furnas et al. [4] to determine whether the vocabulary problem exists in del.icio.us. Because del.icio.us stores every tag (word) that is used to refer to every bookmarked web page (object), we chose a set of analyses from [4] that most closely approximated these parameters, called "Several names per object". We used two

measures that estimate in different ways the "repeat rate", or likelihood that a tag generated by a user is among the tags the system already has stored for that web page. Repeat rate indicates how likely it is for a search on a single tag to succeed. Finally, we also calculated inter-user tag agreement for all users who had bookmarked the 349 web pages [3]. This measure tells us on average, how often random pairs of users generate the same tag for the same web page.

3. RESULTS

The 349 web pages in our sample were bookmarked by 120.23 users (σ = 111.25) on average, who used 3.02 tags apiece (σ = 0.58). These web pages also tended to average 57.21 unique tags (σ = 37.09) associated with them. The three most common tags for each web page represented 45% of all tags applied for an average page, which is greater than the 33% reported by Furnas et al. [4]. Inter-user tag agreement, averaged over the sample, was 0.17, meaning that random pairs of users chose the same tag for the same web page just 17% of the time (σ = 0.10). While this percentage is low, it is higher than the 8% reported in [4] for their text-editing operations dataset.

Repeat rate statistics were calculated in two ways. In the first calculation, "weighted random", when a user searches for a specific page with a single tag, the probability of success depends upon the relative frequencies with which different tags are associated with web pages. Pages associated most frequently with the given tag will be returned most often. The second calculation, called "optimized" in [4], rank-orders the web pages for each tag by frequency, and always returns them such that the highest frequency page is returned first, then the next highest, and so on. Table 1 shows comparison of the repeat rate statistics reported in [4] with those calculated for our sample, limiting the maximum number of tags stored per web page to M.

Table 1. Comparison of repeat rate statistics

	Furnas [4] Common objects	del.icio.us weighted random	del.icio.us optimized lower	del.icio.us optimized upper
M=1	.12	.08	.17	.19
M=2	.21	.15	.28	.31
M=3	.28	.22	.37	.41

The values in Table 1 represent the probability that a user entering a single tag will be successful in their search for a specific web page, depending on how many tags the system is able to store for each web page. Success rates are very similar between our sample and the results reported in Furnas et al. Figure 2, below, illustrates the optimized lower and upper bound values for M=1 through M=40. Success rate increases dramatically between 1 and 10 tags (words) for our sample, and then appears to level off. In [4], success rate appears to level off after about 8 words. These results lead to the conclusion that the "vocabulary problem" does exist, and that selfish tagging, not social, is prevails in `del.icio.us`.

4. CONLCUSION

`del.icio.us` supports both social bookmarking and social tagging. Social bookmarking depends upon nothing more than the public nature of the bookmarks individuals store online. However, the success of "social tagging" depends on consensus for how tags are applied to content. `del.icio.us` is able to store a large

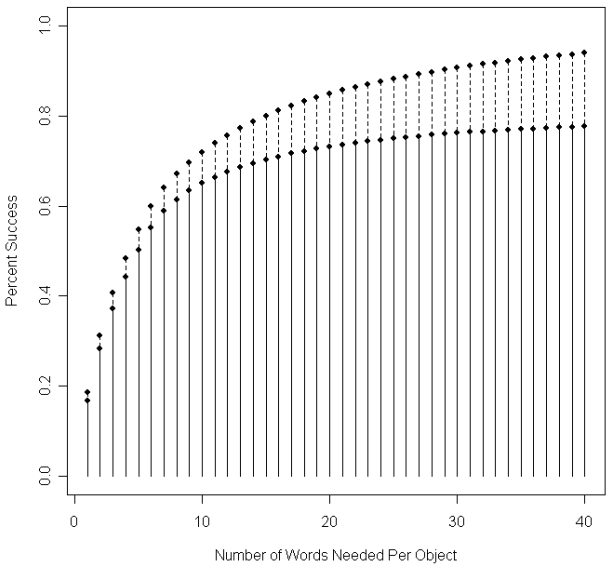

Figure 1. Percent success based on "optimized", as a function of M (number of words stored for each object).

number of tags for every object, which dramatically increases the probability that a sought-after web page will be among the search results returned for a given tag. However, because the "vocabulary problem" is present, tags are applied inconsistently and it is difficult for a user who wants to learn about a particular topic to sort out what she means when she uses the tag, from what others mean when they use it. One can imagine the possibility of a system in which both selfish and social tagging coexist; however, this would likely require a human or algorithmic indexer, or editorial control over tag synonyms and usage [1]. The goal would be to eliminate inconsistency in the way common tags are used, while retaining the rare tags.

5. REFERENCES

[1] Bates, M. J. (1998). Indexing and access for digital libraries and the Internet: Human, database, and domain factors. *Journal of the American Society for Information Science, 49(13),* 1185-1205.

[2] Clark, H. H. (1996). Using Language. Cambridge, England, UK: Cambridge University Press.

[3] Collantes, L. Y. (1995). Degree of Agreement in Naming Objects and Concepts for Information Retrieval. *Journal of the American Society for Information Science, 46(2),* 116-132.

[4] Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1983). Statistical Semantics: Analysis of the Potential Performance of Key-Word Information Systems. *The Bell System Technical Journal, 62(6)*, 1753-1806.

[5] Golder, S., & Huberman, B. A. (2006). The Structure of Collaborative Tagging Systems. *Journal of Information Science, 32(2),* 198-208.

[6] Marlow, C., Naaman, M., boyd, d., & Davis, M. (2006). Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. In the *Proceedings of WWW 2006 Collaborative Web Tagging Workshop*, Edinburgh, Scotland.