



What's Missing from Collaborative Search?

Marti A. Hearst, *University of California, Berkeley*

Research shows that people frequently try to search for information with other people. The fact that no user interface for collaborative searching has yet caught fire suggests that the best parts of the design space have yet to be investigated.

It is fair to say that, according to the literature, tools dedicated to collaborative information seeking have not enjoyed widespread success to date.^{1,2} Surveys, interviews, and observations show that although now even more than ever, people are searching in pairs and in larger groups, they are doing so without the benefit of specialized search tools.³ Instead, they are making do by coordinating via out-of-channel use of communication tools, especially email, texting, phone calls, and social media communication.^{1,2,4}

While a recent study by Meredith (“Merrie”) Ringel Morris found that users were generally satisfied with their most recent collaborative search,² the results also showed that there is much room for improvement. Participants expressed a need to increase group awareness of mutual activities, which is a functionality that has been addressed in collaborative search research systems^{5,6,7} as well as commercial systems. Other participants requested (i) the ability to compare results with others in real time, and (ii) a way to reduce redundant work. Furthermore, there may be

specific categories of tasks that were not uncovered in the study by Morris,² in part because participants reported only their most recent use of collaborative search. The “open answer” portion of the survey suggested that users do desire more, such as support for complex tasks or tasks that extend over longer periods of time.

Research tools that allow for shared content and shared presence in both real time and asynchronous collaboration include SearchTogether,⁶ Coagmento,⁵ CollabSearch,⁷ and Results Space.⁸ These tools allow participants to find, save, and share documents, and see the activities of others in the collaboration group.

A study by Chirag Shah⁵ compared an interface in which pairs of participants could not see any status information, could see their own personal action history, or could see both their own and the action histories of their search partner. This work found that while the different conditions did not change the final search outcome in terms of quantitative evaluation metrics, awareness of the partner’s search history did reduce the number of coordination messages that needed to be exchanged after the search was completed.

Similarly, the work of Hans-Christian Jetter and his colleagues⁹ introduced a highly novel interface for finding hotel reservations in a group, setting constraints via a combination of visualization and haptic displays. A key feature of this interface was that it allowed individuals to set constraints “privately” in one corner of a tabletop display and then combine the constraints in the group publicly in the central view. Conflicts among the constraints were then

adjusted collaboratively until a set of hotels that met some subset of the constraints could be agreed upon. This approach was compared to a standard Web-based faceted interface. The overall outcomes were not distinguishable by the standard metrics, but the communication patterns were seen as having less “noise” when using the interface that was tailored to collaboration, because it was easier to keep track of system state and individual preferences.

Merrie Morris and Eric Horvitz⁶ discussed the friction inherent in having to switch from tools currently being used to a special purpose new tool. To date, people just do not seem willing to move from a standard search tool to another tool for collaborative search. The reason for this may be that there must be enough additional value as yet in the tools offered, and/or they may not yet be easy enough to use, to justify using a specialized tool. If that is the case, the question is, what, if anything, can a collaborative search tool offer that would make it worthwhile to switch to a different tool, or add in a new plug-in?

There is reason to hope that with the right approach, a tool could be developed to support collaborative search successfully. There are some recent prominent examples of new user interface solutions arising and becoming very widely adopted for long-standing problems in coordination and collaboration. For example, for many years there was no popular Web-based way to schedule small-group meetings among people who did not share a calendaring tool; instead, people made use of email chains to schedule meetings. Now online Web polls such as Doodle are a standard solution.¹⁰

As another example, for years there were no simple, high-usability tools available for an organization, such as a government agency, to use to put out a request for ideas and have thousands of lay people participate with comments and votes. The rise of online comment and moderation tools such as UserVoice and Google Moderator, not to mention the rise of third-generation question–answer forum tools such as StackExchange and Quora, show that it is possible to make progress in this space.

WHAT ARE KEY MISSING FEATURES?

The question is, what should the next generation of collaborative search tools be focused on? Below, we consider the special properties of three use cases in which collaboration is likely to be needed in a search task, and what therefore needs to be supported in a collaborative search tool above and beyond what existing tools have already successfully implemented.

SCENARIO 1: SELECTING A FEW FROM MANY SIMILAR CHOICES

Sometimes when people collaborate, it is because they need to make a decision together to meet the various group members’ desires. The classic example is travel planning:

Johnny wants a hotel with a good bar, Khoa wants to be near the beach, Pat does not want to pay very much, Sonali cannot come until the third day, and so on. Or some people engage in “competitive shopping,” in which they split up the work when looking at many similar items offered at different venues to try to negotiate best prices.

For this case, a collaborative search tool should offer structure to let individuals define what their personal constraints or preferences are in some manner, as they search. People should not be required to specify this information up front, as that requires too much cognitive overhead

There is some friction inherent in switching from a standard search tool to another tool for collaborative search.

and because the process of searching is likely to be what reveals reasonable values for those constraints in the first place. Perhaps as the searchers find important constraints, double-clicking on them or swiping them into a special location turns them into a tag that is visible to all in the group as something to pay attention to (price, near-the-beach) in further searches. A smart search engine will eventually learn what the common ones are and automatically recognize these across users, and populate the constraints when search results are retrieved, perhaps using XML microformats or some other common representation.

Research by Kristie Fisher and her colleagues has shown that people can build on simple representations that have already been begun by others,¹¹ and if a search tool can recognize when such a representation is being built, it could suggest it as an organizing tool, or simply match the structure being built with those already seen to make more intelligent suggestions.

SCENARIO 2: COVERING A TOPIC THOROUGHLY

The PhD student researching his or her dissertation work or a paralegal looking through a document trove trying to find all instances of some concept exemplify classic information retrieval challenge problems.¹² Collaborative search should be especially useful for this problem, since it should be amenable to “divide-and-conquer” techniques and since it can be implemented in special-purpose search tools such as academic reference search⁴ and legal document search.

But there are several fundamental functionalities that are missing from current collaborative search tools. For one, the tool should be aware both of what has already been stashed away in the bibliography and what has been viewed by anyone in the group of searchers. It should rerank based on this information, hiding what has already

been assessed (but allowing users to override this setting). Prior research systems, including Querium¹⁵ and Results Space,⁸ have made collaborators' explicit ratings visible, but assessments of these tools are done in laboratories in which users explicitly set up their tasks at fixed starting times, over fixed document collections. Open questions remain about what to do with leftover ratings, implicit information,¹⁴ and new group members joining and leaving the collaboration.

Searchers should be able to achieve a feeling of accomplishment, of walking over the landscape and having their bearing on the terrain as they work, and be able to strategize about who is to do what next over that map.

It may be useful to allow a trio of searchers to work together, with one doing triage using a general query and a general ranking algorithm, another looking at promising documents in more detail, and classifying those that are relevant using a set of agreed-upon categories or tags, and a third using a different ranking algorithm to further search within one of the refined categories, perhaps along the lines suggested by Jeremy Pickens and his colleagues.¹⁵

Such a tool should have a clear depiction of the landscape covered and the landscape yet to be looked at, organized by predetermined categories as well as user-defined keywords, and sortable by citations or usage where available. Searchers should be able to achieve a feeling of accomplishment, of walking over the landscape and having their bearing on the terrain as they work, and be able to strategize about who is to do what next over that map.

SCENARIO 3: DISCOVERING UNKNOWN INFORMATION

By far the most challenging search task is that of trying to help people make a new discovery, such as solving the mystery of why honeybees are dying in North America. For a problem like this, search over known materials is only one part of a multifaceted, wide-ranging collaborative effort. But it is worth considering the role of collaborative search for this type of problem. Any collaborative search tool should always be comparing what has already been added to a community collection to what is currently being viewed to see if the information is redundant.

Text-mining algorithms have long been proposed as an aid to discovery of new information,¹⁶ but data-mining and knowledge-discovering algorithms have not emphasized collaboration for the most part. Collaborative challenges—such as the DARPA Red Balloon Challenge, in which teams

had 10 hours to collaboratively discover the physical position of 10 red weather balloons released before dawn across the continental US,¹⁷ or the Polymath Project,^{18,19} in which the world was invited to collaboratively solve a difficult mathematics proof, and the goal was achieved by 27 people in a matter of months using comments on blogging and wiki platforms—suggest directions forward.

Because people are searching together on a regular basis, there is a need for support for this activity in search user interfaces. The fact that no such interface has caught fire does not suggest there is no need for such a design—but, rather, that the best parts of the design space have not been investigated fully yet. What they are and how to present them remain open questions; this article has made a few suggestions about directions for exploration. ■

References

1. M.R. Morris, "A Survey of Collaborative Web Search Practices," *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 08)*, 2008, pp. 1657–1660.
2. M.R. Morris, "Collaborative Search Revisited," *Proc. 2013 Conf. Computer-Supported Cooperative Work (CSCW 13)*, 2013, pp. 1181–1192.
3. M. Twidale and D. Nichols, "Collaborative Browsing and Visualisation of the Search Process," *Proc. Assoc. Information Management (ISTCT)*, vol. 48, nos. 7–8, 1996, pp. 177–182; http://eprints.lancs.ac.uk/53455/1/twidale_aslib_96.pdf.
4. R. Capra et al., "Tools-at-Hand and Learning in Multi-Session, Collaborative Search," *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 10)*, 2010, pp. 951–960.
5. C. Shah, "Effects of Awareness on Coordination in Collaborative Information Seeking," *J. Am. Soc. Information Science and Technology*, vol. 64, no. 6, 2013, pp. 1122–1143.
6. M.R. Morris and E. Horvitz, "SearchTogether: and Interface for Collaborative Web Search," *Proc. 20th Ann. ACM Symp. on User Interface Software and Technology (UIST 07)*, 2007, pp. 3–12.
7. Z. Yue, S. Han, and D. He, "An Investigation of Search Processes in Collaborative Exploratory Web Search," *Proc. Am. Soc. Information Science and Technology (ASIST 12)*, vol. 49, no. 1, 2012, pp. 1–4.
8. R. Capra et al., "Design and Evaluation of a System to Support Collaborative Search," *Proc. Am. Soc. Information Science and Technology (ASIST 12)*, vol. 49, no. 1, 2012, pp. 1–10.
9. H.-C. Jetter et al., "Materializing the Query with Facet-Streams: A Hybrid Surface for Collaborative Search

- on Tabletops,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 07)*, 2011, pp. 3013–3022.
10. K. Reinecke et al., “Doodle around the World: Online Scheduling Behavior Reflects Cultural Differences in Time Perception and Group Decision-making,” *Proc. 2013 Conf. Computer-Supported Cooperative Work (CSCW 13)*, 2013, pp. 45–54.
 11. K. Fisher, S. Counts, and A. Kittur, “Distributed Sensemaking: Improving Sensemaking by Leveraging the Efforts of Previous Users,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 12)*, 2012, pp. 247–256.
 12. D.C. Blair and M. Maron, “An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System,” *Comm. ACM*, vol. 28, no. 3, 1985, pp. 289–299.
 13. G. Golovchinsky, A. Diriye, and T. Dunnigan, “The Future Is in the Past: Designing for Exploratory Search,” *Proc. 4th Information Interaction in Context Symp. (IliX 12)*, 2012, pp. 52–61.
 14. S. Fox et al., “Evaluating Implicit Measures to Improve Web Search,” *ACM Trans. Information Systems*, vol. 23, no. 2, 2005, pp. 147–168.
 15. J. Pickens et al., “Algorithmic Mediation for Collaborative Exploratory Search,” *Proc. 31st Ann. Int’l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 08)*, 2008, pp. 315–322.
 16. M.A. Hearst, “Untangling Text Data Mining,” *Proc. 37th Annual Meeting Assoc. Computational Linguistics (ACL 99)*, 1999, pp. 3–10.
 17. J.C. Tang, “Reflecting on the DARPA Red Balloon Challenge,” *Comm. ACM*, vol. 54, no. 4, 2011, pp. 78–85.
 18. T. Gowers and M. Nielsen, “Massively Collaborative Mathematics,” *Nature*, vol. 461, 2009, pp. 879–881.
 19. J. Cranshaw and A. Kittur, “The Polymath Project: Lessons from a Successful Online Collaboration in Mathematics,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 11)*, 2011, pp. 1865–1874.
- Marti A. Hearst** is a professor in the School of Information at the University of California, Berkeley, with an affiliate appointment in the Computer Science Division. Her primary research interests are user interfaces for search engines, information visualization, natural language processing, and improving MOOCs. Hearst received a PhD in computer science from the University of California, Berkeley. She is an ACM Fellow and has received an NSF CAREER award, an IBM Faculty Award, two Google Research Awards, an Okawa Foundation Fellowship, and two Excellence in Teaching Awards. Contact her at hearst@berkeley.edu

 Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.



computing|now

GET HOT TOPIC INSIGHTS FROM INDUSTRY LEADERS

- Our bloggers keep you up on the latest Cloud, Big Data, Programming, Enterprise and Software strategies.
- Our multimedia, videos and articles give you technology solutions you can use.
- Our professional development information helps your career.

Visit ComputingNow.computer.org. Your resource for technical development and leadership.

 **IEEE**

 **IEEE computer society**

Visit <http://computingnow.computer.org>

