

From User Needs to Opportunities in Personal Information Management: A Case Study on Organisational Strategies in Cross-Media Information Spaces

Trullemans, Sandra; Signer, Beat

Published in:

Proceedings of the International Conference on Digital Libraries (DL 2014), London, UK, September, 2014

DOI:

[10.1109/JCDL.2014.6970154](https://doi.org/10.1109/JCDL.2014.6970154)

Publication date:

2014

License:

Other

Document Version:

Accepted author manuscript

[Link to publication](#)

Citation for published version (APA):

Trullemans, S., & Signer, B. (2014). From User Needs to Opportunities in Personal Information Management: A Case Study on Organisational Strategies in Cross-Media Information Spaces. In *Proceedings of the International Conference on Digital Libraries (DL 2014), London, UK, September, 2014* (pp. 87-96). IEEE.
<https://doi.org/10.1109/JCDL.2014.6970154>

Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

From User Needs to Opportunities in Personal Information Management: A Case Study on Organisational Strategies in Cross-Media Information Spaces

Sandra Trullemans and Beat Signer
Web & Information Systems Engineering Lab
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
{strullem,bsigner}@vub.ac.be

ABSTRACT

The efficient management of our daily information in physical and digital information spaces is a well-known problem. Current research on personal information management (PIM) aims to understand and improve organisational and re-finding activities. We present a case study about organisational strategies in cross-media information spaces, consisting of physical as well as digital information. In contrast to existing work, we provide a unified view on organisational strategies and investigate how re-finding cues differ across the physical and digital space. We further introduce a new mixing organisational strategy which is used in addition to the well-known filing and piling strategies. Last but not least, based on the results of our study we discuss opportunities and pitfalls for future descriptive PIM research and outline some directions for future PIM system design.

Categories and Subject Descriptors

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

General Terms

Human Factors, Theory

Keywords

Personal information management (PIM), cross-media information spaces, organisational strategies, mixing

1. INTRODUCTION

The desktop metaphor for personal computers was introduced at Xerox PARC in the early seventies. Nowadays, the desktop screen in combination with the tree structure of our file systems still represents the standard way of organising information in digital space, whereas file cabinets

and desks are used in physical space. Already in the eighties, researchers mentioned the lack of information organisation by users, since it is a time consuming and cognitively loaded task [1]. Due to the lacking organisation, we often retrieve information in an ineffective way. The improvement and understanding of the organisational and re-finding activities of personal information is addressed by research on *personal information management* (PIM). A pioneer in the research on how people organise information in their offices is Malone [2]. In his work, filing and piling are formally described as the main organisational strategies. Therefore, Malone is seen as the founder of the most prominent PIM variable, namely the *order* variable which is expressed on an axis ranging from ‘order’ to ‘disorder’ (i.e. from filing to piling). Over the past thirty years, this order variable has been investigated by numerous PIM researchers [3].

In Malone’s approach, information carriers are seen as individual elements (e.g. paper documents) and organisational strategies can be defined to arrange these elements. Malone’s definition of the files and piles organisational units in terms of their specific properties is provided in Table 1.

Table 1: Malone’s definition of files and piles

	Elements titled	Elements ordered	Groups titled	Groups ordered
Files	Yes	Yes	?	?
Piles	?	No	No	?

Files consist of a number of elements which are titled and ordered in a user-specified way such as alphabetically or chronologically. The files (groups) themselves might be explicitly titled and ordered but they do not have to, which is represented by the question marks in Table 1. On the other hand, individual elements (e.g. documents) of a pile do not have to be titled and are definitely not arranged in a particular order. Last but not least, piles (groups) are never titled and do not have to be arranged in a particular order. The filing strategy has been introduced in the digital space via digital file systems where documents are titled and ordered within a folder. Furthermore, each digital folder (group) can be titled and ordered again with the whole hierarchical structure forming a file system.

In a user’s personal information space, it is not always desired to be restricted to the filing requirements of having to label and order the information items due to the cognitive

effort and required time [1], which has to do with the cognitive overload at the moment of describing the element [4]. Additionally, in physical as well as in digital space, filing introduces the classification problem [2, 5]. People have difficulties to decide on the most appropriate classification structure for later retrieval (e.g. alphabetical, contextual or chronological). A second difficulty is the challenge to label an item with enough information about the item itself while still ensuring that the label fits into the classification structure. Frequently, users construct long and complex labels, leading to fuzzy and overlapping categorisations [6]. By squeezing contextual information into the labels, a mismatch between the moment of filing and retrieval can be identified [1, 7].

Users often do not file all their personal information away but instead construct piles on their desks, shelves and on their digital desktop screen. Note that for digital piles we followed Malone’s definition of piles representing a collection of not necessarily titled and non-ordered elements. Therefore, digital piles can, for example, be represented by folders with unlabelled or non-semantically labelled documents or they can be constructed by using specific tools such as BumpTop¹. In addition to the classification problem, the lack of time to process all information and the uncertainty about the future use and value of the gathered information are other factors that lead to the piling of information [8]. Furthermore, piles also provide a much better reminding function than archives [9, 7, 8].

Besides the investigation of filing and piling, various research has been carried out to identify behavioural patterns. Behavioural patterns are often defined by combining the order variable (e.g. degree of filing) and a time variable [3]. For example, the work of Whittaker and Sidner [10] observed three organisational patterns in email. *No-filers* are users who rarely file away emails whereas *spring cleaners* sporadically move emails from their inbox to the specific folder hierarchy. Last but not least, *frequent filers* empty their inbox daily by filing most of their emails. Similarly to this work, other research identified different related patterns based on the order variable in email [11, 12], bookmarks [13, 14] and file systems [14]. Likewise, Boardman’s study defined cross-tool behavioural patterns for email, bookmarks and the file system [14]. The user’s organisational behaviour is expressed by a cross-tool profile and specific profiles were categorised in *pro-organisers* and *neutral-organisers*. These two categories are in fact the “neat” and “messy” behaviours mentioned by Malone [2] but applied in a digital cross-tool setting. In addition to the use of the order variable in defining behavioural patterns in tools as well as in cross-tool settings, more recent work uses the same approach to identify behavioural patterns in how people organise micro-notes [15] and to-do’s [16]. For example, Van Kleek et al. [15] observed behavioural patterns that are related to the work of Whittaker and Sidner [10], including *minimalists* who are closely related to Whittaker and Sidner’s *frequent filers*, *packrats* who are related to *pilers* and *periodic sweepers* that can be compared to Whittaker and Sidner’s *spring cleaners*.

As mentioned earlier, previous research has extensively investigated filing and piling and has used the order variable as a basis for describing organisational patterns [3]. Nevertheless, due to the bottom-up approach of these studies, less is

known about the overall organisational landscape of users in physical and digital information spaces. To which extent do we actually use the organisational units (i.e. files and piles) which are used to describe the order variable and how much do the issues introduced by them bother us? Do we re-find information in physical and digital information space in the same way? By drawing such a generalisation, we might gain insights in the need for enhancements, points of focus for descriptive research and correlations or differences in a user’s personal information space. These insights could then be taken into account when designing future cross-media PIM systems with support for both the physical as well as the digital space. In this paper we address the following questions:

- Is it enough to use only filing and piling as a means to express the order variable or are there additional organisational strategies which are complementary to filing and piling in order to define the degree of order?
- Is there any coherency and/or dependency between these organisational strategies in the physical and digital space?
- Do we use the same re-finding cues after applying an organisational strategy in physical and digital space?

We start by describing the methodology used in our study and then present the results of the study. This is followed by a discussion of new opportunities and an outline of future PIM system design directions.

2. METHODOLOGY

For our study on organisational strategies, we used an embedded multi-case case study research design. The context of our study is given by descriptive research in the domain of personal information management, targeting the issues of organisation and re-finding content as experienced by most of us on a daily basis. Due to our interest in the organisation and re-finding behaviour of personal information in the physical as well as the digital space, we treat these two spaces as individual cases. The first case investigates the organisational and re-finding behaviour in the physical space whereas the second case analyses this behaviour in the digital space. The individual cases are first investigated independently. The units of analysis within each case address the identification of organisational strategies and the ease of re-finding information based on these applied organisational strategies. Second, due to the comparative nature of our research question, a cross-case analysis is performed to identify any coherence and dependency of user behaviour in organising and re-finding information in both spaces.

The validity of our study is ensured by several design decisions. The use of a multi-case design enhances the external validation of the work since both digital and physical information spaces are measured with the same research instruments and compared to identify contradictions and similarities. By analysing both information spaces and their coherency, the analytic generalisation may be extended to the level of personal cross-media information spaces. Note that previous research [8, 14, 10, 17] was limited to the individual analysis of the physical or digital space which makes the contribution of our study of high informative value. Internal validation is provided by the formulation of hypotheses and the use of explanation building in the data analysis

¹<https://code.google.com/p/bumptop/>

stage. Additionally, the carrying out of a pilot study also contributes to the internal validity. Finally, the use of a case study protocol in combination with the centralised data collection further increases the reliability.

2.1 Research Instruments

Our case study took place in two phases. In a first phase, we observed the organisational strategies applied by users in six academic offices. This initial phase gave us some more insights about the used artefacts and lead to the definition of the *mixing* strategy. In the second phase of the case study, we deployed an online survey to gain insights about the degree to which users apply filing, piling and mixing in the physical as well the digital space. Although surveys are mostly used in quantitative methodologies, they can be applied as a measuring instrument in a case study research design [18]. The observations in the first phase also let us define valid questions for the survey. The survey included questions related to the two units of analysis. The organisational strategies (i.e. filing, piling and mixing) were decomposed into several sub-factors such as the use of semi-ordered folders as a factor of mixing. The ease of re-finding is determined by questions such as how easy the users experience re-finding in the digital file system. In addition, the survey included some open-ended questions which were of an exploratory nature in order to, for example, investigate tools that support the re-finding of information in a file system.

After the data collection, we used various non-parametric statistical tests to analyse the survey data. Due to the qualitative nature of the study, most survey questions provide ordinal data (e.g. on a 5-point Likert scale) and a normal distribution could not be guaranteed. Therefore, non-parametrised tests are used to analyse correlations and differences in both individual cases as well as across the two cases in the cross-case analysis. Due to the ordinal character of the data, Spearman’s rho (ρ) is used to identify correlations between factors. All correlations included in the discussion of results are significant (p -value) at a minimal level of 0.05 ($p < 0.05$). Second, to identify a significant difference between factors, the Wilcoxon signed-rank test (z -value), a non-parametric variant of the t -test for dependent samples, has been used with a significance level of 0.05 ($p < 0.05$) defined on the results. Besides these non-parametrised statistical tests, the frequencies and open-ended questions are interpreted via Toulmin argumentation [19], allowing us to form conclusions from non-statistical data by using warrants derived from related work.

2.2 Participants

We have limited our representative population to academic knowledge workers. They are seen as a population which uses a large amount of personal information in order to be informed [20].

Table 2: Distribution of respondents

Profession	#Respondents
Professor	10
Postdoctoral researcher	8
PhD student	35
(Under)graduate student	117

In the second phase of the case study, we have enquired 170 academic knowledge workers including professors, post-doctoral researchers, PhD students as well as undergraduate and graduate university students. Table 2 provides the distribution over all the respondents. Furthermore, the respondents were anonymous and have not been reimbursed in order to avoid a bias in our results. Note that in the presentation of the statistical results, the respondent rate might differ from the initial 170 respondents due to the non-response of some respondents for some questions.

3. ORGANISATIONAL LANDSCAPE

3.1 Mixing in Addition to Filing and Piling

In a first phase, we have investigated filing and piling strategies in six academic offices. Besides the construction of files and piles that has already been observed in previous research [2, 8, 1, 9, 21], we identified other organisational structures as part of our observations. Users had a significant amount of structures which could not be classified as files or piles according to Malone’s work [2]. Consequently, they are not covered by the measurement of the order variable. These additional observed structures included:

- labelled ring binders, folders, letter trays and shelves containing unlabelled and/or unordered documents
- ring binders where the organisation of documents was not consistent over the whole binder
- explicitly ordered documents in the form of piles

Since according to Malone piles cannot be labelled or contain explicitly ordered content, labelled letter trays, ring binders and folders with unordered content cannot be classified as piles. On the other hand, due to the lack of explicitly labelled content and ordering they can also not be classified as files. A second observation is the inconsistent use of labelling and ordering in subparts of ring binders or folders. Users construct labelled or unlabelled subparts in ring binders which are then organised in a different manner. For example, one respondent had a ring binder with four subparts where the first part contained unlabelled unordered content and the second part consisted of unlabelled but explicitly chronologically ordered content. An example of the observed use of labelled letter trays and ring binders which, according to Malone’s definition, cannot be classified as piles or files is illustrated in Figure 1.

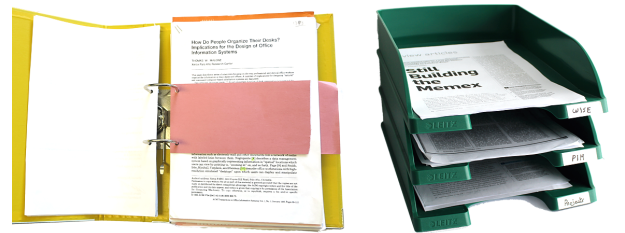


Figure 1: Structures which are neither files nor piles

In the digital space, we have identified the same issue of these unclassifiable structures. In their digital file systems, users had labelled folders containing files that were not explicitly labelled. A common situation which may lead to

such a situation is the import of pictures from a digital camera. For instance, one respondent had a folder named *DisneyLandParis* in their file system, containing images with labels such as DSC00533 with continuously increasing numbers. Similarly to ring binders in the physical space, this structure cannot be identified as a pile since the folder is labelled and on the other hand it is not a file since the images have not been explicitly labelled by the user. Another example is the upcoming use of tagging. In tagging, users categorise documents via tags without having to label the tagged document. The tag can be seen as a “container” including these unlabelled documents which forms a structure that falls beyond files or piles. Due to the fact that these “neither filing nor piling structures” can be found in numerous offices and digital file systems, we provide a formal definition for these observed structures in terms of the dimensions used by Malone to define files and piles.

DEFINITION 1. *Mixing is an organisational strategy which is neither filing nor piling. Mixtures as an organisational unit do not match the definition for files and piles by Malone. They may contain titled as well as untitled elements and elements might be explicitly ordered. A group might be titled and groups can also be explicitly ordered whereby the titles and ordering of elements do not have to be consistent.*

Since previous PIM research often investigated the order variable including filing and piling strategies as mentioned in the introduction, we might wonder if the mixing organisational strategy is of significant value for users and whether it should be seen as an equivalent to the mainly investigated and discussed filing/piling strategies.

3.2 The Use of Organisational Units

In the first unit of analysis forming part of the second phase of our case study, we addressed the question “*To which extent do users apply the mixing strategy besides filing and piling?*”. In order to get a global overview of the degree of use of the mixing strategy—based on the examples observed in the first phase of the case study—we asked respondents to answer questions such as “*To which degree do you have ring binders/folders with explicitly ordered content but where the documents are not labelled?*”, “*To which degree do you use labelled letter trays?*” and “*To which degree do you have ring binders/folders with semi-ordered content?*” on a 5-point Likert scale. To investigate the use of the mixing strategy in the digital space, respondents were asked to which degree they had folders including non-semantically-labelled files as illustrated earlier in the example with the import of digital pictures. The degree of use of filing and piling strategies was analysed in a similar way with several specific questions about file and pile structures in both the physical and digital space. The factors of filing, piling and mixing were constructed via factor analysis with a Cronbach’s alpha larger than 0.6 to ensure internal validation. The three factors were investigated in both cases, the physical as well as the digital space, and the average degree of use was aggregated for each organisational strategy in order to be able to make statements at the level of the cross-media information space. In this way, we now have a better understanding about the general degree of use of organisational strategies. Figure 2 illustrates the use of filing, piling and mixing from ‘not at all’ to a ‘very high degree’ of use.

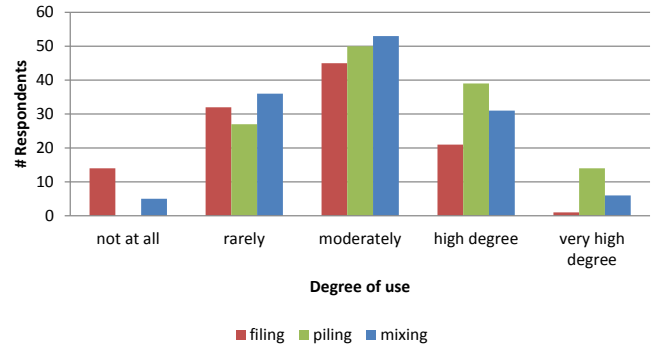


Figure 2: General degree of use of the three organisational strategies

For filing and piling, the diagram reflects the statements made by other researchers [2, 7]. Due to the classification problem, users generally apply less filing than piling. Furthermore, the normal curve is slightly skewed to the left which means that the overall use of filing is less frequent than piling ($z = -3.732, p = 0.000$). Most participants use filing on a ‘rare’ to ‘moderate’ level. Compared to the piling and mixing strategies, a significant number of users do not use filing at all. This is in contrast to piling where all of the 170 respondents had at least some digital piles, piles on a desk or piles in a bookcase. A majority of the respondents pile information from a ‘moderate’ to ‘high’ degree. Although piles impose re-finding issues when a large amount of information is encountered [2, 22], they represent the main strategy used in the organisation of personal information, which can be justified by the affordances of piles. Besides the classification problem as a force to pile information, the lack of time to process all incoming information due to information overload and the uncertainty about the future use and value of the gathered information are other factors that favour the piling of information [8]. Piles show some advantages in these situations but they also provide a much better reminding function than archives as described in the literature [1, 2, 9, 7, 8]. The properties of a pile imply that the elements are not ordered which means that the user needs to browse through the pile. By being confronted with other information elements in the pile, a reminder may be triggered. This reminding function is one of the affordances of paper in a general context [23]. A second function of piles is the preservation of the context of the information. This is in contrast to files, where at the time of filing the context of the label is not preserved [7]. By ordering piles in a contextual way, the context can be used as a cue for later retrieval. Several researchers share these findings and point out that the context cue is one of the main advantages of the human memory for recall [2, 7, 20, 24]. A last function of piles is the spatial property that they provide. The importance of spatial reference is supported by Cole [1] in his research on human aspects in office filing. A file system in a 3D space has spatial reference points which minimally compensates the classification problem. Nevertheless, piles provide much more spatial references than a file system and therefore the recall of an information item in a pile is to a greater extent based on spatial information.

Besides the general degree of use of filing and piling, the results show a normally distributed degree of use regard-

ing the mixing strategy. Some people do not use the mixing strategy at all while others use it to a very high degree. Often mixing is used from a ‘rarely’ to a ‘high’ degree (159 respondents) whereas most users indicated that they use the mixing strategy at a moderate level (53 respondents). This finding indicates that it is worthwhile to have a better understanding of mixing in the future as it is a heavily used organisational strategy in addition to filing and piling. We might attribute such a common use of the mixing strategy to the fact that neither filing nor piling are perfect organisational strategies. They both have their shortcomings, either in the form of the classification problem [2, 7] at the moment of filing or during the re-finding activities as mentioned before [2, 22]. Besides being a compensating strategy, our results regarding the mixing strategy align with the observation that it is better to provide a label to a group of information elements when storing long-term information rather than having a single spatial reference as provided by piles [22]. Mixtures can include more flexible structures where, for example, a letter tray provides spatial reference points to support short-term re-finding activities and users label the letter tray to improve the retrieval of content which is stored for a longer time.

3.3 Coherency and Dependency Between Information Spaces

Our second interest was to investigate the differences, coherency or dependency between the physical and digital information spaces in terms of the filing, piling and mixing strategies. Previous research pointed out the use of filing and piling in physical as well as digital space [10, 14, 25]. In digital space, we are used to the tree structure of our file systems and the desktop screen whereas in physical space we use file cabinets and our desks for the organisation of documents. Nonetheless, there is a lack of research on the coherency and dependency between the three identified filing, piling and mixing organisational strategies in both information spaces. Each space offers its own affordances with regard to the individual organisational strategies. For example, filing might be applied more frequently in digital space due to the spatial restrictions in physical space, such as the limited amount of documents a user may place in a file cabinet. On the other hand, piles might be easier to use in the physical space since the affordances of paper enable the easy grouping of papers documents [23].

In order to gain more insights about the affordances and interplay between both information spaces, a cross-case analysis between the two cases representing the physical and digital space has been carried out. First, we have investigated whether there is a difference in the degree of use of filing, piling and mixing across the physical and digital space. Our results show no significant difference in terms of the degree of use of each organisational strategy in the physical or digital space. This is illustrated in Table 3 where we provide the rank score of each strategy compared over the physical and digital space.

Surprisingly, people do not have significantly less piles, files or mixtures in the digital or the physical space. Figure 3 shows the distribution of the degree of use of the three organisational strategies in both spaces. Although the used Wilcoxon signed-rank test does not show any significant difference in the degree of use of the three strategies in physical and digital space, the diagram shown in Figure 3 indicates

Table 3: Results of Wilcoxon signed-rank test

Strategy	Rank-score (z)	Significance (p)
Filing	-0.659	0.510
Piling	-0.920	0.357
Mixing	-1.616	0.106

some differences to be taken into account. While almost all respondents (110/111) had some degree of physical piles, 30 respondents did not use digital piles at all. Also physical piling has a more moderate use (52 respondents) whereas the degree of use of digital piles is on average equally spread. The same can be observed for the mixing organisational strategy where 30 respondents do not use digital mixing in contrast to 10 respondents who do not use physical mixing. The semantic interpretation and explanation building are still valid besides a statistical test in qualitative research [18] and therefore we may state that there are differences in the degree of use between the two spaces concerning piling and mixing.

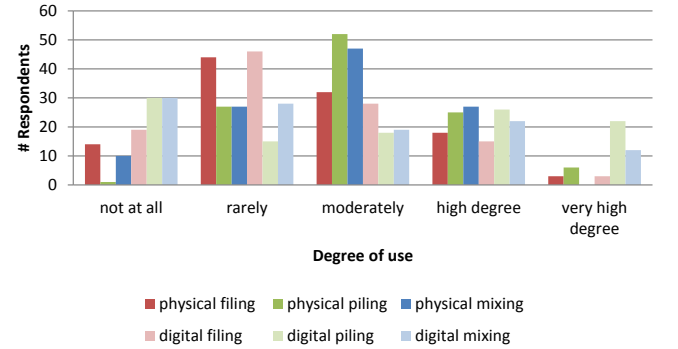


Figure 3: Use of organisational strategies in physical and digital space

Of special interest is the observation of more specific differences in piling. Digital piles are mostly used on the desktop screen whereas physical piles may be on a desk or in shelves. Our results actually reveal that users construct piles on desks and shelves to the same degree without any significant differences. Furthermore, we found a significant positive correlation between the use of piles on a desk and in shelves ($\rho = 0.431, p = 0.000$). This means that the more piles users have on their desk, the more they also have piles in shelves. Besides these findings, we can also indicate that users have at least some piles on the desk or in shelves. Only one respondent had no physical piles at all, whereas 15 respondents had no piles on their desk and 24 respondents had no piles in shelves. This illustrates that the users who had no piles on their desks had at least some piles in a shelf and vice versa. A major difference between desk piles and shelf piles is the kind of information they include. Our results indicate a much lower access frequency for piles in shelves than desk piles. On average, users access shelf piles on a ‘rarely’ to ‘sometimes’ basis (69/90) whereas desk piles are accessed by 70 of 90 valid respondents on a ‘sometimes’ to ‘very often’ basis. Note that the access frequency is one of the most important measurements in determining the category of information such as cold, warm or hot in-

formation [23]. We can therefore state that users store cold information in shelf piles whereas desk piles are used for warm and hot information.

When looking at the filing strategy, users apply several techniques to augment digital as well as physical file systems. In order to help them organise digital content, they use tools such as mind maps, wikis and often also the Mendeley² application for managing research papers. Nonetheless, these tools are not a perfect solution as stated by one of the respondents: *“I tried out many different ones including, for example, Mendeley for papers, but I was never happy”*. A second observation is the use of tools for decomposition and structuring activities in the digital file system as mentioned by a respondent: *“I use Total Commander by which I am able to compare two directories, use enhanced search functions, rename a list of files etc.”*. The previous statement indicates that although previous research already observed the activity of problem decomposition [6] and structuring [26] in the digital file system, users do use extra tools to help them in these activities. In contrast to digital file system enhancements which happen to be digital by nature, the augmentation of physical file systems appears to be physical as well as digital. Several respondents enhanced their physical file systems with digital indexing systems ranging from spreadsheets to digital archiving programs. The two most interesting responses were *“I only use inventories of books in the form of Google document spreadsheets, where I have a column which says location (home/office)”* and *“An Excel sheet to remind me of the files I created and what is inside them and where they are (file number)”*. The first respondent only applies the digital indexing for books which can be justified by the personal value they give to their books. Since we surveyed researchers, books can be of great informative value and therefore the benefit exceeds the effort required to keep such an indexing structure. Nevertheless, the second respondent applies digital indexing to most of the filed information. Since filing does not provide a reminding function as piles do, the respondent squeezes this reminding functionality into the digital indexing structure. Besides the digital indexing of physical file systems, users enhance the physical file system with post-it notes, book cards, stickers on folders as well as notes and they sometimes also use a physical table of contents in ring binders and folders. We can conclude that although users do not apply the physical and digital filing strategy in an extensive manner, once they apply filing, they often also augment their file system for better organisation.

4. RE-FINDING IN CROSS-MEDIA SPACES

Our last unit of analysis investigates how easy users may re-find personal information by the use of the previously described organisational strategies. Jones [5] described the interplay between keeping, organising and re-finding personal information. Re-finding activities are the activities which users execute to find a stored information item which they have integrated in their personal information space before. The main meta-level activities are organising and maintaining the stored items since re-finding becomes ineffective without any organisation of the stored information. Each of the three organisational strategies have their issues when items are added and they imply different re-finding

activities. In addition, these differences in re-finding activities vary across the physical and digital information space. Our results in this section provide an overall view on how easy users re-find information and which cues they apply in each organisational strategy in re-finding activities across the physical and digital space.

4.1 The Ease of Re-finding

By providing a generalised perspective about the difficulties that users experience in re-finding activities based on different organisational strategies, we can identify user needs where PIM research should provide support via augmentation. Figure 4 highlights to which degree users find it easy to re-find based on different organisational strategies.

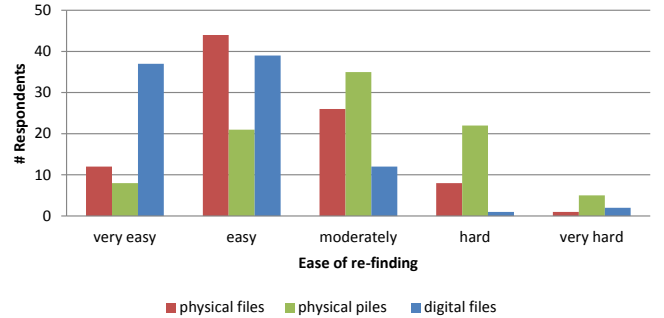


Figure 4: Ease of re-finding

First, we take a closer look at the ease of re-finding with physical filing and piling strategies. Several researchers identified that the use of a classification system enhances the re-finding effectiveness [27]. According to the state of the art, filing is best suited for cold information since the information is more structured in contrast to piling. On the other hand, piles provide more contextual information and spatial awareness for hot or warm information [23]. Nevertheless, we were interested in *how easy users find it to re-find information in files and piles*. Since our previous findings imply that 99% of the respondents have physical piles whereas 87% used a physical classification system to some degree, it is important to see to what extent users really experience difficulties in re-finding. Our results indicate that the effort invested in classifying information items improves the re-finding of information as shown in Figure 4. On the other hand, even though all users have piles, on average they find it ‘moderately’ to ‘hard’ to re-find information in their piles. This indicates that despite the difficulties of re-finding information in piles, the affordances of piles [23]—which include the preservation of context, reminding function and extra spatial reference points—ensure that everybody uses them anyway. The differences between the ease of re-finding in files and piles are significant ($z = -4.389, p = 0.000$).

Finally, in the digital file system it is significantly easier to re-find information than in its physical counterpart ($z = -4.097, p = 0.000$). Most users find it ‘very easy’ to ‘easy’ (37 respectively 39 respondents) to re-find information in the digital file system. A difference is observed compared to the physical file system where most respondents find it ‘easy’ to ‘moderately easy’ (44 respectively 26 respondents) to re-find personal information. This finds its basis in the fact that in digital space one can easily create long labels for

²<http://www.mendeley.com>

Table 4: The identified Spearman’s rho correlations and their significance (ρ -/ p -value) of the degree of use in the three organisational strategies across both information spaces

Variable 1	Variable 2	ρ -/ p -values	Variable 1	Variable 2	ρ -/ p -values
physical filing	importance of meaning of labels	0.345/0.001	physical piling	contextual information	0.221/0.039
	annotations as reminders	0.257/0.011		use of post-it notes	0.206/0.039
	annotating in papers	0.212/0.038		difficulty of re-finding in physical file system	0.315/0.003
	annotating in books	0.208/0.040		difficulty of re-finding in digital file system	0.262/0.013
	annotating in books	0.208/0.040	digital piling	difficulty of re-finding in digital file system	0.262/0.013
	annotating by post-it notes	0.290/0.004		longer labels	0.309/0.002
	use of timestamps for re-finding	0.315/0.002		importance of meaning of labels	0.272/0.009
	spatial awareness	0.224/0.033		actively use of annotations	0.457/0.000
physical mixing	purpose of annotations is re-finding	0.202/0.049		use of digital notes	0.366/0.000
	use of timestamps	0.285/0.005		use of digital annotations	0.264/0.008
	timestamps for re-finding	0.271/0.008		purpose of annotations is reminding	0.254/0.012
digital mixing	purpose of annotations is reminding	0.256/0.011		purpose of annotations is re-finding	0.427/0.000
	use of timestamps	0.202/0.047	digital mixing	timestamps for re-finding	0.234/0.021

files and folders, construct broad and deep hierarchies and navigate through the files [24]. The digital space enables users to squeeze more contextual information about an information item into the file system and re-finding becomes easier based on this extra information [6].

4.2 Re-finding Cues in Cross-Media Spaces

In a second part of this unit of analysis, we investigated the applied cues, namely context, spatial and time cues for each organisational strategy in the re-finding activities. By providing an overview of the used cues in a cross-media information space, a better understanding of the differences between physical and digital spaces can be achieved. Table 5 provides an overview of the used cues over both spaces.

Table 5: Overview of the used cues in organisational strategies for cross-media information spaces

		Context cue	Spatial cue	Time cue
Physical space	Filing	✓	✓	✓
	Piling	✓	✓	
	Mixing	✓		✓
Digital space	Filing	✓		
	Piling	✓		
	Mixing	✓		✓

In the following discussion, we provide arguments based on our statistical results and the state of the art in order to validate this general overview of re-finding cues for the specified organisational strategies. The appropriate correlations of the analysed phenomena and their significance can be found in Table 4, where the first variable represents the degree of use of the organisational strategy in the relevant information space and the second variable represents the identified correlated factors. Note that we only found positive correlations but no negative correlations.

When taking a closer look at the piling strategy, we can observe the importance of context cues in both information

spaces and an additional spatial cue is applied in physical space. People use piles to preserve the context of information items as indicated by previous research [2, 7]. At the same time, they use this extra contextual information when retrieving an item. Several researchers agree that the context cue is one of the main advantages of the recall process in human memory [2, 7, 20, 24]. Furthermore, we found evidence for an increased use of contextual information when more physical piles are used. More specifically, a positive correlation is found between piling and the use of post-it notes for keeping contextual information. Next to the context retrieval from post-it notes, the spatial arrangement of piles may also be of value [22]. This arrangement often reflects categorical information of a pile. For example, papers about a specific topic may be placed in one pile while the pile next to it contains papers about a closely related topic. Secondly, the more piles a user has, the harder the re-finding becomes in a digital and physical file system. A reason might be that users with a significant amount of piles get pretty used to the contextual hints in the re-finding process. As shown in related studies [1, 9, 22, 28, 6] and further illustrated by the extra tools for the augmentation of physical file systems mentioned in the previous section, digital as well as physical file systems do not preserve the contextual information of a classified item. In future work, it might be worthwhile to investigate this hypothetical causality. Since also for digital piles mainly contextual cues are used in retrieval, the same reasoning as for physical piles may hold. However, note that the provision and use of the contextual information is quite different in digital piling. An important difference to physical piles is that elements of digital piles are often labelled. Nevertheless, while this labelling is allowed by the definition of a pile in Malone’s work, the whole pile may not be labelled [2]. To preserve the contextual information, these labels become larger with an increased use of digital piles and the meaning of the label gets used more extensively. Furthermore, annotations are used more actively when the digital pile usage increases. With an increasing number of digital piles, users more often use digital and physical annotations to annotate the included information items. On the other hand, they use these annotations

for re-finding and as a reminder. The phenomenon of annotating digital content reflects the need to preserve contextual information. Nonetheless, in contrast to physical piles, we did not find evidence for the use of spatial cues for re-finding activities in digital piles. Since most current operating systems force users to work with a grid layout on their desktop screen, the spatial arrangement of digital piles is often not possible.

Previous research already defined the importance of contextual information to re-find information items when using a filing strategy. Users squeeze this extra information into labels and use the structure of the file system in such a way that it reflects the overall context of use [2, 4]. For example, different structures (e.g. one for student records and one for research publications) may be used to preserve contextual information in a file cabinet, in order to address one of the issues of the classification problem. At the time of classification, users experience a cognitive overload and need to spend extra time which is not preferable by the end user [1]. In addition to these findings, we were able to identify more specific ways on how users keep contextual information in physical file systems. Besides the use of contextual information in the labels, users extensively annotate physical documents with reminders. Furthermore, an intensive use of annotations in papers, books and the use of post-it notes is observed to keep contextual information. This aligns with our previously mentioned results of the open-ended questions about which augmentations users apply to physical filing systems. A second finding is the use of timestamps in re-finding information items. For example, a user starts their search in the file system by recalling when they had last classified the item. This time-related information provides the starting point for orienteering through the file system. A last cue to support when augmenting a physical file system is the spatial cue. Spatial awareness as a cue in re-finding information in file systems has already been introduced by other researchers [1, 22] and our results confirm these findings.

Last but not least, context cues are less used to re-find information in mixtures but we can still observe a significant positive correlation regarding the use of physical annotations for re-finding content in mixtures. In addition to contextual cues, a time cue is used for information retrieval. Users first annotate information items with timestamps. At the time of re-finding an item, they extensively use these timestamps. Similar to the physical mixing strategy, re-finding information in digital mixing is done via a combination of contextual hints and time hints. Users annotate digital content with reminders. Next to this contextual information, they provide the digital content with timestamps and also use these for re-finding information items in digital mixing. Since the mixing strategy has only been introduced in this paper, we can not yet provide further insights that go beyond the survey results mentioned above.

5. DISCUSSION AND IMPLICATIONS

We have presented our investigation and findings on the organisational strategies applied in physical as well as digital information spaces. The novelty of our findings lies in the generalisation of organisational and re-finding activities across these two information spaces. Although the research context of some studies included both information spaces, often their main goal was to learn from paper documents

to design digital interactions [24]. In contrast, we provide a general overview about a user's organisational behaviour in cross-media information spaces.

In our paper we further provide some fundamental findings through the definition of a new organisational strategy called mixing, which is supplementary to the filing and piling strategies defined by Malone [2]. As shown in Figure 3, users do make extensive use of the mixing strategy in both information spaces. In physical space, the degree of use is similar to the use of piling with a normal distribution on a 5-point Likert scale. Similarly, digital mixing is applied to the same degree as digital piling. Nevertheless, we did not find any correlation or dependency between digital filing, digital piling, physical filing and physical piling, which implies that the majority of individual users will not show this similar degree of use of both organisational strategies. In addition, we could identify that for mixtures users apply context cues and a time cue for information retrieval in both information spaces. They first annotate information items with timestamps which are then extensively used in the re-finding activity. This interaction is totally different from observed re-finding activities in filing and piling where a time cue is not the main retrieval cue. Even though we know a lot about the construction and issues in filing and piling [23, 29, 7, 6], there is a serious lack of these findings concerning the complementary mixing strategy. We think that there is a major opportunity for PIM research in trying to understand the mixing strategy to the same degree as we understand the filing and piling strategies. Furthermore, we might investigate whether the mixing strategy can be further refined into multiple more specific strategies.

It is known that most affordances such as the reminder function or the ease to navigate lose their function if the amount of piles grows within a user's physical information space [2, 23]. In addition, we can now state that re-finding in physical piles is the hardest of all applied strategies. Despite this observation, 99% of the users applied piling which indicates that the affordances of piles are very valuable for these users. Furthermore, when users do use piles to a significant amount, they start to get difficulties in re-finding activities in digital and physical file systems. We might hypothesise that due to the increased use of piles, they use less file systems and therefore find it more difficult to retrieve information from a file system. This hypothetical statement can be rejected due to the lack of a negative correlation between filing and piling. Nevertheless, the observation that an increased use of piles leads to difficulties in re-finding information in digital as well as physical file systems might definitely be a subject to be investigated in future descriptive PIM research. In terms of system design, our observations might guide designers to provide additional support in digital and physical file systems when users apply an extensive piling strategy in physical information spaces.

Finally, our results indicate an extensive use of annotations in information stored in almost all organisational strategies across both information spaces with the purpose of re-finding and as reminders. Only for the physical piling strategy we did not find any correlations with annotations. This could be explained by the fact that piles do have a reminder function as part of their affordances [23, 8, 9] and by the ease of re-finding when the degree of use is low enough [2]. Due to the lack of this naturally supported reminder and re-finding functionality in digital piling as well

as the filing and mixing strategies, users try to add these to the organisational units. Future work might investigate how we can provide better support to users for adding this functionality to the information space.

5.1 Directions for Future PIM System Design

Besides the presented descriptive contributions, our findings can also be used to direct future PIM system design.

Providing Suitable Augmentations for Each Strategy

While the majority of PIM prototypes such as HayStack [30] or Gnowsisi [31] focus on the ability to link all kinds of digital media in order to improve organisational and re-finding activities, there is a lack of innovative approaches to support a user's existing organisational behaviour. Since users have constructed and are used to their own organisational structures—such as files, piles and mixtures—over their lifetime, we should not ignore these behavioural patterns in PIM systems. We should rather try to provide a unification of new innovative personal information interaction techniques and a user's existing organised information space. Due to the lack of dependencies or correlations between the organisational strategies, designers can develop specific augmentations for each individual strategy in both digital and physical information spaces. A first step in this direction has already been taken by Fitchett [32] where the digital file system has been augmented with guidance to improve re-finding. Nevertheless, we might also adapt this guidance for piles and mixtures in the digital file system. We aim for a new type of PIM system with extensible user interfaces which can be adapted for all organisational strategies across tools and information spaces based on a user's profile.

Support for Cross-Media Re-finding Activities

As shown by earlier research [23, 8, 24] and confirmed by our results, a user's personal information space still consists of a significant amount of physical artefacts. Not only do we still use paper to an extensive degree but the presented work also indicates that users also apply the three organisational strategies in physical space to a significant degree. While current research investigates the augmentation of bookshelves [33] or interactive desks with support for interaction with piles [34], we need to push the augmentation for physical workspaces. As observed in the results, all users had some degree of piles on desks and in bookshelves whereas this was the most difficult organisational unit to re-find information in. Additionally, we have shown that users do cross the borders of digital and physical information spaces in order to improve their organisational activities. Physical file systems are augmented with post-it notes or stickers as well as with digital indexes. On the other hand, users make physical annotations to improve the context preservation in all three strategies, over the physical as well as digital space. By recognising this annotation process in the presented work, we can direct future PIM systems to provide support for cross-media annotations.

Support Appropriate Re-finding Cues in the Right Strategy

Due to the generalisation of our work to the domain of cross-media information spaces, we were able to provide an overview of the retrieval cues in both the physical and the digital space. Future PIM systems can use this overview to integrate the appropriate support for the right strategy in their user interface design. Although advanced search func-

tionality is available in the digital information space, it is observed that users only use search as a last resort [35]. In addition, caution needs to be given to the fact that time-based retrieval cues are not often used in digital file systems. While current operating systems provide temporal metadata in order to search for digital media, we should better focus on augmenting the digital file system with contextual information. This viewpoint is in line with previous work where it has been recognised that contextual information is squeezed into the folder hierarchy [6]. On the other hand, the augmentation of physical file systems should integrate this time-based support for re-finding activities, such as when a document has been last accessed or modified, as well as references to documents that have been accessed during the time interval. Additionally, physical and digital mixtures should be augmented with time-based search functionality. However, since we only defined the mixing organisational strategy in our case study, we will need some further research in order to provide causal relationships and insights on how and why users apply these time-related activities when organising and re-finding information.

6. CONCLUSION

We have presented a case study on organisational strategies for personal cross-media information management. In addition to the identification of mixing as a new organisational strategy, we have shown that the mixing strategy should be treated as a complement to the filing and piling strategies which are used to describe the most prominent PIM order variable. Further, the results of our study provide insights about a user's overall organisational landscape and the applied re-finding cues. Last but not least, based on our findings on organisational strategies for cross-media information spaces we identified major opportunities for future PIM research and outlined a number of directions for future PIM system design.

7. REFERENCES

- [1] I. Cole, "Human Aspects of Office Filing: Implications for the Electronic Office," in *Proc. Human Factors and Ergonomics Society Annual Meeting*, Seattle, USA, pp. 59–63, October 1982.
- [2] T. W. Malone, "How People Organize Their Desks? Implications for the Design of Office Information Systems," *ACM Transactions on Office Information Systems (TOIS)*, vol. 1, no. 1, pp. 99–112, 1983.
- [3] O. Bergman, "Variables for Personal Information Management Research," *ASLIB Proceedings: New Information Perspectives*, vol. 65, no. 5, pp. 464–483, 2013.
- [4] D. Kirsh, "A Few Thoughts on Cognitive Overload," *Intellectica*, vol. 30, no. 1, pp. 19–51, 2000.
- [5] W. P. Jones, *Keeping Found Things Found the Study and Practice of Personal Information Management*. Morgan Kaufmann Publishers, 2008.
- [6] W. Jones, A. Phuwanartnurak, R. Gill, and H. Bruce, "Don't Take My Folders Away!: Organizing Personal Information to Get Things Done," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Portland, USA, pp. 1505–1508, April 2005.

- [7] M. Lansdale, "The Psychology of Personal Information Management," *Applied Ergonomics*, vol. 19, no. 1, pp. 55–66, 1988.
- [8] S. Whittaker and J. Hirschberg, "The Character, Value and Management of Personal Paper Archives," *Transactions on Computer-Human Interaction*, vol. 8, no. 2, pp. 150–170, 2001.
- [9] D. O. Case, "Conceptual Organization and Retrieval of Text by Historians: The Role of Memory and Metaphor," *Journal of the American Society for Information Science*, vol. 42, no. 9, pp. 657–668, 1991.
- [10] S. Whittaker and C. Sidner, "Email Overload: Exploring Personal Information Management of Email," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Vancouver, Canada, pp. 276–283, April 1996.
- [11] O. Bälter, "Strategies for Organising Email," *People and Computers XII*, pp. 21–38, 1997.
- [12] S. Whittaker, T. Matthews, J. Cerruti, H. Badenes, and J. Tang, "Am I Wasting My Time Organizing Email? A Study of Email Refinding," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Vancouver, Canada, pp. 3449–3458, May 2011.
- [13] D. Abrams, R. Baecker, and M. Chignell, "Information Archiving with Bookmarks: Personal Web Space Construction and Organization," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Los Angeles, USA, pp. 41–48, April 1998.
- [14] R. Boardman and A. Sasse, "'Stuff Goes Into the Computer and Doesn't Come Out?': A Cross-Tool Study of Personal Information Management," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Vienna, Austria, pp. 583–590, April 2004.
- [15] M. Van Kleek, W. Styke, D. Karger, and m. c. schraefel, "Finders/Keepers: A Longitudinal Study of People Managing Information Scraps in a Micro-Note Tool," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Vancouver, Canada, pp. 2907–2916, May 2011.
- [16] M. Haraty, D. Tam, S. Haddad, J. McGrenere, and C. Tang, "Individual Differences in Personal Task Management: A Field Study in an Academic Setting," in *Proc. Graphics Interface Conference*, Toronto, Canada, pp. 35–44, May 2012.
- [17] D. Barreau, "Context as a Factor in Personal Information Management," *Journal of the American Society for Information Science*, vol. 46, no. 5, pp. 327–339, 1995.
- [18] R. Yin, *Case Study Research Design and Methods*. Sage Publications, 2009.
- [19] S. Toulmin, *The Use of Argument*. Cambridge University Press, 1958.
- [20] A. Kidd, "The Marks Are on the Knowledge Worker," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Boston, USA, pp. 186–191, April 1994.
- [21] B. Kwasnik, "How a Personal Document's Intended Use or Purpose Affects Its Classification in an Office," in *Proc. ACM Conf. on Research and Development in Information Retrieval*, Cambridge, USA, pp. 207–210, June 1989.
- [22] W. Jones and S. Dumais, "The Spatial Metaphor for User Interfaces: Experimental Tests of Reference by Location Versus Name," *ACM Transactions on Information Systems (TOIS)*, vol. 4, no. 1, pp. 42–63, 1986.
- [23] A. Sellen and R. Harper, *The Myth Of The Paperless Office*. MIT Press, 2001.
- [24] O. Bondarenko and R. Janssen, "Documents at Hand: Learning from Paper to Improve Digital Technologies," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Portland, USA, pp. 121–130, April 2005.
- [25] D. Barreau and B. Nardi, "Finding and Reminding: File Organization From the Desktop," *ACM SIGCHI Bulletin*, vol. 27, no. 3, pp. 39–43, 1995.
- [26] S. Henderson and A. Srinivasan, "Filing, Piling and Structuring: Strategies for Personal Document Management," in *Proc. Hawaii Int. Conf. on System Sciences*, Kauai, USA, pp. 1–10, January 2011.
- [27] W. Jones and J. Teevan, *Personal Information Management*. University of Washington Press, 2007.
- [28] S. Dumais and T. Landauer, "Using Examples to Describe Categories," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Boston, USA, pp. 112–115, 1983.
- [29] S. Whittaker, "Personal Information Management: From Information Consumption to Curation," *Annual Review of Information Science and Technology*, vol. 45, no. 1, pp. 1–62, 2011.
- [30] D. Karger, K. Bakshi, D. Huynh, D. Quan, and V. Sinha, "Haystack: A Customizable General-Purpose Information Management Tool for End Users of Semistructured Data," in *Proc. 1st Biennial Conf. on Innovative Data Systems Research*, Asilomar, USA, January 2003.
- [31] L. Sauermann, "The Gnowsiss Semantic Desktop for Information Integration," in *Proc. 1st Intelligent Office Appliances*, Tuscon, USA, pp. 39–42, February 2005.
- [32] S. Fitchett, A. Cockburn, and C. Gutwin, "Improving Navigation-Based File Retrieval," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Paris, France, pp. 2329–2338, April 2013.
- [33] K. Matsushita, D. Iwai, and K. Sato, "Interactive Bookshelf Surface for In Situ Book Searching and Storing Support," in *Proc. Int. Conf. Augmented Human*, Tokyo, Japan, pp. 9–17, March 2011.
- [34] M. Khalilbeigi, J. Steimle, J. Riemann, N. Dezfuli, M. Mühlhäuser, and J. D. Hollan, "ObjecTop: Occlusion Awareness of Physical Objects on Interactive Tabletops," in *Proc. ACM Conf. Interactive Tabletops and Surfaces*, St. Andrews, UK, pp. 255–264, October 2013.
- [35] J. Teevan, C. Alvarado, M. S. Ackerman, and D. Karger, "The Perfect Search Engine is Not Enough: A Study of Orienteering Behavior in Directed Search," in *Proc. ACM Conf. on Human Factors in Computing Systems*, Vienna, Austria, pp. 415–422, April 2004.