

POST-VALUED RECALL WEB PAGES: USER DISORIENTATION HITS THE BIG TIME

JAMES WEN

ABSTRACT

The growth in complexity of the World Wide Web, matched by the increase in user sophistication, has transformed the notion of user disorientation. While the retrieval of visited Web pages has been studied, the particular class of Web pages not initially recognized as being important enough to return to merits closer examination. Such pages, referred to as Post-Valued Recall (PVR) Web pages, tend to sit at the interface between the solitary act of surfing the Web and the social act of interacting with colleagues, who will often trigger the realization that a particular page is, in fact, of value. Results from a user study illustrate the effect of PVR in a social environment, as well as offer design guidelines for browser utilities that could diminish the negative impact of PVR.

James Wen is a professor of Computing and Information Systems at the University of Sunderland, England. jameswen@ureach.com.

A form of user frustration arising from Web navigation will likely only grow with advancing Web technology and increasing user sophistication. As the World Wide Web (or simply, Web) continues its transition from novelty to utility, and as surf behavior becomes more methodically goal-oriented and less serendipitous, the notion of user disorientation seems to be transforming from *How did I get here?* To *How do I get back there?* Put another way, users do not appear to be experiencing a fundamental confusion with respect to paths taken to get to their current page; rather, they are more interested in retrieving pages that they had previously encountered but the location for which they had not previously saved. The nature of such pages—namely that they have been previously encountered—may cause users to misjudge the difficulty with which they can be re-located. The typical impression that such a page should, at worst, take as long to find as it took initially to encounter may be misplaced and could lead to user frustration and potential detrimental consequences in workflow management.

POST-VALUED RECALL WEB PAGES

The term *Post-Valued Recall (PVR)* refers to 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. PVR can occur during or after a surf session. In the former instance, PVR occurs after the user has surfed enough to have established a context within which to judge the value of a Web page. The point at which the user has established a context may not be consciously known, and the decision to start judging Web pages against this context—which itself is built from the Web pages being judged—may be due to a number of factors, such as the user's time constraint, as well as the user's satisfaction that enough material has been viewed so that a proper context has been established.

PVR can also occur after a surf session has ended, because Web pages encountered—even if not tagged as special (e.g., bookmarked)—will remain in the user's mind and can be recalled, if subsequent interaction of a relevant nature, outside of the surf session, should trigger it. PVR can be seen as a natural consequence of extending the act of research from being limited to the solitary act of surfing the Web to one that includes greater social activity in which a user's information space interacts with the surrounding environment. For instance, conversations with colleagues may cause the user realize the importance of a previously visited Web page. Subsequent attempts to revisit such a page may prove frustrating because, in trying to recreate steps used earlier, the user may omit the smallest detail, leading to a fruitless search in the wrong direction.

METHODS

To help better define and analyze PVR, a preliminary user study was designed. To create further tests as well as to design new browser utilities, 12 users—all Web-savvy and ranging from students to professionals—were each given three disparate topics to research on the Web. Except for being restricted to using the Internet Explorer browser, users were encouraged to surf and to use whatever tools they normally would, including bookmarking tools, text editors and paper pads.

All users were interrupted in mid-session with a distractive task—search for some specific information, in a limited time, for a fictitious friend. The purpose of the interruption was to simulate a task that would expose the user to information unrelated to the focus of attention at the time, but which may later be recalled, because it has become part of the user's knowledge.

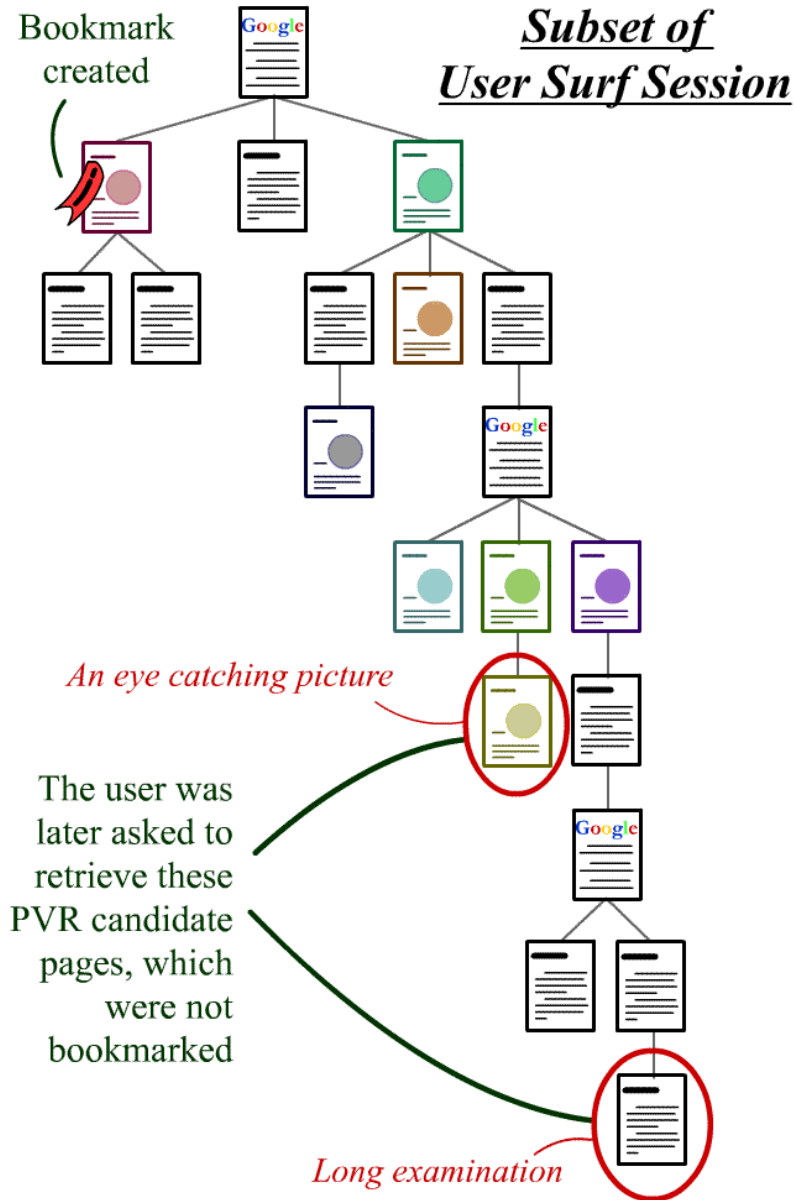
The surf sessions lasted from thirty minutes to an hour. Users were observed through a one-way mirror and recordings were made of the user's face, hands and the computer monitors. Based on the amount of time a user spent on the various Web pages, particular pages were selected as candidate PVR Web pages. For comparison purposes, pages that were bookmarked (or otherwise noted) were also selected. An interview was given after the surf session, which included requests that users retrieve the selected Web pages. If a user could not retrieve a page within two minutes, the page was considered irretrievable.

Figure 1 depicts a small subset of a user's surf session. The root node at the top of the graph represents a point where the user has visited a search engine. In this case, the user considered one of the listed pages to be noteworthy after a brief survey of a couple of its links and tagged the page with a bookmark. No further bookmarks were made in the remainder of the depicted subset. However, two further pages were selected as PVR candidate pages—one because it contained a potentially memorable image, and the other because the user exhibited enough interest in the content of the page to have dwelt upon it for a considerable amount of time, even though it was not directly related to the topic of research.

RESULTS

In the post-session interview, users were asked to retrieve the selected pages, both those bookmarked (or otherwise noted) and the PVR candidate pages. It is interesting to note that virtually all attempts to re-locate PVR candidate Web pages involved some form of retracing. Although it was difficult to recreate a surf session completely—because pure retracing involves re-locating all the previously visited pages—the general strategy of re-creating the original process that led to the candidate page once a temporal landmark is recognized was heavily used. The landmark page

FIGURE 1: SUBSET OF A USER SURF SESSION



could be one that the user recognizes as being part of the original surf path, or it may be one that the user returns to often.

With one exception, all these users used the [Google search engine](#) as a starting point for searches and as a point to return to when a feeling of disorientation set in. This finding is consistent with the notion that users like to return to a known starting point when disoriented in a complex information space (Barret 1988). Interviews revealed that users would rather compromise the associative power of the hyperspace of the Web in favor of a predictable, if more limiting, strategy of the paths offered by [Google](#). The perception seems to be that the access to the Web is not limited by [Google](#), but is organized in a fashion that allows easy retracing later. This holds true, despite the fact that the study indicated that the success rate for retracing to be relatively low (20% success).

Although this study shows bookmarks to be very effective in retrieving Web pages, interviews with users revealed that a very important factor that encouraged their use of the bookmarking system was the knowledge that, by using the test machine, they were not cluttering up their own personal bookmarks. Despite being encouraged to surf as they normally would, almost all users felt they used more bookmarks in the study than they normally would. The primary reason users gave for not using bookmarks was that they did not like to add a bookmark, unless they were assured of its utility in being placed into the bookmarking system, which can easily become overcrowded.

As expected, users failed to recall a number of PVR candidate pages. Although they were able to recall over 70% of images, they were unable to locate those pages in four out of five tries. Textual recalls were slightly lower (50%), but they were more easily retrieved—due, in part, to the fact that textual passages were taken from strings often included in the page titles, which showed up on history lists and bookmarks, as well as in search engines. A few users turned to the history list feature of the Internet Explorer when they could not recall the pages, but they found the feature somewhat difficult to use, as it ordered the Web pages alphabetically.

An example PVR account is provided in Exhibit 1.

PVR IN AN INCREASINGLY DIGITAL SOCIETY

As a phenomenon that creates a relationship between interaction with technology and interaction with society, PVR affects and is affected by both technical and social factors. Because the effectiveness of the Web can suffer as a result of poor transfer rate, connection speed can have great weight on determining the scope of PVR. A slow connection will likely limit the breadth of exposure to information because Web pages may take longer to download, and a user's time—or patience—on a computer may be limited. Furthermore, a slow

Exhibit 1: An Illustrative User Account

When the nature of the experiment was revealed to the test subjects in the post-session interviews, many immediately recalled first-hand experiences of PVR. Here we present an illustrative account that captures a real-world example of the type of situation being addressed and the fact that the social nature of PVR can become a factor that may yield a significant impact in a team environment.

MH was surfing the Web in search of venture capital sources. He encountered a student-funding site that was not applicable to his work but that he knew to be of interest to a colleague, JM. MH sent JM email with the Web site address, and he continued with his work. JM subsequently visited the site, and, to keep his mailbox tidy, deleted the email.

A month later, JM received a request from a third party for information the Web site would have provided. JM recalled the Web site but not its location and asked MH if he could find it again. MH tried to retrace his steps, but he was unable to find the site, which he suspected (but could not recall) might have been found on tangential excursion.

MH and JM then spent an unplanned, unexpectedly long and ultimately unsuccessful three hours trying to find the site. The experience was particularly frustrating, because MH had just wandered on to the site originally without any effort. It was only when MH subsequently recalled the exact name of the funding source that they were able to find the Web page.

Virtually all of the test subjects were able to relate to the notion of PVR; many had their own, similar accounts. Although further tests are planned and needed to better define and quantify the nature of PVR, it appears to be a phenomenon that many Web surfers have experienced, and the phenomenon may have an impact on a greater scale in a digitally-based environment.

connection may encourage greater scrutiny of information already downloaded, since it makes the traversal of links a more expensive operation. Consequently, the impact of PVR may be diminished in such an environment, because a user is able to gain greater familiarity with a smaller sampling of Web pages, presumably resulting in easier recall of pages.

The converse of the above—namely that PVR may increase with greater connection speed—is further exacerbated by the likelihood that users with greater technical proficiency (who tend to be the very ones with access to high-speed connection) may serve to increase the effect of PVR. This is

partly due to the fact that experienced users, due to their efficiency, may likely expose themselves to a greater number of Web pages, thus creating a knowledge base potentially too big to recall. Moreover, the perception of being able to retrieve a previously visited Web page is based, in large part, on a user's confidence in navigating the Web. Being comfortable with the Web may boost the confidence of experienced users of being able to return to pages they have already visited, a task that may actually be a deceptively difficult one.

While users with technical access and experience may often work or socialize in the company of those who evoke PVR situations, the introduction of Web technologies to an ever-wider audience requires a broader examination of the various economic and social considerations relevant to PVR. For example, will the effects of PVR be diminished for a user from an economically disadvantaged background, where access to Web technology and the means to build Web experience are limited? Moreover, does the economic condition affect the sort of social interaction available that may trigger a recall to a Web page that is peripheral to the user's focus of interest, thus affecting the impact of PVR? Or, will those members of the technology-aware community, with little desire to actively pursue technology (e.g., the elderly) regard the potentially novice-level users as conduits of information and pressure them to recall pages they may not have valued originally, thus increasing the impact of PVR?

In a sense, PVR creates a relationship between an individual surfing the Web and group interaction that may likely trigger a recall of knowledge gained from a surf session. Viewed in this manner, PVR sits in a feedback cycle, both affected by and contributing to the specifics of user technology as well as the dynamics of social environments. As users the world over become more Web-savvy and advances in interface design continue to provide greater accessibility (e.g., to users with disabilities), the effects of PVR may become more visible. Unless otherwise addressed, the negative impact of PVR may very well increase as the Web becomes a greater part of society's infrastructure and the lexicon of day-to-day interaction. Introducing tools that will allow users to more easily access the contents of their increased knowledge base can help to diminish the potential negative impact of PVR.

A NEW APPLICATION DESIGN

There is a possible new browser tool that will help users who encounter PVR situations. The tool is an application that keeps a full record of a surf session in a graphical manner so that it can be easily accessed. Its goal is to allow users to retrieve PVR Web pages with greater ease than they do with existing tools. Based on the above observations, design guidelines are formulated with which to create tools meant to offer navigational support by addressing problems arising from PVR

Create a complete but transient tool: Because a user's reluctance to clutter bookmarking systems is one of the fundamental causes of PVR, a utility that does not require modification of a permanent, information space will help alleviate the issue of PVR. Although bookmarking systems are, in fact, not permanent, studies have shown how users tend not to maintain bookmarks diligently, and they are reluctant to delete them (Cockburn and Greenberg 2000; Cockburn and McKenzie 2000). Transient, in this case then, refers to the fact that a user is not committing changes to a personal-information archive. By not requiring the user to take any specific action to mark a page before its value is known, a system can remain in the background and be invoked only when needed.

Allow users to retrace their surf session: Observations suggest that users tend to settle into a familiar methodology that predictably leads to the same results, even after they have gotten lost. In this way, the random associative nature of the Web, which forms much of its appeal, is diminished to allow for a more predictable path of much greater practical use in research. It was noticed that a user's desire to simplify a complex hypertextual space into a simpler linear space is often exhibited in the user's replies that are temporally-based—e.g., “Saw this page around lunch time.”

As a linear and familiar notion, time is perhaps the most powerful parameter used to localize a search space. Both the Internet Explorer and Netscape Internet browsers offer temporally-based history lists. Neither browser has anything more sophisticated than a textual list. In addition, Internet Explorer does not allow temporal breakdown within a day, thus virtually excluding the retracing of surf sessions that rarely last for long, continuous periods. Enhancing temporally-based tools may help utilize those portions of the Web that become part of their knowledge base but, they have yet to be accepted or properly judged for their worth. Offering timelines true to the actual surf session allows users to recreate their surf sessions and therefore to recreate the environment in which they originally found the pages they wished to relocate. Enhancing these timelines with playback facilities allows users to retrace the surf sessions as they originally evolved. Finally, the timelines should be scalable to accommodate the scope of the user's search, allowing the user to quickly navigate from units of days to minutes with a consistent interface.

Use of thumbnails to aid in the recognition of images: Because images are more difficult to search for directly, since they lack the inherent, lexical ordering, offering visual cues, in the way of thumbnails to encourage recall would be of great benefit. Studies have shown that users respond to thumbnails (Hightower et al. 1998; Jones et al. 2001; Kaasten et al. 2001). It is

believed that the incorporation of some sort of visual cue is essential in a browser utility meant to help users relocate previously visited pages.

Based upon the above guidelines, a prototype based on the design shown in Figure 2 has been designed and implemented. The application keeps a complete list of the user's history and represents it graphically, so that the duration spent on a page is easily seen. Thumbnail screenshots of the entire desktop shows all browser windows, as well as other applications that can serve as visual cues to the user. Playback controls allow the user to view the session in compressed time for reviewing the original surf session and finding desired Web pages previously visited.

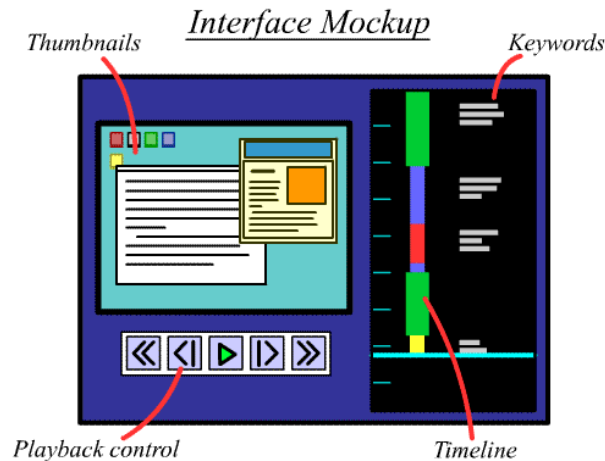
Related work: Considerable work has been undertaken to understand how users return to previously visited Web pages (Catledge and Pitkow 1995; Cockburn and Greenberg 2000; Greenberg and Cockburn 1999; Tauscher 1996) as well as how users can return to Web pages bookmarked or otherwise tagged in the initial visit (Abrams 1997; Jones et al. 2001). Graphical history tools have been proposed to help users return to visited Web pages (Hightower et al. 1998; Wen 1998).

However, graphical history utilities offer a topological rather than a temporal map of a surf session, exhibiting the paths taken and choices made over a complex associative space. Such maps often lose temporal information regarding traversal sequence. Consequently, the user needs to recall the semantic decisions made, which may be a difficult task, almost as complex as navigating through the Web hyperspace.

WebView is a project that embellishes the history list with thumbnail images but, by using a recently ordered list, it fails to keep true to the temporal activities of the user (Cockburn et al. 1999). Netscape has a textually-based history list, whereas the Internet Explorer has a history list sorted by titles, making it difficult to find particular Web pages based on temporal information.

A temporal back button, studied as an alternative to the standard stack-based button, yielded promising results for temporally distant Web pages (Cockburn et al. 2002). However, in its seamless integration into the browser's existing interface, the temporal back button is not the ideal vehicle for recreating a user's complete surf history in a contextual manner, as it discards duplicate Web sites.

In summary, a significant cause of user frustration is caused by Post-Valued Recall Web pages, defined as Web pages whose significance is not appreciated until the pages are inaccessible and need to be relocated. PVR can be caused by external triggers and the (possibly misplaced) judgment of PVR relocation time is often in response to some interested third parties. Thus, PVR often rests on the interface between the solitary act of surfing the

FIGURE 2: INTERFACE MOCKUP

Web and the social interaction experienced subsequently by the Web surfer. The user study described here provides supporting evidence of the impact of PVR in the workplace, and it has yielded some design guidelines with which to create further user tests as well as browser utilities.

REFERENCES

- Abrams, D. 1997. *Human Factors of Personal Web Information Spaces*. Masters of Science Thesis, Department of Computer Science, University of Toronto. Available: <http://www.dabrams.com/research/bookmarks/thesis/default.htm>
- Barret, E. 1988, *Text, Context and Hypertext*. The MIT Press, Cambridge, MA.
- Catledge, L. D. and Pitkow J.E. 1995. *Characterizing Browsing Strategies in the World Wide Web*. World Wide Web 3 Conference Proceedings. Available: http://www.igd.fhg.de/archive/1995_www95/papers/80/userpatterns/UserPatterns.Paper4.formatted.html
- Cockburn, A., Greenberg, S., McKenzie, B., Smith, M.J. and Kaasten, S. 1999. *WebView: A Graphical Aid for Revisiting Web Pages*. Proceedings of the OZCHI '99 Australian Conference on Human Computer Interaction. Available: <http://www.cpsc.ucalgary.ca/Research/grouplab/papers/1999/99-WebView.Ozchi/Html/Webview.html>
- Cockburn, A. and Greenberg, S. 2000. *Issues of Page Representation and Organisation in Web Browser's Revisitation Tool*. Australian Journal

- of Information Systems 7(2). Available:
www.cosc.canterbury.ac.nz/~andy/papers/issuesOzCHI.pdf
- Cockburn, A. and McKenzie, B. 2000. *What Do Web Users Do? An Empirical Analysis of Web Use*. Int. J. Human-Computer Studies. Available:
www.cosc.canterbury.ac.nz/~andy/papers/ijhcsAnalysis.pdf
- Cockburn, A., McKenzie, B. and Smith, M.J. 2002. *Evaluating a Temporal Behaviour for Web Browsers' Back and Forward Buttons*. Available:
www.cosc.canterbury.ac.nz/~andy/papers/backEval.pdf
- Greenberg, S. and Cockburn, A. 1999. *Getting Back to Back: Alternate Behaviors for a Web Browser's Back Button*. Proceedings of the 5th Annual Human Factors and the Web Conference, Held at NIST, Gaithersburg, Maryland, USA, June 3th. Available:
<http://www.itl.nist.gov/iaui/vvrg/hfWeb/proceedings/greenberg/>
- Hightower, R.R., Ring, L.T., Helfman, J.I., Bederson, B.B. and Hollan, J.D. 1998. *Graphical Multiscale Web Histories: A Study of PadPrints*. ACM Hypertext '98 Conference. Available:
<http://www.cs.umd.edu/hcil/pad++/papers/hypertext-98-padprints/>
- Jones, W., Bruce, H. and Dumais, S. 2001. *Keeping Found Things Found on the Web*. Available: www.research.microsoft.com/~sdumais
- Kaasten, S., Greenberg, S. and Edwards, C. 2001. *How People Recognize Previously Seen Web Pages from Titles, URLs and Thumbnails*. Report 2001-692-15, Department of Computer Science, University of Calgary. Available:
http://pharos.cpsc.ucalgary.ca/Dienst/Repository/2.0/Body/ncstrl.ucalgary_cs/2001-692-15/pdf
- Tauscher, L. 1996. *Evaluating History Mechanisms: An Empirical Study of Reuse Patterns in World Wide Web Navigation*. Masters of Science Thesis, Department of Computer Science, University of Calgary. Available: <http://www.cpsc.ucalgary.ca/grouplab/papers/1996/96-Tauscher.Thesis/thesis.html>
- Wen, J. 1998. *Surf Visualization: Seeing More than History*. Available:
www.surfserf.com/papers/surfviz.html.