The Process of Organizing Personal Information

Kyong Eun Oh
Rutgers, The State University of New Jersey
4 Huntington Street New Brunswick, NJ 08901, USA
keoh@eden.rutgers.edu

ABSTRACT

This paper presents preliminary results from an ongoing research study, which explores the process of organizing personal information from a cognitive sociological perspective. To collect data, a short background questionnaire, a diary study, and two post-diary semi-structured interviews were conducted, for each of the participants. The initial analysis of the results showed that there are five stages in the process of personal information organization. Each stage involved different actions, thoughts, decisions, and factors. The findings from this study will deepen our understanding about information organizing behavior and will contribute to the development and design of various personal information devices and applications that support individuals' organizing their information.

Keywords

Information organizing behavior, process of organizing information, personal information management, information organization, classification, categorization.

INTRODUCTION

Personal information is the information one keeps for personal use, either directly or indirectly. In everyday life, people extensively engage in organizing information files. However, while there have been a number of theories and critical findings that broaden our understanding about categorization, information behavior, and personal information organization, few studies examined the process of organizing personal information. Therefore, this research study explored the process of organizing personal information. Eventually, the researcher aims to develop a new model that explains the process of organizing personal information.

BACKGROUND AND RELATED WORK

While numerous previous works presented critical findings in understanding categorization, most of the early categorization theories examined categorization conceptually or investigated classification of objects in the natural world (Aristotle 2007 [B.C.350]; Berlin & Kay, 1969; Rosch, 1978; Wittgenstein, 1953; Zadeh, 1965).

In the case of information behavior research, a number of studies dealt with how people seek new information (Bates, 1989; Belkin, 1980; Kulthau, 1991; Savolainen, 1995; Wilson, 1997). However, fewer studies have focused on how people organize personal information (Jones & Teevan, 2007).

In addition, studies in personal information organization mainly focused on the end results of organization or a specific aspect of the process, rather than examining the process as a whole. To be more specific, the major findings from these studies have been to identify: (1) organizational structures of people's classification systems such as number of files, folders, size of each category, and the depth of those organizational structures (Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy, 2010; Gonçalves & Jorge, 2003; Henderson & Srinivasan, 2009); (2) different types of personal information objects (Barreau & Nardi, 1995; Cole, 1982); (3) different types of personal information organizing strategies (Bälter, 1997; Boardman & Sasse, 2004; Fisher, Bruxh, Gleve, & Smith, 2006; Malone, 1983; Whittaker & Sidner, 1996); (4) criteria that were used in organizing personal information (Barreau 1995, 2008; Case, 1991); and (5) factors that influence personal information organization (Barreau, 1995, 2008; Kwasnik, 1989, 1991).

As shown above, few studies explored the *process* of organizing information, and little is known about how these organizational structures are constructed, what decisions are made during the process, what is happening cognitively, and what factors impact people's grouping and separating of information items during the process of organizing personal information. Most importantly, social influence on the organization process has been not investigated, although personal organization is heavily influenced by society (Zerubavel, 1991; Zerubavel, 1996). Thus, the field needs research that holistically examines the information organizing process from a cognitive sociological perspective.

THEORETICAL FRAMEWORK

This research applies a cognitive sociological perspective to understand how people organize their information. The perspective of cognitive sociology is that the mind is social, and people think not only as individuals and as human beings, but also as social beings that are located in particular social environments (Brekhus, 2007; Zerubavel, 1997). In this view, classification is a social mental act that people perform as social beings. Examples of this social nature of classification are certain distinctions that are made in some societies, but not in others (Zerubavel, 1991). As a matter of fact, people frequently categorize certain objects very similarly to how others around them do. In particular, cognitive sociology uses the expressions "lumping" and "splitting" in explaining the process of category constructions (Zerubavel, 1996). In the process of "lumping", people group things together by focusing on similarities and overlooking differences, and in the process of "splitting", people separate things by exaggerating differences and ignoring similarities. In this view, the world is continuous, but people perceive it as discrete chunks through the mental process of lumping and splitting (Zerubavel, 1991; Zerubavel, 1996). In addition, this process of lumping and splitting is largely influenced by society, especially an individual's thought community, which is a community of people mutually exchanging ideas or maintaining intellectual interaction (Fleck, 1981). Thought communities shape people's perspectives and the ways people perceive the world around them, eventually causing members of those communities to perceive and organize things similarly (Brekhus, 2010; Zerubavel, 1997). This study uses this cognitive sociological view on categorization as a theoretical framework to explore the process of organizing personal information.

METHODS

For the participants of the study, a particular thought community (the academic community) was chosen. In addition, participants in different professional ages were selected because the length of time a person spent in a certain thought community can influence how one organizes information as a member of that thought community. Therefore, this study recruited 23 participants, consisting of 7-9 participants each from undergraduate students, graduate students, and professors. In collecting data, a short background questionnaire, a diary study, and two post hoc semi-structured interviews were conducted. More specifically, the researcher asked participants to record a diary over a week on a given template whenever they organized information in digital forms. Then, in the first interview, the researcher asked how and why participants organized information files as they did, based on the diary entries. Then, after 2-4 weeks, the researcher asked whether there had been any changes made to files or folders that were discussed in the first interview. Interviews were transcribed and coded with a set of categories based on the literature and the researcher's analysis of the literature. Then the researcher started to analyze data to

confirm, extend, and otherwise modify the initial categories. The initial model is presented in Figure 1. This model will be further investigated, modified, and expanded with more analysis of the data.

INITIAL FINDINGS

The initial analysis of the results showed that the process of organizing personal information consists of five stages: (1) initiation; (2) identification; (3) examination/comparison; (4) selection/creation/modification; and (5) categorization. In addition, each stage involved different actions, thoughts, decisions, and factors. It was found that social foundation of participants heavily influenced the whole process.

Initiation

In the first stage, participants initiated the organizing process when they had a file. In this stage, when files were not organized, they felt messiness and initiated organization in an effort to eliminate this messiness. In this stage, the primary decision was whether to organize a file into a category or not, and this decision was influenced by various factors such as future use of the file or number of related files.

Identification

Once participants decided to organize a file, they identified that file so that they could figure out how they should organize it. When identifying a file, typifications occurred. To be more specific, participants typified a certain file by ignoring the uniqueness of the file and regarding it as a typical member of a category. This stage was influenced by factors such as format, purpose, or source of the file.

Examination/Comparison

In this stage, participants examined existing categories to see whether they had relevant categories for an unorganized file. Then, participants compared this unorganized file with organized files in relevant categories so that they could decide where they should categorize the unorganized file. In this process, participants assessed similarities and differences between new and existing files. When examining whether they had a relevant category for the file or not, they considered the main purpose, format, and topic of the file when deciding about relevant categories.

Selection/Modification/Creation

When the participants found that they had a relevant category to organize the file, they selected that category from existing categories. If they did not have an appropriate category to organize that file, they modified the existing category or created a new category to organize that file. In this stage, when participants selected one of the existing categories, they adjusted the mental gap between new and existing files. However, when they failed to adjust this mental gap, they modified the existing category or created a new category for the file.

Categorization

In this stage, participants placed a file into a category. This stage involved several post-categorization decisions such as keeping, moving, re-categorizing, or deleting files. Those decisions were influenced by appropriateness or use of the file.

Social Influence

The influence of the thought community of participants was found in the whole process of organizing personal information. For instance, in the "Identification" stage, files were often recognized and typified by the primary tasks of the academics such as "teaching materials" or "papers to review". In addition, in the "Examination/Comparison" stage as well as "Selection/Modification/Categorization" stage, participants often assessed and adjusted mental gaps between new and existing files based on the academic uses of the file. For example, although both "dissertation survey" and "exam" files were in a Word file (format), created by the same person (source), and created in summer 2011 (time), they were separated from each other because their academic uses were different.

CONCLUSION

This study presented the preliminary results from an ongoing study that explores the process of organizing personal information in a digital form from a cognitive sociological perspective. This study is still in process, and it will be further modified and extended while the researcher analyzes more data. The findings from this will advance knowledge about people's information organization process, of which little is currently known. This research will also lay an empirical foundation for further study of information organizing behaviors. In particular, taking a cognitive sociological perspective has its unique contribution to the field. The results from this research will make direct contributions to the development of devices and interfaces that support individuals' organizing information.

REFERENCES

Aristotle. (2007). Categories (D.M. Edghill, Trans.). Adelaide, South Australia: eBooks@Adelaide, University of Adelaide Library. Available: http://etext.library.adelaide.edu.au/a/aristotle/categories/ (Original work written B.C. 350).

Bälter, O. (1997). Strategies for organizing email messages. In H. Thimbleby, B. O'Connail, & P. Thomas (Eds.), *Proceedings of the Conference of the British Computer Society Human Computer Interaction Specialist Group - People and Computers XII: Vol. 12.* (pp. 21-38). Bristol, UK: Springer.

Barreau, D. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science, 46*(5), 327-339.

Barreau, D. (2008). The persistence of behavior and form in the organization of personal information. *Journal of the American Society for Information Science and Technology, 59*(2), 307-317.

Barreau, D., & Nardi, B. (1995). Finding and reminding: File organization from the desktop. *ACM SIGCHI Bulletin, 27*(3), 39-43.

Bates, M. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407-424.

Belkin, N. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science, 5*, 133-143.

Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R., & Ramamoorthy, A. (2010). The effect of folder structure on personal file navigation. *Journal of the American Society for Information Science and Technology, 61*(12), 2426-2441.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

Boardman, R., & Sasse, M. (2004). "Stuff goes into the computer and doesn't come out" A cross-tool study of personal information management. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Austria, 6*, 583-590.

Brekhus, W. (2007). The Rutgers school: A Zerubavelian culturalist cognitive sociology. *European Journal of Social Theory, 10*, 448-464.

Case, D. (1991). Conceptual organization and retrieval of text by historians: The role of memory and metaphor. *Journal of the American Society for Information Science, 42*(9), 657-668.

Cole, I. (1982). Human aspects of office filing: implications for the electronic office. *Proceedings of the Human Factors Society*, USA, 59-63.

Fisher, D., Bruxh, A., Gleave, E., & Smith, V. (2006). Revisiting Whittaker & Sidner's "Email overload" ten years later. *Proceedings of the ACM CSCW Conference, Canada*, 309-312.

Fleck, L. (1981). *Genesis and development of a scientific fact*. Chicago, IL: University of Chicago Press.

Gonçalves, D., & Jorge, J. (2003). An empirical study of personal document spaces. *Proceedings of the International Workshop on Design Specification, and Verification of Interactive Systems*, 46-60.

Henderson, S., & Srinivasan, A. (2009). An empirical analysis of personal digital document structures. *Human Interface and the Management of Information, 56*(17), 394-403.

Jones, W. & Teevan, J. (2007). *Personal information management.* University of Washington Press, Seattle, WA.

Kuhlthau, C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information and Technology, 42*(5), 361-371.

Kwasnik, B. (1989). *The influence of context on classificatory behavior.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey, New Brunswick.

Kwasnik, B. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation, 47*(4), 389-398.

Malone, T. (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.

Rodden, K., & Wood, K. (2003). *How do people manage their digital photographs*? Paper presented at *the SIGCHI Conference on Human factors in computing systems*, Ft. Lauderdale, FL.

Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. Lloyd (Ed.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Savolainen, R. (1995). Everyday life information seeking: Approaching information seeking in the context of "way of life." *Library and Information Science Research, 17*, 259-294.

Wilson, T. (1997). Information behavior: And interdisciplinary perspective. *Information Processing and Management, 33*(4), 551-572.

Whittaker, S., & Sidner, C. (1996). Email overload: exploring personal information management of email. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Canada*, 276-283.
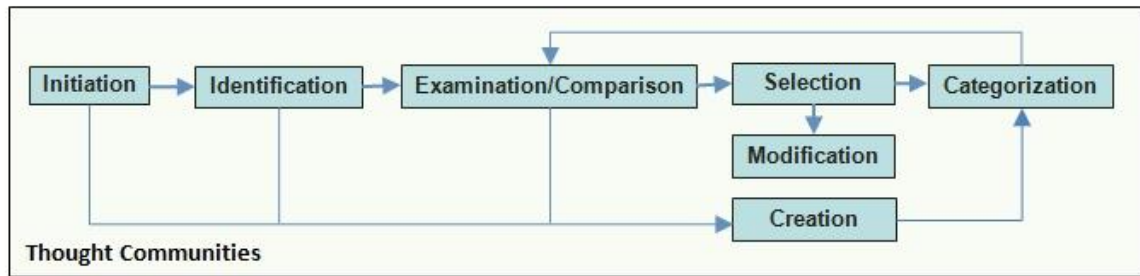
Wittgenstein, L., (1953). *Philosophical investigations* (G.E.M. Anscombe, Trans.). New York, NY: Macmillan (Original work published 1945).

Zadeh, L. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.

Zerubavel, E. (1991). *The fine line: Making distinctions in everyday life.* Chicago, IL: The University of Chicago Press.

Zerubavel, E. (1996). Lumping and splitting: Notes on social classification. *Sociological Forum, 11*(3), 421-433.

Zerubavel, E. (1997). *Social mindscapes: An invitation to cognitive sociology.* Cambridge, MA: Harvard University Press.

	Initiation	Identification	Examination/Comparison	Selection	Categorization
Action (Behavioral)	• Receive file • Download file • Create file	• Focus on certain aspect of file	• Examine existing categories • Compare unorganized file with organized file in relevant categories	• Select existing category • Modify previous category • Create a new category	• Place a file into a folder
Thought (Cognitive)	• Messiness	• Typification	• Remind existing categories • Assess similarities and differences between new and existing files	• Adjust mental gap between new and existing files	• Clean
Decision	• Organize or not	• Which aspect to focus on	• Whether relevant category exists or not	• Whether to select or modify existing category, or create a new category	• Keep • Move • Re-categorize • Delete
Factor	• Future use • Number of related files	• Purpose • Format • Subject • Source • Related person	• Purpose • Format • Subject • Related person	• Purpose • Format • Related person • Source • Number • Necessity of differentiation • Predicted size of the category	• Appropriateness • Use • Time

Figure 1. Personal Information Organizing Process (PIOP) model.