Before and After: User's Knowledge Maturity within Personal Information Management

Azrina Kamaruddin, Novia Admodisastro, Alan Dix

Abstract— Personal Information Management (PIM) is an important discipline that evolves with the technology advancements in their operating systems. The understanding of its four main activities in PIM: organizing, retrieving, deleting and archiving are crucial because the information growth in our desktop increased overtime. The paper presents the study conducted to enhance understanding of one of the causes for core phenomena – user's knowledge maturity which was identified in prior studies. The qualitative approach is taken to deepen our understanding on this cause. As a result, we found that several issues pertaining to time and users' task have increased their knowledge on organizing-keeping information on their desktop.

Index Terms— Users' knowledge maturity, Personal Information System, Keeping behavior, Qualitative analysis, Longititidal approach, Grounded theory

1 Introduction

Personal Information Management (PIM) is an activity that involves information acquisition, retrieving, maintaining and deleting. It has grown over time [1] due to the large increase in digital information that we encounter in our working and personal lives. Boardman [2] classifies it as an "umbrella" that managed digital information such as email, bookmark, files, meetings and appointments in a personal computing environment [3,4]. PIM is an important aspect in our life, where studies have reported that people always consider PIM as a chore [5,6,7,8,9] just like we possessively take control of our belongings. In order, to design tool that asissts users with their information on the desktop, we first need to acquire user's specification and to understand the issues regarding the PIM activities before a suitable design tool is proposed. The application and combination of Human Computer Interaction (HCI) discipline is important for Software Engineer to provide a user-centred design PIM tool for the intended users [9].

HCI defined as a set of "processes and resources that the users employ to interact with the computer" [10]. It is crucial to understand the users' experiences and behavior (a set of processes) when they interact with their computers/machines. In our study scope, we would like to investigate and understand the users' behavior exhibit within the PIM important activities. Several studies reported Malone [5], Lansdale [6], Barreau and Nardi [7], Whittaker and Sidner [8] and Jones et al. [9] discover that people are still struggling to manage their information on their desktop, file, email and bookmark even though they have been supported by latest technology. The challenges in PIM give impact to work productivity [7;11;12] and user experience [13].

- *Azrina Kamaruddin is pursuing a PhD degree program in Human Computer Interaction in Lancaster University, United Kingdom, PH-60389471763. E-mail: azrina@fsktm.upm.edu.my*
- *Novia Admodisastro is currently a lecturer in Information System department, Universiti Putra Malaysia, Malaysia, PH-60389471755. E-mail: novia@fsktm.upm.edu.my*
- *Alan Dix is a professor at Talis Birmingham, UK and University of Birmingham, UK, E-mail: alan@hcibook.com*

Early research works on PIM focus at particular domain such as hierarchy [14], email [15], and bookmark [16]. Gradually, several studies focused on interaction across our Personal Space of Information (PSI) have been conducted and reported such as [14;17;18;19]. However, there is little study on the keeping activities within PIM frameworks.

A study has been conducted as an extension of further investigating one (i.e. user's knowledge maturity) of the causes resulted from prior studies conducted. Table 1 shows a summary of studies that we had conducted to understand the core phenomena identified in re-analysed Study I. We have developed an understanding of keeping activities specifically the core phenomena that we named as the notion of dumping. The notion of dumping is defined as "a procrastination decision [20]". Both decision making and 'dumping' behavior are inter-related [20].

The objectives of this study are:

1. To observe user's maturity in knowledge related to their task.
2. To use the user's maturity knowledge resulted with their frequent task creates better understanding of their action especially managing their PIM.

Table 1. Summary of studies conducted

Study	Method/ Techniques	Analysis Technique	Outcome
Study I	• Interview • Observation • Capture images	Informal analysis	Build preliminary model of PIM user's behaviour
Study II	• Diary study • Simple questionnaire • Interview	Informal analysis	Understanding the issues and methodological
Re-analysed Study I	• Interview • Observation • Sketches	Grounded theory	Build core phenomena and its relationship model

Study III	• Interview • Observation • Partial longitudinal study • Capture images	Grounded theory – until axial coding	Detail understanding Elaborate on the causes model from main model

The paper presents the work that has been conducted in Study III only, which starts with introduction, literature review, methodology, analysis, results and conclusion.

2 Literature Review

There were many studies about information seeking related with domain knowledge such as domain knowledge on search tactic formulation [21], distinguished subject familiarity [22], topic knowledge of the search topic [23]. Recently, a study about finding and refinding based on the effect of task domain by Toms et al. [24] and factors and evaluation of refinding behaviours by Capra and Quinones [5;25]. However, there is yet little study about user's knowledge maturity in the information organization scope.

According to Capra [5;25] domain knowledge is part of knowledge where the user are familiar with the topic. He stated that the user only search information based on their domain in general. Another category of knowledge is stated by Capra [5;25] as task knowledge that measures knowledge about how to do a specific task. The factors that contribute to the task knowledge are familiarity and frequency. Capra [5;25] also reported in his thesis the effects of frequency and familiarity in refinding information.

Our study is about how the user's domain knowledge affects their information management understanding when their domain knowledge increase. We defined the users' knowledge maturity as the users' knowledge about information hierarchy on their desktop. Level of the user knowledge is changed before their knowledge of their task increased and after their knowledge of their task reached certain maturity level. The maturity supports their decision making upon organisation or retrieval of their information. Increasing in user knowledge also help them to decide about their information dumping either in certain location is only for temporary or information is suitable to categorise in specific category.

3 Methodology

This study was conducted nearly a year after the first and second studies. Throughout Study I and Study II, we identified few users clearly mentioned about their experiences and maturity that evolved when handling their new task. This indicates that the participants have increased understanding about their PIM. For example, Participant P002 from Study I reported that her first involvement in the study is as a first year PhD student which is new to the research domain.

The following section 3.1 describes in details the implementation of Study III.

3.1 Longitudinal study

Data collection for a given period of time can be either retrospective that is asking participant to reflect back upon their experiences and attitudes, or contemporary that is collecting data at different times about the current situation. The contemporary longitudinal study involves repeat follow-ups of a single sample, panel or cohort. This is more favoured approach in conducting Study III. Although retrospective data collection is an important part of any study, longitudinal research usually focuses on short-term reflection to avoid deterioration of reliability and validity when asking respondents to reflect back over long time times.

In addition a case study is designed as follow:

- We only select two participants from Study I.
- The two participants that we select are: one uses spring cleaner information management strategy and the other uses no filer information management strategy.
- These two participants used intensive dumping behaviour based on their information management strategy.

This case study consists of some aspects of longitudinal study in terms of time frame and changes over time in people behaviour [30]. The third study is conducted few years after the first and second studies. In order, to observe user knowledge maturity has an impact on user core phenomena, the notion of dumping, we purposely selected two similar users from different information management strategy to be Study III participants. The two participants were selected in Study III because in prior studies they were very co-operative and easy to communicate. We refer the two participants in Study III throughout this paper as S3-01 and S3-02. Further information about the participants is presented in section 3.2.

3.2 Naturalistic Approaches

In this case study, again we gathered information using naturalistic observation. We made the observations at their office desktop where they always do their jobs [31]. The observation on their desktop focused on their dumping files (e.g. folders, single information) that they mentioned earlier in the first study. However, the observations and interview were about their understanding and description of their files when they were more experiences with the task that they performed during study I. When interviewed, we could see the increased of knowledge about their desktop. Hence, they could provide reasons why they acted certain way when it they were at task versus time or urgency.

The interview has been conducted. The questions we asked were as follows:

- What is happening on their desktop (email and bookmark)?
- Why is the transition of desktop interface (email and bookmark)?
- When is the transition happen?
- When so they feel that they are getting matured in understanding their desktop?

- What happen to their information management before, during and after they achieved certain knowledge maturity.

Interview technique was selected when we wish to obtain more detailed and thorough information on the issue of user's maturity that might be gleaned from our GT analysis. We often prepared several questions, but also deviated where necessary in order to maximise the information obtained [26].We recorded our interview and transcribed the interview sheet. The study was conducted at the participants' workspace. Interview the participants was still our preferred approach because we wanted to acquire qualitative data [27] to describe changes in user's information management behaviour. The interview duration is between 1 – 1.5 hours. The images of users desktop screen was captured and compared with the screens captured in Study I.

We audio-recorded the interviewed and captured few images on the participant's computer screen. Then, we transcribed the interview. S3-01 is a non-native English speaker while S3-02 is a native English speaker. There was no translation made for the non-native English speaker transcription.

Participants Selection

The participants in Study III were taken from Study I. We decided to select two participants that we observed and identified from the first study as active participants. In Study I, the participants were referred as P002[1] and P019[2], whilst in Study III we referred them as S3-01 and S3-02 respectively. The participant S3-01 used spring cleaner, while the participant S3-02 used no filer information management strategies when managing information on his desktop.

a) Participants background

Study III was only conducted by interviewing two participants that we purposely selected from Study I. We decided to choose them as the participants because they were very responsive during Study I. Even though, we only have two participants, it would not give problem for us to understand the issues. In fact, we could study them properly and deeply [28]. It is important to know the participants' chronologies personal backgrounds before we proceed with data analysis process. Table 2 shows summary description of the participants.

Table 2. Study III participant descriptions

	Participants Id	
	S3-01	**S3-02**
Level of Competency	High competency	High competency
Information management strategy	Spring cleaner	No filer
Native	Non-English	English native
Sex	Female	Male
At time of Study III	Writing up PhD	Postdoc (data analysis task)
At time of Study I	2nd year PhD	Postdoc (literature review)

S3-01 was a female PhD student. She is also a mother and currently in writing stage of her PhD. She is an international postgraduate student and her native language is Bahasa Melayu. Her husband and children were all in UK. She is a valuable and main contributor to the Study I and Study II. She was in second year PhD when we interviewed her during Study I.

While, the participant S3-02 was a postdoc researcher. He was also our participant that involved in Study I. He had an experienced being a PhD student and this is an advantage for him. He knew that experiences and skills that he gained during PhD was not only about the research area but also skills and experiences managing information to complete the task. He was an international post-doc and his native language is English.

b) Participants Workspace

In this study, the important entity that we observed and analysed was their desktop. Both of our participants were at different stage when delivering their main tasks. S3-01 was nearly completed her PhD thesis whereas, S3-02 was at the analysis stage of his main task. While doing the interview session with them, we also observed their desk. They talked about their desk management especially participant S3-01. During the first interview, we saw her desk was occupied and scattered with papers, articles and books. It was a partially unsorted desk. However, when we observed her desk again in Study III, paper, articles and books were sorted and stacked nicely on her desk. She said that *"Ohm...that is where all information has been sorted and cleared because she had used and referred to that stack few months back, now, might not used it anymore".*

While, participant S3-02 desktop was quite tidy in both studies, however his desktop interface was dynamically changed due to his task. His desktop changed from a cluttered interface to a tidier and categorised interface. When questioned about it? He reported that, "*The reading and finding article stage has stopped and completed at this time, now is the time where you start to get your hands dirty with data. Once a while, you might need to refer to your paper and article so you just find it in the category. I already categorised them under appropriate label.*"

We interviewed both the participants at their workspace. The interview was recorded and transcribed. They were still at their workspace provided by their department. However, participant S3-02 moved to a new bigger office room, in which the room contains one experimental lab with glass window. Nevertheless, the things in the room were not much different from the one we observed in Study I.

While another participant S3-01 was at the same office room but the condition was much neater. This participant

[1] P002 referring to participant who used Window OS as her operating system.
[2] P019 referring to participant who used Mac as his operating system

desk was tidier, because she is nearly finished her thesis writing. Papers were sorted and stacked on her desk, she said that they were done. She reported that she categorised all her printed papers in the folder. The physical folder labels were similar with the digital folder version. The folders were nicely sorted in the folder. However, at the moment folder that was currently used was unordered.

4 STUDY III ANALYSIS PROCESS

The data analysis technique selected was GT. Given that, we have experiences with the technique, we decided to applied few stages in GT. We only interested with categories that related with the participants maturity in their knowledge. We started with open coding in GT or finding concepts and categories [29] that related with users' knowledge maturity. The process is about users' knowledge maturity with time when we integrate using axial code. In this analysis, we aim to:

- Understand users' knowledge maturity in information organization especially in the context of core phenomena of notion of dumping.
- The understanding is arranged towards time factor e.g. before maturity, during maturity and matured.

4.1 Results

The issues pertaining to time and users' task are important properties for users' knowledge maturity that cause the core phenomena the notion of dumping. In this section, we showed the in-depth understanding of users'knowledge maturity on how they behave before and after they achieved knowledge maturity.

Achieving Users' Knowledge Maturity

Users' knowledge maturity is very much related with time and users' task. Most of the participants in studies conducted in our research were PhD students varied in their intakes. Some were in their first year while some were finishing. Besides PhD students, the participants were academicians or researcher. Clearly, in our observation, almost all PhD students are having problems in managing their information on the desktop and adapting with their new role as a PhD student. According to both of Study III participants, they reported that when they were reaching their final stage in PhD they realized that they had discovered many skills such as increased their knowledge vocabulary within their new task and dynamic information management strategies. These are due to the participants increase in knowledge of understanding in their research area.

As a first year PhD student, there are many skills need to be learnt and a lot of adjustment to be made in the research area. Being a research student that dealt with many reading, assessing much information affected their information management. However, when they have reached maturity in their research area or task, their knowledge on information management has also increased. There is no course taught on how to manage information in PSI. The skills is known through participants experiencing themselves and as a result based on what they feel comfortable and applicable about it.

Next sub-section, we presented our analysis based on the users' knowledge maturity evolve with time.

a) Pending Categorisation

Categorizing information is one of the techniques primarily to keep information. However, pending Categorisation is when users are busy in delivering their primary task, categorising is the last option to do. Users' only focus in completing the task/sub-task rather than involving their effort and cognitive to categorise their information.

b) Early categorisation issues

When users are new to the task and knowledge domain, they have vague idea of how their hierarchy tree on their desktop will look like or how they will structure their files and folder due to their shallow knowledge about the new task. They admitted that, they have problems if they quickly categorised information but after they gained understanding of their new task and knowledge they allow retrospective categorisation. They have increased their vocabulary of the knowledge and also their folder labels.

At times, they have problems in retrieving information because they thought that they had moved the files into new folder. This uncertainty of being more knowledgeable is an issue when user is becoming dynamic in their personality pertaining to their categorisation of information. However, both of the participants are aware about their changes in making decisions on categorising their information at early stage or new to their task. When they became matured, they have the desire and tendency to change information in their old folder into new folder based on their maturity.

c) Maturity in Research Knowledge Change Decision Criteria

When users achieved certain level of knowledge maturity, they also have reached maturity in making decision. In addition, earlier technique that was used to make decision also changed or extended. For example, in Study I, one of the participants read the abstract of the paper to decide whether to save the paper or not. When, she becoming more knowledgeable in her research area, she decided that in order to save information, reading an abstract was not enough. She extended her reading technique from abstract section to discussion and conclusion sections in the paper. When, she decided the paper is relevant, then she continue reading the rest of the paper sections and finally save the paper.

d) Better Picture of Personal Space of Information

Getting matured in their research area or task has made the participants to give a better picture and understanding on what they are doing at that time (i.e. achieving maturity). At the beginning of their task and managing information, they do not have a clear picture of how the hierarchy structure will look like, and what information that they need to search, save and etc. When, they nearly achieved understanding of their task and knowledge they

know what they are doing. Even though, they dumped information in a partial category and in 'My Document' or desktop, they know what these information are and why they do that and later they know where the information should go.

One of the participants reported that, he was able to manage the way he did now because of the experiences he had during PhD study. Based on this experience, he applied the information management skill until now.

e) Early Categorisation Advantages

Categorisation makes our PSI looks tidy and organize. User with instant filer strategy definitely categorised their information once they acquired information. However, our participants are from no filer and spring cleaner strategies said that quickly categorisation would cause few problems when they reached maturity. So, some of the participants prefer to dump information and later make a decision of where it should be located/categorised. Their PSI looks cleaned and tidy.

f) Understanding Dumping Action

Before participants becoming matured in their knowledge or research area, their information management especially their information organisation may have been structured or unorganized as in dumping. At the early point of time, participants did not get a clear idea where exactly the information gathered need to be categorised. One of the participants uncertain about the location and did not have a clear picture of where it should be located.

However, when the maturity increased, they understood better their information and where it should be located. They also reported that they knew their reasons why they dumped information. Reasons are compliment with Study I [20] because they are running out of time, they are focusing on their primary task and by thinking of where and what label will distract their main focus on delivering their main task.

When they became knowledgeable in their research area, they knew exactly what is happening in their information management. As a result, they were becoming more knowledgeable on what are going on their PSI. Some of the behaviour that matured is: their judgement on keeping information increased and knew what to find, how to find and where to read in the content. Furthermore, they know that for example this information is only temporarily and later it will be properly categorised and kept. If they came back looking at their scattered information unfiled, they knew why and when it will get filed. In addition, they also know why they dumped the information and how they will organise the information later. As evidence, through our observation for the second time with the participants they can easily said without feeling uncertain about their information.

Excerpt: ...Katakan pasal friendship. Saya dulu tak ada folder pasal Friendship. Actually, I just can just create right away tapi is like I feel like tak ada masa jadi dump saja sini. Katakan saya ada 1 chapter cakap pasal Friendship. Jadi tengah-tengah tulis ke, tengah-tengah baca ke and then I need to get that information. Jadi saya tahu I dumped somewhere here that information [laughing]. So saya cari ajela kat sini. Saya tengok macam ni kalau benda tu tak susun lagi ke dalam folder maknanya saya dump kat sini semuanya dump kat sini, saya pun dump aje la kat sini. Tapi kalau misalnya kalau benda tu katakana dah ada folder. Saya masuk jelah tapi itu yang normal condition tapi kalau saya nak rush ke walaupun ada folder tapi dump juga dekat sini sebabnya nantilah baru buat housekeeping sekarang ni tak buat lagi. Apa tu yang saya dump aje la tapi macam kita tahu macam mana nak cari kat sini, aaa apa tu bila kita dump kat mana tapi mungkin kita tak tahu exactly laa... apa benda yang itu kena manually search la pulak..... (S3-1 non-native English speaker)

The translation:
...If it is about friendship. I don't have any folder about friendship. Actually, I can just create the folder Friendship straight away, but I just don't have time, I dumped the information regarding Friendship somewhere. When I need the information, if by browsing the hierarchy and if it is not there, I know I dumped somewhere outside that particular folder...

5 Conclusion

The study reveals that user's knowledge maturity affects user's PIM especially in the dumping and organisation scope. When users are becoming matured they not only understand more about their task and knowledge but also their information organisation. At this point, even though they dumped their information, they dumped in partially category where they think the information should reside. Later, they will re-categorise the information in a proper category. They also know the information dumped is for temporary purposes. They are uncertain to categorise at the early stage because they are not sure on how the information structure looks in their desktop.

References

[1] C. C. Marshall and W. Jones, Keeping Encountered Information. *Communication ACM*, 49(1): 66-67, 2006.

[2] R. Boardman, *Improving Tool Support for Personal Information Management*, PhD Thesis, Dept of Electrical and Electronic Engineering, Imperial College London, 2004.

[3] T. W. Malone, How Do People Organize Their Desks?: Implications for the Design of Office Information Systems, *ACM Transactions on Information Systems (TOIS)*, 1(1):99–112, 1983.

[4] O. Bergman, R. Beyth-Marom, and R. Nachmias, The User-Subjective Approach to Personal Information Management Systems, *Journal of the American Society for Information Science and Technology*, 54(9), 2003.

[5] R.G. Capra and M.A. Perez-Quinones, *Factors and Evaluation of Refinding Behaviors,* in Proceedings of the Workshop on Personal Information Management (SIGIR'06), 2006.

[6] M. Lansdale, The Psychology of Personal Information Management, *Applied Ergonomics*, 19(1):55–66, 1988.

[7] D. Barreau and B. A. Nardi. Finding and Reminding: File Organization from the Desktop, *ACM SIGCHI Bulletin*, 27(3):39–43, 1995.

[8] S. Whittaker and C. Sidner. *Email Overload: Exploring Personal Information Management of Email*, in Proceedings of the SIGCHI Conference On Human Factors in Computing Systems, pp. 276–283. ACM Press, 1996.

[9] W. Jones, H. Bruce, and S. Dumais, *Keeping Found Things Found on the Web*, in Proceedings of the 10th International Conference on Information and Knowledge Management, pp. 119–126. ACM Press, 2001.

[10] D. Te'eni, J.M. Carey and P. Zhang, *Developing Effective Organizational Information Systems,* Jones Wiley n Sons Publisher, 2006.

[11] A. Sellen and R. Harper, *The Myth of the Paperless Office,* MIT Press, 2001.

[12] W. Jones, *Finders, Keepers? The Present and Future Perfect in Support of Personal Information Management*, First Monday, 9(3), 2004, retrieved March 10, 2004 from http://www.firstmonday.org/issues/issue9_3/jones/index.html.

[13] V. Bellotti and I. Smith, *Informing the Design of an Information Management System with Iterative Fieldwork*, in Proceeding of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, 2000.

[14] R. Boardman, *Category Overlap between Hierarchies in User Workspace*, in Proceeding of the 8th IFIP TC 13 International Conference on Human-Computer Interaction (Interact), 2001.

[15] N. Ducheneaut and V. Bellotti, E-Mail as Habitat: An Exploration of Embedded Personal Information Management, *Interactions*, 8(5):30–38, 2001.

[16] R. Boardman, *Multiple Hierarchies in User Workspace*, in Proceeding of the CHI '01 extended abstracts on Human Factors in Computer Systems, pp. 403–404. ACM Press, 2001.

[17] D. Abrams, R. Baecker, and M. Chignell, *Information Archiving with Bookmarks: Personal Web Space Construction and Organization*, in Proceedings of the SIGCHI conference on Human factors in computing systems, pages 41–48. ACM Press, 1998.

[18] R. Boardman, *Workspaces that Work: The Need for Integrated Personal Information Management*, in Proceedings of the 6th Human Centred Technology Postgraduate Workshop, 2002.

[19] R. Boardman, *Workspaces that Work: Towards Unified Personal Information Management,* in Proceeding of the HCI2002, People and Computers XVI - Memorable yet Invisible, Volume 2, 2002.

[20] A.Kamaruddin and A. Dix, *To 'Dump' or Not To 'Dump': Changing, Supporting or Distracting Behavior,* in Proceedings of the International Confrence on iUser, pp. 49-54, 2010.

[21] M., Wildemuth. The Effects of Domain Knowledge on Search Tactic Formulation, *Journal of the American Society for Information Science and Technology*, vol. 55(3): 246-258, February, 2004.

[22] M.J. Bates, Factors Affecting Subject Catalog Search Success, *Journal of the American Society for Information Science*, vol. 28: 161-169. May, 1977.

[23] B.L. Allen, Cognitive Research in Information Science: Implications for Design. *Annual Review of Information Science and Technology*, 26:3–37, 1991.

[24] E.G., Toms et al., *The Effect of Task Domain on Search,* in Proceedings of the Conference of the Centre for Advanced Studies on Collaborative Research (CASCON'03), IBM Press, pp. 303-312, 2003.

[25] R.G., Capra, *An Investigation of Finding and Refinding Information on the Web*, PhD Thesis, Computer Science and Applications, Virginia Polytechnic Institute and State University Blacksburg, Virginia, February, 2006.

[26] A. Adams and A.L. Cox, "Questionnaires, In-Depth Interviews and Focus Groups", *in C. Paul and A.L. Cox (eds.), Research Methods for Human Computer Interaction*, Cambridge University Press, pp. 17–34, 2008.

[27] M.D, Myers, *Qualitative Research in Business & Management*, SAGE Publication Ltd., 2009.

[28] F.H. Abd Razak, *Single - Person Study: Methodological Issues,* PhD Dissertation, Computing Department, Lancaster University, Lancaster, United Kingdom, 2008.

[29] B.G. Glaser and A.L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine Publishing Company, 1967.