Library Hi Tech

The role of memory in document re-finding

Xiao Xie Diane H. Sonnenwald Crystal Fulton

The role of memory in document re-finding

Xiao Xie
School of Management, Anhui University, Hefei, China
Diane H. Sonnenwald
Royal School of Library and Information Science, University of Copenhagen, Copenhagen, Denmark, and
Crystal Fulton
School of Information & Library Studies, University College Dublin, Dublin, Ireland

Received 6 June 2014
Revised 30 October 2014
Accepted 14 December 2014

Abstract

Purpose – The purpose of this paper is to explore graduate students' behaviour and perspectives regarding personal digital document management, as well as insights into the connections between memory and document re-finding.

Design/methodology/approach – Semi-structured interviews were conducted with 15 graduate students studying information and library science. The interviews were digitally recorded and transcribed. The transcripts were analysed using open and axial coding.

Findings – Participants were overall positive about the importance of managing their digital documents but they had little knowledge about currently available personal information management (PIM) tools. Characteristics of digital documents frequently used by participants to re-find documents include name, subject, storage location, creation time, keyword, document title, document file type, user's location and recency. For participants the act of organizing documents is itself a memory aid. Participants' recommendations for PIM tools include support for information organization and simplistic visualizations that can be customized, e.g., using colour to highlight folders or documents.

Research limitations/implications – The number of study participants was relatively small, and further studies should examine a more diverse participant sample, e.g., to investigate whether tasks influence re-finding. Further studies should also examine PIM with respect to other types of devices and services, including tablets and cloud services.

Practical implications – The results include recommendations for future PIM tool design.

Originality/value – This research identifies documents' characteristics that participants use to re-find documents and the importance of these characteristics. It also examines the usage and expectations of PIM tools in everyday PIM.

Keywords Information management, Information science, Information resources management, Personal information management

Paper type Research paper

1. Introduction

The number of digital documents that individuals create, collect and store for later use continues to increase, and thus personal information management (PIM), i.e., the practice of how individuals create, collect, organize and retrieve their personal information stored in digital formats, has become an important area of research. The ultimate goal of PIM research is to provide assistance to individuals in managing their

The authors thank the study participants and the anonymous reviewers for their helpful comments. This research was supported by Chinese Scholarship Council, UCD CSC Scholarship, and the School of Information & Library Studies at University College Dublin.

Library Hi Tech
Vol. 33 No. 1, 2015
pp. 83-102

0737-8831
DOI 10.1108/LHT-06-2014-0050

personal information. Personal information today comes in many formats, including e-mail messages, digital audio files, paper and digital photographs and paper and digital documents. This paper focuses on PIM with respect to digital documents stored on laptop computers, the most frequent context for personal PIM reported by participants in this study. We explore the practice of re-finding these digital artefacts, and the role memory plays in this process. This paper focuses on one type of digital artefact, documents stored on laptop computers. According to Boardman and Sasse (2004), individuals tend to organize desktop/laptop files more often than e-mail messages or bookmarks. In addition when users need to reuse something, they are more likely to choose documents which are stored on their computer rather than e-mails or bookmarks which are stored on a remote server. This paper builds on these results, focusing on users' re-finding behaviour with respect to documents stored on their laptop computers.

There have been several studies that explore PIM from the aspect of human memory. Bergman *et al.* (2008) show that when individuals need to re-find information, they prefer using navigation strategies that rely on memory. Thus, memory appears to play a very important role in PIM. However, human memory is limited. There is substantial evidence that individuals cannot rely on their memory alone to remember all of their experiences (Elsweiler *et al.*, 2007), and multiple studies have demonstrated that memory problems hinder people's ability to re-find information (e.g. Czerwinski and Horvitz, 2002; Bruce *et al.*, 2004; Jones *et al.*, 2005).

Previous research striving to improve information re-finding efficiency has primarily focused on developing new PIM tools to help individuals more easily and effectively manage their personal digital information. Two examples are Haystack and Stuff I've Seen (SIS). Haystack (Quan *et al.*, 2003) uses semantic web technologies, including the Resource Description Framework data model, to help users manage their personal documents, e-mails and schedules. With Haystack (Karger, 2007) users can store descriptions and relations connected to their digital information objects and then later navigate their digital information like they navigate web pages. The goal is to help users easily re-find the information they need by connecting related information. The system, SIS (Dumais *et al.*, 2003), provides an integrated approach to re-find different types of information. SIS builds a unified index for all digital artefacts that users have stored or accessed on their computer, including documents, e-mail, web pages, calendar entries and media files. SIS enables users to search the index using queries and filters such as type of item (outlook web pages, files), type of file and date.

Very few studies have investigated the use of PIM tools in everyday practice. Recent research (Jones, 2011; Jones and Anderson, 2011) reports that there is a gap between what people actually need and what existing tools can provide, and that information fragmentation becomes worse in today's distributed, cloud-based data environment. There are no existing tools that support PIM across multiple, diverse systems. Our research results also show that individuals either do not know about available PIM tools or they perceive they do not need them. For example, only a few participants said they had heard of Google Desktop, and they stated that they never use it because they do not have a need for it. Thus there is a need to further examine everyday PIM practices in order to design tools that can better support re-finding digital artefacts.

The research reported in this paper was conducted to gain a better understanding of everyday PIM practices, including the role memory plays

in re-finding digital documents. Specific research questions we address in this paper are:

RQ1. What is the role of PIM in everyday life?

RQ2. What characteristics do individuals remember about the documents they are looking for, and which characteristics are remembered most frequently?

RQ3. What additional functionality should PIM tools ideally provide?

In Section 2 of this paper we discuss related research in PIM. Section 3 describes the research method we used in our study, and our results are presented in Section 4. Section 5 summarizes the results and discusses future work.

2. Related work

Our research has been informed through prior research in human memory and PIM, including PIM systems, or applications, that have been developed.

2.1 PIM and human memory

Human memory plays an important role in PIM today. Though many tools, such as Google Desktop Search, Haystack and Semex (SEMantics Explorer), were developed with the goal of improving information re-finding, people continue to rely on their own memory to re-find information in their personal space of information (PSI[1]).

A study conducted by Bergman *et al.* (2008) concluded that when users need to re-find their information, they have a strong preference for navigation which relies on human memory. They estimated that on average 56-68 per cent of re-finding activities conducted by their study population were done using navigation while a much lower percentage were conducted using search tools. Their analysis of log files showed that only 4-7 per cent of retrieval events happened using Windows search tools and 11-15 per cent for Google desktop, Sherlock and Spotlight combined. They also demonstrated that there was no evidence that using desktop search engines lead people to change their filing habits or become less reliant on hierarchical file organization.

It has been observed that memory problems hinder people's ability to re-find information (Czerwinski and Horvitz, 2002; Bruce *et al.*, 2004; Jones *et al.*, 2005). In the "Keeping Found Things Found" project, Bruce *et al.* (2004) demonstrated that people sometimes forget what information they kept in their PSI, and even worse, they forget the location or directory of important information. Some scholars (Czerwinski and Horvitz, 2002; Jones *et al.*, 2005) have suggested that people sometimes forget the information they have kept in their PSI even when they thought that information would be valuable in the future.

Previous research has also shown that context and the passage of time play important roles in PIM. Psychological research shows that people are good at memorizing the meaning of an event, but in general are poor at remembering the details of an event (Sachs, 1967) unless it is a critical event (Flanagan, 1954). Many studies have confirmed these findings (e.g. Fuller *et al.*, 2008; Elsweiler *et al.*, 2011; Madlock-Brown *et al.*, 2012). Furthermore people appear to be very good at remembering things that happened recently, but the key details that people remember about a specific event will diminish, even disappear, over time (Rubin and Wenzel, 1996; Czerwinski and Horvitz, 2002; Fuller *et al.*, 2008).

Fuller *et al.* (2008) suggest that context can be used to link and annotate items mirroring the way people remember items. In their book, Tan *et al.* (2007) describe the use of a SenseCam and personal digital store to support a patient with severe memory

loss due to brain injury. They compared the use of SenseCam with the use of a diary written by the same patient to aid recall of recent events. The results showed that the patient has a much better recall after viewing SenseCam images than after reading their written diary.

So how do people remember documents in order to re-find them? Certain characteristics of information appear to influence what we can remember about information and how we can handle documents which we think will be useful in the future. What specific characteristics of information do people remember when they want to re-find their personal information?

In 1975, psychologist Reicher (1969) evaluated Broadbent's classic models of human information processing, and the results showed that context and structure has a significant influence on human memory and remembering. Many scholars, including Carroll (1982), Barreau and Nardi (1995) and Gemmell *et al.* (2002, 2006), have suggested that standard forms of contextual metadata, such as time, location, people and names, are included in memories. Goncalves and Jorge (2004) asked study participants to describe various documents. The document characteristics the participants mentioned most frequently were: time, place, co-author, purpose, subject, other documents, format, document exchanges, tasks, storage and contents. At a National Science Foundation workshop (Jones and Bruce, 2005), scholars suggested that people remember the following about their personal information: content, context, time, people, storage location, physical characteristics, distinctiveness, encoding effects and recency and frequency. The characteristics, time and content, are common across these studies. However, in an experiment Blanc-Brude and Scapin (2007) found that study participants were able to recall the following characteristics of paper and digital documents most accurately: type or format, visual elements, location and title. Time was only correctly recalled 4.8 per cent of the time. Blanc-Brude and Scapin did not ask participants to recall document content, but rather asked them to recall keywords, i.e., meaningful words within the documents. Participants only recalled keywords accurately 32 per cent of the time. Further research is necessary to identify which characteristics are most effective in helping people re-find their personal information, and what, if any, factors may influence their effectiveness.

2.2 PIM tools

Previous research focusing on PIM tools has resulted in tools and design recommendations for tools to help individuals manage and re-find their personal information. The Compatible Time-Sharing System (Corbató *et al.*, 1962) was the first prototype system to support personal file storage.

In the last two decades, a variety of PIM tools and prototypes have been created, with different ideas put forward by researchers and developers. These ideas can be roughly divided into two categories; those that primarily use metadata and those that primarily employ semantic, or associative, links.

There are many examples for the first category. The prototype system developed by Chirita *et al.* (2005) translates notifications generated by the Linux operating system regarding file modifications into metadata that are attached to the files. The metadata can then be searched when trying to re-find files and e-mail messages. Xiao and Cruz (2005) designed a prototype that incorporates a layered ontology-based framework, comprised of metadata that can be generated by a user or via natural language processing techniques such as text summarization. A prototype created by Murthy *et al.* (2006) allowed users to create superimposed information, e.g., comments, attached

to digital objects that can be searched to assist in re-finding these objects. Fuller *et al.* (2008) designed a PIM prototype that incorporated metadata, such as date, time, source, type, weekday, location, weather, season, duration, and surrounding events, available from systems. The user could re-find personal information by searching the metadata.

The second category of tools primarily employs semantic, or associative, links to support information re-finding. As Aizawa *et al.* (2004) point out remembering the context of our experiences is easier than remembering details about them, so contextual links among objects may be valuable in re-finding objects. The Life Log prototype (Aizawa *et al.*, 2004) strives to record life experiences using sensors and other tools that collect video, audio, acceleration sensor, gyro, GPS, annotations and access to documents, web pages and e-mail messages, and links among them. The prototype strives to capture and use links among the life log data and documents, web pages and e-mail messages to facilitate information re-finding. Haystack (Quan *et al.*, 2003) allows users to add relationships and properties to any objects, including personal documents, e-mails and calendar entries. These relationships can then be searched to re-find objects. Chau *et al.* (2008) designed Feldspar (Finding Elements by Leveraging Diverse Sources of Pertinent Associative Recollection) prototype that allowed users to re-find documents by constructing multi-level associated retrieval queries which support re-finding by re-iteratively specifying objects related, or associated, with the focus of their query.

In summary, memory of specific characteristics of documents appears to influence individuals' information re-finding efficiency. Prior research has identified some of these characteristics. PIM prototype tools have focused on using metadata or semantic links to support re-finding. Further research may provide additional insights regarding PIM tool functionality preferred by users.

3. Research method

3.1 Research design

To investigate participants' PIM behaviour, their memory about digital documents, their use of PIM tools, and their expectations of ideal PIM tools, a qualitative empirical study was conducted. Semi-structured interviews with open-ended questions were conducted with 15 study participants. After an initial interview protocol was designed, three students were invited to participate in a pilot study and provide feedback on the interview. Their feedback was incorporated into the interview protocol.

The interviews took place at a northern European university where the study participants were enrolled as graduate students. Participants were asked to bring their laptop (i.e. their primary mobile computing device) to the interview. They were encouraged not to reorganize their digital documents stored in the laptop or elsewhere before the interview; however, for the sake of privacy they were encouraged to hide their private documents (such as financial and health-related documents). The laptop was used periodically throughout the interview to explore participants' document hierarchy, document organization, document types.

Open-ended questions during the interviews encouraged study participants to discuss their perspectives on organizing and re-finding digital documents, and to illustrate their perspectives on their laptop if they wish to do so. The interview protocol included three parts. The first part asked for demographic information such as gender and age. In the second part, participants were asked to describe their PIM practices, including re-finding practices. Questions included: Do you organize your digital documents? If so, how?

When you need to re-find a digital document, what characteristics do you usually remember about it? And how you usually go about re-finding documents? The third part investigated participants' perceptions regarding PIM tools. Examples of questions include: Do you use any PIM-specific tools? What functionality would you like a PIM tool to have?

The interviews lasted between 44 and 87 minutes, with the average length of 68 minutes. Recording devices were used to record the interview. A digital recorder captured audio, and a camera was used to take a picture of participants' main laptop screen, i.e., the desktop. During the interview, the investigator (the first author) took notes about participants' answers to the questions, and also wrote down details regarding behaviour of participants noticed by the investigator. After the interview, the audio recording from each interview was transcribed.

3.2 Study participants

Interviews were conducted with 15 study participants between the ages of 22 and 35 years, with an average age of 28 (see Table I). Four participants (27 per cent) were male, and 11 (73 per cent) were female. Their experience of using PCs and/or laptops ranged from eight to 20 years (Mean = 14; SD = 4.44). All participants were Master's students in a 12-month library and information science programme. The interviews were conducted during the second semester of their studies.

All participants were volunteers. At the beginning of the interview session they were informed about their rights as a study participant, and reviewed and signed a study consent form. The university's research ethics board reviewed and approved this study.

Currently enrolled master students were selected as the study population because this population uses computers routinely to manage, store, read and create digital documents and other files. Thus PIM has a very close relationship with their daily lives.

Gender	Age	Laptop operating system	Years using PCs and laptops	Number of documents on their laptop	Number of document folders
Male	32	Mac	18	224,800	[a]
Female	24	Win7	10	22,865	12,956
Female	24	Win7	20	6,874	537
Female	30	Win7	20	3,680	321
Female	25	Win7	20	1,287	53
Female	35	Win7	15	976	145
Female	35	Win7	15	970	89
Female	30	Vista	12	583	29
Female	23	Vista	13	536	81
Male	22	Win8	11	408	41
Male	28	Win7	10	289	38
Female	32	Win7	8	259	58
Female	23	Mac OSX	19	201	[a]
Male	26	Vista	10	[a]	[a]
Female	25	Win7	15	[a]	[a]
Average	28		14	20,286	1,304

Table I.
Study participants and their documents

Note: [a]These numbers are not available because the participant's documents were not organized such that they could distinguish their personal documents from system files. Thus we could not count the number of their documents or folders

Information and library science students were selected because information management and information retrieval are important subjects in this discipline. It is taught in courses and is often an important component of professional practice. Thus the study population was knowledgeable about information management and searching.

The number of documents on the participants' laptops ranged from 201 to 224,800 with an average of 20,286. These documents were often organized into folders. The number of folders ranged from 29 to 12,956 with an average of 1,304. All but two participants used laptops running a Microsoft Windows-based operating system. The two participants using Apple laptops had the lowest and highest number of documents on their laptops among all of the participants. During data analysis no discernible differences emerged between participants who used Apple and Microsoft operating systems.

The documents participants kept on their computers were primarily university-related. Most of the participants reported they use their laptops to store digital files. They reported using cloud storage to share files (such as use Google docs or Dropbox when working on group projects) or to create backup copies of files. Some also reported using USB keys and portable disk drives to share and copy files.

3.3 Data analysis

All interviews were transcribed, and Excel was used as a tool to keep track of codes and associated interview excerpts. The pictures of participants' desktop were analysed in conjunction with participants' responses regarding information organization practices, employing a data triangulation approach. After a first round of coding, one co-author reviewed the codes and associated data to improve coding reliability. Another two rounds of coding were conducted to help ensure coding completeness and accuracy.

Four category of codes emerged: the role and importance of PIM in participants' daily lives; characteristics of documents used to re-find documents; information organization – how participants deal with their unorganized documents; and, use of and perspectives on PIM tools. Coding examples are provided in Table II. The coded interview data were used to formulate the results and discussion below. Based on participants' requirements, final suggestions are given for the design of new PIM tools for improved information re-finding efficiency.

3.4 Limitations

There are several limitations which might influence the generalizability of this research. First, the number of study participants was relatively small. All participants were students studying the same subject at one university. A more diverse participant sample may yield additional insights. Second, most study participants were female and thus it is not possible to investigate whether there are gender differences with respect to PIM. Furthermore, due to space limitations, this paper only considers text documents in the context of desktop/laptop PIM. Cloud storage was mentioned several times during interviews, but as a means of sharing files rather than a storage medium; cloud storage is not the focus of this study.

An additional limitation stems from language and cultural differences. All interviews were conducted in English, the participants' first language. The first author, who conducted the interviews, has good command of English, but English is not her first language. While she had studied at the university where the interviews were conducted for nine months before the interviews were conducted, there may be subtle linguistic and cultural differences that arose during the interviews that were missed.

Quote	Code
It's important when I'm looking for information […] I'm not always thinking about it, but when I do think about it, it's nice that […] I know where to look for information I want	The role and importance of PIM in participants' daily lives
I have folders for everything. There's a few files there that don't fit into the folder structure, so I just keep them in my documents [list]. So [the file] "Christmas card" is a Christmas card that I designed and printed for Christmas this year. And then my CV, it doesn't necessarily fit into a particular folder […] hmm, I might move it [into a folder] […] but for the moment, it will stay just on the top of my documents [with] the things I don't really know how to organized	How participants deal with their unorganized documents
Usually it's the name of the actual document. Sometimes […] [for] pdf articles, it would be the name inside the actual article, cause sometimes they have really abbreviated file names […] for stuff I've typed in myself, I will remember the name of the document. [For] the stuff that I have downloaded, it's the title within the document I remember	Characteristics of documents used to re-find documents
I would like it if it was more organized, but I would like it if the organization was more automatic so I don't have to spend time moving the files from one place to another place to keep it organized. If I had a tool that automatically took all of my PDFs about "affordances", and it said, "Would you like me to put these in the affordance folder?" I would say, ok. I want it to be organized, but I want it to happen by itself	Use of and perspectives on PIM tools

Table II.
Coding examples

4. Results

Several major themes emerged from the data analysis. These themes include: the role of PIM in participants' daily lives; characteristics participants use (and do not use) to remember and re-find documents; persistence of unorganized documents, and perspectives of PIM tools, including non-use of specific PIM tools, expectations for PIM tools, and the adoption and use of new PIM tools.

4.1 Role of PIM in participants' daily lives

Participants were generally positive about the importance of PIM, especially with respect to their school work. As several participants reported:

> It's important when I'm looking for information.
>
> When I'm looking for […] school stuff, [it's] very important.
>
> Yes, [PIM is important] for school work and for college work, but in my personal life, no.

Most participants reported keeping only university-related documents on their laptop, and one-third of them had more than 1,000 documents on their laptop. Two-thirds of the participants said they believe they can remember most of the information they keep on their laptop computer. They emphasized that this does not mean that they could write a list to describe all the information they have; rather, it means that they could probably re-find something in a short period of time.

All participants said they organize or re-organize their documents with some regularity because this helps them subsequently to re-find documents. Ten participants (67 per cent) said they typically organize or re-organize their documents at least twice a year. Three participants (20 per cent) organized their documents immediately when

they got new documents and never do it again, and two participants (13 per cent) organized their documents once a week, but only for university work:

> I spend a lot of time organizing the folders on my computer and making sure that everything is where it should be.
>
> I'd like to keep my things organize […] even if the system doesn't necessarily make sense to someone else, it make sense to me, because then I can find it [later].

One-third of the participants said they cannot remember all the information they keep on their laptop, and they also asserted that this is perhaps because their documents are not very well organized:

> [I do] not [remember] everything, but all the important ones I remember. I don't know what they are called. I don't know where I put them, but I know they are there.
>
> No, [I don't remember everything on my laptop,] definitely not. Sometimes […] I could be looking for something, and some stuff comes up, and I go […] "God, what's that?" […] Sometimes stuff comes up […] that I don't remember […] at all.

Somewhat surprisingly the five participants who reported not remembering information did not have a large number of documents on their laptops. The number of documents for three of these participants ranged from 259 to 289, with an average of 261. (The other two participants did not separate their personal files from system files, and thus, their number of files could not be determined.) It appears there is not a close relationship between the number of documents that participants have and the extent of their memory about their documents. It may be that the act of organizing documents is in itself a memory aid, serving to re-enforce or enhance one's memory.

4.2 Document characteristics used to re-find them

During interviews participants were encouraged to talk freely about characteristics of documents that they use to remember and/or re-find documents. Follow-up questions were then asked regarding characteristics mentioned in the literature that the participants had not yet discussed.

As shown in Table II the most frequently used characteristics are: name, subject and storage location. Although most participants asserted that time was very important characteristic, but this is not the first characteristic they use for re-finding. The least frequently used characteristics reported are document type, creation place and recency. Characteristics not used at all, but mentioned in the literature, are author, frequency and weather. In Table II total scores indicating the overall ranking of each characteristic were calculated by assigning a score of three each time a participant mentioned a characteristic as being first in importance, two points when mentioned as being second in importance, and 1 point when mentioned as being third in importance. Each characteristic is described in detail in the following sections (Table III).

The characteristics can be classified primarily into two categories. One category is "use-centred" that includes the following characteristics: name, subject, storage location, creation time, user's location, recency, frequency and weather. All of these with the exception of weather are created and modified through a user's behaviour. The first four of these (name, subject, storage location and creation time) are the characteristics identified as most important by our study participants. The second category is "content-centred". This category includes document title, document file type, keyword and author. These characteristics are typically provided by others (with one exception, i.e., when the user is also the author) and remain unchanged across space and time.

Characteristic	Participants' ranking of characteristic First	Second	Third	Total score for characteristic	Overall ranking
Name	5	2	1	20	1
Subject	3	5	0	19	2
Storage location	4	2	2	18	3
Creation time	1	3	4	13	4
Keywords	1	1	0	5	5
Document title	1	0	2	5	5
Document file type	0	1	2	4	7
User's location	0	1	1	3	8
Recency	0	1	0	2	9
Author	0	0	0	0	
Frequency	0	0	0	0	
Weather	0	0	0	0	

Table III.
Importance of characteristics used to re-find documents

4.2.1 Name of documents. The name of documents has been mentioned as having an important role in PIM in previous research, including Carroll (1982), Barreau and Nardi (1995) and Blanc-Brude and Scapin (2007). Jones and Bruce (2005) consider re-naming documents as an encoding practice as did our study participants.

Participants were asked about their memory of document names. It should be noted that a document's name and title could be different, while sometimes they are the same. For example, when they downloaded an article, the name and the title could be the same; but sometimes our participants would change the name of the document to make it easier to remember. Frequently the system-generated name of a downloaded article was a string of numbers which does not make any sense. According to our participants' responses, they could remember the names of documents more easily than the titles:

> I would go by the name […] I'll completely re-name [the document], and call it something different.

> Stuff [i.e. names] I've typed out myself [is easier to] remember.

Other participants claimed that while they cannot usually remember the exact name immediately, they usually remember words in the name, or they can recognize the name when see it:

> I don't think of the name straight away, but when all of the names are in front of me, I can recognize it.

4.2.2 Subject (content of documents). Subject, or content, was put forward as an important characteristic already by our participants in the pilot study. It was also found to be an important characteristic in previous research (e.g. Goncalves and Jorge, 2004; Jones and Bruce, 2005).

Most of our participants reported that they are easily able to remember the subject of documents they need to re-find. However, they reported that their memory of keywords is usually not as good as their memory of subjects. This is consistent with the findings of Sachs (1967) who stated that people are good at memorizing the meaning of an event, but are not good at remembering the details. One participant explained:

> I always know the subject of the document […] I remember some guy wrote it; I can remember some things about what he says in the document. I can remember the subject. I can remember it's a pdf, but I don't know its name, and I don't know the title.

Although the subject of documents is important to individuals, desktop search tools today typically require individuals to input keywords that individuals do not typically remember.

4.2.3 Storage location (where you store documents). Storage location has been considered an important characteristic in document re-finding in many studies (e.g. Blanc-Brude and Scapin, 2007; Elsweiler *et al.*, 2007; Jones and Bruce, 2005). Barreau's research (1995) showed that people prefer to browse a document storage list rather than remember the exact document name.

Most of our participants reported they prefer to navigate through storage locations to re-find documents. That is, when they need to re-find a document, they go to the folder they believe most likely to contain the document and browse the list of documents in that folder. This result was also reported in Bergman *et al.*'s (2008). A participant explained:

> I usually remember the location. [For example, when] I know I need to find a copy of notes I took in a class, I know they will be on the notes file, and so I go to notes folder. I know which class they are from, so I look at the names. And then I know what date I'm looking for, so I look at the dates. And then that's how I find it, but the first thing I go to is the location.

There was only one participant who said she does not remember locations at all. She admitted that she is not good at organizing her documents.

4.2.4 Creation time (when you create or modify a document). Time is discussed in many PIM articles (e.g. Gemmell *et al.*, 2002, 2006; Dumais *et al.*, 2003; Goncalves and Jorge, 2004). Specific times, such as the time the information was first obtained or created (Jones and Bruce, 2005), and the time the information was last used or accessed (Blanc-Brude and Scapin, 2007) have been identified as playing a role when remembering and re-finding documents.

Our research results show that memory of time is quite strong. This does not mean that our participants can remember the exact time, but often, they remember the time period or time frame. This suggests that time is an important cue when they need to re-find a document. A participant explained:

> A date […] gives me a hint […] if I was looking for something in my downloads, for example, I would be able to find it by the date […] I would know. I downloaded it [a document] like six month ago, so I go right down [the list of downloaded files] to be able to find it.

There was one participant who stated that the memory of time would be helpful if it is an unusual time:

> If I […] [saved a document] either very early in the morning or very late at night […] I'd remember that, but if it's just say, 2 o'clock in the afternoon, I probably wouldn't remember that time, because it's not unusual to finish [and save a document] at 2 o'clock in the afternoon, but it is unusual to finish at 3 o'clock in the morning.

4.2.5 Keywords. Many PIM tools and desktop systems provide keyword search functions to help users re-find documents, e.g., Presto (Dourish *et al.*, 1999), Haystack (Quan *et al.*, 2003) and SIS (Dumais *et al.*, 2003). However, it is often difficult for individuals to remember keywords.

When asked about their memory of keywords, most of our participants said they could not remember keywords, or that keywords do not help them to re-find documents, because they do not use search tools to re-find documents. There were only

three participants who said they can remember keywords, and one of them pointed out that keywords were only useful sometimes:

> Keywords […] sometimes […] [are useful to me], especially if I am looking up articles on a certain topic […] I would remember that [keyword] sometimes, but then that would only be when I'm looking up articles.

4.2.6 Document title (title of papers, news, articles, etc. within a file.). Document title was mentioned in Blanc-Brude and Scapin's research (2007), and their results showed that among all document characteristics, the title has the minimum usage.

Similarly for most participants in our study, document title was difficult for them to remember. However, they reported they might recognize it when browsing a list of files and the file name contained all or part of a document's title:

> Usually I don't actually remember the exact […] title […] that wouldn't be the first thing I remember.

4.2.7 Document file type (word, rich text format, excel, pdf, etc.). Document file type, or format, was identified in Blanc-Brude and Scapin's research (2007) as an important characteristic for remembering documents.

During interviews some participants explained they can remember the document file type because it has a visual aspect. Different document types have different icons, different colours and different interfaces:

> Definitely […] that […] visual thing again – the icon is different, and also […] when you open a pdf, it comes up in a different way, so that is probably a key way that you to remember.

Most participants reported that they remember the file type, but they also commented this was not very helpful when they needed to re-find a document. This is because format does not differentiate their documents; most of their documents are either Word or pdf documents. As a result, they do not use file type as a cue to re-find a document unless it is a unique document type. A participant explained:

> I wouldn't use the type [to re-find a document] […] I haven't got that big a variety of file types. I only have Word [documents] […] I only have 1 or 2 Excel things, and I normally know where they are because they are so unusual.

Another participant mentioned she does not use Microsoft Word, so the characteristic, file type, was not meaningful to her:

> I don't use Word – I use open source, so I don't even pay attention to what kind [of file a document is].

4.2.8 User's location (physical location when a document was accessed, created or last modified). There were two participants who explained that they could remember where they were when they had accessed, created or last modified their documents. Other scholars (e.g. Jones and Bruce, 2005) have identified location as a contextual cue, along with ambient music and time of day. Location was also identified as important in Gonvalves and Jorge's study (2004).

Our results show that the memory of the location can help some participants remember and re-find documents:

> The physical location, I'll remember […] where I wrote it […] I was in this classroom, at this desk with these people […] It helps me remember where I put it actually in the computer as well.

However, most of our participants said their physical location is not a helpful characteristic, because 99 per cent of the time they were at the same place, either on campus or at home, when they created or modified a document.

4.2.9 Recency (how new the document is). Recency was one of the characteristics mentioned in Jones and Bruce (2005). It is generally considered that the newer the information is, the better people will remember it, so it is easier to re-find it. However, in our study only one participant said they would use recency to help re-find a document. Participants explained:

> This is probably my favourite thing to use to be able to find my documents, is my recent items, I love that […] but it does only work for things that you use in the past a week or two.

> It's probably easier to re-find things that are newer because they are refresher in my mind, but I think the structure is so logical that […] [recency] is not […] important.

4.2.10 Author (who created, wrote or modified the document). In the study by Goncalves and Jorge (2004) author was seldom used when describing documents. This result was mirrored in our study. Only one participant said they might remember the author, but it is not a characteristic that they would use when re-finding. Recall that our participants were graduate students who do research in their course of study, and read many books and articles, and write papers referencing those books and articles. Yet the names of authors are not used by this group to re-find documents.

4.2.11 Frequency (how often a document is accessed). Jones and Bruce (2005) considered frequency as an important factor in retrieval from human memory. Generally speaking, if something is used more often, then people might remember it better; however, no one in our study identified frequency as a useful characteristic. Thus frequency appears not to be very helpful when participants need to re-find a document. Furthermore existing computer systems do not have the function to show the most frequently accessed items when you search for something; they only show recently created or modified documents.

4.2.12 Weather (weather conditions when a document was created or modified). Weather as a characteristic has been mentioned as a useful characteristic in multi-media data management (O'Hare *et al.*, 2006, 2007; Elsweiler *et al.*, 2007). Our results indicate that weather is not helpful in re-finding a document. During the interviews participants never talked about weather as a characteristic spontaneously. When asked by the interviewer to consider weather as a characteristic, none of our participants said they would remember, or even take notice of, the weather.

4.3 Persistence of unorganized documents

During the interviews, participants, even those who consider themselves to be very well organized, mentioned that they always have some documents that are not organized, i.e., documents that cannot be organized or put into a folder. This occurs when a document does not "fit with" other documents. A participant explained:

> That [document] is just there, because there's nothing else to file with it yet.

Other participants reported they did not have time to organize uncategorized documents, or they simply forgot to do so. This is consistent with Al-Omar and Andrew's study (2013) in which scholars cited time pressure as one of the reasons that they did not organize their personal information collection. Our participants reported:

> I just don't have time or I just forget about it […] maybe I'm lazy.

> Probably [I'll organize these documents] if I have a lot of time […] it may never happen.
>
> Those haven't been organized yet […] but I will be doing that once I have time.

Some participants reported creating a folder called 'stuff' or 'miscellaneous' to put their unorganized things together in one storage location. They do this because they do not want to see all these documents everywhere, so they put them in a folder as a way to organize them:

> I don't know where to put that, so I just put it in a folder called miscellaneous.
>
> Well, these ones are in a folder, but it's just a bunch of random stuff, because it was all on the desktop, and I didn't want them to be on the desktop.

4.4 Perspectives on PIM tools

4.4.1 Non-use of PIM-specific tools. Most participants explained they do not use any specific PIM tools. They had either not heard about freely available PIM tools, such as BumpTop (Agarawala and Balakrishnan, 2006), or saw no reason to use the tools. They commented:

> I don't […] [know] what tools are there; I don't know what tools are available.
>
> Not really, apart from what the computer gives me [I don't use PIM tools.]

The participants who had heard about Google Desktop but had no inclination to use it:

> I've heard of it. I've never used it […] I don't know anything about it, other than the name, and that exists.
>
> I've read about that actually […] I don't feel like I need it […] I rarely search for things.

Only 20 per cent of the participants stated they used the search feature found in their laptop's operating system to re-find documents. One participant reported using the search feature first; the other two said they only used search first when they were looking for something specific or needed something immediately, but generally they preferred navigation. One person explained:

> If I'm in a rush, I use the search [feature] […] but I […] prefer browsing […] because like I said, the important things that I have, I know where they are and […] if I use that search thing I can often go quite wrong.

All other participants declared that they definitely prefer to use navigation to re-find their documents. They considered search as a last resort, or a backup solution. One participant even said s/he was not sure how to use the search function correctly:

> I'm never sure whether I use it correctly […] I'm not sure whether it has to be the whole title [that I enter], or whether I can just put in a word of the title and it will find it.

These results are similar to prior research results (e.g. Bergman *et al.*, 2008; Van Kleek *et al.*, 2011) indicating that people prefer navigation to search and use keyword searching infrequently.

4.4.2 Preferred PIM functionality. When discussing what functionality participants would prefer PIM tools to have, the most frequently mentioned feature was automatic information organization. Even semi-automatic organization would be welcomed because it takes such a long time to organize digital documents manually. Participants commented:

> When you upload them, that it would [ask] […] do you want these things to go into a folder by date or something […] [and then] do that automatically.

> I don't have good practices when I organize documents, I would like the program to do that for me […] I'd like it if it […] automatically made a lot of folders for me.

> If it saw that I was downloading from […] [the learning management system at my university], and […] would give me a list of […] my folders with documents downloaded from […] [the learning management system], I could just click on a folder and then it would be [stored] in [that folder].

Another important functionality they expect is a better backup capability. Only four participants (27 per cent) said they back up their documents periodically. Most participants do not do a full backup even though some of them have experienced severe information loss. They stated that they would like to do a full backup, but this is not easy to do with existing systems:

> If a new system […] could back up [documents] easier.

> Maybe having an extra back up […] sometimes having an online backup is really good thing to have.

Some participants remarked that they would prefer to have a better visualization or overall view of their document file structure. When they have created many nested levels of folders, i.e., folders within folders within folders, some participants said it is difficult for them to re-find a document. One participant explained:

> A better visualization of the hierarchy, the structure that I have […] sometimes it just get so big, and I would like a better visualization that doesn't get so cluttered or so long.

Several participants suggested the way folders are displayed should become more meaningful to people. E.g., in Microsoft systems all folders look the same, i.e., they have the same colour and shape, irrespective of the folders' content. Different colours could be used, mirroring the practice of using colour of physical file folders and/or labels on folders to differentiate content. Another suggestion was that the size of a folder could vary depending on its importance or volume of its content:

> [The colour of folders is] all […] the same. I think if you could colour-code things, [documents] might be easier to find […] a little bit more intuitive to use I think.

> It would be good if you could flag [the folder] […] maybe highlight them yourself.

One participant mentioned that a disadvantage of the folder search bar in the Microsoft Windows systems was its lack of a spell checker and thesaurus. When a wrong letter is typed in the search field, the search does not offer alternative words. Of course re-finding documents is possible only if the documents exist, because if a user types a keyword which does not exist in the document that s/he needed, then s/he would not be able to re-find the needed document. The participant explained:

> [The search bar in the system] doesn't let you make any mistakes. If it made a suggestion like Google did […] that would be ideal.

Recently Window systems added a feature that displays recent keywords used in searches. The interviews were conducted before this feature became available. We do not know if this feature adequately addresses the issue regarding misspelt keywords.

4.4.3 PIM tool adoption and use. Most participants stated that they would be interested in trying a new PIM tool. They commented:

> If I knew sort of what they did, what their features are, I would be interested in giving them a try.

> I suppose [I'd like to try new tools], it's never any harm to have a better familiarity with new systems.

All participants insisted that any new PIM tool should be very easy to learn. A participant explained:

> I don't have very much patience with them if they are hard to learn.

Participants were asked if a tool was more difficult to learn, but more efficient compared to an easier one, would they still want to give it a try. Only four participants (27 per cent) said they would. A participant explained:

> Taking a lot of time to learn it once is okay for me, as long as after that, it takes less time than it takes me now.

Participants mentioned in order to adopt and use a new PIM tool, the tool should provide relative advantages (Rogers, 2003), that is, provide advantages over their current way of re-finding documents. For example, it should be more effective, provide a more meaningful taxonomy (or organization of one's documents):

> If I can find something that's more effective, that takes less time [then I would use it.] The biggest thing for me is time.

> If it looked more effective or if the taxonomy made more sense to me than what I'm currently be using [then I would use it.]

Other participants reported they had no need currently for a new PIM tool but may have a need in the future if the number of their documents increased. For example participants reported:

> I would like to learn how to use it [i.e. a new PIM tool], but not until I have to use it […] if I had hundreds or thousands of files then I would probably change the way I organize [my files], and the kind of tools I use.

> When I get huge amounts of information that I'm struggling to manage [then I would use a new tool].

5. Conclusion

PIM has become more and more prominent in our daily lives. There have been multiple studies that explored the role of memory in PIM. Our results indicate there is a close relationship between human memory and information organization. There is no standardized method to organize personal documents as there is to organize books in libraries. Yet the largest challenge for individuals is information organization which impacts information re-finding in the future. The act of organizing information is itself a memory aid. Most participants know this, but they do not want to spend time organizing documents. They face a vicious circle; they do not want to organize documents initially when they are created, then their document collection grows. This growth makes it difficult for them to re-find information, but it will take even a longer time to organize all their documents, so they do not. Future research could explore and develop best information organization practices for digital PIM.

Although there have been multiple studies of PIM, there has only been small advances with respect to PIM tools. Most PIM tools developed in research contexts have not been successful enough to be made into commercial products and/or widely used. In this study, 15 graduate students were interviewed about their digital document management, most of them were positive about the importance of PIM in their daily life, but they had never used any PIM tools (other than those provided by their laptop operating system).

We suggest there are several reasons why these PIM tools are not widely used. One reason is the tools' incompatibility with individuals' habits. As Bergman *et al.* (2008) and our study report individuals prefer navigation to search when they need to re-find information in their PSI. For those individuals who do not use search for re-finding, PIM tools that only offer better search features will not provide sufficient advantages to them to merit their adoption and use.

We propose a second reason is that PIM tools do not take full advantage of, or effectively augment, the capabilities found in human memory. Human memory is very important in PIM, especially in re-finding. What kinds of characteristics can an individual remember about the information s/he need to re-find? How do users make use of these characteristics to help themselves become more efficient in their information re-finding?

Our results suggest incorporating additional use-centred characteristics of documents in PIM tools could make document re-finding easier. In our study these are name, subject and storage location and creation time of digital documents. Keywords are used infrequently to re-find documents because they are difficult to remember. Characteristics such as author, weather or frequency were not retained or used at all to re-find documents.

The results verify that information visualization can be useful in PIM; however, participants expressed preference for simplistic visualizations they can control or customize. An example is allowing individuals to use colour as a way to highlight folders or documents. It could also be useful to allow individuals to specify their own organization using individual or standard cataloguing schemes, in addition to the standard name, date modified, file type and size sort order widely available today.

Another tool feature to consider is intelligent and/or automatic or semi-automatic support for information organization. As the results show, most people do not like to spend time in organizing their digital documents, and they would welcome assistance in organizing documents. Because the act of organizing documents aids in re-finding them, a goal is for PIM tools to augment the organization process such that the act of organizing documents provides additional benefits with respect to re-finding documents than it does today. Furthermore tools should consider providing standardized or cross-platform ways for individuals to organize their information, so they can use the same practices in each system, and easily transfer annotations, tags or notes associated with documents when transferring those documents between different systems.

There are many challenges in everyday PIM. This study contributes to the discussion through analysing study participants' perspectives on PIM, everyday PIM practices and PIM tools. Future work should investigate PIM practices with respect to: different systems, including tablets, mobile phones, social networking services and cloud services; different types of digital artefacts, such as audio, photos and videos; and different contexts of use.

Note

1. "At its centre, a person's PSI includes all the items that are, at least nominally, under that person's control. At its periphery, the PSI includes information that the person might like to know about and control but that is under the control of others" (Jones and Teevan, 2007, pp. 10-11).

References

Agarawala, A. and Balakrishnan, R. (2006), "Keepin'it real: pushing the desktop metaphor with physics, piles and the pen", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *ACM*, *New York, NY*, pp. 1283-1292.

Aizawa, K., Tancharoen, D., Kawasaki, S. and Yamasaki, T. (2004), "Efficient retrieval of life log based on context and content", *Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, *ACM*, *New York, NY*, pp. 22-31.

Al-Omar, M. and Andrew, C. (2013), "Finders, keepers, losers, seekers: a study of academics' research-related personal information collections", in Yamamoto, S. (Ed.), *Human Interface and the Management of Information. Information and Interaction Design*, Springer, Berlin Heidelberg, pp. 169-176.

Barreau, D. and Nardi, B.A. (1995), "Finding and reminding: file organization from the desktop", *ACM SIGCHI Bulletin*, Vol. 27 No. 3, pp. 39-43.

Barreau, D.K. (1995), "Context as a factor in personal information management systems", *Journal of the American Society for Information Science*, Vol. 46 No. 5, pp. 327-339.

Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N. and Whittaker, S. (2008), "Improved search engines and navigation preference in personal information management", *ACM Transactions on Information Systems (TOIS)*, Vol. 26 No. 4, pp. 20:1-20:24.

Blanc-Brude, T. and Scapin, D.L. (2007), "What do people recall about their documents? Implications for desktop search tools", *Proceedings of the 12th International Conference on Intelligent User Interfaces*, *ACM*, *New York, NY*, pp. 102-111.

Boardman, R. and Sasse, M.A. (2004), "Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *ACM*, *New York, NY*, pp. 583-590.

Bruce, H., Jones, W. and Dumais, S. (2004), "Information behaviour that keeps found things found", *Information Research*, Vol. 10 No. 1, available at: www.informationr.net/ir/10-1/paper207 (accessed 20 October 2012).

Carroll, J.M. (1982), "Creative names for personal files in an interactive computing environment", *International Journal of Man-Machine Studies*, Vol. 16 No. 4, pp. 405-438.

Chau, D.H., Myers, B. and Faulring, A. (2008), "Feldspar: a system for finding information by association", *Proceedings of Personal Information Management, Workshop at CHI*, *Florence, April*, available at: www-2.cs.cmu.edu/~faulring/papers/feldspar-chi08-pim.pdf (accessed 8 December 2012).

Chirita, P.A., Gavriloaie, R., Ghita, S., Nejdl, W. and Paiu, R. (2005), "Activity based metadata for semantic desktop search", in Aroyo, L. *et al.* (Ed.), *The Semantic Web: Research and Applications*, Springer, Berlin, pp. 439-454.

Corbató, F.J., Merwin-Daggett, M. and Daley, R.C. (1962), "An experimental time-sharing system", in Barnard, G.A. (Ed.), *Proceedings of the May 1-3, 1962, Spring Joint Computer Conference*, *ACM*, New York, NY, pp. 335-344.

Czerwinski, M. and Horvitz, E. (2002), "An investigation of memory for daily computing events", in Faulkner, X., Finlay, J. and Détienne, F. (Eds), *People and Computers XVI-Memorable Yet Invisible*, Springer, London, pp. 229-245.

Dourish, P., Edwards, W.K., LaMarca, A. and Salisbury, M. (1999), "Presto: an experimental architecture for fluid interactive document spaces", *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 6 No. 2, pp. 133-161.

Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. and Robbins, D.C. (2003), "Stuff I've seen: a system for personal information retrieval and re-use", *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *ACM*, *New York, NY*, pp. 72-79.

Elsweiler, D., Harvey, M. and Hacker, M. (2011), "Understanding re-finding behavior in naturalistic email interaction logs", *Proceedings of the 34th international ACM SIGIR Conference on Research and development in Information Retrieval, ACM, New York, NY*, pp. 35-44.

Elsweiler, D., Ruthven, I. and Jones, C. (2007), "Towards memory supporting personal information management tools", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 7, pp. 924-946.

Flanagan, J.C. (1954), "The critical incident technique", *Psychological Bulletin*, Vol. 51 No. 4, p. 327.

Fuller, M., Kelly, L. and Jones, G.J. (2008), "Applying contextual memory cues for retrieval from personal information archives", *Proceedings of Personal Information Management, Workshop at CHI, Florence, April*, available at: http://doras.dcu.ie/15913/ (accessed 15 December 2012).

Gemmell, J., Bell, G. and Lueder, R. (2006), "MyLifeBits: a personal database for everything", *Communications of the ACM*, Vol. 49 No. 1, pp. 88-95.

Gemmell, J., Bell, G., Lueder, R., Drucker, S. and Wong, C. (2002), "MyLifeBits: fulfilling the memex vision", *Proceedings of the Tenth ACM international conference on Multimedia, ACM, New York, NY*, pp. 235-238.

Gonçalves, D. and Jorge, J.A. (2004), "Describing documents: what can users tell us?", *Proceedings of the 9th International Conference on Intelligent User Interfaces, ACM, New York, NY*, pp. 247-249.

Jones, W. (2011), "XooML: XML in support of many tools working on a single organization of personal information", *Proceedings of the 2011 iConference, ACM, New York, NY*, pp. 478-488.

Jones, W. and Anderson, K.M. (2011), "Many views, many modes, many tools … one structure: towards a non-disruptive integration of personal information", *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia, ACM, New York, NY*, pp. 113-122.

Jones, W. and Teevan, J. (2007), "Introduction", in Jones, W. and Teevan, J. (Eds), *Personal Information Management*, The University of Washington Press, Washington, DC, pp. 10-11.

Jones, W. and Bruce, H. (2005), "A report on the nsf-sponsored workshop on personal information management", Report on the NSY PIM Workshop, Seattle, WA, pp. 27-29.

Jones, W., Munat, C.F., Bruce, H. and Foxley, A. (2005), "The universal labeler: plan the project and let your information follow", *Proceedings of the American Society for Information Science and Technology*, Vol. 42 No. 1, available at: http://onlinelibrary.wiley.com/doi/10.1002/meet.14504201185/full (accessed 17 November 2012).

Karger, D.R. (2007), "Unify everything: it's all the same to me", in Jones, W. and Teevan, J. (Eds), *Personal Information Management*, The University of Washington Press, Washington, DC, pp. 96-99.

Madlock-Brown, C., Chin, S.C. and Eichmann, D. (2012), "InfoATV: enhancing PIM systems to support re-finding to (re) discovery", *Proceedings of the Fifth International Workshop on Personal Information Management (PIM 2012)*, available at: http://myweb.uiowa.edu/sichin/pdf/pim2012_infoATV.pdf (accessed 17 December 2012).

Murthy, S., Murthy, U. and Fox, E.A. (2006), "Using superimposed and context information to find and re-find sub-documents", *Proceedings of the SIGIR 2006 Workshop on Personal Information Management, PIM, Seattle, WA*, pp. 10-11.

O'Hare, N., Gurrin, C., Jones, G.J., Lee, H., O'Connor, N.E. and Smeaton, A.F. (2007), "Using text search for personal photo collections with the MediAssist system", *Proceedings of the 2007 ACM Symposium on Applied Computing, ACM, New York, NY*, pp. 880-881.

O'Hare, N., Lee, H., Cooray, S., Gurrin, C., Jones, G.J., Malobabic, J. and Uscilowski, B. (2006), "MediAssist: using content-based analysis and context to manage personal photo collections", in Lew, S.M., Sebe, N. and Eakins, P.J. (Eds), *Image and Video Retrieval*, Springer, Berlin Heidelberg, pp. 529-532.

Quan, D., Huynh, D. and Karger, D.R. (2003), "Haystack: a platform for authoring end user semantic web applications", in Tangmunarunkit, H., Decker, S. and Kessekman, C. (Eds), *The Semantic Web - ISWC 2003*, Springer, Berlin Heidelberg, pp. 738-753.

Reicher, G.M. (1969), "Perceptual recognition as a function of meaningfulness of stimulus material", *Journal of Experimental Psychology*, Vol. 81 No. 2, pp. 275-280.

Rogers, E. (2003), *Diffusions of Innovations*, 5th ed., Free Press, New York, NY.

Rubin, D.C. and Wenzel, A.E. (1996), "One hundred years of forgetting: a quantitative description of retention", *Psychological Review*, Vol. 103 No. 4, pp. 734.

Sachs, J.S. (1967), "Recognition memory for syntactic and semantic aspects of connected discourse", *Perception & Psychophysics*, Vol. 2 No. 9, pp. 437-442.

Tan, D., Berry, E., Czerwinski, M., Bell, G., Gemmell, J., Hodges, S., Kapur, N., Meyers, B., Oliver, N., Robertson, G. and Wook, K. (2007), "Save everything: supporting human memory with a personal digital lifetime store", in Jones, W. and Teevan, J. (Eds), *Personal Information Management*, The University of Washington Press, Washington, DC, pp. 96-99.

Van Kleek, M., Styke, W. and Karger, D. (2011), "Finders/keepers: a longitudinal study of people managing information scraps in a micro-note tool", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY*, pp. 169-176.

Xiao, H., and Cruz, I.F. (2005), "A multi-ontology approach for personal information management", *Proceedings of the 4th International Semantic Web Conference, ISWC, Semantic Desktop Workshop, Galway*, pp. 1-15.

Further reading

Elsweiler, D., Ruthven, I. and Ma, L. (2006), "Considering human memory in PIM", *Proceedings of the SIGIR 2006 Workshop on Personal Information Management, Seattle, WA*, pp. 10-11.

About the authors

Dr Xiao Xie is a Lecturer in the School of Management, Anhui University, China. She was an exchange PhD student at School of Information & Library Studies at the University College Dublin, Ireland from September 2012 to September 2013. Her research focuses on personal information management (PIM), including the relationships between human memory and PIM, how people use digital devices/software to manage their personal information, and the factors influencing PIM tool adoption and use. Dr Xiao Xie is the corresponding author and can be contacted at: 429228948@qq.com

Diane H. Sonnenwald is a Professor of Information Science at the Royal School of Library and Information Science, University of Copenhagen, Denmark and an Adjunct Professor of computer science at the University of North Carolina at Chapel Hill, USA. Her research focuses on socio-technical design of future technologies and information behaviour in a variety of contexts, including the digital humanities, emergency healthcare, academia and industry. She is currently editing a book on theory development in the information sciences.

Dr Crystal Fulton is a Senior Lecturer at the School of Information and Library Studies, University College Dublin. She studies the information behaviour of people in various contexts including online and offline environments, and in particular, her work explores how people create, use, and share information and the impact of this interaction with information on social participation and lifelong learning.