Searching In Public versus Re-finding In Private: Are they the same?

Azrina Kamaruddin

Computing Department
Lancaster University, Lancaster, LA1 4WA United Kingdom
A.Kamaruddin@comp.lancs.ac.uk

Abstract

Introduction

My work is set within the area of personal information management (PIM). More particularly, I am interested in the process of re-finding and the emotional aspect of this. Here the term re-finding is the process where resources, such as web pages, that have previously been seen by a user are later retrieved. Whereas the original process of finding the resources often occurs in the public domain (e.g. WWW), the process of re-finding often uses private collection such as bookmarks. The research will investigate the experienced and novice users. We feel that these two types of users employed different strategy and approach in achieving their information target. We would like to bridge the gap between these two users by having a generic model which both users can benefit from it.

Previous Work

Several studies by [1,2,7,3,9] focussed about re-finding activity. None of the previous study pointed out the emotional aspects of the re-finding process and the differences approach taken by experienced and novice users. The project known as Keep Found Things Found (KFTF) by Jones et al.[7], found that there are several methods that people used to keep track of their web pages for re-use and developed a functional analysis to show how techniques people use depend on anticipated access needs. Capra and Perez [2], found that re-finding consists of two-stage iterative approach to information re-finding, relies heavily on the use of contextual information and domain artefacts and explicitly added artefacts. Haystack project [3], came out with the PIM system which stressed on user's control rather than the application. The prototype system by Dumais et al. [5] and Teevan [10] developed to support users in retrieving their information. Stuff I've Seen(SIS) systems provides a unified index of any of any form of information (email, web page, files, calendar, etc.) and because of the person may has seen the information before, it constitutes a rich contextual cues such as time, author, thumbnails and previews can be used to search for [5]. The Re:Search Engine system helps by preserving their first search results while presenting the new search results by having both old and new results on the same interface. It focuses on consistency in the search results it returns across time [12].

Research Aims

My research is aimed at understanding on these two particular issues: the distinction between the public and private environment in the process of re-finding activity and the issues of uncertainty and its relationship with re-finding process.
Objectives:

This research is divided into three main areas:
1) To understand the distinction between public and private environment in detail especially through the process of searching/re-finding activity (attributes, behaviour, and stages).
2) To understand the emotions and context used in re-finding process – whether the feelings of the uncertainty user felt in both domains compatible with Kuhlthau's information-seeking model results and context used difference or similarity in both domains.
3) To develop or improve an information re-finding model that takes into account in emotions and context distinction between public and private domain in the larger processes in context.
4) To develop a tool (tools) to investigate, evaluate, and refine further understand and effectiveness of the model.

Methodology

This research uses several different methodological techniques in order to provide rich data about the phenomenon of both searching and re-finding activities. These include semi-structured interviews, direct observation, and automated data gathering using a software tool that will record the organization of information in files, emails and bookmarks for each participant. The participants are staffs and students at the Computing Department, Lancaster University and also home users.

Work Completed – Pilot study

Information Search Models Mapping

There are several information search models by Kulthau [8] and Ellis [6]. Information search models provide steps or activities taken to search for information. In general, information search models [8,6] have similar steps started from Initiation – Selection – Exploration – Formulation – Collection – Search closure / Presentation. These steps are true for searching activity. Ellis [6] who is one of the information science researcher proved these steps and suggested searching in WWW using these stages. In one of Teevan's[11] reports points out that several participants think that searching and re-finding are the same. However, our stand, based on our piloted study, we find that searching is not re-find. Our stand is also been support by findings of Capra and Perez [4]. They find that information re-finding is a two-stage iterative approach. The approach is focussed on 1) relocating the source of information (searching) and 2) engaging in a process to re-locate the specific information being sought (browsing). We concluded from our pilot study that the information search models does not 100% map with the re-finding scenario we created. It shows that searching is not re-finding.

Diary study

Before we conducted a real diary study, we piloted the diary on ourselves. We recorded all our activities while searching and re-finding. In order to produce a better task for users to complete their diary, we must make sure that users will provide the data that we want. By implementing it on ourselves, we have the general idea of what users would be recording and what kind of information we are looking for. During our piloted study, we find several interesting output. We presented our initial findings in the table form. Our stand of differences between searching and re-finding also been supported by Capra and Perez [4] research and findings in the mobile domain.

Our initial findings are shown in table 1, the differences between searching and re-finding from our piloted study. However, we need to further investigate more about the differences to wider scope of users. Table 2, presents the differences find by Capra and Perez[4].

Table 1: Differences between Searching and Re-finding.

	Searching	**Re-finding**
Strategies	Plan appropriate keyword. Intuition	Browse the list→ figuring out. Directed focused search.
Methods Results presented	Listing of relevant and precision from the keyword. Results from various places.	Results are from you. The user selects their targeted result from a storage which they think the results in it.
Interface	Short/Long list item presented based on relevancy and precision. Based on user queries.	Hierarchy based.Static
Emotion	Emotions: Uncertain → Satisfied/Dissatisfied	Certain/Confident→Satisfied (Not always the case).
Decision making	To keep or not to keep	Where is it? Which folder?
Contents	Dynamic	Static

Table 2: Differences from Finding and Re-finding by Capra and Perez[4].

Finding	**Refinding**
Uncertainty: Is the information available at all? Do I know the right keyword to use in a search engine?	Certainty: I have seen the information already, but where? What was the keyword that I used in the search engine?
Recognition: Is this the information I am searching for?	Recognition & Recall: Where did I see that? Is this where I saw it? Context is very important.
Strategies: Intuition, search tools, foraging, browsing	Strategies: Directed(focused) navigation

Study 1

Our first study will take place in the middle of October. An initial interview will be held at our participants' workplace in front of their personal computer. An initial interview will provide us detail of user computer and Internet backgrounds. We will ask questions regarding their searching and re-finding activities. An interview, will give us an opportunity to observe their files, emails and bookmarks. An initial interview will give us information about their usual activities and strategies while re-finding. During the initial interview, we will ask their consent to participate in our diary study. Instead of writing down every activity, they just need to write down their unusual activities that they experienced during re-finding activities.

Study 2

In the second stage of our data gathering, we decided to use diary study as one of our observational method. We undertake the diary study in order to understand more about their re-finding activity which we could not acquire during our first interview session. The duration the diary study will be from one to two weeks. Diary studies have high ecological value as they are carried out *in situ*, in the users' real environments. On the negative side, diary studies suffer from the problem that they are tedious for the recorder and they can invoke a "Heisenberg-style" challenge: the process of observing may influence the observations in that journaling tends to add to the interruption of the flow of daily events. Despite these problems, we felt it was overall beneficial to start from ecologically valid data that might reveal interesting patterns of re-finding strategies and, while realizing that there would be imperfections with regard to comprehensiveness and accuracy.

Study 3

Automated log file is not unusual technique to record users' activities. In fact, it will be able to help researchers and users in certain ways. Users will have benefit of progressing with their daily work without any interruption; researchers have the chance to monitor their activities without worrying users would be acting unusual towards their data. The purpose of this log tools, is to find out ways user categorize their files, emails and bookmarks. Few questions than can be learn by employing this technique are:
- Do they apply same method for categorizing files, emails and bookmarks?
- Why?
- Why they mix the categorization and indexation methods?
- How these help them in re-finding?
- What cues do they use to make the categorization?
- How frequent they manage their storage?

Study 4

In Jones et al [7] Keep Found Things Found (KFTF) project, they find that 51.87% users like to "Do nothing to save but access by another web site" methods and 33.18% users create code link in their own web sites. This triggers us with several questions:
- Why people use the strategies?
- Why people prefer to move back to public environment?
- What kind of method users' employ to create link in their web page?
- If they use categorization, is there any similarities and dissimilarities between what they keep in their private and in their public information space?

Conclusion

Understanding searching and re-finding activity for experienced and novice users are not entirely new area in PIM area, but there is no an agreed model that can actually help users and designers towards the development of re-finding systems. Through this research we would like to understand the difference people employed while doing these two activities, what people understand by private and public information space and searching is not re-finding. At the end of the research we hope to achieve:
5) To understand the distinction between public and private environment in detail especially through the process of searching/re-finding activity (attributes, behaviour, and stages).

6) To understand the emotions and context used in re-finding process – whether the feelings of the uncertainty user felt in both domains compatible with Kuhlthau's information-seeking model results and context used difference or similarity in both domains.

7) To develop or improve an information re-finding model that takes into account in emotions and context distinction between public and private domain in the larger processes in context.

8) To develop a tool (tools) to investigate, evaluate, and refine further understand and effectiveness of the model.

References

[1]Abrams, D., Baeker, R., and Chignell, M. Information Archiving with Bookmarks: Personal Web Space Construction and Organization. Proceedings of CHI 1998.

[2]Alvarado, C., Teevan, J., Ackerman, M.S. and Karger, D. (2003). Surviving the Information Explosion: How people Find Their Electronic Information. AI Memo 2003-006, Massachusetts Institute of Technology Artificial Intelligence Laboratory.

[3]Capra, R.G., and Perez-Quinones, M.A. (2003). Re-Finding Found Things:An Exploratory Study of How Users Re-Find Information. Technical Report, Virginia Tech.

[4]Capra, Robert and Pinney, Mary and Perez-Quinones, Manuel A. (2005) Refinding is Not Finding Again. Technical Report TR-05-10, Computer Science, Virginia Tech.

[5]Dumais, S., Cutrell, E., Cadiz, JJ., Jancke, G., Sarin, R., & Robbins, C., D., (2003). Stuff I've Seen: A System for Personal Information Retrieval and Re-use. In the Proceedings of (SIGIR'03), Toronto, Canada. July 28-August.

[6]Ellis, D. (1989). A behavioural approach to information retrieval design. Journal of Documentation, 46, 318-338.

[7]Jones, W., Bruce, H., and Dumais, S. (2001). Keeping Found Things Found on the Web. In Proceedings of the Tenth International Conference on Conference on Information and Knowledge Management. ACM Press, New York, NY, 119-126.

[8]Kuhlthau, C.C. (1993). *Seeking meaning: A process approach to library and information services*. Norwood, NJ: Ablex.

[9]Marchionini, G.N. (1995). *Information seeking in electronic environments*. Cambridge, Eng.:Cambridge University Press.

[10]Teevan, J (2005). The Re:Search Engine Helping People Return to information on the Web. To appear in Proceedings of the 28th Annual ACM Conference on Research and Development in Information Retrieval (SIGIR'05), Salvador, Brazil. August 2005 (Doctorial Consortium).

[11]Teevan, J. and Belkin, N.J. (2005) Don't Drop the Ball: Re-finding Personal Information The PIM Workshop, An NSF Sponsored Invitational Workshop of Personal Information Management (NSF PIM Workshop), Seattle, WA, January 2005. Breakout Group Report.