

Social bookmarking and exploratory search

David Millen, Meng Yang, Steven Whittaker*, Jonathan Feinberg
IBM Cambridge, MA, USA 02142, *Sheffield University, Sheffield, S1 4DP, UK
{david_r_millen,yangmeng,jdf}@us.ibm.com}, s.whittaker@shef.ac.uk

ABSTRACT. In this paper, we explore various search tasks that are supported by a social bookmarking service. These bookmarking services hold great potential to powerfully combine personal tagging of information sources with interactive browsing, resulting in better social navigation. While there has been considerable interest in social tagging systems in recent years, little is known about their actual usage. In this paper, we present the results of a field study of a social bookmarking service that has been deployed in a large enterprise. We present new qualitative and quantitative data on how a corporate social tagging system was used, through both event logs (click level analysis) and interviews. We observed three types of search activities: community browsing, personal search, and explicit search. Community browsing was the most frequently used, and confirms the value of the social aspects of the system. We conclude that social bookmarking services support various kinds of exploratory search, and provide better personal bookmark management and enhance social navigation.

Introduction

In recent years, there has been tremendous growth in shared bookmarking applications. Introduced in 2003, the del.icio.us social bookmark website 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 [see (Hammond, et al., 2005) for a recent review].

These internet oriented social bookmarking services have been adapted for use in large organizations. Examples include the *dogear* (Millen, et al., 2006) and *onomi* social bookmarking services (Damianos, et al., 2007). Both of these enterprise-ready bookmarking services support bookmarking of both internet and in-

tranet information sources, and provide user authentication via corporate directories.

There are two distinguishing characteristics of social bookmarking systems. The first is the use of keywords, or tags, that a user enters to describe the links he or she saves. These tags allow users 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, a limitation of the traditional hierarchically organized folders found in most Web browsers. The second significant characteristic of these social bookmark applications is the social nature of their use. While bookmark collections are personally created and maintained, they are also typically visible to others. As a result, users benefit by getting pointers to new information from others while at the same time getting a general sense of other people's interests.

These new social bookmarking applications are a natural and powerful extension of existing social navigation tools and practices (see, for example, (Dieberger, 2003; Munro, 1999)). They provide a mix of both direct (intentional) navigational advice as well as indirect (inferred) advice based on collective public behavior. By definition – these social bookmarking systems provide “social filtering” on resources from the web and intranet. The act of bookmarking indicates to others that one is interested in a given resource. At the same time, tags provide semantic information about the way the resource can be viewed.

Social bookmarking systems arguably provide support for search activities that range from simple fact-finding to more exploratory or social forms of search. Fact-finding or what is called “known-item” retrieval is supported by traditional application *explicit search* capabilities. Users generate query terms and sift through lists of search results to find the appropriate bookmark (and associated web site). These known-item search tasks are usually characterized by a well understood search problem and reasonable understanding of the search domain.

Known-item retrieval is also supported in social bookmarking applications by *browsing* through collections of one's own (personal) bookmarks, which have been explicitly created, tagged and annotated by end-users. Social bookmarking applications typically allow personal bookmark browsing in one of two ways. The first is by sifting through scrollable pages of bookmarks, and the second is by or by performing a *tag query* of the collection by clicking on a tag.

Social bookmarking tools also support *exploratory* search activities. In exploratory search, the problem definition is less well structured and the emphasis may be on learning or analysis (Marchionini, 2006). One form of this less goal-oriented browsing found in social bookmarking applications is to browse bookmarks by *time*, enabling end-users to serendipitously follow recent bookmark that they find interesting. A second exploratory browsing strategy supported by social bookmarking applications is to explore *popular* bookmarks, where frequency of bookmarking a specific URL is a simple measure of popularity.

There are two other particularly interesting forms of exploratory search supported by social bookmarking services. The first is where end-users click on a visible name and the bookmarks for that person appear. The second is where collaborative tags are used to query the bookmark collection.

While there has been considerable interest in social tagging systems in recent years, little is known about their actual usage. Some work has been done to investigate tag growth and entropy (Golder and Huberman, 2006; Kittur, et al., 2007; Marlow, et al., 2006) and tag choice (Sen, et al., 2006). Little is known, however, about actual usage and the ways in which social bookmarking might help people find information.

In this paper, we present the results of a field study of a social bookmarking service that has been deployed in a large enterprise. We present new qualitative and quantitative data on how a corporate social tagging system was used, through both event logs (click level analysis) and interviews. We are generally interested in understanding the different ways that this social bookmarking service supports different kinds of search. In particular, we are interested in understanding how social bookmarking tools are used to find, refind and explore information resources.

These are important ideas to explore in the enterprise context. Enterprise search has been shown to be ineffective (see, for example, (Mukherjee and Mao, 2004)) and such social search tools may provide a significant alternative method for enterprise information access.

Dogear Social Bookmarking Service

The dogear social bookmarking service is a social bookmarking tool designed to support organizations and large corporations. Adapting social bookmarking to the corporate environment meant enhancing some of the technology's standard features. Rather than allow the use of pseudonyms, dogear requires the use of real names and authentication with a corporate directory. Real name identity allows dogear users to look-up additional information about other people in various corporate databases (e.g., corporate online directory and the enterprise Web). It also facilitates communication between users of dogear since most corporate collaboration tools (such as e-mail, and chat) use real name identities. For example, if someone is looking for a colleague who is knowledgeable about Java, he or she might look to see who has bookmarked articles on that topic and then send an e-mail to get more information. Finally, real name identity may also promote more responsible use of the system. A user who is required to identify himself or herself should be far less likely than a pseudonymous user to post and share links to questionable material.

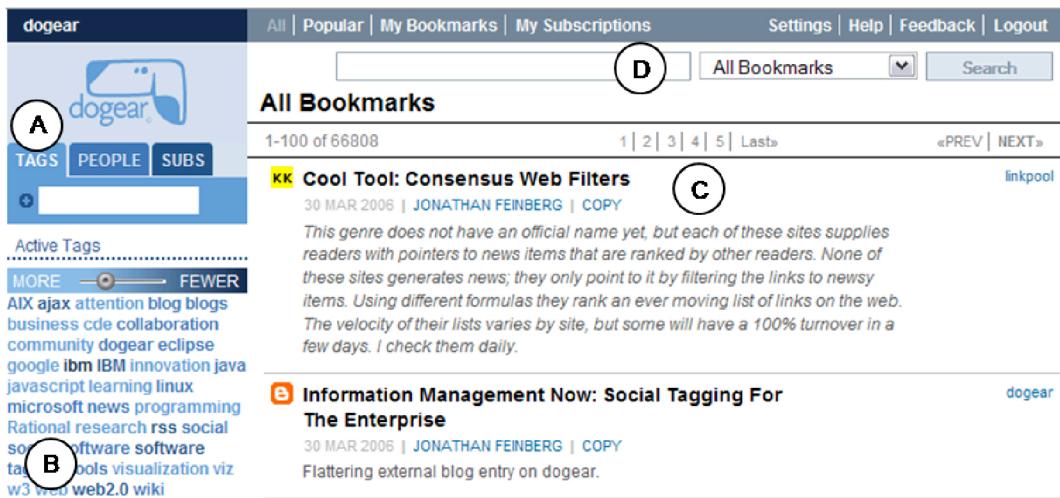


Figure 1. Screen shot of dogear social bookmarking service.

A second distinguishing characteristic of dogear is that it was designed to work behind corporate firewalls. This allows intranet resources to be bookmarked and shared among coworkers (e.g., human resource links, team or project resources, etc.) For example, if an employee wanted information on how to fill out expense report forms, he or she might search dogear using the tag "expense report" to find a list of intranet bookmarks that others in the company found useful.

Figure 1 shows the front page of dogear, featuring the bookmarks most recently added by a dogear user. The tabs at the top of the left-hand column (A) provide: a link to the user's tags; a list of people who have bookmarked the same URLs; and a list of the individual's bookmark subscriptions. The Active Tags (B) area shows an indexed tag cloud with a slider that can reduce the number of tags shown based on their frequency of appearance.

A list of bookmarks runs down the center of the screen (C). Each bookmark includes a title, optional descriptive text, tags, when the bookmark was made, and information about the author. Clicking on the title takes you to the bookmarked site. Clicking on a tag takes you to a list of the other bookmarks that the author has tagged with the same keyword. Clicking on the author's name takes you to a list of the author's bookmarks. Dogear's search (D) lets you search bookmarks by tag, person, or type of collection (all bookmarks/my bookmarks).

Social Tagging of Content

One of the major innovations in social bookmarking services has been the widespread adoption of user-generated keywords (or social tags) that are associated with the web content. In most social bookmarking services, the tag histories are revealed in the user interface in what have been popularly called *tag clouds* (see, for example Figure 2(a)). In the service we studied, 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 browse the bookmark collection via tags is critically important, as *pivot browsing* provides a real and immediate benefit to the end-user for having provided the tags in the first place.

Tag clouds are either system-wide, or specific to one user, depending on the current view. System-wide tag clouds quickly grow to an unmanageable size. In the social bookmarking service studied in this research, the enterprise tag cloud was truncated to include only the most active or popular tags.

In the system studied here, personal tag clouds were also provided and visible to other users of the bookmarking service. The personal tag clouds allowed other viewers to get a sense of the current interests of the other coworkers. The individual tag collections also provide important navigational support as another person's bookmark collection could be browsed by simply clicking on tags.

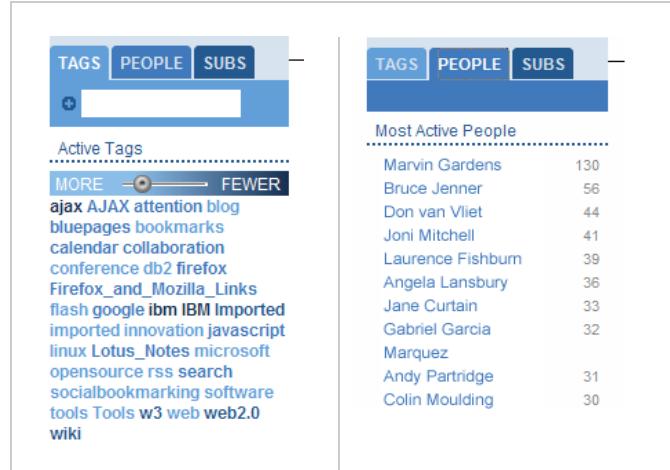


Figure 2(a) tag cloud

Figure 2 (b) People links

The bookmarking service under study also supported direct navigation to bookmarks that are tagged with two or more tags in combination. As the user selects additional associated tags, a new list of tags that co-occur with the previous ones is presented. This allows individuals to easily navigate the bookmark collection via tag clusters.

Social Navigation via People

In addition to browsing with end-user generated tags, it is also possible to navigate the bookmark collection through direct navigation of people links. These links are created using the *author* metadata that is associated with each bookmark. As can be seen in Figure 2(b), the navigation interface presents a separate tab, labeled "people." In the social bookmarking service home page, a list of the most

active users is shown, with a number indicating the number of bookmarks recently created by each user. These “live names” may be clicked and the bookmark collection for that individual is presented.

This navigation by person allows for casual, or intentional, browsing of the bookmark collections of people of interest. End-users who see names of people who they know to be strong information seekers or brokers can navigate directly to see their bookmark collection.

dogear search

The search box in *dogear* works as expected; the user types in a free text query, and a list of bookmarks is returned, in order of relevance. The drop down options allow the user to search by user name (“People”), by tag (“All Tags”), by all fields within the user’s bookmarks (“My Bookmarks”), or within the currently displayed set of bookmarks (“These Bookmarks”). For example, having restricted the currently visible bookmarks to Leon Berg’s “java” tag, one may search within that set for the occurrence of some arbitrary word or phrase in the description or title.

We have also built a Firefox plug-in that detects when a Google or intranet search has been done, sends the same search query to the *dogear* service, and integrates the *dogear* results into the web or intranet search results. This *search insertion* provides a simple form of social search that may help provide a more robust and useful enterprise search capability

Injected results	See 1213 results for "ajax"
Ajax : Ajax Mistakes Arturo Toscaninini (and 1 other person) on Dec 6, 2005 http://swik.net/Ajax/Ajax+Mistakes	
Whats Ajax? - Ajax Patterns Joni Mitchell (and 1 other person) on Dec 4, 2005 http://ajaxpatterns.org/Whats_Ajax#Core_Techniques_and_Technologies	
Ajax Tutorial: Ajax What Is It Good For? Ajax Tutorial XML Training and Ajax XML Application development tutorial and Ajax Javascript Mike Mulligan (and 1 other person) on Mar 7, 2006 http://dhtmlnirvana.com/ajax/ajax_tutorial/	
results	1 - 10 of about 226 for ajax
[PPT] AJAX toolkit ...class libraries Targets AJAX Application Developers and Toolkit Developers. ...inc personality Zimbra Rico and other Ajax tool kits. ...Live Grid Round corners Drag & Sample Button Sample Form AJAX req ...	
[PPT] AJAX toolkit ...The AJAX Toolkit Framework ATF provides and extensible framework and tools for building IDEs for the many different AJAX runtime offerings Dojo Z in the market. AJAX Toolkit Framework 3.7MB Eclipse Update Site ZIP. ...	

The top three search results are bookmarks from *dogear*. The search is a Lucene-based search of the bookmark contents (annotations, title, and tags).

The lower portion of the screen contains search results from the enterprise search engine, which is based on Omnifind search technology.

Figure 3. Example of search results.

A screen shot of the integrated search results for an enterprise search can be seen in Figure 3. At the top of the screen, labeled “injected results,” are the results from the *dogear* query, which have been dynamically inserted into the enterprise search results. The top three *dogear* results are presented. Below the *dogear* results, are the traditional results from the enterprise search.

Every time a search query is initiated, a log event (search) is recorded, and every time one of the *dogear* bookmarks is clicked (from the integrated search results) a log event (click) is created. We are able, therefore, to determine both the number of external searches initiated, and also the click-through rates for the *dogear* bookmarks.

Field Study Results

To answer our research questions, we performed a field study of an enterprise bookmarking service deployed in a large multi-national company. Our general understanding of the use of the enterprise bookmarking service was based on many sources of user data, including log files, the primary bookmark data files, and in-depth interviews with 15 end-users of the service. 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.

After a short friendly trial (~3 months), the system was introduced in mid-2005. By mid 2006, the system contains over 100,000 bookmarks and 250,000 tags created by over 1600 active end-users. The activity analysis presented here is based on log files covering a 12 month period from July, 2005 to July, 2006 (~450K records). The sampling period, therefore, can be considered the first year of system use or adoption.

While the individuals using the *dogear* bookmarking service all worked in a high-tech industry for a large enterprise, they are reasonably diverse in the kind of work that they did the organizations and geography in which they worked. Table 1 show the breakdown of users by organization and geography. Users were distributed across eight major organizations, with more than a quarter coming from the software development organization. While the user base is significantly based in the US, over half of the system users were from other countries.

Organization	Total	%		Geography	Total	%
Software	1701	28%		USA	2564	45%
Consulting	1043	17%		United Kingdom	506	9%
Sales	865	14%		Germany	415	7%
Tech Delivery	613	10%		Canada	356	6%
Systems	392	6%		Japan	311	6%
Services	387	6%		India	259	5%
Research	233	4%		Australia	173	3%
Headquarters	203	3%		China	94	2%
Other	702	11%		Other	966	18%

Table 1. Demographics of bookmarking service users.

Supporting various search activities

In this paper we are particularly interested in understanding how the bookmarking service supports various kind of exploratory search activities. We examined the logfiles for a 12 month sample (June 2005 - July 2006) to get a sense of how often end-users navigated through the bookmark collection using tags, names, and explicit search features. To better understand the results of the end-user actions, we have considered three general kinds of search, which are described in Table 2. These three kinds of search represent different user motivations and strategies. We will attempt to understand these end-user behaviors through a combination of logfile analysis and end-user interviews.

Community browsing	Examining bookmarks created by the community by time (<i>recent</i> bookmarks,), by frequency (most <i>popular</i> sites bookmarked), by <i>people</i> , and by <i>tags</i> .
Personal search	Looking for bookmarks from one's own personal collection of bookmarks.
Explicit search	Using a traditional search box to enter a set of search terms

Table 2 Different kinds of search supported.

The results of our logfile analysis are presented in Table 3. We group the results by the different kinds of search that we described above. We show the number of times that a particular navigation path resulted in a page view, i.e., a click-through to the original information resource that was bookmarked. Finally, we present the percentage of navigation events that resulted in a page view.

Action	# events	# clicks	% clicked
Community browsing			
Recent	77,132	15,621	20%
Popular	9,724	3,408	35%
People	26,535	5,074	25%
People x tag(s)	6,320	1,499	24%
Community Tag(s)	15,941	3,767	24%
Total	135,652	29,369	22%
Personal search			
Mybookmarks	57,236	9,949	17%
My tag(s)	20,980	6,387	30%
Total	78,216	16,336	21%
Explicit search			
Search- Internal dogear	29,961	11,667	39%
Search–plug-in enterprise	70,706	22,861	32%
Total	100,667	34,528	34%

Table 3 Logfile results for various search activities.

The page view percentage might be considered a search “completion” rate, and may be suggestive of different kinds of search strategies. For example, larger click-through percentages might indicate a more goal-oriented search. At minimum, the click-through percentages provide an estimate of the usefulness of the search results that were presented via each of the various search pathways.

Community browsing

The most frequent way to view information is using one of the community browsing mechanisms supported in dogear. The large total for community –related views (135K) confirms the value of the social aspects of the system. At the same time, the generally modest number of click-throughs (~22% of occasions on average) suggests that community browsing is used more for profiling information and other people than for known item retrieval.

Community Browsing: Recent and Popular Bookmarks

As can be seen in Table 3, the single most common way to view the dogear bookmark collection is to look at a listing of the most recently added bookmarks [Recent – 77132 above]. In some respects, this is not very surprising as there is a main dogear menu link labeled “all” and it is recommended in the application “help” to add a browser toolbar link to this dogear page.

Several dogear users that we interviewed indicated that they scan this list of newly created bookmarks with some regularity, much like they would scan news of blog feeds, exploiting the fact that dogear users form a shared community of interest within this organization. For example, one informant said:

The first thing I normally go to in Dogear is *All* because I'm interested in finding out what everyone is doing. And I equate Dogear to basically Slashdot for IBM plus everything else on the Internet. ... I think it's better than Slashdot. Before when Dogear was starting, Slashdot was ... a little superior because there was still a lot of people, more and more people out there finding a lot of interesting feeds. And now that Dogear has achieved critical mass, there are a lot of people within IBM who I know and I respect, they've been picking up lots of interesting things that Slashdot folks haven't been picking up on.

This same informant indicated that he had changed his scanning the *all* list over time as the volume and variability of the bookmark content changed.

I would scan through probably maybe three or four pages a day if I have time. ...when I started with Dogear, I would actually scan maybe more, but the volume was less. And it was a lot easier to pick -- sift through things. But now that there's more bookmarks today, and there's a lot of people from the other parts organization that I don't necessarily care about, I reduced the amount of pages I look at.

There were fewer reports from informants about the use if the “popular” or most active bookmarks, although one person indicated that he was aware of the top three bookmarked sites. There appear to be a relatively large percentage of page views on the popular bookmarks, indicating significant interest in what other people have already found. The click through results (35%) also suggests that, compared with recency, popularity is an effective recommendation about which resources might be worth reading.

Community Browsing: People Search

Of greater interest, perhaps, we note that there is considerable interest in browsing other peoples' bookmark collections. As can be seen in Table 3, the most frequent way to browse another user's bookmarks is by clicking on that person's name (26535). Once again, we also examined the number of times that end-users clicked through on a URL that had been bookmarked by another. The click-through results in Table 3 show that browsing someone else's bookmarks often results in following the URL to the original information source (25 percent of the time). This confirms the utility of social navigation in identifying useful informational resources.

These results suggest widespread curiosity about what others are bookmarking and provide evidence of the kind of explicit social navigation that is taking place within the bookmarking service. These results are significant as they represent a novel form of information browsing within the enterprise. Comments from end-user interviews confirm that the bookmarking service is supporting social navigation. One respondent indicated that she browsed the links of others for different reasons. In one case, it was to infer the current topics of interest for an organization thought leader:

It's usually I'm just looking at the people I know ... thought leaders and trying to get at ..., what is their thinking... And that might come up just through their tags or just through their content, or a combination.

This same kind of people browsing can be used to informally learn about the interest of coworkers. One dogear user said:

"the most value to me is finding people --- who have specific expertise, or experience, or interest ... more than anything it's helped me become familiar with people."

Another respondent indicated that navigating tags was a good way to informally find people with interest or expertise on various topics.

And it just seems like a fabulous way of either finding people who either might know about something or might have bookmarked something that I would be interested in...

This same respondent indicated that another significant benefit is being able to trust the information sources based on knowledge of the people in the group.

...there would be words that I would be thinking in my head, like key words. And if I saw them in a tag cloud, then I would click on them and ... all these things would show up that have potential of being what I'm looking for. And I would do that as opposed to going to Google because in some way, it's a somewhat trusted community already. ... I respect the people in this community and they probably know things that I would be interested in. And I would trust their sources.

Another respondent indicated that she was especially interested in what others with similar job roles were bookmarking.

"And generally,, on technical sales, a lot of the things that I'm looking for are going to be things that are other people that do what I do what I too am looking for. And so if I see a search that somebody else has done and somebody has bookmarked something, there's probably a better chance that that's been useful to me as well."

Community Browsing: Topic Search

Another popular way to browse other users' bookmarks is by selecting a specific tag from the system-wide tag cloud (Community browsing by tags – 15491). One respondent indicated that tag browsing was an efficient way to keep up-to-date on a particular subject.

And sometimes when I'm searching by topic, like if I'm working on – if I have a thought or an idea, or I'm writing something, let's say about attention management, I might want to see what's new on attention, and I'll just search that tag to see what's new, rather than searching my subscriptions or on social networking or task management.

Another informant described different using social searching to do learn about new topics.

So, an example is ...somebody that I work with was telling me that they have suddenly gotten interested in Second Life. Second Life is a virtual reality environment, and I knew that there was a bunch of people in [company] who were tracking it. So, I thought ... let me look and see what other people have found about Second Life because that is something that people who use Dogear are likely to have found.

The same informant described looking through other bookmarks tagged with “second life” until she can ” get an idea of what is this thing? What is this category?” She continues to search until she finds a reasonable description.

Now, in this particular case ... here is something that is a reasonable description. It gives you some idea and, oh, look, it's Business Week. Well, I can click on that because that is going to tell me in plain English what Second Life is.

So, there is some serendipity that occurs within that, but it is really because I am looking for the beginnings of information about something. I am not looking for deep research.

The most popular tags used for topic browsing include specific technologies (e.g., ajax, linux, javascript, and websphere) or emerging topic of interest (e.g., web2.0, wiki, blog). Other popular tag queries seem to be around general terms or broad categories of information (e.g., social-software, programming, architecture, and design). Of some interest is the fact that the tag *dogear* was the fifth most frequently browsed, which may have assisted in adoption and information diffusion about the new social bookmarking tool.

Personal Search

Overall, personal search is the least frequent way for people to access the system – accounting for many fewer accesses than community browsing or explicit search. This is an interesting finding because previous literature has suggested that a primary motivation for social tagging systems is that they provide users with ways to manage and browse their own information, whereas our data show that for dogear at least, personal search is less frequent than other forms of access. In particular this argues for the utility of the social/community features of the system.

As can be seen in Table 3, end-users also browse the bookmark space by looking at their own collection (mybookmarks - 57,236) and navigating through their collection using personal tags (my tags - 20,980). The use of personal tag browsing suggests that tags are serving as a useful way to filter bookmarks once users build up larger collections of bookmarks.

In Table 3, there is also evidence that end-users actually use the bookmarking service to “revisit” information sources that they have previously bookmarked. In fact, when end-users navigate their bookmarks using their personal tags, they click through 30% of the time. This finding is contrary to early research suggest-

ing that classic single-user web bookmarking is ineffective as users seldom turn to bookmarks they have created (Jones, et al., 2003).

Several of the end-users that we interviewed indicated that this support of personal bookmark management was an important characteristic of the service. One end-user highlighted the ability to centrally store bookmarks:

And I find a lot of personal value, or at least personal comfort, in realizing that I saved this stuff somewhere, and it's refreshable, and it's in that location.

Another end-user said that social tags were very helpful in managing the over-load of blog content:

Because I have a lot of blogs that I read and I have found that this (bookmarking service) is just a better way, a quicker way for me to organize them. I will click the word blog, the tag, and then I'll see all the blogs.

The specific need to re-find information was highlighted by another informant I bookmark anything that I think I am going to want to go back to. I think that's the general reason for bookmarks, and often I will, as I am browsing stuff, not bookmark it because I say, oh, I will be able to find it, and then I realize that I am looking for it one day, and I can't find it, so then I will go and bookmark it.

The most popular tags used by individuals to browse their personal bookmark collection were quite similar to those used to browse the community collection described above. Examples of tags used more often for personal bookmark search include: mine, education, learning, career and travel. These tags appear to categorize information resources of personal interest for career planning, training, or personal travel.

Explicit search of dogear bookmarks

Explicit search is overall used more frequently than personal search but less than community browsing. As described above, there is a general search capability provided within the dogear service. As can be seen in Table 3, the internal dogear search capability is the fourth most used way to “explore” the dogear collection (29,961 times). There is also an ability to search the *dogear* bookmarks as part of a more general enterprise search. The number of external search events shows that the external search plug-in is the second most used way to find bookmarks (70,706). Of particular note, is the high click-through rate (39% and 32%) for search actions, which indicates a more purposeful searching of the collection than for either personal or community browsing.

Several dogear users commented on the value of the dogear search that was integrated into the more general search tools (i.e. Google and enterprise search). One benefit was simply being reminded (inline) that related content has been bookmarked on dogear.

Another thing that happens is that even though the Google results are more specific generally, because it's more of a general search than dog ear, I may, depending who the person was that created the bookmark, I may tend to go look at it

Another informant indicated that she frequently did enterprise searches and then decided to redo the search in dogear.

Because a lot of times, I mean especially if it's an internal [enterprise] search, ..., you search for something and you're like, "Man, I can't find it." And then you just jump over to Dogear and you get it.

Other dogear end-users have indicated that an important reason for preferring the dogear results over generic search results was the fact that very inclusion in the dogear collection meant that the search had been actively pre-filtered by a community of interest or trusted IBM colleagues.

Patterns of Search

We were also interested in how patterns of search/browsing differed for different groups of end-users. In order to see whether there were differences in social bookmarking use as a function of bookmarking experience, we looked at the distribution of community versus personal search activities. We compared the distributions for experienced (or heavy) bookmakers, which we defined as individuals having more than 100 bookmarks, with medium bookmakers, which we defined as having between 10 and 100 bookmarks). The results can be seen in Figure 4.

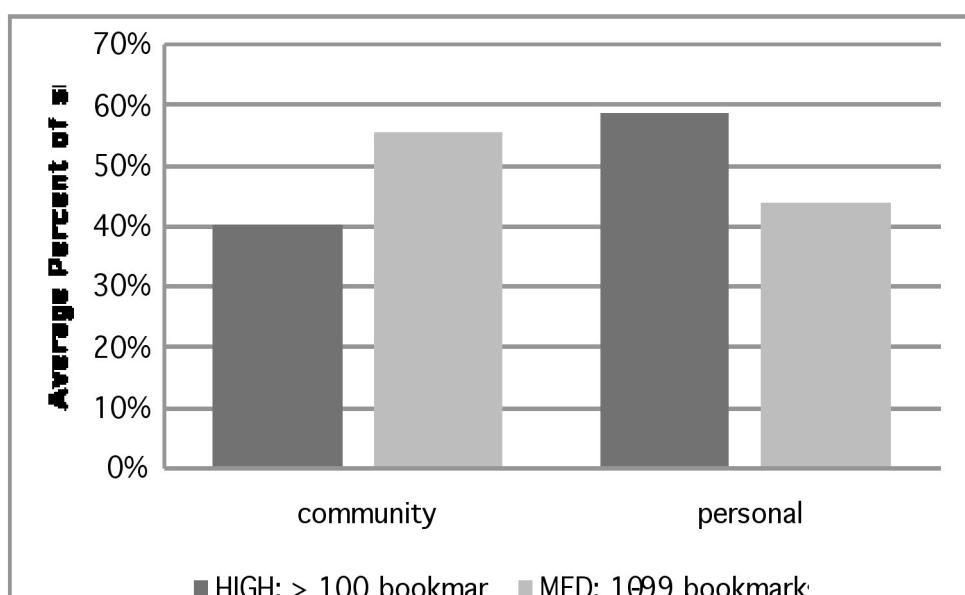


Figure 4. Type of navigation as a function of number of bookmarks.

The interesting result is that individuals with more bookmarks (>100 bookmarks) spend a significantly larger percentage of their browsing time looking at

their own collection of bookmarks. This is consistent with a view that prolific bookmarkers spend time, and presumably derive more value, refinding things that they have previously found. The medium bookmarkers (10-99 bookmarks) spend a larger portion of their time exploring other people's collections. This suggests that less experienced bookmarkers spend more of their search time in learning or discovery search activities.

We would expect to see a change in searching behavior to occur over time as individuals add more and more bookmarks to their personal collection. In particular, we expect that bookmark organization and management will become increasingly important user activities for heavy bookmarkers, which is consistent with earlier reported research on browser-based bookmarks (Abrams, et al., 1998).

Looking for Patterns using Cluster Analysis

As with studies of general web browsing and search ((Heer and Chi, 2002; Sellen, et al., 2002) we expected to observe different search patterns emerge in the use of a social bookmarking service. To better see the patterns of use, we performed a cluster analysis (K-means) for the different types of search activities. We first normalized the use data for each end-user by computing the percentage of each search type (i.e., community, person, topic, personal, and explicit search). The K-means cluster analysis is then performed, which finds related groups of users with similar distributions of the five search activities. The resulting cluster solution, shown in Table 4, shows the average percentage of item type for each of four easily interpretable clusters. To help understand these clusters, we have also provided for each cluster: the number of individuals in the cluster (N), the average number of bookmarks, and the average number of browsing events.

Cluster 1 is comprised of end-users with the largest percentage of *personal* browsing activity (60 %), coupled with the largest average number of bookmarks (155). This cluster is comprised of the kinds of informants cited above who claim to use a social bookmarking service heavily to store a personal bookmark collection and then to refind things that were already found. The individual who indicated that she used dogear to manage a collection of blogs would be a good exemplar of this cluster.

Cluster 2 has the largest number of members (851) and appears to be made up of individuals who are heavy browsers of the community collection (49 %) as well as individuals who spend the largest percentage of time using *explicit search* (23%). This group is an interesting mix. They spend a fair amount of time performing social (community) browsing as well as explicit searching to find information that has already been found and bookmarked. The interview respondent who scans the recent bookmark list, or individuals who use explicit search to locate bookmarks would be exemplars of this cluster.

Clusters 3 and 4 are made up of individuals with less experience using the bookmarking service, as indicated by the lower average number of log events for

these two groups (43 and 66 respectively), and fewer bookmarks (5 and 17 respectively). Cluster 3 has a heavy weighting on topic searches (46%) which is consistent with the informant above who described looking for information about “second life.” Cluster 4, with a heavy weighting on people queries (42%) is suggestive of the informants above who describe looking for information from known experts or organization leaders.

The small number of average bookmarks for clusters 3 and 4 suggests that the initial use of a social bookmarking application is spent performing more exploratory search tasks like learning what has been bookmarked for a topic (e.g., Ajax or Web 2.0) or looking at what a particular colleague has been bookmarking.

	Cluster			
	1	2	3	4
Community	.20	.49	.28	.26
People	.08	.08	.08	.42
Topic	.04	.06	.46	.07
Personal	.60	.13	.07	.13
Explicit Search	.07	.23	.10	.13
N	466	851	238	317
Avg. bookmarks	155	34	5	17
Avg. # events	190	124	43	66

Table 4. Results of cluster analysis.

Discussion

In this paper we have investigated the use of a social bookmarking application within a large enterprise. We have described and examined the design elements that support different kinds of browsing or search activities, including social tags, people browsing, and multiple search interfaces (i.e. both internal dogear search and through the Firefox plug-in). The results of our log file analysis confirm that these navigational elements are used by end-users of the social bookmarking service, while interview comments provide support that both personal and social navigation benefits are being realized. The quantitative and qualitative results presented in this paper show that social bookmarking systems support a variety of exploratory search activities, which help satisfy end-users’ learning and investigative information needs.

The results above show that community browsing is the most frequent way to search/browse for information. Community browsing includes looking at the community collection, looking at co-worker's collections and browsing by topics defined by social tags. This kind of social browsing may be due, in large part, to one of the unique characteristics of enterprise social bookmarking services. End-users of these kinds of services are required to use corporate identities (i.e. no pseudonyms allowed). This enables others in the enterprise to easily recognize a coworker in a list of people tags (see Fig 2(b)) or as a creator of a specific bookmark. Once recognized, it is easy to learn from organizational thought leaders or to learn about the interest of new team members. Furthermore, coworkers provide an important selection and filtering function for information content.

The interview comments indicated that a large portion of this community browsing is in support of what Marchionini would call *learning* and *investigative* search activities (Marchionini, 2006). For example, individuals mentioned following the name links in dogear to learn more about someone and to find out what others are interested in. Interviews also revealed cases in which dogear users followed a tag (or multiple tags) to learn about a particular technology. This is important because it helps promote informal discovery of information and learning about people. This exploration of other people's bookmarks is a very promising way to support lightweight information sharing or knowledge management within an organization.

The second most frequent form of search that we observed was explicit search. The relatively high proportion of click-throughs for search tasks, combined with interview comments, suggest that social bookmarking services provide a good way to capture high value pointers to information sources. This suggests that social bookmarking services, integrated with traditional search engines, have the potential to solve the tough enterprise search problem (Mukherjee and Mao, 2004). The explicit search capabilities appear to be used for what Marchionini refers to as *look-up* and *learning* activities.

The results of this study suggest that social bookmarking services provide significant benefits for managing personal bookmarks. There is substantial browsing of personal bookmarks, which often result in revisits to the original content source, suggesting the superiority of our system over more traditional bookmarking methods. The interview comments suggest that much of the use of personal bookmarks is to refind information, which would support various *look-up* search activities, as described by Marchionini (2006). The cluster results show that personal browsing is most often associated with a larger collection of personal bookmarks.

The results presented here are part of a growing body of work underway to understand how various kinds of collaborative tools can help with exploratory search (White, et al., 2006). There are a number of challenges in finding ways to integrate social interaction with exploratory search. Furthermore, there are a

number of measurement and evaluation challenges that have been identified for these kinds of exploratory search tools (White, et al., 2006).

Design implications and future work

Most social bookmarking services (e.g., del.icio.us) provide a set of common core features. Most provide community browsing by bookmark frequency or most recent bookmarks, as well as pivot queries by tags and names. In addition, all provide application specific search of the bookmark database. We would argue that similar results would be found in a usage level analysis of these systems.

An enterprise-grade social bookmarking service, such as *dogear*, requires the use and authentication of a real name identity. As described above, this allows easy integration of other corporate collaboration tools within the bookmarking service. This is important in that community browsing can be automatically augmented by links to corporate directories, blogs and other information sources. It is possible that community searching will be a richer experience in a corporate bookmarking system.

An enterprise bookmarking service is also different from many bookmarking services found on the web in that the user population comes from a trusted community, and is often bound by an explicit corporate code of conduct. This may change the kinds of social interaction on the site, and very likely change the content of the bookmark collection.

There are a number of implications for the future design of *dogear* (or similar bookmarking services). It seems that although people prefer to browse by recency (i.e. view recent), they actually find more useful information using popularity (as indicated by click-through results). So the interface should be redesigned to respect this. Similarly, for refinding personal resources the “mybookmarks” user interface could better support tag searching by providing a list of recent search terms or most used tags for each individual. And finally, people browsing could be enhanced with group or *friends* support in the user interface. It is easy to imagine articulated social networks supported in this kind of service. The social network could be used for general bookmark displays or bookmark and tag recommendations.

There remain several interesting research questions for social bookmarking applications. First, it is important to better understand how collaborative tagging and social bookmarking can be more tightly integrated with enterprise search. Second, while initial work has been done to understand tag and folksonomy development, there are a number of interesting questions about how to optimize these vocabularies to support various browsing tasks. And finally, there are a number of questions about how these kinds of social software applications are adopted and how communities develop. Social network analytic techniques have been explored and show promise as a way to understand this phenomenon.

References

- Abrams, D., Baecker, R. and Chignell, M. (1998): 'Information archiving with bookmarks: personal Web space construction and organization'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Los Angeles, California, USA, April, 1998, pp. 41-48.
- Damianos, L. E., Cuomo, D., Griffith, J., Hirst, D. M. and Smallwood, J. (2007): 'Exploring the Adoption, Utility, and Social Influences of Social Bookmarking in a Corporate Environment'. In *40th Annual Hawaii International Conference on System Sciences* Waikoloa, HI, USA January, 2007, pp. 86-86.
- Dieberger, A. a. G., M. (2003): 'CoWeb - experiences with collaborative Web spaces.' In C. a. F. Lueg, D. : *From Usenet to CoWebs: Interacting with Social Information Spaces*. Springer-Verlag, New York, 2003, 155--166.
- Golder, S. A. and Huberman, B. A. (2006): 'Usage patterns of collaborative tagging systems', *Journal of Information Science*, 32, 2, 2006, pp. 198-208.
- Hammond, T., Hannay, T., Lund, B. and Scott, J. (2005): 'Social bookmarking tools (I) A General Review', *D-Lib Magazine*, 2005, ISSN: 1082-9873.
- Heer, J. and Chi, E. H. (2002): 'Separating the swarm: categorization methods for user sessions on the web'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Minneapolis, Minnesota, USA, April, 2002, pp. 243--250.
- Jones, W., Bruce, H. and Dumais, S. (2003): 'How do people get back to information on the web? How can they do it better'. In *9th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT 2003)* Zurich, Switzerland, Sept, 2003, 793-796.
- Kittur, A., Chi, E., Pendleton, B., Suh, B. and Mytkowicz, T. (2007): 'Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie.' In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* San Jose, CA, USA, April, 2007, pp. 453--462.
- Marchionini, G. (2006): 'Exploratory search: from finding to understanding', *Communications of the ACM*, 49, 4, 41-46.
- Marlow, C., Naaman, M., Boyd, D. and Davis, M. (2006): 'HT06, tagging paper, taxonomy, Flickr, academic article, to read'. In *Proceedings of the seventeenth conference on Hypertext and hypermedia* Odense, Denmark, August, 2006, pp. 31-39.
- Millen, D. R., Feinberg, J. and Kerr, B. (2006): 'Dogear: Social bookmarking in the enterprise'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Montreal, Quebec, Canada, April, 2006, pp. 111-120.
- Mukherjee, R. and Mao, J. (2004): 'Enterprise Search: Tough Stuff', *Queue*, 2, 2, April, 2004, 36-46.
- Munro, A. J., Hook, K., and Benyon (1999): *Social Navigation of Information Space*. Springer, London, 1999.
- Sellen, A. J., Murphy, R. and Shaw, K. L. (2002): 'How knowledge workers use the web'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Minneapolis, Minnesota, USA, April, 2002, pp. 227-234.
- Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F. M. and Riedl, J. (2006): 'tagging, communities, vocabulary, evolution'. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* Banff, Alberta, Canada, November, 2006, pp. 181-190.
- White, R. W., Kules, B., Drucker, S. M. and Schraefel, m. c. (2006): 'Introduction SPECIAL ISSUE: Supporting exploratory search ', *Communications of the ACM*, 49, 4, 36-39.

White, R. W., Muresan, G. and Marchionini, G. (2006): 'Report on ACM SIGIR 2006 workshop on evaluating exploratory search systems', *SIGIR Forum*, 40, 2, 52-60.