How do graduate students manage their electronic information collections?: Investigating management activities and practices

SooJin Park
School of Library and Information Studies, University of Wisconsin-Madison
600 North Park Street, Madison, WI 53706, USA
park5@wisc.edu

ABSTRACT
As the amount of information people need to handle grows rapidly, it becomes more critical to efficiently manage information. To better understand the personal information management (PIM) process, this study examined key tasks and activities involved in managing a personal collection of electronic documents. Interviews were conducted with twelve doctoral students and data were analyzed qualitatively. The study helped identify management activities in three main tasks -- collecting, organizing, and storing, and three types (A, B, and C) of organizing/managing information collections.

Keywords
Personal information management (PIM), electronic collection management, information behavior, individual differences.

INTRODUCTION
As the amount of information people need to handle grows rapidly, it becomes more critical to efficiently manage information items. In response to the increased demand for personal information management (PIM), research has been done to understand how individuals manage their personal information collections (PICs); however, much work remains to be done (Jones, 2007). While existing studies have identified some strategies for, and patterns of, organizing and maintaining PICs, they are often criticized as being limited to certain applications (e.g., email, file management, or web bookmarks) or workplace settings (Boardman & Sasse, 2004). In addition, such studies have focused on artifacts or by-products of PIM, rather than specific activities that constitute PIM. As a result, how PIM is actually practiced has been overlooked. More in-depth research is needed to investigate what activities are involved in PIM, and how PIM is practiced.

This study is conducted to elicit a fuller picture of PIM practice. The study addresses two research questions: (1) What activities are individuals engaged in while building their information collections? and (2) How differently or similarly are these management activities executed in practice? During the summer of 2009, twelve doctoral students in a large public university were recruited and interviewed: four from the humanities and eight from the social sciences. At the time of data collection, four students were doing coursework, four students had finished their qualifying exams, three students were collecting data for their dissertations, and one student had completed writing a dissertation. Participants’ strategies for constructing and maintaining PICs were investigated. This study focused on personal electronic collections for writing projects in an academic setting.

LITERATURE REVIEW
Some researchers have proposed primary activities of PIM. Barreau (1995; adapted from Soergel 1985) listed the following as components of a PIM system: acquiring, organizing and storing, maintaining, retrieving, and producing information. More recently, Jones (2007) conceptualized PIM by identifying three activities -- finding, keeping, and managing (originally named Meta-level). These activities are similar to Barreau's but differently named and grouped. In this section, studies on PIM activities are reviewed with particular attention given to the keeping and managing activities.

Keeping
Jones (2007) explained the keeping decision is made based on "a rational analysis of alternatives" (p.473) and stated that people choose to keep something when the expected value of keeping it exceeds the value of not doing so. Bruce (2005) argued that a "personal, anticipated information need" is a critical value when creating, building, and managing information collections. Depending on decisions based on such a "personal, anticipated information need," people may take different actions, such as keeping useful information items, leaving (that is, doing nothing to keep) useful and accessible information, or ignoring useless information (Bruce, 2005; Jones, 2007). Details on such potential actions are illustrated in a study that looked at
keeping methods for web documents (Bruce, Jones, & Dumais, 2004).

Managing
Managing includes organizing and maintaining activities. Most studies on PICs management focused on organizing activity in the workplace. Malone (1983) initially identified two main strategies based on individuals’ organizing practices of paper documents in the office: pilers and filers. Focusing on electronic documents, Barreau (1995) investigated contents of managers' information management activities (including acquisition, organization, maintenance, and retrieval) in work place, based on Soergel’s model of information system (1985). Later, Barreau (2008) re-visited some of the previous participants and examined changes in their PIM. Khoo et al. (2007) investigated how people managed electronic files on their office computers and identified specific behaviors of filing, structuring, and maintaining. Studying university members, Henderson (2009) identified three strategies in participants’ filing practices: piling (disorganized), filing (more organized), and structuring (the most organized). In studies examining email and bookmarks, filing and cleaning were the main activities investigated in order to identify individuals’ management strategies. While previous studies identified some important characteristics of PIM, most of them did not provide a complete picture of PIM, because they were based on discretely focused management activities. To fill such gaps in PIM research, the present study aims at investigating all the PIM activities throughout the process of building and managing information collections.

FINDINGS
Management tasks and activities
To manage their information collections in electronic form, participants carried out three main tasks: (1) collecting, (2) organizing, and (3) storing. For the collecting task, participants were engaged in a series of activities that helped acquire and bring information items together, such as gathering, evaluating, format converting, and note taking. After collecting, participants moved on to the organizing task involving a group of activities that helped construct information collections, such as grouping, naming, structuring (expanding, combining, and ordering folders), and maintaining (deleting, re-grouping, and re-naming). The storing task came later and was comprised of activities for keeping information items from potential loss or damage, such as archiving and backing up. The three management tasks and specific activities under these tasks are summarized in Figure 1. While these management tasks look distinctive and separate, they are often interwoven in execution.

Types of information collection management
When management activities in which participants engaged were examined, it was found that some participants spent more time on the activities for the organizing task while others paid little or no attention to the same activities.

Based on the differences, three types (A, B, and C) of organizer for managing information collection are defined.

Type A organizers tended to focus on the organizing task and were very involved in all of the four organizing activities—grouping, naming, structuring, and maintaining. Collected information items were grouped based on various criteria. All of the type A organizers had their own systematic naming rules and consistently applied them. In structuring collections, type A organizers created folders to build or expand the hierarchical structure both horizontally (creating new folders or dividing existing folders at the same level) and vertically (sub-folding by creating or dividing folders in either child or parent levels). Type A organizers also developed the structure by combining folders. Type A organizers were also involved in maintaining activities. While most participants tended to avoid maintaining information items, type A organizers regularly deleted, re-grouped or re-named files and/or folders. Type A organizers considered information collection organization to be critical and useful to their present and future work.

Like type A, type B organizers considered information collection organization to be critically important. However, their organizing practices were limited to some structuring activities focusing on creating or dividing folders, but not combining folders. Furthermore, type B organizers were not involved in any maintaining activities. Although they recognized the usefulness of the combining and maintaining activities, type B organizers reported that they did not spend time on such activities due to limited resources (e.g., time, energy).

In contrast to the other two types, type C organizers tried to minimize their effort in the organizing task as much as possible. They conducted a few basic activities for the organizing task, such as grouping, naming and only creating folders for structuring and no maintaining. Type C organizers tended to use a small number of criteria for grouping. Type C organizers structured collections by adding new folders at the same level of hierarchy (horizontally). They neither divided nor combined folders. Folder structures were flat and shallow because subfolders were rarely created. Type C organizers did not have systematic naming rules. Type C organizers did not act on maintaining information collections at all. However, they
made a great effort in the storing task. Unlike the other two types, type C organizers considered information collection organization to be less useful. In order to allocate more of their “limited” resources to the project itself, they believed that minimizing their efforts on information collection organization was needed.

**DISCUSSION**

**Individual differences in PIC management**

Major individual differences were found in the way in which the organizing task was carried out, specifically structuring and maintaining information collections. In previous studies (Malone, 1983; Whittaker & Sidner, 1996; Bälter, 1997; Abrams, Baecker & Chignell, 1998; Boardman & Sasse, 2004), filing and cleaning activities were key to differentiating types of information organizers in managing files, emails or bookmarks. Depending on whether and when these filing and/or cleaning activities were executed (Whittaker & Sidner, 1996; Bälter, 1997; Abrams, Baecker, & Chignell, 1998), individuals’ styles of information organization were categorized. Furthermore, filing and cleaning activities have been viewed as simple as creating new and deleting old information items. This study, however, further refines the model. The activities of structuring and maintaining identified in this study and the components within these activities provide details in how filing and cleaning are carried out: filing is done via structuring activities (including creating, dividing, and combining files/folders); cleaning is done via maintaining activities (including deleting, re-grouping, and re-naming information items/collections).

Based on the details of structuring and maintaining activities, the process of collection organization can be illustrated further. That is, information collections are initially structured by creating folders. Their structures may be developed through horizontally/vertically expanding folders and/or combining them. The collections are maintained via deleting, re-grouping, and re-naming folders. Through this organization process, individual differences are more clearly delineated. While creating folders is common in collection organization, some participants selectively divide and/or combine folders. Even though many participants tend to keep their collections without doing any maintaining activities, some participants delete, re-group, or re-name old, useless, unrelated information items in order to keep their collections coherent. These identified differences in the process of structuring and maintaining collections distinguishes individual styles of information organization/management more accurately than those in outcomes of PIC management practices, such as file placement and folder structures examined in previous studies (Khoo et al., 2007; Hendersen, 2009). While the later reveals differences, their explanatory power is limited because outcomes cannot demonstrate or explain how the differences have resulted. In practice, file placement and folder structures could be similar, even though different activities are executed.

**Perceived usefulness of information organization**

The study findings reveal that perceived usefulness of information organization affects PIC organization/management practices and leads to the three types of organizers. Previous studies have identified the influence of perceived importance of “information” on differences in organization in various applications (Boardman & Sasse, 2004; Gwizdka & Chignell, 2007): less important information items, such as email messages, tend to be left unorganized than more important ones, such as files in personal computers (Boardman & Sasse, 2004). This study takes one step further to recognize perceived usefulness of “information organization” for accomplishing projects and examines the influence of the usefulness on personal strategies for collection organization. Depending on their perceptions of how useful information organization is for their projects, participants tended to execute management activities differently, especially the structuring and maintaining for the organizing task. The greater the usefulness of information organization to their projects, the more activities within the organizing tasks were executed in practice.

**Support for PIC organization/management practices**

The findings suggest that to improve practices, different methods of support should be offered to the three types of PIC organization/management (Gwizdka & Chignell, 2007). Support that helps to maximize organization/management activities should be provided to type A organizers. For instance, it can be useful to offer customization functions for saving files, creating and placing/structuring folders, retaining collections, and visualizing organization structures. Type B organizers can also benefit from the suggestions made for type A. In addition, for type B organizers, who want to do more work on their information collection organization, but whose practices are constrained by lack of time, assistance with assigning keywords or tags to information items and/or collections could be helpful as an alternative to developing organizational structures. Keywords or tags could better represent the contents of information items and/or collections and could be used for quickly grouping them and for later searching, if needed. Support for type C organizers would be anything that helps them minimize management activities, less structuring and maintaining activities. For instance, more systematic defaults for saving files and folder creation will be helpful for simplifying management activities so that type C organizers can feel they have saved more time and effort and have more time to focus on project accomplishment.

In addition, supporting the naming practice can be useful in general as it is used by all regardless of type. As consistent with previous studies (Kwasnik, 1991; Barreau, 1995, 2008; Khoo et al., 2007), most people in this study tended to assign descriptive names containing both document and contextual properties. In this study, topic, time, and people are identified as the three most commonly used facets in naming. Topic is a facet to represent the subject or “aboutness” properties of information items. Time and
people are used to illustrate contexts in which information items are used and managed. It will be useful to develop naming tools to suggest or incorporate these three facets. The tools can help include more descriptive information in names, support multiple grouping and listing in systematic ways, and also facilitate searching collections. This kind of naming tool can be useful for types A and B, assisting them with the descriptive aspects of PIC management. Such naming tools may also help the Type C organizers to improve information retrieval of less organized collections.

FUTURE STUDIES
The findings of this study suggest promising directions for future PIM research: further investigation into management activities and practices and evaluation of PIM. First, it is useful to investigate the complexity of PIM activities and/or tasks. Research questions could focus on what connections or influences exist among management activities and/or tasks, and how these connections are utilized in practice and whether they affect selecting/developing management strategies. Second, future studies should connect currently identified management activities with searching activities in order to encompass a broader sense of PIM. Future studies can investigate differences or similarities of searching strategies and practices shown by the three types of information management styles and examine what management activities could influence searching activities or vice versa. Third, research is needed to examine individuals’ identified patterns of management practices in different collections or contexts so that the dynamics of individuals’ PIC management practices can be more thoroughly investigated. Finally it is necessary to understand how PIM is evaluated. Future studies can examine what factors are important for PIM evaluation, what criteria are used for the evaluation, and how differently or similarly the three types of information organizers/managers apply criteria to the practice of evaluation. Such research can clarify different functions and utilities of PICs and PIM practices.

REFERENCES