

Hunter Gatherer: Within-Web-Page Collection Making

m.c. schraefel¹, David Modjeska², Daniel Wigdor¹, Yuxiang Zhu¹

¹Department of Computer Science | ²Faculty of Information Studies

University of Toronto

Toronto, Canada

[mc | yuxiang | dwigdor]@dgp.toronto.edu; modjeska@fis.toronto.edu

ABSTRACT

Hunter Gatherer is a tool that lets Web users carry out three main tasks: (1) collect components from within Web pages; (2) represent those components in a collection; and (3) edit those collections. We report on the design and evaluation of the tool, and contextualize tool use in terms of our research goals to investigate possible shifts in information interaction practices resulting from tool use.

KEYWORDS

Interaction design, information gathering, information management, information sharing, attention, collections

INTRODUCTION

Studies of Web-based information interaction (such as [2], [4]) have generally dealt with a page as the smallest unit of interest. Our task analysis indicates: (1) that Web users want to collect information items found within Web pages for a variety of purposes, such personal research, ongoing reference, or resource sharing, but that (2) users only infrequently make such collections, in large part because of poor interaction support for this activity. For instance, bookmarks referencing entire pages often capture more than the desired data; this forces users first to load and then to sift through multiple pages to find desired material. Text editors cause users to shift attention between two tasks - information gathering in the browser and information management in the editor. With editors, users often neglect to label a collected component with a title or URL (from the source page); this makes later access to the original material difficult, reducing the collection's value over time. We developed Hunter Gatherer (HG) for two reasons: first, to support such within-page collection making, and second, through observing tool use, to investigate how user-determined sub-page component capture might affect information interaction practices.

HG (Fig. 1) blends the transparency of bookmark capture for selecting components, with editor support for revising collections. For each component, HG also automates the inclusion of a contextual, editable header and grabs the URL of the source page, so that users can return to the

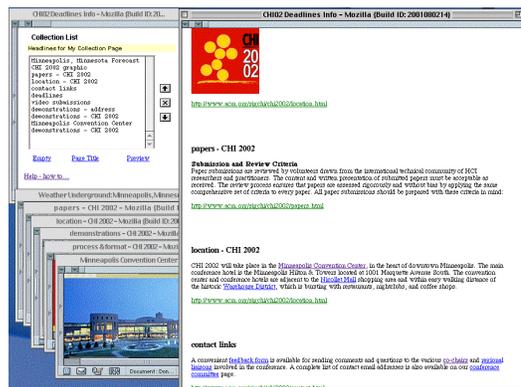


Figure 1. Hunter Gatherer. The Collection Window (foreground) presents all gathered elements, while a List View appears (upper left) for monitoring the Collection as it is being built; lower left are the components' source pages.

source document at any time. Our goal for Hunter Gatherer's interaction design is to let users easily determine the information activity that they wish to focus on: gathering, management or contemplation of the collection.

Hunter Gatherer is the result of iterative user-based design, surveys and evaluation. This paper briefly summarizes related work in Web-based collection management, then describes HG's interaction design, and presents the most recent version of the artifact and its lab evaluation. We conclude with next steps for using the tool to address our larger research questions on possible shifts in interaction practices.

RELATED WORK

When we focus on information gathering on the Web, we highlight the process of "information triage" [2]. While there has been much work done on the management of Web-based document collections, there has been less work on the interaction activity of first capturing the triaged information and then placing it into a collection. Accordingly, our work has focused on the latter process.

With respect to similar, existing research and commercial tools, there are several of particular interest: WebBooks [1], Data Mountain [3], and TopicShop [4]. Each of these tools focuses on the Web page as the smallest unit of capture. Even with browser-integrated editors, such as Front Page, URL information and headers are not captured; file management is generally local. Commercial tools such as WebForia, which do capture components, are platform specific applications requiring management separate from the

browser. Web Squirrel, inspired by Spatial Hypertext, supports the idea of components within pages, but captures neither URLs nor non-textual information. None of this work specifically considers the interaction of moving content from one context (browser) to another (collection).

HUNTER GATHERER DESIGN PROCESS

Based on our task analysis, we determined 3 requirements for Hunter Gatherer. (1) To make the addition of components to collections as transparent as highlighting text. (2) To support user-driven focus shift among component selection, addition, monitoring, and management. (3) To give collected components enough contextual information to provide immediate value to the user.

Description of the Tool and Interaction

HG is browser-based, server-side tool requiring no additional software on the user's part. This integration with the browser minimizes divided attention forced by shifting between browser and editor for gathering and management. HG lets users select parts of Web pages, and by pressing a keyboard command, add the selected component to a collection. The Collection is another Web page. (Fig 1, foreground). Each element in the Collection appears with the URL to its source page, as well as a default, editable header, based on the source page title and keywords in the selection. The user can peripherally monitor the collection as it is built: a background window (Fig. 1, upper left) shows the automatically generated name of each component as it is added to the collection. The list can also be used to sort, delete or rename components (the user can also perform these revisions directly in the Collection page). HG does not copy data into a collection; it creates references for the components instead. The URL for a Collection page contains references to each of its components. Representing Collections in a URL makes them easy to share, e.g., sending the URL via email. Throughout the collection process, the user rather than tool interaction determines task focus: hunt for information, select and add components from those sources, or edit the contents of a collection.

EVALUATION

In order to assess how Hunter Gatherer meets requirements for component collecting, focus shifting and immediate value, we ran a user study to assess the tool's efficiency.

Design and Methodology

We performed a 2x2, within-subjects study to compare HG with a text editor for the task of creating collections. To reduce learning effects, we choose Microsoft Word as a tool familiar to participants. The first factor in the experiment was tool (Hunter Gatherer vs. Word); the second factor was data set (Web pages in a university's Chemistry vs. Physics departments). The study had 12 participants, representing a mix of technical and non-technical undergraduate and graduate students.

To start each session, users received 15 minutes of training with HG. Users were then asked to build two collections, each from a given bookmark set, to be clear enough for use by another person. Tool exposure order was counterbal-

anced. To reduce potential learning effects, we prepared two similar sets of bookmarks, one each on the Chemistry and Physics programs at University of Toronto. The pages in each set were taken from a single Web site. Participants had 5 minutes with each set of 3 bookmarks to familiarize themselves with page contents. Participants then had 15 minutes to build a collection from these bookmarks to (1) explain degree requirements for a subject minor, (2) list and describe required courses, and (3) show current instructors.

Results and Discussion

A one-way within-subjects ANOVA showed a significant effect of tool type (collection time ($F = 5.730$, $p < .040$) in comparing average component collection time using HG and Word. Participants required an average of 6.7 seconds using HG and an average of 10.9 seconds using Word.

In the post-evaluation questionnaire participants commented that sorting components in collections was easier in HG than in Word. Similarly, 10 of 12 users mentioned the automatic capture of a component's URL as HG's strongest feature.

We have met the first design requirement - to make the addition of a component as efficient as copying text from a browser. Hunter Gatherer's selection performance is also significantly better than Word's, and the HG method is also more effective than Word for component addition.

CONCLUSIONS AND FUTURE WORK

Empirical and anecdotal data indicate that Hunter Gatherer is a promising tool for improving user interaction with Web resources. After tool enhancements based on the above results, HG will also serve as a platform for a follow-up, longitudinal study to consider our main research questions: (1) how facilitating capture of user-determined components may change the way users think about information (e.g., becoming active co-creators), and (2) what the consequences of such a shift may be. As a browser-integrated Web-based tool, HG supports wide deployment for considering these research questions in a significant population.

REFERENCES

1. Card, S K., Robertson, G. G., York, W. The WebBook and the Web Forager: an Information Workspace for the WWW. Conf. Proc. on Human Factors in Computing Systems (Van., Canada, April 13-18, 1996), 111.
2. Marshall, C. C. Shipman, F.M., Coombs, J.H. VIKI: Spatial Hypertext Supporting Emergent Structure. Proceedings of the 1994 ACM European Conf. on Hypermedia Technology, 1994, 13-23.
3. Robertson, G., Czerwinski, M., Larson, K. Robbins, D., Thiel, D., van Dantzych, M. Data Mountain: Using Spatial Memory for Document Management. Proc. Of UIST '98, 11th Annual Symp. on UIST, 153-162.
4. Terveen, L., Hill, W., Amento, B. Constructing, Organizing, and Visualizing Collections of Topically Related Web Resources. ACM Trans. on Computer-Human Interact. 6, 1 (Mar. 1999), 67-94.