

The Character, Value, and Management of Personal Paper Archives

STEVE WHITTAKER and JULIA HIRSCHBERG
ATT Labs—Research

We explored general issues concerning personal information management by investigating the characteristics of office workers' paper-based information, in an industrial research environment. We examined the reasons people collect paper, types of data they collect, problems encountered in handling paper, and strategies used for processing it. We tested three specific hypotheses in the course of an office move. The greater availability of public digital data along with changes in people's jobs or interests should lead to wholesale discarding of paper data, while preparing for the move. Instead we found workers kept large, highly valued paper archives. We also expected that the major part of people's personal archives would be unique documents. However, only 49% of people's archives were unique documents, the remainder being copies of publicly available data and unread information, and we explore reasons for this. We examined the effects of paper-processing strategies on archive structure. We discovered different paper-processing strategies (*filing* and *piling*) that were relatively independent of job type. We predicted that filers' attempts to evaluate and categorize incoming documents would produce smaller archives that were accessed frequently. Contrary to our predictions, filers amassed more information, and accessed it less frequently than pilers. We argue that filers may engage in *premature filing*: to clear their workspace, they archive information that later turns out to be of low value. Given the effort involved in organizing data, they are also loath to discard filed information, even when its value is uncertain. We discuss the implications of this research for digital personal information management.

Categories and Subject Descriptors: H.1.1.2 [Models and Principles]: User/Machine Systems—*human factors*; H.5.2 [Information Interfaces and Presentation]: User Interfaces—*evaluation/methodology*

General Terms: Human Factors

Additional Key Words and Phrases: Document management, information retrieval, paper, personal information management, archiving, filing

1. INTRODUCTION

This empirical study investigates the nature of office workers' paper-based information, to explore general issues concerning personal information management. Despite the past importance of paper archives in office work, there are still relatively few studies of their nature and function [Cole 1982; Kidd 1994;

Authors' address: AT&T Labs—Research, Shannon Labs, 180 Park Avenue, Florham Park, NJ 07132; email: {steve; julia}@research.att.com.

Permission to make digital/hard copy of part or all of this work for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

© 2001 ACM 1073-0516/01/0600-0150 \$5.00

ACM Transactions on Computer-Human Interaction, Vol. 8, No. 2, June 2001, Pages 150–170.

Lansdale 1988; Malone 1983]. Some recent work has documented the importance of paper versus digital media in supporting *collaboration* [Frohlich and Perry 1994; Heath and Luff 1996; Luff et al. 1992; Sellen and Harper 1997; Shapiro et al. 1991; Trigg et al. 1999; Whittaker and Schwarz 1995; 1999], and *reading* [Sellen and Harper 1997; O'Hara and Sellen 1997]. None of these studies have addressed the *processing* and *archival* functions of paper, however, which is our concern here.

We therefore address fundamental questions about stored paper data, and archives more generally. Do paper archives continue to play a role in the modern office? What uses do paper archives serve? Do they contain irreplaceable data, or disposable information? What makes paper information valuable? How are archives acquired, maintained, and deleted? And what are worker's general attitudes to, and strategies for, personal information management? More specifically, we examine three hypotheses about the character, handling, and value of personal paper information: these concern *obsolescence*, *uniqueness*, and *filing* behaviors.

The *obsolescence* hypothesis contends that paper archives that were once highly valued are increasingly becoming irrelevant, because of general changes in the nature of office work, along with personal changes in the archiver's job, or interests. One general change is from paper to digital data. For years now, the popular technology press has argued that the paperless office is imminent because of the ubiquity of computers, along with greater ease of distributing and searching digital data. This shift should also be accelerated by the emergence of public digital data stores on the Internet. The third postulated shift is away from work centered on large stores of archival data. The increasingly unpredictable nature and faster pace of modern office work mean that the new breed of "knowledge workers" has vastly reduced reliance on precedent and historical information, with subsequently decreased need for archival data [Barreau and Nardi 1995; Kidd 1994]. Finally archived documents may become less valuable as the archiver's interests, job responsibilities, or company strategy change with time.

The *uniqueness* hypothesis concerns the reasons why people retain paper information. Previous research [Cole 1982; Lansdale 1988] identified three different *types* of paper data: *action information* (to be processed in the next few days), *working information* (relevant to current projects, valuable for periods of weeks and months), and *archival information* (long-term utility, i.e., periods of years, where relevance is not tied to current projects). However, this early work does not address the fundamental reasons underlying the retention of data, especially archival information. What factors influence decisions to keep particular paper documents? Do people only keep frequently accessed data? Do they keep documents that have proved useful in the past, or copies of documents that are unavailable elsewhere? The *uniqueness* hypothesis postulates that people are rational about storage. They therefore tend to retain only unique data, it being inefficient to store duplicates of data available elsewhere.

The *filing* hypothesis concerns strategies for handling paper data. Malone [1983] identified two basic strategies for handling paper: *filing* and *piling*. *Filers* maintain clean desktops, and do not allow papers to pile up. They systematize

their archives (using alphabetical, conceptual, or temporal methods) to support straightforward access to stored data. In contrast, *filers* have messy desktops cluttered with paper piles, making few attempts to organize stored information. Malone speculated that paper-handling strategies are a response to job requirements: predictable work encourages using structured filing systems, whereas less procedural work elicits piling strategies. A commonly held intuition is that *filing* is a superior approach to paper processing, leading to smaller, better organized, and hence more accessible archives. We test whether filing is indeed superior and investigate what factors influence choice of paper-handling strategies.

To address these questions, we collected both qualitative and quantitative data about personal paper information from 50 office workers in a research laboratory, including managers, secretaries, and researchers. We timed our investigation to coincide with an office move, which had important implications for data collection. Workers had all recently sorted through their paper data in preparation for the move. The new offices had slightly less storage space within each office, although extra storage was provided in public locations. This reduction in local storage seemed to motivate careful discarding and sorting of existing data. In interviewing and surveying workers when we did, we capitalized on the fact that they had very recently handled most of their paper data. Furthermore, they had recently been forced to identify criteria for determining what to keep and what to discard. The move also meant that we could collect quantitative data about volumes of data that people moved, and estimates of how much paper people discarded. All workers packed their archives into identical-sized professional movers' boxes, making it easy to quantify and compare different archives. Workers were also given control of the layout, but not the size, of their new offices. We were therefore able to collect data about aspects of people's physical set-up that are relevant to managing paper data.

Researchers may have different archiving habits from other office workers. The nature of their work may require them to access and reference publicly available documents, more often than other workers. Nevertheless, our data should add to general knowledge about the acquisition, handling, and discarding of paper archives. They may also inform us about general attitudes to, and strategies for, handling personal *digital* information, a topic which has received some recent attention [Barreau and Nardi 1995; Fertig et al. 1996; Kidd 1994]. We conclude with a discussion of the implications of our findings for designing tools for processing personal digital data.

2. METHOD

We collected on-line survey data from 50 people with different organizational roles (37 researchers, 9 managers, and 4 secretaries) about how much data they preserved, and their estimates of how much they discarded. We also collected information about each person's new office: amount of storage space (as measured by total number of filing cabinets) and number of working surfaces (furniture was modular so we counted the number of standard 2 ft.-by-5 ft. surfaces). The survey also addressed:

- job title, time in current office, job, and years in full-time employment;
- how people spent their worktime;
- date and reason why they last cleaned out their paper archive;
- different types of paper data workers collected prior to the move;
- frequency of access to different types of data;
- types of data disposed of;
- time to pack, including time to decide what to keep and what to throw away.

We also conducted follow-up interviews with a subset of 14 workers (7 researchers, 6 managers, and one secretary), concerning the same topics, but also to determine more explicitly:

- attitudes to paper archives;
- reasons for keeping and disposing of information;
- strategies for processing paper data.

All interviews were conducted using a semistructured questionnaire and were tape-recorded. There were over 35 hours of recorded interviews, and analyses were conducted on interview transcripts. The survey was administered as a Web form using the company's intranet. The response rate was 43%.

3. RESULTS

3.1 Obsolescence Hypothesis

We predict the following behaviors, for *original archives* (paper data existing before the move), *preserved archives* (kept after the move), and *discarded information* (discarded while preparing for the move):

O1: Old paper archives have decreased value both because of greater availability of public digital archives, and reduced personal relevance due to changes of interest, job specification, or company strategy. Workers will discard large proportions of their *original archives* and have relatively small *preserved archives* after the move.

We did not find good support for this prediction. Even after spending large amounts of time deciding what to discard, workers still retain huge archives after the move. Overall people retain 78% of original archives, with the average preserved archive being more than 18 mover's boxes (equivalent to a pile of telephone directories 63 ft. high).

O2: Old paper archives have decreased value both because of greater availability of public digital archives, and reduced personal relevance due to changes of interest, job specification, or company strategy. *Original archives* will have low perceived value: workers will find it both easy and quick to throw away information.

Again there was not strong support for this prediction. People on average reported spending almost nine hours rationalizing their data in preparation for the move, and the interviews indicate that this is a difficult process:

“Pretty gruesome set of circumstances, when you’re going through this. You feel like you’re casting off your first born.”

The finding that workers expended large amounts of time and effort rationalizing archives while preparing for the move indicates that people ended up with large archives *deliberately* and not because they lacked the time or inclination to sort through their data.

O3: Given that electronic archives and the Internet are relatively recent phenomena, we predict that less-experienced workers will rely more on external digital stores and hence acquire paper more slowly than more experienced workers.

Again there was not good support for the hypothesis. Less-experienced workers (with the company for less than a year) are acquiring paper at 7.21 boxes/year compared with 3.33 boxes/year for more-experienced workers (with the company for more than one year) ($t_{(48)} = 2.09$, $p < 0.05$). Using a simple measure of years in the company, however, neglects the fact that people change jobs. Recent arrivals with long employment histories may arrive with large amounts of data. We therefore calculated a second measure of experience. People who had less than 10 years’ full-time employment were acquiring data at 2.77 boxes/year, and those with more than 10 years employment were acquiring 1.28 boxes/year ($t_{(48)} = 4.02$, $p < 0.001$).

Together these data indicate paper archiving is not the prerogative of older workers, with long-established habits of processing paper. Instead, younger workers acquire paper data more rapidly than their more-experienced counterparts. Less-experienced workers may be unsure of what is relevant, and are consequently conservative in what they keep. Older workers, in contrast, can better identify valuable information and discipline themselves in acquiring new information. An alternative, however, is physical space limitations. As we shall see, spontaneous clean-ups are relatively rare, and more-experienced workers may simply have no room for new documents.

We next analyzed types of discarded information, and found multiple reasons other than obsolescence for discarding information.

3.1.1 Characteristics of Discarded Information. Overall people threw away 22% of their data, but what types of information were discarded? The interviews revealed that at least part of discarded data was obsolete. People’s statements showed the value of certain pieces of information had decreased, and that once-valuable information was now superfluous. People were aware that as jobs, personal interests, or company strategy changes, then the value of particular information decreases:

“I threw a lot of . . . stuff that reflects the history of the company . . . things that seem out-of-date—projects I worked on with people from company A. Since company A is gone, my company A files are all gone. They were useful because they showed you what people in the business units [at company A] were about. But, the company’s changed so much that anything pre-1995 . . . is almost useless.”

But not all discarded information underwent the transition from valuable to obsolete. For example, 23% of discarded data were *unread*. The interviews

suggest two general problems in processing data lead to the accumulation of such superfluous information. People experience problems with *information overload* and engage in *deferred evaluation* causing them to acquire large amounts of paper data that later turn out to be extraneous.

Information overload refers to the fact that people have insufficient time to process all the information they acquire, so that nonurgent information is never processed. Nonurgent data are set aside (often in optimistically named “to-read” piles), accumulating indefinitely, because the same time pressures that prevent complete processing of incoming data also prevent rationalizing (“clean-up”) of archives. Consequently, people seldom discover that these unread nonurgent documents are superfluous until exceptional circumstances (like the current office move) force them to return to the unprocessed data.

“One of the real problems in the age in which we live is I receive more material to read in a day than I can read in a day. And that happens day in and day out, and it also accounts for why the paper overflows. I mean, these file cabinets weren’t chock full when I first got this furniture in 1992. More stuff just keeps coming in, and the pace of things is such that you can’t say ‘I’m going to take a week out, and I’m going to go through my files to clean them.’”

Yet even when people find the time to examine new documents, uncertainty of the future value of information means they are highly conservative and postpone final judgments about utility until some unspecified future date. People deliberately *deferred evaluation* about incoming information, allowing time to pass so as to make better-informed judgments:

“This is sort of like a poor moral hygiene response, but it almost feels like I didn’t feel like making a decision for a while. And I just let things accumulate. . . . Partly, it’s a matter of time. You know; it’s the classic thing. If you don’t use something for a while, you can sort of decide you’re not going to need it.”

In consequence, people retain information of low perceived value just in case it later turns out to be useful. Judgments about potential utility are made more difficult because the value of the same data can change over time. The difficulty of this decision causes people to postpone it while there is archival space:

“Life is too short to come through and sort it, and when it’s fairly recent . . . you really can’t judge what it is that you need and what it is you don’t need. . . . A year from now is when I should probably throw it . . . But will I? It depends on whether I need space or not.”

A major problem with accumulating unprocessed data and deferring evaluation is that users seldom rationalize their archives. Thus, 74% had not cleaned out their archives for over a year. Furthermore very few clean-ups occur spontaneously: 84% arise from extrinsic events such as job changes or office moves. This infrequency of clean-ups means that documents are often not discovered to be superfluous, until they have been stored for some time.

Overall, the evidence for the obsolescence hypothesis is negative. People retain large amounts of *preserved information* they perceive to be of high value, taking large amounts of time deciding what to discard, and newer workers tend to acquire paper data faster than more experienced ones. While there is apparently weak support for obsolescence, with workers discarding about 25% of data,

it seemed that factors besides obsolescence such as *information overload* and *deferred evaluation* contribute to them amassing large amounts of marginally relevant data. Rather than discarding once-valuable information that is now of little utility, much of what people discard is unprocessed information they have never properly evaluated.

3.2 Uniqueness Hypothesis

U1: People will not store personal copies of publicly available data. People will keep only data that are unique, because it is inefficient to keep data that are available elsewhere.

Uniqueness was clearly important in determining whether users would preserve certain documents. Three types of *unique* data accounted for 49% of the archive: *working notes*, *archives of completed projects*, and *legal documents*. Unique data are often written by or highly associated with their archiver.

3.2.1 Unique Data. The first type of unique personal data was *working notes* for current projects, accounting for 17% of people's preserved archives. These were usually handwritten and generated by archivers. Their function was to store context about the current project, recording what had been done, and reminding people of what remained to be done. Because working notes were relevant to daily project activities, they were accessed highly frequently.

"I would be very upset if I were to lose them because when you are right in the middle of a problem, you can think of certain things, which, when you are separated for a period of time from it, you might not be able to think of it immediately. And, sometimes when I look at them I wonder, 'How did I think of that?' . . . That's irreplaceable because it was something that was unique to the problem, unique to the situation, and I don't think I could recreate that again without a considerable expense of time and effort. So, I would be disturbed about losing things like that."

Archives of Completed Projects. Most projects require writing reports. These were almost always preserved, along with unique information about the *content* of the project, such as empirical or marketing data, technical specifications, or code printouts. Together they accounted for 23% of people's data. Reasons for keeping *completed project archives* included answering potential questions and having relevant materials if the company decided to start a related project.

"Some projects get resurrected in different forms, and sometimes, when a new project comes along, rather than, sort of, re-inventing the boiler plate and structure and the kinds of documents that a project needs. . . . I'd rather not keep it, but there's nowhere to find it. So, if I don't keep it, then there's no way to recover it. . . . It's partly personal, partly social. . . . I continue to work on projects, and sometimes documentation like that is of use to the group, just to sort of see what we've done before."

Legal or Administrative Data. A third type of unique data was quasi legal documents that were highly associated with the archiver. These included personal paperwork associated with patents, personnel data (performance evaluation forms, or for managers, employee information), equipment specifications,

contracts, and expense records. Often these data were irreplaceable. Legal and administrative data accounted for 9% of people's archives, and they were not frequently accessed.

For these unique types of data, we probed why people kept *paper* rather than *digital* versions. *Working notes* were originally generated as handwritten paper notes, and were usually preserved in their initial form. Paper notes can be generated easily, and kept around the desktop or carried, which in turn increases availability and reminding. *Legal or administrative data* were also generated on paper—as they often required signatures and their legal status depended on them being preserved in paper format. For *archives*, paper documents were reported to be easier to read, reference, and distribute by hand.

3.2.2 Keeping Documents that Are Not Unique. *Uniqueness* was not the sole criterion for determining whether to archive data, however. Only 49% of people's original archive was unique: 15% was unread, and 36% consisted of copies of publicly available documents. We have already discussed why people acquire unread data, but what explains them keeping paper duplicates of publicly available documents? Four main reasons were given: *availability*, *reminding*, *lack of trust in external stores*, and *sentiment*.

Availability allows relevant materials to be at hand when people need them. Several people mentioned not wanting to experience the delay associated with ordering reference materials, or even accessing them on the Web.

“What is important is how fast I can access that information. . . . My problem has been often I am unable to find where things are quickly, and if it is not quick, you lose interest.”

The importance of immediate access was also revealed in discussions about off-site storage of archival material. During discussions prior to the move, concern was voiced about lack of storage at the new location. In response to this, management offered to provide off-site storage. The majority of interviewees, however, felt that access delays reduced the usefulness of off-site information, revealing the importance of availability:

“I have nothing I would store or put off site. As far as I'm concerned, if it goes that far away, I'll never see it again. That's just a hard way to throw things away. I have put things in [name of local storage company], and I have never, ever had occasion to get things back. All you eventually do is sign off and let them destroy it.”

Reminding Relates to Availability. A personal paper copy prompts people about outstanding actions associated with a document, or simply reminds them they are in possession of that information. Documents in public or digital stores do not seem to support reminding:

“And there's also a reminding issue; if I have a piece of paper . . . it's a physical manifestation or reminder that this is something I can reference and something that is available that I can read, or so forth. Even if it happens to exist inside a bound collection of journals in a library somewhere, my memory may not serve well enough to note that it's there, and there would be the inertia involved in going and finding it.”

People also kept personal copies of public data because they did not trust ad hoc workgroup libraries or other archival institutions to keep the documents they needed:

“I got the ‘LMN Journal’ before that, for quite a while, as kind of a public service to the LMN community here, and I would dutifully put them in [the workgroup library], every once in a while. And, I started realizing that the ones I was counting on were walking. So, I started making copies of the articles I really cared about, and many of the articles in the cabinets in here are those kinds of articles. . . . if I had guaranteed going forward access to a library that I knew would supply my needs, I would be happy to donate them to a worthy home.”

Distrust of external stores also extended to digital libraries such as the Web. People had little faith in the Web as a reliable way to obtain valued documents:

“I’m not sure that I believe that . . . 10 years from now, stuff that is available on the Web now will still be available. I have no reason to believe that. Zero.”

“I haven’t thought about the Web . . . because I don’t have a sense of confidence that it is as comprehensive. If it were like a library, and I could really trust it, then I would throw a lot of this stuff out.”

Indeed, one person already had the experience of discarding paper versions of information which had been on-line but were now no longer available:

“I had already thrown away a great deal and not kept things because it was on-line, and now it’s not on-line. And I’m in trouble.”

This mistrust of external stores was not completely general. Some people placed reliance on known experts as “informal librarians” to store relevant information for them.

“I’m selfish. I say that if . . . colleagues have all this, why should I store it? Why should I store duplicates?”

Some people devolved storage responsibilities to others with superior expertise and higher-quality document collections in the relevant domain.

“I know I may want to look at this again, but I know where to find it. . . . I can go to an expert in the field and ask them for the latest on it, rather than keep this old thing that may be outdated.”

People who relied on others often had sophisticated knowledge about who had relevant information:

“A . . . has a beautiful collection of journals that I don’t subscribe to and don’t keep. He’s the only person in the whole damn work group that I have seen who does that, and he doesn’t store anything else. It’s an extremely valuable collection there. B. has all the ABC’s from day one. I don’t store these. C. has all the DEF proceedings from day one. I threw all of mine away. I don’t even store it. . . . I don’t keep it. It gets thrown away. I used to be a librarian. . . . It’s a shared library system . . . and it’s better than the real library because it’s fairly free of cost.”

In addition to functional reasons, people described *emotional* or *sentimental* reasons for keeping information. These included reviews of their first published paper and specifications for successful research prototypes. People admit such

information has little relevance for likely future activities, but they still cannot part with it, because it is part of their intellectual history:

“[The papers] are, like, by a famous person . . . or things that you might have some sort of stupid sentimental reason for keeping . . . even though you have no need for it, or whatever. Sentiment . . . or something that prevents me . . . Like all the stuff on my thesis. I have piles of stuff from that, but I just can’t throw that out. There are books, like my undergraduate and graduate text books, that I don’t know why I’m keeping them . . . that I could put away because I never . . . Like my book on operating systems or my book on PL1, I can’t throw out for some reason.”

Another potential reason for keeping personal copies of long-term reference documents is that they contain personal annotations. Other research has documented the utility of annotations for focussing attention and improving the process of reading [O’Hara and Sellen 1997; Schilit et al. 1998; Sellen and Harper 1997]. Although a number of people made such annotations, they seemed of little long-term use, however. Most people stated that annotations have transient value, becoming uninterpretable after some time has elapsed.

“I used to keep them [papers] for my valuable comments in the margins. I’ve decided none of my comments are ever valuable, when I reread them.”

3.2.3 What Makes a Publicly Available Document Valuable? Finally people talked about the characteristics of valued publicly available documents. Many documents they kept were research papers, and they gave three main reasons for keeping these. The first was *personal influence*, whether the paper changed the way that they thought about a problem. The second was *citation*, both personal and public. Personal citation means that the archivers themselves refer to the paper frequently. Public citation means that others often cite it. The third criterion was the quality of a paper’s *bibliography*:

“Many papers I had cited, or I had felt had influenced my own work. . . I tried to save those. I also tried to save the papers that have, kind of, stood the test of time and been referenced often and papers that were of a more survey nature that I felt would be good references to go back . . . had large bibliographies.”

3.3 Filing Hypothesis

3.3.1 Filing and Piling. Malone [1983] documented two paper-processing strategies: filing and piling. Filers keep a clear desktop, putting most personal information into structured archive files. With piling, in contrast, much incoming information ends up as working or unprocessed information distributed around the desktop, and less information is archived into filing cabinets. The questionnaire responses and the interviews indicated, however, that the distinction between filers and pilers was one of degree. All respondents filed some information but kept other information in desktop piles. We therefore classified users according to how likely they were to file information. Based on the predominant strategies that people described in our interviews we identified a threshold of 40%. People in the survey who reported more than 40% of their information being either unread or working were classified as pilers, and the

Table I. The Relationship between Filing Strategy and Archive Characteristics

	Pilers N = 15	Filers N = 35	Statistical Significance? (t test)
Time in job (years)	9.3	9.4	No
Time in current office (years)	4.7	3.0	No
Original archives: (# boxes before move)	12.4	20.5	$p < 0.05$
Preserved archives: (# boxes moved)	8.8	16.3	$p < 0.01$
Mean # preserved boxes per year in job	1.8	3.8	$p = 0.06$
# boxes accessed in last year	4.3	4.1	No
% data accessed within last year	41.8	31.3	$p < 0.04$
# boxes discarded	2.1	4.9	No
# reference boxes discarded	2.3	5.7	$p < 0.03$
% of people finding it “easy” to clean up archives	43	7	$p < 0.02$
Time to pack (hours)	9.1	8.6	No
# working surfaces in office	4.2	3.3	$p < 0.05$
# office filing cabinets	8.9	8.8	No

remainder as filers. Table I shows the differences between filers and pilers for different measures.

F1: Filers often classify information as it accumulates. In contrast, pilers can amass information without attempting to systematically organize it, leading to an accumulation of unscrutinized information before the move. We should therefore expect pilers to have more *original archives* and to discard more information than filers.

Contrary to the hypothesis, pilers had smaller *original archives* than filers, ($t_{(48)} = 2.04$, $p < 0.05$). They also had less *preserved information* after cleaning out their archives ($t_{(48)} = 2.55$, $p < 0.01$). Why did filers amass more information? One possible reason is *premature filing*: filers may file information, which turns out later to be of little utility, that they later have to discard. Some managers for example accumulated superfluous data because their secretaries filed unread information without consultation.

“I have lots of internal, technical memos that I just got by default . . . and when I wasn’t looking, J [his secretary] filed them without asking me. And it turned out that I didn’t care for that particular paper . . . There were a bunch of things like that.”

If filers are more likely to file documents of uncertain quality, we might expect them to discard more materials in preparing for the move. This was not true for all documents ($t_{(48)} = 1.58$, $p > 0.05$), but was true for reference documents ($t_{(48)} = 2.17$, $p < 0.05$).

F2: Pilers should acquire information faster than filers because of their tendency to amass unscrutinized data.

We looked at data acquisition rates, in separate analyses of *original* and *preserved information* volumes divided by number of years in current job. Pilers tended to be *slower* to acquire original as well as preserved information, when

we normalize for time in the company: ($t_{(48)} = 1.81$, $p = 0.07$, $t_{(48)} = 1.91$, $p = 0.06$).

F3: Given the effort they address to constructing organized filing systems, filers should find it easier to find data, so they should access their data more often.

We compared access frequency by examining the percentage of documents accessed in the last year. Contrary to the hypothesis, pilers accessed a greater percentage of documents ($t_{(48)} = 2.13$, $p < 0.04$). Why were pilers more likely to access data? The interviews revealed both strategies had strengths and weaknesses. One reported advantage of piling is *accessibility*. With a piling strategy, information can be potentially located in multiple piles, with frequently accessed information moving to the top, and less relevant material moving down the pile:

“I wound up . . . having a multitiered system of stacking up papers, and if I continued to re-reference things, they would tend to be near the top . . . So, periodically, I would sweep of the bottom half.”

Taken to excess, however, piles can dominate not just working surfaces but all areas of the office. In clearing floors, tables, and chairs of their accompanying piles, in order to use them for their normal purposes, urgent piled information can be lost:

“I had kept stuff on this table as sort of the next place I would go to, but R [his secretary] has come in and sort of cleaned my office when we’ve had meetings over there. And, therefore, it was kind of lost. So, I can’t wait to see what I find in those piles . . . probably urgent stuff. The credenza was incoming mail, and I intended to keep that clean . . . and I couldn’t keep up with it. There’s a pile of stuff on the chair and the floor, which was things like articles and things that I had the intention to read soon, or to file. Or, I’d go to conferences and bring back a book bag full of stuff and not have any place to unpack it. So, I’d leave them there, and they’d pile up. So, I must have four or five conferences worth of junk. I have to clear off that chair. That’s another big one.”

Filing does not always guarantee easy access to information however. With complex data, filing systems can become so arcane that people forget the filing categories they have already created, leading to duplicate filing categories. Accessing only these duplicates leads to incomplete retrieval, because information in the original files gets overlooked:

“I had so much stuff filed. I didn’t know where everything was, and I’d found that I had created second files for something in what seemed like a logical place, but not the only logical place. And now, I’m discovering everything . . . Ah! That was there! . . . Or a whole chunk of things over here . . . Or, in some cases, things could legitimately be filed under the business unit or a technology. And I ended up having the same thing in two places, or I had the same business unit stuff in five different places.”

This illustrates a general disadvantage to filing strategies: there are large overheads for constructing, maintaining, and rationalizing complex systems. For some people, filing only takes place if they are provided with secretarial assistance. Without such assistance they fall back on piling as requiring less effort:

“The original ‘A’ to ‘Z’ stuff was nicely organized, largely, because I started this when I was a professor at S., and I actually had a quarter-time secretary. And the secretary could actually create bibliography references for me, and alphabetize and file. And my life stayed very organized, because I had some help. I did it myself for the first three years, maybe, that I was at [company name], and it was just too much effort. So, I wound up, instead, having a multitiered system of stacking up papers.”

A final possible reason why filers access proportionally less of their data is there are finite constraints on how much data one can access. Filers have more data, and consequently access proportionally less. Indeed the absolute amounts accessed by both groups were very similar.

F4: Filers should be quicker and find it easier to rationalize their data in preparing for the move, given the greater care they have taken to organize their data.

There were no differences in packing time ($t_{(48)} = 0.25$, $p > 0.05$). Contrary to our prediction, pilers found it easier to rationalize archives in preparation for the move. Filers were more likely to judge it “easy” to clear out archives ($\chi_{(1)}^2 = 6.32$, $p < 0.02$). Why was this? Although filers discarded more reference information, they generally found it difficult to discard filed documents:

“Once you’ve committed something to paper, and it sits in a file, it requires a separate act of attention and will to take the stuff back out and reduce files to their bare minimums. And, that’s a real problem because you just don’t want to spend time doing it. . . . because you know [documents] have limited lifetimes, but you don’t want to go back through it and do something about it.”

Filers therefore seemed less disposed to discard information they had invested effort in organizing. In contrast unfiled information was much easier to discard:

“I was aggressive with that [unfiled] stuff. . . because it really wasn’t well-organized yet, anyway. So, the cost of throwing it away. . . I wasn’t undoing any organizational cost that I’d really expended. So, it wasn’t too big a thing to stand there with a pile of papers over the trash can and ruthlessly throw them in. One out of every five I may throw into another pile to keep, and eventually, maybe it will earn its place of honor in the ‘A’ to ‘Z’ file. But, this stuff that’s already in here [points to filing cabinet] has already been filtered a couple of times.”

F5: Filers and pilers should have different physical set-ups in their offices. Filers should have more storage space (e.g., filing cabinets), and pilers more surfaces for organizing information.

As predicted, we found that pilers had set up more physical working surfaces in their new offices¹ ($t_{(38)} = 2.02$, $p < 0.05$). Contrary to our expectations there was no overall difference in total number of filing cabinets ($t_{(38)} = 0.20$, $p > 0.05$), although people with more paper data generally had more filing cabinets ($r_{(38)} = 0.41$, $p < 0.05$).

¹We were only able to collect physical information for 40 people in the study.

Table II. The Relationship between Job Type and Archive Characteristics. = signifies no significant difference, < means there is a significant difference, at 0.05 level.

	Researchers N = 37	Managers N = 9	Secretaries N = 4	Statistical significance? (t test)
Time in job (years)	7.5	19.0	5.8	R = S < M
Time in current office (years)	2.8	7.0	2.5	R < M, R = S, S = M
Original archives: # boxes before move	13.8	31.9	26.1	R < M, R = S, S = M
Preserved volume: # boxes moved	10.6	25.2	20.5	R < M, R = S, S = M
Mean # preserved boxes per year in job	2.9	1.7	9.0	R = M = S
% accessed within last year	33.1	17.0	87.2	R = M < S
# boxes discarded	3.6	9.1	7.7	R = M = S
# legal boxes discarded	0.9	3.2	8.7	R = M = S
# reference boxes discarded	5.8	11.9	6.1	R = M = S
% of filers	70	67	75	R = M = S
Time to pack (hours)	7.5	13.3	9.8	R = M = S

3.3.2 *Filing and Job Type.* We also tested hypotheses about job type affecting paper-processing strategies and archival characteristics:

F6: Job type should affect choice of strategy. In the organization we studied, managers in effect carried out two jobs, doing research along with management activities. We expected the pressure of executing two jobs, along with maintaining the additional information needed, would leave little time for organizing data, leading managers to be pilers.

F7: Job type should affect the character of the archive: managers and secretaries should store more data than researchers, as both jobs require access to considerable amounts of administrative data.

Table II shows how archive characteristics relate to job type. Contrary to F6, managers were no more likely than researchers to be pilers than filers ($\chi^2_{(1)} = 0.001$, $p > 0.05$). Personal style may therefore be a more likely determinant of strategy choice than job type. Another possibility, however, is that managers had secretaries who took on the burden of filing. Contrary to F7, managers' archives are generally structurally similar to researchers', in the types of data they preserve, and frequency of access. The only differences are that managers have more data before ($t_{(44)} = 4.99$, $p < 0.0001$) and after the move ($t_{(44)} = 4.71$, $p < 0.0001$), although there is no difference in the percentage of their original archives they discard in preparing for the move. Similarly there were few differences between researchers and secretaries.² One reason why there were no archival differences may be because of researchers' propensity for storing large amounts of personal data. The one major difference between secretaries and the two other groups lay in secretaries' overall frequency of archival access: for secretaries versus managers ($t_{(4,6)} = 9.42$, $p < 0.001$)

²Welch's correction was applied to the t test comparisons, because of the small number of secretaries in the sample.

and for secretaries versus researchers ($t_{(6,4)} = 6.68$, $p < 0.001$). Secretaries may need to access archives more frequently, as their jobs require them to file, record, and track archival data.

What can we conclude about filing strategies? While filers seem to find it easier to access information, they do so at the cost of acquiring large archives which they have to organize and manage. In addition, they are less likely to access a given piece of data, and more likely to acquire extraneous data. One reason why filers find it hard to discard large amounts of information is the organizational effort they have invested in filing information. In contrast, piled information requires less effort to manage, and may be easier to discard. The access data also suggest that in moderation piling has the benefits of providing somewhat ready access to materials as well as reminding about tasks currently in process. In addition pilers, unlike filers, encounter and discard extraneous information while searching for other information. Job type did not seem to determine paper management strategy: while secretaries accessed information more frequently, there were few major differences between researchers and managers, apart from the total amount of information managers stored.

4. CONCLUSIONS

Overall our study generated a number of counterintuitive findings. While studies of other types of office workers are needed before we can generalize these results, we found little compelling evidence that paper had no role in modern office work. People kept large, valued archives, even after investing considerable effort in discarding and sorting information while preparing for the move. The scale, growth rate, and attitudes to paper archives suggest they continue to be a valuable resource for workers, and younger workers were just as likely to rely on paper as older ones. While we did find that people threw away 22% of their archives, this seemed to be for reasons other than wholesale obsolescence. People acquired extraneous paper data, for two reasons: a lack of time to process new information causes unread information to accumulate, and the inability to gauge the future utility of information leads people to save data of questionable utility.

Our findings concerning the obsolescence hypothesis add to the growing literature attesting the intrinsic utility of paper, and contradict the view that the paperless office is imminent [Johnson 1997]. Other research has described the utility of paper for *reading* [O'Hara and Sellen 1997; Sellen and Harper 1997], and for *collaboration* [Bowers 1992; Heath and Luff 1996; Luff et al. 1992; Sellen and Harper 1997; Shapiro et al. 1991; Whittaker and Schwarz 1995; 1999]. The current study showed the value of paper for a different class of task: *long-term memory* and *reference*. Our users had large, valued, archives of paper documents, which they took considerable pains to preserve and manage. Strikingly, even new office workers were compiling archives of paper data, despite the increased availability of digital materials and the growth of the Internet. The value of paper *archives* partly derives from paper's affordances. People keep paper documents because paper is more *available* and better supports *re-minding*. Another rather different reason why paper archives persist is people's

difficulty in processing paper. Paper is still an important mechanism for delivering information in the modern office [Sellen and Harper 1996; Whittaker et al. 1994], but people do not have effective paper-processing strategies, and some incoming paper never gets processed. We return to the problems of processing paper data below. Another hypothesized archival value of paper was *annotation*, but this was not supported by our data. Handwritten notes seem to have short-term benefits (such as tracking current tasks), rather than enhancing the quality of archived paper documents in the long term. Other studies describing the value of paper annotations have examined short-term tasks such as reading and editing, and this may explain discrepancies between our findings and those studies.

A second finding is that (for these workers at least) paper information is valuable *in the long term*. Kidd [1994] argues, that for knowledge workers, paper is a critical method for *acquiring* information, but not for long-term storage. Contrary to this view, however, our users invested time and effort in managing large stores of paper data, and stated that they often accessed substantial proportions of stored data. Furthermore, given the opportunity to discard their data, they did not do so, keeping 78% of their paper data. In addition, much of what they discarded was *unprocessed* information, rather than once-valuable information that has become obsolete.

Workers also experience major problems with processing incoming paper documents, specifically with deciding their future value, and organizing them to facilitate future access. We contrasted two different paper management techniques. While *filing* had the benefits of applying a formal structure to paper data, filers also incurred costs. Consistent with prior research, people found it hard to operate an effective filing structure by applying consistent labels and organizational structure [Kidd 1994; Lansdale 1988; Malone 1983]. In addition, we documented the problem of *premature filing*: in their desire to keep a clear desktop, filers sometimes spent time organizing information that later turned out to be of questionable value. Once information is filed, it seems to be difficult to discard later, given the effort of organizing it. In consequence, filers end up with greater amounts of information than pilers. In contrast, we found a surprising set of advantages to a *piling* strategy. Pilers benefitted from the greater availability of information of recent information. They also experienced fewer overheads in managing their data and found it easier to clean up their archives. The main limitation of their technique was that it did not scale well: pilers found difficulties in accessing information once piles had begun to multiply.

Together with studies investigating strategies for processing email [Bälter 2000; Mackay 1988; Whittaker and Sidner 1996], voicemail [Whittaker et al. 1998; 2000], and information in the computer desktop [Barreau and Nardi 1995; Kidd 1994], our research indicates *general problems* of processing and organizing personal information. Whether information resides in paper documents, email and voicemail messages, or computer files, users experience common difficulties with (a) determining the value of incoming information, (b) deciding whether and how to categorize that data, and (c) in electing whether to keep incoming information available in their immediate workspace or file it away

for later access. These problems are further exacerbated by (a) *information overload* because people receive more information than they can process each day; and (b) the *changing value* of stored information, as work priorities and personal interests shift. Information overload is a pernicious problem: having too much information to process leaves insufficient time to evaluate incoming documents, but there is also a continuous pressure to keep one's workspace clear for working documents. This pressure can lead to premature filing, and the temptation to adopt complex filing systems with potentially low benefits. It also leaves people with little time for rationalizing stored data, as we have seen.

For the *uniqueness* hypothesis, one significant reason for keeping data was that it was irreplaceable (i.e., working notes, project materials, and legal documents), and this accounted for 49% of people's data. However, this was not the only reason why people kept data, and 36% of people's information consisted of personal archives of publicly available reference documents. Important reasons for having personal copies were document *availability* along with *mistrust of storage institutions*, including the Web. Certain individuals established themselves as *informal librarians*—becoming a resource for others who did not care to keep extensive archives [Tushman 1977]. This finding is consistent with much recent work on social filtering where experts evaluate the quality of information for others [Communications of the ACM 1997; Goldberg et al. 1992; Hill et al. 1995; Resnick et al. 1994; Shardanand and Maes 1995]. Finally people kept data for *sentimental* reasons even though they could see little direct use for it.

What are the technological implications of these findings, when we turn to the organization of personal *digital* information? One general point is that a “network computer” providing access to global public databases will not suffice. People want *personal* copies of publicly available data, for a variety of cognitive and emotional reasons. Their mistrust of public stores means that a global repository managed by others would not be acceptable. Other technical implications concern affordances, and we should consider methods to make digital data mimic critical properties of paper. Unlike paper, current digital user interfaces fail to make data available or support reminding, although some research approaches attempt to reduce physical differences between paper and digital worlds. Some recent interfaces mimic important properties of paper in the digital world [Francik 1996; Wolf et al. 1989]. The “augmented reality” approach combines digital and physical worlds, allowing people to interact with digitally aware physical objects [Johnson et al. 1993; Mackay and Pagani 1994; Newman and Wellner 1992]. Other systems do not directly address the affordances of paper, but rather try to provide support for reminding and the maintenance of context. Techniques here include digital document piles [Mander et al. 1992], making communication tasks tangible as “yellow sticky” reminders on the computer desktop [Whittaker et al. 1997], and the “rooms” metaphor that allows users to associate sets of digital documents with different tasks, which regenerates context when switching tasks [Henderson and Card 1986].

Many of our users experienced problems with *processing* incoming information, where the dilemma concerned what to do with incoming documents of uncertain value. These problems are common to digital environments such as

email, voicemail, and the computer desktop [Barreau and Nardi 1995; Kidd 1984; Whittaker and Sidner 1996; Whittaker et al. 1998; 2000]. Should incoming documents of uncertain value be left piled up on the desktop until enough time has elapsed to better determine their value, or should they be filed away to clear the workspace? The decision to pile runs the risk of having papers amassing on the desktop, while decision to file runs the risks of premature filing. In the digital domain there are several techniques that might address these problems. Incoming information might be analyzed in two ways using document-indexing techniques. First, incoming documents might be evaluated in terms of their similarity to the user's digital archive. This would provide information both about potential utility (incoming documents having large overlaps with the existing archive should be valuable), and how it might be filed (along with similar archived documents). But users are also constantly inundated by documents on new topics, which may be different from their archive, but nevertheless of value. How does one make sense of these new data? Clustering analyses could be conducted on incoming information, to identify emerging new categories. Together these techniques might address the problem of information overload. The same techniques might also be used with deferred evaluation: users could "pile up" digital documents for a period of time before applying these analyses.

"Clean-ups" are a prevalent, if irregular feature, of digital and paper processing, occurring when users attempt to rationalize archives or identify overlooked materials requiring action. People experienced problems with cleaning up their archives, and again, document-indexing techniques might be used to evaluate the quality of archived digital documents. We might identify superfluous archived documents, by the fact that they have never been accessed, or because they bear little relation to the remainder of the archive. Other ways to evaluate documents concern their relationship to current user activity. Documents accessed repeatedly over the last few months might be given greater value than those with little relation to recent actions [Rhodes and Starner 1996]. These techniques could calibrate the changing value of information, allowing more informed decisions about what to discard.

We might also explore novel visual representations of personal data [Card et al. 1999]. It is worth noting that the current default system for personal digital information management, the virtual desktop, was invented almost 20 years ago. Furthermore, one underlying assumption of direct-manipulation systems like the desktop is that the number of objects is small [Shneiderman 1982], yet people's personal archives are often vast. In addition, digital filing systems like the virtual desktop emphasize semantic categorization of data. In doing so, they ignore the contextual and temporal dimensions that we know are important for autobiographical memory [Harris and Morris 1984] which would also seem to be relevant for accessing personal data. It should be possible to retrieve documents by remembering their context (e.g., documents worked on concurrently with the target document). This suggests that for digital personal information we should explore novel contextual and temporal representations of information [Fertig et al. 1996; Gifford et al. 1991; Lansdale and Edmonds 1992].

Our results suggest an additional interpretation of personal information management problems reported elsewhere [Cole 1981; Lansdale 1988; Malone 1983]. Those studies offer *cognitive* and *attentional* explanations, arguing that inconsistent filing schemes arise from the cognitive difficulty of deciding how a given document should be categorized, and that piled documents are a *reminding* technique for keeping available information about current tasks. While these factors play a role, our findings suggest that *information overload* is also important. While piles do serve some of the reminding and conceptual scaffolding functions documented elsewhere, our study highlights other reasons for their existence: (a) insufficient time to process all incoming documents, and (b) deferred evaluation. For filers, the cognitive difficulty of categorizing documents is compounded by *overload* which means they may not have read the documents they are filing. The problems people experience in organizing and maintaining their personal archives partially arise, therefore, from a lack of time and attentional capacity for processing the increasing amounts of information they receive. In the digital domain we desperately need tools to help filter and organize this incoming deluge.

REFERENCES

- BARREAU, AND NARDI, B. 1995. Finding and reminding: Organization of information from the desktop. *SIGCHI Bulletin* 27, 39–45.
- BÄLTER, O. 2000. Finding and reminding: Organization of information from the desktop. *SIGCHI Bulletin* 27, 39–45.
- BOWERS, J. 1994. The work to make a network work. In *Proceedings of the Conference on Computer Supported Cooperative Work*. ACM Press, New York, 287–298.
- CARD, S., MACKINLAY, S., AND SHNEIDERMAN, B. 1999. *Readings in Information Visualization*. Morgan Kaufmann, San Francisco.
- COLE, I. 1981. Human aspects of office filing: Implications for the electronic office. In *Proceedings of Human Factors* (Seattle, Wa.).
- COMMUNICATIONS OF THE ACM. 1997. Special issue on Recommender Systems 40, 3 (March), P. Resnick and H. R. Varian, guest editors.
- FERTIG, S., FREEMAN, E., AND GELERTNER, D. 1996. Finding and reminding reconsidered. *SIGCHI Bulletin* 28, 1.
- FRANCIK, E. 1996. Rapid integrated design of a multimedia communication system. In *Human Computer Interface Design*, M. Rudishill, Ed. Morgan Kaufmann, San Francisco.
- FROHLICH, D. AND PERRY, M. 1994. The Paperful Office Paradox. Hewlett Packard Laboratories Tech. Rep. HPL-94-20.
- GIFFORD, D., JOUVELOT, O., SHELDON, J., AND O'TOOLE, J. 1991. Semantic files systems. In *Proceedings of Thirteenth ACM Symposium on Operating Systems Principles*. ACM Press, New York.
- GOLDBERG, D., NICHOLS, D., OKI, B. M., AND TERRY, D. 1992. Using collaborative filtering to weave an information tapestry. *Communications of the ACM* 35, 12 (December), 51–60.
- HARRIS, J. E. AND MORRIS, P. E., Eds. 1984. *Everyday Memory, Actions, and Absent-Mindedness*. Academic Press, London.
- HEATH, C. AND LUFF, P. 1996. Documents and professional practice: 'Bad' organisational reasons for 'good' clinical practice. In *Proceedings of Conference on Computer Supported Cooperative Work*. ACM Press, New York, 354–363.
- HENDERSON, D. AND CARD, S. 1986. Rooms: The use of multiple virtual workspaces to reduce space contention in a window based graphical user interface. In *ACM Transactions on Graphics* 5, 211–243.
- HILL, W. C., STEAD, L., ROSENSTEIN, M. AND FURNAS, G. 1995. Recommending and evaluating choices in a virtual community of use. In *Proceedings of CHI'95 Conference on Human Factors in Computing Systems*. ACM Press, New York, 194–201.

- JOHNSON, A. 1997. Fax trends. *Office Equipment and Products* 25.
- JOHNSON, W., CARD, S., JELLINEK, H., KLOTZ, L., AND RAO, R. 1993. Bridging the paper and electronic worlds: The paper user interface. In *Proceedings of CHI'93 Human Factors in Computing Systems*. New York: ACM Press, 507–512.
- KIDD, A. 1994. The marks are on the knowledge worker. In *Proceedings of CHI'94 Human Factors in Computing Systems*. New York: ACM Press, 186–191.
- LANSDALE, M. 1988. The psychology of personal information management. *Applied Ergonomics* 19, 55–66.
- LANSDALE, M. AND EDMONDS, E. 1992. Using memory for events in the design of personal filing systems. *International Journal of Man-Machine Studies* 36, 97–126.
- LUFF, P., HEATH, C., AND GREATBATCH, D. 1992. Tasks-in-interaction. Paper and screen based documentation in collaborative activity. In *Proceedings of Conference on Computer Supported Cooperative Work*. ACM Press, New York, 163–170.
- MACKAY, W. 1988. More than just a communication system: Diversity in the use of electronic mail. In *Proceedings of Conference on Computer Supported Cooperative Work*. ACM Press, New York, 344–353.
- MALONE, T. 1983. How do people organize their desktops? Implications for the design of office systems. *ACM Transactions on Office Information Systems* 1, 99–112.
- MANDER, R., SALOMON, G., AND WONG, Y. 1992. A “pile” metaphor for supporting casual organization of information. In *Proceedings of CHI'92 Human Factors in Computing Systems*. ACM Press, New York, 627–634.
- MACKAY, W. AND PAGANI, D. 1994. Video Mosaic: Laying out time in a physical space. In *Proceedings of ACM Multimedia*. ACM Press, New York, 165–172.
- NEWMAN, W. AND WELLNER, P. 1992. A desk supporting computer-based interaction with paper documents. In *Proceedings of CHI'92 Human Factors in Computing Systems*. ACM Press, New York, 587–592.
- O'HARA, K. AND SELLEN, A. 1997. A comparison of reading paper and online documents. In *Proceedings of CHI'97 Human Factors in Computing Systems*. 335–342.
- RHODES, B. AND STARNER, T. 1996. The Remembrance Agent: A continuously running automated information retrieval system. In *Proceedings of The First International Conference on the Practical Application of Intelligent Agents and Multi Agent Technology (PAAM'96)*. 487–495.
- RESNICK, P., IACOVOU, N., SUCHAK, M., BERGSTROM, P., AND RIEDL, J. 1994. GroupLens: An open architecture for collaborative filtering of netnews. In *Proceedings of Conference on Computer Supported Cooperative Work*. ACM Press, New York, 175–186.
- SCHILIT, W., GOLOVCHINSKY, G., AND PRICE, M. 1998. Beyond paper: Supporting active reading with free form digital ink annotations. In *Proceedings of CHI'98 Conference on Computer Human Interaction*. ACM Press, New York, 249–256.
- SELLEN, A. AND HARPER, R. 1997. Paper as an analytic resource in the design of new technologies. In *Proceedings of CHI'97 Conference on Computer Human Interaction*. ACM Press, New York, 319–326.
- SHAPIRO, D., HUGHES, J., RANDALL, D., AND HARPER, R. 1991. Visual representation of the flight data strip in air traffic control. *Proceedings of the 10th Interdisciplinary Workshop on Informatics and Psychology*.
- SHARDANAND, U. AND MAES, P. 1995. Social information filtering: Algorithms for automating “Word of Mouth.” In *Proceedings of CHI'95 Conference on Computer Human Interaction*. ACM Press, New York, 210–217.
- SHNEIDERMAN, B. 1982. The future of interactive systems and the emergence of direct manipulation. *Behavior and Information technology* 1, 237–256.
- TRIGG, R., BLOMBERG, J., AND SUCHMAN, L. 1999. Moving document collections online: The evolution of a shared repository. In *Proceedings of ECSCW99 European Conference on Computer Supported Cooperative Work*. Elsevier, Amsterdam, 331–350.
- TUSHMAN, M. 1977. Special boundary roles in the innovation process. *Administrative Science Quarterly* 22, 587–605.
- WHITTAKER, S. AND SCHWARZ, H. 1995. Back to the future: Pen and paper technology supports complex group co-ordination. In *Proceedings of CHI'95 Conference on Computer Human Interaction*. ACM Press, New York, 495–503.

- WHITTAKER, S. AND SCHWARZ, H. 1999. Board meetings: The impact of scheduling medium on long term group coordination in software development. In *Computer Supported Cooperative Work 8*, 175–205.
- WHITTAKER, S. AND SIDNER, C. 1996. Email overload: Exploring personal management of email. In *Proceedings of CHI96 Conference on Computer Human Interaction*, ACM Press, New York, 276–283.
- WHITTAKER, S., DAVIES, R., HIRSCHBERG, J., AND MULLER, U. 2000. Jotmail: A voicemail interface that enables you to see what was said. In *Proceedings of CHI2000 Conference on Human Computer Interaction*. 89–96.
- WHITTAKER, S., FROHLICH, AND DALY-JONES, O. 1994. Informal workplace communication: What is it like and how might we support it? In *Proceedings of CHI94 Human Factors in Computing Systems*. ACM Press, New York, 130–137.
- WHITTAKER, S., HIRSCHBERG, J., AND NAKATANI, C. H. 1998. All talk and all action: Strategies for managing voicemail messages. In *Proceedings of CHI98 Conference on Computer Human Interaction*. ACM Press, New York.
- WHITTAKER, S., SWANSON, J., KUCAN, J., AND SIDNER, C. 1997. TeleNotes: Managing lightweight interactions in the desktop. *ACM Transactions on Computer Human Interaction 4*, 137–168.
- WOLF, C., RHYNE, J., AND ELLOZY, H. 1989. The paper-like interface. In *Designing and Using Computer Interfaces and Knowledge Based Systems*, G. Salvendy and M. Smith, Eds. Elsevier, Amsterdam.