

**Factors that Influence Users to  
Keep and/or Leave Information Items:  
A Case Study of College Students'  
Personal Information Management Behavior**

Dissertation Proposal

by

Christina M. Finneran  
Ph.D. Candidate, Syracuse University

Advisor:  
Elizabeth D. Liddy

Submission for ISI Dissertation Proposal Scholarship

June 2008

# Christina M. Finneran's Dissertation Proposal

## TABLE OF CONTENTS

<b>Description of Research</b>	<b>1</b>
<b>References</b>	<b>11</b>
<b>Dissertation Schedule</b>	<b>13</b>
<b>Budget and Other Support</b>	<b>15</b>

# **Factors that Influence Users to Keep and/or Leave Information Items: A Case Study of College Students' Personal Information Management Behavior**

**Christina M. Finneran**  
School of Information Studies  
Syracuse University  
Syracuse, NY 13244  
cmfinner@syr.edu

## **Abstract**

This dissertation uses a multiple-embedded case study design to investigate Personal Information Management (PIM) behavior within the context of a large project, a college course. The theoretical framework introduces the hoarding literature, and associated affective factors, into the PIM domain to conceptualize users' keeping and leaving behavior. Interviews, direct observation, documentation, and physical artifacts will be used to identify the types of Personal Information Collections students use in a course and to uncover factors that influence their keeping and leaving behavior. Data reduction and analysis techniques will consist of matrices, content analysis and metaphor analysis.

The Personal Information Management (PIM) research stream investigates “the activities a person performs in order to acquire or create, store, organize, maintain, retrieve, use and distribute the information needed to meet life’s many goals (Jones, 2008).” Jones (2008) has developed the concept of a personal information collection (PIC), which is a subset of one’s entire personal space of information (PSI) limited by a particular location or form. “PICs are ‘islands’ in a PSI where people have made some conscious effort to control both the information that goes in and how this information is organized (Jones, 2008, p. 47).” PICs are typically comprised of information items in a particular place or format, e.g., email messages, files on a flash drive, a collection of books. For most project work, users make use of multiple PICs yet most of the empirical PIM literature has studied particular software applications or PICs in isolation. This investigation will take a user perspective to study PIM within the context of a specific project, a college course.

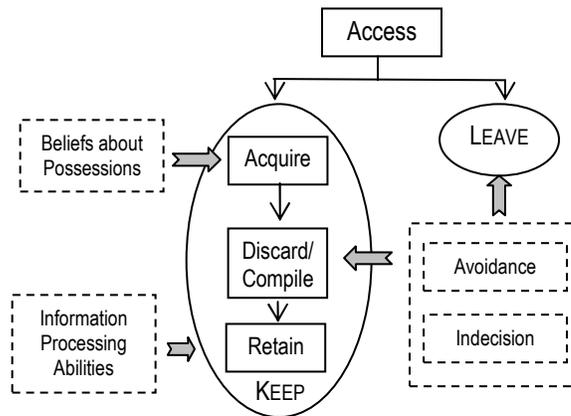
PIM research has primarily focused on how people manage their information: how users develop keeping structures (e.g., Bruce et al., 2004a; 2004b; Jones et al., 2001; Jones et al., 2002; Kwasnik, 1991; Lansdale, 1988) and how users find, or retrieve, items from their systems (e.g., Bruce et al., 2004b; Dumais et al., 2003; Teevan, Alvarado, Ackerman, & Karger, 2004).

Few studies have investigated the underlying motivation of why users decide to keep information in a personal collection given how cognitively taxing it is (Lansdale, 1988) and that an immediate benefit is not always evident (Cole, 1982). In their *Keeping Found Things Found* research, Bruce, Jones and Dumais (2004a) uncovered that in some cases, particularly with websites, users merely “leave” the resource where it is, and expect it will be easily located later if needed.

Though few studies have specifically addressed the motivation underlying these leaving and keeping behaviors, important research has investigated why users select one keeping method over another. Bruce, Jones, and Dumais (2004a) found that functions such as reminding and portability lead some users to particular keeping methods, confirming Malone’s (1983) seminal article. Whittaker and Hirschberg (2001) found that publicly available documents were kept because users had not yet processed and evaluated them, or because such documents enabled fast access, served as reminding devices, were of sentimental value, or were retained because of a distrust of centralized storage. In a Jones et al. (2005) study, project folders were not only used to organize a project but also to break down the steps of a project. Boardman and Sasse (2004) found that the degree of effort and time spent actively filing or managing personal information was influenced by the sense of ownership or investment in the item, the value the user placed on organization in general, and specific Personal Information Collections.

These factors provide some insight into why users may keep information items. In addition to the PIM literature, the hoarding literature can inform the motivation behind keeping and leaving information items, particularly with regard to affective influences. The theoretical framework (Figure 1) of this investigation has incorporated key concepts from the hoarding literature. One of the first contributions is the refinement of the “keeping” concept; it is broken down into the acquisition, discarding or compiling, and retention of information items.

The hoarding literature uncovers potential influences on keeping behavior. Frost and Hartl (1996) found hoarders to be hyper-sentimental about possessions. Their beliefs about possessions have been broken down into four components: Emotional Attachment to Possessions, Memory, Control over Possessions, and Responsibility for Possessions.



**Figure 1.** Theoretical Framework

Information processing deficits have also been associated with hoarding; they are attention, categorization, memory, and use of information to draw conclusions and make decisions (Steketee & Frost, 2003).

Potential influences on leaving and discarding also emerge from the hoarding literature. Frost and Hartl (1996) list emotional distress and avoidance behavior as an aspect of hoarding, along with indecision. Indecision relates to what to save, what to discard, and where to retain it (Steketee, Frost, and Kyrios, 2003).

This dissertation will identify and confirm factors, such as those suggested by the PIM and hoarding literatures, that influence the keeping and leaving of information items. For college courses, Learning Management Systems (LMS) play a particularly important role in the keeping and leaving of course materials. LMSs are specifically designed so that students can access online collections of course materials 24/7, and thus do not necessarily need to keep such items in their own personally managed collections.

The specific research questions for this investigation of students' PIM behavior are:

**RQ1. What Personal Information Collections do students use for their courses?**

- 1.a. What Personal Information Collections do students use for their course?
- 1.b. Among the students within a course, how uniform are the Personal Information Collections they use?

**RQ2. What factors influence college students' personal acquisition of course materials?**

- 2.a. What factors influence students to acquire course materials?
- 2.b. What factors influence students to not acquire course materials?
- 2.c. Which, if any, factors are consistent among the students within a course?

**RQ3. What factors influence the retention of course materials upon completion of the course?**

- 3.a. What factors influence students to save course items?
- 3.b. What factors influence students to discard course items?
- 3.c. How does the Personal Information Collection itself influence students' retention behavior?

**RQ4. What role does the Learning Management System (LMS) play in the course?**

- 4.a. What role do students perceive the LMS to play in the course?
- 4.b. How does the perceived role of the LMS influence a student's acquisition behavior?
- 4.c. Among the students within a course, how consistent is the perceived role of the LMS?

**Method**

Research shows that users' PIM behavior is different depending on the particular information collections or application (Boardman & Sasse, 2004; Bondarenko & Janssen, 2005; Whittaker & Hirschberg, 2001). The broader context of an intellectual activity, such as a college course, is likely to influence PIM behavior as well. Case studies enable the investigation of a phenomenon which can not easily be separated from its context (Yin, 2003). This study will use a multiple-embedded case study design to understand students' PIM behavior within a single

course. A college course will bound each case with a particular intellectual activity and the finite time period of a semester. Students enrolled in these courses will serve as embedded units.

Three cases will be undertaken, to be comprised of one college course within each of the humanities, social sciences, and sciences, as these areas generally make use of different types of learning, course materials, and evaluation techniques. A filtering survey of the enrolled students will be used to select five students, or embedded units, to study closely. Preference will be given to students who are majors in the subject area and in their sophomore year, and thus fully engaged in college and the course. For other individual differences, such as organizational skills and computer skills, a range will be sought among the selected students.

### ***Sources of Evidence (Data Collection)***

This investigation will use interviews, direct observation, documentation, and physical artifacts as sources of evidence. Data collection will occur at both the course and student level. In case studies, the expectation is that, during the data analysis phase, these multiple sources of evidence will converge (Yin, 2003).

At the course level, the goal of collecting multiple sources of evidence is to understand the course, its goals and requirements, and the way information and electronic communication are used and distributed within the course. This information will be used primarily to gain context for analyzing student behavior. Additionally, the data related to the courses will be used to do within-case and cross-case analysis for research questions. The multiple sources of evidence collected will be instructor interviews, direct observation, course documents, and course enrollment data.

The multiple sources of evidence collected about the students will be interviews, physical artifacts, and direct observation. Two semi-structured interviews with selected students, at different times in the semester, will cover not only his/her course materials with the student giving guided tours of their Personal Information Collections (PICs) related to the

course (e.g., binders, notebooks, computer files, flash drive, email, learning management system), but also the importance of the course from the student's perspective. Physical artifacts will be noted during the interviews and will include students' organization structures, such as paper binders or folders, electronic file folders, web browsers, and Learning Management System.

Direct observation will be done via the Learning Management System (LMS). Log data will be collected to measure how frequently students make use of the LMS for this course and how frequently they access particular course materials through it.

### ***Data Analysis***

The data will be analyzed inductively, building from the individual student to the multiple students within the course. First, data analysis of the embedded units (students) within each case will be conducted, then analysis for each course will be conducted, and finally cross-case analysis will be undertaken. The research questions are focused at different analytic levels within the embedded case structure of this investigation, as shown in Table 1. For example, RQ2a will focus on gathering factors first among individual students and then across the students within a course.

The student interviews will be transcribed and content analyzed for factors influencing the acquisition and retention of course materials, thus preserving the students' own words. Content analytic codes will be grouped by courses, contextual information gathered from the instructor interviews, and direct observations will be used to understand how the course may have influenced PIM behavior. The outcome of RQ2a, RQ2b, RQ3a and RQ3b will be a list of factors -- with their context and degree of influence -- that influence students' acquisition and retention of course materials.

**Table 1. Research Questions and Form of Results**

<b>Research Question</b>	<b>Analytical Level</b>	<b>Form of Findings</b>
1a. What PICs do students use for their course?	Within Student; Cross Students	List of PICs used, ordered by frequency of use among all students
1b. Among the students within a course, how uniform are the PICs they use?	Within Courses; Cross Courses	Ordered matrix of PICs by course
2a. What factors influence students to acquire course materials?	Within Student; Cross Students	List of acquisition factors with context
2b. What factors influence students to <u>not</u> acquire course materials?	Within Student; Cross Students	List of non-acquisition factors with context
2c. Which, if any, factors are consistent among the students within a course?	Within Courses; Cross Courses	Matrix of these factors by course
3a. What factors influence students to save course items?	Within Student; Cross Students	List of retention factors with context
3b. What factors influence students to discard course items?	Within Student; Cross Students	List of discard factors with context
3c. How does the PIC itself influence students' retention behavior?	Within Student; Cross Students	Matrix of these factors by PIC; Narrative description of retention behavior and PICs
4a. What role do students perceive the LMS to play in the course?	Within Student; Cross Students	List of dominant student metaphors
4b. How does the perceived role of the LMS influence a student's acquisition behavior?	Within Student; Cross Students	Narrative description of the LMS metaphors and students acquisition behavior
4c. Among the students within a course, how consistent is the perceived role of the LMS?	Within Courses; Cross Courses	Matrix of student metaphors and courses

Matrices enable the researcher to more easily identify important factors and to include contextual information and student quotes in the matrix cells. They will be used to answer RQ1b, RQ2c, RQ3c, RQ4c. Additionally, matrices will serve as a data reduction device (Miles & Huberman 1984) for each student case as shown in the Summary Matrix (Table 2).

**Table 2. Portion of Student Case (S32) Summary Matrix**

COURSE ITEMS	Use	Receive	Store/Access	Save?	Paper	Electronic file on computer	Email	Black-board	Other
Textbook	S32 reads the book, underlines important parts, writes clarifications/summaries in the margins, use post-it notes to write questions. She also brings the book to lecture, and marks which graphs were used in class.	bought	brings to class with her	Yes "Absolutely"	x				
Spiral notebook	Used for in-class notes. "Basically all relevant information finds its way into my notebook."	created excel file created in lab with instructions from the instructor.		Yes, noted that it would relate to her career field, which involves several more years of study beyond college.	x				
Tool for lab calculations	Excel file that has formulas for lab calculations. "there's a lot of answers, there's a lot of questions, um, the lab take-homes, the sections that you can do at home, and before lab, in pre-lab they tell you, like do the calculations down here and I'm betting that a lot of students saw that and said I'm not going to do them, I'm just going to do them by hand because, because I don't know how to set this up anymore. So I'm probably one of the few students who has it and can use it."	S32 copied it from the network file server locally to her laptop hard drive	on personal laptop "it's on my desktop. It's easily accessible. I don't like to search for things that I use quite often."	Yes "It will probably be there [hard drive]."		x			
Emails from instructor	S32 will ask clarification questions top professor via email	via email	in inbox with Mac Mail, can find them	No. Will delete "the information got transferred into my notebook already, so I don't really need them anymore. I keep them around mainly because it's easier to just type in <name> and go to the date, you know, the title, then search around my notebook for it."			x		
Wikipedia	Uses to help understand some concepts that are confusing in the book or for the homework. "I generally try not to use it more than once because I find it handicapping if you do that, you don't memorize it, so I use it kind of as the, 'Check my homework so when I turn it in I get the most points possible' kind of thing."		usually searches for image for a particular compound and then wikipedia will come up. Actually doesn't type into wikipedia directly. Reviews and learns, never prints out.	No.					website

Metaphor analysis will be used to discover metaphors that exist underneath the surface and enhance understanding of the learning management system (LMS), and how it influences students' PIM behavior. Metaphors can help the researcher in data reduction and finding patterns by abstracting each particular situation into a similar conceptual understanding (Miles & Huberman 1984). When possible, the researcher will use the second student interview to confirm and further refine metaphors that emerged from the initial interview.

## Limitations & Contributions

To ensure the rich development of the course cases and the embedded student cases, only three courses and five students within each course can be undertaken within the one semester of this study. The findings will be limited to college courses with similar contexts as those in the selected case courses. Further, extending the findings beyond the college course environment will be problematic. The primary limitation of this study is that it is exploratory in nature. In order to generalize the findings more broadly, further confirmatory studies will need to

be done. However, it is important to conduct exploratory studies to increase our understanding of PIM in the context of a user's project and with information items that a user leaves *in situ*.

This investigation contributes to the PIM literature by expanding the scope of information items included in empirical studies, as well as by developing a new theoretical framework and method. Historically, PIM empirical research has focused on what users keep and control locally; the boundaries between control and access, however, are becoming less clear and even less relevant to users who have 24/7 access to the Internet and portable wireless devices. The PIM literature has addressed this conceptually with Jones' (2008) broad definition of Personal Space of Information, although little empirical research has done so. This investigation begins to address this change in user behavior by focusing on not just what users keep, but also what they decide to leave. The findings of the study will provide an initial set of factors that influence keeping and leaving behavior.

In order to expand the PIM view to what users keep or leave, this dissertation examines the hoarding literature within clinical psychology to bring relevant aspects of such to the theoretical framework. As a result, this work introduces the hoarding literature, and its associated affective factors, into the PIM domain to conceptualize users' keeping and leaving behavior.

Methodologically, this dissertation has designed a naturalistic way to study PIM – within the context of a real-life project (college course), which is shared by multiple users yet experienced by each user individually. The multiple software applications and Personal Information Collections that users handle for the project will thusly be uncovered within their context of use. Additionally, the users are not merely interviewed and observed at a single point in time, given that over time users habits may drift or change (Balter, 1997; Boardman & Sasse, 2004). Users are interviewed twice during the semester and observed via the Learning Management System logs throughout the semester.

These methodological and theoretical contributions are important ones to push the PIM stream of research forward. Additionally, the discovery of influences on keeping and leaving behavior – ranging from those related to the technological tool, to the context of the project, to the individual's cognitive and affective needs – will further our understanding of PIM behavior in a constantly connected world, where 24/7 access may be akin to a local collection.

## References

- Balter, O. (1997). *Strategies for organizing email messages*. Paper presented at the HCI International, San Francisco, CA.
- Boardman, R., & Sasse, M. A. (2004). "Stuff goes into the computer and doesn't come out": A cross-tool study of Personal Information Management. Paper presented at the CHI 2004, Vienna, Austria.
- Bondarenko, O., & Janssen, R. (2005). *Documents at hand: Learning from paper to improve digital technologies*. Paper presented at the CHI 2005.
- Bruce, H. (2005). Personal anticipated information need. *Information Research*, 10(3).
- Bruce, H., Jones, W., & Dumais, S. (2004a). Information behaviour that keeps found things found. *Information Research*, 10(1).
- Bruce, H., Jones, W., & Dumais, S. (2004b). *Keeping and re-finding information on the Web: What do people do and what do they need?* Paper presented at the American Society for Information Science & Technology, Providence, RI.
- Cole, I. (1982). *Human aspects of office filing: Implications for the electronic office*. Paper presented at the Human Factors Society, Seattle, WA.
- Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. C. (2003, July 28-August 1). *Stuff I've seen: A system for personal information retrieval and re-use*. Paper presented at the SIGIR '03, Toronto, Canada.
- Frost, R.O., & Hartl, T. (1996) A cognitive-behavioral model of compulsive hoarding. *Behaviour Research and Therapy*, 34, 341-350.
- Jones, W. (2008). *Keeping found things found: The study and practice of personal information management*. Burlington, MA: Morgan Kaufman.
- Jones, W., Bruce, H., & Dumais, S. (2001, November 5-10). *Keeping found things found on the Web*. Paper presented at the CIKM'01, Atlanta, GA.
- Jones, W., Dumais, S., & Bruce, H. (2002). *Once found, what then? A study of "keeping" behaviors in personal use of Web information*. Paper presented at the American Society for Information Science & Technology, Philadelphia, PA.
- Jones, W., Phuwanartnurak, A. J., Gill, R., & Bruce, H. (2005). *Don't take my folders away!: Organizing personal information to get things done*. Paper presented at the CHI 2005, Portland, OR.
- Kwasnik, B. H. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation*, 47(4), 389-398.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19(1), 55-66.
- Malone, T. W. (1983). How do people organize their desks? Implications for the design of office information systems. *ACM Transactions on Office Information Systems*, 1(1), 99-112.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Steketee, G. & Frost, R.O. (2003). Compulsive hoarding: Current state of the research. *Clinical Psychology Review*, 23, 905-927.
- Steketee, G., Frost, R.O., & Kyrios, M. (2003). Cognitive aspects of compulsive hoarding. *Cognitive Therapy and Research*, 27, 463-479.

- Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). *The perfect search engine is not enough: A study of orienteering behavior in directed search*. Paper presented at the CHI 2004, Vienna, Austria.
- Whittaker, S., & Hirschberg, J. (2001). The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction*, 8(2), 150-170.
- Yin, R. K. (2003). *Case study research: Design and method* (Third ed. Vol. 5). Newbury Park, CA: Sage.

## Dissertation Schedule

### -----Data Collection Preparation -----

December 3, 2007: Submit IRB application to Syracuse University and Bowdoin College

December 3-17, 2007: Research potential courses that use online components

January 2, 2008: Invite Bowdoin College faculty to participate in study

### -----Data Collection Begins -----

January 14<sup>th</sup> - 21<sup>st</sup>: Conduct Instructor Interviews of 3 courses

*January 21<sup>st</sup> Classes Begin*

January 21<sup>st</sup>: First week of each course

Attend first course meeting, explain study and conduct filtering survey

Sign up for course email list

Set up tracking features on Learning Management System

January 28<sup>th</sup>: Input and Analyze the data from the filtering surveys of the 3 courses  
Gather Course Enrollment data from 3 courses

February 4<sup>th</sup>: Select potential students for participation in interviews

February 15<sup>th</sup> - April 9<sup>th</sup>: Schedule and conduct First Interviews of 5-6 students from each of the 3 courses

After each interview: Write notes, Create Student Summary Matrix, and send audio file for transcription

March 3<sup>rd</sup> – April 24<sup>th</sup> : Initial Coding of First Interviews on Acquisition and Retention behavior

*March 8<sup>th</sup>-24<sup>th</sup> Spring Break – Classes not in session*

Create PIE Summary Tables, Continue initial coding of first interviews

April 10<sup>th</sup> - May 10<sup>th</sup>: Conduct Second Interview with students

*May 7<sup>th</sup> Classes End*

May 16<sup>th</sup>:-May 30<sup>th</sup>: Conduct Instructor Debriefing Interviews of 3 case courses

*May 17<sup>th</sup> Exams and Semester End*

May 26<sup>th</sup>: Download sample digital course documents on Learning Management System

May 30<sup>th</sup>: Download semester log data from Learning Management System

----- Data Collection Ends -----

June 1<sup>st</sup> – July 15<sup>th</sup>: Student Case Analysis

- Update Individual Summary Matrices after Second Interview for each of the 16 students
- Review and remove personally identifying information from transcripts
- Review initial codes and create codebook for content analysis
- Content analyze student interviews for acquisition and retention behavior

July 15<sup>th</sup> – August 15<sup>th</sup>: Course Case Analysis

- Create Course PIE Matrices for each of the 3 courses
- Review instructor interviews
- Review course materials for each of the 3 courses
- Review enrollment data for each of the 3 courses
- Analyze usage data of the Learning Management System to determine the number of students using the LMS and the extent of their usage.
- Write up case summaries of each of the 3 courses

August 15-30<sup>th</sup>: Write-Up Results and Conclusions

----- Dissertation Draft Submitted -----

September 1, 2008: Submit Draft Dissertation to Advisor

October 13, 2008: Submit Draft Dissertation to Committee

November 3, 2008: Dissertation Defense

## Budget

### Expenses

	Item	Cost	Subtotal
<b>DATA COLLECTION</b>			
	Gift card incentive for 7 students to be interviewed for the pilot	7 @ \$25/each	\$175
	Gift card incentive for 16 students to be interviewed twice during the semester	16 @ \$20/each	\$320
	Transcription of the 32 (16 x 2 interviews) student interviews + 6 (3 x 2 interviews) faculty interviews	1240 minutes @ \$1.50/minute	\$1,860
<b>DATA ANALYSIS</b>			
	ATLAS ti software	\$175	\$175
	Intercoder Reliability Coders	12 hours @ \$15/hour	\$180
<b>PRINTING</b>			
	Copying of 138 page Dissertation Proposal document with spiral binding and hard stock cover	4 copies @ \$20/each	\$80
	Copying of 250 page Dissertation document with spiral binding and hard stock cover	8 copies @ \$25/each	\$200
<b>GRAND TOTAL</b>			<b>\$2,990</b>

### Other Support

Though I have a professional position at Bowdoin College within the Information Technology Division, none of my dissertation expenses are covered by my employer. I have 2 dependents for which I am responsible for, along with student loan payments.