

HOW CAN A QUANTUM IMPROVEMENT IN PERSONAL AND GROUP INFORMATION MANAGEMENT BE REALIZED?

Roger Tagg and Tamara Beames

School of Computer and Information Science, University of South Australia, Mawson Lakes SA 5095, Australia

Keywords: Personal Information Management, Collaborative Work Support, Information Overload.

Abstract: A number of authors have pointed out that the IT technology used by individuals and groups for general information work support has not advanced very far in the last decade. At the same time the level of information and cognitive overload on individuals has continued to rise. Research has been done in several areas (e.g. text mining and categorization, email threads etc) but a significant improvement in everyday tools has yet to be seen. This paper addresses the current problems, describes a range of issues that need to be addressed, and discusses how it might be possible in the future to advance to a new level of IT support.

1 INTRODUCTION

(Whittaker, Bellotti and Gwizdka 2006) have highlighted a current situation where some software tools, which are primarily used as email clients, are being pressed into service - somewhat beyond what was originally envisioned - as full-function Personal Information Management (PIM) systems. Although products such as Microsoft Outlook do include functions for Task and Calendar management, these functions have yet to be as widely adopted as email. Many users use separate tools for PIM and group information management (GIM) functions such as to-do lists, project management, workflow; they also use separate software for both domain-specific and general office applications. There appear to be few intelligent bridges between these tools.

An often-expressed opinion is that PIM support has not changed that radically in the last ten years. Despite a number of predictions and announcements, the impression is that PIM is still a piecemeal practice. The reasons for this are unclear. One reason could be that the state of the art in recognizing context in free text is still too immature. Another is that the big commercial vendors (e.g. Microsoft, IBM/Lotus) are not prepared to risk a failure in this market.

During the same time, however, the workload pressure on individuals has continued to rise. Globalization and greater competition for business requires increases in both the rate and quality of decision making. The pace of business change forces

more frequent re-learning of processes and tacit knowledge. The increasingly litigious business atmosphere forces organizations to not only have procedures in place, but to have proof that they are being followed. There is greater emphasis on regulatory compliance, quality assurance, and governance. These trends have often led to a culture of "backside covering" - against being blamed for when things go wrong. Economic trends are also forcing organizations to operate with tighter budgets, smaller headcounts - without any extra resources to implement efficiency improvements. For many users, these increasing demands and expectations seem to have outpaced any benefits offered by introducing IT.

This paper aims to review some of the most critical shortcomings in current practice, and to discuss a range of issues, both technical and motivational, involved in overcoming these shortcomings.. Unless these issues are addressed, it seems unlikely that there will be a significant enough advantage to justify users and their organisations changing to a new PIM environment.

Section 2 of this paper describes three of the most serious shortcomings of current PIM support environments. Section 3 analyses a number of dimensions that require addressing in any successful revolution in practice. Section 4 suggests some general conclusions, including how a quantum improvement might be achieved.

2 CRITICAL SHORTCOMINGS WITH CURRENT PRACTICE

The scope of Personal Information Management, as described by Whittaker and others, is quite wide. Messaging is not just email, but also “chat”, SMS, fax – and even phone messages, letters and scribbled notes. Task management is not just self-reminders and emailed tasks, it can include tasks from a workflow, project management or ERP system, or “actions” recorded in the minutes of meetings, which might be stored in shared folders. Personal file management includes many types of files generated by many different application programs; some of which belong to a group in which the user participates. Contact management needs to cover not only people who have email addresses; some contacts are shared among members of a group and the prime responsibility for maintaining them could be in, for example, a customer relations management (CRM) system.

There are also some PIM functions not addressed in the 2006 Whittaker paper. One is the less formal world of “information scraps” (Bernstein et al 2001). Another is “event notification”, where a user may want to be informed when certain conditions arise (Rifkin and Khare 1998). Also missing from PIM support, not for want of effort, are tools that can recognise meaning in text, whether for identifying context or implied tasks.

A more general criticism is that today’s email-based tools are not sufficiently activity-oriented; they give precedence to what someone else wants us to read, than to what we really need to do. This is being addressed by the European DELOS project (Catarci et al 2007) and TV-ACTA (Bellotti et al 2006).

Three particular symptoms of these shortcomings are discussed below.

2.1 Information and Cognitive Overload

Information overload has been widely discussed in the literature, ranging from (Whittaker and Sidner 1996) to, more recently and comprehensively, by (Klausegger et al 2007). In over 40 years in the workplace, the first named author has noticed a relentlessly creeping number of hours one is expected to work per week, and a greater proportion of working time being spent on non-core work and attending to ever-expanding volumes of email.

(Kirsh, 2000) discusses “cognitive overload” as the problem of frequently switching focus from one

job to another. Fragmentation, here defined as not getting a clear run to get things done without interruption, has also been recognized by (Czerwinski, 2006 and Tungare et al, 2006).

The now ubiquitous availability of email itself has sometimes been blamed for information overload. Email misuse, such as spam, excessive bulk emailing,, and use for conversations that would be better done by chat, phone or face-to-face – causes the number of messages to explode. In an attempt to counteract this, practices such as individuals declaring email bankruptcy (Fitzgerald 2004), and of companies setting email-free days (Booher 2007), have arisen.

2.2 Difficulty of Bridging to Other Software and User Interfaces

If a task arrives by email, a user often needs to correlate some data attached to the task with data in other systems, such as a shared database or application system. An example might be that someone asks via email for the status of five students. The data is perhaps in a spreadsheet, a database or the university student records application.

Correlating data about the same “things” is rarely a trivial matter in today’s computers. One has to either use a separate tool to query the spreadsheet or database, or one has to work through an application, often one that only caters for manual interaction through a web form. Wrapping such applications as web services may ease this latter problem. However these solutions tend to be specific to the particular process the user is carrying out.

2.3 Squeezing the Time Remaining to do the “Real” Work

The problems above, together with the increase in non-core activities mentioned in the introduction, mean that users have less time to do what they are really employed to do, which is presumably to contribute his or her intelligence, knowledge and professional skills to further the organization’s business objectives. The result can be failure or poor quality in what the group or organization is trying to achieve.

Many busy workers with administrative responsibilities are primarily driven by what is in their email inbox. A colleague recently was heard to say – “if it isn’t in an email, I won’t get round to doing it”. Using such a reactive work approach based on one’s email inbox is less than ideal, since

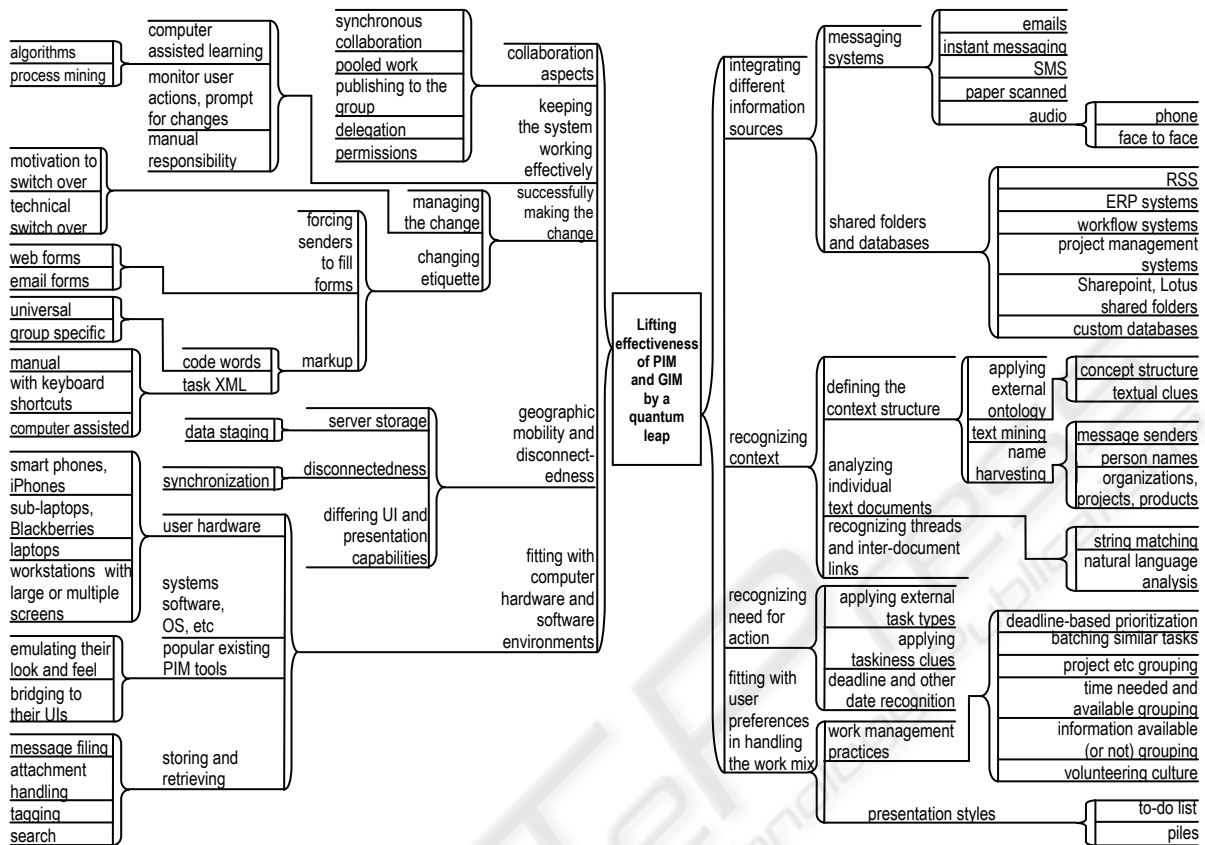


Figure 1: Mind Map of the Issues to be Addressed for PIM/GIM Improvement.

the inbox gives a rather limited and unbalanced view of the what are the important tasks within the user’s current projects and duties.

3 THE DIMENSIONS OF THE CHALLENGE

A mind map, drawn in FreeMind (FreeMind 2008), is presented as Figure 1. The following sub-sections address the major branches of this mind map, proceeding clockwise from the top right.

3.1 Integrating Heterogeneous Information Sources

A typical user receives information from many sources and in different forms. Messages may arrive in emails, IM chat sessions, SMS messages on a mobile phone, paper or audio (e.g. face-to-face communication, phone calls, recorded messages). Likewise, relevant information could be stored in shared folders (e.g. Sharepoint, Lotus), project

management systems, ERP systems or custom databases.

Further systems, such as event handlers, workflow systems and RSS, use rules applied to incoming changes to stored information to send notifications to the appropriate human participants.

The general problem today is that there is very limited integration of this information, which throws a heavy load on the human user. Data is stored in different structures and in different formats – a classic case is the .pst files used in MS Outlook. It is possible to resolve this, with something like the proposed MS Windows file store, (Grimes 2002). However this has still not been released, – and maybe never will.

3.2 Recognizing Context

As information arrives for an individual or group, someone has to resolve the question “what is this all about”. Traditionally this has been done either by the recipient, or by a trained assistant. Context recognition is particularly important in the case of email, where inboxes contain a mixture of items on

widely varying subjects. We identify three aspects to the context recognition problem.

First there is the task of identifying what possible contexts there are, both for an individual and for the groups he or she participates in. These contexts could form an ontology structure, which could be adopted from outside, created manually, or generated from analysis of documents previously received. In addition, an associated set of textual “clues” (including names, is needed to enable the recognition process to deduce that a document is relevant to a context.

The second aspect is that of analyzing documents of incoming information, and determining what context or contexts apply. This requires matching of strings and possibly inference using relationships in the context ontology.

A third aspect is the recognition that one document is related to another (typically earlier) document. A simple example is recognizing threads in email conversations (Bellotti and Thornton 2006). There are also other relationships between messages and documents, such as “attached to”, “new version of”, “comment on” etc.

3.3 Recognizing when Action is Required

Current messaging-based PIM systems are often imprecise about what action a recipient is being asked to carry out.

Currently there seems to be a low uptake – and lack of standardization – of task management facilities in PIM and groupware systems. Not many users actually use the Outlook task facility, which often requires a sender to attach a task to an email. A task list, encompassing tasks arising from all sources, might be a better primary view through which the user is prompted, with lists of emails as a secondary option displayed if required.

Converting free-form emails to “to-do” task items does not appear to be a simple process (Tagg et al 2009) 2008). Message senders use many circumlocutions to request another human to do something. Some level of formality, e.g. marking up in a “Task XML”, or use of agreed code strings (e.g. <deadline>”) might help.

Meetings and appointments are special cases of things we need to take action on, namely to attend them. Adoption is possibly more widespread than for tasks, but depends on there being a critical mass of competent users as well as software that supports a calendar exchange standard.

3.4 Catering for Users’ Work Handling Preferences

While many users tend by default to be driven by their email inbox, this may well not be an effective approach to handling one’s workload.

Alternative approaches to ordering a to-do list include deadline-based prioritization; batching of similar tasks; grouping by project, customer or contract; grouping by the time needed to perform (e.g. all tasks that I can do in less than 5 minutes); grouping by whether the information needed is readily available or whether it has to be rooted out; and pooling of work that can be picked up by any member of a team.

Lists and folder structures (in the style of email inboxes) may not anyhow be a good metaphor. Piles (Lam 2005) have been proposed as an alternative. Dossiers or piles could show icons for associated documents as well as for the messages, and indexes of contents could be included.

Portals allow a user to tailor a combination of visual components that present the information and tools that the user needs to be regularly aware of. There are developments for mobile devices, e.g. ePortal (CRM 2009), but they are not yet widespread for general PIM. More development is needed, particularly in standard paradigms for creating, adding and removing components that are to appear on an individual’s portal.

3.5 Fitting with Hardware and Software Environments

Any solution cannot avoid having to work within existing operating systems and other systems software, usually including databases and communications. It may also have to work, at least temporarily, with existing PIM software and groupware, either by emulating their interfaces or by bridging to the products themselves.

Because of user mobility, any software also has to be implementable on a range of portable computing devices ranging from laptops to mobile phones. (Grundy et al 2002) proposed a form of thin-client groupware oriented to this situation.

One particular issue, already mentioned in 3.1, is to the need for a much more integrated file system, covering – in a consistent way – the tagging, storing and retrieval of messages, attachments, contacts, documents, databases and simple files. Google Desktop and other similar search tools have offered a search facility that caters for files as different as .doc and .pst, but generality to cover networked

storage locations and plug-in devices are not as advanced yet.

3.6 Allowing for Geographic Mobility and Disconnection

Information workers today are becoming increasingly unlikely to stay tied to a single desk, and their networking is often wireless. They use different computing devices at different times and places, so they cannot use tools where working data is stored on a particular client device, unless they carry that device everywhere. Storage of information on a server therefore seems unavoidable, both for back up and for overflow. Synchronizing after a period of disconnection is also an issue.

Automatic indexing of documents (as in e.g. Google Desktop) needs to be extended to remote storage devices, and synchronization between the different computers becomes critical.

3.7 Successfully Making the Change

However good a new PIM/GIM solution may be, the question will arise: “what will it take for a critical mass of users to change to it from their current toolkits?” One can usefully look at quantum leaps in the past, such as spreadsheets, databases, B2C electronic commerce etc. One motivation would be a clearly demonstrable advantage to users, or to bring some new or previously minor niche functionality to the forefront.

However the power of fashions and of “follow my leader” should not be underestimated. If a large organization takes on a new system, or an outstanding public speaker or computer journalist preaches the cause, enthusiasm can become widespread. This power can also be seen from an individual user point of view when there are tools that require all collaborating users to use them (e.g. Facebook).

It is possible that as well as a major software improvement, a change in human-to-human etiquette will be needed. Current culture is to send information to other users without any contextual clues for the receiver. The acceptability of totally free-form emails may have to be reduced for the sake of reducing overload. Use of pre-formatted email forms (as in e.g. an Outlook Forms library), could be prescribed within a group or organization. Senders not following guidelines could be warned that their messages would receive lower priority.

Powerful organizations, or government, can already enforce their customers to use “Contact Us”

web pages instead of email, which shows a distinct possibility for form-based emails to be accepted in practice. However, in a survey carried out in 2006 by members of our research group, one respondent declared that users he knew would never accept forcing small groups or individuals to adopt this approach. Perhaps if users did become accustomed to a form-based system, with space for at least some free-form text to allow for language expression (Guerin 2004), acceptance would be higher.

3.8 Keeping the System Working Effectively

As has been shown with spam filtering, patterns of communication evolve. Peoples’ jobs, and the balance of an organization’s business, can quickly change. If humans are stressed, they may not have the time to revise their ontologies, especially at moments when they most need to do so.

Automatic learning systems, driven by artificial intelligence algorithms, seem desirable, but users need to have confidence in them. It would not be acceptable to users if the category structure was changed too often or too drastically. It might be better to use an agent that can monitor the efficiency of the context recognition tools over time and warn the user of the need to make manual changes. Process mining – the technique of detecting patterns in related operations, messages and documents, could be a useful way of providing such feedback.

3.9 Handling the Collaboration Aspects

Collaborative work brings up a number of needs, including synchronous collaboration, pooled work (see 3.3), publishing information to the group, delegation and permissions to access data.

It is assumed that one user may be involved in many groups, each of which may have their own context structure and preferred practices. So there is a need for any new system to be able to simultaneously link to more than one group work support system.

4 CONCLUSIONS

There have been a number of studies attempting to find improvements in PIM and GIM, e.g. (PIM 2008). The reality seems to be that a piecemeal approach will not be enough to trigger the quantum leap needed, especially when considering the vast number and range of custom-made PIM tools and

add-ons (e.g. to Outlook) already in existence. At least 3 or 4 of the issues raised in section 3 above would seem to be required, in order for things to start happening, and even then one would need to consider how these solutions would be adopted and migrated to in practice.

This is an area where a multidisciplinary approach seems essential. No single researcher, or academic research group, is likely to have the time or resources to build a complete enough toolbox. A well managed consortium, with involvement from commercial interests, might do better.

Otherwise we will all have to wait for Microsoft or IBM to come up with and implement an answer. The need for such improvements has been discussed for at least 5 years. We can wonder why we have seen so few signs of a solution so far. Possible causes are that a) other issues have been more pressing (e.g. spam and information security); b) that the state of the art in recognising context and tasks is not yet good enough to give reliable results; and c) introducing such a wide range of interconnected software is so risky that a commercial organization cannot afford a spectacular failure.

So, the opportunity might be there for an entrepreneur (remember the original Bill Gates?) - or an academic consortium less hamstrung by the constraints of KPIs and research quality assessments.

REFERENCES

- Bellotti V. and Thornton J. (2006), Managing Activities with TV-Acta: TaskVista and Activity-Centered Task Assistant. In *Proceedings of the Second SIGIR Workshop on Personal Information Management (PIM)*.
- Bernstein, M., van Kleek, M., Karger, D. and Schraefel, MC., (2001), Information Scraps: How and Why Information Eludes our Personal Information Management Tools, *ACM Transactions on Office Information Systems*.
- Booher, D. (2007), E-mail: Could You Live Without It? *Booher Banter*, <http://www.booher.com/wordpress/?p=282>.
- Catarci, T., Dix, A., Katifori, A., Lepouras G. And Poggi A. (2007), Task-centered Information Management, In *DELOS Conference on Digital Libraries*., Pisa, Italy
- CRM (2009), Ebridge Software Eportal Now Compatible with Blackberry Storm, Blackberry Bold and Apple iPhone, 12 January, <http://crmhelpdesksoftware.com/index.php?s=erp>
- Czerwinski, M. (2006), From Scatterbrained to Focused: UI Support for Today's Crazy Information Worker.
- In *SIGIR Workshop on Personal Information Management*, Seattle.
- Fitzgerald, M. (2004), Call It the Dead E-Mail Office, *Wired* magazine, 7 June.
- FreeMind (2008), Free Mind Mapping Software, Wiki, http://freemind.sourceforge.net/wiki/index.php/Main_Page
- Grimes, R. (2002) Revolutionary File Storage System Lets Users Search and Manage Files Based on Content, <http://msdn.microsoft.com/en-us/magazine/cc164028.aspx>
- Guerin, B. (2004), Handbook for Analyzing the Social Strategies of Everyday Life, Context Press, USA, Chapter 5.
- Grundy, J., Wang, X. and Hosking, J. (2002), Building Multi-device, Component-based, Thin-client Groupware,, In Proc Australasian User Interfaces Conference, Melbourne, Australia, <http://www.cs.auckland.ac.nz/~johng/papers/auic2002.pdf>
- Kirsh, D. (2000), A Few Thoughts on Cognitive Overload, *Intellectica*.
- Klauegger, C., Sinkovics, R. and Zou, H. (2007), Information overload: a Cross-national Investigation of Influence Factors and Effects, *Emerald Insight*, www.emeraldinsight.com/0263-4503.htm
- Lam, B. (2005), MailStacker: Applying a Pile Metaphor for Email Management, Honours Thesis, School of Information Technology and Electrical Engineering, University of Queensland.
- PIM (2008), CHI Conference Workshop on Personal Information Management, Florence, Italy
- Rifkin A. and Khare, R. (1998) The Evolution of Internet-Scale Event Notification Services: Past, Present, and Future, <http://www.ics.uci.edu/~rohit/wacc>
- Tagg, R., Srinivasan Kumar, R, Gandhi P. and Lalwani, H. (2009) Recognizing Work Priorities and Taskiness in Incoming Messages by using Personal Ontologies with Lexical Clues, submitted for publication)
- Tungare, M., Pyla, P., Sampat, M. and Perez-Quinones, M. (2006), Defragmenting Information using the Syncables Framework. In *SIGIR Workshop on Personal Information Management*, Seattle.
- Whittaker S. and Sidner C. (1996), Email Overload: Exploring Personal Information Management of Email. In *ACM Conference on Computer-Human Interaction*.
- Whittaker, S., Bellotti, V. and Gwizdka, J. (2006), Email in Personal Information Management, *Communications of the ACM* (49) 1.