

The Advantages of a Cross-Session Web Workspace

Natalie Jhaveri and Kari-Jouko Räihä

Tampere Unit for Computer-Human Interaction (TAUCHI)

Department of Computer Sciences, FIN-33014 University of Tampere

{Natalie.Jhaveri, Kari-Jouko.Raiha}@cs.uta.fi

ABSTRACT

Conducting research using the web is often an iterative process of collecting, comparing and contrasting information. Not surprisingly, web-based research tasks habitually span multiple web sessions and involve considerable web page revisitation. Such tasks are not only carried out by researchers, but also by casual web users who, for example, plan vacations and large purchases. Despite the prominence of this activity among web users, existing tools support it poorly. We propose an alternative approach, whereby web-based research tasks are facilitated by a web workspace which represents collected URLs with web page thumbnails. A prototype of our design was developed and studied in an evaluation with 12 participants. Each of the participants adopted the workspace approach instinctively: the workspace was used for web page revisitation, web page comparison, collection overview, cross-session task continuation, and continuous task focus.

Categories & Subject Descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces — user-centered design, graphical user interfaces (GUI), prototyping, evaluation/methodology; H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia — navigation, user issues

General Terms: Human Factors; Design; Experimentation

Keywords: web workspace; web page revisitation; information gathering; thumbnails

INTRODUCTION

The World Wide Web (web) is used for a great variety of pursuits. The traditional notion of web navigation hardly represents the full extent of activities carried out on the web today. For instance, one study on the web activities of knowledge workers, those being individuals who are employed to transform knowledge, found that research-motivated *information gathering* was a more frequent web activity than undirected *web browsing*, which included activities such as browsing online news [12]. Moreover, it was observed that information gathering tasks, which

“involved comparing and contrasting information”, were “often unfolded over time, sometimes spread out over days or even weeks” [12]. Clearly returning to previously visited web pages, known as the act of web page revisitation, plays an important role in information gathering. Past research, which shows web page revisitation to be the most common user action in web navigation [5, 14], is consistent with this suggestion. Despite the great importance of this activity, research has shown on many accounts that available revisitation tools bear a number of limitations [1, 5, 6, 15]. Regardless of these indications, revisitation facilities of common web browsers, namely the Back button, the History list, and the Bookmarks facility, have persisted with little change for more than a decade.

According to the information foraging theory, information seekers will, over time, optimize their strategies to gather maximum quality information with minimum effort [10]. Indeed, studies have shown that more experienced web users employ a number of imaginative strategies to overcome the shortcomings of existing revisitation tools [8, 12, 15]. Whether less experienced users have also adapted creative revisitation strategies was investigated with a 34-respondent questionnaire [7]. It was determined that *casual web users*, those whose profession does not rely on heavy use of the web, have also adopted varied revisitation strategies. Though 8 different strategies were used by the respondents, search, bookmarks, and direct URL (Uniform Resource Locator) entry were the primary revisitation strategies. More importantly, it was confirmed that web-based research tasks are carried out either “sometimes” or “very often” by 94% of the respondents.

With the knowledge that web-based research tasks are commonly carried out by a wide group of web users by means of a number of creative revisitation techniques, some of which are not directly supported by common web browsers, it was our hypothesis that one integrated method may provide improved support for this task. As such, we investigated the use of a web workspace approach for web-based research tasks. A web workspace prototype tool, Session Highlights, was developed and evaluated over 2 sessions with 12 individuals. In this evaluation we aimed to identify the advantages gained with such an approach.

In this paper, after web page revisitation is discussed in further detail, the workspace approach, Session Highlights and the evaluation are presented.

WEB PAGE REVISITATION

Web page revisitation, the act of returning to a previously visited web page, is known to be a prevalent activity in web navigation [5, 14]. When considering activities such as daily visits to a favorite news site, this phenomenon is not surprising. However, web page revisitation within a web browsing session is seldom explicitly distinguished from web page revisitation occurring at some point after the session in which the page was initially visited. We will now make this distinction whereby the former will be called *in-session revisitation* and the latter, *post-session revisitation*.

Typical browsers offer the Back tool to support in-session revisitation and Bookmarks, the History tool, and the URL Auto-complete facility to support both revisitation types. Some of these systems are only effective in a limited few circumstances, while others are known to include clear shortcomings [1, 5, 6, 15]. The inadequacies of Back, Bookmarks, and History in particular have motivated the investigation of numerous alternative tools and approaches. For instance, WebView [6] and WebScout [9] each include in-session and post-session revisitation facilities. Among the users, however, it has motivated a shift towards the use of varied revisitation strategies [7, 8]. Within the scope of in-session revisitation, users attempt to preserve pages using multiple browser windows or multiple tabs within their browser window. Some browsers allow open tabs to persist across sessions, thereby adding post-session revisitation support to this in-session revisitation facility.

WEB WORKSPACE APPROACH

The *information workspace* concept [3], which represents an environment within which one can view, organize and store documents, is the basis of WebBooks and Web Forager [4]. WebBooks, interactive 3D books of web pages, can be stored in Web Forager, a 3D workspace. In fact, a more recent study suggested a specialized version of WebBooks as an information gathering solution [12].

The workspace approach to post-session revisitation support was applied in two subsequent tools as well, Data Mountain [11] and TopicShop [2]. Data Mountain entails a 3D virtual space for collecting web pages [11]. It allows users to place and rearrange web pages anywhere in the space. TopicShop allows users to gather, evaluate and organize web sites topically [2] and like Data Mountain, using spatial arrangement. In user evaluations, each exhibited improvements over the use of Bookmarks.

Our vision of such a workspace differs from these previous implementations. We envisioned a web workspace which, in contrast to WebBooks, Data Mountain and TopicShop, could be opened alongside the browser window when it is needed. Unlike WebView and WebScout, only user-selected URLs would be displayed. A URL would appear in the workspace as a thumbnail and, in contrast to Bookmarks, the user should not be concerned with matters of organization and management; the workspace simply enables the collection of URLs of interest for short-term

research support. Thus such a workspace would support the information gathering process, reserving other tools such as Bookmarks for the preservation of final key URLs.

SESSION HIGHLIGHTS

We designed and implemented a prototype tool named Session Highlights to support the web workspace approach (see Figure 1). The tool was mainly implemented in Java 2 and the Apache Batik SVG Toolkit (<http://xml.apache.org/batik/>) was used to generate the graphics in Scalable Vector Graphics (SVG) format. The prototype is browser-independent and Windows-dependent.

Interaction with the workspace is simple. To add a URL, the user must drag the Address icon (in the address/URL field of the browser) and drop it anywhere in the workspace. This gesture, as used in GroupBar [13], is fast and easy. Upon such a URL addition, a thumbnail image of the web page is added to the end of the chronological list. If necessary, the collection will automatically shift upwards to bring the newly added thumbnail into view. The user can scroll up and down the view using the mouse wheel. Using the Save button, the user can save a collection, open it later and continue building it.

When the mouse cursor is placed over a thumbnail, the thumbnail is enlarged and a tool text listing the web page title and URL is displayed, as shown in Figure 1 with the 4th thumbnail from the top. Clicking a thumbnail opens the associated web page in the most recently accessed browser window.

The layout of the thumbnails is motivated by the aim to include as many thumbnails as possible in view while still maintaining the chronological order of the thumbnails. Thumbnails are connected by a thin line to emphasize the chronological order. The prototype employs a static layout algorithm whereby small collections are presented in a straight vertical line, medium-sized collections are presented in a loose snake layout (as shown in Figure 1) and large collections are presented with a tighter snake layout and smaller thumbnails. In the future, the snake

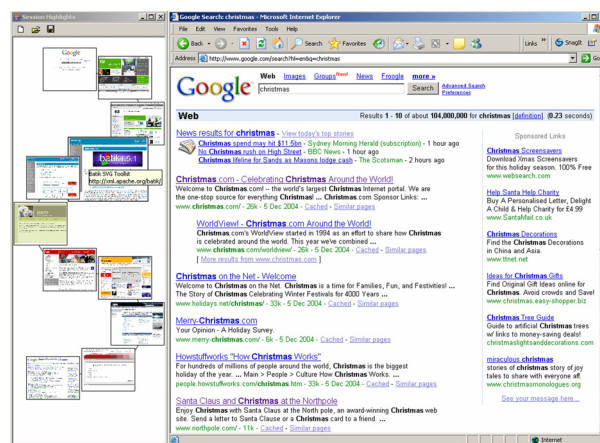


Figure 1: Session Highlights aside a browser window.

layout would dynamically “tighten” along with appropriate thumbnail size changes fitting all thumbnails into the view, but within reason.

Such thumbnail-based, cross-session facilities are included in the tabs of OmniWeb (<http://www.omnigroup.com/>); however, Session Highlights adds to this set a spatial layout and thumbnail enhancements on mouseover.

USER STUDY

Participants

Twelve university students of varied academic backgrounds were recruited for the test. Within this group, the average age was 26.2 (19, 36) years. The participants had 7.8 (6, 10) years of Internet experience and used the Internet for 2.7 (1, 6) hours per day. Among the participants, 5 used Internet Explorer as their primary web browser, 3 used Mozilla or Firefox, 3 used Opera, and 1 used Safari.

Apparatus

The study was carried out in a usability lab on a PC running Windows XP with a 19 inch CRT monitor. The participants used Internet Explorer 6.0 to navigate the web. The screen resolution was 1152 by 864 pixels. The collection tool spanned 303 pixels in width, occupying about 26% of the screen as shown in Figure 1.

A stopwatch was used to time the web sessions. The observations and comments were written down on paper throughout the sessions.

Procedure

The evaluation included 2 sessions on 2 separate days. The participants’ interactions with Session Highlights were logged and each session was screen captured as video.

After the standard consent procedure, the participant was presented with a written instruction set. First, the instruction set detailed the task: To plan a weekend-long trip to Stockholm, considering all of the necessary details of transportation, accommodation and possible activities. Then, Session Highlights was introduced with “The web browser companion tool, Session Highlights is available for collecting web pages of interest to you”, followed by a short description of how it is used.

After the participant was finished reviewing the instructions, the task was summarized verbally and then the participant’s questions were answered. Next, the participant tried using Session Highlights by adding two URLs and then selecting one to open in the browser window, as told in the instruction set. After this short practice, the participant created a new workspace and began carrying out the task. After 30 minutes, the participant was asked to stop and save his or her workspace if it wasn’t already saved.

On a separate day later, ranging from either the next day to 1 week later, the participant returned for the second session. The verbal instruction given then was to “continue planning your trip, keeping in mind that the trip is now approaching

and you must finalize your plans.” The participant carried out this task continuation for 10 minutes. Upon the completion of the 10 minute period, the participant was instructed to return to two pages described to the participant in words. The pages had been selected in advance from the previous session’s log file: the first two distinctive and describable web pages (e.g. “Lonely Planet’s homepage”) were selected.

Finally, the participant was asked to complete a questionnaire addressing background information and subjective impressions.

Results

All 12 participants used the tool to collect URLs. On average, 12 URLs (5, 18) were collected during Session 1. In the second session, 5 of the 12 participants added more URLs (1, 5) to the collections made in Session 1. By the end of both sessions, the total average number of URLs collected amounted to 13 URLs (6, 22).

A multiple collection approach (i.e. multiple files) was employed by 3 participants when carrying out the task. In all 3 cases, the collections were used to create some meaningful grouping of the web page URLs.

Participants were browsing at an approximate average rate of 2 to 3 pages per minute. The collection rates, meaning the number of pages collected per number of pages visited, ranged between roughly 8% and 16% during Session 1.

The participants collected a number of interesting pages such as, search result lists and currency converters. Such pages, which are not usually considered a main outcome of a research task, were used for intermediary task assistance.

In-Session Revisitation

During the first session, 11 participants used the URLs collected for in-session revisitation. Excluding the one exception, the thumbnails were used for in-session revisitation on average 12.4 times (5, 22).

Post-Session Revisitation

At the end of Session 2, each participant was instructed to carry out 2 revisitations to pre-selected URLs collected during Session 1. The success (the completion time and number of mouseovers and clicks to load the correct page) of the revisitation was partly affected by whether the participant had many similar thumbnails, whether the participant had used the workspace to access these URLs often during Session 2, and whether multiple collections were being used. These issues aside, out of the 24 instructed revisitations, 14 were ideal revisitations whereby the mouse moved directly to the thumbnail which was clicked. In addition, 3 were carried out by moving the mouse directly to the thumbnail, viewing the enlarged view and tool tip and then clicking it. Five involved viewing between 2 to 4 different thumbnails with mouseover and then clicking the correct thumbnail. And 2 involved some mouseovers, clicking the wrong thumbnail (of the correct

site but wrong page), and then loading the correct page upon the second thumbnail click. Therefore 22 revisitations were successful with the first click and the remaining 2 required 2 clicks.

Discussion

Clearly, when web users are provided with an easy means for both in-session and post-session revisitation in the context of a research task, they make substantial use of the opportunity. We observed that revisitations facilitated by Session Highlights were carried out: 1) to return to the key hub page to take on a new route when the current trail was uninteresting, 2) to compare two pages (e.g. prices of two different hotels), and 3) for obtaining additional information related to the current page (e.g. to access a currency converter page to check prices on a current page).

When the workspace was partly or even mostly covered by secondary browser windows, participants instinctively dropped new URLs on the portion still in view. When using the thumbnails, the participants were surprised with their own web page recognition abilities. Additionally, the snake layout was considered useful as it “helped to remember the pages by their location in the ‘snake’.” Many participants also commented on their desire to delete, rearrange and annotate the thumbnails in such a workspace.

The data, observations and subjective findings of this evaluation revealed the following advantages of the Session Highlights web workspace approach:

1. web page recognition using thumbnails,
2. web page recognition using spatial and chronological layout,
3. one-click revisitation to web pages of collected URLs,
4. summary of the previous session’s findings allowing easy task continuation in the next session,
5. task-based collections with the possibility for multiple grouping, and
6. continuous task focus as attention was never shifted to URL collection management.

CONCLUSION

The Session Highlights web workspace approach introduced a combined solution for information gathering and revisitation tasks. In the 12-participant evaluation, the approach was instinctively taken into immediate and effective use by the participants.

ACKNOWLEDGEMENTS

We would like to thank the National Technology Agency of Finland for funding this work.

REFERENCES

1. Abrams, D., Baecker, R. and Chignell, M. Information archiving with bookmarks: Personal web space

construction and organization. In *Proc. CHI 1998*, ACM Press (1998), 41-48.

2. Amento, B., Terveen, L., Hill, W. and Hix, D. TopicShop: Enhanced support for evaluating and organizing collections of web sites. In *Proc. UIST 2000*, ACM Press (2000), 201-209.
3. Card, S. K., Robertson, G. G. and Mackinlay, J. D. The Information Visualizer, an information workspace. In *Proc. CHI 1991*, ACM Press (1991), 181-186.
4. Card, S. K., Robertson, G. G. and York, W. The WebBook and the Web Forager: An information workspace for the World-Wide Web. In *Proc. CHI 1996*, ACM Press (1996), 111-117.
5. Cockburn, A., Greenberg, S., Jones, S., McKenzie, B. and Molye, M. Improving web page revisitation: Analysis, design and evaluation. *IT&Society* 1, 3 (2003), 159-183.
6. Cockburn, A., Greenberg, S., McKenzie, B., Jasonsmith, M. and Kaasten, S. WebView: A graphical aid for revisiting web pages. In *Proc. OZCHI 1999*, 15-22.
7. Jhaveri, N. Intermediate and Post-Session Web Page Revisitation Techniques and Tools. M.Sc. Thesis. Dept. of Computer Sciences, University of Tampere, Finland (2004).
8. Jones, W., Bruce, H. and Dumais, S. Keeping found things found on the web. In *Proc. of the 10th International Conference on Information and Knowledge Management* (2001), 119-126.
9. Milic-Frayling, N., Sommerer, R. and Rodden, K. WebScout: Support for revisitation of web pages within a navigation session. In *Proc. of the 2003 IEEE/WIC International Conference on Web Intelligence*, 689-693.
10. Pirolli, P. and Card, S. K. Information foraging. *Psychological Review* 106, 4 (1999), 643-685.
11. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. and van Dantzich, M. Data Mountain: Using spatial memory for document management. In *Proc. UIST 1998*, ACM Press (1998), 153-162.
12. Sellen, A. J., Murphy, R. and Shaw, K. L. How knowledge workers use the web. In *Proc. CHI 2002*, ACM Press (2002), 227-234.
13. Smith, G., Baudisch, P., Robertson, G., Czerwinski, M., Meyers, B., Robbins, D., Horvitz, E. and Andrews, D. GroupBar: The TaskBar evolved. In *Proc. OZCHI 2003*, 34-43.
14. Tauscher, L. and Greenberg, S. How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies* 47, 1 (1997), 97-137.
15. Wen, J. Post-Valued Recall web pages: User disorientation hits the big time. *IT&Society* 1, 3 (2003), 184-194.