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

3rd International Workshop on Collaborative Information Retrieval (held as part of CIKM 2011)
October 28, 2011
Glasgow, Scotland
Background and Motivation

• Goal
  - Gain insight into collaborative information seeking and sharing activities across a range of users, situations, and types of cooperation

• Related studies
  - Morris (2008)
    - 204 knowledge workers
    - 150 mTurkers
  - Turner, Qvarfordt, Biehl, Golovchinsky, & Back (2010)
    - 32 knowledge workers
Method

• Survey using Amazon mTurk + Qualtrics

• Asked details about **specific** recent searches
  - 2 searches in which they shared results found
  - 1 search “in cooperation with other people”
  - 1 known-item (individual)
  - 1 exploratory (individual)

We used the phrase “searches in cooperation with other people” to be more accessible than “collaborative search”

• Asked about **general** practices and frequency
  - For sharing results of searches
  - For cooperating with other people on searches
Demographics

- 452 started, 344 completed & paid
- 64% female, 36% male
- Age: avg = 32, median = 28, stdev = 11

<table>
<thead>
<tr>
<th></th>
<th>Employed</th>
<th>Not employed</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>52</td>
<td>56</td>
<td>108 (32%)</td>
</tr>
<tr>
<td>Not student</td>
<td>150</td>
<td>84</td>
<td>234 (68%)</td>
</tr>
<tr>
<td></td>
<td>202 (59%)</td>
<td>140 (41%)</td>
<td></td>
</tr>
</tbody>
</table>

Of the 108 students: 47% pursuing Bachelor’s degree, 27% Master’s

Occupations: teacher, delivery man, operations specialist, automotive assembler, business manager, scientist, police officer, research assistant, and riding instructor
Frequency of Search Activities

- **known item**:
  - >Daily: 0%
  - Daily: 10%
  - Weekly: 20%
  - Monthly or less: 30%
  - Never: 40%

- **exploratory**:
  - >Daily: 0%
  - Daily: 10%
  - Weekly: 20%
  - Monthly or less: 30%
  - Never: 40%

- **cooperative**:
  - >Daily: 0%
  - Daily: 10%
  - Weekly: 20%
  - Monthly or less: 30%
  - Never: 40%

- **share results**:
  - >Daily: 0%
  - Daily: 10%
  - Weekly: 20%
  - Monthly or less: 30%
  - Never: 40%
Frequency of Search Activities

- **known item**
  - >Daily: 10%
  - Daily: 20%
  - Weekly: 30%
  - Monthly or less: 30%
  - Never: 10%

- **exploratory**
  - >Daily: 10%
  - Daily: 20%
  - Weekly: 30%
  - Monthly or less: 30%
  - Never: 10%

- **cooperative**
  - >Daily: 5%
  - Daily: 25%
  - Weekly: 30%
  - Monthly or less: 20%
  - Never: 20%

- **share results**
  - >Daily: 10%
  - Daily: 20%
  - Weekly: 30%
  - Monthly or less: 30%
  - Never: 10%

Source: Morris 2008
Specific collaborative searches

Searching in cooperation with other people

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.

Think about a situation in which you have done a Web search in cooperation with another person (or several people) and describe it below.

Follow-on questions asked:
One session or multiple sessions?
Sync, async, or mix?
Co-located, remote, or mix?

Answer the following questions based on the situation you described above.
### Common collaboration scenarios

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

Five scenarios account for 75%:

- Single session; synchronous; co-located (25%)
- Multi-session; mix sync/async; mix loc. (14%)
- Multi-session; asynchronous; remote (13%)
- Single session; synchronous; remote (11%)
- Multi-session; mix sync/async; remote (10%)
How many people were cooperating on the search?
How did you communicate with the other people to cooperate on the search? (choose all that apply)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F2F</td>
<td>88% watched &amp; suggested queries</td>
<td>100%</td>
<td>68%</td>
</tr>
<tr>
<td>Phone</td>
<td>49%</td>
<td>~90%</td>
<td>34%</td>
</tr>
<tr>
<td>Email</td>
<td>86% links</td>
<td>--</td>
<td>39%</td>
</tr>
<tr>
<td></td>
<td>60% summary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiki / Web page</td>
<td>15%</td>
<td>~50%</td>
<td>1%</td>
</tr>
<tr>
<td>IM</td>
<td>30%</td>
<td>~79%</td>
<td>21%</td>
</tr>
<tr>
<td>Text Msg</td>
<td>--</td>
<td>--</td>
<td>16%</td>
</tr>
<tr>
<td>SNS</td>
<td>--</td>
<td>~70%</td>
<td>8%</td>
</tr>
</tbody>
</table>
Assigned tasks

• Assigned tasks
  ▪ We noticed a large number in our data set
    • Assigned work tasks
    • Assigned school projects
  ▪ To investigate, three authors coded
    • Assigned (38%)
    • Not assigned (62%)

• Imposed versus self-generated
  ▪ This distinction has been examined in individual information seeking tasks
    ▪ Gross (1995); Bilal (2002)
    ▪ See our ASIS&T 2011 CIS workshop paper
Two self/others dimensions

- **Search**
  - Expressed interest (request for info)
  - Cooperation & Collaboration

- **Share**
  - Self-initiated (e.g. “it was cool”)
  - Perceived Interest

- Self-initiated
  - Proving a point / Resolving a debate

- Promted by others

- Self-initiated

- Promted by others
Questions

• For questions, please contact
  – Rob Capra, rcapra@unc.edu