Using Social Tagging to Improve Social Navigation

David R Millen and Jonathan Feinberg

IBM T. J Watson Research Center
1 Rogers Street, Cambridge, Massachusetts. USA
{david_r_millen, jdf}@us.ibm.com

Abstract. In this paper, we explore the increasingly popular social bookmarking services. These services powerfully combine personal tagging of information sources with interactive browsing, which allows for improved social navigation. We examine the use of a social bookmarking service, deployed in a large organization, to understand how social navigation is supported. We conclude that social tags used in the context of a social bookmarking service are an important way to improve social navigation.

1 Introduction

For several years, there has been work underway to understand how social information can be used to enhance information exploration and discovery and to generally improve information spaces. *Social navigation* is a concept that is generally used to describe navigation that is "driven by the actions from one or more advice providers [4]." Svennson and Hook [18] add that an advice provider can be a person or an artificial agent providing navigational advice. They further point out that social navigation can be either *direct* or *indirect*. Direct social navigation is the direct communication of navigational advice (e.g. in a chat or email) from one person to another; it is characterized by intentional human action or input. Indirect social navigation is when navigational advice is inferred from historical traces left by others. Indirect social navigation involves monitoring and analyzing the behavior of a group of people.

Dieberger [3] notes that an early form of *direct* social navigation support was found in the lists of favorite web sites (a.k.a. "hotlists") found on many personal web pages, and in the recommendations for related sites that were found on many organizational home pages. Hill and Hollan [9,10] provided an example of *indirect* social navigation when they monitored both authors' and readers' online experiences with various documents, and then computed measures of what they called *read wear* and *edit wear*. They then provided navigational support in the form of "attribute-mapped scroll bars" to enable rapid browsing to items that appeared to be of greater interest to most readers. Wexelblat and Maes [23] explored the use of "interaction histories" in a group of tools that allowed navigation based on maps, paths and signposts. In a field study of these tools, they found productivity improvements (i.e. same work with less effort) and high rates of user comprehension of the navigation

models. Yet another early example of indirect social navigation of the web was provided in the PHOAK project [19]. Terveen et al. mined newsgroups to identify web links that were in the discussion, and used text analysis to determine if the link was likely to be a recommendation. They then provided a list of the most-recommended sites for a particular newsgroup.

2 Social Bookmarking Systems

The desire to explicitly share information among small groups, teams and communities of practice has led, not surprisingly, to the development of a number of shared bookmarking systems. Early shared bookmarking systems often used automated techniques to support the creation and categorization of collections of web bookmarks [14, 15, 24]. Other shared bookmarking systems incorporated end-user ratings of web pages [1, 6]. These innovative systems met with some success, although they consistently seemed to fall short of their potential use. Several explanations for their limited success have been offered, including limited privacy protection, little support for end-user tailorability and high requirements for active user participation [13].

Recently, there has been a reemergence of shared bookmarking applications, whose tremendous popularity and growth of use have prompted a second look at this kind of collaborative software. Introduced in 2003, the del.icio.us [2] social bookmark manager was one of the first of this kind of application, and has enjoyed an early and large base of committed users. A flurry of similar offerings has since been unveiled; for a review see Hammond et al [8].

Recently, these internet oriented social bookmarking services have been adapted for use in large organizations. One example is the *dogear* social bookmarking service, which supports bookmarks of internet and intranet information sources, and provides user authentication via corporate directories [17].

These systems share a number of features. First, they allow individuals to create personal collections of bookmarks and to share their bookmarks with others. These centrally stored bookmark collections bring immediate personal benefit by providing a collection that can be browsed from any web-accessible machine.

A second, and significant, enhancement in these systems is the use of keywords or tags that are explicitly entered by the user for each bookmark. These tags allow the individual user to organize and display their collection with labels that are meaningful to them. Furthermore, multiple tags allow bookmarks to belong to more than one category, avoiding one of the limitations of the hierarchically organized folders of bookmarks (or "favorites") found in most web browsers. The use of tags to create an emergent classification system—a "folksonomy"—has been somewhat controversial and is likely to spawn significant research in the short term [7, 16, 22].

The final distinguishing characteristic of these social bookmark applications is the social nature of their use. There is a bias towards increased transparency in these tools. While bookmark collections are personally created and maintained, they are typically also visible to others. A number of user interface elements allow social browsing of the bookmark space. For example, user names are "clickable" links, and,

when a name is clicked, the bookmark collection for that user is presented. This allows someone to get a sense of the topics of interest for a person. Tags are also clickable, and when selected will result in a list of all bookmarks that share that tag. This is a useful way to browse through the entire bookmark collection to see other information sources of interest. We call this ability to reorient the view by clicking on tags or user names, "pivot browsing"; it provides a lightweight mechanism to navigate the aggregated bookmark collection.

These new social bookmarking applications are a natural and powerful extension of existing social navigation tools and practices. They provide a mix of both direct (intentional) navigational advice as well as indirect (inferred) advice based on the collective public behavior. In this paper, we present some results from a field study of the *dogear* social bookmarking service. We will describe the specific design elements of the application that support social navigation, and then provide empirical evidence of enhanced social navigation from a seven month field study of the service.

3 Design for social navigation

The *dogear* social bookmarking service was designed to simply and elegantly display bookmarks within a navigation model that allows users to manage and explore the collection in different ways. The user's bookmark collection is a reverse chronological list of their most recent bookmarks, similar in format to a blog. Each bookmark has a number of pieces of metadata which give the user useful information about its context and content. Other views include recent bookmarks, popular bookmarks, bookmarks to a specific URL, bookmarks for any combination of tags and user, and text search results. A screen shot of *dogear* can be seen in Figure 1.

Fig. 1. Screen shot of *dogear* service. A – tag cloud with browsable tags, B – list view of bookmarks, C – tags for each bookmarks (also clickable).

A number of important design principles for social navigation tools have been described by Forsberg, Hook and Svensson [5] They argue that it is important to have

tools that integrate with other everyday tools, show the presence of others (either synchronously or asynchronously), provide mechanisms to ensure appropriate behavior, build trust, and follow explicit privacy policies. Furthermore, there needs to be a way to personalize the navigation recommendations for individuals.

The *dogear* social bookmarking service has been designed to integrate with everyday work tools. It is a browser-based application, accessible to everyone within a large enterprise, and intended to navigate both intranet and internet web sites. To get started with the *dogear* social bookmarking service, an end-user need only install a browser toolbar button (a few lines of JavaScript) to allow easy bookmarking of web pages. The presence of others is always visible: a navigation bar shows a list of the most active users (described in more detail below), while the *dogear* home page displays the most recent bookmarks, which include the names of the bookmark creators and the dates of creation. To ensure appropriate use for this business-oriented application, all users are authenticated against the corporate directory, and real name identity is used to associate names with bookmarks. To help build an environment of trust and protect privacy, there is a distinction between public (shared) and private bookmarks. And finally, we personalize the browsing experience by providing a "my bookmarks" view, allowing individuals to see their personal bookmark collection.

3.1 Social Tagging of Content

One of the major innovations in social bookmarking applications has been the widespread adoption of user-generated keywords (or tags) that are associated with the web content. The *dogear* design reveals the tag history in the form of what have been popularly called tag clouds (see Figure 1). A slider control allows the tag cloud to be expanded or contracted to reveal more or less of the tag index. Font darkness is used to show more frequently used tags, with a darker font indicating more use.

While human generated keywords as metadata have been available in many applications for a long time, we think that the ability to pivot browse the bookmark collection using tags is important. We believe that the interactive nature of social tags, that is, the ability to click and browse bookmarks based on the tags, is an important design characteristic and provides an immediate benefit to the user for having provided the tags in the first place.

The tag cloud (or tag index) supports easy social navigation in that each of the tags is clickable; clicking a tag leads to a view of bookmarks that are associated with that tag. Tag clouds are either system-wide, or specific to one user, depending on the current view. A system-wide tag cloud would quickly grow to an unmanageable size—after 8 months, the number of distinct tags in the *dogear* service was over 12000. To make the enterprise tag collection manageable, we bound the tag cloud to include only the most active tags.

User tag clouds also provide a frequency slider to allow easy examination of the tags, while a distinctive name label is provided at the top of the tag cloud. The tag cloud for an individual allows viewers to get a sense of the current interests of the collection owner. The individual tag collections also provide important navigational support as each bookmark collection can also be browsed by simply clicking on a tag.

The design of the *dogear* service also allows direct navigation to bookmarks that are tagged with two or more tags in combination, as well as to navigate the bookmark collection through direct navigation of people links.

4 Social Navigation: Results of Field Study

Our general understanding of the use of the *dogear* service was based on many sources of user data, including log files, the data in the bookmark repository, email and blog comments and feedback about the eservice, and a short online survey. In this paper we focus on data from user activity level log files. Included in the log files are user actions (e.g., create, delete, edit a bookmark, bookmark "clicks"), user and bookmark owner identifiers, and a time and date stamp. In addition, we have analyzed aspects of the bookmark collection itself, which provides additional information about the composition of bookmarks (e.g., tag structure information).

The user activity analysis presented here is based on log files covering an eight month period from July, 2005 to March, 2006. During this usage period, 2579 individuals were recorded using the *dogear* service, with 909 (35%) of the participants creating at least one bookmark. To date, 58532 bookmarks have been created in one of four ways: new bookmarks created directly in *dogear*, imported bookmarks from another social bookmarking service, imported bookmarks from a local browser (Internet Explorer or Firefox) and direct copying someone else's bookmark. In the current bookmark collection 35.7% were created within *dogear,* 23.4% were imported from another service, and 38.7 % from local browser bookmarks and 2.2% were copied from another *dogear* user. While it is possible to have private bookmarks in the *dogear* service, the default setting when a new bookmark is created is "shared." To date, 98% of the bookmarks are public/shared.

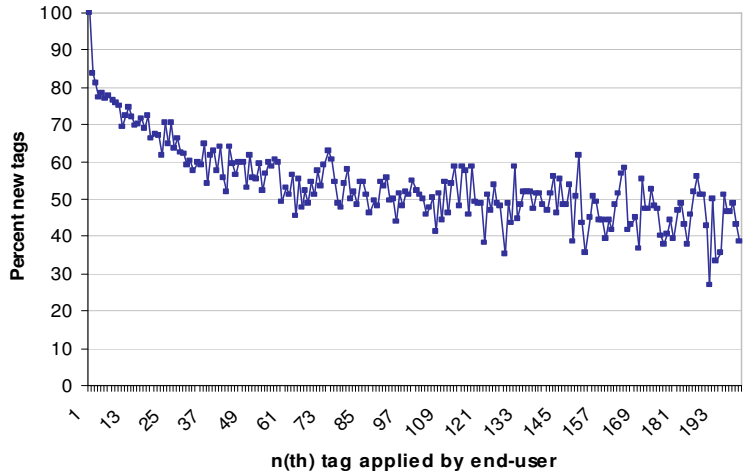

Fig. 2. Distribution of new tags as a function of order of application.

We were particularly interested in the social tagging behavior for end-users of the *dogear* service. To date, 16577 unique tags have been generated for the over 50K bookmarks. The modal number of tags per bookmark is 2, with 71 % of bookmarks having three or fewer tags. There is considerable tag reuse by end-users, which confirms active management of personal tag collections or personal folksononmies. In Figure 2, we show the average percentage of new tags as a function of the number of tags that have been previously entered by an individual. There is a gradual decline in new tags as end-users enter more bookmarks. This gives us a sense that there us significant reuse of tags over time.

4.1 Supporting Social Navigation

We were particularly interested in understanding how the *dogear* service supports social navigation. One kind of evidence for social navigation would be found in the number of times individuals looked at the bookmark collections of other people. In total, 2545 (98.7%) of *dogear* users used tag or people links at least one time to browse and explore the bookmark collection.

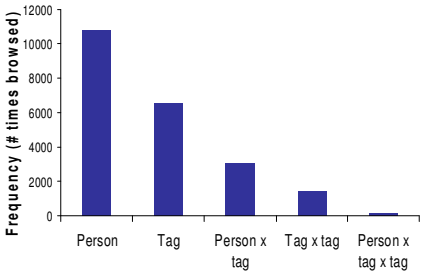

Fig. 3(a). # of times browsed by type.

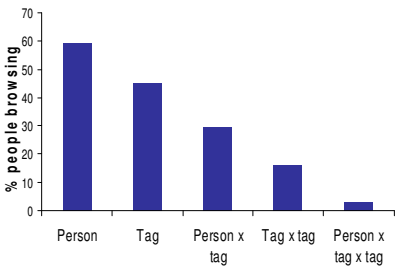

Fig. 3(b). % of users browsing by type..

In Figure 3, we show the number of times end-users looked at the bookmark collections of others as well as the proportion of *dogear* end-users who browsed in each particular manner. There is considerable browsing of the bookmark space by other people, other tags (everyone), and other people's tags. These results suggest widespread curiosity about what others are bookmarking. The most frequent way to browse bookmarks is by clicking on another person's name, followed by browsing bookmarks by selecting a specific tag from the system-wide tag cloud. It is considerably less common for a user to select a tag from another user's tag cloud and there is almost no use to date of the more advanced browsing of tag intersections.

The results reported in Figure 3 provide evidence of the explicit social navigation that is taking within the *dogear* service. We were curious about what kinds of tags were most often browsed. It seems likely that the tags most likely to be browsed would very simply be those that occur most often in the bookmark collection. These frequently applied tags were presumably created by the end-users as they were

expected to be useful for "refinding' the bookmarks at some time in the future. In Table 1, we show the top ten most *browsed* tags as well as the ten most *popular* (i.e. most frequently applied) tags in the collection. Six tags appear on both lists. A reliable positive correlation exists between the frequency that a tag is browsed and the frequency that a tag appears in the bookmark collection is .67 ($p < .001$).

Table 1. Top ten tags browsed compared to top tags applied.

Top 10 Tags Browsed	N	Top 10 Tags Applied	N
tools	162	tools	1771
ajax	159	java	1247
collaboration	128	design	1087
linux	110	ajax	1046
blog	97	linux	1014
java	93	software	986
firefox	90	firefox	984
dogear	85	web2.0	869
eclipse	84	programming	867
websphere	67	blog	858

In addition to understanding which specific tags were browsed most often, we were also curious about whether we could see evidence of tag clustering, which would show that individuals who browsed one tag would be more likely to browse *related* tags. We performed a two-mode social network analysis [21] to understand the tag co-browsing similarity among *dogear* users. We then extracted a single mode network which reveals a network in which the nodes are tags, and the edges indicate that the same person browsed both tags. We have presented a small portion of the tag network in Figure 4. There are clear tag clusters visible in the network diagram. At the center of the biggest cluster is the tag *ajax* with a number of connected tags related to programming terms. There is also a cluster (upper right) with a central node labeled *architecture*. There are two isolated tag clusters, seen at the bottom of the figure, which may indicate a group of people working in a similar area or project.

The results reported above provide evidence of the explicit social navigation that is taking within the *dogear* service. We are encouraged by these results as they represent a novel form of information browsing within the enterprise. While the results indicate that users of the system are looking at the bookmark collections of others, they do not tell us whether or not users are clicking through to those bookmarked sites.

We examined, therefore, the number of times that *dogear* users clicked through on a URL that had been bookmarked by another. During the trial period, 89% of individuals (2291 of 2579) clicked on URLs that had been bookmarked by another person. 74% of the total pages visited (32596 of 44144) were bookmarked by someone else. This provides considerable evidence that the *dogear* service is supporting a high degree of social navigation.

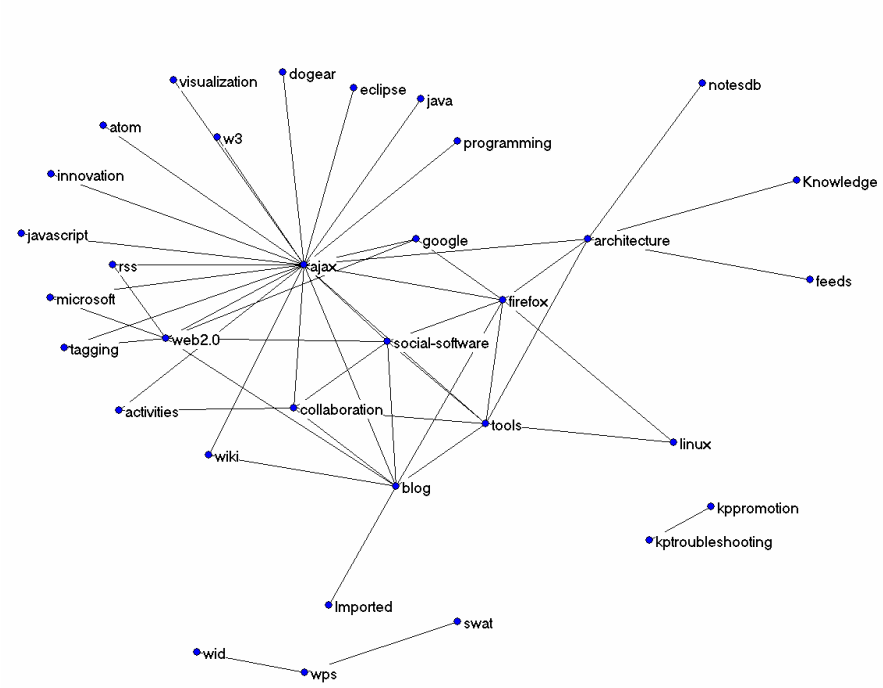

Fig. 4. Tag browsing network for a subset of *dogear* tags.

5. Discussion and Next Steps

In this paper we have described those navigation elements of a social bookmarking application that have been designed to support enhanced social navigation. These design elements afford navigation through the bookmark collection by tags, by people, or by combinations of tags and people. The results of our log file analysis confirm that these navigational elements are in use by a large number of the users of the social bookmarking application under study. Indeed, approximately 60% of the *dogear* service visitors explored the bookmark collection using one or more of the pivot links (tags or people).

The results show users' slightly greater preference for looking at another person's entire bookmark collection than for browsing tag collections. This lightweight mechanism to explore the bookmark space is promising. Furthermore it appears that almost every visitor to the bookmark service clicked through someone else's bookmark to look at the original document source on the internet or intranet. This is also encouraging as a way to promote easy information sharing within the organization.

We are encouraged by the rapid adoption of the *dogear* social bookmarking service and will continue to explore ways to improve the service and learn from the ongoing field trial results. We are convinced that much is to be learned about social tagging

behavior in general, which we have argued is an important aid for social navigation. Much of this work is centered on how tags are used and how they evolve over time [7, 20]. Others are beginning to explore how tags can be used to support navigation [11, 12]. We are extremely interested in team or organizational specific tag vocabularies and how they improve performance of various information management tasks.

A second major area of interest is the integration (and exploitation) of social bookmarking services to aid in enterprise search tasks. While much of recent research in search technology has been focused on automated-tagging of content, we are optimistic that harnessing the social tagging that we have observed in *dogear* service will provide a significant boost in search effectiveness.

7. REFERENCES

1. Balabanović, M., Shoham, Y.: Fab: content-based, collaborative recommendation. Commun. ACM 40, 3 (Mar. 1997) 66-72
2. Del.icio.us; http://del.icio.us/
3. Dieberger, A.: Supporting Social Navigation on the World-Wide Web. Int. J. Human Computer Studies, Vol. 46 (1997)
4. Dourish, P., Chalmers, M.: Running out of space: models of information navigation. In Proc. HCI'94, Glasgow, August (1994)
5. Forsberg, M., Höök, K., Svensson, M.: Design Principles For Social Navigation Tools. In Proc. 4th ERCIM Workshop on 'User Interfaces for All', Special Theme 'Towards an Accessible Web', Stockholm, Sweden, 19-21 October (1998)
6. Glance, N. Arregui, D., Dardenne, M.: Knowledge pump: Supporting the flow and use of knowledge. In Information Technology for Knowledge Management. Springer-Verlag (1998)
7. Golder, S., Huberman, B. A.: The Structure of Collaborative Tagging Systems. Technical report, Information Dynamics Lab, HP Labs (2005)
8. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social bookmarking tools : A general review. D-Lib Magazine, 11,4 (April 2005)
9. Hill, W. C., Hollan, J. D., Wroblewski, D., McCandless, T.: Edit wear and read wear. In Proc. SIGCHI Conference on Human Factors in Computing Systems (Monterey, California, United States, May 03 - 07, 1992). P. Bauersfeld, J. Bennett, and G. Lynch, Eds. CHI '92. ACM Press, New York, NY (1992) 3-9
10. Hill, W.C. and Hollan, J.D.: History-enriched digital objects. In Third ACM Conference on Computers, Freedom and Privacy, San Francisco, CA, ACM, (1993) 917-20
11. Jarrett, A. C. and Dennis, B. M.: NusEye: designing for social navigation in syndicated content. In Proc. Conference on Diversity in Computing (Albuquerque, New Mexico, USA, October 19 - 22, 2005). TAPIA '05. ACM Press, New York, NY (2005) 17-19
12. Jones, W., Munat, C.F., Bruce, H., Foxley, A.: The Universal Labeler: Plan the Project and Let Your Information Follow. In Grove, Andrew, Eds. Proceedings 68th Annual Meeting of the American Society for Information Science and Technology (ASIST) 42, Charlotte, NC (2005)

13.Kanawati, R. and Malek, M.: A multi-agent system for collaborative bookmarking. In Proc. First international Joint Conference on Autonomous Agents and Multiagent Systems: Part 3 (Bologna, Italy, July 15 - 19, 2002). AAMAS '02. ACM Press, New York, NY, (2002) 1137-1138
14.Keller, R. M., Wolfe, S. R., Chen, J. R., Rabinowitz, J. L., Mathe, N.: A bookmarking service for organizing and sharing URLs. In Selected Papers From the Sixth international Conference on World Wide Web (Santa Clara, California, United States) (1997) 1103-1114
15.Marais, H. and Bharat, K.: Supporting cooperative and personal surfing with a desktop assistant. In Proc. 10th Annual ACM Symposium on User interface Software and Technology (Banff, Alberta, Canada, October 14 - 17, 1997). UIST '97. ACM Press, New York, NY (1997) 129-138
16.MacGregor, G., and McCulloch, E.: Collaborative tagging as a knowledge organisation and resource discovery tool. *Library View* (accepted), 55, 5 (2006)
17.Millen, D., Feinberg, J., Kerr, B.: Social bookmarking in the enterprise. In Proc. SICHI Conference on Human Factors in Computing Systems. Montreal, Canada. April 24-27 (2006)
18.Svensson, M., Laaksolahti, J., Höök, K., Waern, A.: A recipe based on-line food store. In Proc. 5th international Conference on intelligent User interfaces (New Orleans, Louisiana, United States, January 09 - 12, 2000). IUI '00. ACM Press, New York, NY (2000) 260-263
19.Terveen, L., Hill, W., Amento, B., McDonald, D., Creter, J.: PHOAKS: a system for sharing recommendations. Commun. ACM 40, 3 (Mar. 1997) 59-62
20.Walker, J.: Feral hypertext: when hypertext literature escapes control. In Proc. Sixteenth ACM Conference on Hypertext and Hypermedia (Salzburg, Austria, September 06 - 09, 2005). HYPERTEXT '05. ACM Press, New York, NY (2005) 46-53
21.Wasserman, S., and Faust, K.: Social Network Analysis: Methods and Applications. New York and Cambridge, ENG: Cambridge University Press (1994)
22.Weinberger, D. Taxonomies to Tags: From Trees to Piles of Leaves. Release 1.0 February (2005) 2-32
23.Wexelblat, A. and Maes, P.: Footprints: history-rich tools for information foraging. In Proc. SIGCHI Conference on Human Factors in Computing Systems: the CHI Is the Limit (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999). CHI '99. ACM Press, New York, NY (1999) 270-277
24.Wittenburg, K., Das, D., Hill, W., Stead, L.: Group asynchronous browsing on the World Wide Web. In Proc. 4th International WWW Conference (Boston MA, December 1995), O'Reilly & Associates (1995) 51-62