
Article:
AL-Omar, M. and Cox, A.M. (2016) Scholars' research-related personal information collections: A study of education and health researchers in a Kuwaiti University. Aslib Journal of Information Management, 68 (2). pp. 155-173. ISSN 2050-3806

https://doi.org/10.1108/AJIM-04-2015-0069

Scholars' research-related Personal Information Collections: A study of education and health researchers in a Kuwaiti University

Abstract

Purpose

The aim of the paper is to explore the character of scholars' research-related personal information collections (PICs).

Design/methodology/approach

The study was based on in-depth interviews and office tours of 17 scholars in Education and Health Sciences in a Kuwaiti Higher Education Institution.

Findings

Scholars' research-related PICs were added to throughout the research life-cycle. They were huge, diverse, hybrid and fragmented. Key factors shaping the collections were the pressure to do research, time pressure in general, quality of space available, technology opportunity, lack of support from central services, the need to collect Arabic material, self-presentation and self-management. Older scholars and non-Kuwaiti nationals experienced the pressures slightly differently.

Research limitations/implications

The study was limited to scholars in two disciplines, in one institution in a developing world context. However the models produced are suggestive of factors involved in shaping of the research-related PICs of scholars in general.

Practical implications

Failures in Personal Information Management are a cause for concern in terms of data integrity and validity of research. Interventions could include training of early career researchers for a life time of collecting.

Originality/value

This is the first study to examine the contents of scholars' research-related PICs and to provide a model of factors shaping them.

Introduction

Scholars are intensive users of information, and study of their work has long been important to information science. For example, the investigation of different scholarly practices of seeking and using literature have been central to the development of Information seeking and behaviour research as a whole (Case, 2012). For the field of Personal Information Management (PIM) study of scholars' behaviour has been less central, yet in the course of their work, scholars generate large collections of information and managing this material must at least in part determine their effectiveness. PIM's main focus today is on everyday office working life, especially management of digital content and the broad principles of PIM such as the difficulties everyone has in managing information effectively apply also to scholars. However, there has been surprisingly little research specifically into how scholars manage their material (Palmer et al., 2009). One of the few exceptions is Kaye et al.'s (2006) fascinating study that developed a model in which different motives for

collecting were found to produce very different types of collection. But there might be other ways of conceiving of scholars' information collections. It is an interesting area because major changes have occurred over the last twenty years, that mean PIM has taken place in a context where the technical affordances have been revolutionised.

Furthermore, this area has become potentially more interesting in the last few years, because of growing concerns around research data management (RDM). Changes in the nature of research especially the emergence of large scale, collaborative e-research has directed attention to how research data is managed within the research cycle (Pryor, 2012; Pryor, 2014). Initial interest revolved around the impact of the "data deluge" in big science, but there is also concern about how data is created and managed in the long tail of small scale research. The study of research data, data scholarship, is an important emerging area of information science (Borgman, 2015). It is becoming recognised that good data management is a basic part of good research practice. Because of funders' requirements, many institutions have become very interested in how practical data management issues are being handled by researchers, hence institutional surveys usually contain questions about volumes and types of data, back up practices and so forth (Pryor, 2014). If we want to help to improve RDM, we should know more about how and why data is stored and created to understand where the critical problems lie. One cannot assume that "data", which is hard to define anyway, is managed separately from other material. As yet studies of RDM and digital curation have tended to neglect the wider context of PIM.

Thus investigating how research-related personal information collections (PICs) are created, maintained and used becomes increasingly important. It is also interesting to explore specific contextual factors that operate for scholars outside western research contexts, which tend to dominate the literature. Thus the study reported in this research investigated scholars' research-related PIM practices in Public Authority for Applied Education and Training (PAAET), a Higher Education Institution in Kuwait.

The paper is set out as follows: a section on related studies explains the role of research in universities and considers the core concepts of PIM and defines the PIC. Previous research about scholars' PIM is reviewed. The nature of the exploratory and qualitative methodology used in the study is then explained. The findings are organised around explaining how the PIC was created, what its main features are and investigating what factors shape this. A discussion considers the specific nature of the Kuwaiti experience.

Related studies

Research has become central to higher education institutions, key both to institutional and individual success (Scott, 2009). It is a core defining activity of many universities; it is also an important source of income. The pressure to do research and increasing culture of evaluation and performance measurement (Jarvis, 2014) or "performativity" (Fanghanel, 2012), creates a pressurised context within which scholars have to work. Productivity is closely monitored. Many of these pressures are manifested at the institutional level, but much of the scholarship on academic research focuses on disciplinary differences as key to understanding variations in scholarly practice (Becher and Trowler, 2001). Certainly the numerous studies of scholars' information seeking and information behaviour recognise great differences between researchers in the metadisciplines of science, humanities and social science, and also by individual discipline within these categories (Case, 2012). Research is a highly complex, non-linear process, a "complicated mix of mundane and seemingly idiosyncratic tasks" (Palmer et al., 2009: 3). Historically, information science has tended to focus on scholarly communication: the publication process and also forms of collaboration and

informal network, including citation patterns (Borgman, 2007). Scholars' seeking for secondary literature has been a recurrent concern in studies of information behaviour. Relatively speaking how scholars manage collections of material has been neglected, but the new interest in data may be a driving force to look more closely at the full complexity of information activities during the research process, including aspects of PIM.

Personal information management has been defined by Teevan as the "user's activities when they acquire, organize, retrieve, and process information in their own spaces" (Teevan et al., 2006: 68). These activities are carried out by a person to complete tasks, either work or non-work related (Jones and Teevan, 2007). According to Whittaker (2011) information consumption should be distinguished from curation. The consumer finds, uses and discards information; the curator, keeps, manages and exploits it. Curation is future oriented. A PIC arises from curation type PIM and it

> *is a collection of information sources and channels that we as individuals have acquired, cultivated, and organized over time and in response to a range of stimuli. The personal information collection is an organic and dynamic personal construct that we take with us into, and out of, the various information events that frame our daily working and personal lives* (Bruce: 2005:1).

A PIC is material kept for future use. People may keep information so that it is available later or simply as a reminder to do something. However, as Jones (2007:473) comments keeping is "difficult and error prone". It is easy to make filing errors and commit inconsistencies in classifying items. While cheaper computer memory has made keeping easier, consequent diversity of data types and fragmentation of material has made it harder to find kept information. Another paradox is that while people see document management as important, they try and spend the minimum time on it (Henderson, 2007).

Not surprisingly, PIM literature increasingly focuses on how people manage digital content, such as email, bookmarks and electronic documents. This takes place in a context of information abundance and problems created by the potential to amass large quantities of information, fragmented over different forms. It is recognised, however, that paper materials remain important (Whittaker and Hirschberg, 2001; Jervis and Masoodian, 2013). According to Whittaker and Hirschberg (2001) people continue to keep print documents for functional reasons such as ease of reading, for their use in collaboration, as a form of memory and reference, as reminders and also for "sentimental reasons" associated with some specific documents. Many of these documents are not unique and many are actually unread. Publically available material are kept because of distrust of such sources. A key issue remains coordinating print and electronic collections (Jervis and Masoodian, 2013).

There is surprisingly little literature explicitly on the PIM of scholars, given how information intensive their work is; most scholars do have a collection (Palmer et al., 2009). An early study by Case (1986) observed that scholars collect from many channels. Notes are also an important part of collections. Discipline was found to be significant, e.g. social scientists collections include more different types of material. This is a relatively isolated study focussed specifically on academics, yet scholars have often been one important part of the sample in wider studies of "knowledge workers", and much of the literature develops relevant insights to the understanding of scholars' behaviour. For example, a large proportion of Whittaker and Hirschberg's (2001) participants were "researchers". Henderson's (2009) study was conducted in a Business school, though it included non-academic staff. Similarly, Bondarenko and Janssen's (2005) sample included some PhD researchers, though it also included people in a wide range of types of office work. The latter authors identify some fundamental aspects of document management that are applicable. Document management and the basic management

of a task are tightly linked. Research tasks differ from administrative activities. There are fewer tasks, urgency is low or undefined and tasks are less repetitive, often unique. Most knowledge work is interrupted before it is completed, but it is less interrupted in research than administration. In research tasks there are also fewer types of document, the flow of documents is slow and they move out of use more slowly. The information in the document is more important than the document itself. The fragmentation of content across the desk, email and digital material is a fundamental management challenge. Printing helps people keep everything together. Visuo-spatial cues, such as the colour of a folder or where something is placed, are important to managing material.

An important exception to the lack of PIM studies explicitly about scholars is Kaye et al.'s (2006) study of 48 academics from a wide range of disciplines in an Ivy League American university. It encompassed all levels of seniority (from research students to professors) and scholars with a wide range of ethnic backgrounds. The principal contribution of the paper was to identify motives for collecting and link these to collection features. The main motives identified were:

- To find material again;
- To build a legacy;
- For resource sharing with others;
- To cope with fear of loss;
- To manage impressions people have of the individual.

These differing motives led to differently structured personal "archives". Thus legacy focuses on storage, while resource sharing implies retrieval by a large number of people. It would be interesting to reflect on the implications for how research data is managed in these different models. However, the implication in the paper that each actual archive was primarily shaped by a single one of these motives, rather than a complex mix, is arguably a bit surprising. Perhaps it would be better to see these as models, with actual collections reflecting a complex range of motives.

There have been a few other studies of scholars. Bussert et al. (2011) found that many scholars amassed large and fragmented collections. How material was discovered often influenced the format in which things were stored. Since the material is usually for own use organisation is idiosyncratic. Pikas (2007) studied the PIM of senior engineers in a research lab, and found they maintained significant PICs. She found they were both pilers and filers, but claimed to be able to find information they were looking for in their collection.

In attempting to build a framework for scholarly information practices as a whole, Palmer et al. (2009) have identified a number of core activities namely, searching, collecting, reading, writing and collaborating. Each are in turn split into specific sub-activities, or primitives. In addition, there are some cross cutting primitives, including notetaking and data practices. Searching for literature seems to retain its primacy in the framework. The approach complements lifecycle studies that seek to identify in particular cases or in particular disciplines how research itself and information practices within them are ordered and inter-linked in the research process, an approach particularly popular for examining the life of research data (e.g. Ball 2012, Higgins 2012). "Collecting" is sub-divided into the primitives of gathering and organising. The authors recognise that this area is under-researched (Palmer et al., 2009). Interestingly, Palmer et al. (2009) mention data gathering under collecting, even though they have a separate cross cutting theme for data activities, reflecting the difficulties in practice of unravelling the different primitives.

As the impression management aspect of the archive found by Kaye et al. (2006) reminds us, it must not be forgotten that as well as functional uses, information artefacts have a potential role in self-

presentation and the management of relations with others. Tian and Belk (2005) observed how personal possessions in workspaces are used to negotiate relations with other people. The academic's office is a setting for interactions with others and the objects in it part of the negotiation of these interactions. Though room size and position may be indicators of status (Harrison & Hutton, 2014) according to Belk and Watson (1998) displays in academics' offices tend to downplay status differences. Academics' offices are very highly personalised; they can be seen as expressions of academic freedom (Belk and Watson, 1998). They also play a role in mood management.

Research questions

In the context of this previous literature, the following research questions were defined for the project:

Question 1: How are scholars' research-related PICs created?

Question 2: What are the main characteristics of research-related PICs?

Question 3: What are the factors that shape them?

Methodology

Research context

The site selected for the study was the Public Authority for Applied Education and Training (PAAET), in the State of Kuwait. Kuwait is a small oil-rich state on the Gulf with a population, at the time of data collection (2012), of around 3 million people. Kuwait's education system developed rapidly in the 20th century. The constitution of 1962 embodied the right to state Education for all citizens. The first university was founded four years later. The PAAET was established as a public university in 1982 with the main objective to " supply the labour market with a national, technically skilled work force in the numbers that could meet the development needs of the country" (UNEVOC Network, 2012). This growing institution has nearly 40,000 students, 2,082 faculty members and 1,141 other staff (UNEVOC Network, 2012). It consists of twelve institutes teaching topics such as telecommunications and navigation, tourism, beauty and fashion. Although a vocational and technical training institution many of its scholars are actively engaged in research, since promotion is based on publishing performance. In April 2012 it was announced that PAAET was to become Jaber University of Applied Science.

Research approach

The study on which this paper is based 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 as they talk about their experience in their own words. The part of the study presented here, explored themes that had emerged from a series of earlier exploratory interviews with information professionals and scholars at the institution. These had revealed a context of scholars struggling with poor library services, this led to an interest in how secondary literature was collected and used through the research process. It then emerged that issues around use and finding of material was tightly linked to the wider issues around storing information. This prompted a further investigation of how secondary literature materials were managed, in the wider context of PICs. The study reported here turned to look in more detail at the research-related PIC.

Data

The data for this paper were interviews with 17 research active scholars based in the Departments of Basic Education, Health Science and the College of Nursing. Potential participants were invited on the basis of knowledge of the institution gained through earlier phases in the research and

examination of online profiles to ensure they were research active. A balance of participants in terms of seniority, experience, gender and nationality were chosen. This made it possible to consider the influence of such variables on PIM practices, though the numbers of participants overall preclude statistical generalisation.

Table 1: Overview of participants

Gender	Nationality	Seniority	Discipline	Place studied PhD	Date PhD	Years of experience
F	Non-K	Non-Senior	Health	Kuwait	1998	9
M	Kuwaiti	Non-S	Health	Egypt	1999	12
F	Non-K	Non-S	Health	USA	1996	40
M	Kuwaiti	Non-S	Health	UK	2005	11
M	Kuwaiti	Non-S	Edu.	Egypt	1996	15
M	Kuwaiti	Professor	Health	USA	1987	35
M	Kuwaiti	Professor	Health	UK	2002	15
M	Kuwaiti	Professor	Edu.	USA	2006	5
M	Kuwaiti	Non-S	Edu.	UK	2004	12
F	Non-K	Non-S	Health	UK	1995	10
M	Non-K	Non-S	Health	UK	2003	6
F	Kuwaiti	Non-S	Edu.	UK	2005	6
F	Kuwaiti	Non-S	Edu.	UK	2002	15
F	Kuwaiti	Non-S	Health	Kuwait	2004	6
M	Non-K	Non-S	Edu.	USA	1999	20
M	K	Professor	Edu.	UK	1999	11
F	K	Professor	Edu.	UK	2006	29

Interviews were held in interviewees' offices and followed an interview schedule with seven main questions and around 30 prompts (see appendix). The interview was built round a tour of the interviewee's office. With consent, photographs of the use of space and technology were taken as an additional form of data (Hartel and Thomson, 2011). Interviews lasted between 35 and 105 minutes, on average 60 minutes. Participants' voluntary informed consent was gained through explaining the project verbally and in a written information sheet; the research was cleared by Sheffield's ethic review process.

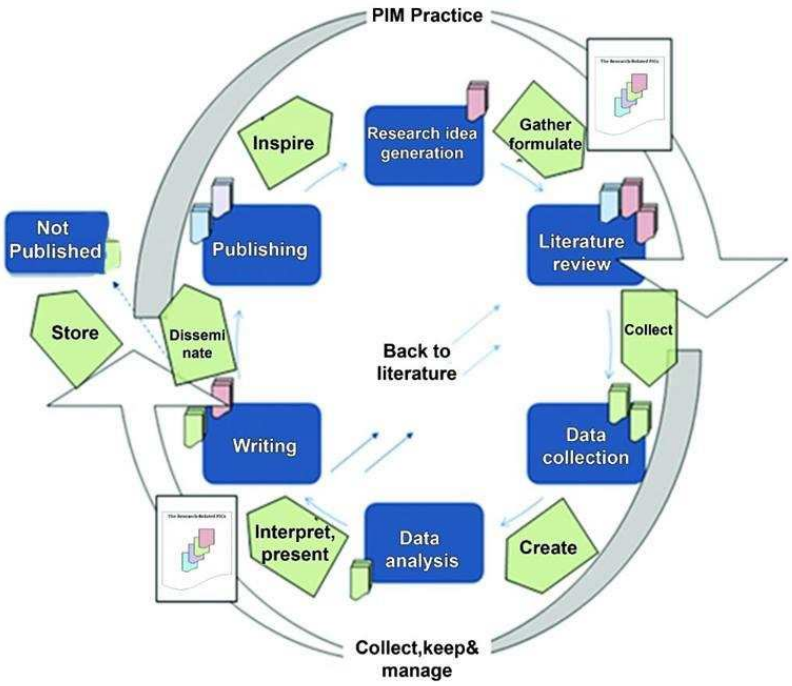
Analysis

The interviews were transcribed in Arabic and then, with the photos, analyzed thematically (Braun and Clarke 2006) to produce a list of codes (in English). Transcripts and photographs were sorted into "proto-themes" 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 final themes. Given that the sample was not representative of a wider population, in reporting results we do not consider the frequency with which something was mentioned in interviews as highly significant or useful to report.

Findings

The creation of the collection within the research lifecycle

Figure 1 presents an overview of how the research-related PIC is created. It seeks to indicate how items are continuously added to the collection throughout the process of research.



In some respects the pattern is that which would be expected: material is accumulated throughout the lifecycle of research, from idea creation through to publication. This cycle is non-linear, e.g. secondary literature is gathered for data collection and then consulted later in the cycle, to write the report on the research and perhaps to respond to reviewers' comments in the course of publication. Future projects may reuse material from a prior project. Some scholars also had a future ideas folder.

The model draws attention to important facets of the PIC which are perhaps obvious to any researcher, but important to state explicitly:

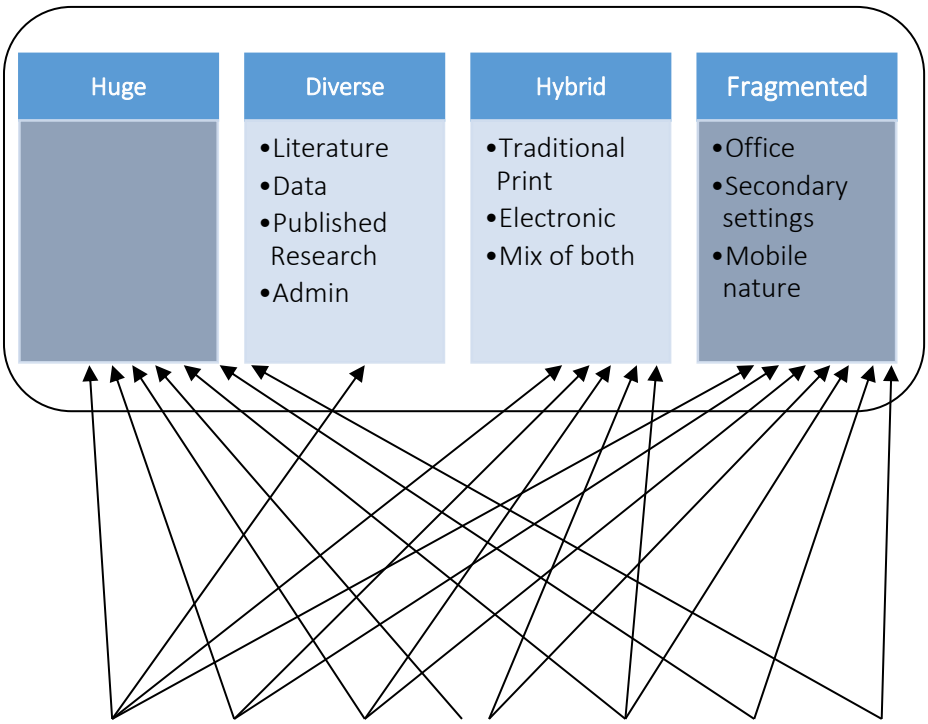
- Researchers work as individuals to manage their own material (even if some of their research is collaborative);
- Collection is an on-going activity;
- Material is organised by project;
- Within project materials, research data is filed with other material;
- Scholars tend to keep things, they usually do not discard material. Material is reused. Material is duplicated in multiple versions (e.g. of data or work in progress) and both print and electronic form. Even after publication all the material is kept.
- The research-related PIC is generally quite distinct from other material, such as teaching-related material.

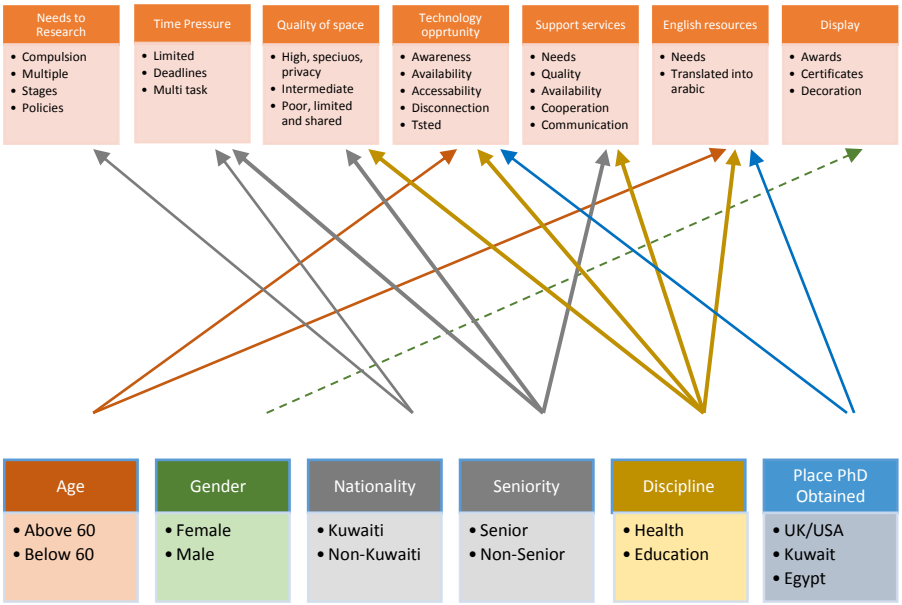
These basic features are very important to understanding the research-related PIC and data management practices in particular.

Characteristics of the collections

The research-related PIC that accumulates can be characterised by four features: huge scale, diversity, hybridity and fragmentation (Figure 2). These features can be linked to a set of key underlying factors, themselves shaped by demographic features of the population of scholars.

{Figure 2 here}





Huge in scale

Through a number of projects each scholar accumulates a lot of material. Enormous piles of printed papers were discovered in several locations in scholars' offices, such as on the desk, on a table near the desk, in storage units and inside the files and box files. Piles of information were also found towering in unusual places, such as on the floor. Envelopes, carrier bags, box files were used to file information. Very large piles were found in storage units outside of the main setting such as in store rooms.

> *As you can see I have loads of files as I am that kind of person who doesn't throw anything away. I don't delete electronic files and I don't throw hard copy versions away either. Even if they have been obsolete for ages I feel that I might need them so I keep them in a folder named Old Files instead of deleting them. (Education scholar)*

Because each project generated so much material, it could rarely be all filed together.

Diversity: four types of material

Creation of material within the lifecycle means there are four basic types of material in the research-related collection:

1. Secondary literature.
2. Research data – raw and analysed; in various formats, but mostly word, excel and SPSS. Raw data was not discarded even when comprehensive summary data existed through analysis.
3. Drafts of publications arising from the research and communications with the publisher.
4. Administrative paperwork associated with gaining institutional approval for the project.

Such diverse material is typically filed or piled together, by project. This mix of material is interesting, broader than found by Pikas in engineers' PICs (2006), reminding us that actual "research" contains elements of administrative tasks as well as research tasks as defined by Bondarenko and Janssen (2005): e.g. gaining consent for a project or dealing with the publisher. This helps to clarify one dimension of the PIM challenge for scholars: they deal simultaneously with tasks

(and related documents) that have different sorts of logic varying in terms of such features as urgency and repetitiveness. In particular it seems significant that research data was managed alongside other material. Researchers usually see keeping data safe as vital because it is unique and costly, if not impossible, to recreate. Yet, here at least, it was kept with a mix of other materials of a less sensitive character. This reflects a rather broad brush stroke approach to managing material. There did not seem to be a definitive form for the research data that could be archived; multiple versions, at different levels of processing and analysis, were kept and seemingly treated equally.

Hybridity

Collections are composed both of print and electronic material. Printed material remains very significant. Often there is duplication of the same material in both print and electronic form.

> *I keep my questionnaires in a brown (paper) envelope. I try to keep them next to each other, I wish to keep them all inside the green (plastic) folder but I'm afraid it's not big enough… actually my data related to my research are divided into two halves one is the hard copy as you may notice stored in the files or envelopes and the other half is on the computer stored on my computer as electronic files. (Health science scholar)*

Material was stored in multiple digital forms: on email, on computer hard drives and flash memory.

Fragmentation

The scholar's collection is physically fragmented across different spaces and in different formats. The main area of research-related work was the office, though because of gender segregation of teaching, some staff had two offices. Within the office there is a clear ordering within the space, with more active material closer to the work desk, less actively used material further away. Archived material may be elsewhere, such as in a departmental store room. But sometimes it was fragmented because physical artefacts worked as reminders.

> *They are all available in the store in the basement of the college. And I am keeping this much here to be honest for another reason: to remind myself about collecting data this year. (Health science scholar)*

Work was continued at home, when targets had not been reached in working hours. Elements of the collection were duplicated in both the office and at home. Some material was organised for mobility, thus one researcher had a lot of material in the boot of his car. Some people said they strongly preferred electronic material because it was more mobile, but there were still issues around the number of devices they had:

> *Yes this is another problem that makes it more complicated: I work on a PC at work and two other laptops at home with another PC. (Health science scholar)*

Scholars made considerable efforts to be organised in what was perceived to be a critical aspect of the research itself, confirming that for researchers (if not for others) collecting information is important (Henderson, 2009; Barreau and Nardi, 1995). Research is carried through via the documents and files that make up the research-related PIC. Scholars often had a file per project (and colour coded this) and tended to keep material in chronological order, with the most recent items at the top, within the file. Yet there was a lack of order of files and information within files; files contained all types of information. Physical filing broke down because it was not large enough for all relevant materials.

> *Here I keep them in a brown envelope. I try to keep them next to each other. I wish to keep them all inside the green folder but I'm afraid it's not big enough. (Health scholar)*

Further, scholars' responses in interviews reflected confusion and anxiety about the PIC. As in previous studies they were disappointed with their own management techniques, while managing to cope adequately (Kaye et al. 2006; Bussert et al. 2011). They often reported their own behaviour inaccurately, e.g. claiming to keep everything in electronic form, but in reality it was observed that they had print copies too. Although they kept secondary literature for later reuse, generally speaking they had difficulty refinding material and tended just to do the search again. Some of the behaviour, such as duplication of material, reflected a sense of anxiety around preserving access to material that did not necessarily result in efficient practices. There was anxiety around research data, because for these scholars, unlike the secondary literature in the PIC, if lost they could not be recreated without a lot of effort. This was probably reinforced because no special arrangements were being made for data as such. They yearned to tidy but could not often find time to do it. From an information management and information security perspective one can see multiple risks in terms of the confidentiality of data, dangers of loss of data loss and data integrity. The situation was strongly suggestive of need for some sort of intervention.

Key factors shaping PICs

Figure 2 summarises the main features of the scholars' research-related PICs and the key factors shaping them.

Pressure to do research

Scholars were required by the institution to do research. Career progress was premised on publication; it was also recognised as important to professional growth.

Time pressure in general

Like academics around the world, scholars faced time pressure to multitask, carrying out their research while also teaching and having administrative roles. A particular characteristic of Kuwaiti academic life is that official working hours are short. Although scholars could work any hours they chose in the office, most conformed to the culture of going home outside official work times. This meant that research had to be accomplished partly at home. Such "secondary" settings were important.

Quality of space available

The quality of space available to scholars was quite variable, partly linked to academic seniority. Because some teaching is gender segregated in Kuwait, scholars sometimes had two offices on the male and female campuses. Where they had limited space this tended to produce more reliance on working at home (and so physical dispersal) and more electronic material (and so hybridity).

> *But I work in several places: two work offices and a home office therefore I always make sure that my work is always saved on flash memory and sent by email to myself to make sure that I can access them from anywhere. (Health science scholar)*

Technology opportunity

Technology was perceived to be an opportunity to ease management of research-related material. Email was used to back up files. Yet the opportunities did not necessarily result in greater effectiveness, at least if measured by ability to refind information. The ease of searching and availability of information caused scholars to keep building up material in their collection without them having a chance to evaluate the information due to time pressure (a factor in PIM recognised by Whittaker and Hirschberg (2001)).

Lack of support from central services

At the time of the study, there were questions over the quality of service provision of electronic resources by the library at this institution. This lack of trust led to scholars to search for literature largely independently of the library collections; indeed they invested their own money in paying for access to research databases. This affected their attitude to material they found: having personally paid for material they were motivated to keep it. Scholars were also distrustful of university computing services, e.g. they used personal email not the institution's email service.

Need to gain English language resources

For most of the scholars interviewed, access to secondary literature in English was key. However, for some others, literature in translation was their main source. In these cases patterns of collecting were distinct.

> *I also usually visit any book fairs, mostly the one held here in Kuwait every year and I make sure to buy the most recent publications in Arabic or English but translated to Arabic because my English is not so good. I sometimes buy 10 – 15. Before I used to depend on the books and resources imported from Egypt but now I do not collect from Egypt with all my respect to them but there are few published works that you think of good value or simplified to our focus in the discipline. Therefore I collect others published for example from Syria, Iraq, Lebanon, as I might find good information. (Education scholar)*

Arabic material in electronic form was also rare: meaning the problem of integrating print and electronic was less evident.

Self-presentation and self-management

To a large extent research-related PICs were kept for the functional purpose of conducting research, and its key characteristics reflect this. However, in some cases how the collection was organised was influenced by other types of use. Material was used or positioned as a reminder to the scholar themselves to work on something.

> *You know I kept these papers here so that I can remember to take them with me in my bag and sometimes I keep my key on top of them. (Health science scholar)*

> *And I am keeping this much here to be honest for another reason, to remind myself about collecting the data for the year. (Health science scholar)*

In addition, for purely personal purposes, without any task-oriented reasons, many scholars liked to display their collections on open shelves rather than hiding them in closed cabinets or keeping them in other locations such as storerooms or other work places. Outputs of research (alongside awards for research) were often kept on display as a way of sharing. If someone visited the scholar's office, they loved to show them their collection, and they were proud of the unique nature of it. Displaying material was not only for welcoming visitors, but was also for sharing with colleagues, and presenting it as a gift for visitors. This perhaps does imply claims of status, unlike Belk and Watson (1998) who focussed on use of other types of more personal possessions and concluded that office use downplayed status differentials.

Demographic factors' impact on PIM

These factors were not experienced in the same way by all interviewees. Older scholars and those who had not done their PhD in a Western country (and who so had less English) tended to rely on Arabic sources which were personally accumulated, in print form. This resulted in large, rather unique, mostly print collections. They typically were less willing to make use of new technologies.

Non Kuwaiti scholars were under more pressure because their annual contractual review turned on meeting publication targets. Senior researchers were under more time pressure, on the other hand they also had more space and support.

There were not very great differences in behaviour by gender. Women did show more concern with comfort of space and attractive displays of outputs. There was also not very much difference between academics from education and health, perhaps partly because both groups tended to do questionnaire based survey research. One difference was that more space in health faculty buildings mitigated some of the problems of finding room for the mass of material; health scholars used technology more.

Discussion

Much of the picture painted in this study, could well apply to research work in any country: both the key features of collections and key factors in shaping them such as time pressure, the quality of space, technology opportunities, a balance between instrumental informational uses and self-presentation/ self-management. Many of the basic features of the collection echo findings of previous studies, e.g. Bussert et al.'s (2011) stress on fragmentation as a key problem for scholars and Kuntz's (2012) work on the impact of time-space compression on use of office space.

Some factors more specific to this institution or Kuwaiti academic culture and a "developing world" context were present, such as what was perceived to be poor infrastructure in terms of library support. However, trust in library support has often been found to be low in developed countries too (Corrall and Lester, 2013). The issue around the status of non-Kuwaiti researchers could be seen as a specific example of issues around seniority, encountered in other contexts. There were very distinct issues around culture of use of time; multiple spaces created by gender segregation; scholars' information seeking in areas where Arabic was more important than English language publications. This reminds us of the need to consider context specific factors as having a significant role in shaping PIM.

Certain key factors would probably complicate the picture if one was to repeat the study in a Western university: increasing large scale, multi-disciplinary and multi-institutional collaborative projects; the new managerialism; more trusted computing and library infrastructures; tightening of legal requirements through Data Protection and Freedom of Information legislation; increasing recognition of data management as part of good research practice, e.g. for replicability. This could make scholars' PIC work dramatically different. For example they might move to a more information consumption model and make dramatically less use of paper; or it could be that scholars in stable environments build collections around sustained long term interests, less around projects as found here. Complex patterns of return and reuse would significantly complicate collection processes. In intensive collaborative work with distant others patterns of managing material would also be likely to be different, though it is doubtful if people have really understood how to use cloud storage effectively in this context (Massey et al., 2014). However, one suspects that many of the results of the current study would be mirrored in many HEIs in "developed" countries.

Kaye et al. (2006) did not find discipline a significant factor shaping the type of collection kept – and only two disciplines were studied here – but we can hypothesise that it should still be considered likely to be a key factor (Case, 1986). Discipline (or sub-discipline) shapes what is the type and scale of data, research methodologies imply different cycles of data collection (in patterns that are different from that presented in Figure 1), some fields rely heavily on shared data within research groups or more widely, types of output are increasingly diverse too. From the model created in the research (Figure 2) it is reasonable to infer these factors would impact the research-related PIC.

Reflecting on Palmer et al.'s (2009) framework for scholarly information practices, the current research has contributed a little to filling in the gap in studies of "collecting". Zooming in on collecting it becomes apparent that "gathering" and "organising" seem rather broad brush stroke primitives. Organising is complex and re-finding and reusing need to be recognised as important aspects of collecting. This research has also suggested that data related activities are woven together with other research related activities at the level of PICs; the boundaries between scholarly primitives are complex. The present study fits in with the trend to examine the research process as a whole, somewhat decentering information seeking. Data practices and the production of outputs may come to be seen as more central information practices, in a more holistic account of research. Zooming in even more closely on particular parts of the collection would reveal their fluid and dynamic character. Data in particular are unfinished representations, temporary reifications of on-going processes, always in need of interpretation. For example, Garrett et al. (2012) found that for artists engaged in practice-based research both data and outputs are "moments of organisation" in a continuing flow of research activity, rather than stable entities. For this reason Borgman (2007, 2015) asks "when" is research data? This prompts us to begin to think about the full complexity of how information activities and primitives are organised during the research process and how collecting and organising fit into this picture. In a digital world artefacts in a collection can play multiple roles; their meaning is relational. As Huvilia et al. (2014) observe objects can take on a new meaning when the wider collection changes. The current study has only touched the surface of this complexity.

The study has sought to review the structural pressures that shape scholarly activities such as collecting. This is an element that seems to be needed to be modelled as the context for the Palmer framework or a lifecycle model, to explain the changing conditions under which scholarly information practices occur. The use of the home office in the context of work life balance issues needs further exploration (Thomson, 2013).

Unlike Kaye et al.'s (2006) study, this investigation focussed purely on research-related material. The study suggested that scholars do distinguish between a research-related PIC and other material they keep. Nevertheless, since they are kept in the same physical or virtual spaces there must be significant interactions between such collections. A fuller picture of scholars' PICs would need to trace the interaction between the research-related PIC and the teaching and other materials scholars accumulate. This could be particularly interesting in institutions that adopt research-led teaching pedagogies (Miller et al., 2012). Since it is recognised that the collection is not only for instrumental uses, but also used for self-presentation, how it is structured and used in the wider context of collections and display of other personal material in offices would also be worth investigation (Belk and Watson, 1998; Tian and Belk, 2005). This echoes Huvilia et al.'s (2014) suggestion that PIM can be better understood by recognising multiple factors or axes that influence behaviour.

Conclusion

In the context of the current concern among funders to ensure good RDM as a key aspect of good research practice, explorations of how scholars currently work are important. This applies not just to data intensive e-research; increasingly there is a concern with the good data governance in all research. The basic patterns in PICs and factors shaping them discovered in this study are plausible. There is an environment of information abundance and it is easier than before to store more information, yet that does not result in more organised collections. Improved PIM practices in general is key to good research through avoiding data loss and ensuring data integrity.

Given the centrality of research to their role and the centrality of material and digital artefacts in the PIC to research, scholars did not seem to be performing optimally. The study did not directly investigate whether and how research had been directly damaged by failures in PIM. But scholars interviewed expressed anxiety and were disappointed with their own performance. Like the respondents in Henderson's (2009) work they are likely to be willing to change.

Basic practical training about PIM to prepare scholars for a lifetime of collecting material seems to be needed. It is required to ensure efficiency in the research process, guarantee data integrity, make possible data sharing and should probably touch on the ownership of data. While the speed of technical change makes it hard to define which technologies might be appropriate to use, basic, generic principles of PIM, largely technology independent, are relevant to all researchers. Active services delivered by information professionals (or support teams trained in information management) from library services or embedded in department. More self-awareness and active planning seem to be needed. While people are motivated to perform tasks, and less to do background management of files that support tasks, the research-related PIC is so central to scholars' roles it seems plausible to suggest that the motivation to attend such training would exist. This can be reinforced by principles set out in funder and institutional good research practice guidelines.

The study also has potential implications for spatial design of academic workspace. Although this is where most scholars perform a life time of work, there is a relative lack of research about it (Harrison and Hutton, 2014). Increasing expectations of collaborative research as well as financial pressure is leading to an increasing trend towards shared office space (Pinder, 2008; Baldry and Barnes, 2012). In this context, and notwithstanding the increasing use of digital information, how scholars' PICs are accommodated in these redesigns will partly shape the success of such initiatives. Offering services proactively to help manage PICs could improve acceptance of more open working spaces.

References

Baldry, C. and Barnes, A. (2012), The open-plan academy: space, control and the undermining of professional identity. *Work employment society*, Vol.26 No.2, 228-245.

Ball, A. (2012). Review of Data Management Lifecycle Models. Bath: University of Bath. http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf (accessed 16-10-2015).

Barreau, D. and Nardi, B.A. (1995), Finding and reminding: file organization from the desktop. *SIGCHI Bulletin* Vol.27 No.3, pp.39-43.

Belk, R.W. and Watson, J.C. (1998), Material culture and the extended or unextended self in our university offices, *Advances in Consumer Research*, Vol. 25 No.1, pp.305-310.

Bondarenko, O. and Janssen, R. (2005), 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.

Borgman, C. (2007), *Scholarship in the digital age: Information, infrastructure and the internet*. MIT press, Cambridge (MASS).

Borgman, C. (2015), *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT press, Cambridge (MASS).

Braun, V. and Clarke, V., (2006), Using thematic analysis in psychology. *Qualitative Research in Psychology* Vol.3 No.2, pp.77-101.

Bussert, K., Chiang, K. and Tancheva, K. (2011), Personal management of scholarly information. In Foster, N.F., Clark, K., Tancheva, K. & Kilzer, R. (Ed). *Scholarly practice, participatory design and the extensible catalog*. ALA, Chicago, pp.123-151.

Bruce, H. (2005), Personal, anticipated information need, *Information Research*, Vol.10 No.3, http://www.informationr.net/ir/10-3/paper232.html (accessed 27-01-2015).

Case, D. O. (1986). Collection and organization of written information by social scientists and humanists: a review and exploratory study. *Journal of Information Science*, Vol. 12 No.3, pp.97-104.

Case, D.O. (2012), *Looking for information: A survey of research on information seeking, needs, and behavior*. Academic Press, San Diego, 2nd edition.

Corrall, S. & Lester, R. (2013), The researcher's view: context is critical. In Watson, L. *Better library and learning spaces: Projects, trends and ideas*. Facet, London, pp. 183-192.

Fanghanel, J. (2012), *Being an academic*. Routledge, London.

Garrett, L., Gramstadt, M., Burgess, R., Murtagh, J., Spalding, M. and Nadim, T. (2012) *JISC funded Kaptur project environmental assessment report. Project Report*. Visual Arts Data Service (VADS), a Research Centre of the University for the Creative Arts. http://www.vads.ac.uk/kaptur/outputs/ (accessed 16-10-2015).

Harrison, A. and Hutton, L. (2014), *Design for the changing educational landscape*. Routledge, London.

Hartel, J. and Thomson, L. (2011), Visual approaches and photography for the study of immediate information space. *Journal of the American Society for Information Science and Technology*, Vol.62 No.11, pp.2214-2224.

Henderson, S. (2009) *How Do People Manage Their Documents? An Empirical Investigation Into Personal Document Management Practices Among Knowledge Workers*. University of Auckland Doctoral Thesis.

Higgins, S. (2012) The lifecycle of data management. In Pryor, G. (Editor). *Managing research data*. Facet: London, pp17-45.

Huvila, I., Eriksen, J., Häusner, E. M., & Jansson, I. M. (2014). Continuum thinking and the contexts of personal information management. *Information Research*, Vol. 19 No.1 http://www.informationr.net/ir/19-1/paper604.html

Jarvis, D.S.L. (2014). Regulating higher education: Quality assurance and neo-liberal managerialism in higher education – a critical introduction. *Policy and Society*, Vol. 33 No.3, pp.155-166.

Jervis, M. and Masoodian, M. (2014), How do people attempt to integrate the management of their paper and electronic documents?, *Aslib Journal of Information Management*, Vol.66 No.2, pp.134 – 155.

Jones, W., (2008), Personal information management. *Annual Review of Information Science and Technology*, Vol.41 No.1, pp.453-504.

Jones, W. and Teevan, J., (2007), *Personal Information Management*. University of Washington Press, Seattle, WA.

Kaye, J. J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I. and Pinch, T. (2006), 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.

Kuntz, A. M. (2012). Reconsidering the workplace: faculty perceptions of their work and working environments. *Studies in Higher Education*, Vol. 37 No.7, pp.769-782.

Massey, C., Lennig, T., & Whittaker, S. (2014). Cloudy forecast: an exploration of the factors underlying shared repository use. *Proceedings of 32nd annual ACM conference on Human factors in computing systems*, (2461-2470).

Miller, A., Sharp, J. and Strong, J. (2012), *What is research-led teaching? Multi-disciplinary perspectives*. CREST. http://www.ukcge.ac.uk/pdf/CREST_What_is_research-led_teaching_web.pdf (accessed 27-01-2015).

Palmer, C., Teffeau, L. and Pirmann, C. (2009), *Scholarly Information Practices in the Online Environment Themes from the Literature and Implications for Library Service Development*. OCLC. http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf?urlm=162919 (accessed 27-01-2015).

Pikas, C. K. (2007), Personal information management strategies and tactics used by senior engineers. *Proceedings of the American Society for Information Science and Technology*, Vol.44 No.1, pp.1-21.

Pinder, J., Parkin, J., Austin, S., Duggan, F., Lansdale, M., Demian, P. Baguley, T. and Allenby, S. (2009), *The case for new academic workspace*. Loughborough University. https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/6037/1/AcademicWorkspace%20report.pdf (accessed 27-01-2015).

Pryor, G. (Ed.) (2012), *Managing Research Data*. Facet, London.

Pryor, G., Jones, S. and Whyte, A.(Ed.s) (2014), *Delivering research data management services*. Facet, London.

Scott, P. (2009), Foreword. In Brew, A. and Lucas, L. (Ed.s) *Academic research and researchers*. Maidenhead: Open University Press, pp.xiii-xviii.

Teevan, J., Jones, W. and Bederson, B. B. (2006), Personal information management. *Communications of the ACM*, Vol.49 No.1, pp.40-43.

Thomson, L. (2013). Situating Information Spaces: The "Office Continuum". In Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI.

Tian, K., and Belk, R. W. (2005), Extended self and possessions in the workplace. *Journal of Consumer Research*, Vol.32 No.2, pp.297-310.

Whittaker, S. and Hirschberg, J. (2001), The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction* (TOCHI), Vol.8 No.82, pp.150-170.

Whittaker, S., (2011), Personal information management: From information consumption to curation. *Annual Review of Information Science and Technology*, Vol. 45, pp.3-62.

UNEVOC Network. (2012), *The Public Authority for Applied Education and Training*[Online].UNESCO. http://www.unevoc.unesco.org/go.php?q=UNEVOC+Network+-+Centre&id=403 (accessed 27-01-2015).

Appendix: interview guide

Part 2: Interview Guide

1. Can you tell me about your research?

Related prompts:

- What area of research?
- The stages of conducting research?
- Using information in different stages of research?
- Research methods?

2. How do you keep track of your referencing in your research?

- Are you using any tool to organize your bibliographies like one called Endnote.. Have you heard about it?
- Have used any?

3. When conducting your last piece of research, can you tell me about the information resources used in that research?

- Where did you get them from?
- Did you use information for the first time or has it been used in other research before?
- Can you describe the way of finding information from your personal collections? Was it easy or difficult to find?

Related to working place (PSI):

4. Can you talk about this room, how much of your research you do in this room?

Related Prompts:

- Talk about the cabinets in the room? How many cabinets are there?
- How are the information organized in those cabinets?
- What are the types of information related to research in the room? (books, articles, … as listed above)
- Does the room contain Piles? or Files? Can you talk about them?
- Any of them related to the research? Format print or electronic?
- Can you talk about the material related to research in particular?
- Why do you keep this material here?
- Why do you keep it in this way?

- If some are related, then where do you keep the rest of each research?
- Can you show me some other research-related information in this room?
- Why do you keep them in this way?
- Why do you keep them in this place?
- Taking some evidence by picture and record the comments on that picture
- Do you keep backup copies of any information in this room in anywhere else?

Related to Personal Information Collection (PIC)

5. Can you please describe your personal information collections related to your research in this room?

- Why do you keep research-related information as part of your personal collections?
- Can you describe how are you storing and maintaining your personal collections to use them in the future?
- Can you talk about any incidents experienced when you tried to find information from your collections and you failed to find it? and what are the reasons?
- Can you talk about incidents experienced when you tried to find information from your collections and you succeed to find it?

6. Do you keep a directory of your collection?

7. Can you describe how you are using the stored collection in your research?

- Do you often search through your collections?
- Do you usually find what you search for from your collection?
- Do you find it useful for your research to build and maintain personal collections?