

The Function and Future of the Folder

ROGER WHITHAM* AND LEON CRUICKSHANK

Imagination Lancaster, LICA Building, Lancaster University, Bailrigg, Lancaster, LA1 4YW, UK

**Corresponding author: r.whitham@lancaster.ac.uk*

Folders are a commonplace metaphor in computing environments, constituting a link to physical work materials and are a key means for individuals to impose order on their digital work materials. This paper presents the findings of a novel qualitative study examining folder use by 12 information workers, using logging to accurately capture how folders were used in individual everyday work over 6 weeks, and challenging participants to work without using folders. Through observation and interviews, the study provides new descriptions of how folders are used and the dependence some study participants had on their folders to think and create, as well as to access files. The findings call into question whether search and recency-based lists of files could fulfil the functional role of folders, identified as key means for individuals to construct and specialize their work environments. Implications are discussed for document management tools, and more generally for operating system design.

RESEARCH HIGHLIGHTS

- Folders are often considered an unwelcome legacy of physical information technologies.
- Views of folder structures can offer valuable perspectives on personal information in themselves, independent of the content of the files they contain.
- Folders can be strongly intertwined with the process of information work, both in the computer and in the mind of the worker.
- Search and recency-based views of files do not eliminate the need to creatively categorize and combine information.
- Eliminating folders and manual organization of files may compromise information worker's ability to control their environment, and in turn thinking and working processes.

Keywords: information retrieval; document management; interaction paradigms; HCI theory concepts and models HCI theory; concepts and methods; HCI design and evaluation methods

Editorial Board Member: Jacek Gwizdka

Received 30 November 2015; Revised 28 November 2016; Editorial Decision 23 December 2016

1. INTRODUCTION

Within *desktop metaphor* computing environments such as Microsoft Windows and Apple OS X, folders are a fundamental resource available for imposing structure on personal information. User-manipulable folders are found in various contexts such as media libraries, web bookmark managers and email clients. Folders provide a means to organize and navigate information, and represent one of the most common and basic metaphors of contemporary digital environments.

Extensive literature within the field of Personal Information Management (PIM) describes and categorizes how individuals use file systems and folders. Folders are known to sometimes have great value to those who use them, but their precise function and potential future is not clear. On one hand, folders are

positioned as a means to an end for storing and accessing information or an inefficient mechanism for laboriously classifying information and later retrieving it (Cutrell *et al.*, 2006a,b; Whittaker, 2011). On the other hand, folders are strongly valued by those who use them and have been shown to prompt action (Barreau and Nardi, 1995), lighten the load of accessing files (Teevan *et al.*, 2004), and have a role in breaking down problems and planning work (Jones *et al.*, 2005). When presented with increasingly powerful search tools, individuals appear reticent to relinquish folders as a means for accessing their own information, even as search has become the *de facto* means of accessing shared digital information (Barreau and Nardi, 1995; Bergman *et al.*, 2008a,b; Boardman and Sasse, 2004; Jones *et al.*, 2005; Teevan *et al.*, 2004).

Motivated by this apparent tension in PIM literature, this paper presents a study that explores the functions of folders in use through detailed observation of the role they play in the hands of information workers in everyday work. To capture this study examines how and why folders are created and manipulated in folder structures, and how and why they are viewed during work within windows in the desktop computing environment. Through naturalistic observation facilitated by specialized software and in-depth interviews, the study provides detailed descriptions of how folders are used, grounded in real-world processes of work. The study software also involves an experimental probe in which study participants attempt to access their files without using folders.

Following a review of previous work, this paper presents the study methodology and findings. The current and potential value of folders are discussed, specifically with reference to the control they offer over the work environment as a whole and the processes of conceptual and practical work that take place within it. Implications are drawn for file organization tools and for the potential of search to support information work.

2. PREVIOUS WORK

The *information item* is often advanced as the basic unit which is subjected to organizational and managerial effort in PIM (Jones, 2007). Folders provide a means to group information items (such as files, emails, bookmarks and folders themselves) and so impose order and meaning on the information. The folders found in today's *desktop metaphor* computing environments metaphorically reference the file folders of physical filing systems, and were designed as a familiar means to enable individuals to tailor the organization of their electronic documents (Smith *et al.*, 1982). The analogous relationship between physical and virtual folders and the act of filing remains commonplace in descriptions of PIM activity; folders provide a means of manually organizing information, and a means of navigating an archive when accessing information in the future.

2.1. Folders for the classification and storage of information items

Various studies have investigated practices around organizing personal information, originating with work conducted by Malone in physical office environments (Malone, 1983). Subsequent work has investigated folder use within desktop computer work environments around files and file systems (e.g. Barreau and Nardi, 1995; Bergman *et al.*, 2010; Dourish *et al.*, 1999; Jones *et al.*, 2005; Teevan *et al.*, 2004), email messages (e.g. Bellotti *et al.*, 2003; Ducheneaut and Bellotti, 2001; Whittaker and Sidner, 1996; Whittaker *et al.*, 2011) and web bookmarks (Abrams *et al.*, 1998; Boardman and Sasse, 2004).

Examining physical and digital document classification practices respectively, Kwasnik and Barreau emphasize the variety of organizational approaches used, both between individuals and between different contexts for the same individual (Barreau, 1995; Kwasnik, 1989). Unlike a shared organizational system which must be understood by multiple people, the personal nature of PIM activities allow for highly individual organizational systems that respond to contextual needs (Kaye *et al.*, 2006; Massey *et al.*, 2014). Barreau describes *satisficing* strategies, in which individuals only engage in the minimum necessary organizational effort to meet their needs (Barreau, 1995). Like the organizational systems themselves, these needs vary between individuals; some prefer to defer organizing items into folders, while others prefer to organize items as soon as they are acquired, leading to the *piler and filer* behaviour categories first described by Malone (Boardman and Sasse, 2004; Kidd, 1994; Malone, 1983; Rao *et al.*, 1994; Whittaker and Sidner, 1996). Boardman and Sasse's work revealed how individual's folder structures evolve and morph over time, with new folders being created regularly during work (Boardman and Sasse, 2004).

2.2. Folders for accessing information items

Once created, folders play a key role in accessing the information items within them. Individuals are apt to know their folders well and in several PIM studies have been shown to be highly successful in retrieving items (Alvarado *et al.*, 2003; Barreau and Nardi, 1995; Bergman *et al.*, 2010, 2012; Boardman and Sasse, 2004). The process of navigating to an item through folders is not always free of errors, but overall rates of success are high (Bergman *et al.*, 2010).

Two decades after Barreau and Nardi's studies individuals' preference for manually navigating folders over using search tools to access files remains a consistent finding of empirical work (Barreau and Nardi, 1995; Bergman *et al.*, 2008a,b; Boardman and Sasse, 2004; Teevan *et al.*, 2004). Jones *et al.* found that the preference for folder navigation was strong for files, but less so for emails and web pages (Jones *et al.*, 2014). When accessing files, though folder navigation is potentially slower than search, it is recognized as providing more information to individuals about the items they encounter, and placing a lower cognitive load on individuals when compared to using search (Bergman *et al.*, 2008a,b, 2013a,b; Teevan *et al.*, 2004). Recent work by Benn *et al.* suggests a neurological basis for the preference for navigation over search, identifying higher planning and attentional demands associated with forming a search query compared with navigation (Benn *et al.*, 2015).

2.3. Folders in planning and coordinating activity

Alongside the functions associated with storing, organizing and accessing files, PIM research also recognizes that folder structures and the items within them can play higher or

meta-level roles in individual work. Malone identified the reminding functions of documents positioned within physical office environments (Malone, 1983), a function also identified by Barreau and Nardi in their studies of file organization with desktop computers (Barreau and Nardi, 1995). Barreau and Nardi describe participants in their studies relying on the presence of items within specific folders to remind them to take action, a strategy identified in subsequent PIM studies (e.g. Bellotti *et al.*, 2004; Whittaker and Sidner, 1996). In addition to reminding of specific actions and specific documents, folders can also provide contextual information relating to activities and tasks. Jones *et al.* describe individuals creating folder structures to decompose problems or projects into discrete categories for information items (Jones *et al.*, 2005, 2006). While such folders might not contain information items when first created their presence as a placeholder offers valuable activity-related information in itself.

2.4. Alternatives to folders

There is discussion within PIM of whether folders are necessary at all in the context of improved search technologies (Cutrell *et al.*, 2006a,b), and automated classification tools (Whittaker, 2011). Enhanced search and filtering tools such as *Stuff I've Seen* (Dumais *et al.*, 2003) and *Phlat* (Cutrell *et al.*, 2006a,b) bring with them the possibility that folders may no longer be required to ensure items can be reliably accessed across time. Earlier PIM work suggested the possibility of automatically classifying items with temporal metadata (Freeman and Gelernter, 1996; Malone, 1983). Kaptelinin proposed a stronger form of automated classification based on explicitly-defined project contexts (Kaptelinin, 2003), though PIM studies highlight distrust of automated classification approaches by users (Bellotti *et al.*, 2003; Whittaker and Sidner, 1996). Alternatives to folders also exist in the form of attribute-based classification tools, such as *Presto* (Dourish *et al.*, 1999) or *Haystack* (Huynh *et al.*, 2002). Akin to the tagging available in *Phlat*, these approaches allow an item to belong to multiple user-defined categories and allow queries to be formed that use these categories. Tags provide an alternative to folders, but are not widely preferred in the desktop computing environments (Bergman *et al.*, 2013a,b).

3. RESEARCH AIMS AND APPROACH

The research presented in this paper aims to (i) investigate in detail how folders are used in everyday information work, and (ii) identify the value folders hold for those who use them. Our approach examines how folders are used to interact with files, both by creating persistent groupings as *folder structures*, and through *folder views* used to browse folders and their contents. The approach reveals not only how folder

structures change over time, but also how folder structures are accessed over time through folder view windows within the desktop computing environment. We seek to provide new, detailed descriptions of folder use, and to respond to the ongoing tension between advances in search technology and the continued preference individuals have for using folders to interact with their files. The work also responds to calls for new methods and data within PIM suitable for describing and understanding interactions with information items (Jones, 2007; Whittaker, 2011).

To meet these aims a study was designed to facilitate naturalistic observation of folder use. Following an approach inspired by Hollan *et al.*'s (2000) cognitive ethnography, the study involves automated recording of interactions with folder structures and folder views, followed up by detailed quantitative and qualitative examination of these records and in-depth interviews with study participants. After a period of pure observation, the study also introduced interventions into the environments of participants who were invited to stop using folder views and instead use alternative means of accessing their files (primarily search and recency-based lists). The interventional element of the study helped identify the importance of folders to participants during their work by identifying circumstances and reasons that required the use of folders.

4. STUDY

The empirical research presented in this paper is drawn from a qualitative study of 12 information workers, based on the naturalistic observation of their folder use during everyday work. The study was conducted between November 2013 and March 2014 and all participants worked or studied at the main campus of Lancaster University, UK. The study focused on the use of folders within participant's personal collections of files, held in the local file systems of their primary work computers. To generate observational data a specialized piece of software was developed and installed on study participants' computers for the duration of the study. Due to resource constraints, the study software was created only for Apple OS X, so the study focused exclusively on users of this operating system (versions 10.6–10.9). Focusing on this operating system had the benefit that there were several well-integrated mechanisms built into the operating system for accessing files without using folder views, particularly *Spotlight* (desktop search), *Smart Folders* (saved Spotlight search queries) and *All My Files* (a list of recently accessed or created files). The study software integrated with *Finder*, the file browsing tool built into Apple OS X, similar to *Windows Explorer* within Microsoft Windows.

4.1. Study process

The study was conducted over three phases (see Fig. 1). At the start participants were fully briefed, an initial interview

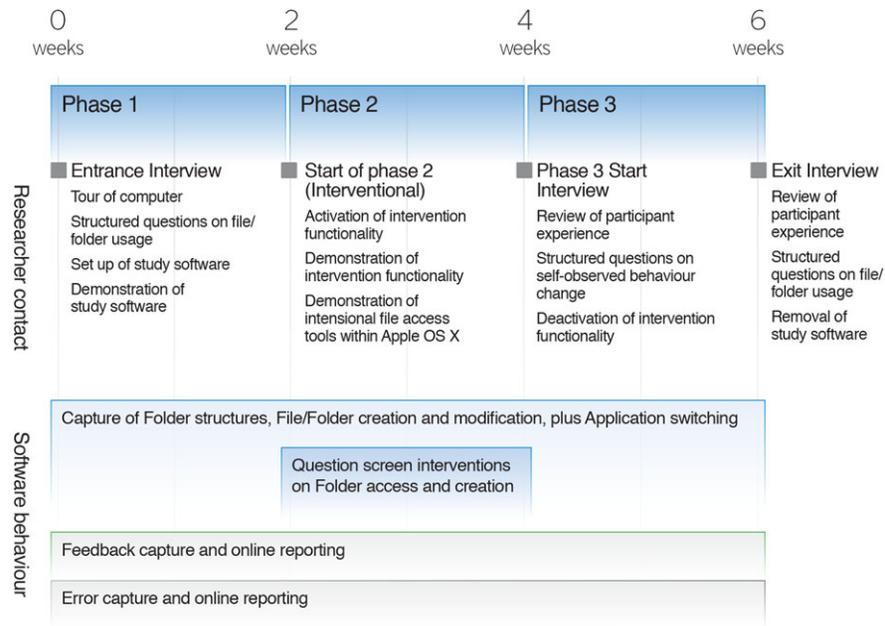


Figure 1. The process of the study.

was conducted, then participants were invited to nominate a subset of their personal files for automated observation during the study. The nomination stage was a key part of the research design as it gave participants explicit control over the observational (and later interventional) elements of the study. Participants were free to nominate any portion of their personal folder structure, but were guided towards their primary working and archival folders, identified in the initial interview.

During all three study phases the study software automatically recorded information about interactions with each participant's nominated files and folders (described in Section 4.2). In the first and third study phases, the study software was purely observational, while during the second phase of the study it also introduced interventions whenever participants viewed or interacted with their nominated folders through Finder (described in Section 4.3). The interventional phase was positioned as an experimental challenge for participants; they were asked to avoid using folders during this phase, but were given complete and explicit freedom to use them as much as they wished, provided they responded to automated prompts (called question screens, see Fig. 3) asking for explanations of their folder use. Participants were also able to disable the automated interventions for fixed periods of time using a 'study holiday' feature built into the study software. When activated the study holiday would suspend all interventions for a chosen period (10 minutes to 3 hours) at the end of which the participant would be presented with a question screen asking them to explain their need for the holiday. The interventions served as a direct prompt that encouraged participants to reflect on why they used folders, and produced

focused, memorable instances of folder use that could be discussed in subsequent interviews.

Each phase of the study began and ended with a visit from the researcher to check the study software was working as expected, and except during the transition from Phase 1 to 2, to conduct a semi-structured interview about participant's file and folder interactions and practices. Interviews were conducted in the normal work environment of participants and ranged from 20 to 85 minutes in length. In the final exit interview a visualization of the data generated across all phases of the study (see Fig. 2) was interactively explored by the researcher and participant. The visualization allowed participant and researcher to reconstruct recorded moments of folder use and discuss the behaviours and dispositions that led to them, facilitating a reflective discussion about the participant's folder use over the length of the study. Data from question screens and other interventions fed into this process, allowing the researcher to direct part of the interview towards moments when participants had needed to use folders during Phase 2 of the study.

4.2. Automated generation of observational data

The study software recorded observational data based on events occurring within each participant's desktop computing environment, storing this data in a local database on the participant's computer. The data generated constituted time-stamped events in the following categories:

- (i) A snapshot of the initial structure of files and folders nominated by each participant for inclusion, followed

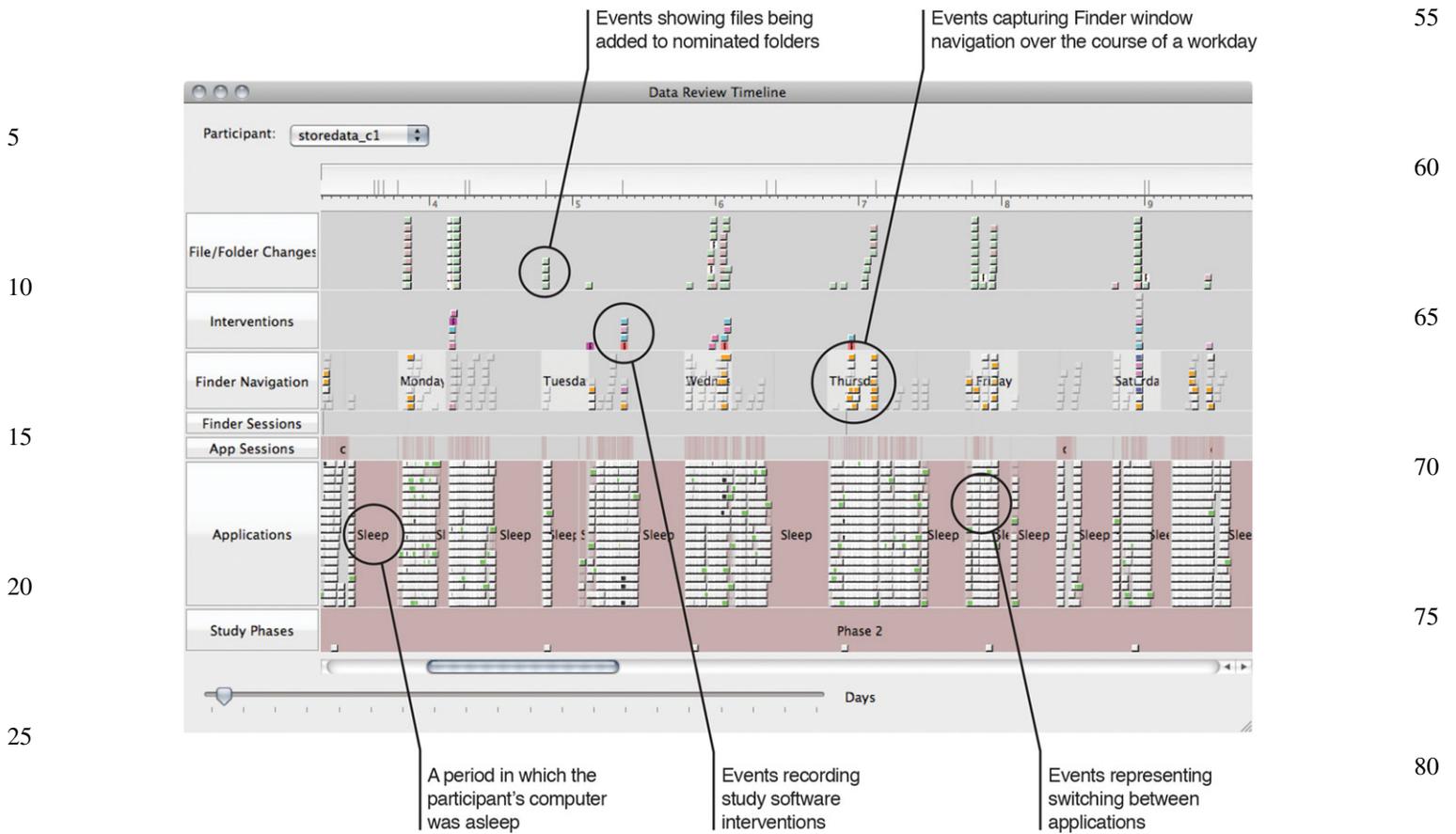


Figure 2. A typical interactive visualization of participant data generated by the study software.

by any changes made to this structure over time. File creation/modification metadata were captured for both files and folders. The names and contents of files were not captured to preserve the privacy of participants and lower barriers to participation in the study.

- (ii) Instances of Finder use within Apple OS X. The titles of Finder windows were captured along with the folder they were viewing (provided the folder being viewed was among those nominated by the participant).
- (iii) Instances of application use, captured through application 'activation' events. The activation of an application occurs when a window of that application receives user input focus within Apple OS X, such as when clicking an application icon in the dock or opening a file within Finder.
- (iv) Environmental events pertaining to the operating system as a whole; power on/off events; sleep and wake events.

This event data was then processed and visualized as a zoomable timeline available to the researcher and a specialized log window that participants could view at any time through the study software. Figure 2 shows part of a visualization relating to one participant; each square element on the timeline could be interrogated to show detailed event information.

4.3. Automated interventions

During Phase 2 of the study interventions were introduced, automatically triggered when participants interacted with their nominated folders. Question screens appeared whenever participants viewed one of their nominated folders using Finder, or when they created a new folder within their nominated folder hierarchies. Figure 3 shows a question screen triggered by viewing a nominated folder using Finder. The screen takes over the participant's computer and requires a response; the participant must elect to either (i) close the Finder window (and access their files another way), (ii) provide an explanation for their need to use folders, or (iii) report the question screen's appearance as a bug. If participants elected to write an explanation for their folder use (of any length) then they were then permitted to use Finder without interventions until they switched to another application. After switching to another application from Finder, all Finder windows displaying the contents of a nominated folder were automatically closed.

4.4. Study software implementation

The study software was created as Apple OS X universal application using Xcode 3. The observational data was captured

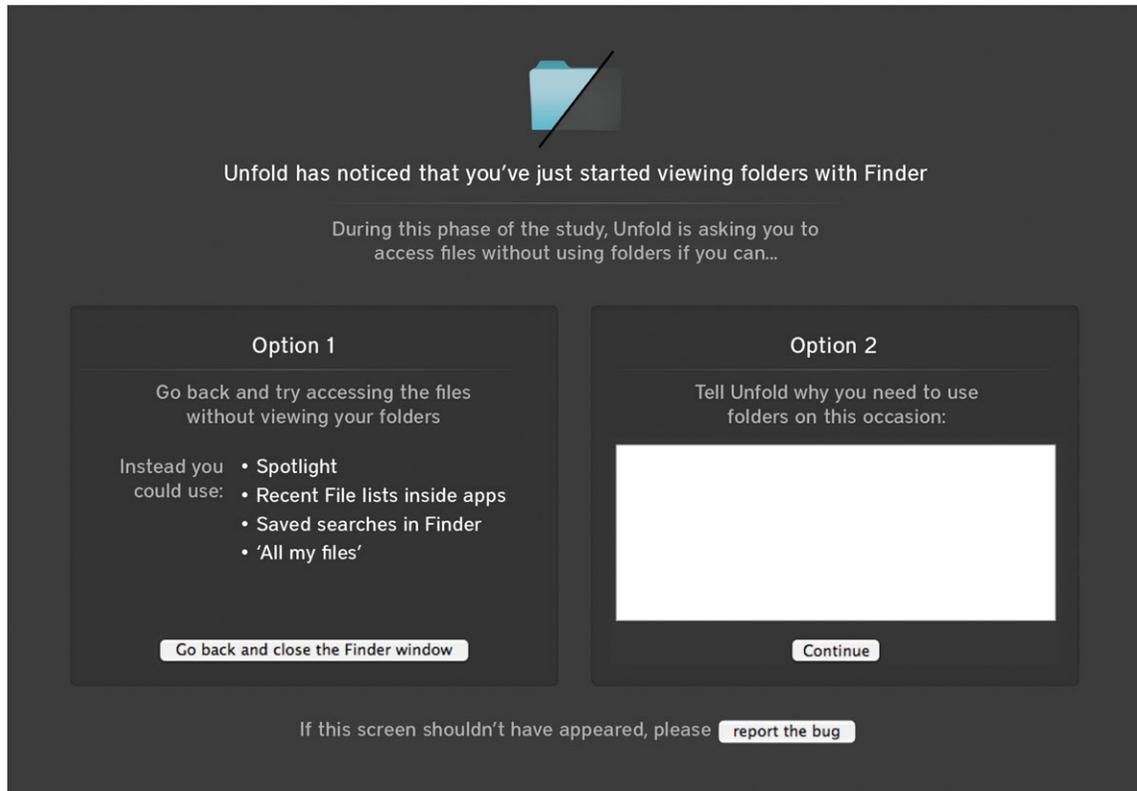


Figure 3. A typical a question screen presented to participants during Phase 2 of the study.

through a mixture of environmental events (NSWorkspace notifications), events from Finder windows (via the Carbon Accessibility API) and file system events indicating changes to files and directories. Data were stored locally on participants' computers using a SQLite database. The software was presented to study participants as a 'status bar' application that appeared throughout the study alongside participant's other background and utility applications. The status bar menu gave access to help documentation, bug reporting, a record of nominated folders and a complete log of all data generated. The software was extensively beta-tested prior to the study.

4.5. Participants

Fifteen participants were recruited for the study, one of whom withdrew due to illness, a second unexpectedly made almost no use of their computer during the study period and a third encountered problems with the study software leading to an incomplete record of their activity. As such we report on findings from 12 participants in this article, shown in Table 1. All participants from whom findings are drawn nominated their main working folders for inclusion in the study, typically their system-defined *Desktop* and *Documents* folders. Table 1 shows the number of files and folders nominated for inclusion by each participant at the start of the study.

All participants were experienced information workers, and all but three of the PhD and Masters students were returning to education after working in commercial contexts. All participants reported regular use of a single primary computer during their work running Apple OS X (10.6–10.9) which became the focus of the study. All were experienced computer users, though some had switched from Microsoft Windows to Apple OS X in the last year and were still discovering features of the operating system (P1, P2, H3, C1). All participants relied on locally stored files, complemented with some peripheral use of cloud services such as Google Drive, Dropbox and Apple iCloud, and all were experienced in using folders to structure and organize files. All participants favoured folder navigation to access files except for participant C4 who favoured Spotlight search.

4.6. Data analysis

Over a 500 days of observational data and 24 hours of interview data were analysed to understand the use participants made of folders and alternative means of accessing files during the study. The analysis distinguished between two types of view based on how the contents of the view were defined; *folder views* and *intensional views*. The content of a *folder view* is defined *extensionally*, that is, by reference to a set of particular items that have been placed inside a folder. In

Table 1. Study participants.

Participant	Age	Sex	Role	Folders included	Files included
P1	25–34	F	Research fellow	183	3401
P2	25–34	F	Lecturer	549	1393
P3	35–44	M	Lecturer	1243	12 068
H1	25–34	F	Masters student	1076	4720
H2	25–34	F	Masters student	186	1389
H3	25–34	F	Masters student	646	14 263
H4	25–34	M	Masters student	768	6011
H5	35–44	M	Masters student	6150	44 698
H6	Under 24	F	Masters student	176	1112
C1	25–34	M	PhD student	362	1047
C2	25–34	F	PhD student	3465	20 307
C4	25–34	M	PhD student	9776	56 994

contrast the contents of an *intensional view* are defined through predication of object properties, such as files matching a particular file type or a search query. Folder views change when items are explicitly added or removed from folders, while intensional views are dynamic and can change without direct interaction, such as a recent file list which is updated automatically.

The transcribed verbal interview data was coded using a data-driven approach to identify the uses of folder views, uses of folder structure and uses of intensional views. Statements about folder use made by participants in interviews were compared against the observational record, allowing triangulation between observational and interview data. Once coded, the interview data was iteratively displayed and reduced to produce categories of folder use, presented in Sections 5.3, 5.4 and 5.5.

The quantitative observational data generated by the study software was analysed to produce metrics to describe participants' behaviour. Metrics such as quantity of files and folders, depths of folder structure and the number of steps in navigation were produced in line with existing PIM research. A new metric was developed to capture view use within Finder, the *Finder Session*. A Finder Session begins when Finder is switched to (activated) and ends when the user switches to another application, or the computer is shut down or enters sleep. Finder Sessions are classified to indicate whether they comprise folder views, intensional views, a mixture of both or neither (null) when no views in the session contained files (such as when viewing file metadata via the 'Get Info' option within Finder).

4.7. Limitations

The combination of naturalistic observation and in-depth interviews used in this study provides detailed engagement with participants, supported by a highly accurate observational record. The duration and depth of engagement gives confidence as to the internal validity of the findings, but the number of participants remains relatively small, and though

consistent with other similar studies in PIM literature (e.g. Alvarado *et al.*, 2003; Bellotti *et al.*, 2003; González and Mark, 2004; Jones *et al.*, 2005) it cannot give external validity to the findings. The sample was sufficiently large to capture and describe in detail a great diversity of folder use, but this cannot be taken to represent a complete spectrum of possible uses and behaviours. The study focuses on document-based information work, but it does so in the context of only a single operating system and a single period of technological development. The approach to observational data generation was highly accurate in terms of capturing events within participants' desktop computing environments, but this record did not extend outside of their computer to capture the use of other materials used in their activity.

The study is based on in-context observation of naturalistic behaviour and the data recorded reflects the differing workloads and activities of participants. As such the quantitative results of the study can provide only a picture of the variety of folder use behaviours, rather than a complete summary of folder use behaviours in general. This is offset by the rich qualitative data generated from interviews, supported by the observational record.

5. FINDINGS

The study revealed a great variety in the use of folder views and folder structures, described and categorized in Sections 5.4 and 5.5. Preceding this we present an overview of participant behaviour based on the quantitative data generated, then compare observations of behaviour across the first two phases of the study.

5.1. Participant behaviour in Phase 1

Table 2 presents an overview of the participant's different levels and kinds of Finder use during Phase 1 of the study. These data have been selected because they provide a

Table 2. Quantitative overview of Finder use during Phase 1.

Participant	Distinct folders viewed	Total finder sessions	Extensional sessions	Intensional sessions	Mixed sessions	Null sessions	Navigation steps (Min,Max) mean
P1	37	81	80 (99%)	0	1 (1%)	0	(0,17) 1.02
P2	28	39	32 (82%)	2 (5%)	5 (13%)	0	(0,16) 1.95
P3	24	69	58 (84%)	3 (4%)	7 (10%)	1 (1%)	(0,7) 0.75
H1	10	70	42 (60%)	14 (20%)	9 (13%)	5 (7%)	(0,4) 0.4
H2	54	55	44 (80%)	0	0	11 (20%)	(0,21) 3.36
H3	29	76	37 (49%)	16 (21%)	12 (16%)	11 (14%)	(0,7) 0.67
H4	153	300	238 (79%)	0	0	62 (21%)	(0,21) 1.77
H5	13	32	27 (84%)	1 (3%)	0	4 (13%)	(0,4) 0.41
H6	28	93	66 (71%)	6 (6%)	13 (14%)	8 (9%)	(0,19) 2.01
C1	63	196	61 (31%)	114 (58%)	11 (6%)	10 (5%)	(0,20) 0.92
C2	26	102	54 (53%)	15 (15%)	11 (11%)	22 (22%)	(0,12) 0.61
C4	3	12	4 (33%)	0	0	8 (67%)	(0,1) 0.08

representative summary of Finder use and its relationship to folder use. Participant H4 and C4 represent two extremes of observed Finder use; H4 viewing many folders very frequently, C4 viewing few folders and doing so rarely. Levels of Finder use did not correspond with the number of distinct folders viewed; participants who used Finder more often did not necessarily view more folders. For example, participants P3 and P2 viewed a similar number of folders, yet participant P3 used Finder almost twice as often.

The data in Table 2 also shows that the participants generally had a strong preference for folder views with the exception of participant C4 who reported minimal use of Finder in general in favour of Spotlight search. Participant C1 did make considerable use of intensional views; however, the observational and interview data show that this was to solve a specific software development problem relying on the mechanics of the view, not for general file access.

5.2. Participant response to Phase 2

All participants except P2, C2 and C4 used folder views less during Phase 2 compared with Phase 1, with participants P3, H1, H3 and H6 showing a reduction of >80%. Overall, the use participants made of folder views within Finder during Phase 2 in comparison with Phase 1 fell into four categories:

- (i) No change in Finder use due to pre-existing preference for intensional views (participant C4).
- (ii) Substantial decrease in use of folder views with adoption of intensional views and other intensional methods of file access (participants P3, H1 and H3).
- (iii) Adoption of intensional views and intensional file access methods, but with continuing use of folder views (participants P1, P2, H5, H6, C2).
- (iv) Limited adoption of intensional views with continued use of folder views (participants H2, H4 and C1) and involving direct circumvention of interventions (participants H4 and C1).

These responses can be explained in part by the relative workloads of participants. For example, participants H1 and P3 both drastically decreased their use of folder views during Phase 2 of the study, but also reported a lower workload than during Phase 1, giving them more time to experiment with intensional file access methods and the freedom to avoid tasks requiring folders. Workload was clearly not the only factor at play however, as participants with high and low relative workloads are found across the above categories of response.

Participants in the third and fourth categories of response were dependent on using folder views during Phase 2. Participants in the third category put effort into avoiding folder use but found they needed to use folders, responding to question screens and taking study holidays in order to do so. All participants in this category used the study holiday feature and accounted for 14 of the 17 study holidays recorded. These participants experienced frustration with the constraints of Phase 2, but were able to accommodate them for a period of 2 weeks with effort.

Participants in the fourth category of response struggled far more with the constraints of Phase 2 and did very little to adopt intensional means of accessing files. These participants needed regular access to their folders during their work and were unable to stop using them. For participant H2 this meant regularly encountering question screens and the use of a long study holiday close to a deadline. Participants H4 and C1 responded more strongly, both circumventing the constraints of the study software during the first day of Phase 2, H4 by moving working files out of his nominated folders and C1 by using an alternative file browser to Finder. These participants reported being unable to accept the constraints of Phase 2 and seeing no other option but to circumvent the study software. Participant H4 was by far the most intensive user of folders amongst all participants and experienced more interventions during Phase 2 than any other participant, despite his circumvention.

The reasons given by participants for using folders during Phase 2 of the study use were explored in interviews

alongside the observational record, feeding into the categories of folder use presented in Sections 5.3, 5.4 and 5.5.

5.2.1. Adoption of intensional views and file access methods

The study design and participant brief for Phase 2 was intended to encourage participants to experiment with and adopt intensional means of accessing files. All participants experimented to some extent with intensional file access methods, although adoption varied greatly. Participants H2 and H4 reported very limited experimentation and adoption while participants P1, P2, P3, H1, H3, H5 and H6 all reported substantial adoption of intensional file access methods. Participant C1 and C4 reported existing use of intensional file access methods and almost no additional adoption during Phase 2. Participant C2 reported some increased use of intensional file access methods but this was restricted to the beginning of Phase 2 and not continued throughout.

The main barrier to adopting search (via Spotlight) as a means of accessing files was the requirements it placed on precise memory of file names and the difficulty of differentiating similarly named items in the search results. Participants were generally very conscious of the risks of incorrectly identifying versions of files, something which the intensional views did little to address. All participants who adopted search without having used it frequently before the study commented on the high level of effort involved in using it when compared with folder views. Participants P1, P2, P3, H3, H5, H6 and C2 reported experiencing or anticipating frustrations in formulating search queries or sifting through search results, participant P3 referring to ‘*labour intensive*’ process of searching and H3 process of trying to ‘*figure out how to find what I wanted to find*’ through Spotlight. Recency-based intensional views, such as *All My Files* and the recent file lists within applications saw some use, with particularly strong adoption by participants P2, H1 and H3. These views were found to be useful, but did not prove reliable; participant H3 reported items not being present when expected, and conversely participant P2 found that irrelevant files would be included due to momentary interactions with them.

In summary, Phase 2 saw most participants experiment with and adopt intensional file access methods and resources

to some extent. Of the participants who did adopt these methods most found some utility in them and were able to use them to reduce their need for folder views. The use of intensional file access methods did not eliminate the need for any of the study participants to use folders at some point during Phase 2, including participant C1 who had an existing preference for using search.

5.3. Uses of folders

Based on the data-driven qualitative analysis of observational and interview data, we present eight distinct categories of folder use. Not all uses were observed for all participants, but each use driving the categories was established with triangulation between the reports from interviews and the observational record. Our analysis separates the uses made of folder views from those made of the underlying, persistent folder structures. This separation reflects the distinct roles folder views and folder structure played in the work of participants and the distinct affordances folder views and folder structure offer. Folder views provide transient perspectives on folders, present only as long as a Finder window remains open, with no persistent effect on the files or folder viewed. In contrast, folder structure provides a means to persistently group and label files that will remain the same across time, but with only partial control over when and how files will be viewed.

The great majority of observed interactions between participants and their folders took place through a folder view: a Finder window representing a portion of their file/folder hierarchy allowing both navigation and manipulation of folders and their contents. Other interactions between participants and their files and folders (such as saving files within applications, or automated creation of files and folders by software) were also observed. Four categories of use for folder views are summarized in Table 3 and described in detail in Section 5.4. Folder structure was observed to have related, but distinct uses, summarized in Table 4 and described in detail in Section 5.5. There are interrelations between the categories of folder view and folder structure use. For example, the categories described in Sections 5.4.4 and 5.5.4 both relate to executing tasks, but the role folder views and folder structure

Table 3. Categories of folder view use.

Category	Description	Identified with respect to participants
Accessing specific items (Section 5.4.1)	Folder views used to access particular files and folders, often to open files in applications	All
Exploring groupings and items (Section 5.4.2)	Folder views used to browse files and folders with some uncertainty of what is being sought or what items will be present	P1, P2, H1, H2, H3, H6, C1, C2, C4
Overviewing materials and groupings (Section 5.4.3)	Folder views used to look at the relations between files and folders, and to bring to mind the activity the files and folder are related to	P1, H2, H3, H5, H6, C2
Supporting task execution (Section 5.4.4)	Folder views used to directly support the practical execution of a task	P1, H2, H4, C2, C4

Table 4. Categories of folder structure use.

Category	Description	Identified with respect to participants
Persistent availability of items (Section 5.5.1)	Folder structure used to ensure files and folders would remain available for access across time	All
Grouping and separating items (Section 5.5.2)	Folder structure used to collate items together in discrete groupings, comprising project-specific groupings, task-specific groupings, project and task-agnostic groupings and groupings for mapping activity components	All
Controlling representation and interaction (Section 5.5.3)	Folder structure used to control the representation of folders and files in folder views	P1, P2, P3, H2, H3, H4, H5, C1, C4
Facilitating task execution (Section 5.5.4)	Folder structure used as an instrumental component in the execution of a task	P1, H4, C1, C4

played in these tasks was distinct: folder views were used to control which information items were present to the user visually and interactionally during tasks, while folder structure was used to preserve, branch and iterate states of work within applications by duplicating files and folders.

The categories of folder view use are not exclusive to particular Finder Sessions; the use made of a given folder view could encompass multiple categories over its life. Folder views were frequently used to access specific files, but these same views could also be used to overview groupings or explore items. Similarly the categories of folder structure use are not exclusive to one another with respect to a particular folder; a folder might be used both to collate items relating to a particular task, and to keep these available for access at a later date. In these findings the terms *task* and *project* are used to refer to different scales and durations of activity. *Task* refers to short-term, discrete activities often completed within a week (Bellotti *et al.*, 2004), while *project* refers to longer term activities comprising multiple tasks (Jones *et al.*, 2006; Kaptelinin, 2003).

5.4. Categories of folder view use

5.4.1. Accessing specific items

Unsurprisingly, folder views were used most commonly by study participants to access known information items (files or folders). This often involved some uncertainty about the precise location or name of the file, something that the process of folder navigation resolved. As participant H6 explained, '*I vaguely remember where we usually would store the file, but I can't remember the file name. I can't remember where we stored it, so I was just navigating myself to find it.*' During Phase 2 of the study, needing to access specific files was a common reason given for needing to use a folder view. For some participants it was very difficult if not impossible to access specific files without a folder view, as P3 put it, '*...having to go to that folder would be because I simply didn't know how to search for those files in order to identify them by another other route*'. Folder views allowed participants to recognize files in context with other files and folders, giving

them confidence they were accessing the intended item. In addition to providing access to specific files in order to open them in applications, participants also used folder views to move items, access item metadata or explicitly determine which application to open files with.

5.4.2. Exploring groupings and items

Distinct from accessing specific items, folder views were also used to explore folders and their contents without a particular item in mind. In examples of this category of folder use participants often had an information need or intent of some kind, but they did not know whether they would find one or more relevant items prior to using a folder view. In two similar examples, participants P1 and P2 both viewed folders to look for teaching material suitable for a new teaching session, doing so by reviewing the folders and files they had relating to past teaching activities. As P1 explained, '*I'm not aiming to retrieve a file. I'm just aiming to get a list of all the files I have.*' This participant reported not knowing if she would find anything suitable, but looking at the list of files and folders helped her think through possible options. Study participants also used folder views to explore items without well-defined information needs, for example participants H3 and H2 both explored their folders for casually looking for inspiration from old images and projects folders. In some cases exploration was needed to discover what files and folders were present (such as when exploring items created by another person) or to check for mislaid items, such as C1, H1 and H6 regular checks of the Downloads or Trash folders for valuable items. In these cases, the folder view provided a finite and complete view of items that once checked, gave certainty to participants.

5.4.3. Overviewing materials and groupings

A third category of folder view use is that of overviewing files and folders to identify patterns within them. In this category of use particular items were less important; it was the broad relationships and patterns between items and groups were of value. Participant H2 described using a view of her activity-related folders like a pinboard or a map, explaining

her need by stating, *'It will be a mixture of needing to see it visually; needing to orientate myself, and then also needing comfort in knowing that there is organised thinking in what I've done'*. C2 described using a view in this way to understand the structure of a personal project and ensure she remembered each part of it, while participant H3 highlighted the value using a view in this way had, explaining, *'...you can see everything all at once and you can see things you might have forgotten or things you might need or gaps. So you can see patterns in your list.'*

Folder views also allowed some participants to compare the structure of past activities, such as H6 who used folder views to review the groupings and items related to multiple past projects. In this case a view of the folder structure relating to a project met her information need and allowed her to identify activities and outcomes from one project that were relevant to another project. During Phase 2 of the study, H6 activated the study holiday feature so she could open multiple folder views and visually compare the structure of activity-related folders, referring to this as akin to having multiple maps open to act as summaries of past projects. Her purpose was to identify the abstract categories of outcome from a past project and make decisions about what to do in a current project, a process she found natural to support using folder views.

In some cases the informational content of a folder view was itself sufficient to meet participant's needs. P1 described using a folder view of her past conference presentations to identify how frequently she had spoken at a particular conference. When describing the process of planning some writing, participant H5 used the term 'weighting' to describe visually scanning a folder view of collated items to determine the relative numbers of items pertaining to particular topics, *'I want visual access across all of them, and what I'm doing is weighting them by their nature.'* His items needed to remain in a single view, *'...it still was manageable group of papers that I could see in one stretch on the screen, but it was enough for me to want to be able to get a sense of their weight; how many did I have in this area? How many did I have in that area?'*. H5 explained that he would not want to encode the groupings he could perceive using folders at that point, instead he wanted to have the freedom to identify and imagine possible groupings and relationships between items.

5.4.4. Supporting task execution

A final category of folder view use relates to the role folder views played in the practical execution of tasks, as opposed to simply providing access to items required for tasks. Beyond simply accessing information or manipulating information items, this category is about the pragmatics of doing work using files and folders. Tasks requiring the creation or manipulation of folders were straightforward examples of this, such as creating a package of structured files to share with another person. For some participants, folder views also

supported activity that spanned across multiple files, applications and windows; here the folder view provided a stable context for tracking and executing each stage of the task. In one example, participant C4 worked through a list of files, viewing each file and renaming it according to his understanding of its content. C4 explained, *'So we had to manually rename each individual file, so there was a lot of renaming going on. So, that would have been switching between applications to see what the image was that we were renaming and then renaming it accordingly.'* In this example, the folder view provided the means to open and rename files, but also represented the state of the task during execution.

For participants who collated items for a task using folders, folder views could play defining roles within task execution. Participant P1, for example, used a folder view as a means to track and facilitate the task of creating a lecture. Working with a folder containing pre-selected material for the lecture, this participant took a study holiday expressly so she could retain visual access to the folder view during the task. P1 explained, *'...I was working on this lecture and I wanted to keep the folder open, so I could go back and forth between the articles that are in that folder, without having to keep all of the articles open...'* The task favoured interacting with the files through a folder view, but P1 also valued seeing the list of items during the task, *'...it helps remind me of the structure. I can look at that list and go: "I've discussed this one and this one, I'm working on these two, and then those will come afterwards."'* The participant recounted a previous occasion when she had failed to include a relevant item in a folder used in the process of creating lecture slides. As a result she had omitted content from lecture without realizing it during the subsequent task of practically creating the slides. Similarly participant H4 would collate together image files to be combined together in a visual collage in a single folder, then rely on a view of this folder during the task of making the collage. In such cases, the folder view acted as a means to access the files, and as a definitive list of items to be combined.

5.5. Categories of folder structure use

5.5.1. Persistent availability of items

All study participants relied upon folder structure to retain the items placed within them for future use. For participants, folders provided the only means of ensuring items would remain available across time and they often found it hard to imagine another means of achieving this. As participant P2 expressed it, *'I created a folder but I did that so I wouldn't lose documents. There wasn't an alternative. I can't use Finder or search to create a new folder.'* In one example participant P2 moved a file from an email message to a folder (in her local file system) to ensure it would be identifiable in the future without the risks she associated with searching her

email. Folders also provided participants with a way to archive materials related to past or completed activity without any certainty they would need access to it again. Participant H1 described her archive as, *'where my documents go to die' and did not expect to resume using them once they had been moved, part of what H3 described as the 'life-cycle' of her files.*

5.5.2. Grouping and separating items

Alongside persistence, a core affordance of folders is to allow items to be grouped and nested, both to associate and disassociate items with one another. All participants used folders to create groupings, and consistent with PIM literature, these were highly diverse in their structure and naming (Barreau, 1995; Boardman and Sasse, 2004; Kwasnik, 1989). Four kinds of grouping were observed:

Project-specific groupings. Collating items around projects was a strategy common to all participants to ensure they could access files in the future and for most was considered simply be *the way* to ensure they could access project-related items across time. Such groupings ensured related items could be accessed again and would be accessed together. The groupings also ensured that groups of items were clearly differentiated by purpose when they were similar in structure or content. Creating project-specific folders was highlighted by participants as a means to reduce the complexity of their folder views during navigation, but also a means to conceptually separate distinct activities from one another. H6 reflected, *'It's kind of having the thing all in one place... makes me feel more secured, like I know that the thing is there, like I can get to it in a second, having the folder open, I mean.'* Similarly P1 expressed, *'By not thinking of my work as one big megalith, but by thinking of discrete units I can work on separate units in particular ways.'*

Task-specific groupings. Some participants also created groupings of items strongly related to particular tasks. A clear example of this was described by participant H4 who used folders to group multiple files he then combined into a single document he referred to as a 'board', *'I might have a lot of images that I've pulled off the internet or I've generated separately, and then I'd create a folder for all those for that particular board, drop everything into there.'* Participant H4 used folders heavily in this way, and was acutely aware of the cost and benefits of doing so, explaining, *'I've got to create a new folder and then put stuff into it, but, say, the following day when I continue to work on the board I know everything is in that one place.'* The contents of the 'board' folder did not define completely the later task, but they represented the completion of a part of it.

Project and task-agnostic groupings. In contrast to groupings for particular projects and tasks, some participants used folders to produce groupings of items related in other ways. Commonly these were top-level divisions in their folder structures, such as folders for particular years or types of activity

(e.g. freelance work, teaching or study). Some participants also had groupings related by content type, such as music or photographs. Groupings independent of specific projects or tasks allowed some participants to draw on a common set of materials across disparate activities, such as P1 who created folder groupings to separate articles she was authoring from those authored by others, then used these across a range of her teaching and research activities.

Groupings for mapping activity components. A final type of grouping and separation included in this analysis exists in interaction with the three described so far: using folders to define and shape the component elements of activity. Participant H2 explained this was part of her strategy for managing projects, *'So it's allowing me to project manage or.. what do I mean.. manage the content of the project is probably the better way to put it, rather than the management.. it's not the project management times; milestones; it's not a Gantt chart; it's just allowing me to place into folders where everything is.'* This participant, previously a professional project manager, likened the practice to creating a collection of project-related materials for handover to another person, though she was well aware that she was the only person working on the project. Participants H2 and H4 both created many folders when beginning a project, many of which were observed to remain empty, but which described anticipated elements the project would come to involve. H4 described these folders as like a *'taxonomy of things'*, representing his understanding of the activity and so, *'de-cluttering my own mind'*. When explored further, H4 explained, *'I guess this stores ideas from me - Things about me and what I've created - so that's an extension of me, of my identity.. Yeah.. it's a way of organising what I'm thinking, or ideas I've collated.'* Over the course of study well over half of the participants created folders related to activity components which remained unpopulated.

5.5.3. Controlling representation and interaction

Parallel with grouping and separating items based on activity or content, this category of folder structure use relates to controlling the experience of viewing and interacting with folders and items within them; controlling the 'look and feel' of folders. Most participants in the study cared that at least some of their folders were free of visual noise. Sometimes this was to make items easier to spot, such as participant C1 who wanted to ensure his system-defined *Downloads* folder was kept free of clutter so he could readily identify valuable items within it. Some participants had tacit rules for how many items a folder could contain before it became less useful, and participant H5 described creating new folders and placing items within them specifically to reduce visual clutter.

For several participants it was important for their folders to be visually consistent as a reflection of their own internal understanding of their work. Participant H2 had highly consistent naming conventions for folders, with important folders

named in capital letters. Participant H4 would number folders to ensure they appeared in a consistent order across time. Folders provided a means for participants to incorporate diverse and sometimes meaningless file names (often downloaded from the web) into a folder structure they could recognize and rely on.

5.5.4. Facilitating task execution

The final category of folder use presented here relates to the practical execution of tasks. These folder structures not only supported task execution, but were a means to take action critical to task execution. Clear cases of this use occurred when manipulating folder structures was either the purpose of the task or a key means to facilitating it, such as when controlling software through folder manipulation (e.g. cloud-based file synchronization software, or software that monitored folders and ingested files placed with them). Folders were also a means for participants to branch and iterate work done in applications. Participants C1, C4 and H4 would duplicate entire hierarchies of files and folders to preserve a particular state of their work. For participants C1 and C4 this provided a means to control a software IDE application, allowing them to test out problem solving approaches without risking losing a working state of the software. Participant H4 duplicated files and folders to iterate work, explaining, ‘...*this is kind of like my own creative behaviour... I might work on something to a certain point and then think: “I’ll save it as it is and keep that as a separate file”, and then I’ll copy it and start working on it again so that I’ve preserved that original thing.*’ This participant retained abandoned branches of his work consistently using folders as a record of the exploration he had undertaken during the activity.

6. DISCUSSION

The research presented in this paper seeks to describe in detail how folders are used in individual information work, and identify the value information workers derive from using them. The way participants used folders fits generally well with observations from previous PIM studies. Participants created folders and placed items in them so they could reliably access them again in the future, matching the keeping behaviours defined in Jones’ and Whittaker’s lifecycle frameworks (Jones, 2007; Whittaker, 2011). Barreau and Nardi’s (1995) categories of personal information (ephemeral, working and archived) fit well with the observations; participants interacted with a relatively small portion of their folders and files during work and maintained archival collections of older files. Groupings made with folders were highly varied and most were strongly activity-focused, consistent with studies of individual information management practices, in particular the use-dependent (Barreau, 1995; Bergman *et al.*, 2008a,b; Kwasnik, 1989) and the context-dependent nature of

groupings (Boardman and Sasse, 2004). Participants generally found accessing their working information easy, consistent with existing studies (e.g. Alvarado *et al.*, 2003; Barreau and Nardi, 1995; Bergman *et al.*, 2008a,b; Boardman and Sasse, 2004), and most showed a strong preference for folder navigation (e.g. Bergman *et al.*, 2008a,b; Teevan *et al.*, 2004).

In the context of PIM activity it is tempting to think of folders as a costly, manual means of organizing and retrieving information that precedes getting real work done. Whittaker (2011, pp. 42) describes the burden of filing new information items and the effort of ‘having to create and maintain appropriate structures that anticipate retrieval’. In articulating an ideal of PIM, Jones suggests that less time and energy spent managing our information through processes such as filing means more time to engage in creative work (Jones, 2007). By investigating how folders are used during work, this study provides examples of the valuable functions that folders and manual organization of information can fulfil beyond keeping and accessing files. Consistent with previous work, the study showed that folders could remind participants of past organizational decisions and of necessary action in the present (Barreau and Nardi, 1995). Study participants created folders that reflected their understanding and intentions for activity, as described by Jones *et al.* (2005, 2006) and these played a part in planning and managing activity. The study also shows the value that representations of folder structures can bring to information work, and the potential for folders to shape and integrate with tasks and the processes of information work more generally.

Rather than a managerial or curatorial stance towards their working information, most study participants had an active, tightly bound relationship with their folders that exploited tacit knowledge and sought to impose structure on both their information and the way they worked. Folder views were used by all participants to access files, but to some participants they also offered valuable perspectives on their past and current activities, allowing patterns spanning files and folders to be identified and exploited. For some participants, folders structured and participated directly in the execution of tasks, rather than existing as a separate precursor to activity. All participants found it straightforward to create and access their working folders, and the majority cared greatly that their folder structures were consistently organized and represented. From this perspective, imposing order on items through folders was not an unwelcome necessity of technology, but instead an important part of some participant’s individual working methods and creative processes.

In the following sections we consider why folders provide functions that extend beyond classifying and accessing information, and discuss the potential of manual organization of information in environments for individual information work. In Section 6.1, we explore the role of that structure plays in work environments, then in Section 6.2 we discuss how folders facilitate control over this structure. Following this in

Section 6.3, we discuss whether folders and could be replaced or eliminated before drawing implications for design in Section 6.4.

6.1. Structure in the environment of work

It is well established in cognitive science that how problems are represented and understood strongly effects the manner (and ease) in which they can be solved (Zhang and Norman, 1994; Zhang, 1997). Moving to more complex problem solving, Hutchins' (1995) work describes how systems that bridge between individuals and their external environments can radically transform tasks. Kirsh (1995, 2001) considers the role of task structure in the context of individual work, describing how experts use the resources in their environments to structure tasks and work more effectively. The structuring that Hutchins and Kirsh describe shapes the actions needed to perform a task, controlling internal cognitive demands and off-loading cognitive work into the environment.

The findings of this study have some clear connections with the idea of structuring problems and tasks. The way some participants used folder views and folder structures fit with Kirsh's (1995) notion of pre-structuring task environments, controlling the information in the environment, and simplifying perception and decision-making during task execution. For example, participants P1 and H4 both used folders to divide up their work on tasks that involved synthesizing multiple distinct pieces of information (respectively, lecture slides and a document). Firstly they collected items together in folders, then later used these folders to access the items and combine them using an application. During the combination stage folder views provided a finite, pre-structured set of materials that defined the task, reducing the number of choices needed and exposing necessary affordances for the work. P1 recounted accidentally omitting an item from a lecture because she had failed to place it in the relevant folder, not realizing this until she was delivering the lecture. She described 'blaming' the folder for the error; she trusted her earlier decisions, as represented in the folder view and relied on the prescriptive role she gave to the folder view in the task of constructing her lecture slides.

The study findings suggest relationships can form between folders, information items, application windows, affordances for taking action and the activity of information work. These relationships were present in tightly-defined, low-level tasks of combination and composition, and in higher-level processes of sense-making, planning and orienting to activity and work, as also identified by Jones *et al.* (2005). When viewing their own project-specific folders, participants could remind themselves of activity they had undertaken in the past; the value of these folder views lay not only in the information explicitly present within them, but also in bringing to mind participant's internal knowledge of the items they saw. Kidd (1994) identified that in knowledge work, the value of

external information lies in its ability to bring about internal change in the worker, not external change in the environment of work. Kidd describes how spatial arrangements of physical documents can function as a 'primitive language' anchoring internal meaning and ideas to external materials. Similarly in this study, folder views were valuable both for their explicit content and for their internal meaning to participants. The same folder view could support a wide range of internal processes; visually searching for items; comparing project activity; looking for out of place items; mulling over ideas for teaching; planning a new piece of writing or a fluid mixture of all of these.

6.2. Folders as work environment structure

As resources within the desktop computing environment, folders offer seemingly spartan control over information items. Yet with only the means to create and label nested groups of files and folders, participants shaped their desktop computing environments to support key conceptual and practical processes within their work. Participants used the basic affordances of folder structure to impose control over how their information items would appear, sometimes with precise rules for naming and ordering items, allowing them to craft representations of past, present and future activity. Folder structure was also a practical tool for doing work, used to spread the effort of tasks across time and to branch and iterate work done within applications.

Looking at the use of folder structure alone reveals some of the practical value of folders, but by capturing the use of folder views moment to moment during work, a more complete picture emerges. Participants used folder views alongside applications windows as practical resources for taking action (e.g. opening files in applications, moving files between folders) and as resources for supporting their internal thought processes (e.g. considering possible materials to reuse, imagining possible new outputs from project). Folder views offered a layer of representational and interactional control through which relevant parts of folder structure could be surfaced into the current state of the desktop computing environment through windows. The link between tasks and windows within virtual work environments is well established with particular tasks involving particular windows and views of information (e.g. Henderson and Card, 1986; Robertson *et al.*, 2000; Smith *et al.*, 2003; Bardram *et al.*, 2006; Robertson *et al.*, 2004). Folder views and structure gave study participants control over windows and so tasks without the need to explicitly define 'activities' or 'spaces' for distinct aspects of their work. Instead folders allowed participants to fluidly adapt their working environments to incorporate visual representations, interactional affordances and informational content relevant to the momentary needs of their work.

The study reveals how critical the support for individual working methods from folders can be; in Phase 2 some

participants were able to sever most of their links to folders for a period of 2 weeks, while others experienced such paralysis from this constraint that they were unable to think and to work without them. The participants who struggled most during Phase 2 were not only dependent on folders to access existing information; their working methods relied on folders to create new things. We suggest that the high value of folders to some study participants is best explained by the close relationships that had formed between the structure of their information, the structure of their work environments and the internal and external processes of their work.

Each use of folders described in this paper could, of course, be readily supported through specialized tools (e.g. project management software, task-tracking tools, workflows within applications). We argue that it is not specific uses and functions of folders that make them valuable *per se*, but instead the way folders connect directly to the materials (i.e. files and their contents), processes, tools and representational states needed by both conceptual and practical aspects of work. Files and folders are part of the central design metaphor of the desktop computing environment and occupy a special niche that allows them to participate in work; the icon that represents a file within a folder view is simultaneously an abstract representation of the document, a set of vectors to manipulate it and a cue to the ideas and meaning it holds for its owner in their work. When manipulating files and folders, some participants were also manipulating the objects, categories and mechanisms of their work, forming a bridge between activity, files and applications.

6.3. Eliminating the need for folders

The question of whether search can reduce the need for folders and other manual means of managing files is ongoing within PIM (Barreau and Nardi, 1995; Bergman *et al.*, 2008a,b; Cutrell *et al.*, 2006a,b; Freeman and Gelernter, 1996; Jones, 2007; Teevan *et al.*, 2004; Whittaker, 2011). This research focuses on what folders do, not how well search could replace them, however it did reveal some interesting strengths and limitations of intensional file access mechanisms.

When attempting to adopt search, even the most committed study participants regularly found themselves unable to remember the names of files, making them reliant on folders. This was a critical barrier for some participants, but not insurmountable; participant C4 was accustomed to using search and accurately recalling file names. As Lansdale identified, enhancements in search technology will theoretically allow an increasing range of *recalled* information to be used in retrieval of specific items, so we can expect search tools to lower this barrier as they improve (Lansdale, 1988). A second, more fundamental barrier to adopting intensional file access methods was encountered by participants when the information needed was not held within a specific item, but instead the relations

between items or groupings. For example, some participants needed to view items in context within folders in order to identify them, or to peruse overviews of multiple folders to identify and compare patterns they had previously imposed. Information needs of this kind could be met through intensional file access methods, but the query used could only be expressed in terms of existing groupings, not properties of items. Searching for folders is a useful shortcut, but it does not remove the need for folders any more than searching for words within files removes the need for the files themselves as discrete bundles of information.

Based on the findings of the study, our position is that search and other intensional tools can provide useful views and ways of accessing files, but cannot give people the means to impose useful order on their work environments and information in the way that folders can. We argue that for search (or other intensional mechanisms) to eliminate the need to manually manage files, the underlying needs for the management activity would have to be fulfilled in other ways. The study showed how folders can be used for classifying items in a taxonomy for later retrieval, but also that they can be used to impose order reflecting structure of work and intentions of the worker, and it was this latter function that was of critical value to some participants.

6.4. Implications for design

The implications drawn from the research fall broadly into two categories, the first relating to the design of tools for managing and organizing personal information, the second relating more broadly to the design of virtual environments for information work.

6.4.1. Tools for organizing and managing working files

We take the findings of the study to show the value that enabling manual control over the structure of working information and the environment of work can bring. The findings show particular value derived from control over the representational and interactional structure of information items, so it is possible to make design recommendations in line with this. Enhanced information management tools could allow items to be arbitrarily reordered, changed in relative size, annotated with imagery and text, locked or otherwise restricted. Designers might also seek to introduce forms of structure that go beyond the nested sets folders allow today (e.g. chains, loops, webs), or even make structure between files optional, added only when needed. Such enhancements could allow individuals to better match the structure of their information to the processes and structures of their work, but this would come with either the cost of maintaining backwards-compatibility with existing file systems, or losing valuable interoperability between different devices and platforms that folders currently offer.

The findings of this study suggest that enhancing folder views may be a less disruptive and more fruitful direction for design. In the study, folder views offered some participants ways to make their environments fit with the practical and conceptual demands of their work, exploiting tacit knowledge and exposing affordances necessary for work. Folder views enabled this to happen in *ad hoc* ways that did not involve any persistent changes to their files or folders; opening and closing folder views alongside application windows was sufficient to support numerous valuable work processes. Enhancements to folder views could build on this, allowing files and folders to be rearranged, exploded or contracted, visualized or summarized, annotated or highlighted, all in a transient layer separated from the underlying persistent folder structure. Enhanced folder views could also better support task execution, by making the selection of items within views more stable, and allowing sequential interactions with files to be controlled. For example, an enhanced view could allow items to be chosen, compared, arranged and ranked, then fed directly into applications, supporting the low-level mechanical process of selecting and opening files, and the higher level conceptual processes of determining which files to select.

Enhancements to views would bring with them new state information (e.g. selections of items, scrollbar positions, states of visualizations) and the study showed how critical this information can be in supporting task execution. Enhanced views would bring the risk of subjecting this new, richer state information to loss, so care would need to be taken to protect this (for example, by allowing the closure of a folder view to be undone). Alternatively, views themselves could become part of the persistent content of folder structure, allowing valued perspectives on files to be stored alongside the files, introducing a new layer of functionality into folder systems without severing the valuable link folders have with underlying filesystems and in turn applications.

6.4.2. Virtual work environments

The findings of this research have implications for how operating systems and virtual environments for work are designed, in particular, for the trends towards application-centricity. Roth et al. (1997) use the term ‘basic currency’ to describe which kinds of objects in a virtual environment are given primacy, differentiating application-centric, document-centric and information-centric models. *Desktop Metaphor* operating systems such as Microsoft Windows and Apple OS X are document-centric, and allow individuals to organize and manipulate files outside of applications. Application-centricity is manifest in the mobile operating systems of today (e.g. Google Android and Apple iOS), but also in the design of desktop applications that de-emphasize or remove the need for individuals to interact with or ‘manage’ files (e.g. email clients, photo managers, music libraries).

In terms of user experience, a shift towards an application-centric approach certainly brings simplicity and removes the

need to wrangle files and folders, but based on the outcomes of this research, there may be other consequences for designers to consider. This study reveals uses of folder views and structures that facilitated valuable, and sometimes critical, creative work that was performed outside of any application. Having access to files and folders independent of applications gave some study participants the means to both conceptually and practically link together their work, from abstracted overviews of projects to controlling and combining the behaviour of applications. The ability to combine together the capabilities of applications in arbitrary and unexpected ways is particularly important for novel, creative work. In application-centric environments applications have to explicitly enable interchange with one another, doing so only in ways anticipated by their creators. Butfield-Addison et al. (2012) survey of tablet use by 507 information workers (96 followed up with interviews) reveals just such a ‘vendor lock-in’ effect at play, limiting interoperability between tablet applications.

Aside from the technical constraints of proprietary formats and applications, the study revealed how folder structures and views can play a critical cognitive role in combining the power of applications. Folders provide a means to collate and structure information in arbitrary ways, determined by the choices of individuals, not the boundaries between applications. Taking Kidd’s (1994) point about where meaning in an information worker’s environment resides, a key feature of resources in the environment is to allow the worker to reconstruct internal meaning, not to capture expressed external meaning. It is reasonable to argue that without a manipulable, unified view of activity that spans *between* applications there is a risk that information work will become restricted, both conceptually and practically, to individual tools (or groups of tools) which have been explicitly linked by their creators, rather than the intentions and ideas of the information worker. Folders are not the only means to meet these needs, but removing them without meeting them in other ways will surely have implications for the viability of virtual environments for information work, particularly if paper and other external resources become less available to information workers.

7. CONCLUSIONS

In this paper, we describe the outcomes of a study of folder use by 12 information workers. The approach used in the research responds to calls within PIM for new methods and detailed investigations of tool use (Jones, 2007; Whittaker, 2011) and captures the use of folder views in addition to folder structure. The findings provide a new perspective on the role that folders can play in information work, extending well beyond the classification of items to facilitate future retrieval, and into sense-making, planning, problem solving and creative synthesis. By examining folder view use, this research reveals the value that over-viewing files and folders can offer, and the

instrumental role that folder views can play in low-level task execution and high-level conceptual work. The interventional phase of the study revealed the dependency some study participants had on their folders in their processes of work, the interventions effectively paralysing some participants from both accessing files, and thinking about their work.

We use these findings to question the role and potential of folders in PIM activity. Rather than a laborious, manual means for organizing information that mimic the limitations physical information storage systems, we suggest that folders and the explicit manual organization of information can form an essential part of work processes themselves. Fixed structures and taxonomies have limitations, but give consistency and control that complements memory and the construction of ideas. Folders provide a medium for this which is connected directly with the documents and environment of information work, without the need for intermediary abstractions. The interactional connectivity between files, folders and the applications and windows of the desktop computer environments provides a lightweight and flexible means to match the content and affordances of the environment with the needs of work. The nature of individual work allows tacit knowledge to play an important role; the environment only need support individual thought and understanding, not capture it externally, as recognized by Kidd (1994).

Folders are one possible way to meet the need to group and organize information. The advent of digital technologies removed the need for physical folders to solve the problem of making documents retrievable in physical space, but the requirement to organize the information in the environment of those doing information work has not disappeared. Arbitrary control over this organization is part of how we think and create new ideas, not just in the world but also internally within the mind. Folders provide a context for this, independent of particular applications and the intent of their designers, something we should hold on to as our virtual environments for work evolve.

7.1. Further work

The limited size of study sample in this research, combined with the productivity and scalability of the observational methods employed, calls out for further research into PIM behaviour around information work. The focus on folder view use in this work was particularly fruitful and the observational method used is readily scalable to a larger sample and other professional contexts or desktop computing platforms. The importance of folder views and representations to study participants suggests further work, both observational and interventional, could explore the design of overviews, representations of task state and *ad hoc* visual 'holding patterns' in information work. The observational approach used in this work could be extended to prototype views integrated into the

existing environments of information workers, also subject to automated observation.

Our interpretation of the findings places primacy on providing manual, arbitrary control to information workers over their work environments. This suggests that further research could productively examine the possibility of a new structural layer within the virtual work environment facilitating manual combination of folder views, intensional views and application windows. Our interpretation also calls for more empirical work examining the cognitive effects of application-centric environment design on information work, investigating whether the requirement for all work to take place within applications effects how people think and use external resources. Further to this, interventional work could explore the potential role of application-independent representations within application-centric environments.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Prof. Alan Dix for the invaluable advice he gave on the design and development of this research.

REFERENCES

- Abrams, D., Baecker, R. and Chignell, M. (1998). Information archiving with bookmarks: personal Web space construction and organization. Paper presented at the Proc. of the SIGCHI conference on Human factors in computing systems.
- Alvarado, C., Teevan, J., Ackerman, M.S. and Karger, D. (2003). Surviving the information explosion: How people find their electronic information.
- Bardram, J., Bunde-Pedersen, J. and Soegaard, M. (2006). Support for activity-based computing in a personal computing operating system. Paper presented at the Proc. of the SIGCHI conference on Human Factors in computing systems.
- Barreau, D. and Nardi, B.A. (1995) Finding and reminding: file organization from the desktop. *ACM SigChi Bull.*, 27, 39–43.
- Barreau, D.K. (1995) Context as a factor in personal information management systems. *J. Am. Soc. Inf. Sci.*, 46, 327–339.
- Bellotti, V., Dalal, B., Good, N., Flynn, P., Bobrow, D.G. and Ducheneaut, N. (2004). What a to-do: studies of task management towards the design of a personal task list manager. Paper presented at the Proc. of the SIGCHI conference on Human factors in computing systems.
- Bellotti, V., Ducheneaut, N., Howard, M. and Smith, I. (2003). Taking email to task: the design and evaluation of a task management centered email tool. Paper presented at the Proc. of the SIGCHI conference on Human factors in computing systems.
- Benn, Y., Bergman, O., Glazer, L., Arent, P., Wilkinson, I.D., Varley, R. and Whittaker, S. (2015) Navigating through digital folders uses the same brain structures as real world navigation. *Sci. Rep.*, 5.

- Q7 Bergman, O., Beyth-Marom, R. and Nachmias, R. (2008a) The user-subjective approach to personal information management systems design: evidence and implementations. *J. Am. Soc. Inf. Sci. Technol.*, 59, 235–246.
- 5 Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N. and Whittaker, S. (2008b) Improved search engines and navigation preference in personal information management. *ACM Trans. Inf. Syst. (TOIS)*, 26, 20.
- 10 Q8 Bergman, O., Elyada, O., Dvir, N., Vaitzman, Y. and Ami, A.B. (2014) Spotting the latest version of a file with Old'nGray. *Interact. Comput.*, , iwu018.
- Q9 Bergman, O., Gradovitch, N., Bar-Ilan, J. and Beyth-Marom, R. (2013a) Folder versus tag preference in personal information management. *J. Am. Soc. Inf. Sci. Technol.*, 64, 1995–2012.
- 15 Bergman, O., Tene-Rubinstein, M. and Shalom, J. (2013b) The use of attention resources in navigation versus search. *Pers. Ubiquit. Comput.*, 17, 583–590.
- 20 Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R. and Ramamoorthy, A. (2010) The effect of folder structure on personal file navigation. *J. Am. Soc. Inf. Sci. Technol.*, 61, 2426–2441.
- 25 Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R. and Ramamoorthy, A. (2012). How do we find personal files?: the effect of OS, presentation & depth on file navigation. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems.
- 30 Boardman, R. and Sasse, M.A. (2004). Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems.
- 35 Buttfeld-Addison, P., Lueg, C., Ellis, L. and Manning, J. (2012). Everything goes into or out of the iPad: the iPad, information scraps and personal information management. Paper presented at the Proc. of the 24th Australian Computer-Human Interaction Conference.
- Q10 Cutrell, E., Dumais, S.T. and Teevan, J. (2006a) Searching to eliminate personal information management. *Commun. ACM.*, 49, 58–64.
- 40 Cutrell, E., Robbins, D., Dumais, S. and Sarin, R. (2006b). Fast, flexible filtering with phlat. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems.
- 45 Dourish, P., Edwards, W.K., LaMarca, A. and Salisbury, M. (1999) Presto: an experimental architecture for fluid interactive document spaces. *ACM Trans. Comput. Hum. Interact. (TOCHI)*, 6, 133–161.
- Ducheneaut, N. and Bellotti, V. (2001) E-mail as habitat: an exploration of embedded personal information management. *Interactions*, 8, 30–38.
- 50 Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. and Robbins, D.C. (2003). Stuff I've seen: a system for personal information retrieval and re-use. Paper presented at the Proc. of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.
- Freeman, E. and Gelernter, D. (1996). Lifestreams: a storage model for personal data. *ACM SIGMOD Rec.*, 25, 80–86.
- González, V.M. and Mark, G. (2004). Constant, constant, multi-tasking craziness: managing multiple working spheres. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 60
- Henderson, D.A. and Card, S. (1986) Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. *ACM Trans. Graph. (TOG)*, 5, 211–243.
- Hollan, J., Hutchins, E. and Kirsh, D. (2000) Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput. Hum. Interact. (TOCHI)*, 7, 174–196. 65
- Hutchins, E. (1995) *Cognition in the Wild* (Vol. 262082314). MIT press, Cambridge, MA.
- Huynh, D., Karger, D.R. and Quan, D. (2002). Haystack: A Platform for Creating, Organizing and Visualizing Information Using RDF. Paper presented at the Semantic Web Workshop. 70
- Jones, W. (2007) Personal information management. *Ann. Rev. Inf. Sci. Technol.*, 41, 453–504.
- Jones, W., Bruce, H., Foxley, A. and Munat, C.F. (2006) Planning personal projects and organizing personal information. *Proc. Am. Soc. Inf. Sci. Technol.*, 43, 1–24. 75
- Jones, W., Phuanartnurak, A.J., Gill, R. and Bruce, H. (2005). Don't take my folders away!: organizing personal information to get things done. Paper presented at the CHI'05 Extended Abstracts on Human Factors in Computing Systems. 80
- Jones, W., Wenning, A. and Bruce, H. (2014). How do people re-find files, emails and web pages? *iConference 2014 Proceedings*.
- Kaptelinin, V. (2003). UMEA: translating interaction histories into project contexts. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 85
- Kaye, J.J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I., Pinch, T. (2006). To have and to hold: exploring the personal archive. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 90
- Kidd, A. (1994). The marks are on the knowledge worker. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems.
- Kirsh, D. (1995) The intelligent use of space. *Artif. Intell.*, 73, 31–68. 95
- Kirsh, D. (2001) The context of work. *Hum.–Comput. Interact.*, 16, 305–322.
- Kwasnik, B. (1989). How a personal document's intended use or purpose affects its classification in an office. Paper presented at the ACM SIGIR Forum. 100
- Lansdale, M.W. (1988) The psychology of personal information management. *Appl. Ergon.*, 19, 55–66.
- Malone, T.W. (1983) How do people organize their desks?: Implications for the design of office information systems. *ACM Trans. Inf. Syst. (TOIS)*, 1, 99–112. 105
- Massey, C., TenBrook, S., Tatum, C. and Whittaker, S. (2014). PIM and personality: what do our personal file systems say about us?

- Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 55
- Rao, R., Card, S.K., Johnson, W., Klotz, L. and Trigg, R.H. (1994). Protofoil: storing and finding the information worker's paper documents in an electronic file cabinet. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. Q1
- Robertson, G., Horvitz, E., Czerwinski, M., Baudisch, P., Hutchings, D.R., Meyers, B., Robbins, D., Smith, G. (2004). Scalable Fabric: flexible task management. Paper presented at the AVI. 60
- Robertson, G., Van Dantzich, M., Robbins, D., Czerwinski, M., Hinckley, K., Risdien, K., Thiel, D., Gorokhovskiy, V. (2000). The Task Gallery: a 3D window manager. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 65
- Roth, S.F., Chuah, M.C., Kerpedjiev, S., Kolojejchick, J.A. and Lucas, P. (1997) Toward an information visualization workspace: combining multiple means of expression. *Hum.-Comput. Interact.*, 12, 131–185. 70
- Smith, D.C., Irby, C., Kimball, R., Verplank, B. and Harslem, E. (1982) Designing the Star User Interface: the star user interface adheres rigorously to a small set of principles designed to make the system seem friendly by simplifying the human-machine interface. *Byte*. 242–282. . April. 75
- Teevan, J., Alvarado, C., Ackerman, M.S. and Karger, D.R. (2004). The perfect search engine is not enough: a study of orienteering behavior in directed search. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 80
- Whittaker, S. (2011) Personal information management: from information consumption to curation. *Ann. Rev. Inf. Sci. Technol.*, 45, 1–62. 85
- Whittaker, S., Matthews, T., Cerruti, J., Badenes, H. and Tang, J. (2011). Am I wasting my time organizing email?: a study of email refinding. Paper presented at the Proc. of the SIGCHI Conference on Human Factors in Computing Systems. 90
- Whittaker, S. and Sidner, C. (1996). Email overload: exploring personal information management of email. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 95
- Zhang, J. (1997) The nature of external representations in problem solving. *Cogn. Sci.*, 21, 179–217. 100
- Zhang, J. and Norman, D.A. (1994) Representations in distributed cognitive tasks. *Cogn. Sci.*, 18, 87–122. 105