

Understanding the Evolution of Users' Personal Information Management Practices

Manas Tungare
manas@vt.edu

Dept. of Computer Science and Center for Human Computer Interaction
Virginia Tech
Blacksburg, VA, USA.

Dissertation Advisor: Manuel Pérez-Quñones, D.Sc.
Research Area: Personal Information Management
URL: <http://manas.tungare.name/>

Abstract. Information is being disseminated much faster than we can assimilate it, leading to information overload. In addition to desktop computers, users use a vast array of other devices to manage their information, which leads to information fragmentation. It has not yet been studied how users adapt their information management practices in response to the introduction of new devices into their personal information ecosystem. As part of my doctoral research, I plan to study this evolution, which is important for the design of next-generation devices and to establish future research directions in personal information management.

1 Introduction

One of the biggest challenges of our time is to create mechanisms to effectively consolidate and control the management of personal information. We have developed amazing capabilities to record, store, and transmit massive quantities of information with minimal effort; however this has relegated us to part-time workers and part-time librarians of our own personal information. In spite of the ease of recording, creating, receiving, storing, and accumulating digital materials, managing and using them sensibly is difficult [1]. With time, the amount of information generated by humans can only increase, while human attentional resources have remained constant [2], thus presenting an alarming future for information workers whose job requires them to stay informed. *Information overload* is defined as occurring when the information processing demands on an individual's time to perform interactions and internal calculations exceed the supply or capacity of time available for such processing [3]. Even worse than information overload is *information fragmentation*, the condition of having a user's data in different formats, distributed across multiple locations, manipulated by different applications, and residing in a generally disconnected manner (e.g. [4]).

To study the problems caused by information overload, information fragmentation and multiple devices, we need appropriate terminology, concepts, and principles. As reviewed in my prior work [5], I have characterized this collection of devices as a *personal information ecosystem* and drawn parallels with biological systems as a way of

understanding the devices that participate in the ecosystem, the relationships among them, the type of user activities that the ecosystem supports, and the equilibrium that must be maintained. We use this ecosystem as a framework with which to study the problem in more detail. Based on our previous work, we have observed that personal information management (PIM) practices evolve rather slowly, much like evolution in biological ecosystems.

As part of my doctoral research, I seek to study how personal information management strategies evolve over an extended period of time and the impact of changes in life patterns on these practices. In one set of studies, I plan to explore how the availability of information management devices (computer, PDA, MP3 player, etc.) affects information management practices: in particular, how new information management practices develop to make use of the newly introduced devices. The second set of studies will examine how changes in work environment influence personal information management practices. Overall, I am interested in exploring the internal and external factors that change the equilibrium of the personal information ecosystem.

2 The Problem

Personal Information [6] includes (1) the information a person keeps for personal use (e.g. files), (2) information about a person but kept by and under the control of others (e.g. health information), (3) information experienced by a person but outside the person's control (e.g. books or Web sites browsed), and (4) information directed to a person (e.g. email). The unique situational aspect of the working environment [7] makes it difficult to study PIM as compared to general information storage and retrieval (ISAR) systems [7].

2.1 Previous Studies in PIM

Many of the studies that have been conducted in the area of Personal Information Management (details in [8]) are limited to how we manage information on a particular device (e.g. desktop), or how we manage a particular type of information (e.g. bookmarks or emails). They have looked at a single snapshot of organizational behavior, not a continuous or long-term inspection of user behavior. The *evolution* of a user's information practices as she progresses from one work context to another, or from one role to another, has not been studied in detail.

2.2 PIM with Multiple Devices

One of the major causes of information fragmentation is that we no longer are restricted to a single device, or a single source of information; most of our information is scattered across multiple devices, such as desktop computers at the office, laptops at home, portable digital assistants (PDAs) on the road, and of course, cellphones. We are not aware of any study so far that has explicitly considered the presence of multiple devices in a user's personal information ecosystem, or examined how they affect a user's information management practices.

3 Research to Date

We reviewed the information fragmentation problem in [9], including its many visible symptoms, primary causes, and philosophical underpinnings. Understanding the need for users to be able to access all of their data, across devices, across collections, and with minimal fragmentation, we designed and developed the Syncables framework [10], which lets applications migrate their task information seamlessly across multiple platforms. Information need not be saved to a file first, and the framework provides a unified addressing scheme for each unit of data stored (e.g. a calendar event, a person's contact information, a file, or a piece of context data such as the current cursor position inside a word processor). Applications can thus directly address data created or manipulated in other applications, and create deep links between data items. These are similar in concept to the information trails proposed by Vannevar Bush in his seminal 1945 paper [11].

We asked users about their information management practices related to different collections, (e.g. files, music, research papers) [12]. Based on the varied responses, it was clear that in spite of useful tools that address the management of information and methods recommended by experts, PIM is still highly individualized. In [5], we discussed some of the design issues in PIM systems for multiple devices and proposed a characterization of a user's many devices in terms of a biological ecosystem.

4 Proposed Research Design

I plan to study how personal information management strategies evolve over an extended period of time and the impact that changes in life patterns have in these practices. Overall, we are interested in exploring which internal and external factors change the equilibrium of the personal information ecosystem. The ecosystem evolves with new processes, new interdependencies between components, and new flow amongst components.

4.1 Exploring Internal/External Factors

We hypothesize that the distinction between filer/piler behavior and several other PIM practices is one that might be initially influenced by personal traits of the users. However, we have observed how users' practices change over time and how some practices are different for the different types of media [12]. For the studies proposed, we plan to:

- conduct basic cognitive tests on the participants (field dependent/independent, assessed using the embedded figure test [13], and a memory test [14]).
- conduct a survey to assess the participants' information management practices. The survey is intended to determine which of the already established 'bins' of PIM practices (e.g. filer, piler, spring cleaner, etc.) do users classify themselves into. (like in [15]).
- conduct the same survey at the end of the subject participation, to assess how their information management practices changed over time.

We want to learn if there is correlation between internal factors (personal traits) and PIM practices. Furthermore, by repeating the survey at the end of the study, we will identify if the practices changed over the time of the study. The studies themselves (see following sections) will provide insight as to what might prompt the change.

4.2 Natural Transitions in Life: Work Environment Change

The effect that a change in work environment has on information management practices has not been studied before. We plan to track several participants as they change work environments to try to answer the following research questions: how do the new information flows at the new work environment influence or prompt changes in information management practices? How is equilibrium maintained and do new practices help the ecosystem reach equilibrium? How much resistance does a user have to changes? What is the relationship between the participant's internal factors and the external factors as it relates to these changes?

4.3 Understanding Long-Term Changes in Users' Workflows

Based on our initial observations [12], and on several findings from the literature, we know that introduction of new devices lead to changes in PIM strategies. What we do not know is how these changes take place. We believe that the changes follow a generic strategy: users first acknowledge they have a problem, then explore different strategies and tools, trying to find a balance between the use of new tools while reducing the change in existing practices (i.e., trying to maintain the equilibrium of their ecosystem). This is followed by trying out an approach, that might not be the one they settle with; there is a period of transition where a hybrid of both, old and new systems is used. Finally, there is a successful deployment of the new strategy, resulting in PIM practices that are either more effective than before, or as good as before but without the failures.

With this study, we hope to address the following research questions: how do the new devices affect what users do? What interrelationships develop between devices? How are pairs of devices are used together (symbiosis)? What types of failures are caused by the use of new technology? How long does it take to adopt new workflows (i.e., ecosystem reach equilibrium)? How is information fragmented over the different devices? Does this cause problems for the participant? How does the participant cope with this fragmentation?

4.4 Participants

We will recruit students transitioning from undergraduate to graduate school, MS students transitioning from graduate school to industry, PhD students transitioning from graduate school to a professional career, and professionals joining a company in the Corporate Research Center at Virginia Tech. We will recruit 2 for each of the above categories, for a total of 8 participants. In the case of the students, we will contact them before the change in job environment to conduct an initial data gathering and interview. For the professionals, we plan to contact them through their companies before they move to the area.

We plan to supply devices to a user (laptop, PDA, MP3 player) and capture data about the user's interaction with them to study the management of different collections: music, video, files, photos, etc. Interviews with individual participants on a weekly basis and monthly focus groups will help understand the users' stories better.

5 Goals and Key Contributions

I expect that these proposed studies will contribute significantly to the state-of-the-art in PIM research and in the development of tools for PIM tasks. In particular,

- the findings from our experiments will help us understand the relationships among various internal and external factors that influence an individual's PIM strategies,
- understanding these factors will help enrich the Personal Information Ecosystems model to guide further studies and experiments in PIM,
- the enriched model will be useful to designers of next-generation PIM systems to create tools that fit well into users' personal information ecosystems, and
- these studies will help companies and universities make natural life transitions easier for new members (students or employees) from a personal information management point of view.

6 Acknowledgments

I would like to thank my advisor, Dr. Manuel Pérez-Quñones, for insightful discussions and his valuable guidance at each step of the way. Prof. Steve Harrison and Pardha S. Pyla participated in many brainstorming sessions which directly or indirectly influenced the current research.

References

1. Czerwinski, M., Gage, D.W., Gemmell, J., Marshall, C.C., Pérez-Quñones, M.A., Skeels, M.M., Catarci, T.: Digital memories in an era of ubiquitous computing and abundant storage. *Commun. ACM* **49**(1) (2006) 44–50
2. Levy, D.M.: To grow in wisdom: Vannevar Bush, Information Overload, and the Life of Leisure. In: *JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*, New York, NY, USA, ACM Press (2005) 281–286
3. Schick, A.G., Gordon, L.A., Haka, S.: Information overload: A temporal approach. *Accounting, Organizations and Society* **15**(3) (1990) 199–220
4. Bergman, O., Beyth-Marom, R., Nachmias, R.: The project fragmentation problem in personal information management. In: *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, New York, NY, USA, ACM Press (2006) 271–274
5. Tungare, M., Pyla, P.S., Pérez-Quñones, M., Harrison, S.: Personal information ecosystems and implications for design. Technical Report cs/0612081, ACM Computing Research Repository (2006)
6. Jones, W., Teevan, J.: *Personal Information Management*. University of Washington Press, Seattle, Washington (May 2007)

7. Barreau, D.K.: Context as a factor in personal information management systems. *J. Am. Soc. Inf. Sci.* **46**(5) (1995) 327–339
8. Teevan, J., Capra, R., Pérez-Quiñones, M.: 3. In: *How people find information*. University of Washington Press, Seattle, Washington (May 2007) 17
9. Tungare, M., Pyla, P.S., Sampat, M., Perez-Quinones, M.: Defragmenting information using the syncables framework. In: *Proceedings of the 2nd Invitational Workshop on Personal Information Management at SIGIR 2006*. (2006)
10. Tungare, M., Pyla, P.S., Sampat, M., Pérez-Quiñones, M.: Syncables: A framework to support seamless data migration across multiple platforms. In: *IEEE International Conference on Portable Information Devices (IEEE Portable)*. (2007)
11. Bush, V.: *As we may think*. The Atlantic Monthly (1945)
12. Pérez-Quiñones, M., Tungare, M., Pyla, P.S., Kurdziolek, M.: A special topics course on personal information management. Technical Report TR 06-26, Virginia Tech Dept. of Computer Science, Blacksburg, VA (2006)
13. Witkin, M., Oltman, P., Raskin, E., Karp, S.: *Group Embedded Figures Test Manual*. Mind Garden, Inc., Redwood City, CA (1971)
14. Crawford, J.R., Smith, G., Maylor, E.A., Sala, S.D., Logie, R.H.: The prospective and retrospective memory questionnaire (prmq): Normative data and latent structure in a large non-clinical sample. *Memory* **11**(3) (2003) 261–275
15. Malone, T.W.: How do people organize their desks?: Implications for the design of office information systems. *ACM Trans. Inf. Syst.* **1**(1) (1983) 99–112