

# Finders, Keepers, Losers, Seekers: A Study of Academics' Research-Related Personal Information Collections

Mashaël Al-Omar and Andrew Cox

Information School, University of Sheffield, UK  
{liq09ma, a.m.cox}@sheffield.ac.uk

**Abstract.** In conducting their research, scholars are not only information seekers, they are information keepers and managers as well. This paper describes a study of seventeen scholars from Education and Health disciplines (College of Nursing and Health Science College) in the Public Authority for Applied Education and Training (PAAET), Kuwait and their research-related personal information collections. A model explaining the size, diversity, hybridity and fragmentation of these collections to immediate and underlying causes is presented.

**Keywords:** Personal Information Management, Human Computer Interaction, Information Behavior, Information Practice, Information Retrieval.

## 1 Introduction

Within their research, scholars engage in “searching, collecting, reading, writing and collaborating” [25]. As a result of these information practices, they build and then have to manage significant Personal Information Collections (PICs) [23], [16]. Such materials accumulate over time [17] and collections can grow to be huge in size, diverse in nature and format [11]. They include books, published works, and web pages, emails and electronic files on a computer’s hard drive [11]. Ensuring that information can be re-found when needed and so “exploited” [17], [27] is an information and data management challenge [10].

Investigation of these issues requires a holistic view of what are often separated out as different fields of study, i.e. Information Behaviour (IB), Personal Information Management (PIM) and Information Practice (IP). Research in this area should link information practices to the research process, and include investigation of the management of research data as well as secondary literature and other information. Studies often focus either on physical or electronic information collections [5], [15], [1], [13], [28] but both should be looked at together.

The purpose of this paper is to describe a study that was carried out to develop such in-depth understanding of scholars’ research-related PICs. The study examined how research-related PICs are created, used and managed within the research process.

It explored the factors that shape their management and sought to evaluate how successful scholars are at achieving the exploitation of information they collect.

## 2 Literature Review

Personal information management (PIM) is about how people “acquire or create, store, organize, maintain, retrieve, use, and distribute the information needed to complete tasks (work-related or not)” [12]. Researchers have investigated PIM in diverse ways such as to develop a system for information retrieval to facilitate information reuse [29] while others investigated information discovery and finding only [6], [7], [20], [4].

Relatively few studies to date have investigated scholars as information keepers and managers, rather the focus has tended to be on scholars information seeking behavior, in the context of their literature review [6], [24], [22]. One exception is Kaye et al. [18] who investigated forty eight scholars in multiple disciplines and ranging in seniority from graduate students to professors by touring their offices and conducting semi-structured interviews. The study found that academics not only store information for the purposes of information retrieval, but also for “*creating legacy, sharing resources, confronting fears and anxieties and identity construction*” [18]. The main uses were similar across disciplines and seniority. A great variety of storage strategies were uncovered in the research. Physical collections of information were stored in anything from custom-built offices to a mobile solution of bags and boxes stored in the back of a car. Digitally, academics developed their own individual way of archiving digital material rather than using available tools and solutions [18]. This is a relatively unique study.

Therefore, the aim of this research was to add to such relatively sparse literature by examining how research-related PICs are created, used and managed within the research process by answering main questions: What are personal information collections of scholars like? How do scholars use their personal information collections in their research? It also explored the factors that shaped their Personal Information Management (PIM) activities such as discipline, seniority, time pressure and the quality of support services.

## 3 Methodology

The study adopted an emergent design and an interpretive-qualitative approach based on in-depth, face-to-face interviews in order to understand the scholars’ world and life [3] as they talk about their experience in their own words [21]. The focus of this paper is to present analysis of seventeen interviews of scholars in the Public Authority for Applied Education and Training (PAAET), Kuwait, from the disciplines of Education and Health preceded by tours of their personal space of information and observation.

In addition to interviews, photographs of scholars’ personal space of information and their information collections were taken and treated as data complementing the interviews and achieve more understanding [26]. The interviews were transcribed

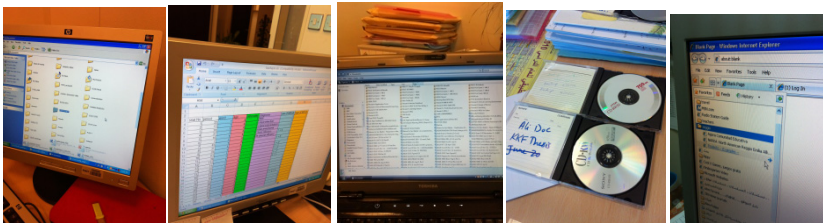
then analyzed thematically with the photographs to produce a list of codes which were then integrated to produce the themes identified in the textual data [19]. Transcripts and photographs were sorted into broad categories known as “proto-themes” [8] in order to allow themes to emerge from the data by categorizing similar topics together. The transcribed interviews were re-read in order to refine the proto-themes into the actual themes [3], [8].

## 4 Findings

The study found that scholars’ research information collections are large, diverse, hybrid, and fragmented (see figure 3 below). Scholars’ personal space of information contained a massive quantity of information related to their research that is stored in different places at different stages of the research. The part of a scholar’s desk where they were working usually contained an accumulation of information related to research, sometimes merged with non-research material such as for teaching, and management tasks. Research is an “off-on process” scholars said, so that when they are in an active research mood, the amount of the information on their desks is larger than when they are less active research mode. Within their working offices, the research collection accumulated in different storage places such as open shelves, drawers and closed cabinets in the form of piles, files, pile of files and randomly stored piles within a file. Information was stored in huge amounts electronically too. Electronic folders were found on scholars’ desktops, external storage devices, and stored virtually in personal E-mail or websites of participants that were created in an effort to support their own research tools of management.



**Fig. 1.** Pile and files of diverse papers stored physically in different ways in the main setting working space



**Fig. 2.** Electronically stored diverse research related information collections

The research-related PIC was found to be typically diverse in type, including four main categories of material, namely: sources gathered for the literature review, research data, publications and administrative papers.

Even after publishing a paper, literature review sources were kept by many scholars. They said that they thought they would need the material in future research in the same subject. It was also kept to share with colleagues in their local network within the department or anywhere across Kuwait. The second type of information was research data. It was found that the original paper questionnaires, for example, were kept together with processed results. Electronic versions were stored on their personal computers. Some scholars found it enough to keep the summary of the data in SPSS or Excel tables, but others preferred keeping the data as originally collected in the form of paper questionnaires. Research publications were also part of the collection, again occurring in both traditional and electronic formats. The final category of material in the collection was paperwork related to the research project.

In addition to its size and diversity, the scholars' collection was hybrid in formats including a complex mix of physical and electronic content, and located in multiple locations that change at different research stages. There was typically no clear strategy of keeping print or electronic or even both. All what was found is a random mix of both with a large amount of redundancy. All four types forming the diverse collection were found in both print and electronic form. Some scholars prefer reading printed versions while others prefer reading electronic. Some would tend to start with one version and continue reading in the other.

The collection is also fragmented. Enormous piles of papers and files were stored in multiple spaces whether in the main or secondary setting, in addition to abundant external storage devices for saving electronic versions such as flash memory and hard disks. Working in different locations necessitated the existence of the collection in each of the locations either by carrying the collection between locations by a mobility solution or like in some case keeping multiple copies here and there. Scholars tried to build an identical copy of the collection in each location, however hard that was to achieve.

The causes for these collection features can be divided into two main categories namely immediate and underlying causes.

The immediate causes found were the need for research, time pressure, quality of space, technology opportunity, support services, English resources and display. Scholars were obliged to conduct research for their career development. Conducting multiple research projects at the same time, along with poor support services, created time pressure and affected the research collection size, hybridity and fragmentation. In order to overcome the time pressure, and accomplish their tasks, scholars tried to manage their time by taking uncompleted tasks to a secondary setting (usually home) rather than conducting all research activities at the workplace. The quality of personal space triggered the failure of management of the research collection. Shared or poor quality offices forced scholars to find different working spaces which encouraged information hybridity and fragmentation. Technology opportunities such as the facilities of internet and ease of tools with the currency of applications were also effectively challenging scholars to maintain control. Such opportunities helped create the

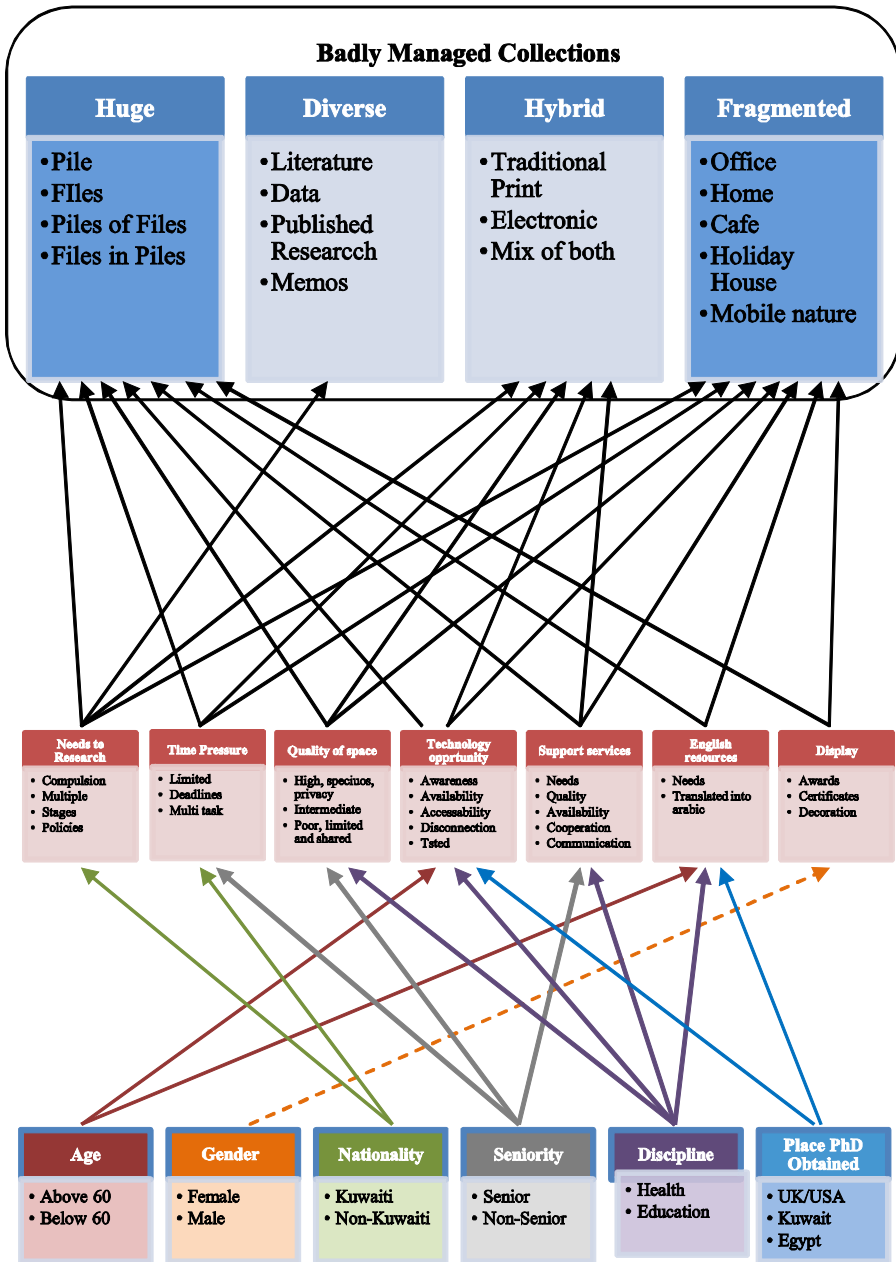


Fig. 3. Features and causes of research-related personal information collections

deluge of information. When participants were asked about the services offered to them by PAAET libraries, all talked about their dissatisfaction about the type and quality of services offered to support their research tasks. Scholars expected more support to aid their information finding, keeping and re-finding tasks. As some of the modules they teach have to be taught in Arabic language, English resources were one of the factors that shaped the features of the scholar's research collections. The research collections were displayed in the working spaces by some scholars for a number of reasons. The displayed collections were not limited to books and research output, but included awards and certificates. Such items were used for several purposes not for just sharing, and confronting fear of loss, but were displayed for remembering, constructing identity and creating legacy as well.

The immediate causes were driven by a number of underlying causes, namely Age, Gender, Nationality, Seniority, Discipline, and the Place PhD obtained. The age of the scholar affected the extent to which scholars took technological opportunities and English resources. Their need of Arabic resources also was affected by the place PhD obtained, as scholars who had graduated from UK and USA had no problem of using English resources, while those who graduated from Arabic universities struggled finding Arabic resources and translating material. It was found that female scholars were more eager to display their research related collections as well as their certificates and awards. Nationality affected the research and time pressure, as non Kuwaiti scholars were working under more pressure to conduct their research in order to renew their contracts. Seniority affected time pressure, quality of space, and support services. Whereas, the discipline affected the quality of space, technology opportunity, support services and English resources.

## **5 Discussion and Conclusion**

This research explored the research-related personal information collections of scholars in a novel context, Kuwait. A model was produced capturing the main factors shaping the collection, and its size, diversity, hybridity and fragmentation. It identifies that such collections are typically made up of four types of information content. The model offers an analysis of both the immediate and underlying causes of this character. Although many of these factors (such as language and support related issues) are relatively unique to the Kuwaiti context, most of the features of PICs in terms of scale, fragmentation and hybridity as well as how the collections are used mirror findings from previous studies e.g. Kaye et al. [18].

If judged by the ability to re-find information scholars' research-related PICs were in many respects a failure. Analysis of the immediate and underlying causes of the character of the collection point to a range of beneficial interventions from improved space, technology support through to training in IM principles.

## References

1. Boardman, R., Sasse, M.A.: Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 583–590. ACM (2004)
2. Bondarenko, O., Janssen, R.: Documents at hand: Learning from paper to improve digital technologies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 121–130. ACM (2005)
3. Boyatzis, R.E.: Transforming qualitative information: Thematic analysis and code development. Sage Publications, Incorporated (1998)
4. Buchanan, G., Cunningham, S.J., Blandford, A., Rimmer, J., Warwick, C.: Information seeking by humanities scholars. In: Rauber, A., Christodoulakis, S., Tjoa, A.M. (eds.) ECDL 2005. LNCS, vol. 3652, pp. 218–229. Springer, Heidelberg (2005)
5. Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R., Robbins, D.C.: Stuff I've seen: a system for personal information retrieval and re-use. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, pp. 72–79. ACM (2003)
6. Ellis, D.: Modeling the information-seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly*, 469–486 (1993)
7. Ellis, D., Cox, D., Hall, K.: A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation* 49, 356–369 (1993)
8. Hayes, N.: *Doing psychological research*. Taylor & Francis Group (2000)
9. Henderson, S.: How do people organize their desktops? In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1047–1048. ACM (2004)
10. Jones, W.: Finders, keepers? The present and future perfect in support of personal information management. *First Monday* 9 (2004)
11. Jones, W.: Keeping found things found: the study and practice of personal information management (2008a)
12. Jones, W.: Personal information management. *Annual Review of Information Science and Technology* 41, 453–504 (2008b)
13. Jones, W., Anderson, K.M.: Many views, many modes, many tools... one structure: Towards a Non-disruptive Integration of Personal Information. In: Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia, pp. 113–122. ACM (2011)
14. Jones, W., Bruce, H., Dumais, S.: Keeping found things found on the web, pp. 119–126. ACM (2001)
15. Jones, W., Dumais, S., Bruce, H.: Once found, what then? A study of “keeping” behaviors in the personal use of Web information. *Proceedings of the American Society for Information Science and Technology* 39, 391–402 (2002)
16. Jones, W., Teevan, J.: *Personal Information Management*. University of Washington Press, Seattle (2007)
17. Kaye, J.J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I., Pinch, T.: To have and to hold: exploring the personal archive, pp. 275–284. ACM (2006a)
18. Kaye, J.J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I., Pinch, T.: To have and to hold: exploring the personal archive. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 275–284. ACM (2006b)
19. King, N.: Using templates in the thematic analysis of texts. *Essential Guide to Qualitative Methods in Organizational Research*, 256–270 (2004)
20. Kuhlthau, C.C.: Inside the search process: Information seeking from the user's perspective. *JASIS* 42, 361–371 (1991)

21. Kvale, S., Brinkmann, S.: *Interviews: Learning the craft of qualitative research interviewing*. SAGE Publications, Inc. (2008)
22. Marouf, L., Anwar, M.A.: Information-seeking behavior of the social sciences faculty at Kuwait University. *Library Review* 59, 532–547 (2010)
23. Meho, L.I., Tibbo, H.R.: Modeling the information seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology* 54, 570–587 (2003)
24. Ocholla, D.N.: Information-seeking behaviour by academics: a preliminary study. *The International Information & Library Review* 28, 345–358 (1996)
25. Palmer, C.L., Tefteau, L.C., Pirmann, C.M.: *Scholarly information practices in the online environment*. Report commissioned by OCLC Research (2009), Published online at <http://www.oclc.org/programs/publications/reports2009-02.pdf>
26. Silverman, D.: *Interpreting qualitative data: Methods for analyzing talk, text, and interaction*. Sage Publications Ltd. (2006)
27. Whittaker, S.: Personal information management: From information consumption to curation. *Annual Review of Information Science and Technology* 45, 3–62 (2011)
28. Whittaker, S., Hirschberg, J.: The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction (TOCHI)* 8, 150–170 (2001)
29. Whittaker, S., Sidner, C.: Email overload: exploring personal information management of email. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground*, pp. 276–283. ACM (1996)