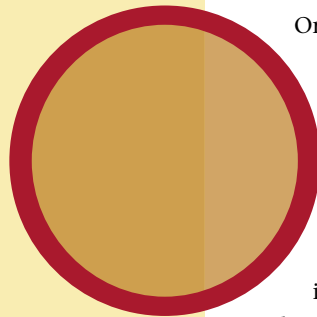


DAVID MILLEN, JONATHAN FEINBERG, and BERNARD KERR, IBM

social bookmarking in the Enterprise

Can your organization benefit from social bookmarking tools?



One of the greatest challenges facing people who use large information spaces is to remember and retrieve items that they have previously found and thought to be interesting. One approach to this problem is to allow individuals to save particular search strings to re-create the search in the future. Another approach has been to allow people to create personal collections of material—for example, the use of electronic citation bundles (called *binders*) in the ACM Digital Library. Collections of citations can be created manually by readers or through execution of (and alerting to) a saved search.

Perhaps the most familiar approach to “refinding” information on the Web has been through the use of personal bookmarks, supported by various Web browsers. For example, the Mozilla browser supports the creation of collections of URLs, which can be annotated using keywords or free-form text, and then sorted on a variety of dimensions (e.g., time last visited, keyword, location). An early study of bookmark use showed that people created bookmarks based on the quality and personal interest of the content, high frequency of current use, and a sense of potential for future use.¹ Furthermore,

social bookmarking in the Enterprise

the number of bookmarks contained in an individual collection grew steadily and roughly linearly, and the use of folders to categorize bookmarks increased as the size of the collection increased. A single level of folders was reported for collections with fewer than 300 bookmarks, whereas larger collections prompted multitiered hierarchies.

The desire to find and share information among small groups, teams, and communities of practice has, not surprisingly, led to the development of a number of shared bookmarking systems. Early systems often used automated techniques to support the creation and categorization of collections of Web bookmarks.^{2,3,4} These innovative systems met with some success, although they seem to fall consistently short of their potential. Their shortcomings include limited privacy protection, little support for end-user customization, and high requirements for active user participation.⁵

SOCIAL BOOKMARKING SYSTEMS

The recent emergence of shared bookmarking applications has prompted a second look at this kind of collaborative software. Introduced in 2003, the *del.icio.us* social bookmark manager was one of the first of its type and has enjoyed an early and large base of committed users.⁶ A flurry of similar offerings have since been unveiled, including a recent beta release of Yahoo's My Web 2.0.⁷

These social bookmarking systems share a number of common features. First, they allow individuals to create personal collections of bookmarks and easily share their bookmarks with others. These centrally stored collections can be accessed from any Web-connected 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 the collection with meaningful labels. 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 final distinguishing characteristic of these social bookmark applications is the social nature of their use. There is a bias toward increased transparency in these tools. Although bookmark collections are personally

created and maintained, they are typically visible to others. A number of user interface elements allow social browsing of the bookmark space. For example, user names are "clickable" links; clicking on a name reveals the bookmark collection for that user. This allows someone to get a sense of the topics of interest for a particular user. Similarly, 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 if it includes information sources of interest. The ability to reorient the view by clicking on tags or user names, called *pivot browsing*, provides a lightweight mechanism to navigate the aggregated bookmark collection.

ENTERPRISE-WIDE SOCIAL BOOKMARKING

The apparent success of Internet-based social bookmarking applications begs the question of whether large enterprises or organizations would also benefit from social bookmarking systems. To investigate this question, at IBM we are designing and developing an enterprise-scale social bookmarking system called dogear. The rest of this article describes the design challenges and early lessons learned from a friendly trial of the technology.

IDENTITY AND TRANSPARENCY

The first significant design decision was whether to base user identity in the application on real names or pseudonyms. We decided to require real-world identity for the following reasons. First, one of the expected benefits of the system is to allow users to make inferences about the interests and expertise of others based on informal browsing of bookmark collections. This discovery of people with shared (or complementary) interests would help to nurture the communities of practice within the enterprise and potentially allow searching and finding experts on specific topics to help solve business problems. Real-name identity would allow users to look up additional information about other users in various corporate databases (e.g., corporate online directory and enterprise Web). In addition, real-name identity would facilitate communication among users of the application since the various corporate collaboration tools (e.g., corporate directories, e-mail, chat,) all use real-name identities. There is also a strong cultural norm

Extensibility by Design in Dogear

In designing dogear, we have considered both the Web site user and the technically sophisticated user who wishes to exploit dogear with computer programs. The needs of both can be addressed simultaneously by adhering as closely as is practical to the architectural style that Roy Fielding dubbed REST (representational state transfer).¹ For the user, the REST style is characterized by transparent, meaningful URLs. In turn, the URL structure suggests an implementation that has proven to be flexible, maintainable, and responsive. Our ability to play nicely with other research and production systems on the intranet has led to early adoption by a number of people and organizations within IBM.

DOGEAR: AN EXAMPLE OF REST

REST specifies that a given resource be represented by a stable URI (uniform resource identifier), such that the same URI will *mean the same thing* over time. This implies that the current state of an application must depend entirely on the URI, and not on opaque, hidden conditions such as cookies. The dogear application state consists of a set of constraints on which bookmarks are currently visible, an output format, a current page number, and the number of links per page.

Each of the machine-readable formats is extended, via XML namespaces, to include dogear-specific meta-data.

Programmatic read/write access to one's own bookmarks is provided by an HTTP API, which resides at /api. Bookmarks are created and modified via HTTP POST using simple name=value query parameters to provide bookmark title, URL, and tags. Bookmarks may be deleted by calling /api?url=url with HTTP DELETE.

QUERY CONSTRAINTS

The results may be constrained through the use of HTTP GET query parameters. Any combination of a person, a tag constraint, and a text search may be specified. The person constraint takes the form of user=userid and restricts the result set to those bookmarks created by the specified person. There are three tag constraints: tag, any, and all. The query tag=foo restricts results to those bookmarks tagged with foo; any=foo+bar+baz will accept bookmarks bearing any of the foo, bar, or baz tags; all=foo+bar+baz permits only those bookmarks bearing all three of the specified tags. Bookmarks are returned in reverse chronological order, except in the presence of a searchText parameter, which causes results

to be ordered by result relevance. In the absence of any constraints, the 10 most recent bookmarks are returned.

OTHER STATE

All queries may include either or both of the page and ps parameters, which specify the results page and page size, respectively.

How Dogear Encodes Application State in URLs

`http://dogear.webahead.ibm.com/html?user=jdf@us.ibm.com&page=2`

output
format

constraints

other
state

`http://dogear.webahead.ibm.com/rss?tag=java&user=jonesn@uk.ibm.com`

OUTPUT FORMAT

The first element of a dogear URL path selects the desired output format. Currently implemented formats include:

- * /html—the Web site itself
- * /rss—RSS 2.0²
- * /atom—Atom 1.0³
- * js—a JavaScript snippet, which, when evaluated, causes HTML to be emitted on the including page
- * /xbel—the XML Bookmark Exchange Language⁴

REFERENCES

1. Fielding, R. T. 2000. Architectural styles and the design of network-based software architectures. Doctoral dissertation, University of California, Irvine.
2. RSS 2.0 specification; <http://blogs.law.harvard.edu/tech/rss>.
3. Atom Syndication Format; <http://atompub.org/2005/07/11/draft-ietf-atompub-format-10.html>.
4. XML Bookmark Exchange Language; <http://pyxml.sourceforge.net/topics/xbel/>.

social bookmarking **in the Enterprise**

within the organization to use more formal names within corporate applications.

Note that although creating personal bookmark collections requires authentication of real-name identity, it is possible to browse the enterprise bookmark collection anonymously. This invisible participation is a form of “lurking” and allows the benefits of the bookmark sharing to extend to the entire enterprise. Lurking behavior is counted in application statistics (e.g., the number of times a bookmark is clicked to access a particular information resource). While lurking in online environments has been characterized as a free ride, it has been shown to provide an important supportive role in the sociability of online environments.⁸

Another important design decision was the degree of transparency to support within the application. Although early public Internet social bookmark applications (e.g., del.icio.us) provided open access to everyone’s bookmarks, there are valid reasons to limit access to some sets of bookmarks. For example, a group may want to create a collection of information sources in support of a project that may be extremely confidential or bound by contract to restricted access. Teams may want to form shared bookmarks that are visible only to the group.

We could also see a need for role- or job-based collections of information sources. For example, managers might benefit from bookmarking and sharing certain information sources, while the sales force would benefit from another.

Finally, we see a need for private bookmarks. An informal survey of potential users of the system indicated that many people had bookmarks that they considered personal and would not want to make public. We did not want to force people to use two bookmarking tools: one for personal bookmarking (e.g., in a browser) and a corporate tool for shared sources. Our initial implementation includes both private and enterprise (public) bookmarks. We have architected solutions to both team- and role-based collections of bookmarks and will roll out these capabilities in the future.

The decisions to use real-name identity and to support private bookmarks were made with some trepidation. The success of social software applications, in general, requires participation to reach critical mass to provide value to users and to ensure a sustainable level of

contribution and vibrant interaction. Real-name identity may discourage some people from using the system, and private bookmarks will significantly reduce the benefits of information sharing among users. For a large enterprise, however, we believe that the ability to reach critical mass will not present a problem.

ALERTING AND DISCOVERY

To encourage ongoing use of a social bookmarking application, users need to be alerted to relevant new information. Indeed, various notification mechanisms have been shown to contribute significantly to sustaining online interaction in other collaboration applications.⁹ We envision several ways to keep users informed of significant new content.

Like the del.icio.us social bookmarks manager, dogear provides an RSS feed on every page. This allows users to subscribe to collections within general bookmark collections, such as the most frequent or most recent bookmarks. In addition, users may subscribe to a bookmark collection belonging to a particular individual, or to a particular tag or keyword (e.g., eclipse). Users may then run a standard RSS reader to monitor new information in a particular collection of bookmarks.

Dogear also exploits collaborative filtering techniques

Dogear Prototype



FIG 1

to screen new bookmarks for those that are predictably of interest to an individual (or a group of individuals). Common interests can be inferred based on a number of observable user actions, including use of similar tags and/or tag combinations, similarity of bookmark (URL) collections, common RSS subscriptions, and click streams that indicate interest in specific kinds of bookmarks. Text analysis of bookmark titles, descriptions, and comments will also be used to determine bookmark relatedness. E-mail notification, system tray status indicators, or small pop-up windows can alert users to new information.

DESIGNING FOR EXTENSIBILITY: ENTERPRISE REMIXING

The design of our system encourages integration—remixing—by exposing its data via a simple and intuitive URL structure and by giving potential exploiters a number of easily parsed data formats from which to choose. For a more detailed description of this approach, see the sidebar, “Extensibility by Design in Dogear.” These URLs are constructed, and their output parsed and transformed, easily enough to permit rapid creation and deployment of useful services that retrieve, aggregate, and represent our data with a minimum of programming.

EXPLOITING THE LINKS

One of the primary benefits of an enterprise social bookmarking application is that it is an explicit

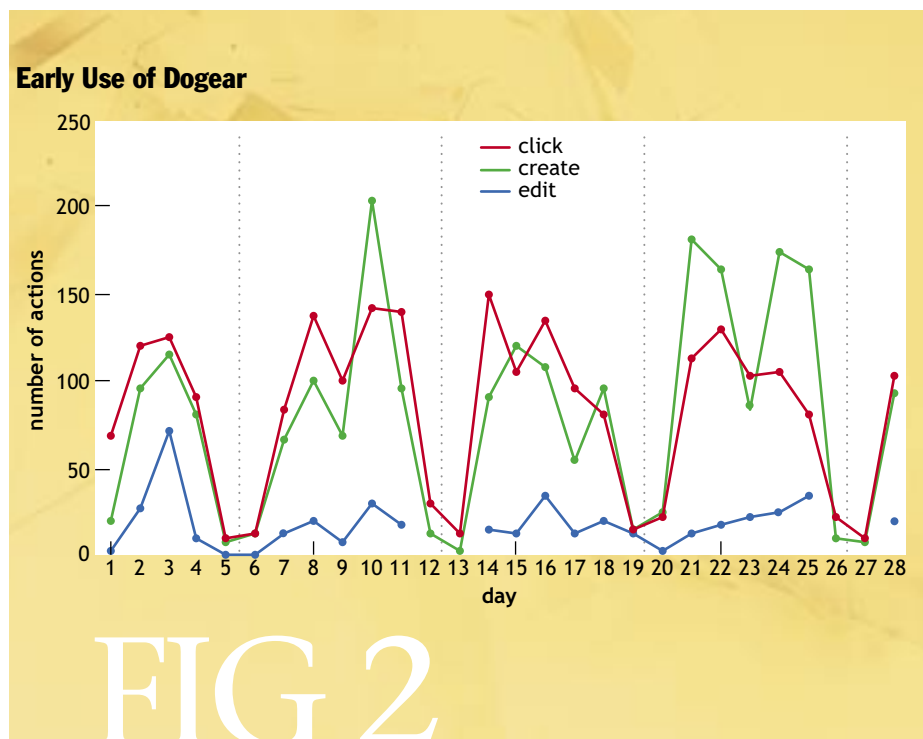
assessment of the utility or value of various Internet and intranet information resources. As an information resource, there are many potential integration points with other corporate applications. Several of these are discussed here.

The first natural integration point would be to use the social bookmark link structure to augment enterprise search applications. Web resources with high bookmark counts are likely to be relevant and useful. Indices of role- or team-based bookmark collections can be specially weighted in role- and team-based search tools. An early variant of this search integration has already been developed as a Firefox plug-in. Intranet searches are augmented with a supplemental search of the bookmark collection, and bookmark matches are presented, as a group, in the search results list.

A second integration would be to combine an individual’s bookmarks with other information sources provided by that individual. A research prototype of an enhanced corporate directory service has already embedded portions of an individual’s bookmark collection in directory information. The idea is that an individual’s bookmark collection is a meaningful reflection of the topics of general interest to that individual. This may be a powerful mechanism to provide dynamic updates to the “current interests” profiles for individuals. In a similar manner, corporate bloggers have

embedded recent personal bookmark collections in their home blog pages.

A third natural integration point would be to place bookmark collections from a group of individuals into Web sites maintained for various teams or projects. We have already observed one team, which is focused on e-learning, embed a collection of bookmarks on a team Web site. Again, using the subscription services supported by the bookmarking application allows automatic updating of bookmarks, promoting timely sharing of important information among team members.



social bookmarking **in the Enterprise**

CASE STUDY: DOGEAR SOCIAL BOOKMARKING SERVICE

We have developed a research prototype to investigate the usefulness of a social bookmarking application for a large enterprise. Figure 1 shows an individual's page in dogear, featuring the bookmarks the individual has most recently added. The tabs at the top of the left-hand column (A) link to the user's *tags*, a list of people who have bookmarked the same URLs, and a list of the individual's own subscriptions. The Active Tags (B) area shows an indexed *tag cloud* with a slider that can reduce the number of tags shown based on the ones that are used more frequently.

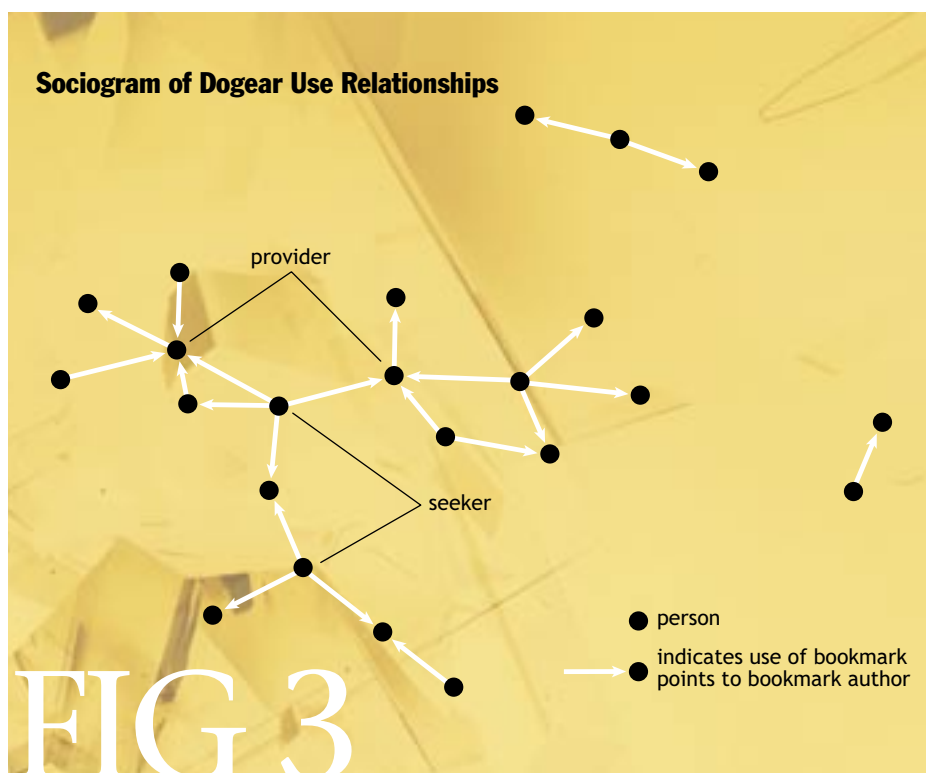
A list of bookmarks runs down the center of the screen (C). Each bookmark includes a title, optional descriptive text, tags, when the bookmark was made, and information about the author. Clicking on the title takes you to the bookmarked site. Clicking on a tag takes you to a list of the other bookmarks 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. You can copy any bookmark and leave a blog-like comment when you click on Comments. Dogear's robust search lets you find bookmarks by tag (D). Any dogear page can be viewed as an RSS or Atom feed and can be embedded as a blogroll via the JavaScript snippet on any other Web page (E).

EARLY RESULTS

The dogear prototype has been in a friendly field trial within the corporate intranet, and several modest announcements about the service availability have been made. We have begun to assess the usage patterns of the early adopters of the dogear application based on log files, and a more

formal user survey and interviews will be completed soon. Figure 2 shows some results from the four weeks of early use. An expected dip in use occurs on weekends, but the general pattern shows a sustained number of bookmark additions (CREATES) with a slight increase in use over time. Also visible is evidence that people are following the bookmarks to the information source (shown as CLICKS in the figure). During this usage period, more than 300 people were recorded using the application, with 42 percent creating bookmarks and 57 percent clicking through to the Web page. We see very little bookmark editing, which would be evidence of bookmark management activities such as refinement of tags or keywords, or additional annotations to the bookmark. This is not surprising as we are early in the trial; others have shown that bookmark management increases as the size of a collection increases.¹⁰

We have also begun to look for patterns of shared interest in various information sources. We have used social network analytical methods to begin to understand the information affinities among dogear users.¹¹ Figure 3



is a sociogram showing which individuals have clicked through to another person's bookmark reference. Clusters of individuals are already visible, as are those individuals who are most likely to be looking for information (seekers) and those who are information providers. We have begun to consider the ways we can use this kind of data in helping to share information among larger groups of people within the organization.

SHOWING GREAT POTENTIAL

Although we have just begun to get feedback on the dogear social bookmarking service, it shows great potential. A significant group of very active early users have already begun to generate excitement about the application through informal communication mechanisms. Indeed, there are already more than 100 mentions of the dogear application in the enterprise blogs. We have already seen a handful of dogear extensions by other members of the organization, which will only serve to increase the benefits of the application. We are excited by the potential for using the dogear application to improve information sharing, expertise location, and support of communities of interest within the enterprise.

Work is already under way to enhance the dogear social bookmarking service in several ways. As the number of shared bookmarks grows, so will the challenge to help users find new and particularly relevant content. We have begun to explore various search enhancements, as well as novel forms of customizable alerting. We have also begun to explore ways to provide better support for teams and communities of practice. These smaller groups may help us to understand what factors are important to scale down the dogear service for small and medium-size businesses. ☺

REFERENCES

1. Abrams, D., Baecker, R., and Chignell, M. 1998. Information archiving with bookmarks: Personal Web space construction and organization. *CHI*, Los Angeles: 41-48.
2. Marais, H., and Bharat, K. 1997. Supporting cooperative and personal surfing with a desktop assistant. *Proceedings of the Tenth Annual ACM Symposium on User Interface Software and Technology*, Banff, Alberta, Canada (October 14-17): 129-138. New York, NY: ACM Press.
3. Wittenburg, K., Das, D., Hill, W. C., and Stead, L. 1995. Group asynchronous browsing on the World Wide Web. *Proceedings of the Fourth International World Wide Web Conference*, Boston, MA.
4. Keller, R. M., Wolfe, S. R., Chen, J. R., Rabinowitz, J. L., and Mathe, N. 1997. A bookmarking service for organizing and sharing URLs. *Computer Networks and ISDN Systems* 29 (September): 1103-1114.
5. Kanawati, R., and Malek, M. 2002. A multi-agent system for collaborative bookmarking. *AAMAS*, Bologna, Italy: 1137-1138.
6. Del.icio.us; <http://del.icio.us/>.
7. My Web 2.0 beta; <http://myweb2.search.yahoo.com/>.
8. Nonnecke, B., and Preece, J. 2000. Lurker demographics: Counting the silent. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands, April 1-6): 73-80. New York, NY: ACM Press.
9. Millen, D., and Patterson, J. 2002. Stimulating social engagement in a community network. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work*. New Orleans, LA (November 18-20).
10. See reference 1.
11. Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. New York and Cambridge, England: Cambridge University Press.

LOVE IT, HATE IT? LET US KNOW

feedback@acmqueue.com or www.acmqueue.com/forums

DAVID MILLEN is a research manager in the Collaborative User Experience group at IBM Research in Cambridge, MA. Through field studies and development of new social software applications, his group studies the social and technological implications of online communities and large-scale collaboration. Millen holds a Ph.D. in cognitive psychology from Rutgers University and has held research and development positions at AT&T Research/Bell Laboratories.

JONATHAN FEINBERG is an advisory engineer in the Collaborative User Experience group at IBM Research, which he joined in September 2003. He considers himself a generalist, with particular interest in "glue programming"—tying together disparate and otherwise mutually agnostic systems. He holds a B.A. in English and American literature from Brown University.

BERNARD KERR is a senior design researcher in the Collaborative User Experience group at IBM Research. His current research focuses on designing and building advanced concept software applications and visualizations for collaboration. Kerr graduated with a M.A. in computer-related design from the Royal College of Art in London, UK. Prior to coming to IBM he held research and design positions at Interval Research Corporation and IDEO Product Development.

© 2005 ACM 1542-7730/05/1100 \$5.00