

Students' Organization Strategies of Personal Information Space

Sharon Hardof-Jaffe¹, Arnon HersHKovitz¹, Hama Abu-Kishk², Ofer Bergman³, Rafi Nachmias¹
{sharonh2, arnonher, nachmias}@post.tau.ac.il, hama@bgu.ac.il, o.bergman@sheffield.ac.uk

¹ Knowledge Technology Lab, School of Education, Tel Aviv University, Israel

² Department of Communication Studies, Ben-Gurion University, Israel

³ Information Studies Department, Sheffield University, UK

This study uses novel data mining methods in order to observe students' personal information space organization strategies in the personal Web space allocated to them on the university servers. The study included 518 users with a personal information space of at least 10 files. Data regarding their personal information space was collected, and a clustering algorithm was applied in order to identify profiles of students' organization strategies. Four clusters were found, refining the classical piling/filing classification: piling, one folder filing, small folders filing and big folder filing. Also, association was found between these profiles and personal information space size. A discussion of these results is provided.

1 Introduction

The relationship between learners and information has changed dramatically in the knowledge age. Students today have an abundance of available information (Salomon 2000) and they are required to choose, organize and retrieve information items in the course of the learning process. Nowadays, many students manage personal information spaces in which they save and from which they retrieve information items needed for their learning, e.g., assignments, articles, class-related hand-written notes, lists, illustrations, presentations, correspondences, and Website links. Personal Information Management (PIM) is an emerging research field exploring individual's activities of acquisition, organization, maintenance, retrieval, and sharing of information (Teevan et al. 2006). These PIM activities have an acute influence on learning processes, and particularly on university students who engage with many information items from various sources. Since PIM activity changes according to context (Krishnan and Jones 2004), the need arises to research how students manage personal information items, what PIM strategies they use and what are PIM characteristics in learning contexts (Barreau 2008; Chang and Ko 2008).

Malone (1982) identified two major organizational strategies for PIM: *piling* and *filing*. Other researchers examined PIM strategies with a variety of tools and in different contexts: in email application, in favorites management, and in the desktop files organization (Abrams et al. 1998; Boardman and Sasse 2004; Fisher et al. 2006;

Whittaker and Sidner 1996) Students' PIM strategies, more specifically, were also researched in a learning context during thesis and dissertations researches (Chang and Ko 2008).

Students manage their information using a variety of spaces (e.g., home desktop, laptop, personal directory in the university network, personal USB flash drive). An online storage space is typically offered to students as part of certain Web-supported course systems (e.g., Moodle, Blackboard) giving them ubiquitous access to their files. Using such environments, students manage their personal information space for their own needs, often in an idiosyncratic way (Kelly 2006). What is unique about this arrangement is that contrary to other environments, this online personal information space affords some data regarding PIM strategies to researchers. That is to say: a structural description of the online space is accessible, and this allows a large-scale study of the nature of students' organizational strategies in such environments.

So far studies that investigated PIM strategies have used traditional methodologies: interviews (Boardman and Sasse 2004; Malone 1982), screen captures (Boardman and Sasse 2004; Whittaker and Sidner 1996), and questionnaires (Abrams et al. 1998). However these methods typically test small number of participants and therefore their external validity is limited. *Data mining* is a set of tools and techniques for discovering unexpected valuable structures in large datasets, and has recently become an emerging methodology in education (Romero and Ventura 2006; Hershkovitz and Nachmias 2009). Data mining methods have been suggested as enabling identification and measurement of PIM activities and personal information space structures for large populations (Chernov et al. 2008; Fisher et al. 2006). The purpose of this study is to use data mining methods in order to identify Personal Information Space Organization Strategies (PISOS) of students using online storage space.

2 Background

2.1 PIM Organization Strategies

Malone (1982) was the first to classify personal information management – in the context of office organization - into two types of strategies: filing and piling. In the piling style,

papers are heaped on top of each other in reverse chronological order, and the pile carries no label. In filing, by contrast, the papers are categorized into physical files which are labeled by these categories. Malone found that piles were useful for small collections both because the user could remember the location of each paper within the pile and because the paper on top of the pile could remind the user to do the associated task. However as the piles grew users could not keep track of their papers and reminder papers were covered by others and lost. Therefore in Malone's experiment filers did better than pilers at the retrieval task. In the digital office, papers are replaced by digital information items (e.g., files and emails), filing is done into folders and directories with labels describing their category, and piling is typically done by heaping the information item in a root directory such as My Documents for files and the Inbox for emails.

The folder hierarchy is the standard mechanism for organizing personal information in digital environments. This mechanism allows users to create a personal classification scheme based on categories and dimensions they see as relevant (e.g., role, project, time). The hierarchic method was the target of some criticism, mainly because of its static nature and poor scalability to large information spaces (Dourish et al. 1999; Fertig et al. 1996; Gemmell et al. 2002; Nelson 1999; Raskin 2000). Unlike the hierarchic method, which requires users to remember the category/location they gave the file at storage time, search allows them to retrieve the file using any information they remember about it (Lansdale 1988). The hope was that search would replace navigation through hierarchic folders and eliminate the need for hierarchic folders (Cutrell et al. 2006; Dourish et al. 1999; Fertig et al. 1996; Raskin 2000). However, Barreau and Nardi (1995) found a preference of navigation over search for file retrieval, and their findings were consistently repeated in later research (Boardman and Sasse 2004; Capra and Pérez-Quñones 2005; Kirk et al. 2006; Teevan et al. 2004). Moreover, no evidence was found that the use of improved search engines leads to more use of the search option or to changes in filing behavior (Bergman et al. 2008).

Research shows that most users employ a mixture of piling/filing strategies (Whittaker and Hirschberg 2001). This binary classification has served as the basis for many other PIM classifications and was extended mainly to describe different filing activities (i.e.,

whether users file their items and when). In the context of email management, Whittaker and Sidner (1996) found three types of users of PIM strategies: no filer - no use of folders, frequent filer – use of folders and clean up inbox on a daily basis and spring filer – use of folders clean up inbox only periodically. These strategies were found to be persistent even a decade later (Fisher et al. 2006), although some parameters of email use had changed dramatically (i.e., archive size, number of folders). Regarding bookmarks management, the filing strategies were observed by one of the following (Abrams et al. 1998): *no filing* (for users who never organize bookmarks); *creation-time filing* (storing a new bookmark in the appropriate category on first accessing the Webpage); *end-of-session filing* (organizing the entire session's bookmarks at once), and *sporadic filing* (sorting out bookmarks occasionally). Whittaker and Sidner (1996) defined two types of dysfunctional folders: the *failed folder* which contains less than three files, and the *too big folder* which contains too many files and becomes unwieldy.

PIM strategies were suggested to be dependent on document type, as shown by Boardman and Sasse (2004) who showed that users often invested more time in organizing files than in email messages and bookmarks. Users were categorized by Boardman and Sasse into *total filers* (most files are filed when created), *extensive filers* (filing is done extensively, yet many items remain unfiled), and *occasional filers* (filing is done occasionally, leaving most items unfiled). In the context of learning, two strategies were observed with reference to the time it takes users to build a new folder: (a) *pre builders* are students who create new folders before they produce any items, and (b) *post builders*, are those who prefer to create new folders after a set of new items is collected (Chang and Ko 2008).

2.2 PIM and Learning

The nature of information has dramatically changed in the digital era, as information is easily accessible, mostly distributed, presented in multiple formats, and hypertext-oriented. As they are learning, students create a personal information space, negotiating between the huge amount of available information - from various sources - and their limited processing abilities at any given time. Students therefore need to acquire Personal Information Management (PIM) literacy in order to efficiently manage their own learning

environment which is associated with the nature of the subject matter and the assignment requirements (Mioduser et al. 2009). PIM is not just a set of practical actions of saving and retrieving information items; it is an integral and a central part of the learning process. Naming a new file, grouping files together, categorizing them under a new folder name and classifying new files into existing folders involves constructive cognitive processes. According to the constructivist approach to learning, knowledge is constructed through a process in which learners actively integrate new knowledge with previous knowledge (Brooks and Brooks 1993).

External influences (e.g., task, environment, and context) have been suggested as a major factor affecting Personal Information Management (PIM), prior to group and individual differences. Among the external influences, the context in which the user organizes her or his documents is a key factor in the documents' creation, classification and retrieval (Barreau 1995; Gwizdka and Chignell 2007).

Although PIM is a central component of the learning process (Bergman, Beyth-Marom, and Nachmias 2003), the relation between PIM and learning has received only little research attention. Students may have specific needs for learning-related activities (e.g., versions management, backup) which require adjustment from PIM in non-learning context. Furthermore, PIM might change as students gradually move from novice to expert level in their domain (Barreau 2008; Chang and Ko 2008).

2.3 PIM Research and Data Mining Methods

Data about how users organize their personal information space have tended to be collected by means of traditional research methodologies, e.g., in-depth interviews, semi-structured interviews, screen captures, and questionnaires (Bergman et al. 2008; Boardman and Sasse 2004). Over the last few years, *data mining* has been suggested as a methodology for PIM research, adding new promises. Data mining is the general term for a set of tools and techniques for finding unpredicted patterns in large databases. Several PIM studies have used such methods, using automatically collected data from personal information spaces and activity log files.

Previous research has been referring to automatically gathered data, personal information items indexing, and log activity of email behavior (Fisher et al. 2006; Teevan et al. 2005). Chirita et al. (2006) applied clustering algorithms to both activity logs of desktop files accessing and the files content, in order to identify groups of files in the same context. Clustering algorithms were also applied by Manco, Masciari and Tagarelli (2008) for automatically classifying email messages according to their content. The above examples demonstrate collection and analysis of large datasets, which would not have been possible using traditional methods. An approach for generating datasets for research was presented by Chernov et al. (2008), according to which activity logs are being kept, holding information about the history of each file or email message. Such a dataset will enable the evaluation of different desktop search tasks and the designing of new search tools.

Previous studies presented PIM strategies with different tools and in various contexts; the learning context is especially interesting since students have to use PIM extensively. This study uses data mining methods to explore Personal Information Space Organization Strategies (PISOS) of students using online personal storage space. The three main research questions are:

1. How do students organize the personal Web space allocated to them by the university Learning Management System, as measured by PIM parameters?
2. What are the students' Personal Information Space Organization Strategies (PISOS)?
3. Is there an association between students' PISOS and their storage size (number of total files in storage)?

3 Methodology

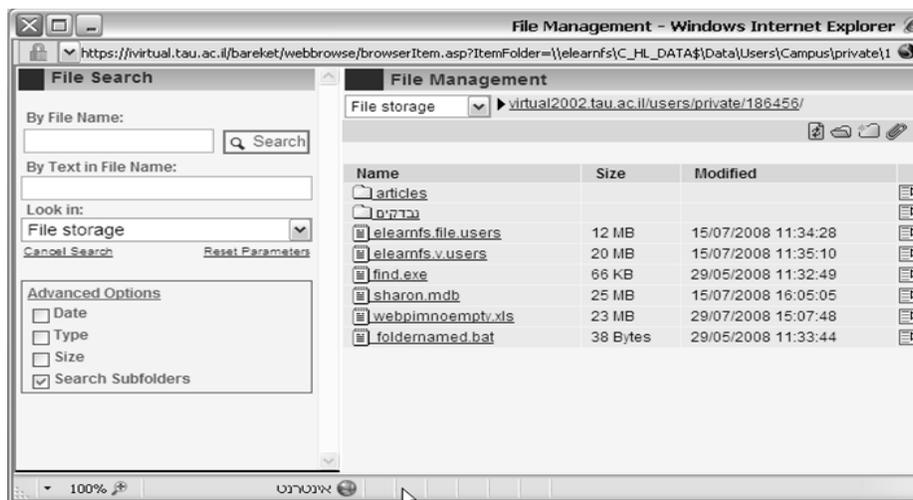
3.1 Research Field

Tel-Aviv University (TAU) is one of the largest research-oriented universities in Israel. Located in the center of the country, it serves about 26,000 students annually. These students are enrolled in about 6,000 courses that are taught by about 2000 instructors in almost every academic discipline. The VirtualTAU project at Tel-Aviv University was launched in the academic year 2000–2001 in response to a government initiative to

stimulate the use of advanced learning technologies in Israeli higher education. Currently, in the project's eighth year, VirtualTAU encompasses over 4,300 courses. The learning management system used by VirtualTAU is Highlearn. Like other learning management systems (e.g. Blackboard, Moodle), Highlearn allows instructors to develop an information base for the course content, to create didactic activities and to use communications tools (Nachmias and Ram, 2009).

One of the components of VirtualTAU is the virtual personal directory for file management (Figure 1), which enables students to store and to manage personal information items on the Web; this environment is accessible from any Internet-connected computer. Users of this environment can ubiquitously upload files, create folders, and retrieve files by navigating or searching.

Figure 1 – The Personal File Management feature in VirtualTAU



3.2 Research Population and Data File

The study was conducted with 2,081 undergraduate students, graduate students and staff who kept information items in their virtual personal directory. Data included the list of files and folders (full paths) for all the users, where each personal information space had a unique random identification. The raw data included more than 70,000 files and folders. Data were collected on August 2008. Since most of the 2,081 students used the storage space for only a few files, during a preliminary preprocessing stage a new data file was

created excluding students with less than 10 files, resulting in 518 students and 48,744 files and folders.

3.3 Procedure

The study was conducted in three stages:

First Stage: Describing students' use of personal information space on the Web. This stage included measuring PIM variables for the group of 518 students, and choosing the PIM variables for the cluster analysis.

Second Stage: Identifying personal information space organization strategies (PISOS). This stage included a Two-step Cluster Analysis of the students into k disjoint groups, in order to classify students according to their personal information space organization strategy. After several iterations, $k=4$ was chosen as resulting in the best fitting clustering.

Third Stage: Association between PISOS and personal information space size. Variance between the four clusters identified in the first stage was examined in order to identify the differences between PISOS groups regarding personal information space size.

3.4 Clustering Variables

In order to examine personal information space organization strategies, four variables were chosen as best representing the differences between the students' strategies:

Files per folder - the average folder size.

Number of files in the largest folder - the largest number of files in one directory (including root directory).

Ratio between number of files in the root and total files – Measures the root directory piling tendency.

Ratio between number of files in the largest folder and total files - the largest folder (not including root directory) divided by the total number of files.

4 Results

4.1 Student Personal Information Space on VirtualTAU

Data analysis was done with the aim of describing how students use the personal information space. Table 1 presents descriptive statistics for the variables used. The average number of files was 80.52 (with a standard deviation of 170.17), and the average for number of folders was 13.58 ($SD = 45.33$). The maximum hierarchy depth on average was 2.59 ($SD=1.42$) and the average file per folder was 16.16 ($SD=23.06$). The ratio between the number of files in the root and the total files was, on average, 0.38 ($SD= 0.4$), and the ratio between the number of files in the largest folder under the root and the total files was, on average, 0.28 ($SD = 0.28$). Dysfunctional folders, failed folders (<2files) and too large folders (>50files), were found in a high percentage of the students population, 56.56% and 9.1% respectively.

Table 1 - Personal Information Space Descriptive Statistics (N=518)

Variable	Mean (SD)	Median
# files	80.52(170.17)	30.50
# folders	13.58(45.33)	3
Hierarchy depth	2.59(1.41)	2
files per folder (user mean)	16.16(23.06)	10.55
Largest folder (includes root) [# files]	27.15(35.34)	16
Ratio between # files in the root and total files	0.38(0.40)	0.17
Ratio between # files in the largest folder (not includes root) and total files	0.28(0.28)	0.20

4.2 Student Personal Information Space Organization strategies

Students in the four clusters were grouped together according to the values of the four PISOS-related variables. The names and descriptions of each cluster are based on the two-step cluster analysis results; means and standard deviation for each variable and for each cluster are given in Table 2.

Table 2 - Means (SD) of the PISOS-related Variables (maximum values are bold, minimum values are in italic)

Cluster name	N	%	Average files per folder	Largest folder (includes root) [# files]	Ratio between # files in root and total files	Ratio between # files in largest folder (not includes root) and total files
Piling	141	27	17.78 (7.33)	22.71 (16.60)	.97 (.08)	.02 (.06)
One Folder Filing	49	9	14.70 (7.49)	18.77 (8.53)	.09 (.11)	.86 (.13)
Small Folders Filing	262	51	6.10 (4.49)	14.52 (11.55)	.18 (.20)	.26 (.16)
Big Folder Filing	66	13	23.10 (7.67)	71.62 (28.00)	.13 (.19)	.48 (.27)
Combined	518	100	12.26 (8.97)	24.42 (24.19)	.38 (.40)	.28 (.28)

Piling - ($n = 141$) is characterized by extreme values of two variables' means (Table 2): *ratio between number of files in the root and the total files* and *ratio between number of files in the largest folder (not including root) and total files*. The mean of *ratio between number of files in the root and the total files* within this cluster is maximal (0.97), indicating that most of these students' files are stored in the root directory. It is not surprising then that the largest folder in the root directory is extremely small relative to the total size, as indicated by *ratio between number of files in the largest folder (not including root) and total files*, the mean of which reaches its minimum in this cluster (0.02). These two extreme values of variables are typical of *piling* organization strategy.

One Folder Filing ($n = 49$), the same two variables as in the previous cluster (Table 2) play an important role as their means take their extreme values in it. In this cluster, the mean of *ratio between number of files in the root and the total files* is minimal (0.09), which can lead to the conclusion that what we have here is a non piling strategy. However, the mean of *ratio between number of files in the largest folder (not including root) and total files* is high (0.86), which demonstrates the existence of a folder holding a very high percentage of the files. What distinguishes this strategy from the previous one is that the files were saved in one main folder out of the root directory – a strategy we may call *one folder filing*.

Small Folders Filing ($n = 262$) has minimum mean values for two variables (Table 2): *average files per folder* and *largest folder (including root)*, i.e., these students have small

folders on average (6.1) and their largest folder is also relatively small (14.52). This suggests that the cluster represents a small folders filing organization strategy.

In **Big Folder Filing** ($n = 66$), the means of the same two variables as in the previous cluster (Table 2) take their extreme values: Both *average files per folder* (23.1) and *largest folder (including root)* (71.62) are maximal. These students' largest folder is the largest of the clusters (71.62). Also, from the mean of *ratio between number of files in the root and total files* (0.13), it might be concluded that about 87% of their files are filed, with one *folder containing about half of their files* (0.48). Therefore, this cluster, which we called *big folder filing*, describes a mixture of filing and piling.

According to this analysis of the clusters, we present the following classification of personal information space organization strategies - PISOS: *piling*, *one folder filing*, *small folders filing*, and a mixture of piling and filing, which we call *big folder filing*. The distribution of the four types in the research population is shown in Table 2; *piling* 27%, *one folder filing* 9%, *small folders filing* 51%, *big folder filing* 13%.

4.3 Association Between PISOS and Personal Information Space SIZE

A descriptive statistics of *personal information space size* for different categories of PISOS is presented in Table 3. It might be noticed that two strategies (*piling*, *one folder filing*) have a small *personal information space size* on average (24.4 and 22.31, respectively). The largest *personal information space size* mean (284.73) was found in the *big folder filing* group. When examining the differences between PISOS groups regarding *personal information space size* (number of files) significant differences were found $F(3,514)=50.16, p<0.01$.

Table 3 – Personal Information Space Size for the different values of PISOS

PISOS	N	Mean (SD)
Piling	141	24.40(20.30)
One folder filing	49	22.31(10.85)
Small folders filing	262	70.18(101.04)
Big folder Filing	66	284.73(369.04)

5 Discussion

The main purpose of this study was to identify different types of personal information organization strategies (PISOS) in learning context, and to do so for a large population using novel data mining methodologies. We found four types of personal information strategies in practice, and we described them as follows: (a) *piling* – keeping most of the files in the root directory; (b) *one folder filing* – filing most of the items in one folder; (c) *small folders filing* – organizing the personal information items in many small folders; and (d) *big folder filing* – a mixture of filing most of the documents in different folders, but still maintaining one sub-folder containing many files, i.e., a hidden pile. The identification of several PISOS, based on the personal information space structure, enabled us to refine and re-examine the traditional dichotomy between piling and filing made by Malone (1982). Previous work that extended the filing\piling distinction by describing changes in these PIM strategies over time and according to various situations (Abrams et al. 1998; Chang and Ko 2008; Whittaker and Sidner 1996). In particular it has been shown that users often apply a mixture of these two strategies depending on the context and the tools they use (Boardman and Sasse 2004).

The differences between these four strategies may have implications for learning since the level of filing which involves naming, grouping, categorizing and classifying is different for each strategy. By naming, grouping, categorizing and classifying the students integrate new information items into an existing archive. This external activity requires an internal cognitive one by which the learners associate their newly acquired knowledge with their existing knowledge (Brooks and Brooks 1993). In this way students create an information repository, which represents their knowledge of the subject matter. In the piling strategy, students neither name, classify, nor categorize any information items. In one-folder filing strategy, the students name only a few folders and don't classify or categorize at all. However, in small-folder filing, and in big-folder filing strategies, students name, classify and categorize items into folders. Further research should study the learning behaviors of each of these four groups.

Another finding is the association between big folder filing strategy and personal information space size. The size of this space is relatively large and it characteristically

has one hidden pile in a sub-folder. Retrieving items from folders that are too large is one of the difficulties with PIM, especially in the case of large storage (Whittaker and Sidner 1996). These hidden piles could also be caused by the student saving many information items without sub category, or from an online course, in his personal information space. How the learning environment affects student PIM should be researched in the future.

PISOS analysis showed that more than half of the participating students used the small folders filing strategy. As this strategy is characterized by the use of small folders (6.1 files per folder on average) this implies the existence of relatively many near-empty folders. This finding might be explained by the possibility that folders were actually created before information items were collected, pre-built according to the student's courses. Pre-building folders were found in previous research examining students' PIM (Chang and Ko 2008). However, when folders are too small they might increase PIM complexity (Whittaker and Sidner 1996), and might suggest the existence of difficulties in classifying the learning material. Further research is needed for examining changes over time in the personal information space, in order to find out whether more items are added to initially small folders.

The automatic large-scale collection of structural data, which was so useful to this research, was the result of new Web-based personal information spaces. PIM is subjective and idiosyncratic (Bergman, Beyth-Marom and Nachmias 2003, 2008), and because PIM research mostly uses qualitative data collection from relatively small populations it might seem that there are as many PIM variations as there are researched users (Kelly 2006). However, using a large research population and data mining techniques, unexpected patterns might arise from the data, suggesting similarities between groups of users. To promote the creation of large datasets, Chernov et al. (2008) have suggested building a repository of PIM activity log files; this then would serve the PIM research community. Since it is likely that there will be problems obtaining participants' consent to trace their PIM activity over time, it might be easier to collect structural data reflecting accumulating activity.

In summary, this research demonstrates how data mining techniques enhance our understanding of students' Personal Information Management activities and reveals how students actually manage personal information items in the university learning system. The four strategies differ in PIM activities that involve cognitive processes valuable for learning: naming classification and categorization. Future research should examine changes over time in students' PISOS; the appearance of these four PISOS in other personal information spaces such as desktop and portable devices; and possible differences in learning behaviors between these four groups.

6 References

Abrams, D., Baecker, R., Chignell, M. (1998) Information archiving with bookmarks: personal Web space construction and organization *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press/Addison-Wesley Publishing Co., Los Angeles, California, United States.

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) From Novice to Expert: Personal information management behaviors in learning contexts, *CHI 2008 Workshop*, Florence, Italy.

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

Bergman, O., Beyth-Marom, R., Nachmias, R. (2003) The user subjective approach to personal information management systems, *Journal of the American Society for Information Science*, 54(9), 872-878.

Bergman, O., Beyth-Marom, R., Nachmias, R. (2008) The user-subjective approach to personal information management systems design: Evidence and implementations, *The American Society for Information Science and Technology*, 59(2), 235-246.

Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N., Whittaker, S. (2008) Improved search engines and navigation preference in personal information management, *ACM Trans. Inf. Syst.* 26(4) 1-24.

Boardman, R., Sasse, M.A. (2004) "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management, *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, Vienna, Austria.

Brooks, J.G., Brooks, M.G. (1993) *In search of understanding: The case for constructivist classrooms, association for supervision and curriculum development*, Alexandria, VA.

Capra, R.G., Pérez-Quñones, M.A. (2005) Using Web search engines to find and refind Information, *Computer* 38(10) 36-42.

Chang, S.-J., Ko, M.-H. (2008) Behaviors of PIM in context of thesis and dissertation research, *CHI 2008 workshop*, Florence Italy.

Chernov, S., Demartini, G., Herder, E., Kopycki, M., Nejdil, W. (2008) Evaluating personal information management using an activity logs enriched desktop dataset, *CHI 2008 Workshop*, Florence, Italy

Chirita, P.A., Gaugaz, J., S.Costache, W.Nejdl. (2006) Desktop context detection using implicit feedback, *In SIGIR 2006 Workshop on Personal Information Management*, Seattle WA, USA.

Cutrell, E., Dumais, S.T., Teevan, J. (2006) Searching to eliminate personal information management, *Commun. ACM* 49(1) 58-64.

Dourish, P., Edwards, W.K., LaMarca, A., Salisbury, M. (1999) Presto: an experimental architecture for fluid interactive document spaces. *ACM Trans. Comput.-Hum. Interact.* 6(2) 133-161.

Fertig, S., Freeman, E., Gelernter, D. (1996) Lifestreams: An alternative to the desktop metaphor, *Conference companion on Human factors in computing systems: common ground*. ACM, Vancouver, British Columbia, Canada.

Fisher, D., Brush, A.J., Gleave, E., Smith, M.A. (2006) Revisiting Whittaker & Sidner's "email overload" ten years later, *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM, Banff, Alberta, Canada.

Gemmell, J., Bell, G., Lueder, R., Drucker, S., Wong, C. (2002) MyLifeBits: fulfilling the Memex vision, *Proceedings of the tenth ACM international conference on Multimedia*. ACM, Juan-les-Pins, France.

Gwizdka, J., Chignell, M.H. (2007) Individual Differences. W. Jones, J. Teevan, eds. *Personal Information Management*, University of Washington Press: Seattle and London, 206-220.

Hershkovitz, A. & Nachmias, R. (2009) Learning about online learning processes and students' motivation through Web usage mining, *Interdisciplinary Journal of E-Learning and Learning Objects*, 5(2009), 197-214.

Kelly, D. (2006) Evaluating personal information management behaviors and tools, *Commun. ACM* 49(1) 84-86.

Kirk, D., Sellen, A., Rother, C., Wood, K. (2006) Understanding photowork, *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, Montreal, Quebec, Canada.

Krishnan, A., Jones, S. (2004) TimeSpace: activity-based temporal visualisation of personal information spaces, *Personal and Ubiquitous Computing* 9(1) 46-65.

Lansdale, M.W. (1988) The psychology of personal information management, *Applied Ergonomics* 19(1) 55-66.

Malone, T.W. (1982) How do people organize their desks? (Extended Abstract): Implications for the design of office information systems, *Proceedings of the SIGOA conference on Office information systems*. ACM, Philadelphia, Pennsylvania, United States.

Manco, G., Masciari, E., Tagarelli, A. (2008) Mining categories for emails via clustering and pattern discovery, *Journal of Intelligent Information Systems* 30(2) 153-181.

Mioduser, D., Nachmias, R., Forkosh-Baruch, A. (2009) New Literacies for the Knowledge Society, J.M. Voogt, G.A. Knezek, eds. *International Handbook of Information Technology in Primary and Secondary Education*. Springer, 23-41.

Nachmias, R., Ram, J. (2009) Insights from a Decade of Campus-wide Implementation of Blended Learning in Tel Aviv University, *The international review of research in Open and distance learning*, 10(2).

Nelson, T.H. (1999) Xanalogical structure, needed now more than ever: parallel documents, deep links to content, deep versioning, and deep re-use, *ACM Comput. Surv.* 31(4es) 33.

Raskin, J. (2000) *The humane interface: new directions for designing interactive systems*, ACM Press/Addison-Wesley Publishing Co. New-York.

Romero, C., Ventura, S. (2006) *Data Mining in E-learning (Advances in Management Information)*, Wit Pr/Computational Mechanics.

Salomon, G. (2000) *Technology and Education in the Age of Information*, Zmora-Bitan, Haifa (Hebrew).

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, *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, Vienna, Austria.

Teevan, J., Dumais, S.T., Horvitz, E. (2005) Personalizing search via automated analysis of interests and activities, *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, Salvador, Brazil.

Teevan, J., Jones, W., Bederson, B.B. (2006) Introduction, *Commun. ACM* 49(1) 40-43.

Whittaker, S., Hirschberg, J. (2001) The character, value, and management of personal paper archives, *ACM Trans. Comput.-Hum. Interact.* 8(2) 150-170.

Whittaker, S., Sidner, C. (1996) Email overload: exploring personal information management of email, *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*. ACM, Vancouver, British Columbia, Canada.