Keeping Found
Things Found
Keeping Found Things Found

The Study and Practice of Personal Information Management

William Jones
I dedicate this book to my wife, Maria Jones Staaf
## Contents

**Foundations**

<table>
<thead>
<tr>
<th>Chapter One</th>
<th>A study and a practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Keeping found things found</td>
<td>4</td>
</tr>
<tr>
<td>1.2 An ideal and the reality</td>
<td>7</td>
</tr>
<tr>
<td>1.3 A brief history of PIM</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Who benefits from better PIM and how?</td>
<td>13</td>
</tr>
<tr>
<td>1.5 A study and a practice</td>
<td>14</td>
</tr>
<tr>
<td>1.6 Looking forward: A map for this book</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter Two</th>
<th>A personal space of information</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Starting out</td>
<td>25</td>
</tr>
<tr>
<td>2.2 What is information to us?</td>
<td>28</td>
</tr>
<tr>
<td>2.3 How is information personal?</td>
<td>33</td>
</tr>
<tr>
<td>2.4 The information item and its form</td>
<td>36</td>
</tr>
<tr>
<td>2.5 Defining a personal space of information</td>
<td>40</td>
</tr>
<tr>
<td>2.6 Making sense of the PSI</td>
<td>45</td>
</tr>
<tr>
<td>2.7 Looking back, looking forward</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter Three</th>
<th>A framework for personal information management</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Starting out</td>
<td>56</td>
</tr>
<tr>
<td>3.2 Perspectives on personal information management</td>
<td>59</td>
</tr>
<tr>
<td>3.3 PIM activities to map between information and need</td>
<td>60</td>
</tr>
<tr>
<td>3.4 PIM-related activities and PIM-related areas</td>
<td>67</td>
</tr>
<tr>
<td>3.5 Weaving PIM activities together</td>
<td>72</td>
</tr>
<tr>
<td>3.6 Looking back, looking forward</td>
<td>75</td>
</tr>
</tbody>
</table>
Preface

This book has its origin in the Keeping Found Things Found (KFTF) project, which began in 2000 at the University of Washington Information School. Funding for the project originally came through a three-year National Science Foundation grant to study how people keep web information for later use and how tools might help. As primary author of the grant proposal, I selected the working title, “Keeping found things found on the Web.” It stuck and was shortened to become the name of the overall project (see http://kftf.ischool.washington.edu/).

Finding information is an important and well-studied problem. Communities of researchers study how people seek information and the ways in which information retrieval systems might help. But, once found, what do we do with it? How do we keep and organize such information for later, and sometimes repeated, use in our lives?

In a first-year study, we observed that people used many methods to keep web pages for later use. Analysis revealed that different methods served different purposes. And no method came close to doing everything a person might like to do to keep information available for future use.

One major challenge was simply figuring out where and how to keep information. At home or at work? On which computer? In which account? In which organization? In what form? As a paper printout? An e-document? In an email message? As a web bookmark? Or perhaps, as an in-line-reference in some document?

Later, questions like these were asked again as people tried to decide where and how to re-find the information (“Where did I put it?”). And, between keeping and re-finding, people also need to determine how best to organize and maintain all the information accessed and accumulated. Is it worth it to organize information into folders? Or maybe labels and tags could be used instead? Or maybe nothing at all, if information could be searched for? But what if basically the same information—a document being revised, for example—exists in many versions. How is one to be sure the most up-to-date or relevant version is the one retrieved?

The preceding are basic questions of personal information management, or PIM.
Mike Eisenberg, while still dean of the University of Washington Information School and also a member of the KFTF project, was the first to suggest the connection between work of the KFTF project and the larger problem of PIM. At first I balked: “Personal information management?”

I knew of Lansdale’s excellent “Psychology of Personal Information Management” article written back in the 1980s (1988 to be exact). Since then, however, the phrase seemed to have become most strongly associated with “PIM tools,” which combine email management with time, task, and contact management.

Leaving aside its associations, what about the phrase itself? Isn’t nearly everything we do informational? Isn’t “personal information management” quite broad? Can there be a meaningful area of study for something as broad and inclusive as that? Yes, yes, and yes.

Yes, nearly everything we do is informational. More to the point, in this information age, much of what we “see” and “hear” about the world is not through direct experience but rather through various forms and channels of information. We watch television, read newspapers and email messages, or surf the Internet. Likewise, many of the ways to effect change in our world are mediated by information. We complete web-based or paper-based applications, send email messages, and write reports. Some of us post information on blogs or to personal web sites. Our days are filled with many acts of information management, large and small.

Yes, personal information management is broad; but its breadth is also part of its point. As I’ve talked to people, formally through the KFTF studies and informally at social gatherings, my ever-growing impression is that important aspects of our personal relationship to the world—as mediated by information and information tools—have been falling through the gaps between established areas of research and development. Feature-laden cell phones, digital cameras, software applications, and various other information tools, even as they help us in some ways, often add extra stress and complications to our lives. One tool organizes and formats information in one way; another tool organizes and formats information in a completely different way. Even though each accomplishes its intended purpose, these tools don’t work together in ways that make sense for our lives with respect to the things we hope to do and strive to be.

Finally, PIM can be, and is becoming, a meaningful area of study with a community of researchers. The process of building a community for the study of PIM may have begun with a Special Interest Group session on personal information management, which was organized as part of the CHI 2004 conference on human–computer interaction. But perhaps the watershed event in the creation of a PIM community was PIM 2005—a special NSF-sponsored workshop held in January of 2005 in Seattle. The participants formed a nexus for follow-on workshops, special issues, and an edited book on PIM.

The community of researchers studying PIM is diverse with respect to primary discipline and technical background. For example, the PIM 2005 Workshop brought together researchers from the fields of cognitive science, human–computer interaction (HCI), information science, artificial intelligence, database management, knowledge management, and information retrieval. Participants had been doing PIM-relevant research all along. But the workshop, and other community-building events that followed its lead, have helped give common expression to PIM problems researchers had each been tackling in separate ways. Events like these have also promoted the exchange of data and ideas in support of PIM solutions.

It is my hope that this book will serve a similar purpose for each of us no matter what our professional, educational, or technical background. We are all involved in a daily struggle to make effective use of information and numerous informational tools. We have been doing PIM all along. The book can help, first of all, by giving expression to our PIM activities and the problems encountered as we try to make sense of all the information in our lives.

The book is also about solutions and approaches to take. It describes steps we can take right now to improve our PIM practices. The book also describes the PIM support we can expect to find available in the not-too-distant future. Most important, the book describes the kinds of questions we need to ask and the considerations we need to apply to a never-ending cavalcade of new tools, new technologies, new systems, new initiatives, and so on. What will work for us in our lives? How can we take charge of our information? And, in doing so, how might we also take greater charge of our lives?

Acknowledgments

The following people have helped, directly or indirectly, with the completion of this book. The ordering of names is arbitrary save for the first:

• My wife, Maria Jones Staaf, gave support at many levels and in many ways even as the process of book completion seemed to take over not only my life but hers as well. The book could not have been completed without her support, encouragement, and understanding.

The book could not have been completed without her support, encouragement, and understanding. 2


2 National Science Foundation (NSF) grant # 0435134—see http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0435134.

3 PIM 2005, a special two-day workshop, was held as part of SIG IR 2006 (http://pim.ischool.washington.edu/pim05/index.htm). As of this book’s writing, planning for another workshop in association with CHI 2008 is underway.


6 For more information on this PIM 2005 workshop and to access its final report, see http://pim.ischool.washington.edu/pim05/home.htm.

7 PIM 2006, a special two-day workshop, was held as part of SIG IR 2006 (http://pim.ischool.washington.edu/pim06/index.htm). As of this book’s writing, planning for another workshop in association with CHI 2008 is underway.
A number of people provided useful feedback on matters ranging from structure and content to style and clarity. I take full responsibility for the book’s remaining failings, but the book is much better for their efforts:

• Bob Boiko, John Wetherbie, and Ryan W. White gave considered, invaluable comments on nearly all chapters in the book.

• Alan Dix, Edward Fox, and Peter Morville provided extremely useful comments on numerous chapters.

• Rick Boardman, Mary Czerwinski, Alan Dix, Jonathan Grudin, Jacek Gwizdka, and one anonymous reviewer provided very useful feedback on the original book proposal that helped me set the book’s direction and coverage.

• Diane Cerra was the personal face of the publisher through much of my struggle to complete the book. She arranged for reviewer feedback from the people just mentioned, as well as its interpretation. Diane was a constant, comforting source of support and guidance.

• Michael Adcock, Andrea Lisabeth Civan, and Predrag V Klasnja provided very useful feedback on several chapters.

• Olle Bäster, Susan Dumais, George Furnas, Jim Gemmell, Jane Glascoc, Lars Johansson, Mike Kelly, Rod Such, and Don Webb gave very useful feedback on selected chapters in the book.

• Olle Bäster, Bob Boiko, Mike Kelly, Peter Morville, Mike Nakahara, and Dan Russell took time out of their already busy professional lives to contribute some of the book’s sidebar content. The book is much better and more well-rounded because of their contributions.

• I am also grateful to Hala Annabi for the many useful references she provided me to current research about information management at the organizational level. I also thank Karen Erickson for her assistance with two of the book’s figures.

• Cheyenne Maria Roduin provided constant, often short-notice, assistance throughout the book’s completion, ranging from research on various topics to copyediting and proofreading. Other people to help me directly with copyediting and proofreading include Andrea Lisbeth Civan, Glenda Claborne, Marguerite Fingnan, Beth Fournier, Shelley K. Minden, and anonymous editors at Morgan Kaufmann/Elsevier.

Other people have provided support without which the book’s completion would have been much more difficult:

• Harry Bruce, as my long-time colleague and co-manager of the Keeping Found Things Found project, provided invaluable support for the project during the book’s completion, even as he also worked to meet the considerable demands of his new position as dean of the University of Washington Information School.

• Kerstin Severinson Ekundh and Olle Bäster were excellent, understanding hosts during my extended springtime visit to the Royal Institute of Technology (KTH) in Stockholm, even though the final stages of the book’s completion were a dominant presence throughout the visit.

• Maria Zemankova, as a program officer at the National Science Foundation, deserves great credit for her vision in seeing the potential of personal information management. Many of the community-building activities listed previously in this preface can be traced, directly or indirectly, to her support and guidance.

Acknowledgments include things as well as people and are a testament to the transforming power of the informational tools and technologies that are a primary topic of this book. For example, the reliable use of email (only occasionally supplemented by the phone) meant that book-related communication could happen over time and across great distances.

The Web greatly increased my reach to relevant material. I was especially grateful for the use of Google’s search service and the Wikipedia to provide useful points of disembarkation into the Web in all its vastness.

Information sources must always be qualified and information validated. This is perhaps especially true for information found on the Web. But the qualification of sources and the validation of information was itself greatly facilitated by the online, Web-mediated access to information collections in the extensive University of Washington Libraries system.

The book was done in its entirety on an aging but very reliable IBM Thinkpad X30 laptop computer. With the benefit of wireless access, large portions of it were completed at home, in my favorite coffee shop, and even at the Kirkland Public Library.

The Web provided me with access not only to information but also to people. For example, I first “discovered” Elizabeth Boling, whose illustrations appear in every chapter of the book, through her work on a professional web site. As of this writing, I have not yet met Elizabeth in person. Our interactions have been entirely through email with the exception of one phone call. Similarly, my interactions with Cheyenne Roduin were exclusively via email (including an email introduction) for a period of nearly seven months before I finally met her in person—even though, for much of this period, we were both in the greater Seattle area.

But even as my work on the book was greatly facilitated by new modes of communication, the serendipitous networking enabled by old-fashioned “face-to-face” contact still played an essential role in its completion. Lisa Liegdon, as noted on the copyright page, did the concept work for the book’s overall design and layout, including its cover and chapter opener spreads. Lisa is a well-known, Seattle-area artist and also a family friend. At the very beginning of my work on this book, I happened to mention my plans to her as we talked to each other at a dinner party. Lisa is interested in PIM and was the one to say, “Let me design your book.” It would not have occurred to me to ask. I have been grateful ever since for Lisa’s involvement: the book would not be the book it is without her.

Finally, it is important to acknowledge the underlying support of the National Science Foundation for the KFTF project and for the efforts by its principle researchers to facilitate an exchange of information among people doing PIM-relevant research. Descriptions of the project and these efforts are based on work supported by the NSF under Grant Nos. 0097855, 0334623, 0435134, and 0534386. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author and contributors and do not necessarily reflect the views of the National Science Foundation.
Contributors

Olle Bälter
Royal Institute of Technology, Stockholm, Sweden
Olle Bälter became known as "Mr. Email" in Sweden in 2002 after a popular book based on his doctoral thesis was published. It has advice on how people can improve their email situation. Olle is currently an associate professor and vice dean of education in the School of Computer Science and Communication at the Royal Institute of Technology.

Bob Boiko
University of Washington, Seattle, WA
Bob Boiko is the author of The Content Management Bible and the new book Laughing at the CIO: A Parable and Prescription of IT Leadership. He has almost 20 years of experience designing and building web, hypertext, and multimedia systems and tools for some of the world’s top technology corporations, including Microsoft, Motorola, and Boeing. Bob is founder and president of Metatotal Services Inc. and is also a faculty member at the University of Washington.

Mike Kelly
Microsoft Corporation, Redmond, WA
Mike Kelly is the director of the Emerging Practices team within Microsoft’s Engineering Excellence group. This team investigates and promotes new software development practices that the company’s product teams can leverage to improve their engineering efficiency and quality. Before joining Engineering Excellence in 2006, Mike worked for 12 years in various development roles on the Microsoft Office team.

Peter Morville
Semantic Studios, Ann Arbor, MI
Peter Morville is the best-selling author of Ambient Findability and Information Architecture for the World Wide Web. He is president of Semantic Studios, a leading information architecture and user experience consulting firm, and he is a founder and past president of the Information Architecture Institute. Peter is also an adjunct faculty member at the University of Michigan. He blogs at findability.org.

Mike Nakahara
Avocent, Redmond, WA
Mike Nakahara is a director at Avocent, a leader in connectivity and IT infrastructure management. Previously, Mike was a program manager in the Smart Personal Object Technology (SPOT) group at Microsoft. In his years there, he worked on a new class of mobile products including Smart Watches and portable weather stations.

Dan Russell
Google, Mountain View, CA
Daniel M. Russell is a research scientist at Google where he works on search quality, with a focus on understanding what makes the company’s users happy when they use the web search engine. From early 2000 until mid-2005, he was a senior research scientist in the User Sciences and Experience Research (USER) lab at IBM’s Almaden Research Center. Dan is perhaps best known for his recent work on the large, interactive IBM BlueBoard system for simple collaboration and for his studies of people’s sensemaking activities.