

# How Do People Get Back to Information on the Web? How Can They Do It Better?

**William Jones & Harry Bruce**

The Information School, University of Washington, Seattle, Washington 98195

[jones@ischool.washington.edu](mailto:jones@ischool.washington.edu)

**Susan Dumais**

Microsoft Research, Microsoft Corporation, Redmond, Washington 98052

**Abstract:** In one study, people are observed to keep web information for later use through many different methods including the use of Bookmarks (or Favorites), self-addressed email, hand-written notes, and paper print-outs. Each keeping method realizes its own constellation of important features or functions. No observed method provides all desired functions. A simple "Add Favorite 2" prototype is developed to support a "super" keeping method designed to realize more functions (reminding, context, multiple points of access) in a single effort. Results were mixed. A second study indicates that people may be increasingly returning to desired web sites using methods that require no explicit keeping behaviour. Especially popular are the use of: 1.) A search service. 2.) Partial completion of a site's web address (and acceptance of a suggested completion to this address). 3.) The Hyperlinks from another web site. Implications for tool support are explored.

**Keywords:** Web navigation, personal information management, information retrieval

## 1. Introduction

How is new information integrated into a person's existing information environment? And how can a person remember to use information later when a need for this information arises? We call this the challenge of **Keeping Found Things Found** or **KFTF**.

We are in the second year of a three-year research project, funded by the National Science Foundation, to understand better how people meet the KFTF challenge on the Web (Jones, Dumais, Bruce, 2002). How do people organize web information for re-access and re-use? How can tools help? Please visit the project web site, <http://www.ischool.washington.edu/~kftf/>, for more information).

The project consists of three overlapping phases: Phase I. Observations and interviews to assess the current situation. Phase II. Modeling and analysis to understand and explain Phase I data and to suggest directions for prototyping. Phase III. Prototyping

and evaluation of promising directions in tool support.

All observational studies take place in the participant's workplace and are video-taped.

## 2. A Study of Keeping Methods

Study 1 involves a one-hour session. Participants first think aloud for 30 minutes as they complete a work-related, web intensive task of their choosing. Participants are then interviewed to explore the reasons for observed actions and the causes for observed problems. Special attention is given to participant efforts to keep and re-find web information. 24 participants (6 researchers, 9 information specialists and 9 managers) have completed Study 1.

Participants use a **large number of keeping methods**. Each of the following methods was used by at least one participant during the **task-based observation** with the intent to keep web information for re-use: 1.) *Send email to self*, with URL referencing web page. 2.) *Send email to others* that

contains a web page reference – with the intent to search the Sent Mail folder or contact recipients later to re-access the web information. 3.) *Print out the web page.* 4.) *Save the web page as a file.* 5.) *Paste into a document the URL for a web page.* 6.) *Add a hyperlink into a personal web site.* 7.) *Bookmark the web page.* 8.) *Write down the notes on paper containing the URL and actions to be taken.* 9.) *Copy to a “links” toolbar so that the web address is always in view in the browser and can be quickly accessed.* 10.) *Create a “note” in Outlook that contains the URL and can be associated with a date.* More generally, create an item in a task management or calendaring system that can then be associated with one or more dates.

Participants sometimes found a web site, commented that it looked useful, but then did nothing overtly to keep it. When asked during the interview, participants pointed to one of the follow “do nothing” methods: 11.) Do nothing and count on *entering the URL from memory* or type in the first part of the address and then accept one of the browser’s suggested completions 12.) Do nothing, and *search for* (find again) the desired web information later on. 13.) Do nothing and *reach the web information later from a known point of access such as a web portal.*

Participants also indicated during the interview that they sometimes used more than one keeping method if they wanted to insure that web information would “stay found”. **Many methods have the effect of transforming information into a different form** – into paper, into an email message, into an e-document, for example. The use of several of these methods distributes web information across organizational schemes introducing a kind of “fail-safe” redundancy.

Other patterns were noted from Study 1: 1.) **Tested and untested references are often “kept” in different ways.** Several participants distinguished between references to new, interesting but as yet untested, unevaluated web information and references to web pages or sites of established value. Different participants used different keeping methods or sometimes the same method in different ways to distinguish “tested” and “untested” web references. For example, one participant placed untested references at the end of her list of Favorites. She reviewed this list of Favorites periodically, either deleting a Favorite or placing higher up on the list, at the top-level or embedded within a folder. 2.) **Some participants went to great lengths to maintain a single organizational scheme.** Examples of this were discussed in the introduction

to this proposal. 3.) **KFTF practices appear to vary according to job position and a participant’s general relationship to web information.** For example, managerial participants in our study indicated that they rarely use the Web directly to conduct research. Instead, they receive their information “second-hand” through communications with others. Email is now used extensively in this communication. Information specialists appeared to make especially conscientious use of bookmark folders as a way to organize web information.

A survey is underway to get a better sense for the relative popularity of methods. Based on the small sample collected so far (21 respondents), indications are that most methods are widely used. All but two of the methods listed above (add to personal web site; write down URL on paper) are used at least once a week by 20% or more of the respondents. The most popular methods are the three “do nothing” methods – each is used by more than 70% of the respondents at least once a week as part of a strategy of web re-use. The next most popular method is bookmarking (used by 57% of respondents at least once a week). Respondents use an average of just over five methods at least once per week.

### 3. A Functional Analysis

A functional analysis helps to explain the overall diversity of keeping methods observed. People differ in the functions they require according to their workplace role and the tasks they must perform. Among the functions that participants valued, based upon comments made during the observations and interviews, are the following: 1.) *Portability of information.* Can participants take the information with them wherever they go? Paper is especially portable. 2.) *Number of access points* – related to but not the same as portability. Could participants access the information from multiple places? From their laptop as well as their desktop computer? From home as well as work? Participants could not access the same bookmarks from several places; participants could access email from several places. 3.) *Context.* Did participants know why they saved a web address? Participants were able, through comments and subject line, to establish a context for a web address sent in an email message. The same was true for web addresses pasted into a document. 4.) *Reminding.* Saving information does little good if we don’t remember to use this information later on and in the right situation. Participants gave high marks for the reminding value of paper printouts and

self-addressed email. Participants gave low marks for bookmarks. Participants commented that they often forgot to use a bookmark until too late.

|                       | Number of access points | Context | Reminding | ... |
|-----------------------|-------------------------|---------|-----------|-----|
| Email to self         | High                    | High    | High      |     |
| Email to others       | High                    | High    | Low       |     |
| Print-out             | High                    | Low     | High      |     |
| Save as file          | Low?                    | Low     | Low       |     |
| Paste URL in document | Low?                    | High    | High?     |     |
| Bookmark              | Low                     | Low     | Low       |     |
| Links Toolbar         | Low                     | Low     | High?     |     |
| ...                   |                         |         |           |     |

**Table 1:** A (partial) functional comparison of keeping methods..

A mapping from functions to keeping methods is partially displayed in Table 1. Although some ratings are preliminary and subject to change, two things are apparent: 1.) **Methods vary widely in the functions they provide.** 2.) **No current method provides all functions.**

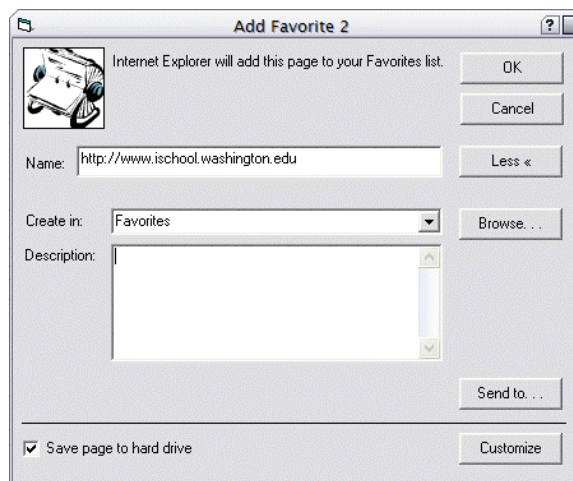
#### 4. A Simple Prototype

A diversity of keeping methods are observed in Study 1. A functional analysis helps to explain why. No current method provides all desired functions. The choice of a given method may then depend upon which functions are deemed most important under the circumstances. If portability of the information is especially important, for example, a person may choose to print out. If it is especially important to be reminded of the information later on, the person may choose to send a self-addressed email.

But suppose it were possible to choose a “super method” that realized all or most of the valued functions? By the logic of the functional analysis, people should eventually come to prefer this new method.

In this spirit, two of students in the Information School, Amanda Whitley and Khoi Duong, worked

on an enhanced “Add Favorite 2” dialog as illustrated in Figure 1.



**Figure 1:** An alternate Add Favorite dialog with added functions.

This new dialog provides all the features of the standard Add Favorite dialog (as provided in the Microsoft Internet Explorer) and provides the following additional features:

- The option to add one or more sentences in a “Description” field. This description then appears as a tool tip when the mouse pointer hovers over the Favorite later on. The description can be used to help establish a **context** for the Favorite.
- The option to email the web address to oneself (or to others) as a **reminder** using the “Send To...” button. The web address and description are automatically placed in the body of the email message. Web information placed in an email message also typically has a **high number of access points**.
- The option to place a newly created Favorite in an exiting folder hierarchy (already established to organize electronic documents).

This prototype is complete and evaluations are underway. However, preliminary results are not encouraging. The prototype works fine. But people in our sample of volunteer users have already fallen out of the habit of using the existing Add Favorite dialog. It is that much harder to establish a new habit of using the prototype Add Favorite dialog. And, admittedly, the advantages of the new dialog are modest and incremental.

## 5. A Study of Re-Finding

**Study 2** involves two one-hour sessions to set up a delayed cued-recall test of a participant's ability to re-access valued web sites. In session 1, participants see a sampling of web sites they have visited recently. For a given web site, participants rate the likelihood that they will have a need to visit this site again over the next 12 months. For sites for which the likelihood 80% or higher, participants then provide a brief description of what they were last doing at the site and what they generally are doing at the site.

These descriptions are then used to generate cues which participants see 3 to 6 months later in session 2. For a given cue, participants are first asked if they recall the web site cued. Participants then attempt to return to the cued web site as quickly as possible by whatever means they chose. Participants are each tested for 14 web sites that are selected to realize a range of access frequencies (from daily access to only once or twice per year). Home pages were excluded from consideration. Nine of a planned 13 participants have completed study 2. The remaining four participants have completed session 1 and will complete session 2 by the end of 2002.

In results so far, participants have had very good success at returning to web site. On only one trial (of 126) did a participant fail outright. (On eight trials, participants did not recognize the web site from the cue; three additional failures can be attributed to problems with the network.) On 93% of the trials, participants successfully returned using the first method they selected. Of these successful "first-try methods" the most commonly observed, in order, were 1.) Direct entry of URL. 2.) Favorites or Bookmarks. 3.) Search using a search service. 4.) Access via another web site. These methods account, respectively, for 42%, 18%, 18% and 16% of the successful first-try methods. (These four methods are also the most popular and most frequently used based upon survey results so far.)

Note that three of these methods, accounting in aggregate for 76% of the successful first-tries, require no explicit keeping method. Some participants indicated that they were increasingly relying on these "do nothing" methods – especially the use of direct entry and searching – to return to web information. To what extent will further improvements in support for these and other do-nothing methods further reduce the need to take

explicit steps in order to keep web information for re-use?

Consistent with the overall success rate for first-try methods, participants took an average of just under 1 minute to return to a web site. But variation appears to be large (though, as yet, unanalyzed). Some trials took as long as five minutes to complete and involved a sequence of methods for re-accessing a web site. For example, one information specialist used the following sequence of methods before finally re-accessing the site: 1.) try to reach from another web site, 2.) use search service, 3.) try direct entry of URL, 4.) look in paper file folder (success). At just over 6%, the percentage of times participants were observed to use three or more methods to access a web site seems small. Nevertheless, participants sometimes expressed considerable frustration at not knowing where to "go" in order to re-access a web site. It is possible that attempts involving multiple methods have a significance and perceived cost beyond their actual numbers or frequency of occurrence.

## 6. Conclusions

- People use a diversity of methods to organize web information for re-access and re-use.
- A functional analysis can help us to understand the diversity of methods observed and their relative popularity.
- Methods differ in the functions they provide.
- No single current method provides all the functions a user may need.
- A "natural" study of people doing tasks in their workplace can be very useful.
- Participants appear to be effective at returning to a site. Success rate is high and the first method chosen usually works.
- Direct entry, search and "hyperlink to" account for 83% of successful first-try methods for re-access. These methods require no keeping activity up front.

## References

- Jones, W., Dumais, S. & Bruce, H...(2002), Once Found, What Then? A Study of "Keeping" Behaviors in the Personal Use of Web Information. *Proceedings of the ASIST Annual Meeting*, 39, (2002), 391-402.