10 Everything Through Email

Steve Whittaker, Victoria Bellotti, and Jacek Gwizdka

10.1 Introduction

Email is one of the most successful and frequently used computer applications. According to marketing surveys, over 4 billion corporate email messages were exchanged per day in 2001, increasing to a projected 35 billion in 2005. Moreover, 97 percent of workers report using email multiple times per week for a daily average of 49 minutes (Gartner 2001; Levitt 2000; Pitney Bowes 2000). For many people, work is interpersonal rather than solitary, and email is the main conduit through which their work and information are distributed. They tend to "live" in email, as evidenced by the sheer amount of time they spend using it (Bellotti, Ducheneaut, Howard, et al. 2005), and by their evaluation of its importance, with 71 percent of people stating that it is "essential" for their everyday work (Pitney Bowes 2000).

Email serves as an information conduit—acting as a delivery channel for different types of information, including documents, slides, contact information, and schedules. Its use as conduit naturally leads to its being used for key PIM functions. People use their inboxes as to-do lists to manage current tasks, their email folders as a repository for archival information, and their email address books to find contacts. They even use email to find and schedule calendar appointments.

Despite its success, there are significant problems with email. Users complain about feeling overwhelmed by the volume of messages they receive, and are concerned about processing incoming messages effectively (Bälter & Sidner 2002; Bellotti et al. 2005; Venolia, Gupta, Cadiz & Dabbish 2001; Whittaker & Sidner 1996). They have difficulties, too, in organizing and managing archives (Bälter & Sidner 2002; Boardman & Sasse 2004; Whittaker & Sidner 1996). They experience severe problems using email to manage tasks, leading them to forget tasks and obligations (Bellotti, Ducheneaut, Howard & Smith 2003, Bellotti et al. 2005; Whittaker 2005; Whittaker, Jones & Terveen 2002). Email's failure to address these PIM problems threatens to seriously reduce individual and corporate productivity.

Why is email so hard to process and manage and why does it generate such serious personal information management problems? And yet why is it still so popular and so essential in PIM? And what can be done to make it better able to support this essential role? This chapter will illustrate how many of email's PIM problems stem from its conduit function. Email is incessant, because it is often the primary delivery channel for work and information (Ducheneaut & Bellotti 2001; Venolia et al. 2001; Whittaker 2005). It is hard to process and organize because it is a mixture of different types of information (tasks, documents, FYIs, meeting scheduling), some of which are important (work tasks) and others unimportant (jokes). And most email is generated by others—making it harder to understand, evaluate, and organize than personally generated information (Boardman & Sasse 2004). These problems are exacerbated by the fact that most email systems have no inbuilt support for PIM aside from folders, so that users have to devise ad hoc ways to manage tasks, find contacts, and organize useful information (Bellotti et al. 2003, 2005; Whittaker 2005; Whittaker & Sidner 1996).

This chapter discusses how people use email to manage their personal information, and describes the tools that can support such behavior. We look at important aspects of PIM through the lens of email: allocating attention, deciding actions, managing tasks, and organizing messages into folders.

10.2 Email Activities and Their Relation to Finding, Management, and Keeping Aspects of PIM

Figure 10.2 shows major email activities and the relations between them. We now describe these activities and identify the PIM problems that each creates. In each case we discuss the extent to which these problems overlap with problems occurring in other areas of PIM.

1. **Allocating attention.** Email messages differ in their importance, urgency, and interest (Dabbish, Kraut, Fussell & Kiesler 2005). They also take differing amounts of time and effort to process. As a result, users do not access email messages sequentially, instead deciding which messages they attend to first (Bälter & Sidner 2002; Bellotti et al. 2005; Venolia et al. 2001).
2. **Deciding actions.** Unlike other information items obtained from the Web or document repositories, email messages may communicate expectations from others of some kind of response. Users have to assess whether this is the case and decide what action to take for a given message. This is similar to, but more complex than, the classic PIM "keeping" decision

(Bellotti et al. 2005; Dabbish et al. 2005; Venolia et al. 2001).

3. **Managing tasks.** Many email tasks are not one-shot and cannot be discharged in a single session (Bellotti et al. 2005; Gwizdka & Chignell 2004; Whittaker 2005). Often this is because they require input from others, or the complex collation of information from multiple sources. We refer to collaborative tasks that require inputs from others as interdependent. Task management relates to re-finding activities. Users need to structure personal information to guarantee it can be found easily and at the appropriate time.

4. **Organizing messages and folders.** As in other aspects of PIM, users have to decide how to organize information. This relates to keeping and management activities. If users decide to file, which folder should they put an item in? Again these decisions are complicated by the fact that much email information is generated by others, sometimes without the context of personal user goals or initiatives. This in turn means that folder organization is often unsuccessful (Bälter 1998; Whittaker & Sidner 1996).

Note that the above list does not imply a strict temporal sequence, with messages first being attended to, decisions made, tasks managed and then filed. As shown in Figure 10.2, email processing can take different forms depending on the nature of the message, its relation to other tasks, and user workload. Simple irrelevant messages may be scanned and deleted immediately. Other messages

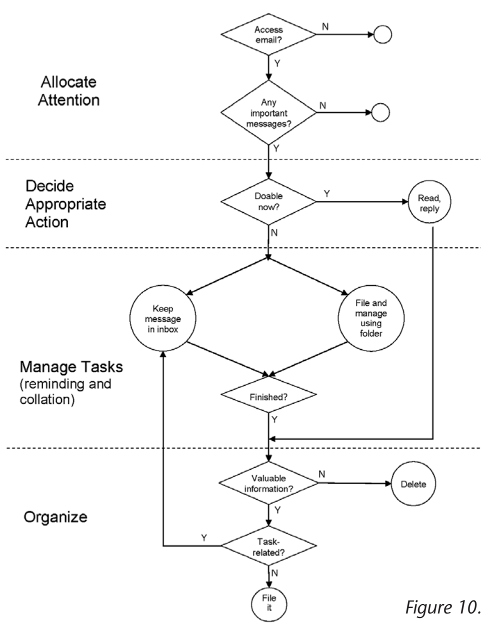

Figure 10.2 – Email processing.

may be left in the inbox, returned to several times as the user scans the inbox, later judged to be irrelevant and deleted. Yet other messages may be scanned and filed if they relate to ongoing projects or interests but do not demand a response. Messages related to interdependent tasks may be filed or left in the inbox as reminders; they may be repeatedly revisited as the task unfolds, or as the user encounters them in the context of accessing their inbox. And older archived messages may be re-accessed if they have relevance for current tasks.

10.3 Understanding Email Tasks

We now review research of special relevance to email's role in PIM. To understand the challenges of managing email, it is important first to document its general properties. We present some general statistics on email volume, and inbox and folder characteristics, followed by a taxonomy of the different types of email message, and how people spend their time processing email.

10.3.1 *Scoping the Problem: Some Basic Email Properties*

Data	Number of Msgs Received/Day	Number of Msgs Sent/Day	Number of Inbox Msgs	Percentage of Total Msgs Kept in Inbox	Total Number of Email Msgs Retained	Number of Folders
#Participants	86	200	206	183	183	244
Weighted Average	43.7	13.7	1424	49.3	2846	38.8

Table 10.3.1 – Summary of Basic Email Statistics. Sources: Balter & Sidner 2002, Bellotti, Ducheneaut, Howard, et al. 2005, Boradman & Sasse 2004, Dabbish et al. 2005, Ducheneaut & Bellotti 2001, McKay 1988, Venolia et al. 2001, Whittaker & Sidner 1996.

Table 10.3.1 summarizes eight studies that have collected general data about email behaviors. One important proviso in interpreting the data is that each of the reported studies reported large individual differences between users, depending on their email processing strategies—and there are large standard deviations for the data presented. Individual differences in PIM are discussed in detail in chapter 12. Note also that these studies have focused on workplace rather than recreational usage, and that the specific context of the workplace probably influences both the volume of emails and the manner in which they are processed.

Somewhat surprisingly, these studies do not suggest a huge increase in email loads over the last 18 years. This may be because these studies researched populations working in high technology or academia where email throughput has always been high, and is perhaps closer to a natural saturation point.

Nevertheless, certain general observations may explain users' complaints about processing and managing email. For example, people have to process a large number (between 38 and 52) of messages each day. They send 31 percent as many emails as they receive—suggesting many messages they receive do not demand a response. On the management side, they keep hundreds or even thousands of messages in their inbox, accounting for about half the total messages they keep. With the proviso that there are large individual differences in how people organize information, these statistics can be explained by the use of the inbox as a "to-do" list for working tasks. On the organizational side, users create on average 38.8 folders.

Message Types

Various attempts have been made to categorize different types of messages that people receive. For example Whittaker and Sidner (1996) describe messages as "to-dos," "to-reads," "messages of indeterminate status," and "ongoing correspondence," based on properties such as message age, length, and whether or not messages are filed or left in the inbox. They also present some prevalence statistics, with "to-reads" accounting for 21 percent of the inbox. Dabbish et al. (2005) present a more systematic analysis of message types. Messages could serve multiple functions, but the overall breakdown was Action requests (34 percent), Information requests (18 percent), Information attachments (36 percent), Status updates for projects/tasks (21 percent), Scheduling requests/responses to scheduling requests (14 percent), Reminders for meetings/actions (16 percent), Social messages (8 percent), and Other (12 percent).

These data clearly show that most messages require the user *to act in some way that has implications for others*—whether this is to carry out an action, respond to a request, or to schedule a meeting. Such interdependence has implications for how effectively others carry out their tasks. Section 10.3.4 details the problems email presents in managing tasks.

Finally, there have been video studies of how people allocate time processing email (Bellotti et al. 2005):

- 54 percent composing messages
- 23 percent reading messages, attachments, and links
- 10 percent filing messages
- 6 percent scanning inbox messages for things to read or deal with
- 2 percent deleting messages
- 2 percent looking for messages in folders
- 2 percent managing attachments

In addition to the obvious dominant activities of reading and writing messages, users dedicate proportionally large amounts of time filing messages. In Section 0 we examine why filing is such a demanding activity and evaluate its success in creating useful structures for future retrieval. Section 0 and 0 discuss how new emails are processed, and Section 0 explains how the inbox is used for task management. We now review research on these key email activities.

10.3.2 *Allocating Attention*

There are two parts to attention allocation. The first involves *when and why* people decide to attend to email, and the second *how* they decide which messages to attend to, and in what order. Allocation of attention is influenced by multiple factors including user workload, available time, nature of the message, identity of its sender, and so on. Deciding which emails to attend to is similar to typical PIM finding tasks (Barreau & Nardi 1995; Lansdale 1988a), as it involves judging items of varying importance to determine which are immediately relevant. It differs from typical PIM finding because items do not have to be actively generated by the user; as they are already in the inbox and do not have to be searched or actively browsed for.

Deciding when to attend to email. Despite email's importance for task delegation, there has been relatively little systematic study of what causes people to attend to their email. Mackay (1988), and Whittaker & Sidner (1996) note how email dominates people's work lives. Similar findings led Ducheneaut & Bellotti (2001) to make the observation that people "live" in email—leading them to characterize it as a "habitat." In Mackay's (1988) study, 13/27 informants stated that they read email "constantly." However as Venolia et al. (2001) observe, people cannot spend their entire work time in email, raising questions about how often they access email, as well as the factors that lead them to switch to email from other activities. Dabbish et al. (2005) report that people check email 19 times per day, so what prompts this? For some, any incoming unread message serves as a trigger. Venolia et al. present survey data from 406 Microsoft employees suggesting that most of their respondents attend to incoming messages as soon as these arrive. Participants also usually access email as soon as they can after they have been away from it, that is, first thing in the morning, or upon returning to their desk after an absence. And because email serves as a task manager, archive, and contact manager, email access may be prompted by the need to retrieve to-dos, reminders, or archived information relevant to a current task (Ducheneaut & Bellotti 2001; Whittaker & Sidner 1996). Access frequency is directly influenced by task management demands; Gwizdka (2004b) reports that the more task-related messages users keep in the inbox the more likely they are to constantly read email, rather then at specific times only.

Deciding which messages to attend to. Users generally don't access messages sequentially in the order in which they arrive, and they often don't access all messages. Bälter & Sidner (2002) observed that their ten users scanned the inbox a mean of 2.3 times to decide which messages to focus on, and Venolia et al. (2001) report that 70 percent of their interviewees processed messages out of sequence. Bälter and Sidner also found that only half their informants accessed all the emails they received. But how do people decide which messages to access first and which to ignore? One possibility is that people focus on *important messages*, and Venolia et al. list a variety of factors that affect perceived importance, such as the identity of the sender, number of recipients, the nature of the message, and whether the message is a reply. Dabbish et al. (2005) found that important messages were those that demanded action, conveyed task status, discussed schedules, or were reminders. Important messages were also more likely to be sent by people with whom the recipient had a strong working relationship, and to be personally addressed rather than to a wide distribution list. They were also highly unlikely to be social messages.

10.3.3 *Deciding Actions*

Having decided *which* messages to attend to, users then have to decide *what* to do about them—for example, whether to respond, file, delete, or defer action until later. In other areas of PIM, users have to make decisions about the information that they encounter, such as whether to keep it or not (Jones 2004). In email, these decisions are more complex, and they may also have important implications for the work of others; that is, when the information concerns an interdependent task.

Dabbish et al. (2005) found that people only respond (or intend to respond, in cases where they defer action) to about 35 percent of the messages they receive. Contrary to their expectations, Dabbish and colleagues also discovered that perceived importance is only weakly correlated with responding; people are only 8 percent more likely to respond to important messages. In contrast, sender characteristics (15 percent), whether the message is personally addressed (20 percent) and, most strikingly, whether the message is social (23 percent) are all stronger determinants than importance in predicting responding. Another notable strategy is that users *defer* responses to 37 percent of messages they intend to respond to. There are several possible reasons for this, including leaving time for reflection, or, with interdependent tasks, communicating with others to gather and collate the information needed for a collective response. Whether people respond to a message immediately or not, their most likely action for a message requiring a response is to leave it in the inbox. Even when they have already replied to the message, users leave it in the inbox 55 percent of the time, whereas if the message has not been responded to, they leave it in the inbox 79 percent of the time. This suggests that people are using the inbox as a reminder to respond—using the inbox for task management as described in the next section.

Dabbish et al. also examined whether incoming messages were filed, deleted, or retained in the inbox. Overall, people filed 27 percent of their messages, deleted 24 percent, and kept the remaining 49 percent in the inbox. People were more likely to keep important messages in the inbox, although message type and sender characteristics did not directly affect retention behaviors. Individual differences were important here—accounting for 48 percent of the variance in the probability of a message being left in the inbox.

Of the messages that don't demand a reply, users file 28 percent of (presumably useful) messages, delete a further 30 percent of (presumably irrelevant) messages, but then retain the remaining 42 percent in the inbox. Why do users retain so many of these informational messages that don't demand a reply? Some messages may be kept for later reading. Immediate time pressure may lead people to defer reading these until they have more time. Whittaker and Sidner (1996) found that 21 percent of inbox messages contained more than 5 screenfuls of text. These to-reads may not be filed because leaving them in the inbox is a way to remind users that they still have to be processed. Other messages are retained because users don't know what to do with them. Rather than investing valuable time in reading a new message at once, users register its arrival, but defer dealing with it until they are more certain of its importance. Rather than delete it immediately, they conservatively retain it, in case it turns out to be important.

Messages are often revisited multiple times, with different decisions taken about how they should be processed. For example, a message might be kept in the inbox or stored in a folder and accessed multiple times until a task is complete, when it might be filed in an archive (Bellotti et al. 2005).

One obvious conclusion is that deciding what to do with email messages is rather more complex than the typical PIM keeping decision. While users do end up keeping the majority of messages, their decisions are more complicated than making a binary choice to keep or delete a message. Having decided to keep a message, users also have to decide what action to carry out on it, such as whether and when to reply to it. There also seem to be two distinct and important types of keeping activity: (1) filing—which is the typical PIM task of categorizing information for future use, and (2) inbox retention—where information is not organized, but kept in a working buffer to remind people it may require some future action. The next section describes this complex management activity.

10.3.4 *Managing Tasks*

A classic PIM activity is that of managing information to meet the demands of specific tasks, and many PIM studies have described how users categorize information in order to facilitate its future retrieval (Jones 2004; Kidd 1994; Lansdale 1988a; Teevan, Alvarado, Ackerman & Karger 2004). Task management in email turns out to have rather different demands however.

Task management involves reminding oneself about current tasks and tracking task status, as well as collating and maintaining information relevant to those tasks. A direct consequence of email functioning as a conduit for work and information is that people use it as a de facto task manager. Many email tasks are too complex or lengthy be executed in one shot (Bellotti et al. 2005; Venolia at al. 2001; Whittaker 2005; Whittaker & Sidner 1996), leading to deferral for 37 percent of messages that require replies (Dabbish et al. 2005).

Deferral is often a direct consequence of *interdependent* tasks, such as those involving others (Bellotti et al. 2005). Two problematic aspects of managing interdependent tasks result from *iteration* and *delays* between messages relating to the task. Iteration arises because interdependent tasks often require multiple exchanges between participants before they can be resolved (Bellotti et al. 2005; Venolia & Neustaedter 2003; Venolia et al. 2001; Whittaker & Sidner 1996). People may need to negotiate exactly what a task involves, or multiple responses may need to be synthesized. Delays occur because collaborators lack the necessary information to respond immediately to address their part of the task. One way to estimate the prevalence of interdependent tasks is by determining how many emails are part of a *conversational thread*, as threads indicate relations and common underlying activities among messages. Threading estimates range from 30 to 62 percent of messages (Bellotti et al. 2003; Fisher & Moody 2001; Klimt & Yang 2004; Venolia et al. 2001).[1] The largest available public email corpus from Enron has 62 percent of messages appearing in threads, with a mean thread length of 4.1 messages, though the median length is only 2 (Klimt & Yang 2004).

Delays and iteration give rise to two critical activities in managing interdependent tasks. Users have to leave themselves reminders about outstanding tasks, and to collate different elements of the same collaborative task. We now look at how people address these problems in order to carry out core task management activities.

People predominantly use email itself to manage tasks arriving in email, rather than relying on a dedicated task-management application, such as a to-do list on a PDA (Bellotti, Dalal, Good, et al. 2004; Venolia et al. 2001). New messages relating to a particular task arrive in email, which, as email is constantly processed, serve to remind users about that task, whereas copying information into a separate application requires additional effort and bookkeeping.

One *reminding* strategy involves setting up a dedicated to-do folder containing reminders about outstanding tasks (Bellotti et al. 2003, 2005; Whittaker & Sidner 1996; Whittaker et al. 2002). However, Whittaker & Sidner (1996) report that 95 percent of users abandoned this strategy because it requires an additional cognitive step. Instead of being reminded about tasks as they access new email in their inbox, people have to explicitly remember to open the to-do folder, and they often forgot to do this. Another strategy is to organize emails relating to current tasks into active folders, returning to these folders when they need to deal with those tasks (Bellotti et al. 2005; Gwizdka 2004b). This has the advantage of grouping messages, allowing them to be more efficiently worked on together. But this strategy only works if users develop the habit of routinely returning to inspect those folders, as most users do with the inbox.

But the most common strategy for managing interdependent tasks is to respond or forward the original message to relevant others, leaving the original message in the inbox as a reminder about the task (Dabbish et al. 2005; Mackay 1988; Venolia et al. 2001; Whittaker 2005; Whittaker & Sidner 1996). This serves to manipulate attention. Users know that they will return to the inbox to access new messages. In the course of doing so, they hope that they will see the reminder and recall the outstanding task. Users may even send themselves emails as reminders, or that contain important information (Jones, Dumais & Bruce 2002). This inbox reminding strategy is far more common than other techniques such as using message flags, classifying messages as to-do items, relying on external applications, or external reminders (Gwizdka 2004b; Venolia et al. 2001).

A second problem with interdependent tasks is *collating* information, or assimilating the disparate pieces of information relating to a common task. Collation is difficult because messages relating to the same activity may come from multiple individuals using different subject lines, and may even end up in different folders. As a result, some users feel that that keeping messages in their inbox makes them more *collatable*, in part because the multiple passes carried out when processing the inbox serve to remind them about its contents (Whittaker 2005).

However, many users experience severe problems with using the inbox for task management. Reminding and collating are highly dependent on *visual scanning*, which becomes less effective as the number of inbox items increases. Most email interfaces present a header list along with message body information for a selected message. Although users differ in how they configure their header list, they tend to operate with 10 to 20 message headers visible on the screen (Whittaker & Sidner 1996; Whittaker et al. 2002). This means, on average, that less than 2 percent of inbox messages are visible.

Because most inbox message headers are not constantly in view, users have to actively remember to scroll back to view older messages. And it is hard to identify specific, related messages scattered at intervals through the inbox among hundreds of other unrelated items. The alternatives of searching and sorting by metadata, such as sender or subject, are only effective when users can remember salient information about the message. For this reason, older inbox messages are often characterized as being "out of sight and out of mind" (Whittaker & Sidner 1996).

And relying on folders or messages threads is little better: Venolia et al. (2001) found that 23 percent of message threads occurred in more than one folder. Using threads is also unreliable, because they rely on a common message subject line—and this turns out to be an inconsistent indicator that messages relate to the same task (Ducheneaut & Bellotti 2003).

A related problem is *obligation management*, or keeping track of the tasks that you "owe" others or that are "owed" to you. Although the strategy of leaving messages in the inbox works reasonably well for reminding about tasks that one receives, it is much less effective when delegating tasks to others. Most email programs keep a copy of sent messages, but these are generally stored in a separate "Sent" folder. Having to remember to access the Sent folder suffers from the same weaknesses as using a to-do folder. Some users attempt to avoid this by copying themselves on every sent message to generate inbox reminders, but this overloads the inbox with duplicate messages for every task (Bellotti et al. 2005; Whittaker 2005; Whittaker et al. 2002).

In conclusion, task management is problematic in email. Although the default strategy is to use the inbox as a to-do list, there are numerous problems associated with doing this, including overlooking reminder messages and inability to collate information. These problems do not have clear analogues elsewhere in PIM.

10.3 Understanding Email Tasks

People dedicate a great deal of time and effort to organizing their email. Many PIM studies have documented the problems of categorizing information in a way that facilitates later retrieval (Barreau & Nardi 1995; Boardman & Sasse 2004; Jones et al. 2002; Lansdale 1988a; Whittaker & Hirschberg 2001). Email presents similar organizational problems.

Creating email folders is a common activity, with users overall filing 27 percent of their messages, although there are large individual differences in the complexity of folder organization. Some users create organizations that include hundreds of folders with complex internal structure; others create few folders (Boardman & Sasse 2004; Venolia et al. 2001; Whittaker & Sidner 1996). One important reason for filing is that messages contain useful reference information such as documents, presentations, or Web links (Bellotti et al. 2005; Ducheneaut & Bellotti 2001; Venolia et al. 2001). For example, Dabbish et al. (2005) estimate that about 36 percent of incoming messages contain useful reference information. Indeed some people tend to use email as a document archive, sometimes using metadata such as sender and date to organize and version different document drafts (Bellotti & Smith 2000). And contrary to Barreau and Nardi's (1995) assertions, users do access old messages and value their archives (Boardman & Sasse 2004; Venolia et al. 2001).

Various studies have examined the *types* of folders that email users create. The most common types of folders relate to work projects, contacts, and mailing lists (Boardman & Sasse 2004). There have been fewer studies of the process of creating folders however, although Whittaker and Sidner (1996) examined the problems users experienced with this, and we summarize their findings here.

Filing is a cognitively difficult task (Lansdale 1988a). Successful filing is highly dependent on being able to anticipate future retrieval requirements. It is hard to decide which existing folder is appropriate, or, if a new folder is needed, how to give it an appropriate and memorable name.

Users may not file messages because failing to remember where information has been filed could be disastrous. And we have already seen that another reason for not filing is to postpone judgments, in order to determine the value of information. Users want to avoid archiving information that later turns out to be useless or irrelevant.

Even when users do decide to file, folders may not turn out to be especially useful. Users may be unable to remember folder names, especially after a time has elapsed, or when they have large numbers of folders. They have to remember the definition of each when filing, and to be careful not to create new folders that duplicate or overlap with existing ones. Duplication detracts from their use in retrieval.

In addition, folders can be *too small* to be useful. A major aim of filing is to collate the huge number of undifferentiated inbox items into a relatively small set of folders containing multiple related messages. Filing is clearly not successful if each folder is small: a folder containing few items has not significantly reduced the complexity of the inbox, nor collated significant amounts of related material. However, Whittaker and Sidner (1996) show that filing often fails: 35 percent of users' folders contain only one or two items. Furthermore, the user has the dual overheads of (1) first creating these folders, and (2) remembering multiple definitions every time there is a new filing decision. Quantitative data illustrate the problems of trying to remember multiple folder definitions. The more folders a user has, the more likely they are to generate "failed folders" containing only one or two items.

Folders can also fail because they are *too big*. When there are too many messages in a folder, it becomes unwieldy. It is difficult to find relevant messages in a large folder, as the relationships between different messages in the folder become tenuous, and one of the main benefits of keeping them together is much reduced.

There has been less study of how users access their email archives. Some studies claim that users report few difficulties in locating information in their archives. Boardman and Sasse (2004) found otherwise: users experienced considerable problems when asked to find archived messages. The most common strategies for accessing archival information are to browse folders or to sort messages by sender or subject. Search is much less commonly used, although this may reflect the inefficiency of search in many email applications, and recent developments in desktop search should change this (Boardman & Sasse 2004; Dumais, Cutrell, Cadiz, et al. 2003; Venolia et al. 2001). Dumais et al. (2003) looked at common patterns of search over archives, finding that recency was a very strong predictor that a message would be accessed, and that access by sender or recipient name was another common search strategy.

In summary, various studies have shown that, in common with other areas of PIM, email users experience problems in organizing their information. They spend large amounts of time constructing structures for future retrieval that may not be greatly effective in supporting access.

10.4 Techniques to Support PIM in Email

Having characterized users' main email activities and PIM problems, we now discuss how current technologies address these issues. From our prior discussion it is clear that there are strong interrelations between different email activities. Thus, a particular technology such as automatic filing may help both with organization of email, and with task management; for example, by reducing inbox size it allows users to better focus on outstanding messages. And techniques that analyze message content and headers might help users to allocate attention to important messages as well as to decide what action to apply to a message. Table 10.4 shows a breakdown of different techniques and the PIM activities they support.

Technique		Email Activities Addressed
Spam filtering		Attention Allocation
Personal filtering	sender specified	Attention Allocation, Deciding Actions
	user created	Attention Allocation, Deciding Actions
	machine learning	Attention Allocation, Deciding Actions
	pre-defined rules	Attention Allocation, Deciding Actions
AI message summarization		Deciding Actions
Information structuring		Task Management
Embedded task management support		Task Management
Workflow systems		Task Management
Assisted filing		Archive Management, Task Management
Search		Archive Management

Table 10.4 – Techniques for Supporting Key Email Activities.

10.4.1 *Allocating Attention*

Two main approaches have been taken to help users allocate attention more effectively. Both rely on filtering, but each has a different aim. *Spam filtering* aims to remove irrelevant information from the inbox. *Personal filtering* takes the opposite approach in attempting to identify, highlight, or prioritize important information in the inbox.

Spam filtering. Spam has become a major problem for most email users, with AOL estimating that it blocks 1.4M spam messages each day—an average of 22 per user per day (AOL 2003). Many organizations find it necessary to apply filters at the email server to remove suspect content. Spam detection has been relatively successful (if not yet completely accurate) because spam messages often have distinct properties such as large distribution lists, predictable headers, and somewhat predictable message content (Cranor & LaMacchia 1998). But the race between those who develop spam filters and the spammers evolving techniques to defeat them remains an ongoing contest, with recent spamming techniques involving the generation of more plausible message content.

Personal filtering. Personal filtering relies on various techniques, depending on how filtering rules are created. One technique relies on the sender to specify information that allows the user to better judge message utility. For example Information Lens (Malone, Grant & Turbak 1986) or the Co-ordinator (Flores, Graves, Harfield & Winograd 1988) allow the sender to specify the priority of the message or its projected action. Another technique used by many commercial email clients allows the user to define rules to prioritize incoming messages, using information about the sender, recipients, message content, subject line, and so on. Yet another (predominantly research) approach is to use machine learning to automatically induce prioritization rules. Horvitz, Jacobs and Hovel (1999) and Metral (1993) built systems that track users' actions to learn the properties of messages the user attends to and in what order, using this information to identify new messages that are likely to be important. Finally, systems can provide predefined rules that users can tune to meet their needs. For example, Marx and Schmandt (1996) use information from calendar appointments, outgoing messages, phone calls, and the user's Rolodex to infer message importance. Bälter and Sidner (2002) propose a much simpler set of predefined rules that identify urgent messages, important senders, and personal messages.

Personal filtering has not been successful, however. It is not commonly used despite being prototyped almost 20 years ago, and being available for several years in commercial products (Bälter & Sidner 2002; Bellotti et al. 2003; Whittaker & Sidner 1996; Whittaker 2005). One problem is that writing filtering rules is a programming task—which most users find both hard and time-consuming. And often the users who are most in need of filters have the least time to dedicate to writing rules. Systems that provide predefined general rules simplify the programming task, but this approach has not been successful either. Although machine learning and bespoke filters do not require users to write rules, trust is a more serious concern with such approaches. Users are unconfident that rules will operate correctly (Pazzani 2000). In particular, they worry that misdefined rules will lead important messages to be overlooked (Whittaker & Sidner 1996). And whatever the approach to defining filtering rules, these must be modified frequently to reflect changes in people's work activities—making rule maintenance an important issue (Whittaker & Sidner 1996). In sum, while personal filtering may work for restricted message types with predictable characteristics (e.g., originating from specific newsgroups or users), this technique has limited utility for many messages that users receive.

10.4.2 *Deciding Appropriate Message Actions*

Recent research has developed techniques that suggest actions for new messages by analyzing past user actions on similar messages. As much of this work is at a preliminary stage, we postpone analysis to section 10.5, where we discuss future email trends.

10.4.3 *Managing Tasks*

People have proposed three approaches to task management: (1) *information structuring*, which seeks to impose organization on the inbox by clustering together messages related to the same task, (2) *embedded support for task management*—where techniques are used to recognize and organize task-related information-building in direct support for reminding and collation, and (3) *workflow*, which seeks to organize messages in terms of underlying task sequences and the roles of the people executing the task.

Information Structuring

Most *information structuring* work proposes novel visualizations for inbox message threads, imposing order on the undifferentiated inbox. Thread detection clusters related messages, allowing users to collate task information and work on related messages together.

Various approaches have been taken to visualizing inbox threads. These range from combining thread components linearly (Bellotti et al. 2003), to constructing complex tree structures with subthreads (Venolia & Neustaedter 2003). Most approaches are based on information derived from message subject lines (Rohall, Gruen, Moody & Kellerman 2001; Venolia et al. 2001; Venolia & Neustaedter 2003; Wattenberg, Rohall, Gruen & Kerr 2005). But topic drift and replying practices mean that the subject line is at best a weak indicator of relations between elements of a specific interdependent task (Whittaker 2005). One way to address this is to allow users to manually intervene, by combining different threads or redefining them according to their own view of what is related (Bellotti et al. 2003, 2005). But while these hybrid approaches are promising, most have not as yet been subjected to detailed evaluation, and to be effective they may have to be tuned to an individual user's email or to the user's workgroup.

Embedded Support for Task Management

Bellotti et al.'s (2003, 2005) prototype, TaskMaster, provides embedded support for tasks based around the notion of *thrasks*. TaskMaster is designed to present a task-centric view of email, addressing the problems of *collation* and *reminding*. The thrask model supports collation as follows: any related incoming messages (replies in a thread, along with associated files or links) are automatically grouped together, using message content and metadata. Each thrask is represented in the inbox by its first message's subject line, providing a compact representation that serves a reminding function. Clicking on the subject line reveals the relevant set of messages, allowing the user to work on the task. In a field trial involving nine users, the system performed well, with users finding many of the embedded task-management features compelling.

Whittaker, Swanson, Kucan, and Sidner (1997) also provided direct support for task management in the TeleNotes system. TeleNotes supported interdependent task management by mimicking people's current use of physical desktop piles and spatial location. Messages relating to a common task are organized into piles that appear on the users' computer desktop. This serves not only to remind users about outstanding tasks but also helps with collation, as related materials are clustered together. The system was positively evaluated by eight users in an extended field trial.

A slightly different approach to task management was taken in the ContactMap system (Nardi, Whittaker, Isaacs, et al. 2002; Whittaker, Jones, Nardi, et al. 2004). Instead of organizing the user interface (UI) in terms of messages involved in a task, it organizes in terms of the *people* involved in that task. The UI represents a social network of the important people in the user's work and social life, and messages (and tasks) can be accessed by clicking on individuals or groups involved in that task. It also supports reminders and alerting, and was positively evaluated in laboratory and field trials.

Workflow Systems

Workflow systems assume that organizational tasks have a predictable structure and that this is associated with different work roles (Prinz 1993; Winograd 1994). For example, a purchase order may have to be initiated by an employee, approved by a manager, and processed by the purchasing department before an item is ordered and delivered to the employee. In principle, workflow systems address *interdependent task management*; they support reminding as well as message collation. In practice, however, they have three major limitations: *additional work, coverage*, and *lack of integration*.

Workflow systems are hard to set up, as workflows are often hard to define, and they often introduce *additional work* for senders. One reason users rejected early systems such as Information Lens (Malone et al. 1986) and Co-ordinator (Winograd & Flores 1986) was that these required senders to add additional information to the initiating message (Mackay 1990). The *coverage* problem arises because many interdependent tasks lack the predictability needed for the workflow approach to succeed. Workflow is effective for tasks that have predictable structure—but, as we have seen, most interdependent tasks have an evolving structure and require iterative negotiation for their solution. This makes them inappropriate for workflow tools (Bowers, Button & Sharrock 1995; Dourish, Holmes, MacLean, et al. 1996; Suchman 1993, 1997). And even when tasks are amenable to workflow, there are *integration* issues. With a few notable exceptions (Borenstein 1992; Borenstein & Thyberg 1988), most workflow systems are not well integrated with email clients, so that users have to switch to a separate application—introducing extra cognitive overhead.

10.4.4 *Organizing Messages into Folders: Assisted Filing*

Organization is a major problem for email users, and several agent-based systems attempt to provide *assisted filing* (Boone 1998; Cohen 1996; Mock 2001; Segal & Kephart 1999; Takkinnen & Shahmehri 1998). These systems use machine learning techniques to automatically elicit the defining characteristics of existing folders, based on message headers and content. These definitions can then be used to classify each inbox message, suggesting the folder it best matches. Some of these algorithms have been tested offline on message corpora, showing that they can categorize inbox documents with a reasonable degree of success—achieving 85 percent accuracy (Cohen 1996).

Although these results are promising, none of the systems has been tested in real-usage contexts, and it may be that (as with the filtering techniques described above) users are unwilling to trust systems that automatically file their messages. It has been noted that users feel uncomfortable about automated mail classification, since they may not agree with the system's decision (Pazzani 2000). And if a message is automatically filed they may be unable to relocate it. MailCat (Segal & Kephart 1999) improves classification performance and partially addresses the problem of users not knowing where content has gone by presenting its best three guesses about where to place the message. Users then actively select their best choice. Crawford's i-ems intelligent classification application also allows user intervention if the system suggests inappropriate folders (Crawford, Kay & McCreath 2002).

While these recent designs address issues of user trust, another serious problem is that *assisted filing* may not cover important aspects of filing. Assisted filing classifies inbox messages into *existing* folder categories, whereas a major user filing problem lies in defining *new* folders (Whittaker & Sidner 1996). A related problem concerns individual differences in email handling, where a subset of users either do not file messages at all, or make rudimentary use of folders (Bälter & Sidner 2002; Boardman & Sasse 2004; Whittaker & Sidner 1996). Assisted filing cannot work for them because they have few folders for the system to learn. A final problem concerns messages that are hard to classify in part because they belong in multiple folders. With some exceptions (e.g., NotesMail) most systems force users to classify messages into a single folder.

10.5 Looking Forward: The Future of Email and PIM

We now speculate about the future of email. There are some aspects of email that we do not expect to change in the next few years. These include (1) *list views*, because they are convenient for viewing, archiving, and sorting; (2) *folders and sorting*, because even if the system can help search, people will still need multiple ways to organize and find something; (3) *information overload*, because email continues to be an easy means to distribute one message to many people and ever more collaborative work processes are being moved online; and (4) *prevalence of attachments*, because messages often concern discussion of (and work around) other content. We do anticipate changes in email and PIM for our four activities however.

10.5.1 *Allocating Attention*

One recent development that promises to affect email is the popularity of other communication and publishing applications, such as instant messaging (IM), blogs, Wikis, and the Web. A large proportion of email traffic consists of messages that do not necessarily demand a reply (Dabbish et al. 2005). So, instead of using email to distribute such information, it might be better to publish it in a blog, Wiki, or on a Web page, reducing overall email traffic and allowing users to focus their attention on personal emails that demand a response. Similarly, one might migrate urgent messages out of email into more interactive applications such as IM. TeleNotes (Whittaker et al. 1997) does this by integrating email with IM. It presents IM messages on the desktop with alerting, but the same messages can also be viewed from within email if the user chooses. IBM's Activity Explorer (Muller, Geyer, Brownholtz et al. 2006) is another unified communication application, blurring the boundaries between email chats and shared workspaces, and offering integration of six types of objects: message, chat, file, folder, annotated screen shot, and to-do item. Although these developments are promising, there are significant research questions raised by such melded communications. For example, if information is published rather than emailed, how are people alerted about important new information? Clumsy implementations of alerts (e.g., sending these through email) replicate the original problem by increasing email volumes. And if we have information distributed in different communication applications this may contribute to fragmentation—making it harder for users to collate related information. On a different note, we expect that advances in machine learning and text processing will lead to the development of effective methods to prioritize emails, identifying those that are urgent, and allowing users to better focus on important messages.

10.5.2 *Deciding Actions*

Deciding appropriate message actions is a difficult challenge for automated support, requiring insight into the meaning and importance of email. Given the current state of the art, we hardly want our intelligent email systems deciding to sign us up for distribution lists or to follow sponsored links to purchase products. But recent attention has turned to content-based techniques that identify task-related aspects of messages, so that they can be more easily acted upon by people.

Some of this new work simply focuses on summarizing messages or threads so that it is easier to extract key information (Abu-Hakima, McFarland & Meech 2001; Muresan, Tzoukermann & Klavans 2001; Rambow, Shrestha, Chen & Laurdisen 2004; Tzoukermann, Muresan & Klavans 2001). Other work goes further and attempts to identify the relationship between messages and activities that the user is engaged in. One approach is to recognize messages that are part of a standard automated workflow (such as purchasing a book online) and what steps they represent in such processes (Kushmerik & Lau 2005). Another is to combine analysis of senders' and recipients' addresses and message bodies to identify messages that are related to a common task, making it easier for the user to monitor it (Dredze, Lau & Kushmerick 2006). Yet another is to identify the actions expressed in the text of the email body and extract these for the user's optional addition to a to-do list (Corston-Oliver, Ringger, Gamon & Campbell 2004). More advanced systems suggest possible actions on a message. Horvitz, Jacobs, and Hovel (1999) and Metral (1993) built systems that track the history of user actions to learn which messages the user reads, deletes, responds to, or files—suggesting appropriate actions for new messages based on these inferred rules. And Carvalho and Cohen (2005) use machine learning to identify messages requiring an action, such as requests or commitments.

With continued advances in machine-learning techniques, we expect to see more such systems that assist users in determining appropriate actions on messages. So far, however, this work is somewhat preliminary, and even where there is a functional email client, problems of trust mean that how they act upon the message content still remains at users' discretion. Even as these techniques improve, they will still require both careful integration with existing systems and transparency of operation in order to engender users' trust in their effectiveness. And one important limitation of all this new work is that little of it has been evaluated with real users.

10.5.3 *Task Management*

Task management is currently *the* critical unresolved email problem for users, although there are a number of promising trends here. We expect future research will continue to generate novel visualizations of email, to help provide greater structure to the inbox (Kerr 2003; Viegas, Boyd, Nguyen, et al. 2004; Whittaker et al. 2004). And various novel email clients already provide embedded support for task management, including collation and reminding (Bellotti et al. 2003, 2005; Wattenberg et al. 2005; Whittaker et al. 2004; Whittaker et al. 1997). While only one of these, ReMail (Wattenberg et al. 2005), has been fully integrated into a functioning product, in all cases user trials suggest that they have features that support critical email problems.

What may significantly improve the performance and acceptability of these systems is developments in text processing and machine learning, allowing systems to automatically infer users' tasks (Canny 2004; Dragunov, Dietterich, Johnsrude, et. al. 2005; Dredze et al. 2006). Other similar techniques will improve inbox organization, by collating related information (Khoussainov & Kushmerick 2005; Surendran, Platt & Renshaw 2005; Wan & McKeown 2004). These clustering techniques not only help with task management by collating messages relating to outstanding tasks, but they may also clear the inbox by suggesting new folders. While clustering has been applied elsewhere to document collections and Web pages (Voorhees 1986; Willett 1988; Zamir, Etzioni, Madani & Karp 1996), as yet it has not been fully evaluated on email. And other research directed toward identifying different message functions should help task management, for example by recognizing messages containing obligations, or standard messages that can be dealt with via autoreply (Carvalho & Cohen 2005). Of course these developments give rise to important issues that need to be addressed by careful interaction design. They introduce automated processes into a critical application where the cost of algorithmic error without human oversight is high.

10.5.4 *Organizing Messages into Folders*

Another recent innovation that promises radical change in organizational practice is the emergence of *desktop search*. Dumais et al. (2003) developed Stuff I've Seen, a prototype system that addresses the PIM finding problem by providing search that operates across different applications. This has been emulated by commercial desktop search from both Google and Microsoft. Unlike prior desktop or email search, these systems are fast and operate across entire archives, retrieving emails, documents, presentations, spreadsheets, or other items that relate to the user's query. Desktop search tackles a number of important email problems. It supports *collation* for task management—allowing information items concerning the same task to be retrieved and worked on together. This addresses the significant problem of reuniting originating emails with their attached documents when the documents have been filed in a forgotten location. It should also reduce users' filing burdens by retrieving information regardless of where it has been filed.

While desktop search undoubtedly addresses some email problems, it nevertheless suffers from some important limitations. One significant question concerns its accuracy, and information retrieval researchers have suggested that new techniques will be needed to customize existing search tools so that they can work effectively on private as opposed to public collections (Dumais et al. 2003). Some have speculated that desktop search means that users will no longer file emails, finessing problems of defining and managing folders, but this is unlikely to be the case. Firstly, browsing rather than search remains many users' preferred way to access personal information, even when search is available—in part because of users' familiarity with the structure of their own file systems (Barreau & Nardi 1995; Boardman & Sasse 2004; Teevan et al. 2004). Folders may also be retained because they provide important semantic information that would be hard to replace by search. For example, folder structure may replicate the relations between different subtasks of the project (Jones, Munat & Bruce 2005; Teevan et al. 2004). Most importantly, while search addresses the *finding* problem it does not address other critical aspects of task management, such as reminding oneself that certain tasks remain outstanding. It also cannot help users to determine which messages they need to allocate attention to, or to decide what actions need to be applied to a message.

10.5.5 *Other Long-Term Research Questions*

There are other long-term developments that promise to affect our four main email and PIM activities. For example, alternative attentional models could fundamentally change email usage by imposing payments on sending messages, or rewarding recipients for reading messages (Kraut, Morris, Telang, et al. 2004). Implementing this would surely reduce email volumes, by addressing attention allocation and, more indirectly, helping task management. Other key developments may be at the system level, where people have proposed radical solutions to PIM and email problems that demand changes in application architectures, by integrating all PIM activities into email, or by migrating information out of email into dedicated applications (Whittaker, Bellotti & Gwizdka 2006). There are important trade-offs here. Centralized systems reduce information fragmentation, but may not provide direct support for key PIM functions. Less radical but still important issues concern the question of how we can provide more direct support for the individual differences that are pervasive in PIM and email (see chapter 12). Finally, despite the centrality of email in our working lives, we lack good theoretical models of asynchronous communication and computer-mediated communication (CMC) that might drive new designs, providing new alternatives to our mission-critical but still problematic systems. More work is needed here to address these critical problems.

NOTES

1. These threading figures generally underestimate interdependent tasks. They ignore new interdependent tasks delegated to users who have yet to deal with them, and they do not register tasks forwarded to others. Furthermore, people are often casual when replying: sometimes failing to reply using re. They may occasionally overestimate interdependence; e.g., when someone initiates a new task by replying to an old message from another user (Erickson & Kellogg 2000; Herring 1999).