
Information behaviour that keeps found things found

[Harry Bruce](#), [William Jones](#)

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

and [Susan Dumais](#)

*Microsoft Research, Microsoft Corporation
Redmond, Washington 98052, USA*

Abstract

This paper reports on a study that the researchers call: 'Keeping found things found on the Web' or 'KFTF'. The research focuses on the classic problem of ensuring that once a useful information source or channel has been located, it can be found again when it is needed. To achieve this goal, individuals engage in information behaviour that the research team refers to as keeping and re-finding. The research study observed both types of information behaviour. To study keeping, the researchers designed an observational study to record what people do in their offices when they are searching or browsing the Web and they find information they want to keep for re-use. This behaviour was observed and then analysed for its underlying purpose in the first phase of the KFTF study (the keeping study). In the second phase of the study (the re-finding study), the researchers designed a delayed recall observation which required participants to re-find information on the Web that they had located during the observations of phase 1. This delayed recall study focused upon observations of re-finding information. Finally, the researchers conducted a survey to validate and augment the data from the keeping study. Two hundred and fourteen individuals participated in the survey.

Introduction

In this complex information world of *the information age* the fundamental importance of, and the role of information in our daily lives is widely acknowledged. People regularly locate, encounter and have access to far more information than they can possibly process. It is common to speak of a *glut of information* and of *information overload*. Information plays a significant role in our daily professional and personal lives and we are constantly challenged to take charge of the information that we need for work, fun and everyday decisions and tasks. In response to this challenge, individuals create a personalized subset of the *information world* that they can use when they are faced with information needs. This subset of the information world is a personal information collection. A personal information collection is defined as the space we turn to first when we need information to do a task or pursue an interest. It is a collection of

information sources and channels that we as individuals have acquired, cultivated, and organized over time and in response to a range of stimuli. It is an organic and dynamic personal construct that we take with us into, and out of, the various information events that frame our daily working and personal lives. Personal information collections include content in various forms (documents, Web pages, mail, notes, calendars, address books, etc.), structures for representing and organizing this information (folder hierarchies, piles, lists, etc.), and pointers to information (people, links, Favourites, etc).

Keeping (leaving) and re-finding information

The ideal of building a personal information collection is that once we find useful information, it will be organized so that it is readily at hand when a later need for it arises. Individuals approach the ideal of a perfect personal information collection through their ability to distinguish useful from non-useful information. The assumption of the personal information collection is that an individual will keep useful information and discard information that is not useful.

The personal information collection, therefore, includes information sources and channels that an individual has kept. Keeping acts are interventions by the individual. He or she identifies information that is useful and then engages in an act that relates to collecting, representing, indexing, cataloguing, classifying, and storing the information. The individual may not use these (information) terms to describe his or her behaviour but the goals of these interventions remain the same—to keep the information on-hand, available and accessible for a time when it will next be needed. It is very important to note that the personal information collection also includes information sources and channels that have been *left*. In this case, the individual recognizes the information is useful (and that it will be useful in the future) but s/he also concedes that it is in a space or place that can easily be located again. The individual makes a conscious decision to leave the information *in situ*. The individual makes this decision assuming that he or she will be able to find this information again in the place where it was originally located or encountered. This is *leaving* behaviour. A regularly used Website is one example of an information source that an individual might choose to *leave* (because he or she can find it later) rather than *keep*. Information that is *kept* and information that is *left* belong equally to the conceptual space that we refer to as the personal information collection. This choice can be seen as an instance of a signal detection task ([Jones, 2004](#)), that is, whenever an individual locates or encounters an information source or channel, s/he faces an essential choice, whether consciously exercised or not—should steps be taken to keep the information for later access and use? Should this information be included in my personal information collection? The choices made in the act of keeping the information and the reasons for doing so may vary greatly but the essential choices remain the same: to keep the information, to leave the information where it is (it is useful but I can find it when I need it) or to ignore the information (it is not useful and it is not likely to be useful at any time in the future). The goal is that once we have found, encountered or acquired useful information, it should never be lost or forgotten and it is readily at hand when a later need for it arises. We refer to the challenge inherent in this goal as *keeping found things found*. Whether the useful information is kept or left, the overall goal is to increase the likelihood that, whenever the information is needed, the individual will remember where it is and be able to re-find it. The ultimate test of keeping and leaving, therefore, is observed in the success or failure of re-finding.

Previous research

Building, managing and using a personal information collection is behaviour that relates to personal information management. There has been a number of studies that help us to understand the practices of managing the information that helps us in our daily lives ([Bellotti & Smith, 2000](#)). These studies provide insight into how and why individuals choose to keep and re-find print based information by organizing this information into piles and files ([Cole, 1982](#); [Lansdale, 1988](#); [Malone, 1983](#); [Mandler, et al., 1992](#)). Other studies identify the information types that individuals tend to organize into piles and files ([Barreau, 1995](#)); [Kwasnik, 1991](#); [Sellen & Harper, 2002](#)) and the affect that the number of files has on an individual's abilities to remember the files names and locations ([Jones & Dumais, 1986](#)).

Research into the function of piles and files has also focused on electronic documents. Most computer users have support for the creation of different folder hierarchies into which electronic information can be filed. But if the desktop is often obscured by various open windows, research by Kaptelinin ([1996](#)), for example, observed that the accessibility of each item is reduced. Researchers have also studied the organization of piles and files in relation to e-mail use Whittaker & Sidner ([1996](#)) observing that the filing decisions (for e-mail)– which folders to create, what to name them, how to organize them, etc., –are fundamentally difficult regardless of the item being filed. Filing takes time and the folders that are created today may prove to be ineffective or even an impediment to access to the information in the future. In addition, subjects in this study reported an '*out of sight, out of mind*' problem that items placed in a folder were sometimes forgotten until well after the period of their usefulness had passed. Like their paper-based counterparts, folders of electronic documents such as e-mail messages may contain too few or too many items to be effective or they may be simply obsolete ([Tauscher & Greenberg, 1997](#)).The e-mail inbox provides pile-like functions of accessibility and visibility but these functions are clearly reduced as the number of items in the inbox increases and especially for older messages that scroll out of view.

There is also an extensive body of research from which we can speculate upon the role of a personal information collection in the information seeking and use of particular professional groups. Generally, these information behaviour studies tend to identify a characteristic set of information needs for the group being researched, and then observe the way that members of this group locate the information that is needed for specific aspects of their work or the practice of their profession (there are many examples such as [Kulthau and Tama, 2001](#); [Hertzum and Pejtersen, 2000](#); [Owen, 2003](#); [Buchanan, 1993](#); [Watson-Boone, 1994](#); [Sutton, 1994](#); [Ellis, 1997](#); [Leckie, et al., 1996](#); [Meho and Haas, 2001](#); [Ashill and Jobber, 2002](#); [Smith and Culkin, 2001](#)). Recent reviews of this body of research can be found in Pettigrew, *et al.*, ([2001](#)) and, with specific reference to use of the Internet and the World Wide Web in Bruce ([2002](#)). It is impossible in the scope of this paper to adequately summarize this body of work but the findings of this type of research tend to include the reported preference by members of these professional groups to seek information internally (from a personal information collection – which includes their contacts with colleagues and experts in their field). These professionals also tend to follow habitual patterns of information seeking, and are primarily concerned with the time and cost involved in finding the right information.

Keeping found things found (KFTF)

Thus, a lot of data has been assembled from previous research. But, what tends to be lacking is research that examines how individuals manage information they find or

encounter with the purpose of using that information later. As stated earlier, the authors of this paper see the goal of this information behaviour expressed in the concept of a personal information collection. The challenge for the individual is—Keeping Found Things Found. The aim of this study was to understand better how people manage information for subsequent re-access and re-use (keeping-leaving and re-finding). The study focused on this behaviour in relation to information found or encountered on the World Wide Web. The KFTF study included a keeping study, a re-finding study and a survey.

The keeping study

The aim of the keeping study was to understand the diversity of keeping methods (including leaving methods) that people use in their efforts to manage Web information for re-use. In this study, the researchers also analysed the underlying function of the observed keeping methods. The researchers wanted to understand the underlying reasons for the diversity of keeping and leaving methods. How do methods compare, for example, with respect to the functions they provide and, are there important variations in the use of various methods?

Participants

Participants were drawn from three professions: researchers, information professionals (including librarians) and managers. It was assumed that members of each of these professions depend heavily on the timely availability of information for the performance of their jobs. However, the traditional orientation to information and information gathering differs between the three groups. A total of twenty-four people participated in the keeping study: six researchers, nine information specialists and nine managers.

Method

Information professionals and researchers were observed and interviewed in their workplace in a session lasting approximately an hour. A few days before the observation session, participants were asked to complete an e-mail questionnaire designed to establish background information concerning education level, job responsibilities, experience with computers and the Web, etc. In this questionnaire, each participant was also asked to list at least three work-related, Web-intensive, *free-time* tasks that they might like to work on over the next week should they have a half-hour or more of unscheduled time. During the subsequent observational session, one of these tasks was selected, by agreement between the observer and the participant. The participant then spent the next thirty minutes working on this task. Participants were instructed to think aloud while performing the task. An *over-the-shoulder* video recording was made to capture screen content (at very coarse resolution), the participant's hand movements and the participant's think-aloud commentary. Participants were asked to handle office interruptions (phone calls, visitors, etc.) as they normally would. They were also encouraged to do what they would normally do in the face of serendipitous discoveries (e.g., Web pages of relevance to other aspects of their lives such as upcoming vacations, purchases, health insurance, child care, etc.). The observer did not speak except to answer questions of procedure or, as needed, to remind the participant to continue to think aloud.

Immediately following the completion of the task, the observer reviewed the participant's

actions and asked the participant to fill in the gaps or to explain certain actions. Participants were asked to discuss other keeping-leaving methods they might use in other situations of Web use. Participants were also asked to discuss the strengths and weaknesses, the estimated costs (e.g., execution time) and benefits (e.g., success rate) of various methods.

The research team found that this procedure would not work well for managers who had agreed to participate in the KFTF study. Several interested, willing managers had agreed to participate but scheduling an hour of their time proved difficult. More importantly, each of our managerial participants indicated that they seldom, if ever, spent a half-hour in a predominantly Web-based task. Participants indicated they might read e-mail, return phone calls or work on a report. But a half-hour Web task would be somewhat contrived. Consequently, for managerial participants, the research team conducted brief (usually less than half an hour) interviews, either over the telephone or in person.

Results

A large number of keeping methods was observed. Each of the following methods was used by at least one of the participants of this study:

1. send an e-mail to oneself, with a URL referencing a Web page;
2. send to others an e-mail that contains a Web page reference (and then search the Sent Mail folder or contact recipients to re-access the Web information). An instance of sending e-mail containing a URL to someone else is not necessarily an instance of this method. For example, all information specialists indicated that they sent mail to others as part of their job (as information mediators). However, no one in this group indicated that sending mail to others was part of a strategy to keep found things found. On the other hand, we do count the e-mail containing URLs that managers indicated they send to subordinates. Managers send this mail with the expectation of getting back information, albeit in processed, reduced form, later on;
3. print out the Web page;
4. save the Web page as a file;
5. paste the URL for a Web page into a document;
6. add a hyperlink to a personal Website;
7. bookmark the Web page;
8. write down 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; and
10. create a note in Outlook, which 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.

These ten keeping methods imply one or more companion method for re-finding. Keeping methods and methods for re-finding combine to form a strategy for keeping found things found. For example, creation of a bookmark has a companion re-finding method of using the bookmark facility later to locate the bookmark and, through the bookmark, re-finding the underlying Web page. Sending a Web reference to oneself by e-mail has a companion method of periodically checking the inbox. Sending a Web reference in e-mail to someone else may have the companion method of searching through the Sent Mail folder if the need for the Web reference arises later.

As stated earlier, when individuals locate or encounter useful information they may decide to keep or leave the information as a strategy associated with its future re-access and re-use. Leaving behaviour is, therefore, associated with keeping found things found. In the keeping study, participants revealed a number of *doing nothing* or leaving

methods. In each case, participants in the study indicated during the interview that they expected to be able to return to the Web information that they located but they would do this later. The leaving methods observed were:

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; or
13. do nothing and reach the Web information later from a known point of access such as a Web portal.

During the de-briefing interview that followed each observation, a number of participants indicated that they sometimes used more than one keeping method if they wanted to insure that Web information would *stay found*. A number of the methods observed, in fact, have the effect of transforming the Web information into a different format: into paper, into an e-mail message, or into an e-document, for example. This observation led the research team to engage in an analysis of function in an attempt to understand the reasoning behind an individual's choice of the various methods observed. Data from the observations and interviews indicate that several functions appear to influence the choice of method:

1. **portability of information:** users can take the information with them wherever they go;
2. **number of access points:** the user can access the information from multiple places;
3. **persistence of information:** the Web page, with the same URL and content will remain the same into the future;
4. **preservation of information in its current state:** the interactive, design and layout qualities of the Website will be preserved;
5. **currency of information:** the information can be refreshed to reflect the most current updates to the content;
6. **context:** the user is able to establish a context for why a Website was kept in the form of reminders about its use and likely future use;
7. **reminding:** the user is reminded about the information by the keeping method;
8. **ease of integration:** the method helps the user to integrate new information or new references with ongoing projects and existing organizational schemes;
9. **communication and information sharing:** the method makes it easier to share information with others; and
10. **ease of maintenance:** the user will need to maintain and update his or her personal information collection; does the keeping method support or facilitate maintenance processes?

For each of the functions listed above, the research team attempted to give a simple high, medium or low rating for selected keeping methods. These ratings are summarized in Table 1.

	Portability	No. of access points	Persistence	Preservation	Currency	Context	Reminding	Ease of integration	Speed of re-find	Ease of maintenance
e-mail to self	low	high	low	med	high	high	high	med	low to med	med
e-mail to others	low	high	low	med	high	high	low	low?	low	high
Print-out	high	high	high	low	low	low	high	med	low	med
save as file	med?	low?	high	high	low	low	low	med?	low	med
paste URL in document	low	low?	low	med	high	high	high?	high?	low	high
personal Website	low	high	low	med	high	high	high?	high	med	high?
search	low	high	low	med	high	low	low	?	med	high
direct entry	low	high	low	med	high	low	low	?	high	high

bookmarks	low	low	low	med	high	low	low	low	high	low
history	low	low	low	med	high	low	low	low?	low	?
links on toolbar	low	low	low	med	high	low	high	low	high	med

Table 1: A functional comparison of different methods of keeping Web information for re-use

Table 1 should be regarded primarily as an illustration of the functional approach. It is not intended to be complete, either with respect to keeping methods or the functions which might influence an individual to use each method. Reading through the data in the table, we see that paper, for example, is rated high for portability. Two participants in the study, in fact, indicated that printouts made it possible to work with the Web information during *dead times* (for example, while commuting on a bus or waiting for a meeting to begin). On the other hand, a printout cannot preserve the interactivity of a Website (hence a low rating on preservation), nor can the printout be updated readily (meaning a low rating on the currency). E-mailing a Web address (or the Web page itself) is clearly an effective way to share information and for several participants in the study, printouts of Web information were also an effective way to share information (hence, both methods score high on communication and information sharing). On the other hand, bookmarks, as currently implemented in both Internet Explorer and Netscape Navigator, cannot be readily shared (this method scores low on communication and information sharing). The reliability of these ratings may be somewhat contentious but the functional analysis is useful for the kinds of questions it encourages us to ask. For example; Does a keeping method have a reminding function? Does it enable us to provide a context for the Web information? Does it integrate well with our existing organizational schemes? Does it allow us to access the Web information we wish to keep from home and on the road as well as at work? The functional analysis also indicates that the methods observed in the keeping-leaving study vary widely in the functions they provide. It also reveals that none of the methods observed manages to provide all of the desired functions.

The re-finding study

The overall aim of keeping or leaving information is that, later, the individual will be able to access and use the information if it is needed. The test of keeping and leaving methods, in other words, is in the ability of the individual to re-find the information. In the re-finding study, the KFTF research team set about observing the various methods that participants used for re-finding information that they had located or encountered on the World Wide Web.

Participants

The information specialists, researchers and managers who participated in the keeping study were asked if they would participate in the re-finding study. Twelve members of the original sample agreed to participate in this follow-on study.

Method

The re-finding study involved two one-hour sessions to set up, and then conduct a delayed cued-recall test of the participant's ability to re-access Websites. In session one, each participant was presented with a sampling of Websites that he or she had visited

recently. This listing was obtained from the participant's History List. For each Website, the participant was asked to rate the likelihood that he or she would have the need to visit this site again over the next twelve months. For sites where the likelihood was rated as 75% or higher, participants were then asked to provide a brief description of what they were last doing at the site and what they generally do at the site. These descriptions were used to generate cues which were presented to the participants three to six months later in session two. For a given cue, participants were first asked if they recalled the Website. They were then asked to return to the cued Website as quickly as possible by whatever means they chose. Each participant was tested for fourteen Websites. These Websites were selected according to a range of access frequencies (from daily access to only once or twice a year). Home pages were excluded from the re-call test.

Results

The data in Table 2 report the rates of success for the re-finding trials observed. A re-finding was rated as an objective success if the participant got back to the Website (the actual URL) as listed in the delayed recall test. It was rated as a subjective success if the participant felt that he or she had been able to return to the Website. Objective success was rated by the observer and subjective success by the participant. The data in Table 2 indicate that all participants rated all their re-finding trials as a subjective success. This was true for all Websites irrespective of whether the participant frequently or infrequently accessed the site (access frequencies – low, medium or high) according to the data collected in session one.

Access frequency	Objective success	Subjective success	Trial time	No. of methods tried
low	90%	100%	1:42	1.2
medium	98%	100%	1:07	1.2
high	100%	100%	0:47	1.0

Frequencies of Web access in a typical week range from low (sites rarely or never visited in a typical week), medium (sites visited one to three times in a typical week), or high (once or more per day).

Table 2: Re-finding performance when cue is effective at reinstating the Website and the task across frequencies of Web access in a typical week

In 93% of the trials, participants successfully returned to the site using the first method they selected. Of these successful first-try methods the most commonly observed methods in order, were:

1. direct entry of URL (42%)
2. Favourites or Bookmarks (18%)
3. search using a search service (18%), and
4. access by another Website (16%).

It is important to note that three of these methods, accounting for 76% of the successful first-attempts at re-finding the Website in the trial, require no explicit keeping method. In the keeping study, some participants had indicated that they were increasingly relying on *doing nothing* or *leaving* methods for getting back to useful Web information.

Consistent with the overall success rate for first-try methods, participants took an average of just under one minute to return to a Website. There was great variation

however: some trials took as long as five minutes to complete and involved a sequence of methods for re-accessing the Website. One information specialist, for example, used the following sequence of methods:

1. try to re-find from another Website;
2. use search service;
3. try direct entry of the URL; and
4. look in paper file folder,

before finally re-accessing the site. It must be noted, however, that participants used three or more methods to re-find the cued Website in only 6% of the recall trials.

Survey

The keeping and re-finding studies of the KFTF project were supplemented by a survey. The survey aimed to collect data on the behaviour that was observed in the field studies; to determine, from a larger sample of participants, if the list of observed keeping methods was complete and how frequently people use each method. The survey was also aimed at finding out why certain keeping methods are used, to supplement or validate the functional analysis conducted as part of the keeping study. The survey was administered online from [the KFTF Website](#) for a period of one month (January 2004).

Sample

During this period 214 individuals responded to the survey. A majority of respondents were within the age range of forty to fifty-nine years (60.75%) and female (65.89%). The most common occupations represented in the sample were librarian (49.53%); researcher (19.63%); manager (9.81%) and student (4.67%). Other professions included editor, systems analyst, office manager, programmer, Web developer and account manager.

Keeping methods

The survey instrument asked respondents to indicate whether they had ever used each of the thirteen methods that were identified in the keeping study (see Figure 1). Most of the respondents had used many of these keeping methods. The data appearing in Figure 1 ranks the methods in order from those least used to the methods most commonly used by individuals in the sample.

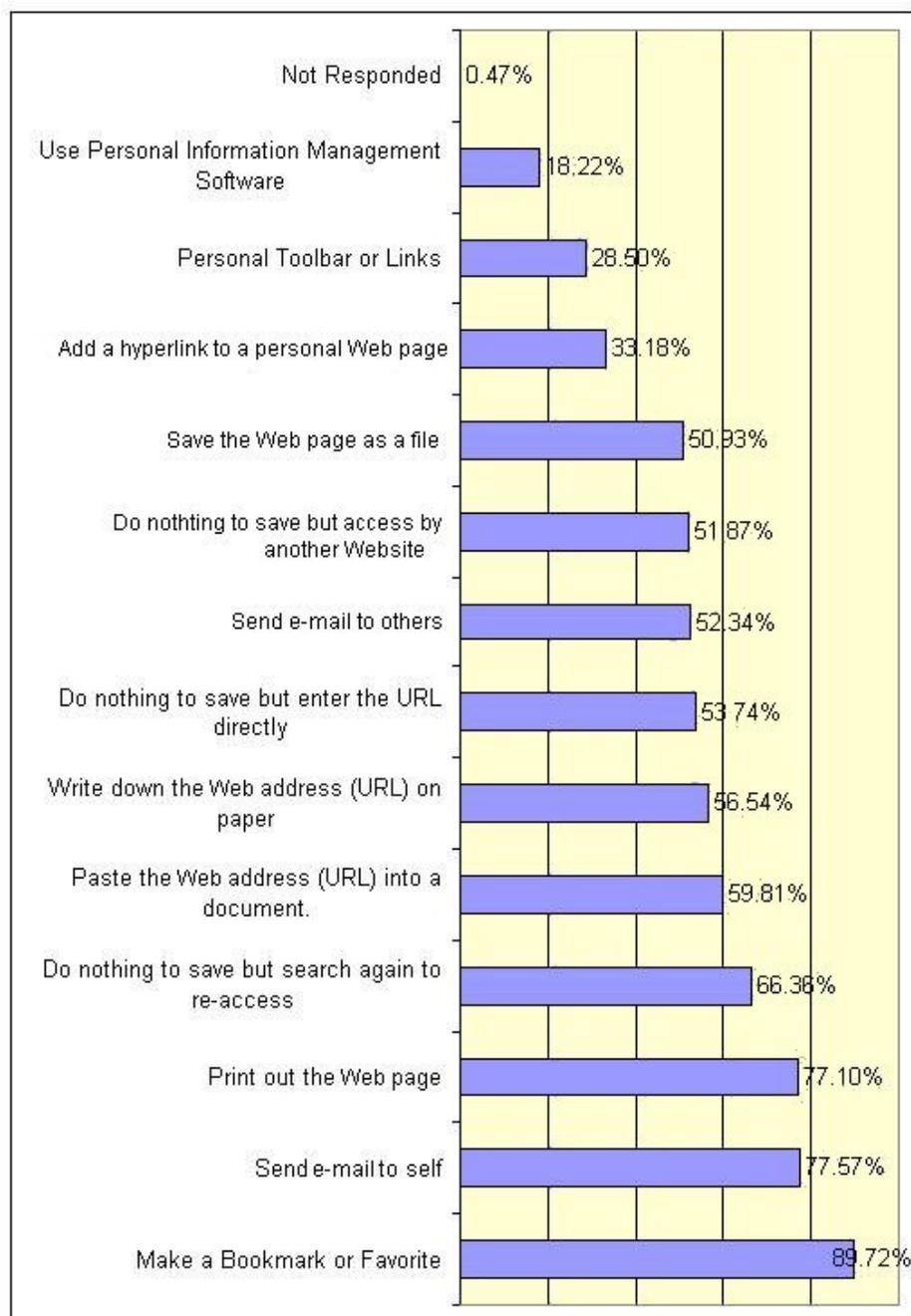


Figure 1: Percentage of respondents who have used a keeping method at least once.

Respondents were then asked to indicate how often they used each keeping method in a typical week using a six-point scale (1-Several times a day, 2-Daily, 3-Two or more times a week, 4-About once a week, 5-Less than once a week, and 6-Never). The data in Table 4 summarize how frequently participants use a keeping method over a typical week. Creating a Bookmark or entering a URL as a Favourite was reported as the most frequently used keeping method in a typical week for participants in the sample. On average, respondents reported that they made a Bookmark or Favourite two or more times a week (see Table 3). The next most frequently used methods are actually leaving methods. Participants indicated that they chose to, *Do nothing to save but search again for re-access* and *Do nothing to save but enter the URL directly*, about once a week.

Keeping Method	Median
Send e-mail to self	5

Send e-mail to others	5
Print out the Web page	5
Save the Web page as a file	5
Paste the Web address (URL) into a document	5
Add a hyperlink to a personal Web page	6
Do nothing to save but search again to re-access	4
Do nothing to save but enter the URL directly	4
Make a Bookmark or Favourite	3
Do nothing to save but access through another Website	5
Use Personal Information Management software	6
Personal toolbar or links	6
Write down the Web address (URL) on paper	5

Table 3: Usage frequency of 'keeping methods' used in a typical week

In question 1 of the survey (the data in Table 4), participants in the sample were asked to select only those keeping methods they had used at some time. Data from the survey revealed that most participants had used approximately seven of the keeping methods listed. When asked which methods were used at least once a week, the data from the sample revealed that participants use on average five of the thirteen methods with this degree of regularity (see Table 4).

	Mean	Median	Mode	Std. Dev.	Max	Min
Methods ever used	7.16	7	8	2.79	13	1
Methods used at least once a week	4.87	5	6	2.54	12	0

Table 4: Average number of 'keeping methods' used

It is interesting to note that the number of keeping methods used at least once a week was consistent across the various groups represented in the sample (see Table 5).

Occupation	Mean	Median	Mode	Std. Dev.	Max	Min
Librarian, information professional	4.96	5	3	2.64	12	0
Researcher	4.64	5	6	2.15	9	1
Manager	5.24	6	6	1.79	8	2
Student	5.3	4	4	2.62	10	1
Other	4.5	4	4	2.45	10	1

Table 5: Comparison of numbers of 'keeping methods' across occupations

To specify this observation further, the data were examined to find out what proportion of the participants in the sample reported using specific methods at least once per week. These proportions were then ranked to identify the five to seven keeping methods that are likely to comprise the standard repertoire for participants. The data appearing in Figure 2 show the top seven keeping methods.

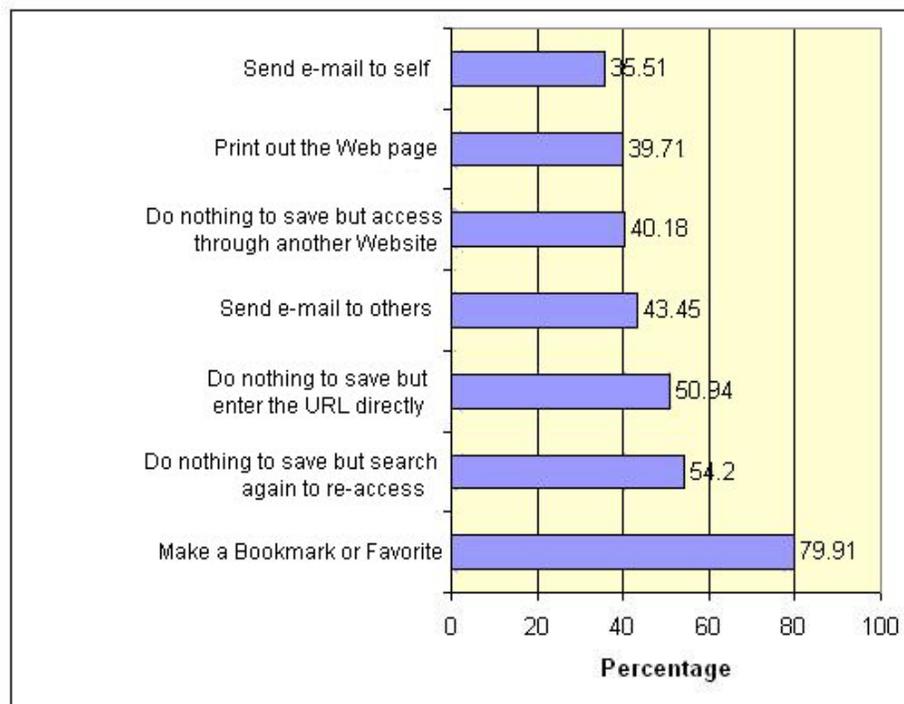


Figure 2: Top seven keeping methods as ranked by proportion of participants using the method at least once a week

The data were then analysed in an effort to reveal the five to seven methods most likely to comprise the repertoire of keeping behaviour for the occupational groups represented in the sample. These data appear on Table 6.

Method	Librarians	Researcher	Manager	Student	Others
Number of respondents	106	42	21	10	35
Send e-mail to self	40.57	30.95	33.33	20	31.43
Send e-mail to others	46.23	40.48	57.14	30	34.29
Print out the Web page	43.40	38.10	23.81	60	34.29
Save the Web page as a file	20.75	23.81	38.10	40	34.29
Paste the Web address (URL) into a document	35.85	26.19	47.62	40	31.43
Add a hyperlink to a personal Web page	18.87	16.67	9.52	20	17.14
Do nothing to save but search again to re-access	51.89	59.52	52.38	70	51.43
Do nothing to save but enter the URL directly	50.94	52.38	57.14	70	40.00
Make a Bookmark or Favourite	83.96	64.29	95.24	70	80.00
Do nothing to save but access through another Website	38.68	30.95	57.14	50	42.86

Use Personal Information Management software	10.38	28.57	19.05	10	17.14
Personal Toolbar or links	21.70	30.95	9.52	40	25.71
Write down the Web address (URL) on paper	33.02	21.43	23.81	10	14.29

Table 6: Comparison of frequency of keeping method across occupations

The five most commonly used keeping methods for each occupational group in the sample in rank order were:

- Librarian
 - Make a Bookmark or Favourite
 - Do nothing to save but search again to re-access
 - Do nothing to save but enter the URL directly
 - Send e-mail to others
 - Print out the Web page
- Researcher
 - Make a Bookmark or Favourite
 - Do nothing to save but search again to re-access
 - Do nothing to save but enter the URL directly
 - Send e-mail to others
 - Print out the Web page
- Manager
 - Make a Bookmark or Favourite
 - Do nothing to save but access through another Website
 - Do nothing to save but enter the URL directly
 - Send e-mail to others
 - Do nothing to save but search again to re-access
- Student
 - Make a Bookmark or Favourite
 - Do nothing to save but search again to re-access
 - Do nothing to save but enter the URL directly
 - Print out a Web page
 - Do nothing to save but access through another Website

It is significant to note that, of the five most commonly used keeping actions for each of the occupational groups represented in the sample; at least two of the methods regularly used require no keeping behaviour. For students, three of the five top-ranked keeping actions fall into this category. As stated earlier, the KFTF research team has been interested in studying the behaviour of keeping and leaving information for re-access and re-use. The popularity of the *Do nothing* or leaving action that appears in the data collected from the survey sample, suggest that participants in the sample are reasonably confident of their ability to remember information left *in situ*, and their ability to find that information again.

Participants in the survey were also asked to describe what influenced their decision to use a particular keeping method. The participant was asked to comment on all thirteen of the methods identified in the keeping study. The data were analysed in light of the functions that emerged from the keeping study and responses were categorized as reported in Table 1. In coding these data, we found that a participant's response could refer to several functions. Each function identified for a selection of keeping methods appears in Table 7. For example, 'There is a graphic on the Web page, or I want a physical reminder for reference or to share'. This response was counted for *Reminding*

and the other for *Preservation of information in its current state*.

	Port-ability	Number of access points	Persist-ence	Preser-vation	Currency	Context	Remind-ing	Ease of integ-ration	Communi-cation	Ease of mainten-ance
E-mail to self	-	83	-	-	-	2	25	17	-	-
E-mail to others	-	2	-	-	-	-	-	-	123	-
Print	34	-	19	7	-	5	12	21	22	-
Save as file	-	-	35	10	-	-	3	17	2	-
Paste URL in document	-	-	-	-	-	20	4	35	15	-
Personal Website	-	9	-	-	-	-	-	8	18	-
Write down URL	-	-	-	-	-	-	13	-	5	-

Table 7: Functions that influence a participant's decision to use a *keeping* method

Examples of the reasons that participants gave for using specific keeping methods appear in Tables 8 to 15, which are presented in the [Appendix](#).

Conclusion

A number of conclusions can be drawn from the empirical studies reported here. It is clear from the data collected by the keeping study and the survey, for example, that individuals use a range of methods to keep and organize information that they have found on the Web and want to re-access and re-use. The keeping methods most regularly used are:

1. Make a Bookmark or Favourite;
2. Do nothing to save but search again to re-access;
3. Do nothing to save but enter the URL directly;
4. Send e-mail to others;
5. Do nothing to save but access through another Website;
6. Print out the Web page; and
7. Send e-mail to oneself.

Each of these methods is used because it has a particular set of functions that help the individual to keep the found information available for use later. But no one method observed in the study presented the user with every function that he or she might need. Many participants, therefore, used several keeping methods and selected the method that was most appropriate (and functional) for the purpose that the information was likely to serve in the future. Generally people have a repertoire of between five and seven keeping methods that they are likely to use as regularly as once a week. These methods remained fairly consistent across the various professional groups participating in the survey.

The most common, regularly used method for keeping information was to make a Bookmark or Favourite but the second most common method was to do nothing: in this case the participants indicated that they would find the information again if they left it where it had been found or encountered. The reason for the choice of this method is revealed by the data in the re-finding study. When faced with the challenge of re-finding information, most people are successful. The re-call trials conducted during the re-finding study demonstrated that re-finding success rates were high and the first method chosen by each participant usually worked. It was interesting to note that two of the top five most commonly used methods that participants chose for re-finding Web information actually required no keeping methods.

It is clear from the data collected by the KFTF study that people continue to be challenged by their own limitations and the limitations of the tools they use for building, managing and using a personal information collection. Fifty years ago Vannevar Bush in the concluding paragraphs of *As we may think* reinforced this concern with the comment that:

...civilization can only continue with the great experiment of its continued development and find fulfillment and joy... if [we]... can reacquire the privilege of forgetting the manifold things... [we do]... not need to have immediately at hand, with some assurance that [we] can find them again if they prove important'. (Bush, 1945: 8)

Half a century after Bush penned these words, we are still (as individuals and members of social and professional groups) confronting the challenges of our overloaded and busy information-based lives with fragmented attention, imperfect memory, and limited information skills and literacies. Effective information seeking and use continues to rest fundamentally with the individual and with his or her ability to create, manage and use a personal information collection. Further research by scientists who study human information behaviour is key to resolving these challenges.

The KFTF study is, therefore, continuing its investigation into the compelling and widespread (individual) challenge of managing the many forms of information that play so fundamental a role in our daily working and personal lives. A second survey has been mounted on the KFTF Website to collect additional data on the functions of keeping methods. The project has recently entered a prototype design, building and evaluation phase and further empirical work is planned for the exploration of personal information collections and the methods people use to represent and organize information that they have assembled to support the tasks and interests that they wish to pursue.

Acknowledgements

This paper owes a debt of gratitude to graduate assistant and MLIS student Hayden Bass, and research assistant and Ph.D. student Ammy Phuwanartnurak, and to the suggestions of the anonymous referees.

References

- Ashill, N. & Jobber, D. (2002). Defining marketing information needs: an exploratory study of senior marketing executives. *Journal of Business and Management*, **8**(2), 165-179.
- Barreau, D.K. & Nardi, B.A. (1995). Finding and reminding: file organization from the desktop. *SIGCHI Bulletin*, **27**(3), 39-43.
- Bellotti, V. & Smith, I. (2000). Informing the design of an information management system with iterative fieldwork. In D. Boyarski & W. Kellogg (Eds.), *Proceedings of the DIS 2000*

conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 227-237). New York, NY: ACM Press.

- Bruce, H. (2002). *The user's view of the Internet*. Lanham, MD: Scarecrow Press.
- Bruce, H., Jones, W. & Dumais, S. (2004). The PAIN hypothesis. Submitted for publication.
- Buchanan, H.S. (1993). Library services and health care administration. *Library Trends*, **42**(1), 62-82.
- Bush, V. (1945). As we may think. *Atlantic Monthly*, **176**(1), 101-108.
- Case, D.O. (1988). The use of anthropological methods in studying information management by American Historians. In C.L. Borgman and E.Y.H. Pai (Eds.), *Proceedings of the First Annual Meeting of the American Society for Information Science (ASIS '88) Information technology: planning for the next fifty years* (pp. 187-193). Medford, NJ: Learned Information, Inc.
- Case, D.O. (1991). The collection and use of information by some American historians: a study of motives and methods. *Library Quarterly*, **61**(1), 61-82.
- Cole, I. (1982). Human aspects of office filing: implications for the electronic office. In R.E. Edwards (Ed.), *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 59-63). Santa Monica, CA: Human Factors Society.
- Ellis, D. (1997). Modeling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, **53**(4), 384-403.
- Hertzum, M. & Pejtersen, A.M. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing and Management*, **36**(5), 761-778.
- Jones, W.P., (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, **9**(3). Retrieved June 29, 2004, from http://www.firstmonday.dk/issues/issue9_3/jones/index.html.
- Jones, W. P. & Dumais, S. T. (1986). The spatial metaphor for user interfaces: experimental tests of reference by location versus name. *ACM Transactions on Office Information Systems*, **4**(1), 42-63.
- Jones, W., Bruce, H., & Dumais, S. (2001). Keeping found things found on the web. In H. Paques, L. Liu, D. Grossman (Eds.), *Proceedings of CIKM'01: the tenth international conference on Information and knowledge management* (pp. 119-126). New York, NY: ACM Press.
- Jones, W., Bruce, H. & Dumais, S. (2003). How do people get back to information on the Web? How can they do it better? In G.W.M. Rauterberg, M. Menozzi & J. Wesson (Eds.), *Proceedings of INTERACT 2003: the 9th IFIP TC13 international conference on Human-computer interaction* (pp. 793-796). Amsterdam: IOS Press.
- Jones, W. P., Dumais, S., Bruce, H. (2002). Once found, what next?: a study of 'keeping' behaviors in the personal use of web information. In E.G. Toms (Ed.), *Proceedings of ASIST '02, the annual conference of American society for information science and technology* (pp. 391-402). Medford, NJ: Information Today, Inc.
- Kaptelinin, V. (1996). Creating computer-based work environments: an empirical study of Macintosh users. In *Proceedings of the 1996 ACM SIGCPR/SIGMIS conference on Computer personnel research* (pp. 360-366). New York, NY: ACM Press.
- Kuhlthau, C.C. & Tama, S. (2001). Information search process of lawyers: a call for 'just for me' information services. *Journal of Documentation*, **57**(1), 25-43.
- Kwasnik, B. H. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation*, **47**(4), 389-398.
- Lansdale, M. W. (1988). On the memorability of icons in an information retrieval task. *Behaviour & Information Technology*, **7**(2), 131-151.
- Leckie, G.J, Pettigrew, K. & Sylvain, C. (1996). Modeling the information seeking of professionals: a general mode derived from research on engineers, health care professionals, and lawyers. *Library Quarterly*, **66**(2), 161-193.
- Malone, T. W. (1983). How do people organize their desks: implications for the design of office information-systems. *ACM Transactions on Office Information Systems*, **1**(1), 99-112.
- Mander, R., Salomon, G., Wong, Y.Y. (1992). A 'pile' metaphor for supporting casual organization of information. In P. Bauersfeld, J. Bennett, G. Lynch (Eds.), *Proceedings of the ACM SIGCHI conference on Human factors in computing systems* (pp. 627-634). New York, NY: ACM Press.
- Meho, L. & Haas, S. (2001). Information-seeking behavior and use of social science faculty

- studying stateless nations: a case study. *Library and Information Science Research*, **23**(1), 5-25.
- Miller, G.A. (1968). Psychology and information. *American Documentation*, **19**(3), 286-289.
 - Owen, D. (2003). Information-seeking behavior in complementary and alternative medicine (CAM): an online survey of faculty at a health sciences campus. *Journal of the Medical Library Association*, **91**(3), 311-321.
 - Pettigrew, K., Fidel, R. & Bruce, H. (2001). Conceptual frameworks in information behavior. *Annual Review of Information Science and Technology*, **35**, 43-78.
 - Sandstrom, P. (1994). An optimal foraging approach to information seeking and use. *Library Quarterly*, **64**(4), 414-449.
 - Sellen, A.J. & Harper, R.H.R. (2002). *The myth of the paperless office*. Cambridge, MA: MIT Press.
 - Smith, D. & Culkin, N. (2001). Making sense of information: a new role for the marketing researcher? *Marketing Intelligence and Planning*, **19**(4), 263-271.
 - Sutton, S.A. (1994). The role of attorney mental models of law in case relevance determinations: an exploratory analysis. *Journal of the American Society for Information Science*, **45**(3), 186-200.
 - Tauscher, L.M. & Greenberg, S. (1997). How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies*, **47**(1), 97-137.
 - Watson-Boone, R. (2003). The information needs and habits of humanities scholars. *RQ*, **34**(2), 203-214.
 - Whittaker, S. & Sidner, C. (1996). E-mail overload: exploring personal information management of e-mail. In M.J. Tauber (Ed.), *Proceedings of the ACM SIGCHI conference on Human factors in computing systems* (pp. 276-283). New York, NY: ACM Press.

Appendix

Examples of the reasons given by participants for using specific keeping methods

Reason	# Responses
Number of access points <ul style="list-style-type: none"> • I do this when I'm going to use it at another site • I may find a site at work that I wish to access when I get home later that evening. • If I'm at someone else's computer when I see it. • It enables me to access the link from different computers. • When I find sites at work that I want to use at home or vice versa. 	83
Reminding <ul style="list-style-type: none"> • Used a reminder to investigate the site further at a later date. • Easy to retrieve and remember to complete task. • If I'm away from home, and I want to remind myself to visit a site later, this is a good method. • When I want to remind myself to do an activity or project that the Web information supports. 	25
Ease of integration <ul style="list-style-type: none"> • It is a simple matter to file the e-mail in a subject file of like resources. • To record for workflow purposes. • I save e-mails in categorized folders in Outlook Express. If I have an appropriate folder, I'll e-mail a Web page to myself along with 	17

<p>the URL.</p> <ul style="list-style-type: none"> • If I have a folder set up in my e-mail for that topic, I will send the URL so that all information on that topic is in one place. 	
<p>Context</p> <ul style="list-style-type: none"> • I can make comments about the site. I can click and open a site from e-mail. • When I also want to make a note about the site, not just record the URL, but also remind myself of some context or particular thought I had about the site. 	2
<p>Others (unclassifiable)</p> <ul style="list-style-type: none"> • New York Times articles only. • If the URL list is long (from a site). • I know it's there somewhere! • Usually when I'm isolated from everything but my e-mail. 	21

Table 8: E-mail to oneself

Reason	# Responses
<p>Communication and information sharing</p> <ul style="list-style-type: none"> • This is my primary way of sharing sites with others. • To notify friends of interesting pages. • I want to share the information with my friends or a specific interest group. 	123
<p>Number of access points</p> <ul style="list-style-type: none"> • Usually incident to moving links from office to home. • Sometimes, if I find a good URL at home I will e-mail the URL to my work address to use there. (I am counting my work address as "others.") 	2
<p>Others (unclassifiable)</p> <ul style="list-style-type: none"> • To save me time while browsing the Web - if there's a "click to send this page," I love that. Then at my convenience I can re-review the page and bookmark if interested. • Convenience. • Reliable. 	7

Table 9: E-mail to others

Reason	# Responses
<p>Portability</p> <ul style="list-style-type: none"> • Mostly do when the followup is offline. 	

<ul style="list-style-type: none"> • When I'll be traveling and won't have easy access to the internet or a computer. • Read on the train. 	34
<p>Communication and information sharing</p> <ul style="list-style-type: none"> • I want to show it to my fellow librarians. • Mainly to hand to the patrons doing research on school projects. • Want to bring it home to my kids for them to use later. 	22
<p>Ease of integration</p> <ul style="list-style-type: none"> • When doing research at work for articles to be written. • To add to a topic file that I keep (e.g., teaching resources, health/food topics of personal interest). • I do this when the Website is a useful part of an ongoing project, and will need to be filed with the project documents. 	21
<p>Persistence of information</p> <ul style="list-style-type: none"> • I do this when the information on the page is likely to change, will be used regularly or is not amenable to saving. • To have it in case site disappears. Like plans or circuit diagrams I plan to use at a later date. • If the information is transitory and I want to keep a copy. 	19
<p>Reminding</p> <ul style="list-style-type: none"> • If I don't have time to read it and it's really important to me. • Something I want to look at and may reference for future use. • As a reminder to reread the material. 	12
<p>Preservation of information in its current state</p> <ul style="list-style-type: none"> • There is a graphic on the Web page. • When the formatting, or organization, is the important thing. • Only if the mix of pictures and text are important to be kept as is. 	7
<p>Context</p> <ul style="list-style-type: none"> • Usually only if I need to annotate the page or take notes. • I can make any notations on the paper to trigger why I wanted to use this site in the first place. • When I want a visual to add in a (paper) file to remind me WHY I wanted to save it in the first place. 	5
<p>Others (Unclassifiable)</p> <ul style="list-style-type: none"> • Whole forests have gone down for me to keep printing. I felt bad until I read that study about what people do with all the paper on their desks. • When the URL is especially long. • Useful for FAQs, knowledge bases, where searches can be difficult. • Yes, this is helpful. I'm still paper-oriented. • Infrequently access that page. 	11

Table 10: Print a Web page

Reason	# Responses
Persistence of information <ul style="list-style-type: none"> To capture information that will be needed for a short term only, but I don't want it to be subject to change or deletion at the server side. When I want to look at it in greater depth and I'm afraid it might disappear. When I think the Web page may disappear. 	35
Ease of integration <ul style="list-style-type: none"> To add to a file in my documents. If I am working on a project, it is worthwhile to save the file in my project folder. Ideas that I want to incorporate into something I am writing. 	17
Preservation of information in its current state <ul style="list-style-type: none"> When I am interested in the page as a whole - the layout, images or other components, in addition to the text info. If I need to have the same content as when I viewed it. Visuals are important or whole Web look, example for my own Web designing, key topics. 	10
Reminding <ul style="list-style-type: none"> If the page holds an immediate or special interest, I save it as a desktop file, then discard or otherwise file later. I save a Web page as a file when I think I will need it again within the day. On my desktop (Mac) if it's something that will be needed in the next week. 	3
Communication and information sharing <ul style="list-style-type: none"> If I think it is a topic I will be sending others repeatedly. Keep in shared folder as online documentation. 	2
Others <ul style="list-style-type: none"> For offline browsing. I find it easier than using software that does that. I do this a lot because my bookmark file won't accept any more entries (AOL). For information I will use again, but which is too long to print. 	11

Table 11: Save as a file

Reason	# Responses
Ease of integration	

<ul style="list-style-type: none"> • When writing reports I will insert a link to reference the source of the information. • I do this when creating reports that will be shared with others so that they can access the site as well. • When I already have a document relating to the subject. 	35
<p>Context</p> <ul style="list-style-type: none"> • I use this sometimes when I am doing research on something specific. I will cut and paste the URL and then write some notes about the site. • Allows me to add comments. • This is how I save links in the context of other information (e.g., a list of graduate schools I'm considering attending, along with links to their sites). 	20
<p>Communication and information sharing</p> <ul style="list-style-type: none"> • To share with fellow employees. • Sending a link to a group in my contact list. • Rarely do this unless I have to share it with a wide audience or am at a temporary machine. 	15
<p>Reminding</p> <ul style="list-style-type: none"> • Reminders to go back when time, lists to send to others later, reminders to put on Website. • If I am in a hurry, I will save a URL as a draft e-mail, and store it in my e-mail software - not very organized, difficult to maneuver in, though. 	4

Table 12: Paste URL into a document

Reason	# Responses
<p>Communication and information sharing</p> <ul style="list-style-type: none"> • When I want to share it with students. • For links related to my Websites that others might find useful too. • Places I want other people to go to. 	18
<p>Number of access points</p> <ul style="list-style-type: none"> • Almost always. I click 'post to Weblog' and then I just search my blog later, from any machine. • Easy access. • I copy my list of bookmarks to "Yahoo! Bookmarks," so that I may have remote access to them. 	9
<p>Ease of integration</p> <ul style="list-style-type: none"> • When related to the topic of my Web page. • When appropriate to a document I am working on. 	8

Table 13: Add a hyperlink

Reason	# Responses
Easy to remember URLs (e.g., common URLs and short URLs) <ul style="list-style-type: none"> • Use when very commonly used or well known or easily findable (from a browser) Website. • When the URL is short and obvious, I am more likely to just search again. • If it's a catchy or easy title (Web page) and I am certain I can find it easily. 	14
Easy to search <ul style="list-style-type: none"> • If it seemed easy to find or is common information. • I know I will be able to find easily. • When the result comes up at the top of the search and the terms used in the search represent the name of the organization. 	31
Powerful search engine (e.g. Google) <ul style="list-style-type: none"> • Since the advent of Google, I rarely, if ever, add any links to my Favourites. • Sometimes, this is easier when topic phrases via Google are known to me to produce what I want on the first page of hits. • Sometimes I use the major search engines, e.g. Google, Vivisimo, etc., to search the same topic again just in case there are brand new sites with better and more relevant info that I might have missed. 	24
Rarely visited Website <ul style="list-style-type: none"> • For things I don't expect to need very often. • If it's a site I'll use rarely. 	8
No intention to revisit the site <ul style="list-style-type: none"> • If I don't think I'll need it again, so I don't bookmark it - of course, the following week, I always want it! • When I didn't know I'd want the page again. • When I think that I will not need to access the page again. 	10

Table 14: Do nothing but search

Reason	# Responses
Easy to remember URLs (e.g., common URLs and short URLs) <ul style="list-style-type: none"> • Only short-titled sites, such as cnn.com. • Fairly often, for easy-to-remember addresses. • Simply named sites -- colleges and universities, news sources, etc. 	39

Frequently visited Website	
<ul style="list-style-type: none">• Some are used so often and are memorized so don't need to write them down.• I do this for many sites that I use every day.• If I have a very-frequently used site that is simpler to type than to locate my bookmark.	43

Table 15: Do nothing but enter the URL directly

Find other papers on this subject.

Articles citing this paper, [according to Google Scholar](#)

How to cite this paper:

Bruce, H., Jones, W. & Dumais, S. (2004). "Information behaviour that keeps found things found." *Information Research*, 10(1) paper 207 [Available at <http://InformationR.net/ir/10-1/paper207.html>]

Check for citations, [using Google Scholar](#)

10897
[Web Counter](#)

© the author, 2004.
Last updated: 14 August, 2004



[Contents](#) | [Author index](#) | [Subject index](#) | [Search](#) | [Home](#)
