Investigating Features in Support of Web Tools for Information Gathering

Anwar Alhenshiri
Dalhousie University
anwar@alhenshiri.com

Carolyn Watters
Dalhousie University
cwatters@dal.ca

Michael Shepherd
Dalhousie University
shepherd@cs.dal.ca

Jack Duffy
Dalhousie University
Jack.duffy@dal.ca

Abstract

This paper presents a study that investigated three features to improve how users perform information gathering tasks on the web. These features were based on recommendations that were developed in a previous work. The results of the previous study indicated that tools supporting information gathering could be improved by: 1) keeping track of references to information as it is gathered for the task 2) keeping the task information integrated between sessions; and 3) providing integrated search, browsing, and editing capabilities. The results of the study described in this paper using a prototype indicate that these features improve performance.

1. Introduction

Prior to the emergence of the web, information seeking models were introduced in the literature (Ellis, 1989; Marchionini, 1995; Choo, et al., 1998). Those models identified the sequences of activities users perform to locate the information needed. Following the emergence of the web, researchers focused more on the concept of a complete task and on identifying the kinds of complex tasks users perform on the web.

In this context, task is used to describe the broader information goal of the user. Different models have been built to categorize the types of user tasks on the Web (Ahn et al., 2008; Tao, 2011). Broder (2002) identified three different kinds of tasks: navigational, informational, and transactional tasks. Similarly, but with different labeling of Web tasks Sellen (2002) categorized user tasks on the web as: information gathering, finding, browsing, transacting, communicating, and housekeeping. Building upon the work of Broder (2002), Rose and Levinson (2004) inferred users' goals based on the tasks performed as: navigational, informational, and resource based. Kellar, et al. (2007) developed a model of user tasks on the web that identified: fact finding, information gathering, transacting, and browsing as the main kinds of tasks users perform based on the results of a field study. In each of these studies, information gathering was found to be a very common task.

After categorizing the kinds of tasks users performed on the web, researchers began to examine more closely some of those tasks further, including information gathering.

Information gathering tasks, earlier labeled informational, are complex, highly search reliant, often require more than one session, and typically result in an information product; such as notes or a report.

The task of information gathering has been found to represent between 48% and 61.25% of all the tasks users perform on the web (Broder, 2002; Rose and Levinson, 2004). Amin (2009) identified many of the defining characteristics of information gathering. The task of information gathering on the web was chosen for investigation in this research for several reasons. First, information gathering typically requires collecting information from different sources. Second, information gathering requires the completion of subtasks requiring multiple applications and tools. Finally, information gathering typically requires multiple sessions to complete.

The effectiveness of current web tools to support users dealing with information gathering tasks has been shown to be problematic for users (Alhenshiri et al. 2010).

In a previous study work, Alhenshiri, et al. (2012a) developed recommendations for the design of web tools intended for information gathering including: support the re-finding of pages, save information between sessions, and integrate the management and organization of information related to the task.

Based on those recommendations, a prototype was built for use in the study described in this paper. The purpose of the study was to compare specific features built in the prototype to support the recommendations against the conventional use of a browser for information gathering tasks. The research questions included:

1. Are automatically generated thumbnails of accessed pages more effective in keeping track of web pages than conventional methods such as bookmarks, or copying and pasting links into text files.

978-1-4799-2504-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HICSS.2014.121

2. Is a single integrated state of an information gathering session more effective for information gathering over multiple sessions than conventional strategies such as saving pages, saving information in files, bookmarking, and so forth.
3. Is the use of a single application within the browser that supports searching and managing data more effective than the use of multiple applications (i.e. the web browser and a text editor) in the case of the subtask of managing and organizing information.

The remainder of this paper is structured as follows. Work related to the investigations of the task of information gathering is illustrated in Section 2. Section 3 explains the research study. Section 4 provides a detailed discussion of the study results. Section 5 concludes the paper.

2. Related work

Researchers have categorized the tasks users perform on the web and information gathering is consistently identified as a very frequent task. Information gathering tasks involve collecting information possibly of different types from different sources to achieve an overall goal (Alhenshiri, et al., 2012a). Information gathering tasks are mostly search-based as shown by Kellar, et al. (2007) and Amin (2009). Information gathering was recognized as the most frequent task goal for users who are re-finding information on the web (Kellar, et al., 2006) and even for users involved in a search (Rose and Levinson, 2004).

Earlier research (Alhenshiri, et al., 2010b, 2012b) identified subtasks that are typically part of the overall task of information gathering. The core subtasks identified were: finding information sources, finding information, managing information, handling multiple sessions, and re-finding information. A model was created of the relationship of those subtasks to the overall task as shown in Figure 1.

Information gathering tasks have been studied over the past few years as part of examining user interactions on the web for searching and navigation, for example Kules, et al. (2008). Researchers have investigated general aspects of the information gathering task. For example, Yamada and Kawano (2009) used sections in web pages located for an information gathering task to extract links to other pages. The target pages were considered a part of the user plan for the task and suggested to the user to continue gathering related information. In a similar approach, Bagchi and Lahoti (2009) used hyperlink connectivity among web pages to assist users in gathering information on the web. They argued that providing links to pages currently being viewed by the user can facilitate the process of information gathering. However, the only subtask of information gathering considered in these two studies was locating web information, i.e. finding.

Dearman, et al. (2008) investigated the subtask of finding sources of information during information gathering tasks. Re-finding information on the web has been investigated with respect to locating previously found results (Tauscher and Greenberg, 1997, Mackay and Watters, 2008) and for monitoring web sources of information (Kellar, et al., 2007). Issues with how users deal with information gathering and how they manage their time for the task were discussed in the work of Murphy (2003). Tao and Li (2009) addressed the problems of information mismatching and overloading during information gathering using concept-based personalized techniques. They suggested that improvements are needed for the representation and acquisition of user profiles in personalized web information gathering. Finally, Zilberstein and Lesser (1996) looked at decision making as an intermediate step in information gathering tasks.

The research conducted prior to the study discussed in this paper attempted to model the subtasks comprising the overall task of information gathering on the web. The subtasks are shown in Figure 1 and described as follows:

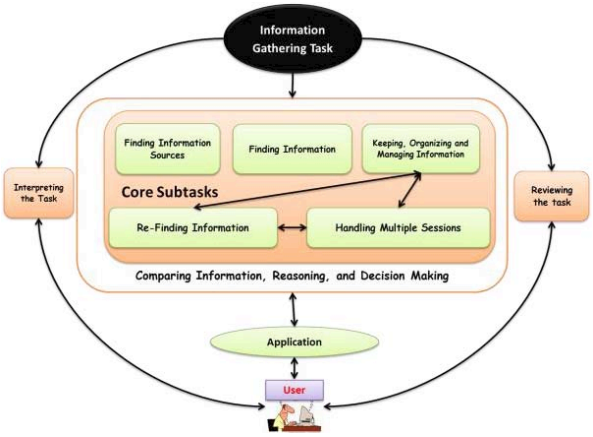


Figure 1. A Model of the Information Gathering Task

2.1. Core Subtasks

1. **Finding information sources.** This subtask involves activities intended to located websites and pages that have the potential of being considered for collecting information for the task.

2. **Finding information.** A continuation of the previous subtask is the subtask of finding actual information on web pages located. Information may involve parts of pages such as text, pictures, and so on.
3. **Keeping, organizing, and managing information.** This subtask involves preserving information organized as required in the task for: either working on the same task in subsequent sessions; or finalizing the task requirements. It also involves other managerial activities such as moving, copying, and editing objects.
4. **Re-finding the task information.** Activities regarding relocating information sources by revisiting links to web pages and sites comprise this subtask.
5. **Handling multiple sessions.** This subtask involves activities to manage the task information and context for restarting the task in subsequent sessions.

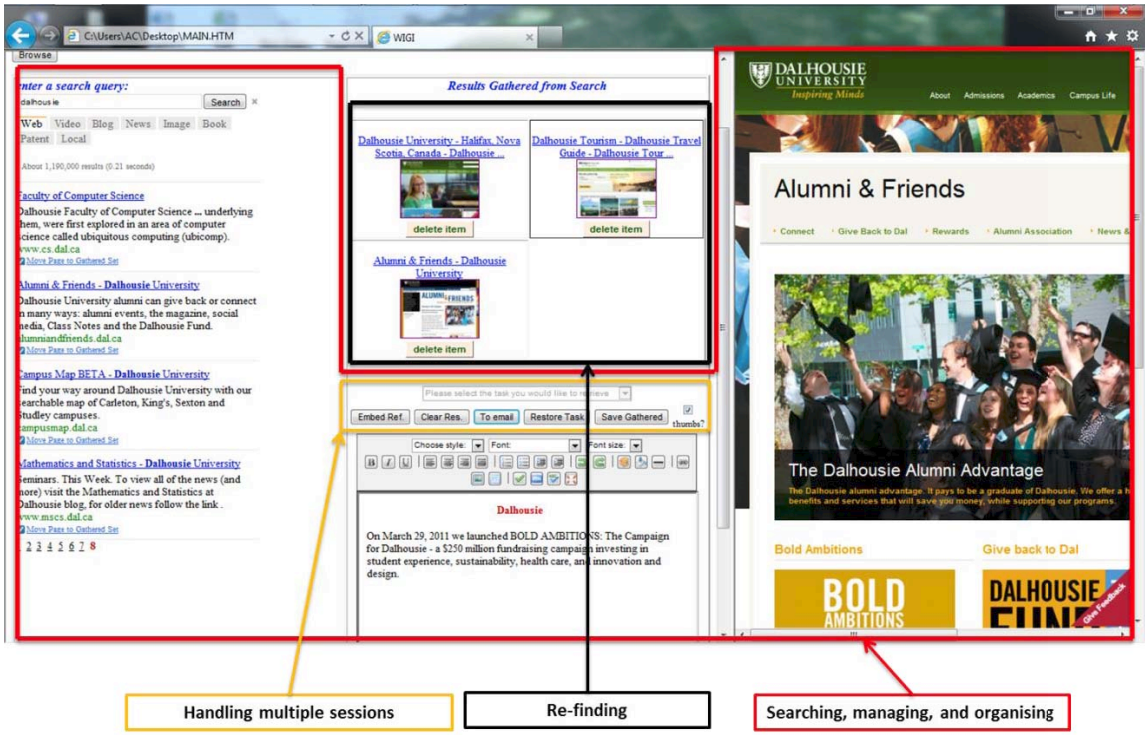
2.1. Other Subtasks

1. **Comparing information, Reasoning and Decision Making.** This subtask involves user behavioral activities that concerns comparing information sources, comparing information, and decision making for selecting information appropriate for the task at hand.
2. **Interpreting the task.** The interpretation of the task results in the choices of tools and kinds of information to gather for the task.
3. **Reviewing the task.** This subtask involves activities to ensure the completion of the task requirements or the session requirements in a multi-session task.

Following the proposal of the task model, the features of tools to support information gathering needed to be validated and best practices to be established to help users meet the challenges of this frequent task. The research discussed in this paper is an investigation of the effectiveness of recommendations for the design of features in web tools intended for gathering information from the web. The features developed and tested in a prototype are compared in the study to the use of a conventional web browser.

3. Research study

A prototype interface called WIGI (**W**eb **I**nformation **G**athering **I**nterface) was designed and implemented to investigate specific features for particular subtasks of the information gathering task as identified in a previous study. The recommendations developed in Alhenshiri, et al. (2012a) identified the following subtasks as highly relevant: re-finding information, handling multiple sessions, and managing and organizing information. WIGI, shown in Figure 2, consists of three main parts illustrated along with the features implemented as follows.



1- ***Re-finding Information: the Reference Tracking Area.***
   a. Users can keep track of every URL clicked in the search results.
   b. They can click each URL during any session within the same task.
   c. Links clicked from the search hit list are captured and shown to the user associated with the thumbnail as recommended in the works of Morgan and Wilson (2010) and Teevan et al. (2009).

2- ***Handling Multiple Sessions: the Control Bar.***
   a. Users can save the session information including: the tracked links, the information collected in the editor, and the links embedded as references as one integrated unit representing the task.
   b. Users can restart the task in subsequent sessions and have the information gathered in previous sessions retrieved as one integrated unit.

3- ***Searching, Managing and Organizing the Task Information: three panes within the browser window.***

   a. **The Embedded Editor**
      i. Users can drag and drop information from web pages into the editor.
      ii. They can add their input to the task using the editor.
      iii. Users can format the information in the editor as required in the task.
      iv. They can embed references into the information gathered in the editor.
   b. The Browsing Area
      i. Users can browse search results (pages) and typed in URLs on the same display along with editing, searching, and reference tracking.
   c. **The Search Area**
      i. Users can search the web for information using a search engine.
      ii. They can track every search hit clicked to appear in the reference tracking area.
      iii. They can browse search hits on one display along with the list of hits being viewed.

3.1. Study population and tasks

Thirty participants were recruited for the study. All of the participants were computer science students from Dalhousie University. Of the participants, 15 users were males and 15 were females. Fifteen participants were graduate students while the remaining were undergraduate students. Participants in the study were between the age of 18 and 30.

The study used four different information gathering tasks each of which had two parts (e.g. Task 1_a and Task 1_b). The reason for splitting each task into a sequence of two related parts was to provide a context in which participants might find some advantage in re-finding information for $Task1_b$ that was found or kept during $Task1_a$. The tasks were created using principles described in the work of Kules, et al., (2008). A focus group was used to ensure that the tasks are at the same level of complexity as described in the work of Alhenshiri, et al. (2012a). An example of one of those tasks is provided in Table 1.

Table 1. Task Example

Task (Part a). First Session	Task (Part b), Second Session
Somebody told you that the number of farms in Nova Scotia is decreasing. On the web, you decide to look for factors behind the drop in the number of farms in Nova Scotia. How accurate is what you heard? Provide links to at most five pages that contain information you found useful. Also, provide a copy of the actual information you located on those pages. You will need to come back to the information and web pages you found for this task.	Now, we would like you to compare the situation in Nova Scotia to that of the province of Ontario. On the web, find out about the status of farms in Ontario. Does the situation regarding the drop of the number of farms in Nova Scotia apply in Ontario? Why/Why not? Provide links to at most five web pages you found useful for comparing Nova Scotia to Ontario. Also, provide any information you find particularly helpful in the comparison.

3.2. Study design

The design of the study was complete factorial and counterbalanced. Four different tasks were used in the study (from the work of Alhenshiri, et al., 2012a). Every task had the same chance of being used in the study. The order of distributing the tasks over the tools (WIGI or browser) and participants was random. Every participant performed a total of two tasks with one task (divided into two parts that were nested and not crossed) executed on WIGI and one task on the ordinary browser. Both the browser and WIGI had the same chance of being used first. The browser used in the study was Internet Explorer (version 9). This browser was selected due to the need for using ActiveX components. The study had four conditions: two processes (browser + WIGI) and two tasks.

3.3. Study methodology

Each participant was randomly assigned two of the four tasks. The study was conducted over two sessions. On the first day of the study, each participant signed

the consent form after being introduced to the study and after explaining the participant's role in the study. Then, the participant was given a short training session on WIGI (five to ten minutes). The participant then completed an online pre-study questionnaire. After completing the questionnaire, the participant performed the first part of the first task on either WIGI or the browser. Then, the participant was given the first part of the second task to complete on the tool (WIGI or browser) the participant did not use for the first task.

Each participant returned to complete the second session of the study on the next day. First, the participant completed the second part of the first task on the same tool (WIGI or browser) they used for the first part of the first task. Following completing the post-task questionnaire for the first task, the participant completed the second part of the second task on the same tool they used for the first part of that task (which they completed in the first session). Afterwards, the participant completed a post-task questionnaire for the second task. Then, the participant was interviewed shortly to answer questions related to the way the participant completed the study with regard to why certain tools and strategies were used.

3.4. Study results

During the study, 5436 activities were logged. Of the activities, 2539 activities were logged while using the browser and 2897 activities were logged while using WIGI. The following is a description of the results with respect to each of the three subtasks considered in the study.

3.4.1. Re-finding information. The study logged activities related to re-finding information in the case of the browser and WIGI. The re-finding activity used in the analysis of the study data was revisiting references (links to web pages) to information accessed in the first session.

On WIGI, users made 185 (3.4% of the total activities) re-finding activities and on the browser nine re-finding activities (0.16% of the total activities). The difference between the two cases was significant using ANOVA ($F\ (1, 58) = 14.15, p < 0.0005$). The results also showed that those users who performed frequent re-finding activities on WIGI seldom revisited any links when using the browser. Consequently, the difference cannot be attributed to individual preferences. The average number of re-finding activities on WIGI was 6.17 (SD=8.51) and the average on the browser was 0.30 (SD=0.79).

When asked, participants indicated that when using the browser they did not attempt to re-visit links they kept in separate text files or emails because they had to copy the links from the files or emails and paste them in the browser address line. This process would have involved switching between two applications in addition to the copying and pasting. On the other hand, WIGI users had the links tracked and accumulated automatically in the reference tracking area as thumbnails that were visible and clickable at any time. Thus, it was easier to revisit links to view content in the browsing area.

3.4.2. Handling multiple sessions. Every user used the *save gathered* feature of the WIGI system control bar to save the state of a session for recall later in the next session. During the second session, every user restarted the task using the *retrieve task* feature of the WIGI system. None of the users used any files or emails to handle multiple sessions with WIGI.

On the browser, users used four different strategies to handle multiple sessions. Twenty six participants (26/30) created text files (using either MSWord or Notepad) to keep the task information and restart the task in the subsequent session. Four users (4/30) created 15 bookmarks. However, the same users reopened the bookmarks they created only 10 times. Four users (4/30) created email drafts to keep the information for subsequent sessions. Two users (2/30) saved complete pages to be used in the second sessions. Interestingly, neither of those two users re-opened the pages they saved.

The difference between the number of users who used the *save gathered* feature in WIGI (30/30) and the number of users who used text files to keep the task information in the case of the browser (26/30) was significant (z-test, $z=2.15, p<0.04$). The text file was the most frequently used feature for saving information between sessions while using the browser.

3.4.2. Searching, Managing and organizing information. The use of the search feature was not different between the systems but differences were found in how users managed and organized the task information. The participants used different strategies on each tool (WIGI or browser). The following is a comparison of the most frequent managing and organizing strategies found in the study.

1. Copying and Pasting Information

An important activity related to managing and organizing in information gathering tasks is copying and pasting information during the task for later use. The study logged copying and pasting information from web pages into an editable space where the user could collected relevant information during the task. Since some copying activities were not completed by the user (no subsequent pasting took place), the study

counted the successful pasting activities. The study recorded 330 pasting activities (6.01% of the total activities).

On WIGI, users performed more copying and pasting than they did on the browser (214 times vs. 116 times). The average number of pasting activities on WIGI was 7.13 (SD=5.84) while it was 3.87 (SD=4.75) on the browser. The ANOVA test results ($F(1,58)= 5.6, p<0.02$) show a significant difference between the number of pasting activities on WIGI and the number of pasting activities in the case of using the browser. It is worth noting that in the case of using the browser, the user had to copy information from web pages into a separate application such a text editor or an email draft.

2. Typing Information

While gathering information, not only do users copy information from web sources, but they often provide their own input and perform re-phrasing such as is needed to produce notes or even a report. Using WIGI, the participants did not type in information as frequently as they did in the case of using the browser. A total of 198 typing activities occurred while using the browser (Mean=6.6, SD=5.06). The total number of typing activities while using the browser was 292 (Mean=9.73, SD=6.44). The ANOVA test results ($F (1, 58) =4.40, p<0.05$) showed a significant difference between the number of typing activities on WIGI and those performed on the browser. This seems to be related to the relative ease in the WIGI case of dragging text from a web page directly into the editing area.

In addition, the data indicate that the difference between the use of WIGI and the browser with respect to the typing activities cannot be attributed to individual differences. Individual users behaved differently using the two systems indicating that the difference may be attributed to the system used.

3. Formatting Information

A relatively minor activity was the actual formatting in the editable space or file of information collected during the tasks, which included headings in the text, fonts and colors, moving objects within the gathered information (within a file, an email draft… etc.), and resizing objects such as images. There were 572 formatting activities logged during the gathering process. While not a dominant activity, formatting as an activity is closely related to the substance of information gathering as an activity that most often results in a written result.

On WIGI, users used the formatting features in the embedded editor built into WIGI to format the task information 521 times (9.58% of the total activities, Mean=17.37, SD=17.63). The embedded editor provided several formatting features such as font formatting, tables, and image formatting. On the browser—using other applications—users performed a total of 51 formatting activities (0.93% of the total activities, Mean=1.7, SD=2.88) during the study. The difference between the number of formatting activities on WIGI and the browser was statistically significant according to ANOVA ($F (1, 58) =23.08, p<0.0001$).

The correlation between the activities of typing and formatting was considered to see whether or not users who typed in more information (as opposed to pasting) did more formatting. The results of the Pearson Product Moment correlation test showed that in the case of WIGI, the correlation was not significant ($r = -0.11, p=0.53$) while it was significant and positive in the case of the browser ($r = 0.77, p<0.0001$). This indicates that in the case of WIGI, users who did not type much in the first place also did not perform much formatting since WIGI allowed them to copy and paste the information with its original formatting (as later explained by the users). In the case of the browser, however, the correlation explains that as users did more typing, they followed with more formatting.

3.5. A comparison of the number of actions on WIGI and the browser

The number of steps (actions) required to perform subtasks using WIGI was compared to those taken to complete the same activity using the browser. Substantial differences were observed. For example, users needed only one click to re-find a page using the thumbnail view in the reference tracking area compared to multiple steps on the browser unless the page was already open. To handle multiple sessions on WIGI users simply used the *save gathered/retrieve task* feature to save the state of the session information and to retrieve that information to restart the task. The participants employed a variety of strategies to cope with multiple sessions in the case of the browser. For managing and organizing information being collected by the user on WIGI, the user could simply drag the information from any web page to copy and paste it to the managing window. On the browser, however, the user typically needed to explicitly "copy" and "paste" into an external application, such as a word processor or email.

4. Discussion

The study revealed interesting differences in the way the participants performed the three subtasks identified for the study; re-finding information, handling multiple sessions, and managing information.

Re-finding information was more effective on the prototype interface compared to the strategies participants used on the browser. Reference tracking through the use of thumbnails, which included an image, the URL, and the page title, was used significantly more than keeping track of links using the browser, by copying and pasting links into text files and emails or bookmarks. All users rated the reference tracking feature embedded in WIGI as effective.

Handling multiple sessions was more effective on WIGI than on the browser. The *save gathered* and *retrieve task* features prevented the loss of information over multiple sessions. Of the participants, 20% (5 participants) lost the task information from the previous session while using the browser which never occurred when using WIGI. These *save gathered* and *retrieve task* features were the only strategies participants used handle multiple sessions on WIGI. On the browser, participants made use of bookmarks for keeping links to web pages yet never came back to open 33% of the bookmarks they had created. Furthermore significantly more use was made of the *save gathered* and *retrieve task* features of WIGI than any strategy employed on the browser including the use of text files, the most frequent strategy used for handling multiple sessions on the browser.

Managing and organizing information was more effective using the prototype WIGI than using the browser and other complementary applications. Having the embedded editor along with the search and browsing areas in one display lead to significantly more saving of information on WIGI than on the browser. Saving information is a core activity for information gathering as is managing that data. Moreover, users of WIGI performed significantly more formatting activities to manage and organize the information for the task than they did in the case of using the browser. At the same time, there was significantly less actual typing on WIGI than in the case of using the browser. The difference between the typing activities could not be attributed to the users but rather to the system used. Twenty nine users (96.7%) rated the ability to edit, format, search, and browse the information in one display as effective. They indicated that this feature eliminated the need for switching among different applications. Twenty five participants (83.4%) rated the ability to embed references into the editor as effective.

5. Conclusion

This paper discussed a study to investigate particular features to improve the user experience of information gathering on the web that was based on the recommendations of an earlier user study. The recommendations focused on three subtasks: re-finding information, handling multiple sessions, and managing and organizing the task information. The study showed that the features designed to respond to those recommendations achieved significant improvements over the ordinary web browser and other tools used to complete the task.

Re-finding by revisiting links from a previous session for the task was enhanced by using thumbnails to keep visual tracks of pages opened. Handling multiple sessions was improved by automatically saving the state of the current session including all information collected by the user thus eliminating the loss of information in subsequent sessions. Managing and organizing information was improved by providing a single view of the task; combining access to searching, browsing, re-finding and editing features. That would lead to eliminating switching among tools.

Among the limitations of the study is that all participants were computer science students. This kind of users does not reflect all information gatherers on the web. In addition, the study used simulated tasks that may not reflect all possible realistic information gathering tasks.

The results of this study support the recommendations proposed in our earlier study and provide some validation for our model of information gathering task. Our results also provide guidance in the design of applications and browser tools to better support information gathering tasks for users.

Further research will require a larger field study to continue to investigate user strategies in more complex contexts including using small-screen devices, which are projected to surpass traditional personal computer projects by a ratio of 4-to-1 by 2015 (http://www.vertic.com).

The problems associated with accomplishing a complex task on the web—such as information gathering—have not yet been fully investigated.

6. References

1. Ahn, J., Brusilovsky, P., He, D., Grady, J., and Li, Q. "Personalized Web Exploration with Task Models." *The 2008 WWW Conference*, Beijing, China, April,, 2008, 1-10.
2. Alhenshiri, A., M. Shepherd, C. Watters, and J. Duffy. "Web Information Gathering Tasks: a Framework and Research Agenda." *The International Conference on Knowledge Discovery and Information Retrieval (KDIR2010)*. Valencia, Spain: Springer, 2010b.
3. Alhenshiri, A., Watters, C., Shepherd, M., Duffy, J. "Building Support for Web Information Gathering Tasks." *The 45th Hawaii International Conference on System Sciences*. Grand Wailea, Maui, Hawaii, 2012a. 1687-1696.

4. Alhenshiri, A., Watters, C., Shepherd, M., Duffy, J. "Investigating Web Information Gathering Tasks." *The ASIS&T75 annual meeting.* Baltimore,MD, 2012b. 26-30.
5. Amin, A. "Establishing Requirements for Information Gathering Tasks." *TCDL Bulletin of IEEE Technical Committee on Digital Libraries* 5, no. 2 (2009): 1-10.
6. Bagchi, A., and G. Lahoti. "Relating Web Pages to Enable Information-Gathering Tasks." *ACM Conference on Hypertext and Hypermedia.* Torino, Italy, 2009. 100-118.
7. Broder, A. "A Taxonomy of Web Search." *ACM SIGIR Forum* 36, no. 2 (2002): 2-10.
8. Choo, C. W., B. Detlor, and D. Turnbull. "A Behavioral Model of Information Seeking on the Web--Preliminary Results of a Study of How Managers and IT Specialists Use the Web." *The Annual Meeting of the American Society for Information Science*. Pittsburgh, PA, USA: ASIS, 1998. 25-29.
9. Dearman, D., M. Kellar, and K. N. Truong. "An Examination of Daily Information Needs and Sharing Opportunities." *2008 ACM Conference on Computer Supported Cooperative Work*. San Diego, CA, USA, 2008. 679-688.
10. Ellis, D. "A Behavioural Approach to Information Retrieval System Design." Journal of Documentation 45, no. 3 (1989): 171-212.
11. Kellar, M., C. Watters, and M. Shepherd. "A Field Study Characterizing Web-based Information-Seeking Tasks." *Journal of the Amercian Society for Information Science and Technology* 58, no. 7 (2007): 999-1018.
12. Kules, B., W. Wilson, M. C. Schrafel, and B. Sheiderman. "From Keyword Search to Exploration: How Result Visualization Aids Discovery on the Web." School of Electronics and Computer Science, University of Southampton, Southampton, UK, 2008.
13. Mackay, B., and C. Watters. "Exploring Multi-session Web Tasks." 2008 ACM Conference on Human Factors in Computing Systems. Florence, Italy, 2008. 4273-4278.
14. Morgan, R., and Wilson, M. L., "The Revisit Rack: Grouping Web Search Thumbnails for Optimal Visual Recognition." *The Annual Meeting of the American Society for Information Science*, Pittsburgh, PA, USA, 2010. Vol 47, Issue 1, 1-4.
15. Murphy, J. "Information-Seeking Habits of Environmental Scientists: A Study of Interdisciplinary Scientists at the Environmental Protection Agency in Research Triangle Park, North Carolina." *Issues in Science and Technology Librarianship*, 2003: No. 38.
16. Sellen, A. J., R. Murphy, and K. L. Shaw. "How Knowledge Workers Use the Web." *the SIGCHI Conference on Human factors in Computing Systems.* Minneapolis, Minnesota, USA: ACM, 2002. 227-234.
17. Tao, X., and Y. Li. "Concept-Based, Personalized Web Information Gathering: A Survey." *3rd International Conference on Knowledge Science, Engineering, and Management.* Vienna, Austria, 2009. 215-228.
18. Tao, X., "A Personalized Ontology Model for Web Information Gathering." *IEEE transactions on knowledge and date mining*. Vol 23, issue 4, April, 2011, 496-511.
19. Taucher, L., and S. Greenberg. "How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems." *International Journal of Human Computer Studies* 47, no. 1 (1997): 97-138.
20. Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., Ramos, G., Andre, P., and Hu, C. "Visual Snippets: Summarizing Web Pages for Search and Revisitation." 27th International Conference on Human Factors in Computing Systems. Boston, MA: ACM, 2009. 2023-2032.
21. Yamada, S., and H. Kawano. "Information Gathering and Searching Approaches on the Web." *New Generation Computing* 19, no. 2 (2009): 195-208.
22. Zelberstein, S., and V. Lesser. "Intelligent Information Gathering Using Decision Models." Computer Science Department, University of Massachusetts, Boston, Massachusetts, 1996.