Milestones in Time: The Value of Landmarks in Retrieving Information from Personal Stores

Meredith Ringel[1], Edward Cutrell[2], Susan Dumais[2], Eric Horvitz[2]

[1]Stanford University, 353 Serra Mall, Stanford, CA, USA

[2]Microsoft Research, One Microsoft Way, Redmond, WA, USA

merrie@cs.stanford.edu, {cutrell, sdumais, horvitz}@microsoft.com

Abstract: We describe the design and analysis of timeline visualizations for displaying the results of queries on an index of personal content. The visualization was built on top of a personal search engine that provides a unified index of all the information a user has seen, including web pages, email, and documents. Results of searches are presented with an overview-plus-detail timeline visualization. A summary view shows the distribution of search hits over time, and a detailed view allows for inspection of individual search results. In a user study, we explore the value of extending a basic time view by adding public landmarks (holidays and important news events) and personal landmarks (photos and important calendar events).

Keywords: Timeline, landmarks, event journal, episodic memory, search, time-centric visualization.

1 Introduction

People employ a variety of strategies when searching through personal emails, files, or web bookmarks for a specific item. One strategy is to narrow the scope of the search by considering the time an item was viewed or modified. Although exact dates may not be remembered, people often recall the relative times of important events in their lives (*e.g.*, their children's birthdays, exotic travel, prominent news events such as the 9/11 attacks or the assassination of JFK). We explored the effects of providing important events as context to support searching through personal content.

Our interactive visualization provides a timeline-based presentation of search results, anchored by both public (news, holidays) and personal (appointments, photos) landmark events. Search results are provided by a new personal indexing and search system named *Stuff I've Seen* (SIS). SIS indexes the full text and metadata of all the documents, web pages, and email that a user has seen in order to provide a fast and easy way to search over personal content (Dumais et al., 2003).

We first review background research on episodic memory and timeline visualizations. Then we present a design that overlays personal and public landmarks on search results. Finally we present findings gathered during a user study about the value of adding landmark events to a default timeline view.

2 Related Work

The psychology literature contains abundant discussion of *episodic memory*, a conception of memory that holds that memories may be organized by *episodes*. Episodes include information such as the location of an event, who was present, and what occurred before, during, and after the event (Tulving, 1983). Research also suggests that people use routine or extraordinary events as "anchors" when trying to reconstruct memories of the past (Smith et al., 1978). Huttenlocher proposes that the time of a particular event can be recalled by framing it in terms of other events, either historic or autobiographical (Huttenlocher & Prohaska 1997).

In other related work, a study of memory about office activities within a desktop computing environment (Czerwinski & Horvitz, 2002) showed that people forgot a significant number of computing tasks they had performed one month in the past. However, when prompted by videos and

photographs of their work during the target time period, they were able to recall significantly more of the tasks that they had performed and were able to more accurately remember the sequence of those tasks. More generally, research on encoding specificity (Tulving et al., 1973) emphasizes the dependency between encoded content and cues that are used to retrieve memories. Memory also depends on the reinstatement of not only item-specific contexts, but also on more general context capturing the situation surrounding events (Davies & Thompson, 1988).

There is a large body of research on the presentation of results for efficient searching. This work includes studies of visualizing search results in a matrix where rows and columns can be ordered by a variety of user-specified parameters (Nowell et al., 1996), work on 2D and 3D interfaces for displaying search results (Sebrechts et al., 1999), and research on displaying categorical, summary, and/or thumbnail information with search results (Dumais et al., 2001; Dziadosz & Chandrasekar, 2002).

Our project centers on probing the value of timelines and temporal landmarks for guiding search over subsets of personal content. Our visualization leverages key ideas about episodic memory by annotating a basic timeline with personal and public landmarks when displaying the search results.

Time is a common organizational structure for applications and data. Plaisant et al.'s (1996) LifeLines takes advantage of the time-based structure of human memory by displaying personal histories with a timeline. Kumar et al.'s (1998) work on digital libraries uses timelines to visualize topics such as world history and stock prices, as well as metadata about documents in the library, such as publication date. Rekimoto's (1999) "time-machine computing" leverages the time-centric nature of people's activities by allowing users to find old documents via "time-travel" to a prior version of their desktop where the target items were present. Fertig et al.'s (1996) LifeStreams presents the user's personal file system in a timeline format.

A number of projects have focused on collecting and making available histories of events in browsing and reminding applications. "Forget-Me-Not" (Lamming et al., 1994) is a ubiquitous computing system that serves as a memory augmentation device by gathering information about daily events from other devices in the environment, and allowing perusal and filtering of those records. Meetings with coworkers (time, location, and names of people present), phone calls, and emails are examples of the type of data collected and available as memory cues. "Save Everything" (Hull et al., 2001) has a similar approach, collecting various data about documents and then allowing querying using personal metadata such as the manner of a document's acquisition (*e.g.*, fax versus email versus photocopying) or the relevant activities occurring at the time of the data's acquisition. Minneman and Harrison's Timestreams (Minneman et al., 1997) use everyday activities (*e.g.*, speaking, drawing sketches, typing notes) to index into audio and video streams.

In contrast to earlier efforts, our system uses a mix of personal and public landmarks as memory cues. We explore whether such context provides useful navigation cues for efficiently searching personal content. Prior efforts separately explored timeline-based visualizations, contextual cues for retrieval, and other methods for increasing search efficiency. We pursue a synthesis of these ideas via use of the metaphor of a timeline combined with contextual cues in searching over personal content.

3 Visualization

To investigate the value of annotating timelines with event landmarks, we developed a prototype that provides an interactive visualization of results output by SIS. The visualization, displayed in Figure 1, provides both overview and detail about search results. The left edge of the display shows the *overview timeline*, whose endpoints are labeled as the dates of the first and last search result returned. Borders between years are also marked on the overview if the search results span more than one year. Time flows from the top to the bottom of the display, with the most recent results at the top. The overview provides users with a general impression of the number of search results and their distribution over time. The highlighted portion of the overview corresponds to the subset of results that are expanded in the detailed area of the visualization. Users can interact with the overview timeline as if it were a scroll bar, by grabbing the highlighted region with their mouse cursor and dragging it to a different section of the timeline, thus changing the segment of time that is displayed in the detailed view.

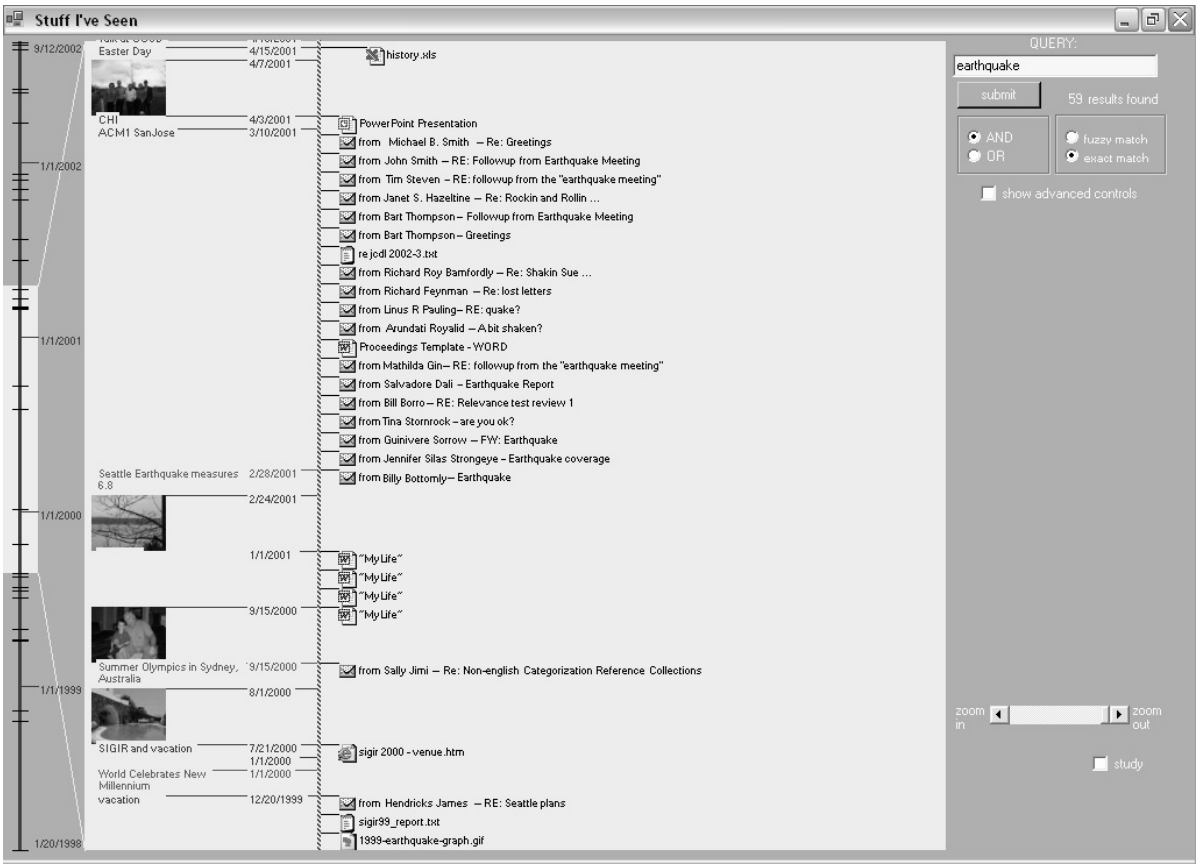


Figure 1: A screenshot of the timeline visualization. The overview area at the left shows a timeline with hash marks representing the distribution of the search results over time. The highlighted region of the overview timeline corresponds to the segment of time displayed in the detailed view. To the left of the detailed timeline backbone, beyond basic dates, context is provided with landmarks drawn from news headlines, holidays, calendar appointments, and digital photographs. To the right of the backbone, details of individual search results (represented by icons and titles) are presented.

Next to the overview we show *date and landmark* information. Landmarks appear to the left. Four types of landmarks may be displayed to the left of the dates: holidays, news headlines, calendar appointments, and digital photographs. Each type of landmark appears in a different color. Dates appear to the right, nearest the stippled line we call the timeline *backbone*. The granularity of dates viewed (hours, days, months, or years) depends upon the current level of zoom.

The detailed portion of the visualization shows a zoomed-in section of the timeline, corresponding to the slice of time highlighted in the overview area. To the right of the timeline backbone, each search result is positioned at the time when the document was most recently modified (for most files) or the time an email message was received. An icon indicating the type of document (html, email, word processor, etc.) is displayed, as well as the title of the document (or subject line and author, in the case of email). Hovering over a search result pops up a summary containing more detailed information about the object. This includes the full path, a preview of the first 512 characters of the document, as well as *to*, *from*, and *cc* information in the case of email messages. Clicking on a result opens the target item with the appropriate application.

3.1 Public Landmarks

Public landmarks are drawn from events that a broad base of users would typically be aware of. All public landmarks are given a priority ranking, and only landmarks that meet a threshold priority are

displayed. For our prototype, all users saw the same public landmarks, although we hope in future versions to explore letting users customize these; for instance, a user could add religious holidays that are important to them, or identify news headlines that they don't deem as important landmarks.

Holidays
We obtained a list of secular holidays commonly celebrated in the United States, and the dates those holidays occurred from 1994 through 2004, by extracting that information from Microsoft Outlook's calendar. Priorities were manually assigned to each holiday by the authors, based on their knowledge of American culture (*e.g.*, Groundhog Day was given a low priority, while Thanksgiving Day was given a high priority). Holidays and priorities could easily be adapted for any culture.

News Headlines
News headlines from 1994–2001 were extracted from the world history timeline that comes with Microsoft Encarta, a multimedia encyclopedia. Because 2002 events were not available in the latest release, the authors used their own recollections of current events to supply major news headlines from that year.

To prioritize the news headlines, 10 Microsoft employees (none of whom were participants in our later user study) rated a set of news headlines on a scale of 1 to 10 based on how memorable they found those events. The averages of these scores were used to assign priorities to the news landmarks.

3.2 Personal Landmarks

Personal landmarks are unique for each user. For our prototype, all of these landmarks were automatically generated, but for future versions we will allow users the option of specifying their own landmarks.

Calendar Appointments
The dates, times, and titles of appointments stored in the user's Microsoft Outlook calendar were automatically extracted for use as landmark events. Appointments were assigned a priority according to a set of heuristics. We increased an appointment's priority proportionally with the duration of the event, as longer events (such as conferences or vacations) seemed likely to be particularly memorable. For similar reasons, appointments designated as "out of office" times received a boost in score. Being flagged as a "tentative" appointment lowered priority, while being explicitly tagged as "important" increased the assigned priority. Several of the heuristics we used were initially identified in a study by Horvitz et al. (2003) on Bayesian models of memory landmarks. As an example, the study showed that recurrent appointments would be unlikely to serve as memory landmarks. Thus, we lowered the priority of meetings marked as recurrent.

Digital Photographs
Our prototype crawled the users' digital photographs (if they had any). The first photo taken on a given day was selected as a landmark for that day, and a thumbnail (64 pixels along the longer side) was created. Photos that were the first in a given year were given higher priorities than those which were the first in a month, which in turn were ranked more highly than those which were first on a day. Thus, as the zoom level changed, an appropriate number of photo landmarks could be shown. We did not explore more sophisticated algorithms for selecting photos to display, but we hope to explore techniques such as those developed by Graham et al. (2002) or by Platt (2000) in future iterations.

4 User Study

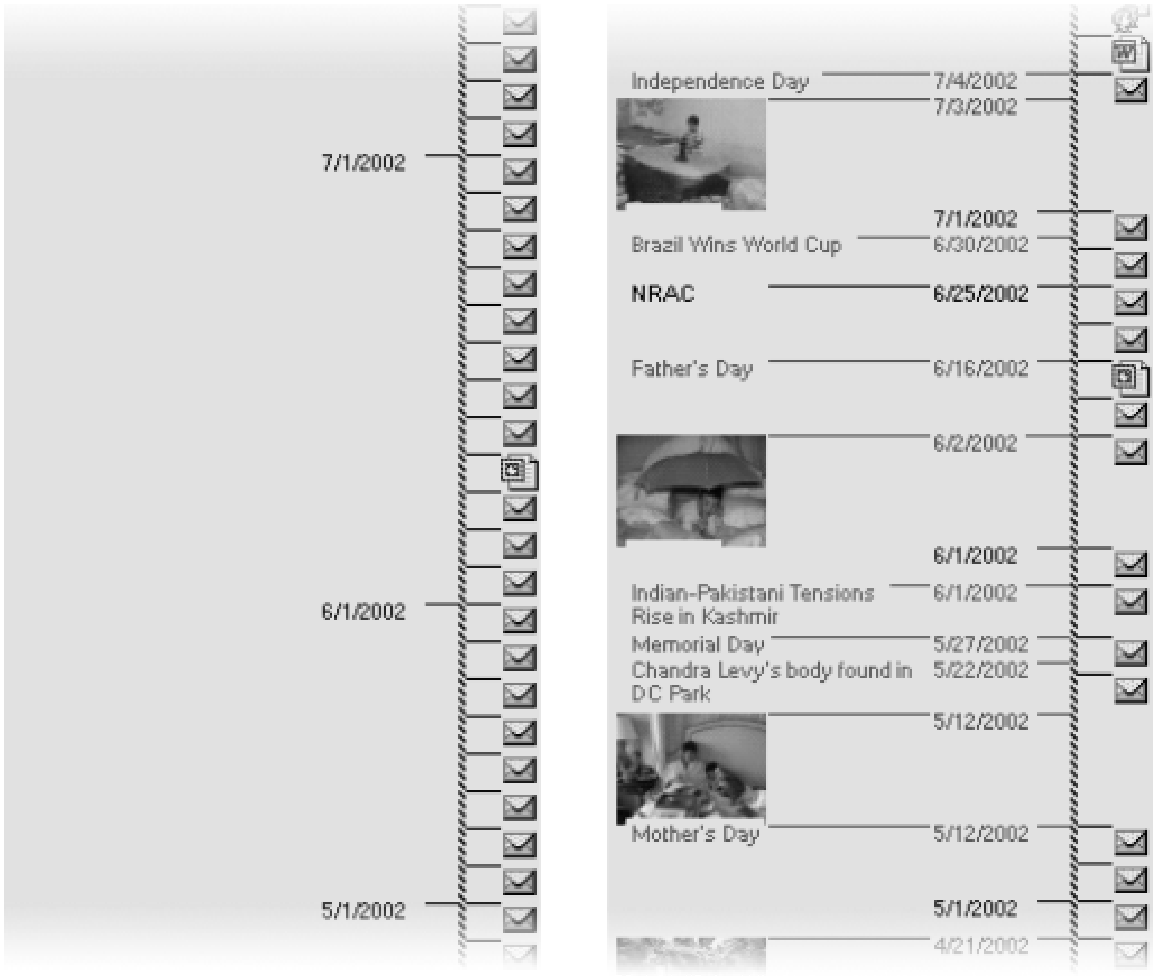
To evaluate the value of displaying landmark events, in addition to dates on the timeline visualization, we conducted a user study, gathering both quantitative and qualitative data.

4.1 Participants

Twelve Microsoft employees participated in the study. The subjects were all males, ranging in age from twenty-five to sixty.

4.2 Preparation

The day before their session at our usability lab, we sent subjects a .pst file (a repository of Microsoft Outlook email messages). This file contained a collection of messages that had been sent to a large number of people in the company (*e.g.*, announcements of talks, holiday parties, promotions, etc.). Although we knew that everyone had received these messages, we did not know whether individual participants had retained such mail; the .pst file was sent to guarantee that the target items were included in their index. The file, which contained 110 messages, was merged with each person's regular



(a) **(b)**

Figure 2: (a) The "Dates Only" experimental condition shows only dates to the left of the timeline's backbone. (b) The "Landmarks + Dates" experimental condition has a timeline that displays landmarks (holidays, news headlines, calendar appointments, and personal photographs) in addition to basic dates.

mail store, In the end, subjects' stores ranged from 5,844 to 70,469 messages, based on the differences in the amount of email messages they had retained.

4.3 Method

Participants were asked to fill out a questionnaire gathering demographic information as well as information about their searching and filing habits and about the ways they remembered information. Next they read a tutorial and performed two practice searches using the timeline interface. They were encouraged to take their time and to ask questions about the system. The experiment began after the tutorial was completed.

The experiment was a within-subjects design. Each participant completed a series of tasks using two different interfaces. For half of the tasks, they saw their search results presented in the context of a timeline annotated only by dates (Figure 2a), and for the other half they saw the timeline annotated by calendar appointments, news headlines, holidays, and digital photos (if they had any stored on their computer), in addition to the basic dates (Figure 2b). The conditions were counter-balanced to avoid

learning effects, so that half of the participants experienced the *landmark* condition before the *dates-only* condition, and the other half experienced the conditions in the reverse order. To avoid ordering effects, the sequence of questions was randomly changed for every pair of participants.

All subjects performed the same 30 search tasks. The tasks involved finding items included in the .pst file we had installed earlier. Fifteen questions were performed using each interface. For each task, we provided participants with a pre-formulated query to issue to the system, and instructed them not to change this query. We chose to use pre-set queries because our goal was to test different timeline presentations, and we did not want to confound the results by differences in how well users were able to formulate queries. Thus, we chose queries that would ensure that the target item would appear, but that were broad enough so that many other results from the participants' large stores of mail and documents were also likely to appear.

Once a query had been issued, users could navigate the timeline and inspect the search results by looking at the icons and titles, hovering for popup summaries with more detailed information, or clicking to open the actual document. When they had found the target item, they clicked a large button marked "Found It," and were automatically presented with the next task and query. If they were unable to locate the target item, there was also a button marked "Give Up," which allowed them to proceed to the next question. During the experiment, software logged all the details of their interaction, including the number of search results returned for each query, the number of landmarks of various types that were displayed, and information on the users' hovering, clicking, and overall timing of interactions.

After completing all of the tasks, subjects filled out a second questionnaire asking for feedback about the usability of the software, the utility of the timeline presentation and the various types of landmarks, and for free-form comments.

In summary, each of the 12 study participants was exposed to two experimental conditions—using the timeline with dates and landmarks, and using the timeline with dates only. In each condition, participants used the visualization to find specific targets by searching through the results generated by a series of fixed queries.

5 Results

5.1 Search Time

Analysis was performed on the median search times for each participant. We use median rather than mean search time to help mitigate the typical skewing of human performance times. A paired-sample *t*-test of the median search times for each participant indicated that times for the landmark condition were significantly faster than the date-only condition, $t(11)=2.33$, $p<0.05$. Figure 3 shows the effect graphically. We plot the average of each participant's median search time (±standard error about the mean). For the landmark condition this average was 18.37 seconds, while for the dates-only condition the value was 24.25 seconds.

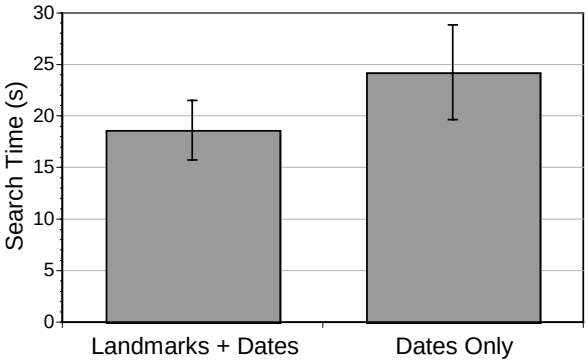


Figure 3: Median search time with the use of landmark events displayed on the timeline (left) was significantly faster than median search time when only dates were used to annotate the timeline (right).

5.2 Questionnaire

Participants completed questionnaires at the beginning and end of the experiment. Questionnaires included demographic information, a number of questions using a 7-point Likert scale (a score of 1= "Strongly Disagree" and 7 = "Strongly Agree,"), and free-form questions.

Pre-Questionnaire

Before seeing our visualization, subjects answered a series of questions about their current strategies for locating documents (Table 1). The three most highly rated attributes for searching were *topic*, *people* and *time*. Existing search tools support access by topic and people, but provide less support for time-oriented search. Our visualization helps remedy this

by allowing keyword-based search to generate an initial set of results, coupled with a rich timeline display for navigation among results.

I often find documents or email based on…	**Mean Response**
…when they happened	4.4
…who I was working with	5.3
…what topic they covered	5.5

Table 1. Users' assessment of their search habits before initiating the study (7-point scale).

Before the study, subjects were also asked to rate the importance of different types of landmarks for recalling events (Table 2). It is interesting to note that personalized events received higher ratings than public events.

Event Type	**Mean Rating**
World events	3.5
Holidays	4.3
Personal events (birthdays, vacations, etc.)	4.5
Work events (meetings, due dates, conferences, etc.)	5.2

Table 2. Subjects' ratings, in advance of the study, of the importance of several types of events for recall (7-point scale).

Post-Questionnaire

After finishing the experiment, participants evaluated the general usefulness of our timeline interfaces (Table 3). Participants generally found the time-based presentation of results useful. The vertical presentation of the timeline was well received; some users suggested allowing the option of reversing the flow of time such that more recent search results were displayed near the bottom.

Users generally found the overview provided in the visualization to be useful (one user commented, "I liked the way the little horizontal lines showed bursts of activity. That way I could figure out what time period stuff happened."). However, some users found it confusing to navigate through the search results by selecting a section of the overview timeline.

Question	**Mean Response**
The time organization was useful for these tasks	5.6
I liked having time be presented vertically.	5.9
The order (recent items at the top) made sense to me.	5.2
The overview timeline (at the far left) was useful to me.	5.4

Table 3. Participants' feedback on the visualization after completing the study session (7-point scale).

6 Conclusions and Future Work

We developed and evaluated a timeline-based visualization of search results over personal content. We augmented a basic timeline view with public (news headlines and holidays) and personal (calendar appointments and digital photographs) landmark events, in hopes that this added context would aid people in locating the target of their search. A user study found that there was a statistically significant time savings for searching with the landmark-augmented timeline compared to a timeline marked only by dates. Additionally, we gathered important feedback about the way users believe that they remember and search for events and about their reactions to our visualization. We believe that this work demonstrates the potential value of adding global and personal context to the presentation of search results, as well as suggesting directions for future study.

An avenue for future study is to explore the value of different types of events—*e.g.*, running separate "personal landmarks" and "public landmarks" conditions in addition to the two conditions explored here. In addition, there are opportunities for investigating more generally when timeline-centric views are most useful for finding target results of interest. It is likely, for example, that the distribution of items over time returned for a particular query will influence the overall utility of a timeline view for finding items.

There are a number of other opportunities for refining the prototype. Users reported some difficulty in navigating the timeline and we would thus like to improve the control of navigation via better coupling of zooming and translation in time. We are also interested in refining the heuristics for selecting and ranking landmarks, and in exploring different types of summary landmarks. For example, shading segments of the overview timeline different colors to indicate years or seasons within a year could prove fruitful. Also, landmarks related to the search results themselves could be identified, such as key attributes about the content and structure of documents. In addition to passively displaying landmarks, we hope to allow users to combine landmarks and more traditional search terms in the formulation of a query, enabling users to search "by landmark," *e.g.*, expressing such a query as, "show me all documents that I composed right before the project review with my manager" or "show me all emails I received the week of the earthquake."

Acknowledgments

We are grateful to the participants in our user study for their time and suggestions. Special thanks to Mary Czerwinski and Gina Venolia for their feedback on the prototype, and to Raman Sarin, Gavin Jancke, Paul Koch, and Johnson Apacible for their technical assistance.

References

Czerwinski, M., and Horvitz, E. An Investigation of Memory for Daily Computing Events. *Proceedings of HCI 2002*, 230-245.

Davies, G. and Thomson, D., eds. *Memory in Context: Context in Memory*. Wiley: Chichester, England, 1988.

Dumais, S., Cutrell, E., and Chen, H. Optimizing Search by Showing Results in Context. *Proceedings of CHI 2001*, 272-284.

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., and Robbins, D. Stuff I've Seen: A System for Personal Information Retrieval and Re-use. *Proceedings of SIGIR 2003* (to appear).

Dziadosz, S., and Chandrasekar, R. Do Thumbnail Previews Help Users Make Better Relevance Decisions about Web Search Results? *Proceedings of SIGIR 2002*, 365-366.

Fertig, S., Freeman, E., and Gelernter, D. Lifestreams: An Alternative to the Desktop Metaphor. *Proceedings of CHI 1996*, 410-411.

Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T. Time as Essence for Photo Browsing Through Personal Digital Libraries. *Proceedings of JCDL 2002*, 326-335.

Horvitz, E., Dumais, S., Koch, P. Harnessing Models of Memorability in Computing Applications (submitted for publication).

Hull, J. and Hart, P. Toward Zero Effort Personal Document Management. *IEEE Computer*, March, 2001, 30-35.

Huttenlocher, J., and Prohaska, V. *Reconstructing the Times of Past Events. Memory for Everyday and Emotional Events.* Mahwah, NJ: Lawrence Erlbaum Associates (1997), 165-179.

Kumar, V., Furuta, R., and Allen, R. Metadata Visualization for Digital Libraries: Interactive Timeline Editing and Review. *Proceedings of DL 1998*, 126-133.

Lamming, M., and Flynn, M. "Forget-me-not" Intimate Computing in Support of Human Memory. *Proceedings of Friend21*, 1994.

Minneman, S. and Harrison, S. Space, Timestreams, and Architecture: Design in the Age of Digital Video. *Proceedings of IFIP 1997 Workshop on Formal Design Methods for CAD*, 1997.

Nowell, L., France, R., Hix, D., Heath, L., and Fox, A. Visualizing Search Results: Some Alternatives to Query-Document Similarity. *Proceedings of SIGIR 1996*, 67-75.

Plaisant, C., Milash, B., Rose, A., Widoff, S., and Shneiderman, B. LifeLines: Visualizing Personal Histories. *Proceedings of CHI 1996*, 221-228.

Platt, J. AutoAlbum: Clustering Digital Photographs Using Probabilistic Model Merging. *IEEE Workshop on Content-Based Access of Image and Video Libraries 2000*, 96-100.

Rekimoto, J. Time-Machine Computing: A Time-centric Approach for the Information Environment. *Proceedings of UIST 1999*, 45-54.

Sebrechts, M., Vasilakis, J., Miller, M., Cugini, J., and Laskowski, S. Visualization of Search Results: A Comparative Evaluation of Text, 2D, and 3D Interfaces. *Proceedings of SIGIR 1999*, 3-10.

Shum, M. The Role of Temporal Landmarks in Autobiographical Memory Processes. *Psychological Bulletin*, *124*, 1994, 423-442.

Smith, S., Glenberg, A., and Bjork, R. Environmental Context and Human Memory. *Memory and Cognition* (1978), 6(4), 342-353.

Tulving, E. *Elements of Episodic Memory*. Oxford University Press, 1983.

Tulving, E. and Thomson, D. Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review 80*, 1980, 352-373.