The User-Subjective Approach to Personal Information Management Systems

Ofer Bergman

Tel Aviv University, The Open University of Israel

E-mail: oferb@post.tau.ac.il

Ruth Beyth-Marom

The Open University of Israel

E-mail: ruthbm@oumail.openu.ac.il

Rafi Nachmias

Tel Aviv University

E-mail: nachmias@post.tau.ac.il

This is a preprint of an article published in Journal of the American Society for Information Science and Technology, 54 (9), 872-878. Final electronic version available at http://www3.interscience.wiley.com/cgi-bin/issuetoc?Type=DD&ID=104530821

Abstract

In this paper we suggest a user-subjective approach to Personal Information Management (PIM) system design. This approach advocates that PIM systems relate to the subjective value-added attributes that the user gives to the data stored in the PIM system. These attributes should facilitate system use: help the user find the information item again, recall it when needed and use it effectively in the next interaction with the item. Driven from the user-subjective approach are three generic principles which are described and discussed: (a) The subjective classification principle stating that all information items related to the same subjective topic should be classified together regardless of their technological format; (b) The subjective importance principle proposing that the subjective importance of information should determine its degree of visual salience and accessibility; and (c) The subjective context principle suggesting that information should be retrieved and viewed by the user in the same context in which it was previously used. We claim that these principles are only sporadically implemented in operating systems currently available on personal computers, and demonstrate alternatives for interface design.

Introduction

Personal Information Management (PIM) systems are information systems designed for individual personal use. Though PIM is discussed in the literature primarily in the context of its use in a work environment, its significance and applications are far broader.

The management of personal information is one of the central components of an individual's learning processes. While learning, people collect information items (e.g., books, notes, files, bookmarks, emails) and store them outside of their cognitive system (e.g., bookshelf, desktop). Usually an item is briefly reviewed by the person's cognitive system, and then cataloged, tagged and put aside for possible retrieval in the future. As the information to which a person is exposed expands, the necessity to store more and more information outside the cognitive system increases, as does the need for effective mechanisms for organizing, storing, retrieving and using this information (Barreau, 1995; Whittaker & Sidner, 1996). Effective retrieval has been referred to as the need for "keeping found things found" (Jones, Bruce, & Dumais, 2001).

One could expect the current sophisticated personal-computing systems to be able to offer solutions to this need, and supply individuals with efficient and friendly platforms for personal information management. In fact, operating systems and other PIM technologies try to do so. However, their design often ignores the fact that a PIM system is organized and thereafter used by one person: The same person who screens, classifies and then stores the information, is the one who retrieves it.

After a short literature review on the limitations of current PIM systems, we suggest, in the present paper, a user-subjective approach to PIM system design. This approach advocates that PIM systems should relate to the subjective value-added attributes given to the information by

the user. These should overcome some of the limitations and facilitate use of the system: they should help the user find the information item, recall it and use it effectively in the next encounter with it. Derived from the user-subjective approach are three generic principles which will be described and demonstrated: (a) The Subjective Classification Principle - all information items related to the same subjective topic should be classified together regardless of their technological format; (b) The Subjective Importance Principle - the subjective importance of information should determine its degree of visual salience and accessibility; and (c) The Subjective Context Principle - information should be retrieved and viewed by the user in the same context in which it was previously used. We will claim that these principles are only sporadically implemented in currently available PIM tools and operating systems. Alternative approaches to PIM technology interface design will be suggested. In this, we hope to contribute to the study of users' PIM system activities and to the design of effective tools that will empower individuals in their PIM activities.

PIM Systems: Some Limitations

Though intended to decrease the load on memory and facilitate retrieval, current PIM systems very often fail to do so. Users of PIM systems have difficulty remembering where they placed their personal information, be it in a real physical office (Malone, 1983) or in a computer environment as reflected in their files (Barreau, 1995; Barreau & Nardi, 1995), their email messages (Whittaker & Sidner, 1996), and important bookmarks (Abrams, Baecker & Chignell, 1998; Jones et al., 2001). These same studies show that many users of PIM systems do not use procedures available in the PIM system, but rely on alternative strategies, not necessarily more efficient: they pile up papers instead of filing them in the office environment, they keep hundreds of emails in their inbox instead of organizing them in folders, and make self-email links instead

of using bookmarks or the history option. As a result, users often express guilt feelings about being disorganized in such situations (Bellotti & Smith, 2000).

Why is it so difficult to “keep found things found” in a PIM system? A possible explanation is that PIM system designers fail to recognize the unique requirements of personal information management as compared to general information management (GIM). In his value-added model, Taylor (1982; 1986) suggests that GIM systems, which are created for many individuals, should rely on general attributes that are added to the information items to make them useful for users. Thus information professionals (e.g., librarians, website designers) should add attributes which are objective and general to cater the different needs of many users who vary in their profession, education, socio-cultural environment as well as the context in which they want to use the information. On the other hand, we believe that since a PIM system serves only one user, it should make use of subjective attributes; those that are given to the information by the user while interacting with it. These attributes are often episodic and idiosyncratic in nature, meaningful to the person himself but possible meaningless to others (Kidd, 1994; Sellen & Harper, 2001).

Kwasnik (1991) analyzed the descriptions of eight faculty members who were asked to describe how they organize documents in their offices. She found that a minority (30%) of the attributes described were document-related (e.g., author, form, topic, title), while the majority (70%) were attributes related to the interaction between the user and the information (e.g., situation attributes, disposition, time, cognitive state). Thus, users base their natural organization more on subjective attributes than on general “objective” ones. However, PIM systems are often not designed to allow friendly reliance on subjective information attributes, as we will show. The

user-subjective approach to PIM proposed here will suggest design principles aimed at addressing the above limitations by implementing users' subjective attributes in PIM systems.

The User-Subjective Approach to PIM System Design

Our user-subjective approach suggests that PIM systems should be able to accommodate the subjective value-added attributes that the user gives to the information in order to facilitate the user-system interaction: Help the user find the information item again, recall it, and use it effectively in the next encounter. This approach puts forward three generic principles that relate to three information-item attributes: The category to which the item belongs (the subjective classification principle); its importance (the subjective importance principle) and the context in which it was previously processed (the subjective context principle).

The Subjective Classification Principle

The subjective classification principle suggests that all information items related to the same subjective topic should be classified under the same category regardless of their technological format. A topic is a subjective value that is added to the information item by the user when storing the information. The same information item may relate to different topics for different users. For example a bookmark of an Internet site about the Rockfort caves could be classified under "Trip to France" for one person, under "Geology Course" for another and under "Food" for a third. PIM systems should help users classify their personal information items according to subjective topics regardless of the item's original general classification, and, more important, regardless of the item's technological format. Text files, emails, bookmarks, to-do lists and other information items belonging to the same subjective topic should be classified together. How this can be done will be discussed below.

The Subjective Importance Principle

The subjective importance principle suggests that information items should be characterized by their subjective importance[1] and that this attribute should determine their visual salience and accessibility. The importance of an information item is determined by the user relative to the importance of other information items. Subjective importance does not rest in the information itself, since what is priceless to one person can be worthless to another. In office environments, workers place information which they perceive as important where it can be seen (Kidd, 1994; Malone, 1983; Sellen & Harper, 2001). In the computer environment, when a user indicates that an information item, or part of it, is important to her or him, the system should react accordingly: such an item should be noticeable and accessible. Items with low subjective importance should be less eye-catching so as not to distract the user's attention. We will demonstrate below how to take into account what can be called the "deletion paradox": while information items with no subjective importance distract the user's attention and time, it takes time and attention to review them in order to make sure that they are no longer needed.

The Subjective Context Principle

The subjective context principle suggests that information should be retrieved and viewed by the user in the same context in which it was previously used. Research has shown that information is better recalled when it is stored in the context in which it was learned (Baddeley, 1982; Tulving & Thomson, 1973), and that learning depends on the situation in which it takes place (Brown, Collins & Duguid, 1989). The notion that information should be understood in its context, is widely accepted in the field of information science (e.g., Medin & Schaffer, 1978; Saracevic, 1999; Schamber, 1995) and in particular in the study of PIM systems (Barreau, 1995; Kwasnik, 1991; Lansdale, 1988). However, when using PCs as PIM systems, context does not

usually become an attribute of information items. Context should be captured and added to information items when saved for future use according to the subjective context principle.

These three principles are not independent: For example, the subjective classification principle and the subjective importance principle are context-related in the sense that how one classifies an information item and how important it seems are partly determined by its context. Furthermore, the decisions regarding the subjective importance of an information item may depend on the decisions regarding its classification. These interdependencies should affect future efforts aimed at implementing some of the ideas offered here.

The Implementation of the User-Subjective Approach in PIM Technology Interface

The user-subjective approach is generic in the sense that its principles can be implemented in various PIM technologies such as organizers, concept maps and web companions.[2] This section will demonstrate the implementation of the user-subjective approach on a PC's operating system, which serves as a PIM system when used to organize, save and retrieve personal information (Barreau, 1995). This system was chosen not only because it is the most common PIM technology, but also because it serves as a platform for various technologies that deal with personal information (e.g., word processors and browsers). When implementing the subjective approach in the PC operating system, we aim to apply it to these other technologies as well.

Like all systems, PIM systems can be designed to utilize various degrees of automation (Shneiderman, 1998). On one pole of the automation scale is the adaptive approach which holds that a system should keep track of a user's performance and adapt its behavior to suit the user's needs. On the other end of the scale is the direct manipulation approach that calls for user control over the behavior of the system. In implementing the subject-oriented approach in PIM systems, there is a temptation to use the adaptive approach because, as we will show, many subjective

attributes can be automatically deduced from the user's interaction with the information, thus relieving the user of much technical work. However, Shneiderman (1998) warns us that extreme adaptive systems could surprise the user and change the system in ways that s/he doesn't expect. The user may not be able to interpret what has happened, to predict the next change, or to restore the system to its earlier state. Thus, designers of PIM systems need to find a balance between automation and user control for each specific implementation of the user-subjective principles.

Implementation of the Subjective Classification Principle

Handling an information item in a general information management system forces users to acquaint themselves with a given classification system, and use their knowledge of the system when retrieving information items. In an effort to ease the cognitive load on the user, Taylor (1986) suggested a user-driven model in information system design (complementary to the technology-driven and the content-driven models), in which system design is based on an analysis of user characteristics:

> In a common-sense way, we do know a fair amount about the information behavior of different groups, for example, managers, engineers, lawyers, and students. We have not, however, developed suitable methods of organizing those data; nor have we ways of translating those characteristics into system design. This process is inhibited in part by our dependence upon technology as a solution to all problems, and to our tradition of content or subject analysis as the sole means of access for all clients or users (pp. 25-26).

A personal information management system should be certainly based on a user-driven model; however, not necessarily one determined by the designer for a specific group of many users but by the individual user for personal use. The personal classification scheme created by individuals will, after its implementation, enable easy classification and retrieval: users will store

the information in a place that “makes sense” to them, allowing for quick retrieval when necessary.

The most frequently used PIM technology -- the PC operating system -- is technology-driven and therefore suggests a technological rather than a content- or user-derived classification. The PC’s hard drive is divided first according to the technological formats of the information item and only then by personal topics. Take, for example, the case of Jane, a chemistry student using a Windows operating system: Word document files she created for her chemistry course assignments are located in her “My Documents” folder; bookmarks of Web pages that she used to gather information for the assignments are located elsewhere in her browser; emails she exchanged with her professor and classmates are in her mailbox folders, and her assignments are located elsewhere in a separate to-do list.

According to the subjective classification principle, information items that relate to the same subjective theme (chosen and defined by the individual) should be clustered together regardless of their technological format. Such organization would allow the user to see all the project related information items together, and make the search for a specific information item much easier, and involve fewer distractions from irrelevant items. Some applications, such as the MS Binder, Aladdin System’s DragStrip and Raton Laveur (Bellotti & Smith, 2000)[3] are designed to support clustering of information items of different technologies according to topics that the user determines. However, to use these applications, one needs to index the information items intentionally after they have already been arranged in a different way in their folders, a process that requires time and cognitive effort. Most users do not use these applications, and thus their PIM systems afford no connection between information items which differ in their technology format.

According to the subjective classification principle, information items should be labeled, used, stored and retrieved according to a personal classification scheme. Jane could classify all the items relevant to her chemistry course under "Chemistry Course" and put all relevant items (be they bookmarks, email messages or word processing documents) into that category. To make the system even more friendly and efficient for the individual, the operating system might then be able to adopt an adaptive approach and designate a labeled topic as "active," thus giving its items priority in the process of labeling, storing and retrieving information. Such an operating system feature, might increase the accessibility of all information items belonging to the project on which the user is working at a given time.[4]

Implementation of the Subjective Importance Principle

Most GIM systems do not include a feature that allows the user to add an importance tag to stored information since what is important to one person is not necessarily important to another. However, when exposed to new information, the user usually determines how important that information is. Thus, PIM systems should be able to convey the importance of an information item as perceived by the user. Subjectively important items should be easily accessible and noticeable, while subjectively irrelevant and unimportant information items should not distract users.

High subjective importance. The way current operating systems deal with highly important information is based on either a direct manipulation approach (e.g., allowing the user to put the most important files and folders on the desktop) or the adaptive approach (e.g., keeping track of recent documents and allowing the user to return to them using the History feature). Adopting an adaptive approach to highly important information items requires an operational definition of subjective importance such as : "What was dealt with most recently is the most important" or

"What was frequently dealt with, within a fixed period of time, is important". As previously mentioned, highly important information items (determined automatically by the system or intentionally by the user) should also be noticeable to the user. Perhaps files should be listed by default according to how recently they were accessed, rather than by the file name which is the current default listing.[5]

Within an information item, the user often needs to specify which sections are more important and which are less important since rapid and effortless accessibility to the important sections may be desired. Here a combination of the adaptive approach and the direct one can be applied. According to the direct approach, the user needs to specify which sections are subjectively more important. S/he can do this explicitly by highlighting these sections, or implicitly by annotating them (on the assumption that sections that are worth writing about are subjectively more important). According to the adaptive approach, the system could automatically create an index to allow easy access to these sections. Several technologies allow the user to do this for different technological mediums (for example, the Adobe Acrobat creates such index for electronically published text (Klemes, Epstein & Illovitch, 2002), Synopsus for digital movies (Bergman, Beyth-Marom, Hadar & Dekel, 2000; Dekel & Bergman, 2000), and MIXIM Web Raveler for web pages (Kensler & Rebelsky, 2000)). These technologies allow the user to mark and annotate sections of information items while automatically creating an index on the assumption that the user will wish to return directly to these sections.

Low subjective importance. In the past, the main problem with unimportant information was that it occupied highly valuable storage space. Today, when storage space is no longer a major obstacle, it seems that the main problem that relates to unimportant information is a cognitive one: when searching for a specific information item, irrelevant and unimportant items

distract the user's attention. One way to avoid this is by deleting outdated or irrelevant information. However, deciding what to delete requires attention and is time consuming and is often repeatedly postponed in the belief that "I'll do it when I am not so busy". Barreau (1995), who interviewed managers about their PIM system use, noted that their maintenance seems to be guided by the philosophy "if it ain't broke, don't fix it". As a result, critical inspection of one's folders so as to decide to delete whatever is no longer necessary is often performed only at times of crisis such as a hard disk crash. This behavior reflects what we call the "deletion paradox": In order to decide which information items should be set aside or deleted to avoid distracting attention during the next encounter, full attention is needed. An adaptive operating system design should be able to set aside old unused items without denying the user the access to them when needed. Files that have not been accessed for a long time (according to a time threshold defined by the user) can be manipulated graphically by making them smaller or faded in color

Implementation of the Subjective Context Principle

The subjective context principle suggests that PIM technologies should be designed to relate to the context of an information item during interaction with it. The context will help the user to recall the information when it is next needed. Contextual characteristics may be external, internal or temporal.

External context. The external context of an information item refers to the other items that the user dealt with while interacting with a specific information item. The linkage between all the information items that were activated while working on a specific item may be viewed as that item's working environment. The PC operating system should be able to create a link between items in that environment. In this way, when viewing an information item again, less effort is needed to reconstruct the mental processing involved in its creation, and the user will suffer from

less memory load, since other relevant information items will be accessible from the item's working environment.

For example a user creates a word processing document and copies information from a Web page source and pastes it into the document. The item's working environment, as perceived by the user when creating the document, should be preserved so that when the user accesses the document again, its working environment (the Web page) is accessible. However, in current PCs, linkage between information items (e.g., the Word document and the Web page) is not stored in the system. Thus, the user has no straightforward way of returning to the entire relevant working environment.

Internal context. Internal context relates to the user's thoughts while interacting with the information item. In most encounters with an information item, there is some cognitive processing on the part of the user: the item triggers thoughts relevant to the item, responses concerning its relevance, significance, reliability, its association to other items, questions, etc. All these constitute the internal context of an information item and contribute to the construction of new or revised knowledge (as defined by the constructivist point of view).

When re-accessing an item, it makes no sense to return to the item as it was originally perceived in the user's first encounter with it, but the user should be able to return to its "revised/reconstructed" form, including the earlier individual responses to it. When saved, these cognitive processes have a positive effect on information recall (Kiewra, 1985).

PIM technologies and, most importantly, the operating system should allow users to write annotations about whatever they read, to automatically provide easy access to these annotations, and to present them in the context of the original information. Current PC operating systems do not offer such general capability.

Temporal context. Temporal context pertains to the state in which the user left the information item when s/he last interacted with it, and to his or her working plans regarding that information. Efficiency in the working environment of the PIM system might improve if users were able to trace their previous steps and to plan further ones. Subjects of email messages are in bold font if an email has not been opened and in normal font if it was screened. Links are blue when they haven't yet been accessed and turn purple when activated. In both cases, the design is adaptive: the technology automatically changes the font's attributes in response to user interaction with the information. These temporal attributes given to information items allow users to trace their past activity and to know, when returning to an email message or a webpage, whether or not they have already interacted with it.

To complement this, we suggest that technologies should also allow users to mark the information items that they plan to work with in the future, such as email messages which users plan to answer later, and links within a text that will be accessed only after completing the reading of the text. Such marking will enable users not to forget email messages which were temporarily ignored and will allow them to read a text that contains links without interruption.

Summary

In his value-added model, Taylor (1986) discusses information professionals who add value to data in GIM systems to enable friendly and useful activity by GIM users. In adding these values, information professionals need to consider users because it is they who will benefit from the added value when retrieving and using the information. In the case of PIM systems, information is acquired, organized, stored, and later retrieved, recalled and reused by the same person. Thus, the added value of information items clearly needs to be subjective -- relating to the individual user. The user-subjective approach suggests that PIM systems should capture the

value-added attributes relating to previous user-information interactions so that users can utilize these in their next interaction. Such attributes, while useful and meaningful to the user, are often worthless and insignificant to others.

Three user-subjective principles were defined and exemplified in this paper: the Subjective Classification Principle, the Subjective Importance Principle and the Subjective Context Principle. Although they don't require sophisticated technology or a complex interface for their implementation, they rarely appear in PIM technologies. To allow for such implementation, software designers should apply the user-subjective approach and its specific principles in planning and designing interfaces for PIM technologies.

Research should aim to test empirically the user subjective approach suggested: it should inverstigate how users describe their PIM system and the way they behave in their PIM environment. We predict that users will talk about subjective attributes of their personal information but will not necessarily be able to use these attributes when working with their PIM technologies because of their limited design. Research should also examine whether, and how, implementation of the user subjective principles affect user-information interaction and users' information processing (e.g., learning and memory). Research results should guide interface designers in their contribution to the construction of user-friendly and efficient PIM technologies.

References

Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: Personal web space construction and organization. Proceedings of CHI '98 (pp. 41-48). Los Angeles CA: ACM Press.

Ahituv, N., & Neumann, S. (1986). Principles of information systems for management (2nd ed.). Dubuque, IA: Wm. C. Brown Publishers.

Baddeley, A. D. (1982). Domains of recollection. Psychological Review, 89 (6), 708-729.

Barreau, D. K. (1995). Context as a factor in Personal Information Management systems. Journal of the American Society for Information Science, 46, 327-339.

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

Bellotti, V., & Smith, I. (2000). Informing the design of an information management system with iterative fieldwork. Symposium on Designing Interactive Systems, 2000, 227-237.

Bergman, O., Beyth-Marom, R., Hadar, D., & Dekel, A. (2000, June). From “learning-by-viewing” to “learning-by-doing”: A video annotation educational technology tool. Paper presented at ED-MEDIA 2000, Montreal, Quebec, Canada.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, (January-February), 33-42.

Dekel, A. & Bergman, O. (2000). Synopsus: A personal summary tool for video. Extended Abstracts of CH’2000 (pp. 4-5). ACM Press.

Jones, W., Bruce, H., & Dumais, S. (2001). Keeping found things found on the web. Proceedings of CIKM’2001, 119-126.

Kensler, C., & Rebelsky, S. A. (2000). Web raveler: An infrastructure for transforming hypermedia. Proceedings of Ed-Media 2000 (pp. 479-484). Norfolk, VA: Association for the Advancement of Computing in Education.

Kidd, A. (1994). The marks are on the information worker. Proceedings of CHI'94 (pp. 186-191). Boston, MA: ACM Press.

Kiewra, K. A. (1985). Investigating notetaking and review: A depth of processing alternative. Educational Psychologist, 20 (1), 23-32.

Klemes, J., Epstein, A., & Illovitch, T. (2002). Computer-assisted instruction and multimedia for students with learning disabilities at the Open University of Israel. Proceedings of Ed-Media 2002 (p. 52). Norfolk, VA: Association for the Advancement of Computing in Education.

Kwasnik, B. H. (1991). The importance of factors that are not document attributes in the organization of personal documents. Journal of Documentation, 47, 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 desk? Implications for the design of office information systems. ACM Trans. on Office Information Systems, 1 (1), 99-112.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Saracevic, T. (1999). Information science. Journal of the American Society for Information Science, 50, 1051-1063.

Schamber, L. (1995). A user-based cognitive approach to modeling highly dynamic information problem situations. Proceedings of the 58th Annual Meeting of the American Society for Information Science, 32, 157-162.

Sellen, A. J., & Harper, R. (2001). The myth of the paperless office. Cambridge, MA and London: MIT Press.

Shneiderman, B. (1998). Designing the user interface (3rd ed.). Reading, MA: Addison-Wesley.

Taylor, R. S. (1982). Value-added processes in the information life cycle. Journal of the American Society for Information Science, 33, 341-346.

Taylor, R. S. (1986). Value-added processes in information systems. Norwood, NJ: Ables.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80 (5), 359-380.

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. Proceedings of CHI'96 conference of Computer Human Interaction (pp. 276-283). New York: ACM Press.

Author Note

We would like to thank Amnon Levav, Dr. David Mioduser, Edna Yafe, Ainatte Inbal, and Shlomit Zarchovitz for their helpful comments. This paper was written as part of Bergman's Ph.D. studies at Tel Aviv University.

Correspondence concerning this article should be addressed to Ofer Bergman, The Knowledge Technology Lab, School of Education, Tel Aviv University, Tel Aviv 69978 Israel, or by e-mail to: oferb@post.tau.ac.il.

Notes

1. The term “subjective importance of information” used for PIM systems should not be confused with the term “subjective value of information” which is used in general information management systems and measures how much money the user is willing to pay for the information (Ahituv & Neumann, 1986).

2. A detailed description of PIM tools is available at http://www.ms.lt/ms/projects/toolkinds/organize.html

3. Raton Laveur (Bellotti & Smith, 2000) has a unique implementation of the subjective classification principle. Information items of different technologies are presented in the form of an email viewer. These items can be grouped together for functional reasons such as a meeting. Instead of a topic, the group has a representative item that leads to the rest of the items in the group. Each of the items can also be defined as a to-do item, conforming to the subjective (temporal) context principle discussed below.

4. Patent pending.

5. Notice also that in order to find a file by name one needs to remember what it was called which is much more demanding than recognizing it on a list (Lansdale, 1988).