

Collaborative Information Seeking by the Numbers

Robert Capra, Javier Velasco-Martin, Beth Sams

School of Information and Library Science
University of North Carolina at Chapel Hill

rcapra@unc.edu, jvelasco@unc.edu, bsams@unc.edu

ABSTRACT

In this paper, we present results of an on-line survey conducted in early 2010 to inquire about collaborative information seeking activities and behaviors in the context of Web searches. We recruited participants using the Amazon Mechanical Turk service to gather responses from people with a wide range of backgrounds. We present results about the frequency of collaborative Web searches, number of sessions involved, concurrency, location, collaborative group size, and communication methods used to collaborate. We draw comparisons with prior work and highlight collaborative search scenarios that are important for tools and systems to support.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; K.4.3 [Computers and Society]: Organizational Impacts

General Terms

Measurement, Experimentation, Design, Human Factors

Keywords

Collaborative information seeking, information sharing

1. INTRODUCTION

Collaborative information seeking (CIS) and collaborative information retrieval (CIR) are varied activities that can occur for many different purposes, in many contexts, and among many different groups of collaborators. In this paper, we report on results from an online survey to gain insights about “everyday” collaborative information seeking in the specific context of Web search. We recruited participants from a wide variety of backgrounds using the Amazon Mechanical Turk service and asked questions about their individual and collaborative Web search behaviors. Our results extend prior research on CIS and CIR practices – especially studies by Morris [10] and Evans and Chi [6] – and provide additional insights about the types of situations and behaviors that occur in collaborative and cooperative searches. Results from our study provide information that will be useful to researchers and designers of CIS systems.

2. RELATED WORK

Morris [10] conducted a survey of “204 knowledge workers at a large technology company” and found that over half (53%) had cooperated on searches with other people. Morris also

documented common collaborative activities such as watching over someone’s shoulder as they search, dividing up parts of a search, and using instant messaging to share search results synchronously. Evans and Chi [6] conducted a critical-incident style survey of 150 users of the Amazon Mechanical Turk service to inquire about users’ “most recent search experience” and found that 31% of the searches were motivated by an external prompt (e.g., another person). Evans and Chi also reported that searchers engaged in a number of social interactions before, during, and after their searches. Capra et al. [2] conducted 30 in-person interviews about users’ current practices and needs for sharing information from searches and found that many users relied on simple “tools-at-hand” such as text editors and email to take notes and exchange results.

There are many dimensions that can be considered in CIS, including: concurrency (synchronous vs. asynchronous), location (co-located vs. remote), and intention (implicit vs. explicit) [8][11][16]. In addition, Morris, Fisher, and Wigdor [12] outline a design space that includes collaborative style, group size, location, and device ecology. All these dimensions represent factors that may influence the nature of the collaboration and the needs of collaborators in an information seeking process.

Studies of collaborative information seeking behaviors are often conducted with information workers, library patrons, and students, although some studies have examined CIS in families and other specific populations [11, section 2.1]. Information workers and students have frequent needs to engage in CIS and are logical users to study. However, Morris’ [10] survey of cooperative search behaviors among knowledge workers motivated us to wonder if similar patterns would be found among users in different contexts and among different sets of collaborators. Evans and Chi [6][7] began to address such questions in their survey of Amazon Mechanical Turk users and our work extends this line of investigation.

The research we present here is part of our effort to extend our current understanding of CIS behaviors and practices across a range of users, situations, and types of cooperation on search [2][3][4]. Understanding these dimensions and how they influence CIS will benefit designers and developers of tools to support CIS activities.

3. METHOD

Like Evans and Chi [6][7], we conducted an online survey and recruited participants using the Mechanical Turk service (MTurk). One of our primary reasons for using the Mechanical Turk was to reach a set of users outside the traditional knowledge worker and student populations. Using the MTurk also allowed us to recruit a reasonably large set of responses in a short period of time. For more information about using the Mechanical Turk in HCI research, interested readers are referred to Kittur, Chi, and Suh [9] and Ross et al., [13].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CIR’11, October 28, 2011, Glasgow, Scotland, UK.

Copyright 2011 ACM 978-1-4503-0951-6/11/10...\$10.00.

3.1 Survey

We developed and pilot tested a survey with sections about several types of searches and search behaviors including: 1) individual known-item searches, 2) individual exploratory searches, 3) searches that involved sharing found information with other people, and 4) collaborative/cooperative searches.

In the survey, we used the term “cooperate” instead of “collaborate”. We did this in part to use more accessible language, but also to include a slightly wider range of “searching together” activities than might be implied by a strict reading of “collaborative search”. Recently, Denning and Yaholkovsky [5] outlined levels of “working together” including information sharing, coordination, cooperation, and collaboration and Shah [14] suggested ways to apply these models to CIS. In prior research, we have discussed CIS behaviors on a similar continuum from loose information sharing to intentional collaboration [2][3], but we have also proposed considering CIS instances along two separate dimensions: the motivation for the search (e.g., self-motivated versus external), and the motivation for sharing results [3].

Our primary interest in the survey was to understand dimensions and practices related to collaborative/cooperative searches and sharing information from searches. Sections on individual known-item and exploratory searches were included mainly to provide points of comparison. In this paper, we focus primarily on results from the section about cooperative searches.

For each of the sections, we developed carefully worded descriptions to help participants understand the distinctions among the different types of searches. For example, the section on sharing was titled “Sharing results of Web searches” and started with the text, “*Sometimes when you do a Web search, you may decide to share some of the results you find with other people*”. The section on cooperative searches was titled, “Searching in cooperation with other people” and began, “*Next, we would like you to think about some times that you have done Web searches in cooperation with other people. This could be while you were in the same location, or at different locations. You might have been doing the searching at the same time as the other people, or at a different time.*”

The first page of our survey was an informed consent document. Participants indicated that they met our eligibility criteria (over 18 years old, and had done Web searches that involved sharing and cooperation) and that they agreed to participate by clicking a button to go to the first page of the survey. Participants were free to answer or skip any questions on the survey, so the response counts vary slightly for each question. The survey took approximately 15 minutes to complete.

3.2 Participants

In addition to the eligibility criteria mentioned above, in our Mechanical Turk posting, we included two MTurk-imposed restrictions on participation: 1) the Turkers must be registered in the United States, and 2) the Turkers must have a reputation score of 95% in the MTurk system.

A total of 452 people began the survey and 344 gave usable answers to the main questions. Of these, 64% were female, 32% were either full- or part-time students, and 59% were employed at least part-time. Participants ranged in age from 18 to 68 years ($M=32, SD=11$). Ninety-one percent (91%) of the respondents reported taking at least some college courses and 56% had completed a 4-year college degree. Participants came from a wide

variety of backgrounds including: teacher, delivery man, operations specialist, automotive assembler, business manager, scientist, police officer, research assistant, and riding instructor..

4. RESULTS

4.1 Frequency of Searches

The survey had sections with questions on four types of search activities (known-item search, exploratory searches, sharing results, and cooperating on searches). In each section, we asked participants to indicate how frequently they did the corresponding type of search on an ordinal scale with five choices: multiple times a day, daily, weekly, monthly or less, and never. Figure 1 summarizes the responses to these questions, showing the relative percentages of each response.

Participants often conducted known-item and exploratory searches, with 63% and 49% (respectively) indicating that they did these types of searches daily or more frequently. Sharing results of searches daily or more was also common (36%). Cooperation on search was less common, with 2% indicating multiple times a day, 7% daily, 29% weekly, 58% monthly, and 4% never. Interestingly, these results are slightly higher than those reported by Morris [10]. Knowledge workers in her survey reported cooperating on searches daily (1%), weekly (26%), monthly (49%), and yearly (25%). The differences in these frequencies may be due to differences in the samples between the two studies, or may be due to increased tendencies to engage in cooperative searching from November 2006 when Morris conducted her survey to February/March 2010 when our data was gathered.

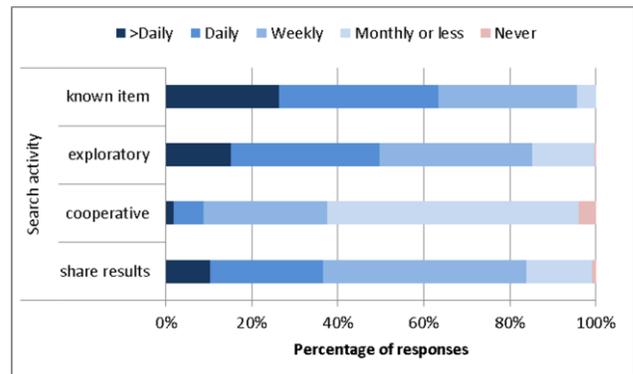


Figure 1. Frequency of Types of Searches

4.2 Sessions, Concurrency, Location

In the section of the survey about cooperative searches, we asked participants to think of a specific situation in which they had done a Web search in cooperation with other people. Participants were asked to describe the specific situation in a text box and then answer a series of fixed-choice and open-ended questions about the specific situation. We asked: if the search was completed in one session or multiple sessions; if people worked on the search at the same time, on their own time, or a mix of both; and whether the people working on the search were located in the same room, different locations, or a mix of both. Table 1 summarizes the responses for the 334 respondents that provided answers to all three questions.

Table 1 shows several interesting results. First, there is about an even split between single session (47% = 37% + 4% + 6%) and multiple session searches (53% = 9% + 16% + 28%). Multiple session searches may be indicative of more complex, exploratory

searches and are likely to have different collaborative needs versus single session searches. Interestingly, within the single session searches, over half (and 25% overall) were synchronous and co-located. Such situations would benefit from tools to help users share and record results, but may not require the same communication support needed for situations in which collaborators are remotely located.

Next, we consider the locations of the collaborators. Table 1 shows that overall, in 38% of the cooperative searches the collaborators were co-located, and in 42% they were in different locations (remote). Morris' (2008) findings were that 22% were always co-located, 12% were always remote, and 66% experienced both types. Different needs are often present for co-located versus remote collaboration – for example, Amershi and Morris [1] found that many co-located searches were fairly brief. Our results indicate that both scenarios are common and should be supported.

Furthermore, for 20% of the scenarios described by our participants, there was a mix in which some collaborators were co-located and some were remote. Morris and Teevan [11] note that most CIS systems “are designed around exclusively remote or co-located scenarios” but that “mixed-presence arrangements... are also possible”. Our results suggest that mixed-presence scenarios can be fairly common and that CIS systems should consider the needs and requirements to support these situations.

Table 1 shows that there are five predominant scenarios that account for almost three-quarters (73%) of the cooperative searches described by our participants:

- Single session; synchronous; co-located (25%)
- Single session; synchronous; remote (11%)
- Multi-session; asynchronous; remote (13%)
- Multi-session; mix sync/async; remote (10%)
- Multi-session; mix sync/async; mix locations (14%)

These results illustrate a range of types of cooperative searches, and suggest the need for different types of tools to support different types of CIS.

Table 1. Cross-tabulated frequency counts

		Co-Loc	Remote	Mix	Σ
ONE SESS	Sync	82 (25%)	37 (11%)	3 (1%)	122 (37%)
	Async	0 (0%)	11 (3%)	4 (1%)	15 (4%)
	Mix	10 (3%)	5 (1%)	4 (1%)	19 (6%)
MULT SESS	Sync	19 (6%)	10 (3%)	2 (1%)	31 (9%)
	Async	4 (1%)	44 (13%)	6 (2%)	54 (16%)
	Mix	12 (4%)	33 (10%)	48 (14%)	93 (28%)
	Σ	127 (38%)	140 (42%)	67 (20%)	334 (100%)

4.3 Number of people cooperating

Group size has been identified as a design dimension for CIS tools [12]. In our survey, we inquired about this, asking, “How many people were cooperating on the search?” Participants reported an average of 3.04 collaborators (median = 2, SD=3.32). Of the 325 total responses to this question, 187 (58%) reported two

collaborators, 66 (20%) reported three, and only 26 (8%) reported more than four. Figure 2 shows a graph of the number of collaborators reported by each participant, sorted by increasing size. We note that three outliers of 20, 37, and 40 collaborators were omitted from the graph for presentation. A small percentage of participants reported one collaborator, suggesting that this question might have benefited from wording to “include yourself” in the count of collaborators. The group sizes reported by our participants are slightly larger on average than those reported by Morris [10]. She reported that 81% of her respondents indicated groups of two, 19% reported three or four, and none reported larger sizes. Across both studies, the median group size of two is comparable and in both studies the sizes were generally smaller than five.

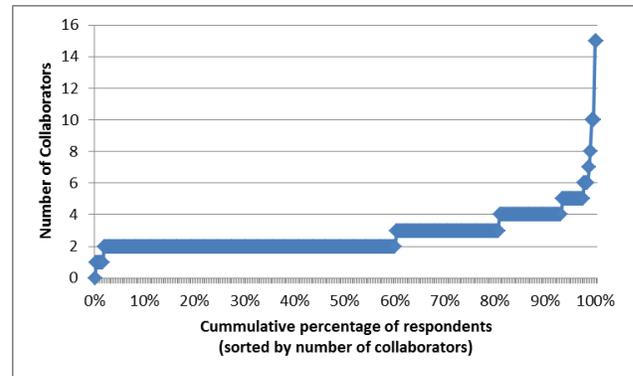


Figure 2. Distribution of the number of collaborators

4.4 Communication tools used

We wanted to understand more about the communication channels used by the collaborators and asked, “How did you communicate with the other people to cooperate on the search? (choose all that apply)”. Face-to-face communication was the most prevalent (68%), followed by email (39%), phone (34%), instant messaging (21%), and text messaging (16%). Social network services (e.g. Facebook, Twitter, etc.) (8%) and writing a blog post/wiki/web page were less common (1%) methods to support collaboration on the search.

As points of comparison, we consider two surveys. Morris' [10] survey also inquired about methods used to cooperate on searches. Turner et al. [15] surveyed 32 knowledge workers at a small US corporation in May 2008 and again in May 2009 about their general use of communication tools in the workplace. Table 2 summarizes the data from our survey and the two related studies.

Table 2. Comparisons of communication tools

	Morris (2008) – CIS use	Turner (2010) – Work use	Our survey – CIS use
F2F	88% watched & suggested queries	100%	68%
Phone	49%	~90%	34%
Email	86% links 60% summary	--	39%
Wiki / Web page	15%	~50%	1%
IM	30%	~79%	21%
Text Msg	--	--	16%
SNS	--	~70%	8%

Face-to-face communication is common across all three surveys, but the Turner study reports higher use of Web, instant messaging (IM), and social networking services (SNS) – perhaps an artifact

of the communication ecology at the small corporation studied. Compared with Morris' survey of CIS, use of email, phone, and IM are comparable, but slightly lower in our sample. These differences may be due to differences in the samples, or in the nature of the collaborations – Morris's participants were knowledge workers describing collaborative search situations in a work environment whereas our participants had a more varied set of backgrounds and collaborative search scenarios. Overall, these results support Evans and Chi's [7] findings that people communicate with friends before, during, and after searches across a variety of search types and situations.

4.5 Assigned work tasks

As we reviewed the data, we noticed that a distinct group of search tasks described by participants were done as part of assigned work or school projects. To investigate this, the three co-authors individually coded each response as either "assigned", or "not assigned" (along with a code for missing/problematic responses). The "assigned" classification included all assigned tasks in both work and school contexts. Based on our preliminary analysis of this data, 38% of the searches were coded as "assigned", and 62% were "not assigned". In future work, we plan to look for differences between these groupings in terms of other dimensions such as the number of collaborators, concurrency, and location.

5. SUMMARY

In this paper, we presented results from an online survey about collaborative and cooperative search behaviors from participants with a range of backgrounds. Our results illustrate important scenarios that designers of CIS systems should consider. Single-session, synchronous, co-located search is highlighted as a common scenario, as are mixed-presence and mixed-concurrency situations. A majority of collaborative searches in our study involved two people and only 8% of groups were larger than four. Face-to-face communication, email, and talking by telephone were commonly used to coordinate cooperation on searches, and instant messaging was to a lesser degree. Our results extend and update those from Morris' 2008 [10] survey of knowledge workers and suggest additional areas for exploration such as differences between assigned and non-assigned collaborative search tasks.

6. ACKNOWLEDGMENTS

This work was supported by the National Science Foundation grant IIS 0812363. We thank Gary Marchionini, Chirag Shah, and Katrina Muller for their valuable help and advice on this project. We also thank our participants for sharing their experiences.

7. REFERENCES

- [1] Amershi, S., & Morris, M. R. (2009). Co-located collaborative web search: understanding status quo practices. In *Proc. CHI 2009 Extended Abstracts*. ACM Press, New York, 3637–3642.
- [2] Capra, R., Marchionini, G., Velasco-Martin, J., & Muller, K. (2010). Tools-at-hand and learning in multi-session, collaborative search. In *Proc. CHI 2010*. New York: ACM Press, 951-960.
- [3] Capra, R., Velasco-Martin, J., and Sams, B. (2010). Levels of "Working Together" in Collaborative Information Seeking and Sharing. In *Collaborative Information Seeking. Workshop held at GROUP 2010*.
- [4] Capra, R., Muller, K., and Velasco-Martin, J. (2010c). Classifications of Collaborative Search. In *Proceedings of the 2nd International Workshop on Collaborative Information Retrieval*. Held at ACM CSCW 2010.
- [5] Denning, P., and Yaholkovsky, P. (2008). Getting to "We". *Communications of the ACM* 51(4): 19-24.
- [6] Evans, B. M. and Chi, E. H. (2008). Towards a model of understanding social search. In *Proc. CSCW '08*. ACM Press, New York, 485-494.
- [7] Evans, B. M., & Chi, E. H. (2010). An elaborated model of social search. *Information Processing & Management*, 46(6), 656-678.
- [8] Golovchinsky, G., Qvarfordt, P., and Pickens, J. (2009). Collaborative Information Seeking. *IEEE Computer* 42(3): 47-51.
- [9] Kittur, A., Chi, E. H., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. In *Proc. CHI 2008*. ACM Press, New York, 453-456.
- [10] Morris, M. R. (2008). A survey of collaborative web search practices. In *Proc. CHI '08*. ACM Press, New York, 1657-1660.
- [11] Morris, M. R., & Teevan, J. (2009). Collaborative Web Search: Who, What, Where, When, and Why. *Synthesis Lectures on Information Concepts, Retrieval, and Services*. Morgan & Claypool.
- [12] Morris, M. R., Fisher, D., & Wigdor, D. (2010). Search on surfaces: Exploring the potential of interactive tabletops for collaborative search tasks. *Information Processing & Management*, 46(6), 703-717.
- [13] Ross, J., Irani, L., Silberman, M., Zaldivar, A., & Tomlinson, B. (2010). Who are the crowdworkers?: Shifting demographics in mechanical turk. In *Proc. CHI 2010*. ACM Press, New York, 2863–2872.
- [14] Shah, C. (2008). Toward Collaborative Information Seeking (CIS). *Proc. of Collaborative Exploratory Search Workshop at JCDL 2008*. Pittsburg, PA, June 20, 2008.
- [15] Turner, T., Qvarfordt, P., Biehl, J. T., Golovchinsky, G., & Back, M. (2010). Exploring the workplace communication ecology. In *Proc. CHI 2010*. ACM Press, New York, 841-850.
- [16] Twidale, M. B., Nichols, D. M., & Paice, C. D. (1997). Browsing is a collaborative process. *Information Processing and Management*, 33(6): 761-783.