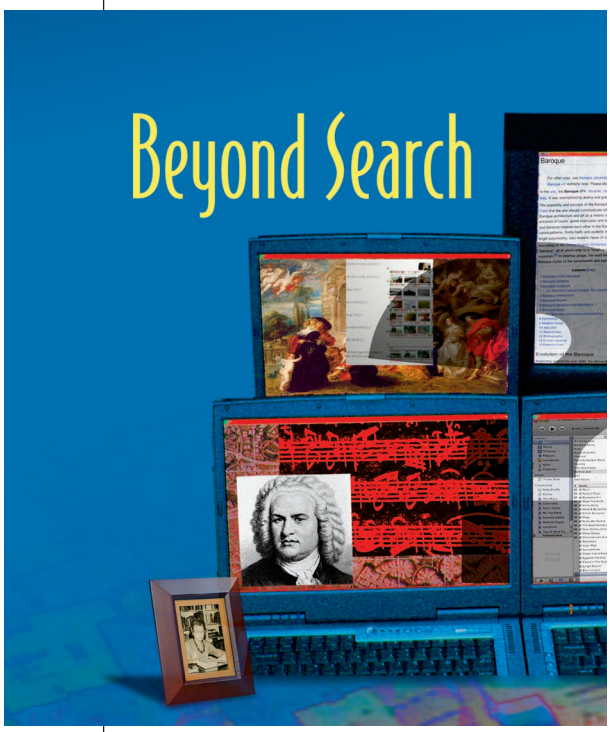


COLLABORATIVE INFORMATION SEEKING

Gene Golovchinsky, Pernilla Qvarfordt, and Jeremy Pickens, *FX Palo Alto Laboratory*

An examination of the roles and dimensions of collaborative search reveals new opportunities for information-seeking support tools.

People often work together when searching for information, but traditional information-seeking systems do not support such behavior, causing people to adopt workarounds such as over-the-shoulder surfing or sending links by e-mail to compensate for a lack of adequate tools.[1] This observation has led several research groups to explore various aspects of collaboration in support of information seeking.[2-4] The goal of this research is to create tools that support collaborative search based on explicitly shared information needs.

INFORMATION-SEEKING CHALLENGES

Collaboration and information seeking can assume many forms. The needs and search activities of John and Gail described here illustrate some of the challenges for collaborative information seeking.

Medical information

John is a middle-aged accountant who recently underwent successful surgery for colon cancer. He knows that he can have a long life as a cancer survivor as long as he maintains a healthy lifestyle. He continues to learn about the disease, joins online support groups, and seeks tools that allow him to find and organize information about colon cancer. As part of his support group, he offers encouragement and advice to others.

John joins an online community to identify a group of peers—fellow cancer survivors, researchers, doctors, and other motivated and knowledgeable people—with whom he can coordinate his ongoing investigations. Searching independently, they form a trusted search community[5] that makes automatically generated recommendations more coherent since the community has similar information needs. They also form synchronous collaborations in small groups to explore and share information about new developments and track ongoing research in the field.

John is lobbying his insurance company and hospital to expand their post-treatment counseling services to include online search help from registered nurses who specialize

0018-9162/09/$25.00 © 2009 IEEE Published by the IEEE Computer Society

in particular diseases. Working through a dedicated portal, John would be able to get feedback from his RN counselor on the results he is finding on his own and with his peers. The RN would have access to John's medical record (the details of which might not be available to John in a useful form); the RN could use this record to identify information relevant to his condition.

John would like a tool to manage the information set related to his cancer, both to maintain a record of what he has learned and to publish a subset of this information—along with his commentary—to help his family and friends understand his condition and perhaps draw them into his circle of collaborators.

Collaborative research

Gail is a neuroscientist who leads a multicampus team working on new techniques for using viruses as vectors for neuronal therapies. Gail also consults for a large biotech company. She scans a dozen journals regularly, participates in online forums, and has a postdoctoral fellow who conducts systematic searches of the biomedical literature. Gail's research group maintains a database of papers, and she spends much of her time refinding and integrating new ideas she gains from preprints, discussions with colleagues, and results from her lab work into the database. She needs tools that integrate search activities across diverse collections of public and personal information streams and allow flexible organization and analysis.

We classify information-seeking systems along four dimensions: intent, depth of mediation, concurrency, and location.

Gail often relies on her clinical colleagues to identify and interpret newly published research results and directs less experienced colleagues to search on topics they did not initially consider. She often wishes that they could share particular search sessions as well as the summaries she receives. Sometimes in meetings, Gail and her colleagues discuss particular searches and run additional searches during the discussion.

Gail accesses the biotech company's library of technical reports and preprints through a secure Web portal. While this is a useful resource, she is frequently confused by the organization of the subcollections and the company-specific terminology and metadata. Having online access to a reference librarian during search sessions would make her more productive.

We use these scenarios to consider how information-seeking support tools can enhance collaboration.

FORMS OF COLLABORATION

In information-seeking literature, the term collaboration refers to a variety of forms of mediation, communication, and coordination[6] that might not have as much in common as the name suggests. To distinguish among the various forms of computer-supported collaboration for information seeking, we classify such systems along four dimensions: intent, depth of mediation, concurrency, and location.[7]

Intent: Explicit versus implicit

Explicit information seeking occurs when two or more people set out to find information on a topic based on a declared understanding of the information need, which might evolve over time. In our scenarios, explicit collaboration occurs when Gail interacts with clinical colleagues or works with the librarian, and when John receives help from the RN. In contrast, implicit intent characterizes collaborative filtering systems that infer similar information needs based on users' actions or opinions. Recommendations such as "People who bought this product also bought …"[8] and responses to queries such as "What are my coworkers reading that I also need to be aware of?"—which use aggregate statistics to make suggestions—are examples of implicit collaboration. Such recommendations are useful only if others previously indicated similar information needs, as in the case of John and his online community. The "Better Exploratory Search through Social Bookmarking" sidebar provides a discussion of implicit search.

Depth of mediation

Depth of mediation is the level at which collaboration occurs in the system—user interface versus search engine back end. The level of mediation affects how aware a system is of the contributions of different people and how it uses those contributions to influence searches. Collaborative filtering systems, for example, keep track of each user's data separately, before aggregating it to make specific recommendations. Cerchiamo[4] uses collaborators' relevance judgments to influence the ranking of search results. Systems such as SearchTogether,[3] on the other hand, distinguish among users only in the user interface; the search engine component is unaware that multiple people have contributed queries, saved documents, and so forth.

Concurrency

People can collaborate synchronously or asynchronously. Synchronous collaboration implies the ability of people to influence each other in real time, as Gail does when she meets with her colleagues to search and discuss the results. Asynchronous collaboration describes situations where previous searches, either personal or aggregated from a community, influence later searches, as when John receives recommendations based on results found by others in his peer group. Systems can support

BETTER EXPLORATORY SEARCH THROUGH SOCIAL BOOKMARKING

David Millen, *IBM T.J. Watson Research Center*

An ongoing challenge facing workers in large organizations is to find valuable information and then to refind that information when it is needed again. Equally challenging is finding effective ways to share the information with colleagues who have similar needs and interests. Increasingly, large enterprises are experimenting with social bookmarking systems to support workers in these challenging search and sharing tasks.

One early tool developed for enterprise use was the dogear social bookmarking system.[1] Developed and initially deployed at IBM, workers could bookmark important information resources and easily annotate each bookmark with keywords or social tags, creating a personal folksonomy over time. Field studies showed that workers often used these social tags to browse or rediscover information in a bookmark collection, significantly supporting information reuse.

Colleagues also benefit from these bookmarking applications due to the social nature of their design and use. Design details encourage public (visible) sharing of bookmarks, which coworkers can then search and browse. Social networks made up of workers with common interests are discernable from both shared tags and bookmarks, and these social networks can then be used to browse, filter, or track information resources of mutual interest. Furthermore, early dogear users expressed conscious tagging strategies that benefited coworkers by promoting the community or supporting common tag use among teams.

Perhaps even more important than the benefits of bookmark discovery provided within the social bookmarking application is the significant improvement in enterprise search enabled by integrating search and the bookmark data. A Firefox plug-in that integrated dogear bookmarks and an enterprise search application was well received by dogear users and offered significantly better search results for enterprise content. In many respects, when employees bookmark an intranet resource, they are performing a rating task, explicitly signaling that the resource is interesting and valuable. Follow-on studies within IBM show bookmarked resources are more relevant and more often clicked than other intranet resources. Other approaches to collaborative/social search are highlighted in this issue, including "Information Seeking Can Be Social" by Ed H. Chi.

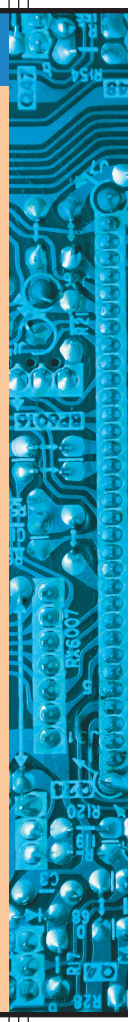
Current research is under way to more fully understand the use of social bookmarking applications in exploratory search activities. Social bookmarking applications are a natural way to collect, store, and manage information resources over the longer periods of time typically associated with exploratory search. Better tools are needed, however, to help workers identify important relationships across resources and to detect emerging trends. Integration with other enterprise tools such as new authoring environments—for example, wikis and blogs—may be desirable.

The bookmark collection for a community or organization is an important shared asset, and finding ways to better leverage this asset will be important.

Reference

1. D. Millen, J. Feinberg, and B. Kerr, "Social Bookmarking in the Enterprise," *Queue*, Nov. 2005, pp. 28-35.

***David Millen** is a research manager at IBM T.J. Watson Research Center, where his group explores new social software for large organizations. Contact him at david_r_millen@us.ibm.com.*



both synchronous and asynchronous collaboration: In SearchTogether, for example, users can save search results for others who might not be online, or they can run a query and examine search results together.

Location

Finally, collaborators might work in the same place at the same time, which allows them to communicate in a variety of ways not necessarily mediated by the computer, or they might be distributed, increasing opportunities for collaboration but decreasing the fidelity of possible communications.

ROLES IN COLLABORATION

Collaboration implies that different people can play different roles in the overall human-computer system. When working with others, Gail and John can divide the task in different ways depending on the roles they and their collaborators assume. How tasks are divided depends on the nature of the task, people's expertise, and the capabilities of the system that mediates information seeking. Roles might be implicit in the functionality of the interface, such as specifying queries and making relevance judgments, or they can be more explicit, for example, people using different interfaces for different subtasks.

Peer

The most obvious role is the peer. All collaborators use the same interfaces to control the system and coordinate their activities. This is the most common situation with existing (nonmediated) tools. Collaborators use their systems independently and combine their results manually, as Gail and her colleagues do. Experimental collaborative search systems such as Fischlár-DiamondTouch[2] and SearchTogether[3] implement peer search roles. Fischlár-DiamondTouch uses a multitouch surface to allow groups of people to manipulate search results in parallel. In SearchTogether, each person can specify queries, examine results, and so on.

Domain A expert/domain B expert

A variation on the peer role is collaboration between people with symmetrical user interfaces but different domain knowledge, for example, when Gail collaborates

with physicians. Mediation can increase the chances that users will recognize documents relevant to both sets of expertise.

Search expert/search novice or domain expert/domain novice

One common asymmetry in user skill and experience is the degree of expertise or familiarity with a domain and with search tools. One way to support different levels of familiarity with tools is to allow expert searchers to customize their interfaces with more sophisticated functionality. For example, one person might use an "advanced search" interface, while another uses a simple text box. Asymmetry in domain expertise requires a more nuanced approach. For example, SearchTogether can treat experts' and novices' relevance judgments differently to produce more reliable results.

Search expert/domain expert

The pairing of search expert/domain expert roles introduces true asymmetries in contributions of team members. This configuration facilitates collaboration between a skilled searcher and a person with a complex information need, as when Gail works with the reference librarian. The expert searcher knows how to select collections and formulate queries against those collections, but can only make rudimentary relevance judgments based on a description of the information need provided by the domain expert. The domain expert, on the other hand, has a better grasp of the evolving information need and can evaluate retrieved documents.

Another configuration of this role combination pairs a layperson with an information need with an expert domain coach, as when John collaborates with the RN. The layperson acts as a search expert, performing most of the querying and iterative refinement, while the domain expert makes occasional query suggestions and positive and negative relevance judgments. Other examples of this collaboration include student/adviser and law clerk/attorney. Suitably mediated, these roles allow people with different skill sets to produce results not easily obtained independently.

Prospector/miner

While the other roles focus on various combinations of the searchers' expertise, prospector/miner roles[4] focus on searchers' activities during the search. These roles allow one user to search broadly and the other to search deeply. The prospector generates many queries to explore the collection and makes a few relevance judgments for each result set before moving on to the next query. The miner makes detailed relevance judgments on documents identified by the prospector's queries. One example of these roles is when Gail suggests new search topics for her colleagues. In some ways, these roles parallel the search expert/domain expert roles, but rather than being driven by users' knowledge, the specialization into roles is driven by decomposing the information-seeking task into subtasks and developing specialized user interfaces for each subtask.

Other role combinations are also possible, and roles are not limited to pairs. Multiple peers can collaborate trivially, a search expert can mediate information seeking of multiple collaborating domain experts, and any number of prospectors and miners can work together.

Most complex tasks improve with collaboration. While social search has flourished over the past several years and ad hoc collaboration has become common in a variety of information-seeking situations, existing tools do not effectively support these tasks. The need for workarounds suggests that a real gap exists between users' desire to collaborate and the capabilities of the tools they use. The dimensions of collaboration and roles can serve as a design framework for systems that support explicit collaboration. We expect that tools for explicit collaboration will slowly move from the research domain to specific vertical markets, such as legal or medical research, then to mass-market acceptance. C

References

1. M.R. Morris, "A Survey of Collaborative Web Search Practices," *Proc. 26th Ann. SIGCHI Conf. Human Factors in Computing Systems*, ACM Press, 2008, pp. 1657-1660.
2. A.F. Smeaton et al., "Físchlár-DiamondTouch: Collaborative Video Searching on a Table," *Proc. SPIE Conf. Electronic Imaging-Multimedia Content Analysis, Management, and Retrieval*, SPIE, vol. 6073, 2006, pp. 8-19.
3. M.R. Morris and E. Horvitz, "SearchTogether: An Interface for Collaborative Web Search," *Proc. ACM Symp. User Interface Software and Technology* (UIST 07), ACM Press, 2007, pp. 3-12.
4. J. Pickens et al., "Algorithmic Mediation for Collaborative Exploratory Search," *Proc. ACM Special Interest Group on Information Retrieval* (SIGIR 08), ACM Press, 2008, pp. 315-322.
5. P. Briggs and B. Smyth, "Trusted Search Communities," *Proc. 12th Int'l Conf. Intelligent User Interfaces* (IUI 07), ACM Press, 2007, pp. 337-340.
6. C. Shah, "Toward Collaborative Information Seeking (CIS)," *Proc. 1st Int'l Workshop Collaborative Information Retrieval*, 2008; http://workshops.fxpal.com/jcdl2008/submissions/tmpE1.pdf.
7. G. Golovchinsky, J. Pickens, and M. Back, "A Taxonomy of Collaboration in Online Information Seeking," *Proc. Int'l Workshop Collaborative Information Retrieval*, June 2008; http://www.fxpal.com/?p=abstract&abstractID=454.
8. G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76-80.

Gene Golovchinsky *is a senior research scientist at FX Palo Alto Laboratory. His research interests include interfaces for information seeking, dynamic hypertext, and pen-based computing. He received a PhD in industrial engineering (human factors) from the University of Toronto. Contact him at gene@fxpal.com.*

Pernilla Qvarfordt *is a research scientist at FX Palo Alto Laboratory. Her research interests include human-computer interaction and computer-mediated communication, particularly in tools supporting collaboration and meeting environments. She received a PhD in computer science from Linköping University. Contact her at pernilla@fxpal.com.*

Jeremy Pickens *is a research scientist at FX Palo Alto Laboratory. His research interests include collaborative information seeking, multimedia analysis, and multimedia information retrieval (video, music). He received a PhD in computer science from the University of Massachusetts, Amherst. Contact him at jeremy@fxpal.com.*