

Personal Information Management (PIM) 2008

Jaime Teevan
Microsoft Research
teevan@microsoft.com

William Jones
University of Washington
williamj@u.washington.edu

Robert Capra
University of North Carolina
rcapra3@unc.edu

Abstract

This paper reports on the Personal Information Management (PIM) 2008 workshop held at the CHI 2008 conference. The workshop addressed many challenges that must be overcome for personal information management to become a natural and effortless part of people's everyday information experience. The topics explored included how to build good PIM tools, how to develop sound methodologies so that PIM can be studied and understood, how to account for the influence of other people on an individual's PIM practices, and how to understand the affect of widespread mobile information access on PIM.

1 Introduction

Personal information management (PIM) is the study and practice of the activities that people perform to acquire, organize, maintain, and retrieve information for everyday use. PIM is a growing area of interest as we strive for better use of our limited personal resources of time, money, and energy, as well as greater workplace efficiency and productivity. In an ideal world, we would always have the right information, in the right form, with the right context, right when we needed it. The purpose of the PIM 2008 workshop was to explore how people in the real world manage the massive amounts of information they encounter daily, and discuss how tools can bring real-world information interactions closer to the ideal.

Researchers exploring PIM-related topics participate in a large set of disparate conferences, including those related to information retrieval, human computer interaction, and information science. The PIM 2008 workshop summarized in this paper was held at the ACM SIGCHI conference in Florence, Italy, as part of a series of meetings organized every year and a half to bring these communities together. Prior workshops have occurred at the University of Washington in 2005 (an NSF sponsored workshop) and at SIGIR in 2006. The next workshop will take place during ASIS&T 2009. These meetings provide an opportunity for PIM researchers to share research findings on a regular basis and for a broad research community to engage in the study of personal information management.

2 Day 1: Understanding PIM

The first day of the workshop focused on understanding what activities and strategies people use to manage information in an information-saturated world.

2.1 Understanding PIM

The workshop began with panel presentations exploring how people perform PIM activities, moderated by Deborah Barreau. Rob Capra, in a presentation titled *Studying Elapsed Time and Task Factors in Re-Finding*, talked about characterizing re-finding tasks by information need. The process of re-finding can vary greatly depending upon task characteristics such as whether the target is exact (e.g., “the web site for Joe’s Pizza”) or fungible (e.g. “any web site for Joe’s Pizza”). He also discussed the importance of the amount of time elapsed between initially finding an item and trying to re-find it. Tim Coughlan’s presentation titled *PIM for Creative Practitioners* explored the role of creativity in PIM. He noted that creativity is difficult to schedule or plan, and that creative thoughts often occur when people are mobile or engaged in other tasks. This creates challenges for the expression and capture of creative ideas. Melanie Kellar (*An Overview of Web-Based Monitoring*) described research into Web content monitoring as an increasingly prevalent web activity. People often repeatedly visit web sites, for example, to stay tuned concerning important changes of status (e.g., in flight arrival/departure times) or for distraction and recreation (e.g., Facebook or Twitter).

A common theme that emerged in all of the presentations was the importance of considering people’s personal information management practices in the greater context of their lives, as people spend most of their time away from conventional desktop or laptop computers. This theme was echoed in the discussion, which focused on the capture and use of personal information beyond the desktop. Participants explored the trade-offs between the clutter that can be created by capturing information and the benefits that can come from it.

Two examples discussed were handwritten notes and the unanticipated value of captured information. Although nearly all workshop participants took notes during the workshop, the majority of the notes were made on paper, and most participants never expected to actually return to their notes. Sense-making was cited as a reason for why people may take notes they expect not to re-use. In this case, the process of capturing the information is more valuable than the captured information.

Participants also discussed how information captured for one purpose can provide unanticipated value for other purposes. William Jones related a story a colleague told him about an outing to a zoo. His colleague’s daughter arrived at the zoo with a jacket and lost it over the course of the day. Because her father had been taking pictures throughout the trip, he was able to go back through the photos to determine where his daughter last had the jacket and retrieve it. The unintentional capture of superfluous information provided value to the family.

2.2 How PIM is Personal

The second session of the workshop consisted of short presentations and small group discussion looking at the “personal” aspect of “personal information management.” In a presentation titled *How is Information Personal?*, William Jones discussed several meanings for “personal information,” including: 1) information controlled or owned by us, 2) information about us, 3) information directed towards us, 4) information sent to others by us, 5) information experienced by us, and 6) information relevant to our needs. He argued that the different senses can form a yardstick by which to evaluate tools, organizational schemes, and overall PIM strategies.

Unique aspects of dealing with personal information were discussed further in several breakout groups. Deborah Barreau led a group that explored what makes the fields of information management and personal information management different. The group concluded that PIM is unique because its goals are different, with the study of information management typically restricted to the collection and distribution of information. William Jones ran a breakout group exploring who we trust with our

personal information. They explored a number of schemes for storing personal information, including on the desktop, in the cloud, or on a local storage device. The group discussed one possible path of gradual change from the desktop to the cloud: Storage space on the Web could initially provide backup functionality for a local storage device, but become the primary storage location as services and supporting tools develop.

In addition to diving deep into the nuances of “personal” information, workshop participants also considered how the experience of information management can be made “personal” through personalization. Andreas Rath’s talk on Context Aware Knowledge Services discussed a personalization system, DYONIPOS, which attempts to automatically identify a user’s work task and contextualize different types of knowledge services accordingly. A break out group lead by Paul Andre on the topic discussed the meaning of “personalization” and how it differs from “automation”.

Special focus was given during the workshop to the ongoing revolution in mobile and Web-based tools for information management. Manas Tungare led a breakout group exploring how the management of personal information is different off the desktop. The group discussed how many people now rely primarily on mobile computers for their personal information management. Other people do not rely on a single device, but rather access, organize, and manage their personal information through any device that provides access to the Web.

2.3 Research and Design Methodologies

The third and final session of the first day discussed research and design methodologies for studying PIM. Olha Bondarenko talked about methodologies for understanding PIM in a presentation titled *How Do People “Read” Their Desk*. She summarized the use of a triad elicitation method (part of a *repertory grid* methodology), and then led a breakout group on the topic. The group discussed the importance of studying diverse sets of people in real-world situations, and debated the value of understanding current personal information management practices for building future tools, since the lessons any learned are inherently biased by existing tools.

Ofer Bergman talked about design methodologies for PIM in his talk entitled *The User Subjective Approach*. He described the application of the user subjective approach, which is one of the few design methodologies developed specifically for designing personal information management tools, to the creation of several prototypes. In the following breakout group on design methodologies, the group questioned whether PIM system design should focus on applications or on the personal information items being managed.

There were two talks in this session on methodologies for evaluating PIM tools, *Now It’s Personal! Evaluating PIM Retrieval Tools* by Daniel Goncalves, and *Evaluating PIM Using an Activity Logs Enriched Desktop Dataset* by Sergey Chernov. Goncalves discussed the limitations of traditional methodologies of tool evaluation as applied to evaluation of PIM tools. He then discussed a methodology used in the development of a personal document retrieval tool and stressed the importance of time as a contextual element in the retrieval process (e.g. “the document I wrote around Christmas of last year”). Chernov’s talk explored the use of activity logs and enhanced desktop datasets. Chernov argued that while methods evaluating retrieval effectiveness are well-developed for the public stores of information, evaluations involving desktop retrieval tools are neither repeatable nor comparable. He proposed that desktop datasets, enhanced with logs of actual use by an individual over a multi-month period, could provide a basis for the evaluation and comparison of desktop retrieval tools. A breakout group discussing evaluation methodologies further explored issues of evaluation, including what measures should be used for PIM evaluation, who should be studied, and how. Workshop participants

as a whole expressed an interest in creating test beds that could be used across research groups to support the comparison of results, but acknowledged creating these test beds is challenging. One first step in the right direction may be the creation of common tasks for study, as is done for TREC tracks.

3 Day 2: Supporting PIM

Focus in Day 2 of the workshop shifted from understanding how people manage their personal information to exploring promising solutions.

3.1 PIM Tools: Finding v. Organizing

The day began with the presentation of two papers about tools for finding personal information. Duen Chau presented *Feldspar: A System for Finding Information by Association*. Feldspar lets users create queries by association. For example, a person may remember that a webpage was mentioned in an email from a particular person, even in the absence of remembering particular keywords about the page. This paper led to discussion of how personal information search systems can support natural ways of expressing their information target, with some people suggesting that the process is iterative. Liadh Kelly continued the discussion of finding and re-finding in her talk *Applying Contextual Memory Cues for Retrieval from Personal Information Archives*. She conducted a study capturing computer activity, events, and locations for one participant for six weeks, and found evidence that using content and context together helped in relocating information.

Following these presentations were two talks about tools to support personal information organization. In a talk titled *Ontology Creation and Visualization for a PIM System*, Alan Dix suggested using personal ontologies to describe domains of interest and to help users organize their personal information. Ji-Lung Hsieh presented *A Web-based Tagging Tool for Organizing Personal Documents*. The system allows users to add tags to documents in their personal collections and then use these tags as part of search queries.

After the paper presentations, Michael Bernstein led a group discussion on the topic of finding versus organizing. To start the discussion, he queried the room about how many people regularly used a desktop search tool. Although 90% of the participants did, only one person said he was willing to give up traditional hierarchical folders as an organizational tool because he felt search supplanted his need to organize. William Jones confirmed that in his research he has found only about one in 20 people are willing to give up folders in favor of search.

The group discussed different organizational metaphors, including files and tagging, with an emphasis on the value each metaphor provides for different PIM behaviors. The behaviors focused on were primarily finding and organizing, but also included reminding, handling interruptions, and dealing with task management. Collections with a lot of meta-data, such as email, appear to be searched differently than meta-data poor collections. Tagging or automatically extracting meta-data from information items may affect people's interactions with those collections.

3.2 The Disappearing Desktop

As mentioned earlier, special focus was given in the workshop to personal information management that occurs beyond the desktop via mobile devices and the Web. In the second session of Day 2, participants explored the implications of the disappearing desktop on personal information management through panel presentations and discussion.

Several of the panel presentations dealt with the issues that make personal information management in a mobile environment different from personal information management on the desktop. Manas Tungare,

in his talk titled *Thinking Outside the (Beige) Box: PIM Beyond the Desktop*, explored which device important messages should be delivered when a person is working on multiple devices. Daniel Robbins presented a proposal for unifying multiple PIM applications on a mobile device in his discussion of *TapGlance: Designing a Unified Smartphone Interface for PIM*. He argued that mobile devices may be a good place to begin exploring solutions to fragmentation because such devices typically have fewer applications, less data, and only one user.

Unification of data was a common theme in the presentations. Andreas Komninos, in a panel presentation about *Holistic PIM: Managing PI for a Nomadic Generation*, argued that to support mobile scenarios, information needs to travel with people and be available from multiple locations. Wolfgang Woerndl, in a presentation of *SeMoDesk: Towards a Mobile Semantic Desktop*, discussed how to apply principles of the Semantic Desktop, where meta-data is assigned to information items and used to relate items, to mobile devices.

Following the presentations, Jens Dittrich led a group discussion. To start the discussion, he asked participants to look into the future and imagine what non-desktop-centric tools will look like in ten years. The discussion quickly focused on people's ability to gain access to personal information in a variety of settings. People questioned whether information should be device-centric or stored in the cloud. None of the workshop participants reported currently relying on the cloud exclusively. Some people discussed the influence of future hardware, such as bigger portable displays and battery power, on mobile PIM, and others questioned whether there is value to distributing PIM among different devices or consolidating it onto one device. Regardless of the future of mobile technology, participants agreed that social aspects of PIM on mobile devices will remain as important in the future as they are now. Divided attention will always be an issue, and the availability of plausible deniability (e.g., cell phone dead spots as a reason for not answering the phone, or spam-filters as an excuse for ignoring an email) will likely continue to provide a benefit.

3.3 Fragmentation

As was touched upon in the session on mobile PIM, a grand challenge for personal information management is dealing with fragmentation. Personal information is fragmented across electronic documents, email messages, paper documents, instant messages, photographs, etc. Each type of information is organized and used to complete different tasks and to fulfill disparate roles in an individual's life. The session on fragmentation began with two panel presentations, *The Semantic Desktop as a Foundation for PIM Research* by Leo Sauermann, and *ZOIL – A Cross-Platform User Interface Paradigm for PIM* by Hans-Christian Jetter. Sauermann suggested fragmentation could be minimized by building relationships between existing files (or file fragments) on a person's desktop. Jetter presented a design proposal for a system, called ZOIL, which provides different views into information.

Following the panel presentations, Steve Whittaker began discussion by asking why fragmentation happens, and why researchers have not fixed it yet. A few participants argued that some fragmentation can be good, or that it may be a necessary aspect of application innovation and specialization of applications. Text was cited as a successful example of a unifying format, with many applications being able to export or import text. However, while universal, text does not preserve relationships between artifacts. One problem highlighted by trying to unify relationships is that a single namespace is often used to refer to all sorts of information artifacts. It was suggested that by making applications namespace- and URI-aware, designers could expose artifacts and relationships using RDF to achieve a semantically separated but shared representation. The conversation then shifted to a discussion of how external programs (such as, for example, Task Tracer) might be able to create cross-application models

by observing user behaviors across open applications. The idea is appealing, but also presents many challenges.

3.4 Group Information Management

The study of personal information management cannot succeed by considering a person in isolation from the various groups in which that person works and lives. In the final session of the workshop, participants heard several talks as part of a panel on group information management (GIM). As seen in the panel, group information management can refer to many things, including how other people influence an individual's personal information management, how people share personal information, and how groups manage shared group information. Prabhakar Tadinada presented a system for sharing personal information within a group in his talk titled *Semantic 'LS' – An Approach for Personal and Private GIM*. The system reduces fragmentation across group members' information by allowing access based on content rather than storage location. Chris Middup, in a talk called *Managing the Adoption of PI in Co-located Work Groups*, described a study and model of how groups synthesize personal information into group information, and in doing so develop a shared understanding of the artifact.

Tara Whalen highlighted several challenges with GIM in her talk titled *File Sharing and Group Information Management*. Through a survey of how people share slides, she found that people have a hard time keeping track of who has an item, and that they sometimes accidentally share private information. Anthony Collins also found that privacy is important in his study of *Collaborative PIM with Shared, Interactive Tabletops*. He presented a system where information is exported to a tabletop device and shared among users of the tabletop.

Following the panel presentations, Kirstie Hawkey led a discussion on GIM. Privacy in GIM, which came up often in the talks, was also a major topic of discussion. Once personal information is released to a group, some control over it is lost. Privacy was also discussed in the context of group interaction with personal information. For example, accessing personal information in a group setting via desktop search may expose private information in a way that accessing it via folder navigation would not. Similarly, a person's notes taken during a meeting may contain private information that can be inadvertently shared when the notes are shared. Offices and organizations may have ad-hoc rules or practices that have been established over time for managing and sharing GIM, but these may not be codified. There are also issues of the level of detail or summarization that is used when sharing information in a GIM setting, and more summarization may be desirable in GIM.

4 Summary

In contrast to previous PIM workshops, the PIM 2008 workshop included less of an emphasis on defining terminology, identifying related fields, and building a common background, and more discussion of active research. This change perhaps reflects the fact that the study of personal information management is starting to mature as a field. Common topics of interest included the value of tagging versus using folders, searching versus browsing, and capturing information versus comprehending it. Discussion was generally tool or solution focused, although participants highlighted the continuing need to understand users' PIM needs as tools continue to evolve. The tools and solutions considered were generally holistic in nature. They focused on reducing, rather than increasing, fragmentation, and recognized that individuals do not manage information for its own sake. Instead, personal information management was seen as a means to an end – whether the end is a completion of a well-defined task or something less well understood and harder to articulate. We believe that this workshop was a productive step in bringing people closer to completing their personal information tasks.

Further information about the workshop, including pointers to the accepted papers, can be found at <http://www.pim2008.org>. Two recent books on the topic of personal information management, *Personal Information Management* edited by William Jones and Jaime Teevan, and *Keeping Found Things Found* by William Jones, provide additional background.

The workshop was made possible thanks to the efforts of many people. Jaime Teevan and William Jones chaired the workshop. Ofer Bergman coordinated reviews, and Ofer Bergman, Danyel Fisher, and Steve Whittaker were review area chairs. Robert Capra was the volunteer coordinator, Antonella Poggio, the local area chair, Jens Dittrich, the website coordinator, and Jacek Gwizdka, the publicity and online activities coordinator. Deborah Barreau will lead PIM 2009, to be held in conjunction with ASIST in Vancouver in the fall of 2009. Look for a call for papers come spring of 2009.