

Unifying PIM Research: Fostering a Connection Between Descriptive PIM Studies and Prescriptive Outcomes

Anne R. Diekema
Utah State University
2830 Old Main Hill
Logan, UT 84322
+1 435 797-2697

anne.diekema@usu.edu

ABSTRACT

Descriptive Personal Information Management (PIM) studies inform us about PIM behavior and their findings should guide the design and development of PIM tools to support the behavior under study. Unfortunately, judging from the literature, descriptive studies do not always provide useful recommendations and PIM tool research is often carried out separately. This paper discusses what appears to be a possible research dichotomy and ways to bring the research back together. Three solutions are suggested: 1) PIM workshops where both types of studies are presented and researchers meet should be important venues for dissemination of results, cross-fertilization between different research areas, and collaboration between researchers; 2) A bridging methodology to translate research findings explicitly into design criteria could bring research and practice closer together; and 3) A general PIM framework based on the three essential PIM activities (finding/refinding activities, keeping activities, and meta-level activities).

General Terms

Management

Keywords

Personal Information Management; Research; Methodology

1. INTRODUCTION

In the midst of studying the personal information management (PIM) of teachers by interviewing them in their natural environments, the nagging “so what?” question emerged. Sure, PIM researchers are finding interesting behavior and we have very rich data to analyze, but how can we translate our findings into something applied such as a tool, technique, or strategy to improve PIM practices [13]? This dichotomy between researching people and building and studying systems appears to be reflected in the PIM literature. Some studies focus exclusively on specific applications or systems (e.g. [2, 17, 18]), while others analyze people and the informational tasks they encounter in their daily lives (e.g. [29, 25]). Fewer studies appear to make strong, explicit connections between descriptive studies and resulting PIM tools,

(e.g. [6] and the Keeping Found Things Found project and related studies [8]). One of the reviewers of this paper suggested that perhaps this dichotomy is artificially created by publishing requirements, leading to papers that are either about behavior or about tools, but when one examines PIM projects as a whole both strands of research appear. The division between research that focuses on users versus systems is a well-documented phenomenon in the field of information retrieval. This separation has led to the increasing divergence of two research communities, much to the detriment of the field as a whole [8, 10, 21].

The system side of information retrieval, also known as the physical or Cranfield paradigm [5], is mainly concerned with the creation, improvement, and evaluation of retrieval algorithms. A good retrieval algorithm efficiently finds relevant documents while ignoring documents that are not on topic. The main thrust of this empirical research is evaluating system performance based on a test collection of documents, queries, and related relevance judgments. Systems run batches of queries and their results are scored using standard metrics [24, 26]. The user side of information retrieval, also known as the cognitive paradigm, studies and models information seeking and retrieval practices of people [9]. The cognitive researchers criticize the systems side for their unrealistic retrieval experiments that do not take real users and situational relevance criteria into account [22]. The systems side is also critiqued for not incorporating user studies findings, which have implications for systems design. The system researchers, for that matter, seem unaware of the cognitive research. That said, the failure of the cognitive side to provide concrete recommendations makes it difficult to incorporate user studies in actual system design [21].

According to [5], the duality observed in information retrieval research is preventing the field from developing “a powerful body of theory” (p. 45). Given that PIM research has the same two areas of focus - people (managers, teachers, people with HIV/AIDS, students, etc.) and systems (email, the Internet, desktop search, note taking tools) - it is likely that some of the same dichotomies affect PIM research. That leaves the question: How can we pull both areas of focus together in order to create more usable systems and PIM practices?

2. LIKE SHIPS IN THE NIGHT: RELATED BUT SEPARATE RESEARCH AGENDAS

The problem of having what appear to be related but separate research streams in a field is by no means unique to information retrieval. A similar problem was recognized by [27] who label this as the *dissemination problem*, although the issue goes beyond dissemination alone. They write in the context of Human Computer Interaction (HCI): “In sum, although we lack basic

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference'10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010...\$10.00.

understandings of current users, tasks, and technologies, the field is encouraged to try out even more radical solutions without pausing to do the analysis and investigation required to gain systematic understanding. Furthermore, even when a useful body of knowledge does exist for a core task, the HCI community does not have institutions and procedures for exploiting this knowledge” [27, p. 80-81].

While I don't think the situation is quite this dire in the PIM arena, there do seem to be problems with disseminating descriptive research results so they can inform tool or strategy development. After completing our labor-intensive naturalistic studies, we often wistfully send our findings out into the world with general endorsements such as “PIM software developers will benefit from an understanding of how teachers manage their information” [3, p. 189], and leave it at that.

One of the reasons for the separate research streams undoubtedly originates in the different backgrounds and skill sets of the researchers. The information retrieval system research tends to come out of computer science while the cognitive research originates in the library and information sciences and other social sciences. PIM researchers comfortable with social science research methodologies are not necessarily well-versed in prototyping and system design and development, and vice versa.

3. POSSIBLE SOLUTIONS

[27] suggest several solutions to the lack of integration of research findings on core HCI tasks. The authors recommend workshops to discuss and disseminate research on these core tasks and also propose stricter reviewing practices to ensure the incorporation of existing research into new work (a practice known as generativity, see [23]). The role of education in spreading these practices is also stressed [27]. In their book *The Turn*, the authors [10] seek to bring the two sides of information retrieval together through extension of the cognitive viewpoint by consistently incorporating the technological and user information behavior sides. For PIM the most rational approach to bring together descriptive studies and tool and application research is perhaps through workshops, methodology, and creating a general framework of PIM tasks.

3.1 Workshops

Workshops are great venues to disseminate research information because they are smaller and more intimate than conferences. Also, workshops tend to have a much narrower focus, bringing together researchers working in the same or closely related areas. Fortunately, PIM researchers meet regularly at workshops and have plenty of opportunity to learn about current user research (<http://pimworkshop.org/>). As long as all types of PIM researchers keep attending these workshops, the conditions for collaboration and dissemination are ideal. The combination of PIM workshops with larger conferences attended by different kinds of researchers (information retrieval, human computer interaction, information science, computer supported cooperative work) is another productive way to foster collaboration in various related areas.

The first PIM workshop was sponsored by NSF and took place in 2005, followed by a 2006 workshop as part of the Special Interest Group Information Retrieval (SIGIR) conference. Workshops at the Computer Human Interaction conference and the American Society for Information Science & Technology (ASIS&T) annual meeting followed in subsequent years. The latest workshop is scheduled for 2012 as part of the Computer-Supported Cooperative Work conference (CSCW). To improve integration of

previous research findings into new research, review guidelines for future workshops could include a generativity criterion.

3.2 Methodology

Achieving a deeper understanding of PIM behavior before developing tools is perhaps the most logical sequence of PIM research [15] although there are studies that begin with a prototype that is then tested by real users (e.g. [1]). Can methodology help the integration of the PIM behavior findings into tool design development and evaluation? One way might be to follow [7] who call for a systematic approach to PIM research with formalized PIM tasks, methods, behaviors, and mappings between PIM contexts and strategies. Another way that methodology may help research integration might be the creation of a bridging methodology. In the conclusion of their book on PIM, the authors [13] stress the importance of the development of methodologies of a relatively young field like PIM. Yet, the methodologies discussed are divided into two categories: descriptive studies (for studying people's PIM behavior) and for prescriptive evaluations (measuring the effectiveness of tools) [19, 16]. Perhaps what is needed is some sort of bridging methodology that can translate the prescriptive study findings into actionable tool design considerations for tool development as well as tool evaluation. Obviously these criteria should go beyond the often generic recommendations typically mentioned in many PIM papers. A model for bridging methodologies can be found in user-centered design which encompasses system design methodologies with various levels of user involvement [4, 20].

3.3 General PIM Framework

Another way to integrate both sides of PIM research is to develop a general PIM Framework where descriptive research and tool development meet. Work by [11, 12] has already identified essential PIM activities such as finding/refinding (resulting from an information need), keeping activities (concerning incorporation of information into a person's personal information space), and meta-level activities (related to organizing and managing the information in the personal information space) that could form the basis of such a framework. Based on these activities researchers could organize and present their findings and tool design implications in such a way that they are easily accessible. The framework might need to be extended with a miscellaneous activities category to capture new and unusual activities that might arise. An extension might allow for individual differences in the creation of personalized and customized PIM tools. Alternatively existing frameworks as proposed by [10] and [28] can be adapted to guide PIM research.

4. CONCLUSIONS

The strong user-system dichotomy in the field of information retrieval does not currently seem to be as problematic in PIM. That said, there is a potential rift between descriptive studies of PIM behavior and the design and development of PIM tools. Unfortunately, judging from the literature, descriptive studies are not always clear in providing useful recommendations, while PIM tool research is often carried out separately. Workshops where both types of studies are presented should be an important venue for dissemination of results, cross-fertilization between different research areas, and collaboration between researchers. This is already happening, but the connection between research and practice needs to be more intentional. Another way to bring the research together is to create a bridging methodology to translate research findings explicitly into design criteria and by creating a

general PIM framework based on the three essential PIM activities [12] as a common language.

5. REFERENCES

- [1] Adar, E., Karger, D., & Stein, L. A. (1999). *Haystack: per-user information environments*. Paper presented at the Proceedings of the eighth international conference on Information and knowledge management, Kansas City, Missouri, United States.
- [2] Bellotti, V., Ducheneaut, N., Howard, M., & Smith, I. (2003). *Taking email to task: the design and evaluation of a task management centered email tool*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, Ft. Lauderdale, Florida, USA.
- [3] Diekema, A. R., & Olsen, M. W. (2011). *Personal Information Management Practices of Teachers*. Paper presented at the ASIS&T 2011, New Orleans, LA, USA.
- [4] Draper, S. W., & Norman, D. A. (2009). *User centered system design : new perspectives on human-computer interaction*. Boca Raton, FL.; London: CRC Press.
- [5] Ellis, D. (1992). The Physical and Cognitive Paradigms in Information Retrieval Research. *Journal of Documentation*, 48(1), 45-64.
- [6] Elswailer, D., Baillie, M., & Ruthven, I. (2009). On Understanding the Relationship Between Recollection and Refinding. *JODI: Journal of Digital Information*, 10(5), 1-31.
- [7] Gwizdzka, J., & Chignell, M. (2007). Individual Differences. In W. P. Jones & J. Teevan (Eds.), *Personal information management* (pp. 206-220). Seattle: University of Washington Press.
- [8] Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, 52(1), 3-50.
- [9] Ingwersen, P. (2001). Cognitive information retrieval. [Book Chapter]. *Annual Review of Information Science and Technology*, 34, 3-52.
- [10] Ingwersen, P., & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dordrecht: Springer.
- [11] Jones, W. P. (2007). Personal Information Management. *Annual Review of Information Science and Technology*, 41(1), 453-504. doi: 10.1002/aris.2007.1440410117
- [12] Jones, W. P. (2008). Keeping found things found : the study and practice of personal information management. Amsterdam; Boston: Morgan Kaufmann Publishers.
- [13] Jones, W. P., & Teevan, J. (2007). *Personal information management*. Seattle: University of Washington Press.
- [14] Keeping Found Things Found: Personal Information Management (<http://kftf.ischool.washington.edu/>)
- [15] Kelly, D. (2006). Evaluating Personal Information Management Behaviors and Tools. *Communications of the ACM*, 49(1), 84-86.
- [16] Kelly, D., & Teevan, J. (2007). Understanding What Works: Evaluating PIM Tools. In W. P. Jones & J. Teevan (Eds.), *Personal information management* (pp. 190-204). Seattle: University of Washington Press.
- [17] Kleek, M. G. V., Bernstein, M., Panovich, K., Vargas, G. G., Karger, D. R., & Schraefel, M. (2009). *Note to self: examining personal information keeping in a lightweight note-taking tool*. Paper presented at the Proceedings of the 27th international conference on Human factors in computing systems, Boston, MA, USA.
- [18] Minkov, E., & Cohen, W. W. (2010). Improving Graph-Walk-Based Similarity with Reranking: Case Studies for Personal Information Management. *ACM Transactions on Information Systems*, 29(1), 4:1-4:52. doi: 10.1145/1877766.1877770
- [19] Naumer, C. M., & Fisher, K. E. (2007). Naturalistic Approaches for Understanding PIM. In W. P. Jones & J. Teevan (Eds.), *Personal information management* (pp. 76-88). Seattle: University of Washington Press.
- [20] Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.
- [21] Saracevic, T. (2008). *Information Science: Where does it come from and where is it going?* Presentation given at Syracuse University.
- [22] Schamber, L. (1990). A Re-Examination of Relevance: Toward a Dynamic, Situational Definition. *Information Processing and Management*, 26(6), 755-776.
- [23] Shulman, L. S. (1999). Professing educational scholarship. In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 159-165). San Francisco: Jossey-Bass.
- [24] Spärck Jones, K. (1981). *Information retrieval experiment*. London; Boston: Butterworths.
- [25] Veinot, T. (2010). A multilevel model of HIV/AIDS information/help network development. [Article]. *Journal of Documentation*, 66(6), 875-905. doi: 10.1108/00220411011087850
- [26] Voorhees, E., Harman, D. K., National Institute of, S., & Technology. (2005). *TREC : experiment and evaluation in information retrieval*. Cambridge, Mass.: MIT Press.
- [27] Whittaker, S., Terveen, L., & Nardi, B. A. (2000). Let's Stop Pushing the Envelope and Start Addressing It: A Reference Task Agenda for HCI. *Human-Computer Interaction*, 15(2/3), 75-106.
- [28] Wilson, T. D. Information behaviour: An interdisciplinary perspective. *Information Processing & Management*, 33, 4 (1997), 551-572.
- [29] Xin-Yu, J., & Hsiao-Tieh, P. (2010). An Exploratory Study on the Re-finding Behavior on the Web. (English). *Journal of Library & Information Studies*, 8(1), 29-57.