Intermediate and Post-Session Web Page Revisitation Techniques and Tools

Natalie Jhaveri

University of Tampere
Department of Computer Sciences
Interactive Technology
M.Sc. Thesis
September 2004

University of Tampere
Department of Computer Sciences
Interactive Technology
Natalie Jhaveri: Intermediate and Post-Session Web Page Revisitation Techniques and Tools
M.Sc. Thesis, 49 pages, 6 appendix pages
September 2004

Web page revisitation, the most commonly performed activity in web navigation, is the primary concern of this thesis. This subject is subdivided into two categories, intermediate revisitation, which occurs during the web browsing session and post-session revisitation, which takes place at some point after the session of initial visitation. The strategies and existing facilities belonging to each are discussed in detail.

A questionnaire assessing the post-session revisitation techniques of less experienced web users was conducted and the results were compared with previously reported techniques of experienced web users. The varied approaches of facilitating revisitation, coupled with the shortcomings of existing tools, motivate the design of our prototype revisitation tool, Session Highlights. The design of this tool is presented along with the results of an observational user study in which five participants carried out a web-based research task using the tool. The participants expressed great fondness for the tool's visual representation of URLs and for the availability of such a web workspace. Moreover, when carrying out their task, they were instantly able to use the tool efficiently and effectively to address their intermediate and post-session revisitation needs. Based on the observations and subjective results of the study, design improvements and subsequent features for Session Highlights are established.

Key words and terms: Web navigation, web page revisitation, web-based research, information visualization, thumbnails, information workspace.

Table of Contents

1. Introduction 1
2. Web Navigation and Web Page Revisitation 2
    2.1. History 2
    2.2. Present Situation 4
        2.2.1. Web Browsers 4
        2.2.2. Current Web Navigation Trends 5
        2.2.3. Behaviour of the Back Tool 6
3. Characterizing Revisitation 6
    3.1. Why Subdivide Revisitation? 7
    3.2. Intermediate Revisitation 8
        3.2.1. Multiple Window Navigation and Tabbed Browsing 10
        3.2.2. SmartBack 10
        3.2.3. Scratchpad 12
        3.2.4. Combined Revisitation Solutions 12
    3.3. Post-Session Revisitation 14
        3.3.1. History List 15
        3.3.2. Bookmarks 15
        3.3.3. Passive versus Active Strategies 16
        3.3.4. Past Research on Post-Session Revisitation Techniques 17
        3.3.5. Proposed Solutions 19
        3.3.6. Web Users 21
4. Empirical Study of Post-Session Revisitation 21
    4.1. Interviews With Casual Web Users 21
    4.2. Questionnaire 21
        4.2.1. Respondents 22
        4.2.2. Results 22
        4.2.3. Relevancy of the Post-Session Revisitation Task 24
        4.2.4. Bookmark Collections 24
        4.2.5. Knowledge Workers versus Casual Web Users 26
5. Prototype Revisitation Tool 26
    5.1. Conception of Session Highlights 26
    5.2. Design and Visual Layout 27
        5.2.1. Basic Design 27
        5.2.2. Thumbnails 28
        5.2.3. Snake Layout 29
        5.2.4. Interaction 31

5.3. Implementation Technologies 32
5.3.1. Scalable Vector Graphics 33
5.3.2. Apache Batik SVG Toolkit 33
5.4. Comparison with Existing Tools 33
6. Evaluation of Session Highlights 34
6.1. Method 34
6.1.1. Participants 35
6.1.2. Apparatus 35
6.2. Procedure 36
6.3. Results 37
6.3.1. Usage Scenarios, Selected Tasks, and Resulting Collections 37
6.3.2. General Observations 38
6.3.3. Working Set of Web Pages 39
6.3.4. Combined Intermediate and Post-Session Revisitation 39
6.3.5. Supported Web-based Information Retrieval 40
6.3.6. Overview of Document Type 40
6.4. Post-Usage Questionnaire 40
6.5. Discussion 41
7. Conclusions and Future Work 42

References 45

Appendices 50

1. Introduction

The World Wide Web (web) is used for a great variety of pursuits. In fact, the activities of web users these days are far more extensive than the term web navigation tends to portray. The web imparts its users with amenities of information, communication, entertainment and much more. Users thus engage in a great variety of web-based activities including, browsing, retrieving information, downloading media, emailing, etc. Such activities are facilitated by an assortment of tools and services, such as web browsers, search engines, ftp clients, peer-to-peer networks and web-based email clients. However, as the web's user group expands and the overall dependency on the web mounts, the task of supporting the needs of web users becomes far more challenging.

In the web's early years, its user population was mainly comprised of computer-savvy individuals. Today the user population of the web is far more diverse. A great variety of professions, ages and nationalities are represented among its users. Accordingly, a variety of web usage objectives and strategies have emerged. 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 among their participants than undirected web browsing, which included activities such as browsing online news [Sellen *et al.*, 2002]. Moreover, they observed that information gathering tasks, which "involved comparing and contrasting information", were "often unfolded over time, sometimes spread out over days or even weeks" [Sellen *et al.*, 2002].

Despite such changes, facilities supporting web use have undergone comparatively little adjustment. In fact, the basic functionalities of today's common web browsers bear little difference from those of ten years ago. The main facilities, namely, the Back button, the History list, and the Bookmark tool, have persisted for more than a decade. In addition, a web page is still most commonly represented by its title and/or Unified Resource Locator (URL), also known as its address. It has been shown on many accounts that the activity which is supported by these three web browser mechanisms, *web page revisitation*, is the most common user action in web navigation [Tauscher and Greenberg, 1997a; Cockburn et al., 2003]. In spite of the great importance of this activity, the tools available to support it are each known to include setbacks. Not surprisingly, more experienced users have been observed to employ a number of creative alternative techniques to address their revisitation needs.

It is the hypothesis of this thesis that web users are not currently provided adequate facilities for carrying out web-based research tasks and the associated revisitation. As a solution, a prototype tool, *Session Highlights*, which serves as a web workspace for lightweight web page collection, was designed and developed.

The background of this research includes two main themes: strategies of web page revisitation and support for web page revisitation. This terrain sets the context for the two main research problems addressed in this thesis:

1. Do users employ techniques to facilitate post-session web page revisitation when carrying out web-based research tasks?
2. Is the proposed visualization tool a viable solution for this revisitation need?

A questionnaire-based survey was conducted to consider the first research problem and an observational study and subjective evaluation was conducted to address the second. In addition to the presentation of these results, the following chapters will include a detailed discussion on web page revisitation and a description of our proposed solution, Session Highlights.

2. Web Navigation and Web Page Revisitation

The web offers its users an overwhelming amount of information. Navigating through this ever-expanding information space and finding specific information is undeniably a challenging task. Search engines have certainly served to ease the process of finding information. However, web navigation remains a rather problematic pursuit. This exact concern has motivated the Semantic Web initiative, a concept of a structurally-enhanced version of the current web, which can be better processed by computers and yield a more beneficial information source [W3C Semantic Web, 2004].

Web navigation includes both the acts of web page visitation and revisitation. *Web page visitation* describes the act of visiting a web page for the first time, while *web page revisitation* describes the act of returning to a web page that has been visited previously by the user [Xu *et al.*, 2001]. This thesis concentrates on the latter.

2.1. History

The disorientation and cognitive overhead associated with viewing and navigating hypermedia networks was already identified in 1989 [Utting and Yankelovich, 1989]. At that time, Utting and Yankelovich had already proposed their design of Web View, a network map of paths and documents. They maintained that with enhanced spatial and temporal context, the user is provided with orientation cues, thereby reducing cognitive overhead.

In 1995, the act of *backtracking*, which by definition was accomplished using the "Back Command", was noted to be a heavily used web navigation technique by Catledge and Pitkow [1995]. In fact, in their log analysis of web use, which included three weeks of web activity of 107 users, it accounted for 41% of all web browsing interactions. Thus, they were the first to highlight the importance of this activity in web navigation.

Within the same period, a surge of related research surfaced. Identifying the deficiencies of web browsers and proposing navigation history visualizations to reduce the disorientation associated with web navigation were common research endeavours. Domel [1994], for instance, presented WebMap, a two-dimensional graphical navigation map to provide users with contextual web navigation. The layout of the network-structured map changed depending on the topology information of the web session. For example, the system would change its layout from a traditional tree-type layout to a circular, more spider web-like layout to better illustrate the content of the map. By double-clicking on a node in the map, the user could open the corresponding web page in a Mosaic browser window.

Ayers and Stasko [1995] proposed MosaicG, their extension to the web browser Mosaic, which presents a visualization of the user's web traversal history. In the view, traversed web pages are depicted with thumbnail images and presented in a tree-structure. Their motivation was to assist users in finding their way back to previously viewed web pages.

Brown and Shillner [1995] presented DeckScape, an experimental browser offering navigation through a "deck" of collected web pages, as opposed to "depth-first navigational model" of other browsers.

In 1996, Cockburn and Jones [1996] discussed usability problems associated with web page revisitation and proposed a graphical browser named WebNet as a solution. The entire "web subspace" of the user's browsing session is displayed in a network-structured map. As the user navigates from page to page, the representation of the subspace is dynamically updated to show already visited pages, the current page, and possible next pages.

Despite these contributions, by 1997, basic facilities of web browsers had experienced very little change. A second log-analysis, including 5-6 weeks of web activity of 23 participants, was conducted by Tauscher and Greenberg [1997b]. In their analysis, they defined the *rate of web recurrence*, also known as the *web page revisitation rate*, as the "probability that any URL visited is a repeat of a previous visit, expressed as a percentage." This value was simply based on the total number of URLs visited less the total number of distinct URLs visited. With this new metric, they reported the rate of web page recurrence to be 58% for their own participants and using the data of the Catledge and Pitkow study [1995], found the recurrence rate of 61%. In their report, Tauscher and Greenberg also noted the tendency of web users to access only a few pages frequently as well as navigate within small clusters of related pages. They also identified a strong *recency effect* in users' revisitation patterns, meaning that the most recently visited pages are the same pages that are most likely to be visited next. Therefore, among their participants, they found that the Back button was used frequently, accounting for 30% of all navigation actions.

Thus, in the earlier years of the web, the importance of web page revisitation had been identified and confirmed. The period also experienced a surge of research endeavours to support the act of web navigation and web page revisitation with visualization. For instance, a variety of tools allowing web users visualize web subspaces emerged. A number of these tools are described and compared in a 1997 study, Cockburn and Jones [1997]. However, as many of them are based on old technologies, the results are not completely pertinent today.

2.2. Present Situation

Clearly the web and its associated technologies undergo numerous changes within relatively short periods of time. Web navigation trends and technologies of 10 years ago hardly apply to those of today. As such, it is critical to consider more recent web navigation trends.

2.2.1. Web Browsers

Web browsers and the facilities that they offer delimit the trends among web page revisitation. Thus it is necessary to consider the relative popularity of currently available web browsers.

The global distribution of web browser usage based on the estimates for June 2004 [Refsnes Data, 2004], is illustrated in Figure 1. Though such statistics can be distorted for a variety of reasons, they are relatively consistent with those presented by others. At the least, we can use these to conclude that Internet Explorer is by far the most popular web browser and the second most widespread is Mozilla. Neither of these web browsers is entirely new however; Internet Explorer has historical roots NCSA Mosaic, as does Netscape Navigator, upon which Mozilla is based. Undoubtedly, NCSA Mosaic and Netscape Navigator were each prevalent web browsers in previous eras of the web.

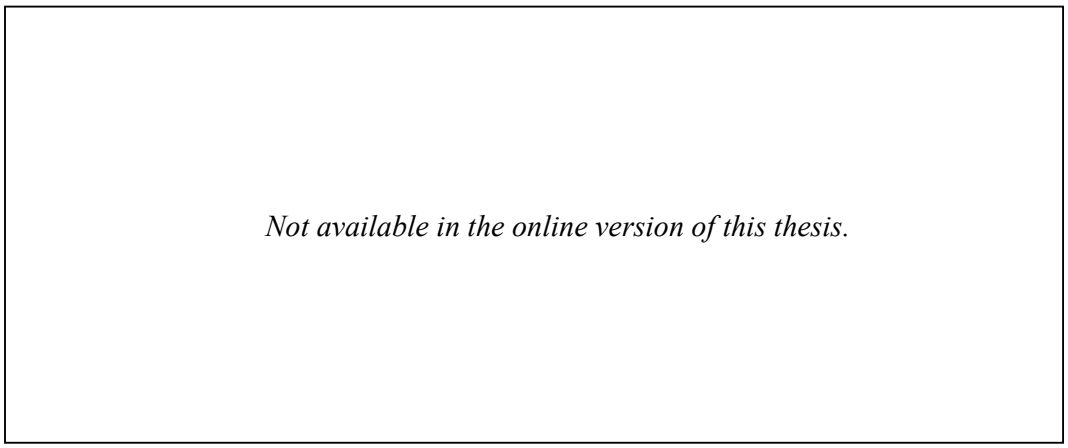

Figure 1. Web browser usage: Internet Explorer (IE), Opera (O), Mozilla (Moz), Netscape Navigator (NN).

Interestingly, compared to figures of the previous year (July 2003: *Not available in the online version of this thesis*), Internet Explorer usage has decreased, while Mozilla and Opera, relatively new web browsers are gaining popularity. Yet, unlike Internet Explorer, which comes installed by default with all Windows operating systems, Mozilla and Opera each require the user to download and install the tool independently. This shift may suggest that users are becoming more active in their pursuit for a better web navigation experience.

Indeed, Mozilla and Opera build upon the common default set of web browser features with enhancements such as *tabbed browsing*. Tabbed browsing is derived from the graphical user interface technique called tabbed document interface, in which the parent window contains the child windows as tabbed panes. With this facility, which is displayed in Figure 2, users are able to load multiple web pages in synchronous without creating multiple browser windows. Furthermore, as illustrated in the second tabbed pane of the figure, the user can view one page while the other is loading. As web page loading is known to be a time-consuming part of web navigation [Byrne *et al.*, 1999], tabbed browsing offers a method for more efficient web navigation, which is especially valuable in conditions with slow Internet connections.

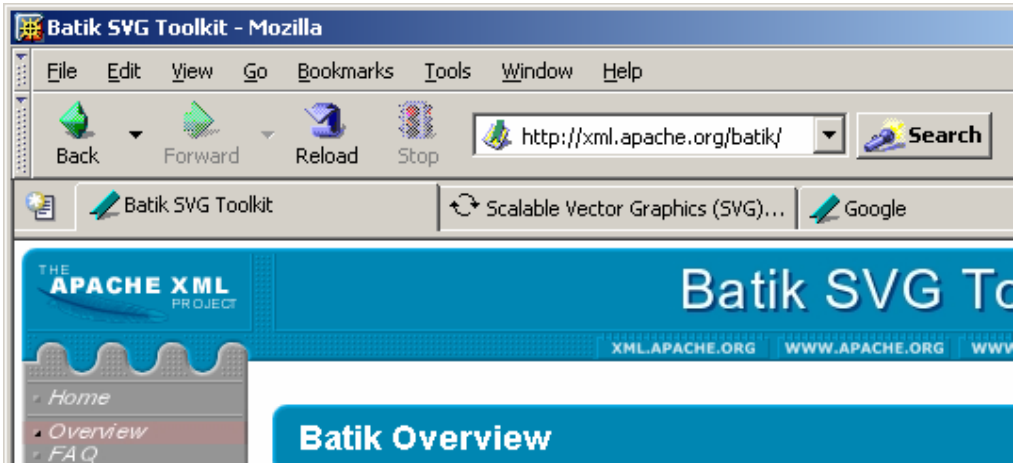

Figure 2. Mozilla web browser: Multiple web pages opened using tabbed browsing.

To simplify the discussion of web browsers and their facilities, only the most widely used web browsers, Internet Explorer and Mozilla, will be addressed when web browser facilities are discussed hereafter.

2.2.2. Current Web Navigation Trends

Web page revisitation is still a prevalent navigation activity. A relatively recent client-side log analysis conducted by Cockburn and McKenzie [2001] established that the act of web page revisitation is more heavily engaged in than earlier. Their 17-participant, 119-day data resulted in an average web page revisitation rate of 81%, with individual

participant rates ranging from 61% to 92%. This is a considerable increase from the 58% and 61% revisitation rates reported earlier [Tauscher and Greenberg, 1997a].

When one considers the fact that web sites are by nature dynamic, this phenomenon is not entirely surprising. Activities such as daily visits to news sites certainly contribute to the prevalence of web page revisitation. Certainly one would expect sufficient, if not exceptional, support for this evidently dominant activity in common web browsers. Unfortunately, this is not the case. One key revisitation tool, the Back tool, has been shown on many accounts to be misunderstood by web users [Cockburn *et al.*, 2003, Cockburn and Jones, 1996, Nielsen, 1990]. The Back tool, which includes both the Back button and the Back drop-down list, uses a somewhat problematic stack-based system. The problem associated with the stack-based system is described next.

2.2.3. Behaviour of the Back Tool

Consider the navigation scenario demonstrated in Figure 3 where each boxed-letter represents a web page and each arrow represents a traversal. In this scenario, the following navigation route is presented: A to B to A to C to D. Due to the stack behaviour of the Back tool, once the user returns from "B" to "A", the "hub page", and continues to "C", "B" is popped off the stack. Therefore, at this point, "B" can no longer be accessed using the Back tool.

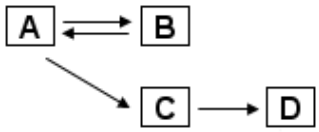

Figure 3. Illustration of a non-linear navigation route.

In many circumstances users can successfully navigate to a previously visited page using the Back tool. However, in cases when the stack-based behaviour takes effect and the user cannot return to the desired page, the user typically dismisses this failure as a "mysteriously lost page". Such users naively trust the Back tool without questioning the chain of events. This is because the Back tool is, outside of this one issue, a very easy, quick and effective tool [Greenberg and Cockburn, 1999]. In addition, because of the strong recency effect described earlier [Tauscher and Greenberg, 1997b], it is very likely that the user can access a desired page with only a few, cognitively undemanding clicks of the Back button. That is, assuming it still remains on the stack.

3. Characterizing Revisitation

As stated earlier, web page revisitation is defined as the act of returning to a web page that has been previously accessed by the user [Xu *et al.*, 2001]. 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 referred to as *intermediate revisitation* and the latter, *post-session revisitation*.

3.1. Why Subdivide Revisitation?

The reason for subdividing web page revisitation is twofold. The first justification lies in the divergent objectives served by the revisitation facilities of common web browsers. The second relates to the fact that the intentions, techniques and strategies of today's web users are different from those of the web users of ten years ago. This shift must also be reflected in the characterization of web page revisitation. These two reasons will be discussed in greater detail next.

The distinction between intermediate revisitation and post-session revisitation is necessary given that common web browsers provide different facilities for each. For example, the Back tool, which, as defined earlier, includes both the Back button and the drop-down list, only supports intermediate revisitation. After the user ends his or her session by closing the web browser window, the list of visited sites is cleared and no longer available for subsequent browsing sessions. For this reason, both Internet Explorer and Mozilla include a History tool which stores the user's browsing history over a longer period of time, thus supporting post-session revisitation. In Internet Explorer, a visited web page is represented by its domain name and URL in the History tool. In Mozilla, a web page is represented by its title, URL and the date when it was last visited. In both web browsers, the listed history can be presented in a variety of ways, such as sorted by date. In addition, the sites can be grouped, for example, by day, whereby the resulting groups may be "Today", "Yesterday", and "Last Week".

A mechanism related to the History tool, named Auto-complete, suggests visited URLs in a drop-down box below the URL or Address entry field when a URL is being typed in. These suggestions, which are dynamically proposed based on the letters being typed into the URL field, are retrieved from the History list. Therefore, when the user clears the history, Auto-complete will no longer have a list of URLs to suggest. When the History list is populated, however, this tool effectively supports both intermediate and post-session revisitation.

Web browsers also include a post-session revisitation tool that facilitates selective URL collection. The name of this tool varies between browsers however. It is referred to as Favorites in Internet Explorer, Bookmarks in Mozilla (and Netscape) and in Mosaic, it was called the Hotlist. For ease of reference, this tool, which is standard to all web browsers, will be referred to as *Bookmarks.* In addition, the act of adding a URL to a *Bookmark collection* will be referred to as *bookmarking*. Bookmarking allows user to collect URLs of interest and, with the use of personalized URL naming and folders for grouping, manage the collection.

Justifiably, Bookmarks, History and Auto-complete are all available within the same session when the page was initially visited; however, it is their persistence beyond that initial session which qualifies them as post-session revisitation mechanisms.

The distinction between intermediate and post-session revisitation is further necessary when considering changes in web navigation strategies. As recognized by Newfield, Sethi and Ryall [1998], to handle the immensity of information available on the web, users have shifted their web navigation strategies. Instead of engaging in *undirected web browsing*, as was often the case previously, more and more individuals use the web to perform *directed searches*. When a user carries out a directed search, his or her goal is to "extract information about a predetermined topic" [Newfield *et al.*, 1998]. Such a directed search can be, for example, a fact-finding task or a subject-based exploration. Indeed recent information retrieval research indicates the need for facilities of search history preservation [Aula and Käki, 2003]. Newfield, Sethi and Ryall further assert that the inadequacies of current web browsers are perhaps a result of their being designed to support the original main activity of web users, that being undirected web browsing.

Likewise, web page revisitation must also be considered outside the realm of undirected web browsing. From the perspective of directed searches, which may result in the collection of web links, revisitation can occur either during the current session or at some time later. A related issue, "web page recall" is discussed in a recent paper by Wen [2003]. He introduces the concept of Post-Valued Recall (PVR) as the "interest a user may have in recalling information—in this case, a Web page—whose value is not recognized until some time after its initial retrieval." He asserts that PVR can take place either during or after a web session.

To study the effects of PVR, Wen [2003] conducted a 12-participant study of ordinary web usage sessions. He selected PVR candidate pages based on the length of time the participant had spent on the page. He analyzed the amount of time it took for the participant to return to PVR pages as well as pages that they had bookmarked, or otherwise made note of during their web session. He concluded that PVR is a significant cause of user frustration and offered a set of design guidelines which address problems arising from PVR.

3.2. Intermediate Revisitation

As discussed, users are offered the Back tool to address their needs of intermediate revisitation. However, what do users do if the page is no longer available in the Back tool list? Do users employ strategies to hold onto "possibly useful" pages?

By means of observation and a quick show-of-hands survey in a class of approximately twenty, a generalization of strategies for taking note of an interesting web page was established. Users typically employ the following methods to deal with this dilemma:

1. bookmark the page, just in case it becomes of some interest later on,
2. use the Back button to retrace the sequence of web pages until the desired one appears,
3. use the Back drop-down list to select the name of the desired web page, or
4. open a new browser window for a new web link so that this interesting page is still left available in another browser window.

The first option can result in an extensive and hard to manage bookmark collection. The second and third are unreliable as the page may have been popped off the stack-based list, and the last, though least problematic, can result in an unmanageable amount of open windows. Obviously one clear solution is not available.

Intermediate revisitation obstacles, such as that which was just recounted, have spawned the development of a variety of web navigation techniques, such as gesture-based navigation [Moyle and Cockburn, 2003; Opera Software 2004], tabbed browsing, and breadth-first navigation facilities [Newfield *et al.*, 1998], included along with the conventional depth-first traversal mechanisms. Moreover, it has motivated the investigation of temporal revisitation systems, which, in contrast to the current stack-based system, maintain a complete, chronological list of all visited web pages.

A study conducted by Cockburn *et al.* [2002], in which a temporal-based system was evaluated against the Back tool, concluded that "the relative efficiency of the interfaces differ across different types of navigational tasks." The temporal system was less efficient than the Back tool for returning to parent pages, such as page "A" in Figure 3 presented earlier. However, for larger traversals, the temporal system was highly inefficient when used with the Back button and highly efficient when used with the Back menu. In the case of the Back button, the user must click the button once for every backward traversal and thus requires many clicks to return to a web page visited many pages in the past. In contrast, the Back menu was enhanced by the temporal system because it provided direct page access. Based on the findings, Cockburn *et al.* [2003] suggested that "web page revisitation could be made more efficient, if browsers encouraged users to make greater use of direct page access methods such as the Back menu, rather than relying on multiple clicks of the Back button." Web browsers do include a mechanism that stores visited pages temporally: the History list. This tool, which qualifies as a post-session revisitation tool, will be discussed in the next section however.

In the earlier years of the web, numerous web navigation history visualization tools were developed to provide users with both an overview of their navigation history and a means for direct page access. More recently, a number of alternative tools and techniques have surfaced to address the problems associated with intermediate revisitation and these will be presented next.

3.2.1. Multiple Window Navigation and Tabbed Browsing

Even if users do not know the exact reason behind it, they are at least aware that a visited web page will not necessarily be accessible using the Back tool. This uncertainty has perhaps instigated the use of the *multiple window navigation* technique. This technique involves opening a web link in a new window, thereby preserving the previous page in the initial window. Though this technique comes with the drawback of having several windows open towards the end of the session, it is certainly a more reliable technique for guarding pages of interest. Numerous search engines have reacted to this trend by providing an option to open a search result in a new browser window.

This technique is so widespread that several current web browsers, such as Mozilla, Opera, and Netscape Navigator 7.2, include an enhanced facility for multiple window navigation, tabbed browsing. As mentioned earlier, this feature allows users to open multiple web pages as separate tabbed panes within one parent web browser window. One Mac OS X web browser, OmniWeb, shown in Figure 4, represents tabs with thumbnails of corresponding web pages. As it is known that users are able to identify web pages more accurately from thumbnail images than from titles or URLs [Kaasten *et al.*, 2002], this is likely a beneficial improvement over the more traditional use of web page titles in tabs.

Figure 4. Tabbed Browsing in OmniWeb: Tabs (left) include both web page thumbnails and titles [OmniGroup, 2004].

3.2.2. SmartBack

One study suggesting a supplement to the Back button to better facilitate intermediate revisitation proposes SmartBack [Milic-Frayling *et al.*, 2004]. This feature marks certain key pages from the user's navigation history and provides quick access points

for their re-access. The key pages that are marked include: "(1) hubs, (2) results of form submissions, including search result pages, and (3) beginnings of navigation trails; such as typed-in URLs, bookmarks or links from similar features that help the user access pages on the Web."

In their study two SmartBack designs are discussed: *SmartBack button* and *SmartBack Link Bar*. The SmartBack button, which is illustrated in Figure 5, stores a list of all the key pages in sequential order. The SmartBack Link Bar, on the other hand, displays the last 4 to 5 pages as selectable link buttons in the link bar, that being on the same level as the Back button in Internet Explorer. When the mouse cursor is placed over the link, a thumbnail image of the page appears and clicking the link opens it in the browser window. In both cases, as shown in the Back drop-down list in Figure 5, the key pages are distinguished with icons, a green SmartBack icon for hub pages, a magnifying glass icon for search result pages, and an "ABC" icon for typed-in URLs.

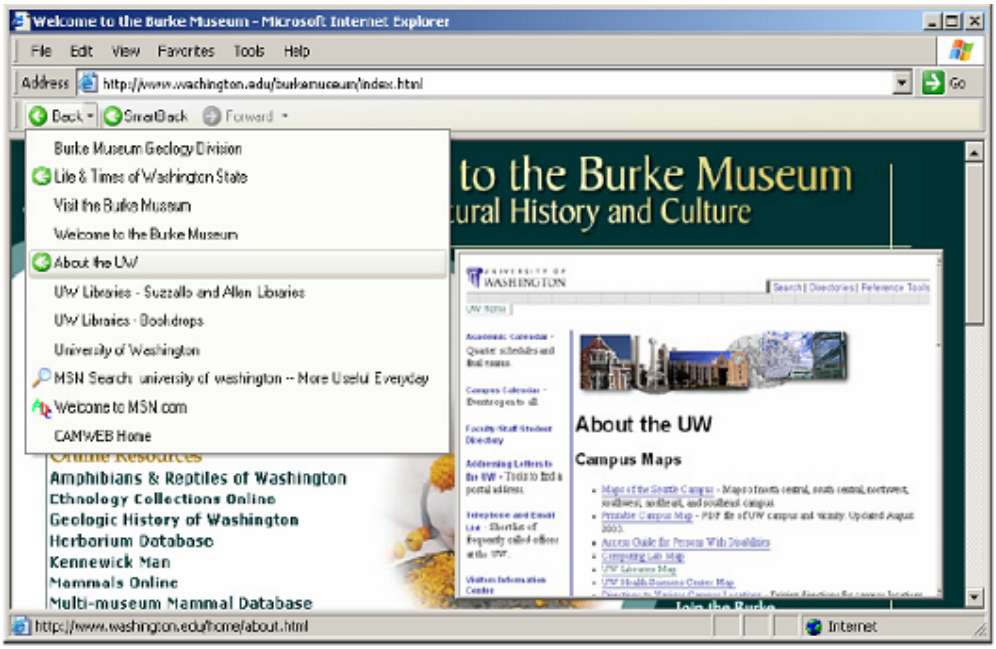

Figure 5: Back menu in the SmartBack interface: Icon-enhanced links (left) and thumbnail appearance on link mouse over (right) [Milic-Frayling *et al*., 2004].

Milic-Frayling *et al*. [2004] conducted a user study that was set in a workplace with 7 employees and in a home with 4 family members. They found that participants adopted the tool as a means for accessing hub pages. However, they found that learnability improvements are still required for the SmartBack design. They also discovered the need to investigate a more engaging version of SmartBack, in which users would be able to select their SmartBack targets themselves for future re-access.

3.2.3. Scratchpad

Newfield, Sethi and Ryall [1998] also developed a prototype with the goal of "making hypertext navigation faster" and "easier". Their prototype, Scratchpad, promotes an alternative web navigation strategy, *breadth-first navigation*. They remark that current web browsing facilities, which endorse depth-first traversal, permit the well-known stack-based behaviour problems. As shown in Figure 6, Scratchpad reveals the next available web pages in a Pending Links list, thus enabling breadth-first navigation. More importantly, this tool includes the concept of temporary bookmarks. This facility, called Dogears, enables easy revisitation to *dogeared* (*i.e.,* marked) web pages. To dogear a certain page, the user clicks a button at the top of the browser window. As "dogears do not persist between sessions and are intended to act as lightweight bookmarks" [Newfield *et al*., 1998], users can bookmark without being concerned with flooding their collection.

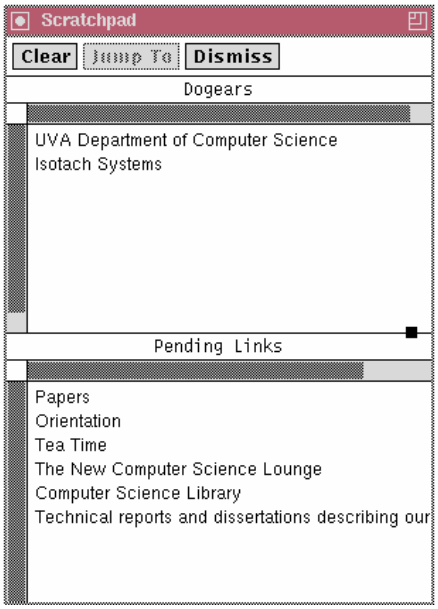

Figure 6. Scratchpad showing two visited pages marked by the user and six links automatically listed as possible next web pages [Newfield *et al.*, 1998].

3.2.4. Combined Revisitation Solutions

One integrated intermediate and post-session revisitation solution is also offered among some tools. For instance, Opera web browser users have the possibility to continue from their previous tabbed browsing session. Therefore, the tabs left open at the end of a web browsing session are saved and can optionally be re-loaded in a next session, thereby saving the user's previous web session. This feature thus extends the intermediate revisitation support of tabbed browsing to additionally support the user's post-session revisitation needs.

Another integrated solution is proposed by Cockburn, Greenberg, Jones, McKenzie, and Moyle [2003]. They conclude their extensive revisitation journal article with a proposal that the Back and Forward buttons, Bookmarks and the History tool can be "beneficially integrated into one revisitation resource." Building on this notion, they present WebView, a collection of several Netscape Navigator-based prototype interfaces, which aim to "improve the effectiveness of direct web page access mechanisms."

In contrast to the traditional stack-based approach, WebView presents users with a complete temporal list of all previously viewed web pages. Of the WebView interfaces, only one is presented as an example [Cockburn *et al.*, 2003]. This example WebView interface includes a temporal list of small thumbnails integrated into the Back/Forward menus, as illustrated in Figure 7. When the user's cursor is placed over a thumbnail, the thumbnail is enlarged. As illustrated in Figure 8, the individual thumbnails include "dogears" as cues: the top-left dogear reveals information about the number of visits to a page and a bottom-left dogear indicates that the page is bookmarked. Therefore, since the WebView list interface includes all visited web pages, including those of previous sessions, it acts much like chronologically-ordered, visually-enhanced History list. A preliminary evaluation with 7 graduate Computer Science students showed that WebView improves the efficiency of some navigational acts [Cockburn *et al.*, 1999].

Figure 7. An example WebView interface: A temporal list of visited links are represented with zoomable thumbnails in the Back menu [Cockburn *et al.*, 2003].

Figure 8. A WebView thumbnail showing visitation frequency (top dogear) and indicating bookmark status (bottom dogear) [Cockburn *et al*., 2003].

Another combined revisitation prototype that was recently proposed is WebScout [Milic-Frayling *et al*., 2003]. This prototype enhances Internet Explorer with SessionNavigator for intermediate revisitation and HistoryExplorer for post-session revisitation. In this system, web navigation sessions are divided into smaller units called WebTrails. A new WebTrail is specified when the user requests a certain page. Such a request includes: typing a URL into the Address bar, selecting a link from the bookmark collection, or selecting a link obtained through search. The SessionNavigator, shown in Figure 9, depicts WebTrails in both a linear, temporally-ordered path included as a toolbar and a tree-structured graph displayed in a separate window. The user is free to use either or both of the views in which visited web pages are represented by thumbnails. Conversely, HistoryExplorer, offers support for post-session revisitation with a search-enhanced, option-rich History list.

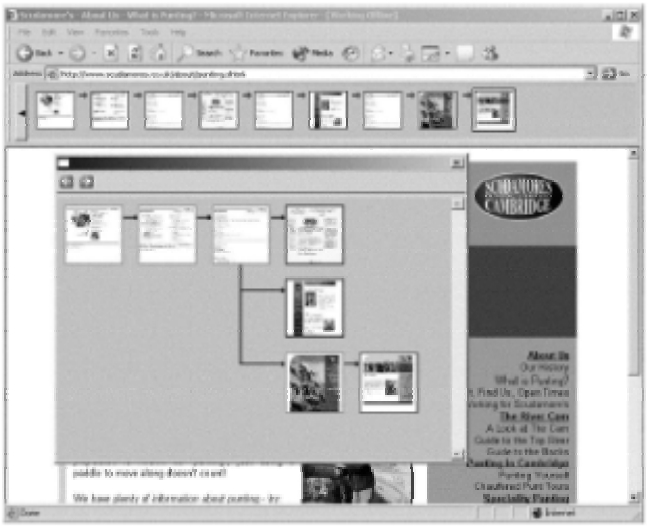

Figure 9. WebScout: SessionNavigator [Milic-Frayling *et al*., 2002].

3.3. Post-Session Revisitation

As stated earlier, web browsers offer two facilities to explicitly support post-session revisitation, the History list and Bookmarks. Arguably, each of these tools can also support intermediate revisitation. We have already reasoned this categorization for the History list and we will confirm this categorization of Bookmarks later in this section.

Despite the availability of these two tools, web users employ a great variety of strategies to ensure their own ability to return to a web page after their web session. Obviously, with the nomadic tendencies of many web users, History list and Bookmarks fall short for at least one reason, they are not portable solutions. However, leaving this issue aside, even in the earlier days of the web, these two facilities were seldom used. In Tauscher and Greenberg's study [1997b], interactions with the hotlist, otherwise known as Bookmarks, accounted for only 3% of all navigation actions and interactions with the History list amounted to less than 1% of all navigation actions.

History lists, Bookmarks, other post-session revisitation approaches and their relative benefits and drawbacks are discussed next.

3.3.1. History List

The History list is a well-known post-session revisitation tool included in all conventional web browsers. In fact, history lists are not exclusive to web browsers. They are incorporated in many tools, such as the recently used files lists of file-based applications. However, in the context of web-related activity, the History list is widely known as a very unpopular tool. This was demonstrated in Tauscher and Greenberg's study [1997b], in which interactions with the History list amounted to less than 1% of all navigation actions, and later observed again by Byrne and his colleagues [1999]. Furthermore, as was observed in Wen's study of PVR [2003], the History list is at times used as a last resort for finding and returning to a page that the user recalls having visited. Using the History tool requires scanning a list of many web pages that are listed by their title and/or URL. This is both a time-consuming and cognitively demanding activity.

JasonSmith and Cockburn studied the relative retrieval efficiencies of four history lists of varied structure: linear, tree, chunking, and two-pane [2003]. Linear represents a list sorted with the most recently visited page listed first. Tree signifies a hierarchically structured list where the URLs are first divided into days and then subdivided into sites. Chunking represents a linear list that shows pages of a previous day with a different background color. Two-pane corresponds to a list which was grouped so that the days were selected from the left frame and the associated URLs are shown in the right frame. They found the "tree" listing to be significantly inefficient compared to the other three: "linear", "chunking", and "two-pane". They suggest to History list designers that as chunking and two-pane don't offer an advantage over linear, a simple linear list is perhaps best suited for history lists. Despite this suggestion, Internet Explorer's History tool does not offer a basic chronological list of visited web sites.

3.3.2. Bookmarks

The Bookmark tool facilitates selective web page collection. A key study of personal bookmark collections and related trends, by Abrams, Baecker and Chignell [1998], is

based on a set of 322 surveys. They determined that their respondents use bookmarks for long-term storage "rather than as caches for frequently used information." The related question offered four types of bookmark usage strategies: temporary, archival, publishing, and collaborative. The respondents were requested to rank each using a 7-point scale, where 1 was "not important" and 7 was "very important". While temporary, publishing, and collaborative each received mean ratings of importance of approximately 3, archival received a mean rating of 6. Furthermore, they comment that "bookmarks are also created when people want to defer reading an interesting page until a future session, possibly because they are too busy dealing with a current problem." Not only did they find that bookmarks are used for archival purposes, but they found that bookmarks are commonly left unvisited for months. Therefore, although the bookmark facility can support intermediate revisitation, it is predominantly used as a post-session revisitation tool.

Research shows that users employ a variety of Bookmarks usage strategies [Cockburn *et al., 2003*]. While some web users bookmark pages sparingly, others include a substantial number of web pages in their bookmark collections. The cost of a large bookmark collection is the increased time required to locate the web page within the list. For this reason, bookmark management is an activity supported by all bookmark facilities. Users can thus use folders to group URLs in a way that is meaningful to them. For large collections, however, this is not necessarily enough for enabling the efficient retrieval of web pages. Others such as Lucid Step Software [2003] have reacted to this shortcoming with Bookmark-related enhancements. They have developed a commercial tool, NetVisualize, which enhances Bookmark collections of Internet Explorer or Netscape Navigator with thumbnail images. Without such visual cues, users are limited to identify a web page based on its title. Likewise, the web site HotScripts.com [2004] includes a great variety of scripts and applications that extend current Bookmark tools with enhancements such as link icons and search facilities. Despite these enhancements, fundamental problems associated with bookmarks would still remain. Though users rearrange their bookmarks, they are reluctant to delete them and this behaviour produces large collections with numerous invalid links [Cockburn *et al*., 2003].

3.3.3. Passive versus Active Strategies

There is one key difference between the History list and Bookmarks. The History list includes all visited web pages without confirmation from the user, while Bookmarks are selectively collected by the user. Consequently, the History list offers a passive revisitation solution, while Bookmarks provide an active method for supporting revisitation [Newfield *et al.*, 1998].

Bookmarks are rather demanding of the user. If one simply adds a link to his or her bookmark collection without providing a meaningful name or a place for it in a folder

with a clear theme, the task of locating the link again in the collection can be demanding. However, performing such activities in the midst of a web session diverts the user's attention away from his or her main task. Moreover, research reveals that some users are reluctant to add web pages to their bookmark collection in fear of cluttering it [Abrams *et al.*, 1998; Wen, 2003]. This is understandable as organization requirements grow with the size of the bookmark collection [Abrams *et al.*, 1998]. Considering these drawbacks, it is reasonable to suspect that users rarely rely on Bookmarks as a sole method for returning to web pages of interest.

3.3.4. Past Research on Post-Session Revisitation Techniques

Though terminology related to this subject varies, there exists considerable research related to web user techniques for enabling post-session revisitation. In the earlier years of the web, Pirolli and Card presented the concept of *information foraging*, referring to "activities associated with assessing, seeking, and handling information sources" [1995]. They argued that, like animals foraging for food, information seekers employ a cost-benefit analysis, whereby the cost is the time required to find certain information. Accordingly, they asserted that design endeavours should be focused on optimizing the user's time rather than enabling the collection of more information.

More recently, Jones, Bruce and Dumais have presented their research topic as the problem of *Keeping Found Things Found* (KFTF) [2001, 2002, and 2003]. Throughout their publications, they refer to the related techniques as "methods of web re-access and re-use" as well as methods of "keeping relevant or potentially relevant information for subsequent use". Likewise, Sellen, Murphy and Shaw address *information gathering*, meaning the use of the web to "purposefully research a specific topic for various reasons" [2002]. As they found that 40% of the information gathering activities observed exceeded one session and that the "participants used various ways of saving the interim results", such as bookmark folders, post-session revisitation plays a key role in this activity. Though both of the information gathering and KFTF observations include static web page collection techniques, such as printing a web page or saving web pages as files, the majority includes relevant post-session revisitation techniques.

Both of these research groups studied the web-related activities of heavy web users in the workplace context. Though the definition of such experienced professionals differs between the research groups, we can generalize that both groups considered individuals who are employed to carry out tasks that regularly involve web-based information. Hereafter, such individuals will be referred to as *knowledge workers*.

In the KFKF study, the information re-use methods of their knowledge worker participants, which included researchers, information professionals (e.g., librarians) and managers, were presented [2001]. These included:

1. Sending an email including the URL to oneself,

2. Sending an email including the URL to others (making the URL available from the Sent folder of the email client),
3. Printing the web page,
4. Saving the web page as a file,
5. Pasting the URL into a document,
6. Adding the URL to a personal web site,
7. Searching for the desired information (again),
8. Entering the URL directly and possibly making use of the URL Auto-complete facility of the web browser, and
9. Bookmarking the web page.

Though explicit post-session revisitation tasks were not included in the study, they found that among their 11 participants, only one used bookmarks and none used the History tool in the observational session. However, the preliminary findings of a subsequent observational study in which participants were asked to recall a cued web site, direct URL entry (8) was employed most often, while select from Bookmarks (9) and search for the page (7) were the second and third most commonly used methods [Jones *et al.*, 2003]. Again, from these studies, we can note the trend that Bookmarks are only made when the page is definitely needed again, but the History tool is still seldom used.

Sellen and her colleagues [2002] classified the activities of knowledge workers based on the findings of their 24-participant diary study. By clustering activities of similar characteristics and objectives, they determined 6 classes of activities:

- Finding: Using the web to find something specific, such as the price of a product.
- Information Gathering: Using the web to research a specific topic and gather relevant information, such as looking for job opportunities.
- Browsing: Using the web for non goal-orientated purposes, such as reading the news, a magazine, checking what's new on a hobby-related site.
- Transacting: Executing a transaction on the web, such as paying a bill.
- Communicating: Participating in chat rooms or discussion groups by means of the web.
- Housekeeping: Checking or maintaining the correctness and functionality of web resources.

Based on the frequency of engagement in these classes of activities, they found that 35% of the actions were related to Information Gathering. This class included the greatest frequency of activity. The next largest activity class frequencies were Browsing, accounting for 27% of all activities and Finding, accounting for 24%.

Thus, post-session revisitation is a key activity of knowledge workers and it is carried out using a variety of methods. However, whether this activity and the associated methods are also common to other web users is unknown.

3.3.5. Proposed Solutions

Improving support for post-session revisitation is a common goal of many researchers. In many cases, visualization is a key trait of these solutions. The various tools proposed are presented in this section.

The designs of WebBooks and Web Forager, proposed by Card, Robertson and York [1996] build on two concepts, the *information workspace* [Card *et al.*, 1991] and *information foraging* [Pirolli and Card, 1995]. The information workspace paradigm proposes the use of a workspace for information access, whereby the user can view, organize and store documents. Such a tool is meant to support the experienced information seeker's approach to gather maximum quality information with minimum effort, as predicted by the information foraging theory. Using the metaphor of a physical book, WebBook represents collected web pages as pages in a 3D interactive book. WebBooks can then be stored in Web Forager, a 3D information workspace.

Data Mountain, which entails a 3D virtual space for collecting web pages [Robertson *et al.*, 1998], also adheres to the concept of the information workspace [Card *et al.*, 1991]. Data Mountain, shown in Figure 10, allows users to place and rearrange web pages anywhere in the view space.

Figure 10. Data Mountain with a collection of 100 web pages [Robertson *et al.*, 1998].

In an experiment comparing the use of Data Mountain for storing and retrieving web pages to that of Internet Explorer's Favourites tool, it was found that Data

Mountain users retrieved pages more efficiently and with a greater success rate. In fact, in a subsequent study, participants were able to rely on their spatial memory so effectively that, when the thumbnail images were removed, they were still able to retrieve pages successfully [Czerwinski *et al.*, 1999].

Amento [2001] addresses the problem of inadequate support for "topic management", which he defines as "the task of gathering, evaluating, organizing, and sharing a set of web sites for a specific topic". As a solution, Amento and his colleagues have developed the tool shown in Figure 11, TopicShop [2000]. This tool is designed to assist users in the evaluation and collection of web sites. By permitting users to group or "pile" web sites anywhere in the Work Area (middle window) using drag and drop, TopicShop also employs spatial arrangement. Textual annotations can be made to individual sites or to entire piles in the Work Area. To further distinguish groups, each is shown with a distinct colour. Correspondingly, each site appears in its group's colour in the rightmost view, Site Profiles. In the Site Profiles view, users can perform individual or group-side operations such as moving or renaming. The upper left window displays a large thumbnail of the most recently selected site to provide greater detail.

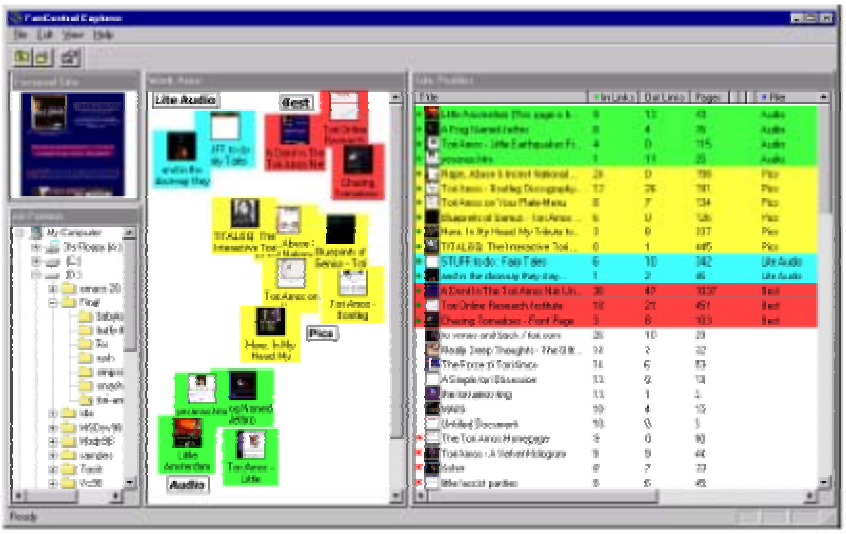

Figure 11. TopicShop containing a collection of four groups [Amento *et al.*, 2000].

They evaluated TopicShop against the combined use of the Yahoo! search engine and Bookmarks. The study included 40 participants, who were required to select the best 15 web sites related to a certain topic, group the selected sites and annotate the groups. They found that TopicShop participants had selected web sites of significantly higher quality than those selected by the Yahoo! participants. As for task completion time, the TopicShop participants took approximately 72% of the time of the Yahoo participants. Furthermore, the TopicShop participants had browsed 67% as many sites as the Yahoo! participants.

3.3.6. Web Users

Although past research identifies as number of post-session revisitation strategies and potential solutions, the majority of the research is based on the activities of knowledge workers [Jones *et al., 2001, 2002, and 2003;* Sellen *et al.*, 2002]. Whether or not the identified techniques apply to less experienced web users remains unknown however. Not only is it uncertain whether these *casual web users* employ the same techniques, it is unknown as to whether such research-oriented web activities are even of relevance to them. With every individual on earth being the target audience of the web, knowledge workers account for only a fraction of its audience. Casual web users are unlikely to receive training, are less likely to adapt to advanced techniques and are more likely to reject tools with poor usability. As such, their web needs and post-session revisitation approaches are of at least equal, if not greater concern.

4. Empirical Study of Post-Session Revisitation

To study the post-session revisitation strategies of casual web users and the relative importance of the task, a questionnaire-based survey was conducted. The primary aim of this survey was to build upon previously observed tendencies of knowledge workers and finally address the first research question of this thesis.

4.1. Interviews With Casual Web Users

Two individuals, who were both contacts of this researcher, were selected as interviewees. These two individuals were selected because neither was a knowledge workers and both could express themselves well in English.

The first interview [Koivikko, 2004] was carried out as an informalized interview [Järvinen, 2001, pp. 131] and the second [Dare, 2004] as a structured interview [Järvinen, 2001, pp. 131], which was recorded with the interviewee's permission. The interviewees' background information was collected, and their post-session revisitation strategies and related rationales were discussed in great detail.

4.2. Questionnaire

Based on the previous findings [Jones *et al*., 2001, 2002, 2003; Sellen *et al.*, 2002], and the two interviews [Dare, 2004 and Koivikko, 2004], a questionnaire was developed [Järvinen, 2001, pp. 132-136; Kitchenham and Pfleeger, 2002] to identify the relative popularity of post-session revisitation techniques using structured questions. Moreover, we were interested to test our theory that web-based research tasks are also carried out by this user group. The inclusion of what Jones, Bruce and Dumais refer to as "do nothing" methods among the possible post-session revisitation strategies was of key importance. "Do nothing" methods include methods such as return later via search or direct URL entry, whereby the user does not actively collect, but instead simply relies on certain methods for revisitation when necessary. The questionnaire is presented in

Appendix 1. As the questionnaire was implemented as a web form, the actual layout is slightly different than that shown in Appendix 1. For instance, text boxes were provided for entering responses and options were presented in drop-down boxes so that the user could see his or her selection with certainty.

After pre-testing the questionnaire with 2 respondents and adjusting the questionnaire accordingly, 47 individuals were requested, by email, to complete the web-based questionnaire. As the aim was to investigate the web page revisitation techniques of casual web users, the selected individuals included all of the contacts of this researcher who qualified, such as high school teachers, sales persons, school children, retired seniors, etc. The instructions for accessing and completing the questionnaire were included in the email.

The respondents were requested to answer the questions from the perspective of only one workstation. For example, if the individual was responding from home he was not to include his web activities of work in his answer. Thus, the answers were exclusive to only one workstation, providing the respondent with one clear context.

4.2.1. Respondents

Of the 47 individual requests for participation, 34 responses were received, resulting in a response rate of just over 70%. The respondents consisted of 14 males and 20 females with a mean age of 29.5, ranging from 12 to 64 years of age. The Internet experience of the respondents ranged from 4 years to 10 years, with a mean of 7.6 years. As shown in Part B of the questionnaire (Appendix 1), the respondents were asked to answer the questions from the perspective of only one location. This location varied between respondents as follows: 17 at home, 14 at work, 1 at school, 1 at another's home, and 1 in an internet café.

4.2.2. Results

The respondents ranked each of the eight common web page revisitation strategies listed to the right of Figure 12. Based on their own experiences of conducting research on the web, such as planning a little trip for next weekend, or doing some background research for a digital camera purchase, respondents were asked to rank each technique with "never", "rarely", "sometimes" and "very often" to represent their own web revisitation activities. The results are depicted in Figure 12.

Based on these results, it is apparent that the technique most depended on for post-session revisitation is returning to the web page using search. It was considered to be used "very often" by 50% of the respondents and "sometimes" by 29.4% of the respondents, together demonstrating its use by 79.4% of the respondents. Following close behind is the use of bookmarks, a strategy employed by 67.6% of the respondents ("very often" by 23.5% and "sometimes" by 44.1%). Entering the URL of the desired web page was the third most popular technique of the respondents, accounting for

58.8% (23.5% “very often” usage and 35.3% “sometimes” usage). Of the eight techniques, the two least popular were self-emailed URLs (“never” 35.3% + “rarely” 35.3% = 70.6% usage) and saving URLs in a document (never 33.4% + rarely 47.1% = 80.5% usage). These findings are illustrated in Figure 13, in which the techniques are presented in order of popularity.

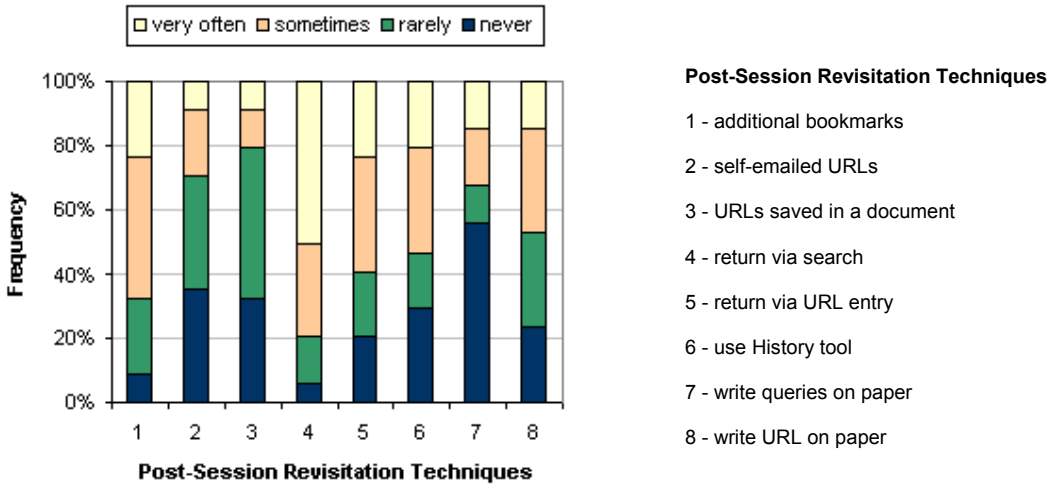

Figure 12: Frequency ratings of the post-session revisitation techniques.

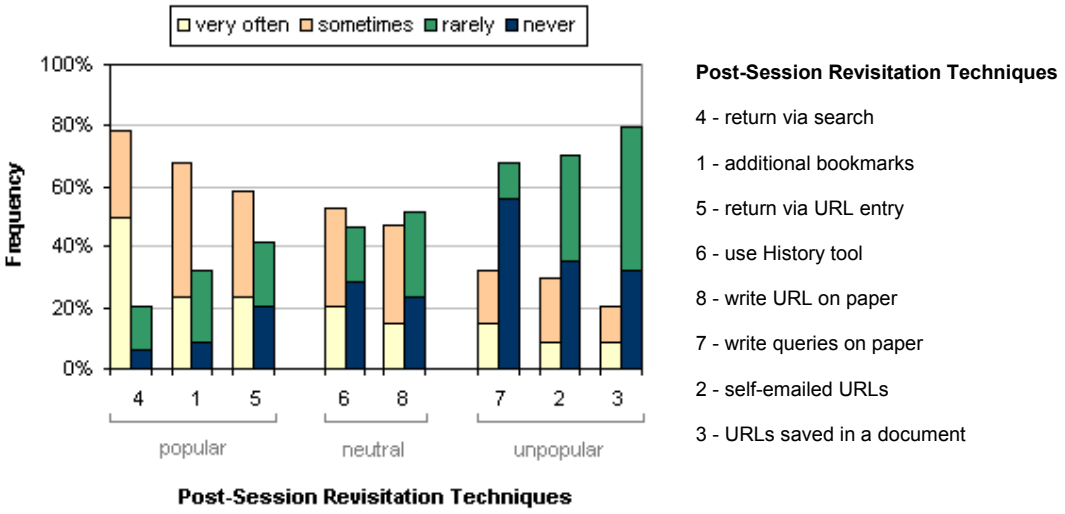

Figure 13: Post-session revisitation techniques ordered by relative popularity. Answers are grouped as positive (very often and sometimes) and negative (rarely and never).

Though the multiple browser window approach is an intermediate revisitation strategy, it was also included with the above mentioned strategies in the questionnaire. Interestingly, this was the most popular strategy of all. Of the respondents, 67.6% replied that their web session results in multiple browser windows “very often”. In addition, 26.5% answered that this occurs “sometimes”. Only 5.9% responded with

"rarely" and 0% selected "never". Therefore, out of the 34 respondents, every single respondent employs the multiple browser window approach to some extent.

4.2.3. Relevancy of the Post-Session Revisitation Task

The key question of the questionnaire, "How often do you think that you engage in this type of task (i.e. as listed earlier planning a little trip for next weekend, doing some background research for a digital camera purchase, etc.)?", aims to determine how often the casual web users engage in web-based research tasks. After having analyzed their own revisitation strategies while completing the previous questions, respondents were in the mind set to consider how relevant this type of task is to them.

Of the respondents, 0% responded with "never", 5.9% with "rarely", 44.1% with "sometimes" and 50% with "very often". When considering that the positive responses together total to 94.1%, it is obvious that this type of task is in fact a routine part of web usage for casual web users as well.

4.2.4. Bookmark Collections

Less fundamental, though still of interest, is the size of the respondents' bookmark collections. The bookmark collection question included instructions on how to determine the number of files and folders in their collection. As this is a relatively advanced request, this question was presented as an optional question. Of the 34 respondents, 9 omitted an answer to this question. Of these 9 respondents, 1 was responding from an internet café and 1 from school, which denote valid empty bookmark collection circumstances. The remaining 7 respondents who did not respond to this question are thus removed from the data represented in Figure 14.

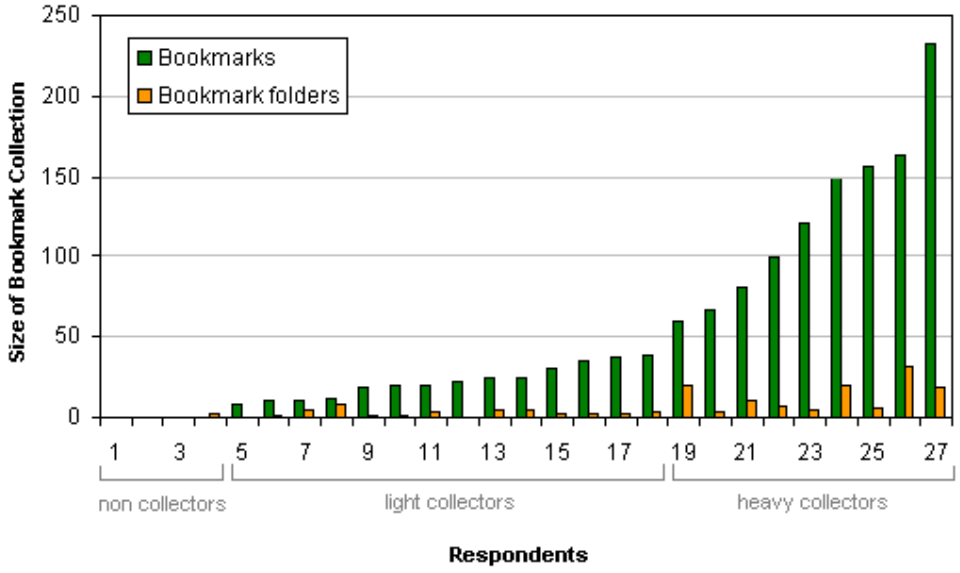

Figure 14. Number of bookmark files and folders of each questionnaire respondent.

Of the remaining 27 respondents, 23 reported having bookmarks. This approximately 85% bookmark usage falls slightly short of the 94% bookmark usage reported in a 322-respondent questionnaire study by Abrams *et al.* [1998]. The Abrams *et al.* study used pre-defined size categories (0 bookmarks, 1-10 bookmarks, 11-25 bookmarks, 26-100 bookmarks, 101-300 bookmarks, and 300+ bookmarks) to retrieve bookmark collection size details from their respondents. For comparative purposes, our own findings are organized accordingly and presented with the findings of the Abrams *et al.* study in Figure 15. A certain consistency is apparent between these two accounts of bookmark collection sizes. Though, a majority of the Abrams *et al.* respondents have bookmark collections sized between 11 to 100 bookmarks, while the collection sizes of our respondents are somewhat more evenly distributed across the categories.

Figure 15. Bookmark collection sizes reported in the Abrams *et al.* study [1998] compared with those of this study.

Of our questionnaire's 23 bookmark-collecting respondents, 22 used folders in their collection. Therefore, a great majority of the bookmark collectors engage in some level of bookmark management. However, as few use folders extensively, it appears that they tend to avoid employing complicated organization structures.

Among the respondents, it is perhaps possible to identify 3 distinct groups of bookmark collectors: *heavy collectors* with over 60 bookmarks, *light collectors* with fewer than 40 bookmarks, and *non collectors* with no bookmarks. Of this group, 9 can be classified as heavy collectors, 14 as light collectors and 4 as non collectors.

Cockburn and McKenzie's four-month analysis of web activities (October 1999 to January 2000) also considered bookmark collections; however, the participants, being Computer Science staff, were knowledge workers [2001]. Of their 17 participants, 15 qualify as heavy collectors, 1 as a light collector, and 1 as a non collector. Additionally, 2 of the heavy collectors had over 500 bookmarks. As well, 12 of their 17 participants had more than 10 folders, 4 of them having over 20 folders. This is again quite different than those reported in this study, where only 4 of the 27 respondents had over 10

folders, of which only 1 had over 20 folders. Understandably, knowledge workers have greater reason to collect and archive more information than casual web users; this fact is evident in this comparison.

4.2.5. Knowledge Workers versus Casual Web Users

Thus, it is obvious that, like knowledge workers, casual web users engage in web-based research activities. Moreover, both groups (actively) employ strategies or (passively) rely on "do nothing" methods for returning to the key pages found during their session.

One inconsistency between the methods of the knowledge workers and the casual web users is the use of the History tool. Though our respondents indicated a variety of dependence on the History tool, a surprising portion, approximately 20%, claimed to use the History tool very often. In the studies of knowledge workers, however, the History tool is used rarely, if at all. It is possible that with intensive

Aside from the History tool, the use of revisitation techniques is rather consistent among the two groups. For instance, the preliminary findings of a recent study by Jones *et al.* [2003], consisting of 21 participants, indicated that the most popular re-access method was direct entry of URL, the second was using Bookmarks, and the third, searching using a search service. Furthermore, some of their participants had noted that they are increasingly relying on "do nothing" methods. Certainly, these "do nothing" methods are portable, easily accessible, convenient, don't require management and don't clutter existing bookmark collections. However, they are unreliable and at times, cognitively demanding to execute. Still, knowledge workers and casual web users are making greater use of these methods. In fact, our respondents are more likely to rely on returning to a web page via search than they are to use either of the post-session revisitation tools included in their web browsers. This is perhaps the clearest indication of the usability problems of currently available post-session revisitation facilities.

5. Prototype Revisitation Tool

Clearly knowledge workers and casual web users alike struggle with a great variety of strategies to support their revisitation needs. To build upon our findings on web navigation and address these needs, we have developed Session Highlights, a prototype revisitation tool. The motivation behind this tool, its design and the general development details are discussed in this section.

5.1. Conception of Session Highlights

The motivation behind the design of this tool is the notion that, though web-based research tasks can span several sessions [Sellen *et al.*, 2002], the information sought is often only needed in the short-term. For instance, an individual who is using the web to plan a trip may need to access the related links frequently during the month before the trip, but never again after the trip. Such a user may be reluctant to clutter his or her

bookmark collection with such provisional links. On the other hand, a workspace of collected links can provide rapid access to key links when they are needed and can be easily dismissed when no longer needed.

We believe that such a tool should allow users to take note of a page of interest, with minimal effort and free of the burden of collection management. Furthermore, a rich set of descriptors, beyond the simple title used in Bookmarks, would be used to represent the web page in the collection. Most importantly, the tool would present users with an information workspace that acts as an "immediate storage area" [Card *et al.*, 1991], leaving Bookmarks as a secondary storage tool. With a resulting collection of a session, the user could proceed to:

- review and compare the content of collected pages,
- save the collection and share it with others,
- save the collection as a session summary to permit session continuation at a later time, and
- add the key pages of the collection to his/her bookmark collection, for long-term storage.

Thus, such a tool offers a workspace in which users can mark interesting web pages, without the need for opening multiple web pages, memorizing search queries, re-evaluating retrieved search results, cluttering a bookmark collection with unneeded URLs, or memorizing URLs.

5.2. Design and Visual Layout

"Exploring information collections becomes increasingly difficult as the volume grows" [Shneiderman, 1996]. This is certainly a key concern in the design of our tool, Session Highlights. To address these challenges, we took advantage of information visualization techniques. Some well-known visualization techniques are employed, such as the use of *focus+context* [Spence, 2001, pp. 116-120] and *details-on-demand* [Shneiderman, 1996]. Moreover, we introduced a novel layout approach, the *snake layout*, for presenting several thumbnails in a size-limited view. In this section, these techniques and the use of thumbnails as a representation of web pages are discussed along with the fundamental design details of the application.

5.2.1. Basic Design

Session Highlights is a separate application window designed to be used as a web browser companion. A collected URL is represented by a thumbnail image, which is created with a screen capture of the associated web page. Therefore, Session Highlights is a web browser-independent application. When the web user starts Session Highlights, it opens as a narrow window on the left side of the screen. Thus the web browser window is preferably located on the right side of the screen. In this current prototype, a screen capture of the entire browser window is included in the thumbnail. As such, the

application works ideally when the browser window is free of frames, such as the History tool in Internet Explorer, which opens as an embedded left frame.

The next development step for this tool includes web browser integration. Then, like the other revisitation facilities, it could be opened in the left frame, offering a solution to the design concern described above. Such integration is perhaps most achievable with an open-source web browser, such as Mozilla. As such browsers are gaining in popularity, extending on Mozilla for example, would have the potential to be of significance to the web user population.

5.2.2. Thumbnails

Thumbnail images are commonly used to represent web pages in visualization tools related to the web. MosaicG [Ayers and Stasko, 1995], PadPrints [Hightower *et al.*, 1998], Domain Tree Browser [Gandhi *et al.*, 2000], WebScout [Milic-Frayling *et al*., 2003], SmartBack [Milic-Frayling *et al*., 2004], OmniWeb [OmniGroup, 2004], TopicShop [Amento *et al*., 2000], and WebView [Cockburn *et al.*, 2003] are all web tools that use thumbnails to depict web pages. Data Mountain [Robertson et al., 1998], the 3D document management tool, also represents web pages with thumbnails. However, as described earlier, it was shown that Data Mountain users were just as successful in retrieving collected pages when the thumbnail was removed and only the spatial location and mouse-over title text were available as cues for identifying the web page [Czerwinski *et al.*, 1999].

In a relatively recent study, Kaasten, Greenberg and Edwards [2002] investigated the recognition of previously visited web pages from thumbnail images. Table 1 summarizes the recognition rates found using thumbnail images of the web pages. In this table, select levels of recognition are presented along with the associated size required of the thumbnail image. For instance, we can see that an exact web page could be recognized from a 208 × 208 pixel image 80% of the time. Moreover, the thumbnail size required to recognize an exact web page is larger than that for simply recognizing the web site to which the page depicted belongs. This is understandable as web pages that belong to a certain web site are often very similar, and so more detail would be required to help distinguish between them. This effect can be seen in Figure 16 where thumbnails of two pages belonging to the same web site are illustrated. It is also important to note that a 100% recognition rate was never achieved. For recognizing a web site, the highest recognition rate is 92% and for recognizing the exact web page, it is 90%.

These findings show that the thumbnail sizes required for decent recognition rates are actually very large, and even then, a perfect recognition rate is not achievable. Based on the results of this study, zooming is a critical feature for visualizations that include thumbnails. While small thumbnails can provide overview and context, users

will need to zoom or otherwise enlarge thumbnails to recognize the associated web pages.

Recognition rate		Thumbnail size (pixels)	
		Web site	Exact page
Minimal:	15%	32 × 32	48 × 48
Low:	30%	48 × 48	80 × 80
Medium:	60%	96 × 96	144 × 144
High	80%	160 × 160	208 × 208
Maximum:		**92 %**	**90%**

Table 1. Thumbnail Recognition Rates [Kaasten *et al.*, 2002].

Figure 16. Thumbnails of two different web pages, sized at 130 pixels by 100 pixels.

Based on their findings, Kaasten *et al.* [2002] concluded that thumbnails are in fact a slightly better representation of web pages than titles or URLs. This is primarily for the reason that "while thumbnails have marginally less instant recognition, they received generally better overall ratings." [Kaasten *et al.* 2002]

Others outside of research have also taken note of importance of visual representations of web pages. For instance, an extension being developed for the Mozilla Firefox web browser incorporates previewing thumbnails of web pages, Amazon products and Stock charts into the result list of the Google search engine (http://www.google.com) [Ackroyd, 2004]. And, as mentioned earlier, a commercial Mac OS X web browser OmniWeb, represents web pages with thumbnails in its tabbed browsing facility.

5.2.3. Snake Layout

In the design of this prototype, we aimed to include as many thumbnails as possible in view while always maintaining the chronological order of the collected web pages. This dual need inspired the *snake layout*. In this layout, thumbnails span both the horizontal and vertical space of the view while being connected by a thin line to emphasize the chronological relationship.

Currently, the prototype employs a static layout algorithm whereby one of three layouts is applied depending on the size of the collection. A basic *vertical line layout* is

applied for small collections (1 to 6 thumbnails), a *snake layout* for medium-sized collections (7 to 22 thumbnails), and a *tighter snake layout* for large collections (23 or more thumbnails). With the first two layouts, when space is abundant, the thumbnails are sized at 100 pixels by 84 pixels. In the tighter snake layout, however, they are reduced to 74 pixels by 60 pixels. Though a number of factors were involved in determining these specific confines, the main ones include: 1) maximizing vertical and horizontal space of the Session Highlights window within the boundaries of a typical 14-inch display, 2) minimizing the amount of scrolling required, and 3) optimizing the possibilities to gain an overview of the collection. For instance, a collection of 6 thumbnails fills the entire height of the Session Highlights window in a typical 14-inch display. With the snake layout, however, 10 to 12 entire thumbnails typically fit into the view and with the tighter snake layout, 24 to 27 thumbnails are fully presented. The snake layout is displayed in Figure 17 on the left, and the tighter snake layout is displayed on the right.

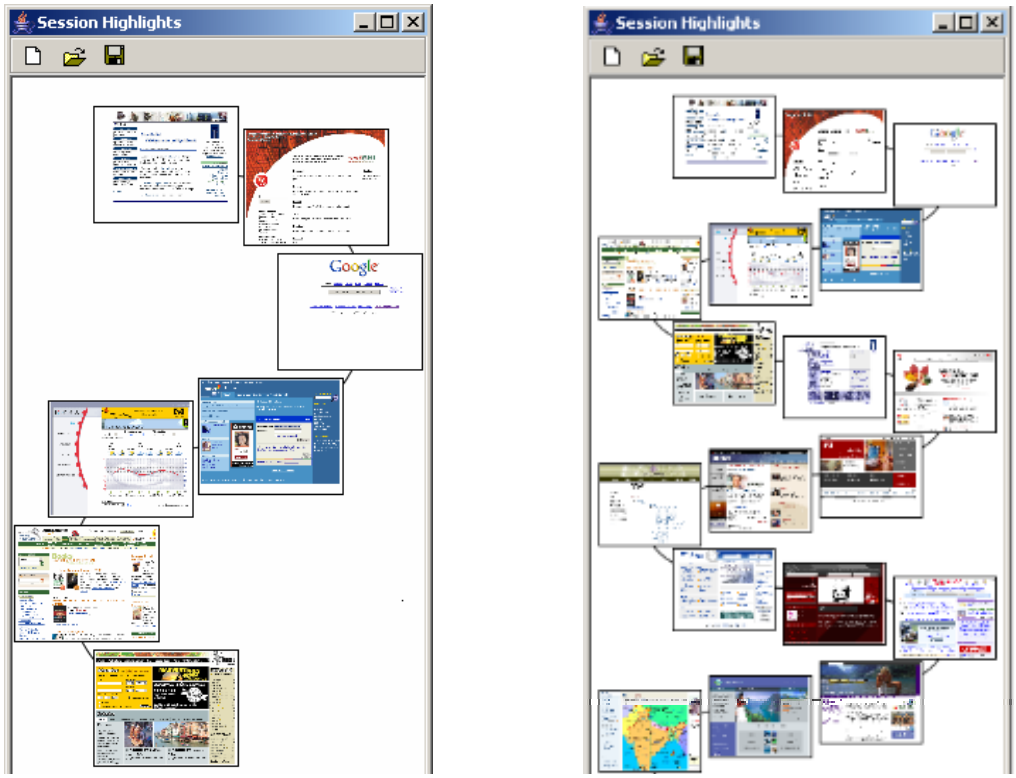

Figure 17. Two Session Highlights windows depicting the snake layout of a medium-sized collection (left) and that of a large-sized collection (right).

Though a static layout was sufficient for the initial prototype implementation, certainly a dynamic algorithm, which maintains the entire collection within the view and thus provides users with a continual collection overview, will be necessary in the future. As North [1997] remarks, a stable, dynamic layout can be of great value for certain applications which, for example, include dynamic data or include interactive

graph editing possibilities. With the simple chronological approach of the snake layout, we believe that such rearrangement would not disturb the user's internal (mental) map of the collection.

In the implementation of a dynamic algorithm, certain minimum and maximum thumbnail sizes must be included. Otherwise, for example, extremely large collections would result in unrecognizably small thumbnail images. Furthermore, at points when such limitations are reached, the use of supplementary visualization techniques will be critical. As well, though the current algorithm commences calculations based on the user's screen size, details such as minimum and maximum thumbnail sizes and the snake layout width could then be configurable, allowing users to define details in accordance with their own screen resolutions and visual preferences.

5.2.4. Interaction

Interaction with the prototype is simple and intended to be intuitive. To add a web page to the collection, the user can drag the URL/Address icon, shown in Figure 18, onto any location of the view. Once the icon is dropped on the view, a thumbnail image of the page is added to the end of the snake. The snake will automatically shift upwards to bring the newly added thumbnail into view if necessary. The user can scroll up and down the snake structure using the mouse wheel. Since the screen must include both the tool and the web browser window, screen width is in short supply. Therefore, a vertical scrollbar was not included in the collection tool.

Figure 18. URL/Address icons as seen in the URL/Address bars of Internet Explorer (top) and Mozilla (bottom).

By placing the mouse cursor over the thumbnail image, a tool text listing the web page title and URL is displayed and the thumbnail is enlarged, as in WebView [Cockburn *et al.*, 2003]. This details-on-demand [Shneiderman, 1996] technique provides users with enhanced focus on the selected web page while maintaining the context of its position in the list. For the purpose of always maintaining context, in other words, to avoid covering adjacent thumbnails, the thumbnail enlargement is relative to the default size of the thumbnail. The size difference between a default thumbnail and an enlarged thumbnail of small and medium-sized collections is illustrated in Figure 19 and that of large collections is illustrated in Figure 20.

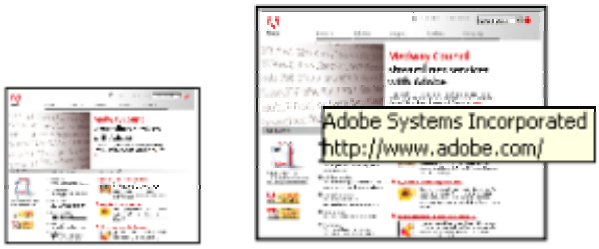

Figure 19. Default thumbnail (left) and enlarged thumbnail with tool text (right) as displayed in small and medium-sized collections and shown to scale.

Figure 20. Default thumbnail (left) and enlarged thumbnail with tool text (right) as displayed in large collections and shown to scale.

To re-access a collected web link, the user must click on the thumbnail of the desired web page with either the left or right mouse button. This opens the associated web page in the most recently opened web browser window.

At any point in time, the user can start a new blank collection, open a previously saved collection, or save a current collection. These three facilities are available using the toolbar buttons located at the top of the application window.

Thus, our approach combines the management-free web page marking approach of Scratchpad [Newfield et al., 1998] and multiple window browsing, with the visual, persistent collection approaches of other research groups [Card et al., 1996; Robertson *et al.*, 1998; Amento et al., 2000; Cockburn *et al*., 2003]. Moreover, it provides "direct page access", as was suggested earlier as a means for improving revisitation efficiency [Cockburn *et al.*, 2003].

5.3. Implementation Technologies

The prototype was mainly implemented in Java 2, with the screen capture functionality written in Delphi 6. As the screen capture portion uses Windows system calls, the prototype is a Windows-specific application. However, since the main body of the tool was developed in Java, platform independence can be achieved rather easily by developing a screen capture component for each platform.

The Java 2 language was selected for the common reasons: it is faster to implement a working prototype in Java in comparison to, for example, C++, it is platform independent and there are many good, free, open-source components available for Java.

5.3.1. Scalable Vector Graphics

The visualization is drawn using the Scalable Vector Graphics (SVG) format. SVG is a language for describing two-dimensional vector graphics in Extensible Markup Language (XML). SVG allows for three types of graphic objects: vector graphic shapes (e.g., paths consisting of straight lines and curves), images and text [W3C Scalable Vector Graphics, 2004]. As well, SVG supports a generous set of effects, such as transformations, transparency, etc. Being a high quality graphics format, SVG is well-suited for displaying interactive visualization. Furthermore, as it is an application of XML, it is highly compatible and thus offers many integration possibilities. For instance, metadata could be collected and saved along with the thumbnails to provide search possibilities and visualization adjustment based on thumbnail interaction history.

5.3.2. Apache Batik SVG Toolkit

The Apache Batik SVG Toolkit was employed for dynamically exporting Java graphics into SVG format. This toolkit was a key technology in the development of this prototype. Batik is a Java technology based toolkit which can be used by applications for viewing, generation or manipulating images in SVG format [Apache, 2002]. By using Batik to assist with the generation of SVG, more of the implementation effort can be allocated to other development tasks.

In addition, the Document Object Model (DOM) API was used to generate the SVG document in order to provide the flexibility for future changes, such as user interface changes. As an additional benefit, the use of DOM API simplifies the management of the entire document structure.

Like many open-source technologies, the Batik toolkit suffers from poor documentation and thus its use entails a steep learning curve. Ultimately, though using the toolkit caused the creation of the base SVG document and its elements to be a lengthy process, the inclusion of visualization and interaction possibilities was facilitated by its use. Moreover, its support for exporting generated content as an SVG document will be of use in the future development of Session Highlights when providing for the distribution and accessibility (from multiple workstations) of URL collections. In this case, other users would not require Session Highlights to view the collection. Any user with an SVG viewer web browser plugin, a viewer which is gaining in popularity, would be able to view and use the resulting collection.

5.4. Comparison with Existing Tools

Though Session Highlights includes similarities with existing revisitation tools, it is unique for a number of reasons. To highlight these differences, a comparison of Session Highlights and the other alternative revisitation tools is presented in Table 2. As the features listed in the table are only those of relevance to Session Highlights, the listing is by no means exhaustive. Moreover, the presence of a feature in a certain tool may be

a matter of interpretation or not entirely applicable to the tool in question. Therefore, this table should be considered as a simplified summary rather than as a comprehensive report.

	SmartBack	ScratchPad	WebView	WebScout	WebBook	Data Mountain	TopicShop	OmniWeb	Opera tabs	Session Highlights
Intermediate revisitation support	●	●	●	●				●	●	●
Post-session revisitation support			●	●		●	●		●	●
1-click URL access while browsing	○							●	●	●
Thumbnails enhanced with text	●		●	○		●	●	●		●
Enlarged thumbnail view	●		●				●			●
Selective URL collection		●	●		●	●	●	●	●	●
Management-free	●	●		●				●	●	●
URLs can be annotated with text			●				●			
Supports multiple collections			●		●	●	●			●
Spatial arrangement of URLs						●	●			
Serves as an information workspace					●	●	●	●	●	●

Table 2. Comparison of proposed revisitation tools and Session Highlights (● = included, ○ = only partially included).

6. Evaluation of Session Highlights

A two-stage observational user study was conducted to observe the learnability and effectiveness of Session Highlights, as well as collect subjective opinions on the users' experiences with the tool. By considering *learnability*, meaning "the ease with which new users can begin effective interaction and achieve maximal performance" [Dix *et al*. 1998, p. 162], we hope to determine the shortcomings of our design. More importantly, by observing *effectiveness*, that being the "accuracy and completeness with which specified users can achieve specified goals in particular environments" [Dix *et al*. 1998, p. 192], we aim to determine whether the tool in fact enables revisitation.

6.1. Method

The observational user study was comprised of a 20 minute web-based research session and, on a subsequent day, a 5 minute continuation of the task ending with an instructed revisitation task. In an attempt to uphold external validity, provide an appealing task, and include clearly defined objectives, the participants carried out a small research task of their own selection. At the end of the second session, the participants completed a survey of questions concerning their background as well as their subjective experiences with the tool.

6.1.1. Participants

Five participants, 2 female and 3 male, with moderately varying computer expertise were recruited from the staff of the Computer Science department of the University of Tampere for the study. The age of the participants ranged from 24 to 41, with a group mean of 30.4 years of age. The group, which included 3 Internet Explorer web browser users, 1 Mozilla user and 1 Opera user, was somewhat representative of the actual browser usage distribution. Most importantly, each of the participants was entirely unfamiliar with Session Highlights. We believe that, despite this group having on the whole far more computer and web experience than the typical web user, for an initial study of this nature, the observations and feedback gained are still of value in evaluating the basic approach and design of the tool.

6.1.2. Apparatus

The study was carried out in a usability lab on a PC running Windows XP with a 15 inch LCD monitor. The participants used Internet Explorer 6.0 to navigate the web. The screen resolution was 1024 by 768 pixels. The collection tool occupied 302 pixels on the left side on the screen, while the web browser, spanning 722 pixels of the 1024 total width, filled the remaining portion of the screen to the right of the tool. Thus the tool occupied nearly 30% of the screen, as shown in Figure 21.

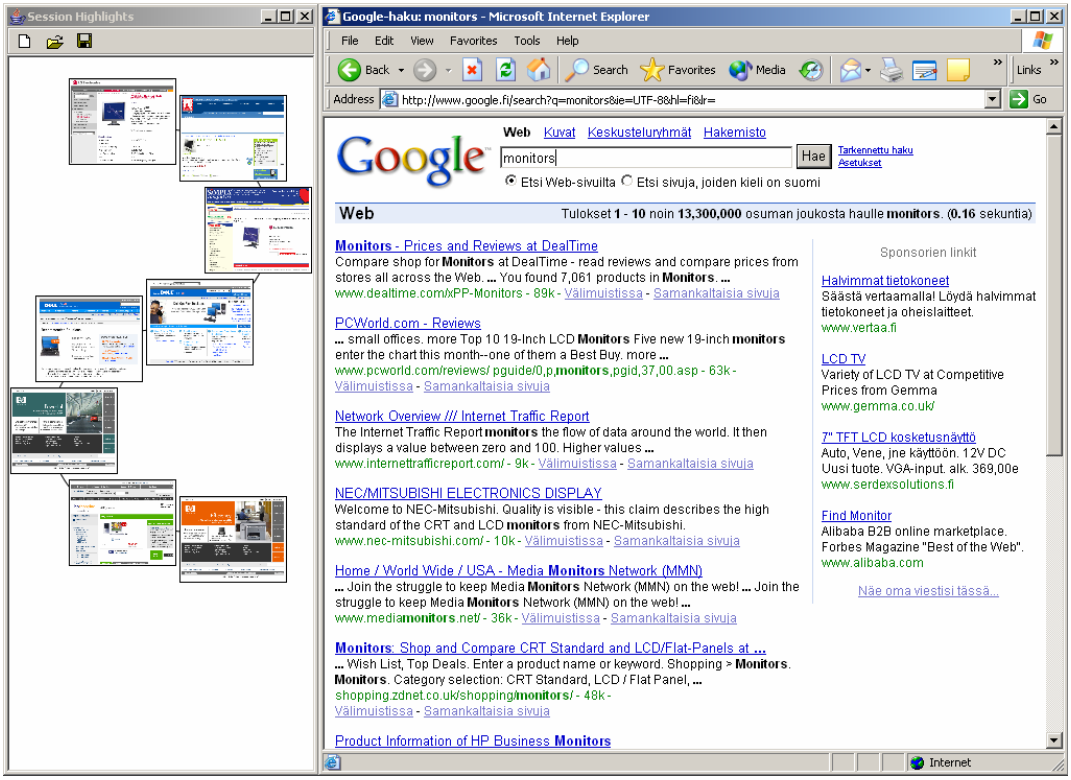

Figure 21. Screen layout as proportioned in the evaluation sessions.

The stopwatch of a mobile phone was used to determine when the participant's web session time was complete. The observations and comments were written down on paper throughout the sessions.

6.2. Procedure

Before starting, the participant was asked to sign and date a consent form which conveyed the purpose of the study, the fact that interactions with the tool would be logged and that the anonymity of the participant will be maintained in the report of the study. Subsequently, the written instruction set was summarized verbally and then given to the participant to read. First listed in the instruction set was the participant's role: "You will use the web to gather some information related to a certain topic." This was followed by 3 example tasks, a description of the tool, and the time length of the web session. The tool description was as follows: "For collecting pages of interest, ***drag and drop*** the ***URL icon*** (i.e. Address icon) of the web page onto the Session Highlights window on the left. To re-access a collected page, click on its thumbnail image and it will open in your most recently opened browser window. You can save your collection by clicking the Save toolbar button. You can also start a new blank collection or an already existing collection using the toolbar buttons."

In the instructions, the participants were asked to select an information gathering task of their own interest. If they did not have such a task in mind, they were able to select from the provided example tasks:

- You are planning a biking weekend in Åland Islands. You need to make a schedule and itinerary, check the ferry options for traveling there, find accommodation options, and look into bike rental options.
- You are planning to buy a barbeque grill. You want to compare different brands and models as well as prices and quality.
- You are interested in taking a whale-watching trip and you want to find out which location/country might be best for this trip and how to go about the trip (e.g. ferries, accommodation, etc.).

To maintain a relaxed setting for the participants, the evaluation was carried out with minimal formality. Neither an audio nor video recording of the session was taken. Instead, the users' interactions with the tool were automatically logged. When an acceptable information gathering task had been decided on, the participant was allowed to try adding a link to Session Highlights, which was already opened along with the web browser. The participant was then instructed to click the "New" toolbar button to start a new workspace and then start browsing. The first stage of the evaluation ended after exactly 20 minutes. At that point, if the participant hadn't already saved the collection, he or she was instructed to save the workspace.

On the second day, the tool and web browser were set up as before. The participant was instructed to continue where he/she left off previously. In order to do this, the participant needed to open his/her previously saved file. After five minutes of continuing the research task, the participant was instructed to return to a certain page, selected from the previous session's log file and described to the participant in words. The page selected was the one which could be best described in words, such as, "the City magazine's article on biking in Ahvenema."

Lastly, users were asked to complete a questionnaire consisting of questions related to their background information and their subjective impressions of the tool. The full questionnaire is included as Appendix 2.

6.3. Results

The user study yielded a number of interesting usage strategies and valuable feedback. These findings are discussed in the following section.

6.3.1. Usage Scenarios, Selected Tasks, and Resulting Collections

Of the five participants, one (participant E), due to scheduling difficulties, completed the entire evaluation in one large session. Two of the participants (C and D) completed the second session after one day and two (participants A and B) after two days.

All of the participants were able to select engaging information gathering topics. In fact, the participants were so engaged in their tasks that most were slightly startled when told that their session time had completed. Two of the participants (B and D) selected the first example task, planning a biking trip in the Åland Islands, each claiming that it was of great interest to them. The others selected online furniture shopping (C), planning a weekend trip to another city (E), and collecting information related to a narrow scientific research topic (A).

Though a variety of usage scenarios and tasks were included in this study, the collection sizes did not fluctuate to a large degree. Table 3 summarizes the relative sizes of the participants' final URL collections. Interestingly, one of the participants used multiple collections to further organize the page gathered for the task.

	Size of Collection (number of URLs)				
	Collection 1	Collection 2	Collection 3	**Total # of URLs collected**	Total # of distinct sites
Participant A	15			**15**	14
Participant B	8	5	2	**15**	12
Participant C	9			**9**	4
Participant D	12			**12**	4
Participant E	8			**8**	6

Table 3. The final collection sizes of each participant (by URL and distinct web sites).

As only interactions with the tool were logged, the participants' web navigation data was not recorded. Therefore, the number of URLs visited by each participant

during the sessions is unknown. Generally, however, it seemed that the participants had added about half of the web sites that they had viewed. This collection rate may be considered rash, possibly resulting from the novelty of the tool. On the other hand, it is also possible that this was truly the amount of web pages of relevance to the participants. For instance, when participant E realized that thumbnails couldn't be deleted, she abandoned her 4-URL collection and started a new collection. However, the new collection contained 3 of the 4 URLs included in the initial collection.

6.3.2. General Observations

The participants were able to take the application into immediate use. The participants who used the Back button for web page revisitation were most successful with the system. Participant A, who obviously manages revisitation through the use of multiple browser windows, experienced some inconvenience however. When this participant opened a link in a new window, it opened with the same width and height as the other window, but on the leftmost side of the screen on top of Session Highlights. Therefore, this participant was forced to drag every new browser window to the right side of the screen to render the collection tool visible. This was indeed a key inefficiency in the participant's information gathering session. This new window location problem does not occur with Mozilla, which opens new windows either approximately one centimetre to the right or to the left of the current page, depending on its location on the screen. Furthermore, this problem doesn't exist when using tabbed browsing.

The most obvious usability issue of the tool appeared when the participants requested the ability to delete and re-arrange thumbnails. Of the five participants, three tried or asked for the possibility to delete thumbnails in the collection and three tried or asked for the ability to re-arrange or group the thumbnails using drag and drop. One of these participants even stated that the chronological ordering is not meaningful and that it would be preferable if the thumbnails would appear in the exact location where they were dropped in the view.

One participant expressed concern for the lack of vertical scrollbar. This participant stated that the size and location of the scrollbar slider were important indicators. Without this feature, she felt that she lost the ability to determine the size of the collection and her current position within the document. This participant expressed great preference for the large collection layout in which more thumbnails appear within the view and a better overview is presented.

For the last task, the revisitation task, four out of five users used the collection tool to return to the requested page. The fifth participant had recently visited the requested page and thus simply selected it from the Back menu.

Participants had clearly enjoyed their experience with Session Highlights. For example, though commentaries were not explicitly requested, participant E commented that the tool was "very easy to use" and that "it was nice to actually see the pages", as

opposed to referring to them by their titles or URLs as other revisitation tools require. Incidentally, this participant hadn't even noticed that the thumbnail was enlarged when the cursor was placed over it. Clearly, this enlargement hardly posed concerns of irritation or distraction.

6.3.3. Working Set of Web Pages

Some of the participants commented that this tool was useful for assembling a "working set" of web pages, which can later be compared and selectively included in one's bookmark collection. For instance, participant E commented that she "might use it for pre-scanning a great amount of pages/links and save as bookmarks only the pages that are found useful." Participant C stated that it would be useful for short-term collection needs: "I might use it in an information gathering task where I need temporary/short-term collecting of web pages, instead of bookmarks." In fact, during the session, this participant had noted that Session Highlights was "useful for online shopping. First you can collect what you like and then later compare prices".

6.3.4. Combined Intermediate and Post-Session Revisitation

Most of the participants used their collected web page to return to an earlier point in their navigation trail during their first session. Within the session, users rarely waited for the thumbnail enlargement or read the title and URL in the tool tip. Since they had recently seen the pages depicted in their collections, they were able to rapidly recall the page. Thus, even within the session, Session Highlights offered a rapid method for revisiting collected pages. For the pages that weren't collected most participants used the Back button or drop-down list for revisitation. However, the need for the Back tool had diminished considerably, as participants had generously added interesting pages to their collections.

For a clearer illustration of Session Highlights enabling intermediate revisitation, we will recount an observed scenario: One participant, who was evaluating furniture, had added the IKEA homepage to his collection. After having navigated deeply into the web site, he simply clicked the IKEA homepage thumbnail to return to where he started. This saved him the time of multiple Back button clicks or opening the Back drop-down list and locating the page in the list. In addition, when he wished to compare products, he was able to evoke collected pages of the IKEA site that would no longer have been available in the Back tool stack.

Therefore, though the main intent of Session Highlights is to enable easy post-session revisitation, it also grants the user with a straightforward intermediate revisitation method.

6.3.5. Supported Web-based Information Retrieval

One innovative technique observed during the evaluation was the use of Session Highlights for preserving search queries and search result lists.

Two of the participants added their search result list to their collection. After having pursued a certain link, each selected their result list from Session Highlights to continue evaluating their retrieved results. In this manner, neither participant had to return to the result list using the Back tool, use a separate browser window to preserve it, or recall the original query. In fact, when these participants returned on the subsequent day, they were able to continue their task where they left off by simply selecting their search results thumbnail. Thus Session Highlights encouraged the collection of search result lists and thereby enabled both intermediate and post-session search result list revisitation.

6.3.6. Overview of Document Type

Participant A, who had pursued a narrow scientific research topic for the evaluation, provided unique comments regarding Session Highlights advantage over Bookmarks. This participant noted the ability to distinguish between the various content types of the collected web pages, such as PDF files, reference lists and organization charts using Session Highlights. In Bookmarks, where the page is represented with only a default title or a user-defined name, such a distinction is seldom possible.

6.4. Post-Usage Questionnaire

Though the questions of the questionnaire were in English, the participants were encouraged to complete the form in whichever language they preferred (English or Finnish). Four of the questionnaires were completed in English and one in Finnish.

Overall, the answers to the subjective rating questions, shown in Table 4, were positive and consistent with what was observed during the sessions. Of the results, the most convincing were for question 1 indicating that the tool was easy to use. The second most indicative result came from question 6, revealing that the tool was considered an efficient means for web page revisitation. The results of question 5 show that the participants found the tool relatively useful for web page revisitation.

	always	often	sometimes	rarely	never
1. Did you ever find this tool (Session Highlights) hard to use?				3	2
2. Did you prefer to look at the thumbnail with your mouse cursor over it so that it was enlarged?		3		1	1
3. Did you read the thumbnail tool tips containing the title and URL of the web page?		1	2	1	1
4. Did you find it hard to recognize the web pages from the thumbnail images?			2	3	
5. Did you find this tool useful in returning to web pages of interest?	1	1	3		
6. Did you find it inefficient to use this tool for returning to web pages of interest?			1	2	2

Table 4. The participants' subjective ratings of Session Highlights.

Responses to questions 2 and 3 illustrated that participants' experiences with the thumbnails varied; while some read the tool tip text, some didn't and while some preferred to look at the thumbnail in an enlarged state, some didn't. Since these features are activated on request, these indicate that their design is well-suited for the demand of these features.

The positive subjective ratings were further supported by the participants' answers to the remaining questions. The first question addressed the participant's engagement in this sort of task. All of the participants agreed that they "often engage in this type of web-based information gathering task." Each described a different combination of strategies for managing the outcome of this task, though the conditional use of bookmarks was included by all.

Each of the participants indicated a clear preference for the snake layout over any other layout, such as a straight vertical line. One of the participants did note that though the snake layout "packs more information into one view", it is does so at the cost of space on the horizontal dimension. Another participant noted that it would be further enhanced with tabbed views to enable grouping.

The participants included many helpful design issues in their responses to the other comments question. These will be included in the discussion presented in the next section. The most common comment, given by 3 out of the 5 participants was desire to edit the collection, meaning delete and re-arrange thumbnails. Finally, participant E noted how useful the tool was for travel planning because related pages appear together, making them "easier to find". She contrasted this approach with Bookmarks in which the pages might have been categorized/grouped differently. This comment supports the utility of such a workspace, which promotes a task-based approach to URL collection.

6.5. Discussion

The evaluation was a success in that it resulted in a numerous suggestions and observations. These will be of great use in the further development of Session Highlights.

As mentioned earlier, participants often tried or ask if it was possible to delete or rearrange the thumbnails. This combined with the multiple requests for thumbnail deletion and rearrangement possibilities in the questionnaire responses demonstrate that the possibility to manage the collection was the key missing feature of this prototype. Though we suspect that deletion will be needed less as the novelty of the tool wears off and additions are perhaps less rash, the option to delete is an always an essential user interface attribute.

Interestingly, only one of the five participants expressed concern for the browser window width limitation associated with this method. This participant noted that with this layout, one needs to use horizontal scrollbars to see the full width of certain web pages in the browser window. As well, only one of the five participants stated that they

would prefer the mouse over thumbnail enlargement to result in even larger thumbnails. He commented that sometimes the current enlargement wasn't enough for viewing logos or reading text. This same participant also suggested a magnifying glass function for viewing select thumbnail sections in greater detail.

Based on the comments of participant A regarding the need for a vertical scrollbar and the preference for the tight snake layout of thumbnails, it is obvious that the tool would benefit from a dynamic layout rearrangement algorithm. This would maintain the entire structure of thumbnails within the view at all times.

Unlike the rest of the participants who maintained only one collection file, participant B created a total of three different collections and switched between the three collections often throughout the sessions. Clearly, this was participant B's method for grouping his findings. Accordingly, the inclusion of tabbed views for grouping web pages in Session Highlights would be a key feature. In fact, this was a popular suggestion among the participants and furthermore, already considered as a future feature for the tool. This study confirmed the importance of this feature. With such a facility included, Session Highlights users could then save the entire collection as one file or each tab individually.

The suggestion for a tabbed view, combined with the frequent requests for rearrangement possibilities indicate that the participants would prefer to apply a meaningful arrangement to their collection. Moreover, it is interesting to note that textual annotation was never requested by any of the participants. Clearly, visual and spatial presentation of the URLs was adequate for the participants. Therefore, Session Highlights' default management-free approach is useful for the base collection stage; however, to support the next stage of research tasks, that being evaluation and comparison, management options, such as re-arrangement, deletion and tabbed views must be available.

7. Conclusions and Future Work

This thesis offers three contributions to the field of Human-Computer Interaction. The first is a greater understanding of the web page revisitation needs and techniques of casual web users. The second is the snake layout, a novel layout for visualizing a URL collection in a horizontally-limited space. The third is Session Highlights, a prototype web workspace and post-session revisitation tool.

To address the initial research problem, "Do users employ techniques to facilitate post-session web page revisitation when carrying out web-based research tasks?", previous research on the post-session revisitation strategies of knowledge workers was considered. To extend on these findings by assessing the approaches of casual web users, a web-based questionnaire was conducted. The results indicated that among our respondents web-based research is a common web activity. In fact, 50% claimed to

engage in the activity “very often”, while about 44% claimed to engage in it “sometimes”. The most popular techniques used to enable the revisitation of web pages found while carrying out a research task included: 1) relying on a search engine to retrieve the web page, 2) bookmarking the web page, and 3) relying on recalling the URL, with the possible assistance of the Auto-complete mechanism. However, research indicates that each of these techniques can be problematic. Thus it was anticipated that a more effective revisitation facility to support this task would be beneficial to web users.

To build on these findings, we designed and developed a visualization-enhanced revisitation tool, Session Highlights. This tool provides users with a saveable workspace within which they can collect links to interesting web pages in a management-free manner. The added links are represented by thumbnails and presented chronologically in a manner referred to as the snake layout. This unique layout maintains continuity among the list while also making efficient use of the limited space.

Session Highlights received encouraging results in an initial observational user study. In fact, the tool was found to be more useful than originally expected at the outset. Not only did it serve as a post-session revisitation tool, but it was also used instinctively by the participants for direct-access intermediate revisitation. Furthermore, the tool’s workspace approach provided the participants with a valuable summary of their session’s findings. Clearly, Session Highlights, despite being an early prototype, performed as an easy and effective web page revisitation tool. Moreover, observations and participant feedback have offered valuable indications of how to further improve this tool.

In conclusion, we have validated the importance of the web-based research task and its inherent association with web page revisitation. Though such tasks often entail frequent intermediate revisitation and post-session revisitation, current web facilities lack appropriate support for this activity. However, when users are provided with a web workspace, such as Session Highlights, in which they are able to first gather and then later assess interesting web pages, this task is eased.

This thesis presents the foundation for a variety of possible subsequent research endeavours. The most obvious next step includes the further development of Session Highlights. As revealed during our observational study, the key requirements include:

- drag and drop thumbnail rearrangement
- thumbnail deletion
- tabbed panes for thumbnail grouping
- a dynamic snake layout algorithm
- explicit support for varied visualization techniques, such as a fisheye view, a magnifying glass technique, and/or greater zooming facilities
- integration with web browser(s)

Moreover, the separation and continued investigation of intermediate and post-session revisitation is essential in supporting the vast information processing requirements of web users. Though it has been ten years since revisitation was first identified as a fundamental and yet troublesome activity in web navigation, there remains great need for improved revisitation facilities.

References

[Abrams *et al.*, 1998] David Abrams, Ronald Baecker, and Mark Chignell, Information archiving with bookmarks: Personal web space construction and organization. In: C. M. Karat, A. Lind, J. Coutaz and J. Karat (eds.), *Proc. of the 1998 ACM SIGCHI Conference on Human Factors in Computing Systems*, 41-48.

[Ackroyd, 2004] Edward Ackroyd, GooglePreview 0.2, Mozilla Firefox Extensions for Search Tools. Available as http://update.mozilla.org/extensions/showlist.php?type=E&application=firefox&category=Search%20Tools&numpg=.

[Amento, 2001] Brian Amento, User interfaces for topic management of web sites. Virginia Tech, Dept. of Computer Science, Ph.D. Thesis. September, 26th, 2001. Also available as http://scholar.lib.vt.edu/theses/available/etd-12022003-105949/unrestricted/amento-dissertation-final.pdf.

[Amento *et al.*, 2000] Brian Amento, Loren Terveen, William Hill, and Deborah Hix, TopicShop, enhanced support for evaluating and organizing collections of Web sites. In: *Proc. of the 13th Annual ACM symposium on User interface software and technology*, 201-209.

[Apache, 2002] Apache Batik 1.1 Documentation. Available as http://xml.apache.org/batik/.

[Aula and Käki, 2003] Anne Aula and Mika Käki, Understanding expert search strategies for designing user-friendly search interfaces. In: P. Isaías and N. Karmakar (eds.), *Proc. of IADIS International Conference WWW/Internet 2003*, 759-762.

[Ayers and Stasko, 1995] Eric Ayers and John Stasko (1995). Using graphic history in browsing the World Wide Web. In: *Proc. of the 4th International World Wide Web Conference*.

[Brown and Shillner, 1995] Marc H. Brown and Robert A. Shillner, A new paradigm for browsing the web. In: *Proc. of the 1995 ACM SIGCHI Conference on Human Factors in Computing Systems,* 320-321.

[Byrne *et al*., 1999] Michael D. Byrne, Bonnie E. John, Neil S. Wehrle, and David C. Crow, The tangled web we wove: A taskonomy of WWW use. In: *Proc. of the 1999 ACM SIGCHI Conference on Human Factors in Computing Systems,* 544 - 551.

[Card *et al*., 1991] Stuart K. Card, George G. Robertson, and Jock D. Mackinlay, The Information Visualizer, an information workspace. In: *Proc. of the 1991 ACM SIGCHI Conference on Human Factors in Computing Systems*, 181-186.

[Card *et al.*, 1996] Stuart K. Card, George G. Robertson, and William York, The WebBook and the Web Forager: An information workspace for the World Wide

Web. In: *Proc. of the 1996 ACM SIGCHI Conference on Human Factors in Computing Systems*, 111-117.

[Catledge and Pitkow, 1995] Lara Catledge and James Pitkow, Characterizing browsing strategies in the World Wide Web, In: *Proc. of the 3rd International World-Wide Web conference on Technology, tools and applications*, 1065-1073.

[Cockburn *et al.*, 2003] Andy Cockburn, Saul Greenberg, Steve Jones, Bruce McKenzie, and Michael Molye, Improving web page revisitation: Analysis, design and evaluation. B. Shneiderman, J. Lazar, M. Ivory (eds.), Special Issue on Web Navigation Skills, *IT&Society* **3** (1), 2003, 159-183.

[Cockburn *et al.*, 2002] Andy Cockburn, Bruce McKenzie, and Michael JasonSmith, Pushing back: Evaluating a new behaviour for the back and forward buttons in web browsers. *International Journal of Human-Computer Studies* **57** (5), 2002, 397-414.

[Cockburn *et al.*, 1999] Andy Cockburn, Saul Greenberg, Bruce McKenzie, Michael JasonSmith, and Shaun Kaasten, WebView: A graphical aid for revisiting web pages. In: *Proc. of the Australian Conference on Computer-Human Interaction*, 15-22.

[Cockburn and Jones, 1997] Andy Cockburn and Steve Jones, Design issues for World Wide Web navigation visualisation tools. In: *Proc. of the 5th Conference on Computer-Assisted Research of Information*, 55-74.

[Cockburn and Jones, 1996] Andy Cockburn and Steve Jones, Which way now? Analysing and easing inadequacies in WWW navigation. *International Journal of Human-Computer Studies* **45** (1), 1996, 105-129.

[Cockburn and McKenzie, 2001] Andy Cockburn and Bruce McKenzie, What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies* **54** (6), 903-992.

[Czerwinski *et al.*, 1999] Mary M. Czerwinski, Maarten van Dantzich, George G. Robertson, and Hunter Hoffman, The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. In: A. Sasse and C. Johnson (eds.), *Proc. of Interact '99, the 7th IFIP Conference on Human-Computer Interaction*, 163-170.

[Dare, 2004] Tanya Dare, interview on June 8th, 2004.

[Dix *et al*. 1998] Alan Dix, Janet Finlay, Gregory Abowd, and Russell Beale, *Human-Computer Interaction, Second Edition*. Prentice Hall Europe, 1998.

[Domel, 1994] Peter Domel, WebMap - A graphical hypertext navigation tool. In: *Proc. of the 2nd International WWW Conference*.

[Gandhi *et al.*, 2000] Rajiv Gandhi, Girish Kumar, Ben Bederson, and Ben Shneiderman, Domain name based visualization of web histories in a zoomable

user interface. In: *Proc. of the 11th International Workshop on Database and Expert Systems Applications*, 591-600.

[Greenberg and Cockburn, 1999] Saul Greenberg and Andy Cockburn, Getting back to back: Alternate behaviors for a web browser's back button. In: *Proc. of the 5th Conference on Human Factors and the Web*.

[Hightower *et al.*, 1998] Ron Hightower, Laura Ring, Jonathan Helfman, Ben Bederson, and Jim Hollan, Graphical multiscale web histories: A study of PadPrints. In: *Proc. of the 1998 ACM Conference on Hypertext*, 58-65.

[HotScripts.com, 2004] HotScripts.com, Bookmark Management. Available as http://www.hotscripts.com/PHP/Scripts_and_Programs/Bookmark_Management/.

[JasonSmith and Cockburn, 2003] Michael JasonSmith and Andy Cockburn, Get a way back: Evaluating retrieval from history lists. In: *Proc. of the 4th Australian user interface conference on User interfaces 2003*, 33-38.

[Jones *et al.*, 2003] William Jones, Harry Bruce, and Susan Dumais, How do people get back to information on the web? How can they do it better? In: *Proc. of Interact 2003, the 9th IFIP Conference on Human-Computer Interaction*. Also available as http://kftf.ischool.washington.edu/Jones,%20Bruce,%20Dumais-Interact2003.pdf.

[Jones *et al.*, 2002] William Jones, Harry Bruce, and Susan Dumais, Once found, what then?: A study of "keeping" behaviors in personal use of web information. In: E.G. Toms (ed.), *Proc. of the 65th ASIST Annual Meeting*, 391-402.

[Jones *et al.*, 2001] William Jones, Harry Bruce, and Susan Dumais, Keeping found things found on the web. In: H. Paques, L. Liu and D. Grossman (eds.): *Proc. of the 2001 ACM CIKM 10th International Conference on Information and Knowledge Management*, 119-134.

[Järvinen, 2001] Pertti Järvinen, *On Research Methods*. Opinpaja, Tampere, 2001.

[Kaasten *et al.*, 2002] Shaun Kaasten, Saul Greenberg, and Christopher Edwards, How people recognize previously seen WWW pages from titles, URLs and thumbnails. In: X. Faulkner, J. Finlay and F. Detienne (eds.), *Proc. of Human Computer Interaction 2002, BCS Conference Series*, 247-265.

[Kitchenham and Pfleeger, 2002] Barbara A. Kitchenham and Shari L. Pfleeger, Principles of survey research part 3: Constructing a survey instrument. *Software Engineering Notes* **27** (2), 2002, 20-24.

[Koivikko, 2004] Jaakko Koivikko, interview on May 30th, 2004.

[Nielsen, 1990] Jakob Nielsen, The art of navigating hypertext. *Communications of the ACM* **33** (3), 1990, 297-310.

[North, 1997] Stephen C. North, Dynamic Graph Layout and Information Visualization. In: *Proc. of the 1997 CODATA Euro-American Workshop on Data and Information Visualization*. Available as http://www.codata.org/archives/1997Vis/sp12.htm.

[Lucid Step Software, 2003] Lucid Step Software, NetVisualize Favorites Organizer. Available as http://www.netvisualize.com/.

[Newfield *et al.*, 1998] Dale Newfield, Bhupinder Singh Sethi, and Kathy Ryall, Scratchpad: Mechanisms for better navigation in directed web searching. In: *Proc of the 11th Annual ACM Symposium on User Interface Software and Technology*, 1-8.

[Milic-Frayling *et al*., 2004] Natasa Milic-Frayling, Rachel Jones, Kerry Rodden, Gavin Smyth, Alan Blackwell, and Ralph Sommerer, SmartBack: Supporting Users in Back Navigation. In: *Proc. of the 13th International WWW Conference*, 63-71. Also available as http://www.www2004.org/proceedings/docs/1p63.pdf.

[Milic-Frayling *et al*., 2003] Natasa Milic-Frayling, Ralph Sommerer, and Kerry Rodden, 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.

[Milic-Frayling *et al*., 2002] Natasa Milic-Frayling, Ralph Sommerer, and Robert Tucker, MS WebScout: Web navigation aid and personal web history explorer. In: *CD-ROM Proc. of the 11th International World Wide Web Conference*. Also available as http://www2002.org/CDROM/poster/170/.

[Moyle and Cockburn, 2003] Michael Moyle and Andy Cockburn. The design and evaluation of a flick gesture for 'Back' and 'Forward' in web browsers. In: *Proc. of the 4th Australasian User Interface Conference*, 39-46.

[OmniGroup, 2004] OmniGroup, OmniWeb web browser. Available as http://www.omnigroup.com/applications/omniweb/.

[Opera Software, 2004] Opera Software, Opera web browser. Available as http://www.opera.com/.

[Pirolli and Card, 1995] Peter Pirolli and Stuart Card, Information foraging in information access environments. In: *Proc. of the 1995 ACM SIGCHI Conference on Human Factors in Computing Systems*. Also available as http://www1.acm.org/sigs/sigchi/chi95/Electronic/documnts/papers/ppp_bdy.htm.

[Refsnes Data, 2004] Refsnes Data, Browser Statistics. Available as http://www.w3schools.com/browsers/browsers_stats.asp.

[Robertson *et al.*, 1998] George Robertson, Mary Czerwinski, Kevin Larson, Daniel C. Robbins, David Thiel, and Maarten van Dantzich, Data Mountain: Using spatial memory for document management. In: *Proc. of the 11th Annual Symposium on User Interface Software and Technology*, 153–162.

[Sellen *et al.*, 2002] Abigail J. Sellen, Rachel Murphy, and Kate Shaw, How knowledge workers use the web. In: *Proc. of the 2002 ACM SIGCHI Conference on Human Factors in Computing Systems*, 227- 234.

[Shneiderman, 1996] Ben Shneiderman, The eyes have it: A task by data type taxonomy for information visualizations. In: *Proc. of the 1996 IEEE Symposium on Visual Languages*, 336-343.

[Spence, 2001] Robert Spence, *Information Visualization*. Addison-Wesley, 2001.

[Tauscher and Greenberg, 1997a] Linda Tauscher and Saul Greenberg, 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.

[Tauscher and Greenberg, 1997b] Linda Tauscher and Saul Greenberg, Revisitation patterns in World Wide Web navigation. In: *Proc. of the 1997 ACM SIGCHI Conference on Human Factors in Computing Systems*, 399-406.

[Utting and Yankelovich, 1989] Kenneth Utting and Nicole Yankelovich, Context and orientation in hypermedia networks. *ACM Transactions on Office Information Systems*, **7** (1), 1989, 58–84.

[Wen, 2003] James Wen, Post-Valued Recall web pages: User disorientation hits the big time. *IT&Society* **1** (3), 2003, 184-194.

[W3C Scalable Vector Graphics, 2004] W3C Scalable Vector Graphics (SVG) 1.1 Specification. Available as http://www.w3.org/TR/SVG/

[W3C Semantic Web, 2004] W3C Semantic Web. Available as http://www.w3.org/2001/sw/.

[Xu *et al.*, 2001] Geoffrey Xu, Andy Cockburn, and Bruce McKenzie, Lost on the web: An introduction to web navigation research. In: *Proc. of the 4th New Zealand Computer Science Research Students Conference*.

Appendix 1

Questionnaire

PART A

1. How long have you been using the Internet for? (approximately, in years)

2. What is your age? (This will only be used to compute the group average. So your age will not be used against you in any way!)

3. What is your profession? (e.g. sales representative, biology student, retired secondary school teacher, etc.)

4. What is your email address? (This will not be sent to anyone. This is just in the case that I am so fascinated with your answers that I want to contact you to ask you a few more.)

PART B

The following questions relate to your Internet usage methods. Please answer them ONLY from the perspective of the location where you MOST use the Internet (i.e. at home, at work, etc.)

5. What is this location that you will be referring to when answering these questions? (e.g. home, work, library, school, someone else's home, etc.)

6. How long have you been using this computer for? (i.e. How long have you been collecting web site bookmarks on this computer for? An approximation is good enough. E.g. 6 months, 2 years, etc.)

7. List some of the most typical tasks that you carry out on the Internet at this location? (e.g. email, read the news, order books, health-related research, etc.)

PART C

8. Imagine yourself spending a few hours of the day surfing the web with the aim of exploring a certain topic (e.g. planning a little trip for next weekend, or doing some background research for a digital camera purchase, etc.) Towards the end of this web browsing session, would you have:

a) many web pages open at the same time (i.e. you have opened many web browser windows)?

never rarely sometimes very often

b) one or more additional bookmarks added to your Bookmarks/Favorites list?

never rarely sometimes very often

c) one or more web addresses (i.e. URL's) in an email to yourself?

never rarely sometimes very often

d) one or more web addresses (i.e. URL's) saved in a document?

never rarely sometimes very often

e) hope that you can always use a search engine to find them again when you need them?

never rarely sometimes very often

f) hope that you will remember the web addresses (i.e. URL's) of the relevant home page(s), when you need them again? (This includes selecting from that list based on your browsing history, which appears automatically from the Addresses box.)

never rarely sometimes very often

g) hope that the locations you've visited today will be stored in the History tool of your web browser? (If you don't know what this is then your answer is "never".)

never rarely sometimes very often

h) some queries (i.e. search key words) written down (i.e. with a pen or pencil) somewhere?

never rarely sometimes very often

i) some web addresses (i.e. URL's) written down (i.e. with a pen or pencil) somewhere?

never rarely sometimes very often

j) something else? Please specify:

9. How often do you think that you engage in this type of task (i.e. as listed earlier planning a little trip for next weekend, doing some background research for a digital camera purchase, etc.)?

never rarely sometimes very often

PART D

10. How many folders and links are saved in your Bookmarks/Favourites tool?

If Internet Explorer is your web browser, you can check this rather easily in the following way:

- Go to the folder named "Favorites" in your computer (e.g. using Windows Explorer). This folder is, for example, probably located in the following location on your computer C:\Documents and Settings\Natalie Jhaveri\Favorites

- Right-click on the folder (may be marked with a star icon) and select "Properties" from the drop-down list.

- Take note of the number of files and the number of folders listed with "Contains". E.g. "Contains: 119 Files, 30 Folders"

If you don't want to do it this way, you can select the "Favorites" or "Bookmarks" menu in your web browser and actually count the contents. Or, you can simply skip this question.

Number of files:

Number of folders:

Appendix 2

Session Highlights Usage Survey

Background Information:

What is your age? _____

What is your gender? male female

What is your profession? _____

How many years have you been using the Internet for? _____

On average, how many hours per day do you use the Internet for? _____

What web browser do you use primarily?_____

Questions:

Circle the option that closest reflects your opinion.

1. Did you ever find this tool (Session Highlights) hard to use?

   always often sometimes rarely never

2. Did you prefer to look at the thumbnail with your mouse cursor over it so that it was enlarged?

   always often sometimes rarely never

3. Did you read the thumbnail tool tips containing the title and URL of the web page?

   always often sometimes rarely never

4. Did you find it hard to recognize the web pages from the thumbnail images?

   always often sometimes rarely never

5. Did you find this tool useful in returning to web pages of interest?

   always often sometimes rarely never

6. Did you find it inefficient to use this tool for returning to web pages of interest?

always often sometimes rarely never

7. a) Do you often engage in this type of web-based information gathering task?
   b) If so, do you actually collect the "potentially useful/important" pages that you find?
   b.i) If you do collect these web pages, how do you go about collecting them?
   b.ii) If not, do you ever need to return to them later? If so, how do you do it?

8. In what cases, if any, might you imagine yourself using such a tool? (e.g. instead of using some other tool or technique?)

9. Did you like the "snake layout" of the thumbnails or would you have preferred some other layout (e.g. a straight vertical line)?

10. Do you have any other comments about the web page collection tool or your experiences with it?