Using card sorts to understand how users think of personal information

Paul Thomas
CSIRO, Canberra
paul.thomas@csiro.au

David Elsweiler
University of Regensburg
david@elsweiler.co.uk

ABSTRACT

Understanding how users think of personal information, and how they mentally categorise or classify the objects they work with, should inform the design of personal information management (PIM) or personal retrieval systems. However, most investigations of this topic predate widespread multimedia, websites, and social media—objects that a contemporary PIM or retrieval system should work with.

We describe a pilot study that has used a variant of card sorts to elicit categories for personal information such as files, email, tweets, and websites. Our early results suggest that there are common categorisations which are not yet supported by PIM software, but which might reward further work. Our results also suggest that—with some caveats—card sorts are useful for understanding users' categories.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces
General Terms: Human Factors
Keywords: Facets, classification, card sorts

1. INTRODUCTION

Tools for personal information management (PIM) and search support the archival, retrieval, and management of "personal" data: the files, email, photos, videos, and other digital objects a person creates or uses [15]. Several studies show that PIM can be challenging [5, 11, 15] and it has been suggested that tools could be easier to use and more useful, if the way they represent objects matches the way users think of them [10, 20].

We are interested in how users think of the wide range of digital objects they interact with—the objects conventionally considered by PIM tools, objects less commonly considered such as websites and applications, and newer objects such as messages from social media. There are three linked questions:

1. What properties do users think "personal" digital objects have? That is, in which ways do users think of the objects they use?
2. Can we expose these properties in a PIM or search tool? Can the properties of an object be determined algorithmically? How should the properties be presented?
3. Assuming we can expose some or all of these properties, would we expect that to make management or retrieval easier?

At present, we are considering the first question. In particular, in this work, we have experimented with card sorts to elicit users' own descriptions of personal information.

2. CLASSIFICATION AND TOOLS

Past work has investigated the properties users assign to files, and elicited categorisation schemes. This work has not however considered as wide a range of object types as we do here; we may expect that with different types, sources, and quantities we would see different categorisations. Existing PIM and file management systems also support, or impose, particular faceting schemes.

2.1 Classifications

There is a rich tradition, in information science and information behaviour, of studies that try to understand how people organise, classify, and think about their information—that is, how people understand their information independent of any particular software capabilities or restrictions.

Three studies of note are by Cole, Kwasnik and Case. Cole [7] studied how 30 office workers classified their document collections. Six aspects of documents were important in filing decisions: "type", "form", "volume", "complexity", "functions", and "levels of information". Similarly, Kwaśnik [19] examined the categorisation behaviour of eight researchers and identified seven dimensions: "situation", "document", "disposition", "order/scheme", "time", "value", and "cognitive state". Case [6] investigated the behaviour of twenty historians and identified three main factors by which objects were classified in offices; "ease of access", "form" and "topic". While there is considerable overlap in the findings of these studies, particularly the criteria "form" and "topic", the studies all predate the rich digital landscape we have today, and focus on physical information objects.

Other research relating to our work has tried to learn about how people think about digital information by investigating how they behave with information in practice. For example, people have been shown to organise email messages and files based on projects [17, 23] and prefer to refind objects by location than using search facilities [2, 3]. These kinds of studies provide strong hints at how people may think about

digital information, but are influenced by the tools they have available to them.

More recently, Gonçalves and Jorge [13] asked participants to tell stories about three of their personal documents by describing each, from memory, in terms of its features, its content and the context in which it was created or used. It was discovered that time, location, and purpose of the document were the most common attributes used in stories. Similarly, Blanc-Brude and Scapin [4] used semi-structured interviews to examine participants' recollection of their documents. They found that location, format, time, keywords and associated events were remembered most frequently, but many of these attributes, particularly keywords, time and location were often only partially remembered or the recollections offered by the participants were incorrect. Both of these studies add a rich understanding of how people perceive their documents by examining a small number of documents in great detail, but do not explore how documents are related.

Our aim here is to add to and complement this previous work by using a technique that can deal with rich variety of information objects we interact with today; be tool agnostic; and allow insight into how different documents can be associated in different ways. We would also like to understand whether this has any impact on the design of PIM tools.

2.2 Tools

Tools for desktop search typically expose not just filesystem attributes such as name, size, and timestamp, but also extracted metadata. For example, Phlat [8] uses title, date, author, recipient, media type and tags; Haystack [1] has extensible facets but the authors have discussed media type, people named in email, text in a document, and URL.

More elaborate PIM tools have exposed other attributes to support different interactions. Some, such as Lifestreams [12], have supported time-based browsing and searching; an interface to Stuff I've Seen [21] extended this by indexing documents according to contemporaneous events. Other tools have taken a more personal view of time, or document lifecycle, and supported information management by context. Here, objects are organised according to tags for the context in which they are used [16, 18] or grouped according to patterns of use [9].

These systems offer variety of projections, across a number of media and storage types, but it is hard to know whether these match the way people naturally think of their objects. Alternative presentations that were natural for users, and easy to implement, would be worth further thought.

3. METHOD

In this work, we have experimented with repeated single-criterion card sorts to elicit users' mental categorisations. Cards represented digital objects on each participant's computer.

3.1 Card sorts

Repeated single-criterion card sorts—or just "card sorts"—are a common technique for eliciting users' categorisations (see e.g. Rugg and McGeorge [22] for an overview). Compared with interview-based techniques, card sorts are less flexible but are very lightweight: in our experience participants grasped the idea very quickly, many found it enjoyable, and the entire protocol took little time. Coding card sorts for later analysis is also relatively straightforward.

In a typical exercise, each participant is given a number of cards, each representing an object or a concept. They are asked to partition these cards according to any criteria they like; the criteria used for the sort, the categories (piles or sets), and the cards in each category are recorded. This is repeated several times, with participants suggesting a different criteria each time. For example, given cards labelled as follows:

1. pig; 2. chicken; 3. snake; 4. horse; 5. spider

a participant may sort cards according to the criterion "raised on a farm", with cards 1, 2, and 4 in category "yes" and cards 3 and 5 in category "no". A second sort, according to the criterion "where eaten", might have cards 1 and 2 in category "almost everywhere", card 3 in category "Asia", card 4 in category "Asia and Europe" and card 5 in category "don't know".

Records of the sorts may then be analysed with qualitative or quantitative methods.

3.2 Our approach

For this early experiment, we recruited a convenience sample of ten participants from two institutions. All were heavy computer users.

As preparation for the experiment participants were asked to select several information objects they had seen, used or created in the recent past. An "information object" was defined by giving as examples computer files, emails, websites, tweets or Facebook updates, documents or articles read, photographs or images, videos, music, and computer applications. However, participants were not restricted to these objects and could choose anything they wanted using these as a guideline. We encouraged participants to label 10 to 15 cards, which we believe balances the need for broad coverage with practical limits on participants' time.

While choosing objects, participants were asked to create index cards with the name of each (or some other reminder of its identity or contents).

With cards made, each participant was introduced to sorts using a set of cards with pictures of buildings; they were taken through some example sorts which included criteria clear from the pictures themselves (colour, material), criteria which were not immediately clear (insulation), criteria which were subjective (good place for a party). "Can't tell" or "don't know" categories were included in these examples.

The participants' own cards were then used for repeated sorts. Participants were asked to make piles according to a criterion of their choosing, and we noted the criterion (sometimes this was implicit), the categories used, and the cards in each category.

After collecting individual classifications from all participants, the full dataset, i.e. the criteria used to associate information objects, was analysed qualitatively using an affinity diagramming technique. This is a group-based process, which allows the discovery and validation of patterns in the data [14]. The researchers, as a team, looked for patterns in the data and grouped related criteria; we then related the formed groups in a way that creates a hierarchical coding scheme.

4. RESULTS OF THE SORTS

The results from this pilot are promising. Card sorts elicited a variety of criteria; some of these differ from those seen before, and many are not well supported by PIM or retrieval tools.

Group	Participants
document lifecycle	3
events	1
object's form	7
object's affective qualities	2
object's cost and value	3
object static/dynamic distinction	2
people and community	4
properties of associated tasks	4
topics covered	5
work/leisure distinction	10
(three other object-related groups)	3

Figure 1: Groups of criteria at the top of our heirarchy. Numbers are the number of participants who used each criteria at least once—note that some participants used some criteria, or criteria in the same group, more than once.

4.1 General observations

Our ten participants provided 64 sorts, a median 5.5 sorts each (first/third quartile 5.0/8.0 sorts each). At the leaves of our hierarchy, there was in general little overlap: 13 of 25 criteria were used by only one participant. However, 12 were used by two or more participants, 8 by three or more, and one criteria (discussed below) was used by every participant for at least one sort. Figure 1 summarises how many participants were represented in each top-level group, that is each group at the top of our hierarchical coding.

The single most common criterion was a distinction between objects used for work and objects used for leisure—all our participants used this criteria, and typically early on. Following this there were four common groups: to do with the object itself, especially the form (data type and other surface features), which 7 participants used at least once; the topics an object is connected with (5 participants); the properties of tasks associated with an object (4); and criteria describing people and community (4). A striking finding is the diversity in the criteria derived by the participants. Although we were able to group the criteria into 13 distinct high-level cateorgies, only two of these, work/leisure and form, were named by more than half the population.

After our initial grouping, which was based on labels proffered by participants and not on any statistics of the sorts themselves, there were no clear correlations between criteria—that is, "work" did not look the same as "important" or "Word files" as "makes me angry".

On our analysis, five of Kwaśnik's seven groups were represented in our data: situation, document, time, value, and cognitive state. However, they were very unevenly distributed: criteria we classified as "situation" were used by all ten participants, for one or two sorts each; "document" was used by nine participants, for a median 3.5 sorts each; while at the other end of the scale, "time" was used by only three participants and "cognitive state" was used by two participants, once each in each case.

4.2 Criteria, groups, and tool support

In many previous studies two groups of criteria—form and topic—were found to be central, and our data reinforces this. There are also, however, notable contrasts.

Work/leisure. To the best of our knowledge, previous studies of classification behaviour have not found a work/leisure distinction. However, every participant in our sample used this criteria. This may be because we most of the objects in our study were digital, not paper documents—it is very easy to mingle work- and leisure-related objects online—but it is clearly important and is not explicitly supported by PIM tools.

One participant reported that he used two top-level folders in his file system, and two email accounts, to keep work and leisure information separate. No other participants reported as clear a distinction, however. It should be possible in a PIM/search tool to tag files, or e.g. learn a classifier, to help maintain this distinction. Distinguishing work from leisure contexts might also allow different technologies to be used in each case.

People. Our participants did associate their objects in terms of specific people, but not in the way we might have expected. Rather than linking objects to particular, specified people, our dimensions relate to relationships with the community: "popular with many people", for example, "things I will/won't talk about", or "involvement of other people". Unfortunately it is not clear how a PIM or search tool could support this sort of classification.

Task, time, and workflow. Users in our study did not group objects by particular tasks—objects related to task A, to task B, etc—but four users did group objects according to whether an object had an associated task, and by properties of that task (state, importance, and cost or difficulty). This could be used to extend the work of Jones and his colleagues [16], who advocate project organisation, but do not allow tasks within projects to be annotated with properties such as cost or importance.

Time was mentioned by four participants. However rather than categorise objects according to time of use (or receipt), as supported by a number of tools, three participants derived criteria from the lifecycle of an object. Criteria such as "when I need to act on this" or "when this is important" will change over time for each object. This is related to Cole's "level of information" dimension. Only one participant used objects' importance to an event, at a particular time, as a criteria. Tools which support an explicit notion of document or task lifecycle, or approximate this e.g. by recording patterns of use, would presumably suit these participants.

5. DISCUSSION OF THIS APPROACH

The results above suggest that card sorts, in this variant, are useful for eliciting criteria: it does seem possible to gain some insight into how users think of personal objects, and how we might support this. This pilot has, however, highlighted some limitations.

The objects represented by each participant's cards were familiar—that is, they tended to choose objects they had used recently or frequently. They were also selected for sharing, since although we did not record the card titles we did see them. We cannot be sure that the chosen objects represent the sorts of things users may search for in a PIM system, and of course they are not representative of *un*familiar objects. We could instead choose objects from a participant's computer, for example by choosing randomly from the file

system and labelling cards with file icons and names, and similar. There is a tradeoff, however: if participants did not recognise these objects, the only possible criteria would be file icons and names, and we would learn little. By allowing users to choose their own objects the cards hopefully acted as prompts for other, richer, associations.

On a related point, some media types, such as video or audio, are difficult to represent on cards. It is not clear what this means for eliciting criteria. We are possibly unlikely to get criteria such as "out of focus" (for photos) or "scratchy bit in the middle" (for audio), but participants' familiarity with the objects may mitigate this to some extent.

There are of course properties that are not captured by this method: links between documents, for example (except implicit links of the type "sorted into the same pile"). It is also possible that our presence, and the apparatus we used, made it hard for participants to think naturally. They may have been inspired to create other categories if prompted (as in Gonçalves and Jorge [13]); on the other hand, our approach has the advantage that we can see which properties were immediately obvious.

We also note that the objects chosen varied greatly from participant to participant, and this may have played a role in the criteria that were chosen—although we did see some overlap, possibly there would have been more if the objects were more similar. We are considering constraining participants more in future, for example by prompting them to make a certain number of cards for each media or perhaps having subpopulations sort a shared set of cards, e.g. emails, web pages etc. they have all seen or received.

6. CONCLUSIONS

We hope to extend this work by scaling to a larger group of participants, but we will consider some methological changes: constraining the objects chosen, for example, or careful prompts to elicit more classifications. Nonetheless card sorts, in the variant here, have proved useful for starting to understand how users think of "personal" digital objects from a wide range of sources and media. Some classifications were both common and expected, but we did observe interesting differences both with the criteria found in earlier studies and with the criteria exposed in PIM and search tools.

7. ACKNOWLEDGEMENTS

We would like to thank our participants for their time.

8. REFERENCES

[1] E. Adar, D. Karger, and L. A. Stein. Haystack: Per-user information environments. In *Proc. CIKM*, pages 413–422, 1999.

[2] D. K. Barreau and B. Nardi. Finding and reminding: File organization from the desktop. *ACM SIGCHI Bulletin*, 27(3):39–43, 1995.

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

[4] T. Blanc-Brude and D. L. Scapin. What do people recall about their documents?: Implications for desktop search tools. In *Proc. IUI*, pages 102–111, 2007.

[5] R. Boardman and M. A. Sasse. "Stuff goes into the computer and doesn't come out": A cross-tool study of personal information management. In *Proc. CHI*, pages 583–590, 2004.

[6] D. O. Case. Conceptual organization and retrieval of text by historians: The role of memory and metaphor. *JASIST*, 42(9):657–668, 1991.

[7] I. Cole. Human aspects of office filing: Implications for the electronic office. In *Proc. Human Factors Society*, 1982.

[8] E. Cutrell, D. Robbins, S. Dumais, and R. Sarin. Fast, flexible filtering with Phlat—personal search and organisation made easy. In *Proc. CHI*, pages 261–270, 2006.

[9] P. Dourish, W. K. Edwards, A. LaMarca, J. Lamping, K. Petersen, M. Salisbury, D. B. Terry, and J. Thornton. Extending document management systems with user-specific active properties. *ACM Trans. Inf. Syst.*, 18(2):140–170, 2000.

[10] D. Elsweiler. *Supporting Human Memory in Personal Information Management*. PhD thesis, The University of Strathclyde, 2007.

[11] D. Elsweiler, M. Baillie, and I. Ruthven. What makes re-finding information difficult? A study of email re-finding. *Proc. ECIR*, 6611:568–579, 2011.

[12] E. Freeman and D. Gelernter. Lifestreams: a storage model for personal data. *SIGMOD Record*, 25(1):80–86, 1996.

[13] D. Gonçalves and J. A. Jorge. Describing documents: what can users tell us? In *Proc. IUI*, pages 247–249, 2004.

[14] J. Hackos and J. Redish. *User and Task Analysis for Interface Design*. 1998.

[15] W. Jones. Personal information management. *Annual Review of Information Science and Technology*, 41(1):453–504, 2007.

[16] W. Jones, H. Bruce, A. Foxley, and C. Munat. The universal labeler: Plan the project and let your information follow. In *Proc. ASIST*, 2005.

[17] W. Jones, H. Bruce, E. Jones, and J. Vinson. How do people keep and re-find project-related information? In *Proc. SIGCHI*, 2010.

[18] V. Kaptelinin. UMEA: translating interaction histories into project contexts. In *Proc. CHI*, pages 353–360, 2003.

[19] B. H. Kwaśnik. How a personal document's intended use or purpose affects its classification in an office. In *Proc. SIGIR*, pages 207–210, 1989.

[20] M. Lansdale. The psychology of personal information management. *Appl Ergon*, 19(1):55–66, 1988.

[21] M. Ringel, E. Cutrell, S. Dumais, and E. Horvitz. Milestones in time: The value of landmarks in retrieving information from personal stores. In *Proc. INTERACT*, pages 184–191, 2003.

[22] G. Rugg and P. McGeorge. The sorting techniques: A tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems*, 22(3):94–107, 2005.

[23] S. Whittaker and C. Sidner. Email overload: exploring personal information management of email. In *Proc. CHI*, pages 276–283, 1996.