

HCI Browser: A Tool for Studying Web Search Behavior

Robert Capra

School of Information and Library Science
University of North Carolina at Chapel Hill
rcapra3@unc.edu

ABSTRACT

We introduce the HCI Browser, a Mozilla Firefox extension designed to support studies of Web search behaviors. The HCI Browser presents tasks to the user, collects browser event data as the user searches for information, records answers that are found, and administers pre- and post-task questionnaires. The HCI Browser is a configurable tool that researchers can use to conduct studies and gather data. It is especially well suited for “batch mode” laboratory studies in which multiple participants complete a study at the same time, but work independently. The HCI Browser is open-source software and is available for download at: <http://ils.unc.edu/hcibrowser>

Keywords

Data collection tools, Web search behaviors.

INTRODUCTION

Laboratory studies of Web information seeking often involve presenting tasks to participants and observing their behaviors as they use a Web browser to search for information requested by the task. Researchers may be investigating a new Web interface to information, or may be investigating search behaviors on the open Web. In these types of studies, researchers are often interested in factors such as the Web pages visited, links clicked, the amount of time spent on each page, and the use of browser interface elements such as the back button or typing in the URL address bar. In addition, researchers may wish to administer brief questionnaires before or after each task to gather information about participants’ experiences.

Researchers have used a variety of tools to help support the data gathering and have noted the challenges involved with capturing naturalistic user behaviors for Web search (Keller et al. 2008). We reviewed a number of these approaches (Capra, 2009) and found that they did not meet the needs we had for a study of how users refind information on the Web. Specifically, we needed a tool that would: 1) integrate with an existing Web browser to provide a familiar browsing experience to participants, 2) record a wide variety of user interactions with the Web pages and the browser interface, 3) provide support for administrative aspects of conducting a laboratory study (e.g. administering

pre- and post-task questionnaires, recording the answers participants found for the tasks, and managing other details such as closing any opened browser windows before the start of the next task). To support these needs, we developed the HCI Browser as a Firefox extension and have released it as freely available open-source software (<http://ils.unc.edu/hcibrowser>). In developing the HCI Browser, we built off our previous experience building an instrumented web browser (Capra, 2008) and also utilized concepts from the open-source Lemur Query Log Toolbar project (Lemur, 2010).

HCI BROWSER OVERVIEW

The HCI Browser is designed to: present experimental tasks to users, collect and log browser event data, manage the opening and closing of windows as needed, and to administer optional pre- and post-task questionnaires.

When the HCI Browser is started a dialog box is shown that prompts the experimenter to enter a session number, participant number and starting task number. The session number and participant id are used to label the log data that is recorded for the session and can be any string that the experimenter wishes to use.

Next, a configurable introduction screen is displayed that is a useful place to provide instructions and general information to the participant before beginning the tasks.

When the participant clicks "OK" on the introduction screen, an optional set of pre-task questions can be presented (Figure 1). These questions are configurable and three question types are supported:

- MultipleChoice – displays the question text followed by a vertical list of the choices
- LikertType – displays the question text followed by a horizontal list of the choices (note: this can also be used for semantic differential scales)
- OpenAnswer – displays the question followed by a free-response text box

After answering the pre-task questions, the participant is taken to the main browser window which displays the task in the toolbar area (Figure 2). This is a standard Firefox browser and the participant can search, browse, and navigate as usual. The task and controls are displayed in the toolbar area at the top of the window so that the participant

can always see the task and so that web pages designed to fit standard resolutions do not require horizontal scrolling.

Figure 1. Pre-Task Questions – Example Form

In Figure 2, a participant has navigated to a particular web site that has information requested by the task. Participants can indicate that they have "Found an answer on this page", or are "Done with answers for this task" (cutoff in Figure 2) by pressing buttons on the toolbar. The number of answers submitted for this task are displayed in the lower left of the toolbar, along with an indication of the maximum number of answers they may submit (these are configurable).

Figure 2. Task Presentation Integrated with Toolbar

When the participant finds an answer and clicks the "Found an answer" button, the toolbar changes to allow them to enter information about the answer found. Additional controls allow the participant to change their mind about a

found answer. When they are done finding answers or have found the maximum number of answers for a task (configurable), then the system will automatically close all opened tabs and windows and display an optional set of post-task questions using an interface similar to that of the pre-task questions.

When the participant has completed the post-task questions, the program will automatically advance to the next task. When the last task is reached, a message is displayed letting the participant know that they have completed all the tasks.

DATA LOGGING

In addition to logging data from the questionnaires, the HCI Browser monitors and logs of a wide array of browser events while participants work on the tasks (see Figure 3 for an example log file excerpt). The current version logs: pages loaded, links clicked, window and tab focus changes, opening/closing of windows and tabs, back/forward button clicks, URLs typed in the address bar, scrolling, and history/bookmark menu activity. A new log file is automatically created for each task, and log entries include a timestamp, session number, participant number, and task number.

OPEN SOURCE CODE

The HCI Browser is available as open-source code. For additional screenshots, documentation, downloads, and updates visit: <http://ils.unc.edu/hcibrowser>

REFERENCES

Capra, R. (2008). Studying Elapsed Time and Task Factors in Re-Finding Electronic Information. In *Personal Information Management 2008 CHI Workshop*, April 5-6, 2008, Florence, Italy.

Capra, R. (2009). The HCI Browser Tool for Studying Web Search Behavior. In *Proceedings of the Third Workshop on Human-Computer Interaction and Information Retrieval (HCIR 09)*, Oct 23, 2009, Washington DC.

Keller, M., Hawkey, K., Inkpen, K., and Watters, C. (2008). Challenges of Capturing Natural Web-Based User Behaviors. *International Journal of Human-Computer Interaction*, 24(4), 385-409.

Lemur Query Log Toolbar. (2010). <http://www.lemurproject.org/querylogtoolbar> Accessed July 16, 2010.

1248198389580	21-7-2009	13:46:29	S4	P27	T1	intask	Focus	http://news.google.com/nwshp?client=firefox-a&rls=org
1248198390167	21-7-2009	13:46:30	S4	P27	T1	intask	LoadCap	http://news.google.com/nwshp?client=firefox-a&rls=org
1248198401531	21-7-2009	13:46:41	S4	P27	T1	intask	Scroll	http://news.google.com/nwshp?client=firefox-a&rls=org
1248198411929	21-7-2009	13:46:51	S4	P27	T1	intask	RClick	http://news.google.com/news/url?sa=t&ct2=us%2F1_0_s_0
1248198413706	21-7-2009	13:46:53	S4	P27	T1	intask	AddTab	http://news.google.com/nwshp?client=firefox-a&rls=org

Figure 3. Example Log File Excerpt