

SearchPanel: A browser extension for managing search activity

Simon Tretter
University of Amsterdam
Amsterdam, The Netherlands
s.tretter@gmail.com

Gene Golovchinsky
FX Palo Alto Laboratory, Inc.
3174 Porter Drive
Palo Alto, CA
gene@fxpal.com

Pernilla Qvarfordt
FX Palo Alto Laboratory, Inc.
3174 Porter Drive
Palo Alto, CA
pernilla@fxpal.com

ABSTRACT

People often use more than one query when searching for information; they also revisit search results to re-find information. These tasks are not well-supported by search interfaces and web browsers. We designed and built a Chrome browser extension that helps people manage their ongoing information seeking. The extension combines document and process metadata into an interactive representation of the retrieved documents that can be used for sense-making, for navigation, and for re-finding documents.

1. INTRODUCTION

Broder *et al.* [3] proposed a taxonomy of web search that included transactional and navigational searches in addition to the more traditional (from an IR perspective) informational searches. To this taxonomy we might add re-finding [17] [5], the task of locating a previously-found document. From a theoretical perspective, it is not clear whether re-finding is a different kind of search activity or an orthogonal dimensions. Regardless, while major web search engines offer simple and efficient interfaces for navigational and transactional searches, relatively little support is available for more complex informational search or re-finding.

These seemingly neglected activities are not unimportant, however: Teevan *et al.* [17] reported that 39% of queries are re-finding queries; furthermore, 20-30% of searches represent open-ended informational needs [13]. Related, Qvarfordt *et al.* [11] found query overlap rates of 50-60% in exploratory search, and suggested that awareness of this overlap may be useful in supporting more efficient searching behavior. Thus we decided to explore ways in which searchers' interactions with search engines could be enhanced to support these more complex information-seeking tasks.

We created a web browser extension that enriches common web search engine interfaces and addresses important deficits with respect to open-ended (exploratory) search and re-finding. Our extension visualizes search results to help users find the right document or documents by visualizing metadata of the retrieved pages.

Following Golovchinsky *et al.* [7] we distinguish *document metadata* from *process metadata*. Document metadata – dates of publication, titles, hosting web sites, etc. – are basic characteristics of documents that are independent of the means by which these documents were retrieved. Process metadata, on the other hand, characterize aspects of

documents in relation to the searcher's activity: how many times was a document retrieved, whether it was viewed before, etc. This kind of information can help searchers to remember, understand and plan their search processes.

The browser plugin enhances the searcher's ability to use process metadata to understand their search results and to plan subsequent activity by displaying surrogates for the current set of retrieved documents. We represent prior retrieval state, whether a document was opened, and whether it was bookmarked in an integrated overview that appears at the side of the browser window. We also make it possible for searchers to examine multiple documents without returning to the search results or using multiple tabs.

The remainder of this paper is organized as follows: we review the relevant related work, describe the browser extension, and conclude with a discussion of the design space.

2. RELATED WORK

There are two broad categories of related work: the management of search history and the representation of search results. Refinding has received increasing attention recently. While the browser implements some history mechanisms, these are typically not well-suited to users' needs [15]. El-sweiler and Ruthven [5] described different patterns of re-finding; Teevan [16] proposed a mechanism for merging previously-found and newly-retrieved documents. More explicit management of search history has also been investigated in the literature; see [7] for a succinct summary.

Information overload due to large numbers of results is a common problem in information seeking [2]. This problem can be addressed in a variety of ways. MetaSpider [4] uses a 2D map to display and classify retrieved documents. Grokker [8] uses nested circular and rectangular shapes to present results and also shows them in a hierarchical grouped way. Sparkler [12] uses a star plot for the result presentation, where every star represents a document.

One potential issue with the systems above is that the overall organization of the interface itself may induce usability problems. Complex interfaces allow more individual settings to be specified by a user, but simple interfaces allow a broader spectrum of users to use them. This tradeoff is not trivial to handle, and as we see nowadays, most Web search interfaces tend to be quite simple.

Supporting the searcher's decision making process can be crucial for effective search performance for complex information needs. This support can take the form of enhanced surrogates for documents. One type of information often used for this purpose is document metadata (author, date,

images of the document, etc.). Even *et al.* [6] has shown that the decision making process can be highly improved by adding process metadata (in our case information that is related to the search process) to the user interface. Research has shown that presenting simple tasks in a slightly different way may help the user to understand how the search is performing and what can be done to gain better results [18]. One common example of incorporating process metadata in web browsers is the practice of changing the color of a traversed link anchor.

Spoerri [14] showed that users can benefit from different or additional visualizations of web search results. However, none of the techniques above have been integrated by major search engines into their main interfaces. In some cases, extension developers have enhanced the user experience of web search. Examples include: SearchPreview[9] that fetches screen shots of the result pages and shows them directly next to the each search result. Bettersearch[1] is a Firefox extension that performs a similar task, but also enriches the result page with more features and links. For example, this extension allows users to open a result in a new tab, or adds links to a search result to quickly show the web page on the "Wayback Machine"¹. WebSearch Pro [10] is also a Firefox extension that adds the ability to look up a text by highlighting it on a page. Another feature is drag&drop zones to search for things directly from any website.

3. BROWSER EXTENSION

To compensate for the deficiencies of SERPs we created a browser extension called SearchPanel. This extension combines document and process metadata in a visual representation of search results to help people manage their information seeking. We chose the browser extension approach rather than creating a proxy for several reasons. While both offer the potential of parsing and augmenting SERP and document pages, a browser extension has some advantages. It scales better with respect to storing user history data. It ensures a higher level of data privacy, since data that might potentially reveal user interests (e.g., query keywords, selected URLs, etc.) can be logged as hashed values. Finally, it has access to bookmarks and local browsing history.

3.1 Design space

When performing search tasks, searchers may need different kinds of information to support their information seeking. We represent the design space as consisting of three categories of activities: search activity, navigation activity, and organization activity.

Historically, web UI support for the search process, or search activity, has been focused on query formulation and understanding the current query. Web browsers offer limited support for comparing current results set with earlier activity by marking the visited status of documents.

When engaged with a search task, users need to shift their attention between the SERP and the retrieved pages. In some cases, the searcher does not find the desired information in a retrieved document, but rather in links to other documents containing relevant information. This navigation activity can be an important part of the information seeking process.

¹The Wayback Machine is a service that provides access to archived and historical versions of web sites.

Table 1: Design space: Activities and supporting features related to document and process metadata. "Doc" refers to document metadata and "Proc" to process metadata.

Activity	Feature	Doc	Proc
Search	perform search	yes	no
	switch engine	no	yes
	results list	yes	no
	visit status	no	yes
	visualize no. of visits	no	yes
Navigation	access results	-	-
	mark current result path	no	yes
	<i>identify results</i> : preview snippet	yes	no
	<i>identify results</i> : favicon	yes	no
Organization	bookmarking	no	yes
	organize bookmarks	no	yes

When searchers find useful web pages, they may wish to save those documents for future access. More specialized search engines sometimes support this capability directly, but it is most often supported only by the browser's bookmarking capability.

We can consider these search and sense-making activities in light of the kinds of information required to satisfy them. In particular, Table 1 shows when document and process metadata might be pertinent for the different categories of search activities. A representation of the number of visits to a retrieved result (process metadata) could be used by a searcher to decide how to interact with that result. In a re-finding sub task, for example, searchers might want to ignore newly-found documents or pages that were not opened.

The purpose of the search panel is to complement the SERP and to be available when exploring search results; we wanted the design to be simple and unobtrusive but still convey useful information. Some features (e.g., organizing bookmarks) listed in Table 1 are too complex to be integrated into the extension. Others, such as favicons, while seemingly trivial, may still provide useful information for navigating search results.

3.2 Implementation

SearchPanel displays automatically on the right side of the browser window when it is enabled (Figure 1). The right side of the content page has been chosen because this location is frequently free of document content. In cases of overlap, its vertical position can be adjusted manually to accommodate page content that may be occluded.

SearchPanel displays immediately after a search has been performed on a supported web search engine (currently, they are Google, Google Scholar, Yahoo, Bing and Microsoft Academic Search). SearchPanel remains visible even if the searcher follows links from retrieved documents. In addition, searchers can return directly to the original query, or re-run it on a different search engine.

A short tutorial page is displayed at installation, and can also be reached through the option menu. This page also allows logging (see 3.2.4) to be disabled, and can be used to delete the recorded history.

3.2.1 Document metadata

SearchPanel displays several kinds of document metadata. Documents are represented by bars arranged in order corre-

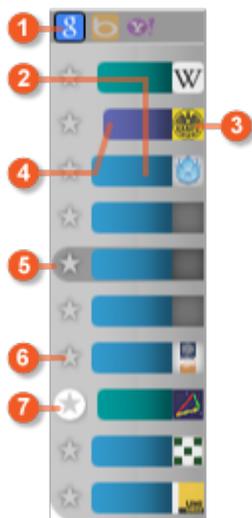


Figure 1: SearchPanel control annotated to show important aspects. ① search engine selector; ② bar representing a newly-found page; ③ favicon representing the site from which the page was retrieved; ④ bar representing page that has been visited; ⑤ highlighted bar based on cursor position; ⑥ bookmark indicator; ⑦ currently-selected page.

sponding to the retrieved list; clicking on a bar is equivalent to clicking on a link on the SERP. Almost all websites have icons (favicons) to help re-identify the web page quickly; these icons are shown to the right of the bar (see Figure 1, item ③). A tooltip with the title of the document is added to each bar as well. We considered identifying other metadata such as document MIME type, but that would incur the overhead of a separate HTTP request for each document. At least initially, we chose not to pursue this strategy.

3.2.2 Process metadata

Process metadata is also incorporated into SearchPanel. First, the icon of the search engine that ran the search is highlighted in the top bar (item ①). Other icons represent available comparable search engines. Clicking on one of these icons re-runs the query with the selected search engine. Search engines are grouped into two categories (web search and academic research) and only the relevant ones are shown. The current selection (highlighted with a black border) links back to the search result page if the user navigates to one of the retrieved documents.

Each bar can have one of three different colors, depending on the link history. If a link has never been retrieved before, the state of the link is "new" and the color will be teal. Results that have been retrieved by prior queries but have not been clicked on are colored blue. Visited links are colored violet. The local browser history is examined to retrieve the link status. This allows us to incorporate page views that occurred before SearchPanel was installed.

Each bar's length reflects the frequency of retrieval of the corresponding page. The more frequently a page has been retrieved, the shorter the bar gets (item ③). The retrieval history is stored locally in the browser for privacy reasons and can be deleted through SearchPanel's option page.

In SearchPanel, the bookmarking function serves two purposes (item ⑥ in Figure 1). First, searchers can click on



Figure 2: Highlighting of snippet on the SERP when mousing over SearchPanel.



Figure 3: Snippets of other pages are shown on a document page when mousing over SearchPanel.

the star to bookmark the corresponding page. Second, previously bookmarked documents in the SERP will show a yellow star next to them. This allows to re-find a web page quicker, as the user does not need to navigate to a document to know if they have previously bookmarked it.

3.2.3 Navigational support

The selection indicator (see item ⑦ in Figure 1) indicates the currently-selected result page. If a link on a result page is clicked, the page indicator will stay on the last retrieved document page to indicate that navigation started with it. Hovering over the result highlights the associated bar (item ⑤), and also highlights the corresponding snippet in the SERP (Figure 2); the SERP is scrolled as necessary to bring highlighted snippet into view. Conversely, when the mouse is over a snippet on the SERP, the related bar jiggles left-right to reinforce the connection between the two.

When the user navigates off the SERP to a search result, SearchPanel remains active. Clicking on bars navigates among the retrieved documents, bypassing the intermediate step of reloading the search results. When the mouse is over a bar in SearchPanel, the SERP snippet of that result will be shown. This can be seen in Figure 3, where a preview of the Wolfram Alpha snippet is shown. If the snippet is not available, a tooltip with the document title is shown instead. Both of these features should make it easier and more efficient to navigate the search results without necessarily creating a large number of tabs in the process.

3.2.4 Logging

The extension was created to study people's information seeking behaviors. The goal of the project is to understand how people use the web when looking for information to improve their search experience. Therefore logging of user activity was necessary. To encapsulate it from the basic functionality it was designed as plugin that could be connected or disconnected from SearchPanel. It collects information related to the use of SearchPanel for the purposes of statistical analysis of patterns of behavior.

To maximize searchers' privacy, no personally-identifying information is saved. Queries and found URLs are recorded as MD5-hashed values only. This allows us to identify re-

curing queries and documents, without being able to read the content of the query or to observe which pages people view. Specifically, the following information is recorded:

- The IP address and the time the event was logged
- When a search result was clicked and where this happened (SearchPanel or SERP)
- Hash strings that represent the queries and found web pages.
- Time spent with the mouse on different interface parts (SearchPanel vs SERP)
- Various actions related to the extension (adding bookmarks by clicking the start, moving it, etc.).

4. NEXT STEPS

After an in-house pilot deployment, SearchPanel has been made available through the Google Chrome store. The goal of the deployment is to understand whether the extension helps people with their search tasks, and to assess the relative utility of document vs. process metadata. We also expect to collect a dataset that characterizes people's browsing and searching behaviors in terms of patterns of retrieval and re-retrieval, search result navigation, etc.

5. CONCLUSIONS

Web search engines are used for many different kinds of search tasks. While navigational and transactional uses of search engines are well-supported by current interfaces and algorithms, searchers are left to their own devices for more open-ended information seeking and re-finding. We created a Google Chrome browser extension to help people manage their search activity. We explored the design space of document and process metadata related to the wide range of activities searchers may engage in during information seeking. The extension keeps track of retrieval, page visits, and bookmarking, and integrates traces of these activities with document metadata to give people a more complete impression of their search activity. An upcoming deployment will explore the effect that this extension has on how people interact with search results.

6. REFERENCES

- [1] ABAKUS. Bettersearch a firefox addon for enhancing search engines. <http://mybettersearch.com/>, 2010. [Online; accessed 06/06/2013].
- [2] BAEZA-YATES, R., RIBEIRO-NETO, B., ET AL. *Modern information retrieval*, vol. 463. ACM press New York, 1999.
- [3] BRODER, A. A taxonomy of web search. *SIGIR Forum* 36, 2 (Sept. 2002), 3–10.
- [4] CHEN, H., FAN, H., CHAU, M., AND ZENG, D. Metaspider: Meta-searching and categorization on the web. *Journal of the American Society for Information Science and Technology* 52, 13 (2001), 1134–1147.
- [5] ELSWEILER, D., AND RUTHVEN, I. Towards task-based personal information management evaluations. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (New York, NY, USA, 2007), SIGIR '07, ACM, pp. 23–30.
- [6] EVEN, A., SHANKARANARAYANAN, G., AND WATTS, S. Enhancing decision making with process metadata: Theoretical framework, research tool, and exploratory examination. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on* (2006), vol. 8, IEEE, pp. 209a–209a.
- [7] GOLOVCHINSKY, G., DIRIYE, A., AND DUNNIGAN, T. The future is in the past: designing for exploratory search. In *Proceedings of the 4th Information Interaction in Context Symposium* (New York, NY, USA, 2012), IIX '12, ACM, pp. 52–61.
- [8] HONG-LI, Q. A novel visual search engines: Grokker. *Journal of Library and Information Sciences in Agriculture* 8 (2008), 047.
- [9] KG, P. U. . C. Searchpreview, the browser extension previously known as googlepreview. <http://searchpreview.de/>, 2013. [Online; accessed 06/06/2013].
- [10] MARTIJN. Web search pro, search the web the way you like... <http://websearchpro.captaincaveman.nl>, 2012. [Online; accessed 06/06/2013].
- [11] QVARFORDT, P., GOLOVCHINSKY, G., DUNNIGAN, T., AND AGAPIE, E. Looking ahead: Query preview in exploratory search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in Information Retrieval* (New York, NY, USA, 2013), SIGIR '13, ACM.
- [12] ROBERTS, J., BOUKHELIFA, N., AND RODGERS, P. Multiform glyph based web search result visualization. In *Information Visualisation, 2002. Proceedings. Sixth International Conference on* (2002), IEEE, pp. 549–554.
- [13] ROSE, D. E., AND LEVINSON, D. Understanding user goals in web search. In *Proceedings of the 13th international conference on World Wide Web* (2004), ACM, pp. 13–19.
- [14] SPOERRI, A. How visual query tools can support users searching the internet. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on* (2004), IEEE, pp. 329–334.
- [15] TAUSCHER, L., AND GREENBERG, S. How people revisit web pages: empirical findings and implications for the design of history systems. *Int. J. Hum.-Comput. Stud.* 47, 1 (July 1997), 97–137.
- [16] TEEVAN, J. The re:search engine: simultaneous support for finding and re-finding. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2007), UIST '07, ACM, pp. 23–32.
- [17] TEEVAN, J., ADAR, E., JONES, R., AND POTTS, M. A. S. Information re-retrieval: repeat queries in yahoo's logs. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (New York, NY, USA, 2007), SIGIR '07, ACM, pp. 151–158.
- [18] WANG, T. D., DESHPANDE, A., AND SHNEIDERMAN, B. A temporal pattern search algorithm for personal history event visualization. *Knowledge and Data Engineering, IEEE Transactions on* 24, 5 (2012), 799–812.