PIM Research Protoypes Ideas and What Has Been Transferred to Mainstream Software

Matjaz Kljun
m.kljun@lancaster.ac.uk*

Alan Dix
dixa@comp.lancs.ac.uk *

September 2010

Abstract

Personal Information Management (PIM) is a study on how people acquire, organize, maintain, retrieve, archive and discard information for various reasons and needs in physical and digital worlds. Many PIM tools are available for managing information on our desktop computers. And many research prototypes tried to augment or replace them. The development of these tools was based on the knowledge drawn form the field of psychology, human computer interaction, information retrieval and the research in the PIM field. Different metaphors and ways of organizing were introduced. However, little has been transferred to mainstream products. Most of these these prototypes were not extensively tested but results of the most looked prominent. This paper a classification of PIM research prototypes, solved issues, what has been developed and learnt, what has been transferred in mainstream applications and a quick look in to the future on where the development and research prototypes might be heading.

This paper provides an early classification and is a basis for another in depth analysis of supported means and trends in PIM prototypes research (based on PIM activities - see later publications on http://pim.famnit.upr.si/blog/). A lot of text from this paper is left out in the new version. Especially the analysis of what have been transferred to available software for consumers. Hence a technical report as this could prove of use for other researchers.

1 Introduction

Researchers have long studied how people manage different types of documents and several problems have been found in both physical and digital domains. A lot of research has been done to study how people manage files [11][120], email [167] and web bookmarks [1] as separate entities. It is not surprising that PIM prototypes tried to overcome problems people had with managing separately information of different types. In the last decade, when most of the information got digitalized and with the advance of portable devices, a problem of how information fragmentation affects the task performance has become prominent [23, 16, 154]. Many prototypes tried to unify the information space on personal computers and beyond. A focus also shifted to a research on how people manage their information together in physical and digital domains [74, 24] (although

*InfoLab21, Lancaster University

it seams that there was always interest in providing digital capabilities in managing of physical documents), as well as how they manage information on different devices and across different tools on these devices [154]. In the last few years managing personal information on, from and with the help of the web is included in studies as well [106].

Some authors have already classified PIM prototypes. Boardman classified research prototypes in four groups: (1) tools that improve management of specific information type (files, email, web bookmarks) where integration of other information types is not a goal, (2) tools that try to improve integration between distinctive tools, (3) tools that are focused on managing one information type but embed additional support for managing other information types as well, and (4) tools to manage different information types within a new, unified design [21]. Karger's classification focused on information integration from the point of underlying technology [82, 84]. He divided PIM tools in three main groups: (1) visual unification, (2) unification based on standard data types like copying and pasting text from one application to another and (3) unification based on metadata (grouping of information items, cross referencing like on the web, attributes and relations).

This paper is not focusing on integration alone and describes more than 100 prototypes divided by information types and information integration. The first three sections (2, 3, 4) describe prototypes that addressed problems in managing each information type separately: emails, files and web bookmarks and history. Next section (5) is dedicated to prototypes that integrate different information types to overcome managing problems of semantically related information, while section 6 looks at other prototypes that are not focusing on three main information types. In each section, prototypes are listed in a time order whenever possible (with a year in brackets following the name), but are also grouped and compared together by implemented functionalities to overcome different existing drawbacks. Each of these sections and subsections also discusses what is used in (or have been transferred to) mainstream applications. The paper concludes with an overview of what we gained from these prototypes and a glance in the possible future of PIM research and prototypes (7).

2 Email and other communication tools

Email has become one of the most popular asynchronous communication tool and computer application. Most of us spend significant amount of time behind the email client checking, reading and replying to email messages [47, 96]. Sometimes even a quarter of a working day tasks include email [47]. One of the problems people face with email is email overload. Concerns about information overload has been already expressed in the 80's [40] and it was later studied by several researchers [96, 167, 56]. Below we divided email prototypes in five groups which sometimes overlap as well: (1) prototypes that provide some sort of automation or help with rules and filers and focus on organization, (2) prototypes providing context clues for easing retrieval, (3) prototypes in support of tasks, (4) prototypes with a different visualization of a inbox and (3) prototypes with other information types embedded.

2.1 Rules, filters, agents and automatic classification

Information Lens (1987) addressed speed of communication (with provided templates) and information overload (with rules and filters to automatically file emails). It also tried spotting possibly relevant information (from e.g. mailing list emails) and automatically forward it to others [100]. One potential problem with rules and filters is that users do not like them in general [167, 13]. Automatically filed information can easily be missed and automatically sent information deprives users of a control of what is sent out. Tapestry (1992) introduced user and collaborative based email filtering to help users organize email [63].

Email client Maxims (1994) learned from users' actions and it could prioritize, delete and archive messages on behalf of users, if its prediction confidence level was above a threshold level (in this case it also generated a note message for the user) – otherwise it suggested possible actions [98]. Automatic classification was also used in email client Re:Agent (1998) with text mining based learning capabilities [25]. CAFE (1998), introduced three different modes of automatic email management based on the time users have at hand: busy, cool, and curious mode [149]. Segal and Kephart were not in favour of automatic actions and suggested that rules and filters should be applied only after emails are read [139]. Their email client Mailcat (1999) processed all incoming email though an adaptive classifier and proposed three most probable folders to which emails could be filed (with a 80% to 90% success rate).

Another approach was used in the Apple Data Detectors (1998). Its main feature was to make it easier for users to copy information from email to other tools. Apple Data Detectors could parse a selected region of text in an email for structured information like dates, postal and email addresses, phone numbers, etc. If, for example, a selected text contained a date, the text could be placed in a calendar [107].

Rules and filter are present in almost all email clients, although, to our knowledge, it is not known to what extent these are used. Automation is usually used for detecting spam, however, the new Google Mail feature ”Rriority inbox” incorporates some of the ideas from mentioned prototypes for sorting email in inboxes. ”Priority inbox” learns and sorts email based on users' actions like: which email is read, replied to, the sender, etc. Data Detectors are also used in Apple's PIM suite (iCal, Mail and Address book). Some email clients (like Thunderbird) parse email text and detect if a user wants to send the attachment while some clients (e.g. Thunderbird, Gmail) allow one click ”archiving” from the inbox. So agents and automation are getting implemented in mainstream email clients, albeit very slow (over two decades after email has started to be used in everyday life and not only in academia).

2.2 Contextual clues for retrieving and reminding

Context in which information is viewed, classified and remembered was found very important for later access and retrieval [153, 89]. Several prototypes help retrieving activity by providing relevant information, different views to inbox and conversations.

Email client Mona (1993) could show context of each conversation (email thread) graphically [38]. Similarly Thread-based email client (2001) joined emails and replies in the threads and participants using this prototype liked the possibility to refer back to previous messages in conversations [157].

BiFrost (2002) organized email in inbox by categories which changed based on the context. A first category contained potentially important emails based on a day and calendar inputs. A group of emails from people that are important were in the next category. Then there was a category of email replies, a category of emails sent to many people, etc [9]. It was proven that category interfaces help users find relevant information on the web quicker than with plain list interfaces [49] and participants of the BiFrost study also liked the way categories brought organization to the inbox. Bifrost could be also mentioned in the previous group for its automation sorting and text parsing, but its focus was more on providing information in context.

Sudarsky's prototype (2002) automatically ordered email (which freed users of filing) in a hierarchy in three possible ways [148]: (1) based on domain names of senders' addresses, with top domains (.com, .net, .org, etc.) as the first hierarchy level, (2) based on the company's or organization's names retrieved from a domain (companyname.com and companyname.org would be grouped together) and (3) based on senders' last name. The main interface window showed email of a selected hierarchy node as squares, organized in a grid by time and sender (based on TimeSpace mentioned in the next subsection 2.3). Messages in this view could be additionally filtered with a keyword search which coloured relevant email or if the number of relevant results was low, additional clues like email content, senders photo and number of recipients were shown. This part of the interface could also show conversation threads in a similar graph as the Mona prototype but also ordered by time. Sudarsky's prototype could also be mentioned with automatic classification prototypes, but its focus was to provide contextual clues to email retrieval.

EzMail (2004) had a multi-view interface to assist email management. It provided (1) visualizations for individual messages, their properties and user made annotations, and (2) visualizations of messages as components of threads, providing contextual information of conversations (with capability of replying simultaneously to several messages at the same time) [134]. ReMail (2004) provided even more contextual clues like a time scale, a graphical representation of email threads (called Thread Maps or Thread Arch), list of people involved in a conversation, annotations, discovery of dates and time, showing relationships between messages in the same folder (called Message Maps), annotations and selective displays [132, 86, 131].

Some of these concepts such as time scale (e.g. Thunderbird search results) and email threads (e.g. Gmail and some Thunderbird's extensions such as ThreadVis or Gmail Conversations) are implemented in available email clients. Several clients also support easy one-click annotating (or flagging) which provides contextual clues for easier retrieval or tagging (labelling) for multi-categorizing email. Some even have rules for automatic tagging (Gmail) which does not move emails out of sight, but rather give clues to what might be important (e.g. tag all emails from a certain domain) which, to some extent, allows grouping email in the inbox.

Other research prototypes mentioned in the next two subsections also provide several clues for reminding and retrieving email, but their focus was on supporting tasks or visualizing hidden patterns in email archives.

2.3 Task support

Email has been always extensively used for other activities besides its primal role - communication [96, 47]. People manage tasks, delegate work, backup files, send themselves reminders and use email for several other activities which were not envisioned by email creators. This flexibility of email has its drawbacks as well, as important emails can easily be buried in inbox with less important ones. The diversity of usage has led to several prototypes.

Raton Laveur (2000) integrated calendar, to-do list and notes in a regular inbox to provide support for tasks reminders [14]. To reduce clutter (of mixed information in the inbox) and cognitive effort (to file information), it allowed grouping and searching for information with users' defined and predefined queries (unseen, completed to-dos, etc.). TaskMaster (2003) also had task management incorporated in the email client where each task (called thrask by authors as a combination of words thread and task) could include a list of emails, URL's and attached files [13]. All this information could be viewed in an internal viewer without leaving the email client. Each thrask could have associated clusters of actions and a due date (additional attributes). The due dates were presented with graphical elements which were always visible and changed accordingly to approaching due dates. The tool was positively accepted by participants but was never developed in a fully functional email client. One of the drawback of both tools is added complexity and cognitive load with several organizing mechanisms.

An interesting approach was used in TimeStore (1997) - a time-based email visualization [175]. This email client visualized emails as dots on a two dimensional grid with time on horizontal axis and contacts on vertical axis. In this view a user could see rhythms of conversations (on which days a particular person sent emails or emails were sent to), but the subject of each email was obscured. Users could create additional notes on each email to support tasks. This time-based visualization provided a reminding function with visually emphasised emails (tasks) that moved too far away in the past and a list of contacts that helped as social reminders. TimeStore vision was to completely abandon the traditional inbox view and free users from filing. The tool inspired another grid based email prototype called TaskView (2002). This prototype provided a matrix view of email messages with dates on the top axis and email subjects on a side (day/subject as opposed to day/sender in TimeStore) [65]. While both prototypes provide a better overview of tasks in emails, they also lacked the regular inbox view in which users can search for their emails by subjects, dates, senders and other standard email attributes.

All prototypes in this group could be also mentioned in the above section as they provide contextual clues and reminders, but their main focus is on tasks. Some present email clients have tasks management embedded in email interface. Outlook version 2007 has a task list, to which emails can be dragged to as new tasks. Gmail has an extension called Taskforce which splits emails in three categories: information, action and broadcast. Tasks can be created from its interface or from email, can be shared and can have files attached as well. Thunderbird also has several extensions that provide reminding capabilities of to-do's and tasks integrated in a calendar. But there is still room for improvement of design of these tools, integration (e.g. calendar with tasks in Outlook) and reminding functionalities.

2.4 Different visualizations of the inbox

The idea behind most of presented visualizations is to provide a view to a missing context by revealing habits, relationships and other data and patterns that are hidden within the email archives. Most of these prototypes are not focused on email management at all and would only work as an add on to present email applications. Nonetheless they allow users to reflect upon their PIM past. Most of these prototypes would also fit into context category as well, as they provide additional context to email browsing via various means, but their visual presentations are very different from traditional list view of the inbox and email folders. Such prototypes have been divided in three main categories [29]: thread-based, time-based and social-network based visualizations.

Instead of visualizations based on social networks we will describe a group of prototypes that visualize inbox through contacts (which can be based on social networks or not). And we are adding two other categories: a metaphor based visualizations and will also just mention art visualizations. As said before, some of these prototypes fall in more than one category (eg. a mix of a social network evolving through time or threads on a time scale) and will be mentioned a few times.

Thread-based vizualization

We already mentioned some threaded visualizations such as the Mona client, Thread-based email client, EzMail, Sudarsky's prototype (see section 2.2) and TaskView (see section 2.3). We also mentioned reMail client which provided several contextual clues. During the development of the reMail (2001), authors came up with several visualization ideas such as threads on a time scale and coloured thread nodes to distinguish relations of senders to receiver [132]. At the end Remail used Thread Arcs (2003). The Thread Arcs tool visualizes a chain of email nodes that are related to each other, where messages are nodes ordered by age from left to right and arcs link each child to its parent [85].

Another prototype called Zest (2002) extracted initial questions and replies and constructed threads based on content as opposed to whole emails (http://zesty.ca/zest/) [173]. Already mentioned EzMail incorporated several visualizations with a focus on individual messages, message properties, and messages in a thread. Threadmap (2005) used a treemap algorithm [72] to display and help identify positions, arguments, and evidences of conversations (using special conversation convention) [174]. Although even authors acknowledge that this layout is not good to visualize long conversational threads (estate limits of the depth and breadth of the tree) and long email (because narrow columns of text are hard to read). Their prototype showed only a few levels of a thread but allows users to dig in. This prototype was intended to use in newsgroups where threads can grow really big and occasional reader would quickly find all relevant positions in a conversation (http://zesty.ca/threadmap/). Zest and Threadmaps were meant for large scale conversations; however, they could be used for visualizing threads in inbox as well.

Time-based vizualization

TimeStore, Sudarsky's prototype and TaskView were mentioned in previous sections (2.2, 2.3). The grid layout of email was meant to completely replace list view and free users of filing. TimeStore and Sudarsky's prototype with contacts on a vertical axis revealed the rhythms of conversations with each contact. While TaskView showed timely ordered email in threads with email subjects on a vertical axis. Timely ordered email threads with coloured emails as nodes in a graph were used in a predecessor of reMail [132]. This prototype used also other time based visualizations such as reduced resolution email overviews on a calendar like interface with dates on the vertical axis on the top and columns of emails underneath.

Some prototypes did not even show email. PostHistory (2004) interface was divided into two main panels: the calendar panel on the left (with entire calendar year), which showed the intensity of email exchanges over time, and the contacts' panel on the right [159]. Intensity of conversations was shown with squares placed on days in a calendar, where size of each square told the amount of email being sent and received (bigger the square heavier the email traffic) and color the directness (brighter the color more email was directed to the recipient as opposed to more recipients or mailing lists). The contacts panel had a user's name in the centre and names of other contacts placed around it - more email they exchanged, closer they are to the user.

FaMailiar (2004) also visualized email data over time. It revealed communication rhythms and patterns displaying email on a time plot (shown as different shapes based on directness and contact's intimacy with user). It also showed some statistics such as daily email averages, daily quality of emails, frequency of email exchanges, comparative frequency of email exchanges, etc. [103]. Similarly, a visualization from EmailViz showed “rhythm of a relationship” (2005) or how email activity with each contact changes over time. Several graphs were used to visualize a start of an active relationship, rhythm (intensity) of conversations and how it eventually dies away [115]. Mailview (2005) also plotted emails in a chronologically based visualization which enabled users to analyse and observe conversations' trends over time and perceive emails with similar features in a filtering and zoomable interface [57].

The Mountain (2005) prototype revealed how our conversations or relations with other people change over time. It could reveal life changes such as switching jobs, retiring, graduating or similar life events. It visualized an email archive with layers ordered by time on a time scale. Each layer represented a different person. Bottom layers represented contacts from early conversations, while the top layers of the mountain represented more recent people in the archive. The thickness of each layer referred to how recently the person has been in contact with a user. Layers could also be highlighted, causing the first words of every email exchanged with this person to appear on the screen [158].

Similar to the Mountain, Themail (2006) also showed words on the time scale [162]. This prototype parsed content of email and extracted keywords which were displayed on a screen in time ordered columns, coloured and sized based on the frequency of use and conversation they belong to. It basically visualized topics of conversations with each contact and showed how conversations differ between contacts.

Contacts-based visualization

Four of above mentioned prototypes (TimeStore, contacts panel of PostHistory, "rhythm of a relationship" and the Mountain) used contacts in relation to time, where time was a component that drove visualization. In this section visualization's focus is on contacts. Most of the presented prototypes tried to visualize a network of connections between people found in email archive(s). There are two main approaches: (1) a user perspective on a personal social network built from a point of view of one user (either automatically or manually) or (2) the whole network visualization with no central spot. We will describe prototypes with a personal approach but will also mention prototypes that visualized bigger networks.

Personal Map (2003) draw a social network with a user in the centre, surrounded by groups of contacts based on their co-appearance in emails and radially arranged so that the most important contacts (based on quantity of the exchanged email) are nearest to the center [54]. Participants in a study liked the view of their social networks (even with errors) but they mostly stated they would use it for curiosity or sending out group email. In a similar fashion, Social Network Fragments (2004) also arranged contacts according to their closeness, which was derived from co-occurrence of the addresses in email [159]. But it added a time dimension, so one could watch the evolution of the network – when people appear, when they become linked to others, when they fade into the background.

On the contrary (to the rest of prototypes in this group), developers of ContactMap (2004) discovered in a series of development circles, that users want to build their social network by themselves (with a help of a computer) and did not like automatically built networks [166]. ContactMap is a social desktop where thumbnail images of contacts are spatially managed and grouped by the user. This social desktop provided the support for two distinctive tasks that usually happen in a shared working environment: (1) reminding about social commitments (such as TimeStore with a focus on people) and (2) social data mining for expertise or advice. Emails could be written directly by selecting appropriate option on the contact's thumbnails, while new emails were shown as envelopes on a contacts' thumbnails.

Another set of visualizations from EmailViz used treemap, scatter plots and histogram timelines (2006) to reveal some patterns in communication [116]. Treemaps presented hierarchically ordered set of boxes within boxes based on email domain names (Threadmaps used this approach to visualize conversation threads). Size of each contact's node was proportional to the number of messages authored during the selected time period and the intensity of a color showed the number of unread messages currently in the email store. "Correspondent Crowds" scatter plots revealed the dynamics of communication and were generated based on the number of messages sent to the the contact against the number of messages received from the contact. The third visualization was a histogram called AuthorLines [161] that revealed the difference between the number of messages in conversation threads authored by the user (presented as bubbles above a centre dividing line) and number of replies initiated by others (presented as bubbles below a centre dividing line) on a weekly basis. The histogram contained 52 (weeks in a year) strips of bubbles below and above the central line where bubble's size increased for every message authored in that week in that thread.

The rest of the prototypes visualize networks from a 3rd person's perspective such as the whole company's network with no inherent centre. Visual Who (1995) is one of the first examples of an email based whole network visualization. It used an organization's mailing lists to depict the members affiliations [43]. Conversation Map (2000) analysed text of messages and let users navigate discussions and threads by (1) social networks using patterns of response such as mutual citations, (2) semantic networks of discussions with possible relationships among them, and (3) a set of discussion themes [133]. The Enronic (2005) prototype visualized an inferred social network [67]. It introduced number of interactive features such as highlighting parts of the network with hovering over, panning and zooming, clicking on nodes revealing all email of one person, clicking on edge between two contacts revealing all email traffic between these two people, pie charts on edges showing how email between two people has been categorized, etc. Small-World Email Networks (2007) also used several types of visualization: (1) visualization of a social network on the surface of a sphere to reveal the relationships between different groups, (2) a 2.5D hierarchical visualization method combined with the centrality value of nodes to analyse important people, (3) a 2.5D visualization method to analyse the evolution of email relationships between people changing over time (similar to Personal Map), and (4) an ambient display method for finding social circles derived from the email network [60].

Metaphor-based visualization

Some prototypes tried to introduce different metaphors to email management. Email nodes (2004) let users manage their email in a spatial zoomable environment [41]. Emails were organized in ordered piles (called nodes) based on threads, topics, projects, etc. Zooming closer to nodes or email revealed more content and mail thumbnails used coloured borders to additionally distinguish them. Piles were also used in Mailstacker (2005) [90]. The main drawback is scalability of such systems as with vast amounts of daily received emails the workspace could easily get cluttered.

A natural metaphor was used in Anymails (2007) where email were categorized in 6 (predefined) categories and each category was represented with 6 species which differed in color and form [71]. Email (animals) were shown on a surface for a certain period of time which could be adjusted by user in a time line. The age of email could be distinguished by the size and opacity of animals, while status (read, unread, replied to) was shown with number of hair and speed (from hairy and speedy unread email to almost static replied to email). This prototype could reveal density of received email by groups in a selected time period (http://carohorn.de/anymails/). In the Mail Garden (2007) email in inbox was arranged horizontally by time and visualized as trees. The height of a tree reflected the length of the email it represented. As Anymails, it had a time line on the bottom of the interface where users could select a time period (http://www.kjenwilkens.com/projects/mail-garden). Both prototypes could also be categorized as time based visualizations.

Magnet Mail (2009) used natural metaphor of magnetism [29]. It helped users browsing and searching email on a zoomable spatial area where searching keywords were represented as magnets and email as envelops. Dragging magnets attracted or repulsed email in the spatial visualization area, according to their value.

Art-based vizualization

These visualizations are focused entirely on art and are not revealing hidden patterns and do not allow managing email. Such prototypes are Spam Recyler (http://www.spamrecycling.com/), Spam Architecture (http://www.sq.ro/spamarchitecture.php), SpamPaint (http://www.kingcosmonaut.de/spampaint/), Spam Plants (http://www.sq.ro/spamplants.php).

While different visualizations or views were positively accepted by participants in the studies, users usually did not want to abandon inbox view. Especially with prototypes that still allowed some email management such as ContactMap, Time Store, Time View. Such visualizations are often seen as a supplement to existing email management for easier navigation, search and revelation of some patterns.

Some attempts were made to bring social networks to email. Personal Map was partially integrated with Microsoft Outlook and enabled users to compose email to groups of people. One of the problems was dynamically changing network as names appeared in different places at different times – this is fine for exploring interesting pattern, but for composing mail such dynamics seem annoying. Xobni also provided social network and search for Outlook users. Google also tried to integrate social networking (Google buzz) with email. There are many aggregators that try to bring several web 2.0 services together. One of them is Vodafone 360 that brings together all contacts from phone, email, instant messenger and social network accounts and aggregates social network feeds.

Besides social networks, there are other available views of one's email archive. Gmail has several add-ons such as "Graph your inbox" which generates some statistics of usage. There are even more interesting approaches. The 3D mailbox turns email management in to a game within a 3D digital world where emails are visualized as people and aircrafts http://www.3dmailbox.com/ while 0boxer turns cleaning the Gmail's inbox into a competition.

2.5 Embedding other information types

Three of the above mentioned prototypes integrated management of other information types but the main focus of these tools is still email (TaskMaster, Raton Laveur) or communication (ContactMap). Embedded calendar, address book and to-do notes are also common in several email clients. Some clients also incorporate agents that can facilitate entries from email to calendar or email to address book (such as the above mentioned Apple PIM suite). These embedded systems are daily used by millions of users.

Integrating email with other information types and formats can provide easier retrieval. It has been found that to some extent people want to manage email with other information types: particularly files and web pages (note that contacts and calendar are usually a part of email clients). Boardman and Sasse found that email hierarchy and a file hierarchy overlap especially where information is managed by projects [23]. To overcome this problem some people save email messages as files in a file hierarchy or print emails to be able to manage them with paper documents [24]. But such emails can lose context of a conversation thread. Even so, most users want to keep email, files and web bookmarks separated to some extent because of their different acquisition type and lack of control over email flow [23]. Integration prototypes are discussed in section 5.

3 Files

Several studies have been carried out to better understand how users manage their files [120, 11, 10, 45]. In these studies, hierarchies used to manage files were criticized for various reasons: files can only be at one place at a time, organizational structure cannot easily change over time, dynamically organized collections (according to projects, time, content, etc.) are not possible, automation in organization and archiving is missing, naming and categorizing require high cognitive efforts, and file attributes such as names, creation/modification/access times, size, etc. are more helpful to computers than users.

Several of these issues have already been addressed in mainstream file managers. Some features (aliases, virtual folders) exist for a while now, but it is not known how and if they are used in practice, while some features (automatic stacks or arrangement views) are fairly new and it is not known yet how they affect information management. Some researchers argue that most of the practices users employ and problems they have, are related to the technology and artefacts that underlie the managing activities [55]. The static nature of hierarchies has also its supporters, who argue that it makes orienteering easier [22]. Orienteering is what people seem to prefer as a navigation technique for retrieving personal and also publicly available information [11, 18]. Discussion in this section will focus on prototypes for file management only. Several cross tool prototypes also tried to solve mentioned problems, but are presented in section 5 as they are also addressing information fragmentation.

3.1 Attributes and multi-categorization

All mentioned issues have lead to several research prototypes. Some prototypes tried to address rigidity of hierarchies and user unfriendly attributes. Semantic file system or SFS (1991) was developed to provide access to files by associations via virtual folders [62]. This folders corresponded to associative queries over automatically indexed set of attributes. SFS was an extension to the tree-based Unix file system and its main focus was not on PIM but on enriching a command line interface. Nonetheless similar ideas lead to several other cross-tool prototypes that provided associative access to information [64, 32, 121] and those of semantic desktop [83] which will be presented in section 5 .

DomainView (1999) introduced dynamic structures based on (computer and user defined) attributes where categories were dynamically formed by search queries or manually defined by users [8]. Categories could also overlap and be nested. TagsFS (2006) also introduced file attributes and fluid hierarchies, where on each orienteering step only relevant folders are shown to the user [20].

User defined attributes were fully supported in already late BeOS (now known as Haiku). To some extent, Windows Explorer in Windows 7 lets users define attributes (author and title) for certain file types such as Office documents, and it lets adding tags to images. Since Windows Vista, it supports saving search results as virtual folders. The same feature is available in OS X Finder file manager. Such folders change dynamically based on the search query. Windows explorer also supports automatic grouping for certain file types such as stacking photos by month or music files by authors using additional metadata of these file types (Exif, id3 tag).

3.2 Spatial management and piling

Piles were first mentioned as a possible computer metaphor by Malone in 1983 [99] and later developed by Mander et. al (1992) [102]. Malone argued that piles require less cognitive effort to be managed , compared to categorizing information. He also argued, that spatial management better served memory for locating documents as do hierarchies. One of the earliest systems that allowed spatial management of (public) documents on a zoomable interface was Dataland (1979) or Spatial Data-Management System [108]. Its successor CCA introduced sub-workspaces (virtual rooms in a large spatial area), where users moved to when they zoomed in close enough [27].

Dynapad (2004) introduced non stacked piles of photos' and documents' thumbnails on a zoomable interface to achieve higher visual recognition and remindability, and to overcome spatial limits (such as Dataland) [12]. It supported implicitly organized piles and clumps which could be created by encircling a set of thumbnails with a mouse pointer. It also provided a view to a timely arranged thumbnails and it supported sub-piling. Archy (2005) also provided a zoomable desktop where information (thumbnails) could be zoomed in and directly edited (no need for applications) and was always in a persistent state [168]. Bumptop (2006) enhanced piles by moving them to a 3D environment and introduced physics to piles' interaction [2].

Even early file managers allowed spatial management of document icons on a desktop and in folders (and some users use this feature [120]). This feature still exists in OS X Finder and Gnome Nautilus (which are so called spatial file managers) while it was disabled in Windows Explorer on Windows 7 (where auto-arrangement can not be turned off anymore). Windows Explorer in Windows 7 also tries to hide a hierarchy behind so called libraries which aggregate multiple storage locations. During the navigation it does not automatically expand hierarchy on the left pane by default, which hides it from the users (showing only a content of one folder at a time). It also introduced (as we mentioned above) arrangement views which let users stack files such as photos or music based on metadata (month, author) which are shown as piles in a folder.

3.3 Searching and contextual clues for retrieving and reminding

Piles and spatial management provide contextual clues for retrieving and reminding. Although, organizing information was as important as retrieval in these tools. A tool focusing only in retrieval, by providing a compiled list of related documents to a currently used one was Remembrance agent (1996) [125]. Other similar prototypes also addressed information fragmentation and will be presented later [170, 50].

The VIBE system (1993) visualized documents returned by a search engine on a spatial surface together with search keywords [112]. Document icons were placed closer to the searched word based on word count. The system also looked for semantic relationships between documents and used indexing to extract all relevant information. Automatic indexing was also used by already mentioned Semantic File System. Several indexing techniques were later developed and tested in file search engines such as GLIMPSE (1994). GLIMPSE provided full-text search, approximate matching which allowed misspelling and support for regular expressions [101]. Dynamic queries (1998) were also implemented on a top of a file system, besides being successful in dynamically visualizing data from databases. Dynamic queries provide easy real time manipulation (a kind of faceted search) of information visualization with widgets (sliders, buttons, etc.) [94]. While these are moved and clicked, appropriate information is shown on the screen. The problem with dynamic queries over a file system is that a file system is not a set of semi-structured information and the only two variables that can be manipulated are size and time. Although for some file types that contain integrated meta data (e.g. music and photos), dynamic queries make sense.

There are several other tools for searching semi-structured collections of data (or tagged information) such as Envision (1996) [111] which presented search results in a matrix where axes held documents' characteristics, or a multi-faceted bidirectional search interface called mspace (2003) [137]. A more general system that supported navigation of semi-structured information is Magnet (2003) which integrated keyword search and tags navigation, provided suggestions of similar results, and kept track of a browsing history [143].

Thumbnails can also facilitate searching and reminding. Of the so far listed prototypes only Dynapad used thumbnails. Several present file managers show thumbnails and preview documents. Indexing and full text search are also a standard, but often used only when all other searching possibilities fail [18]. However there is little or no support of semantics in file managers (one exception is Konqueror with implemented Nepomuk semantic desktop which will be presented in subsection 5.3), and additional context clues (more tools addressing this issue are presented in subsection 5.2).

3.4 Time dimension

Already mentioned Dynapad and Envision provided a time scale view over information items. Before these two, time based ordering was used in the Perspective Wall (1991) [27] where search results were placed in a fish-eyed grid ordered by time and file types [97] (this visualization was used in The information visualizer presented in subsection 3.6). MEMOIRS system (1992) also tried to organize information (documents and diary entries) as events on a timeline interface [91]. This design supported (1) episodic retrieval by exploiting autobiographical memory, and (2) retrieval by recognition (based on visual appearance) by allowing items to be colour coded.

TimeScape/TimeMachine (1999) also exploited the time dimension, although for a very different use. TimeScape was a time based desktop manager which preserved past states of a , functioning as an automatic archiving system since every past state of the desktop could be accessed. It also allowed creation of future states that served as reminders [122]. An application also called the Time Machine is a part of the OS X, but it server as a backup solution and does not act as a time desktop manager.

Another time based application is Time2hide (2008) [93]. It helped users to keep computer desktop uncluttered with hiding old and unused icons (fading effect) using general and per icon hiding settings. Similar behaviour was also implemented in TimeScape/TimeMachine where deleted files and old post-it notes faded away. Time dimensions used in other cross tool prototypes such as Lifestreams and MyLifeBits are presented later in section 5.

Time based organization is used in Nemo file browser, which shows files in a calendar interface based on creation, access and modification time. It also supports faceted search to address clutter on a week, month and year views. In all major file managers, a list of files can be ordered by time (similar to email or browsing history where default organization is by time), size and type on a folder basis. Besides a list view, major file manages provide a thumbnail view as well. However, other visualizations and browsing possibilities are not present.

3.5 Agents and automatic classification

Automation for filing information was used in a data mining prototype Athena (2000). It could discover topics and organize files by them, reorganize documents by giving it some as example, route documents based on current folders' content and find misplaced documents [3]. Interestingly there are no other prototypes to our knowledge that supported automated filing in file hierarchies on behalf of users. Perhaps because file hierarchy grows incrementally and its growth is under users' control, which is not the case with email. As with email, any possible future automation feature should be implemented very carefully and let users have a control of what is happening in their information spaces.

3.6 Different visualizations of the file storage

Several above mentioned prototypes tried to change visual perspective and management practices used in hierarchy file system. These were piles on a spatial surface (Dataland, Dynapad, Archy, BumpTop) and time based grid view (Envision, Perspective Wall, MEMOIRS) and visualization of search results (VIBE).

Other prototypes used a file hierarchy to explore several visualization techniques. The information visualizer (1991) was a 3D version of Rooms system (presented later in subsection 5.5) that allowed user to visualize their information in various 3D hierarchical visualizations (such as cone trees) with a zooming interface, searching capabilities and other information retrieving capabilities (such as Perspective wall - a grid of time scale and information types [97]) [27]. Probably most known 3D file managers are SGI File System Navigator (made popular by film Jurassic Park) and 3DOSX (2002) [34]. However, there is no proof if a 3rd dimension brings any advantage over present 2D hierarchies.

More abstract hierarchy visualizations are not suitable for everyday information management, although they have some advantages for certain tasks. Treemap (1991) uses a size of files and visualizes them as rectangles where the biggest file takes the biggest space on screen [72]. There are also a 3D versions such as StepTree (2004) [19]. Another is SunBurst (2000) which radially lays items in a hierarchy, with the top of the hierarchy at the centre and deeper levels farther away from the centre [147]. These visualizations are good for spotting big files or directory depth, but are not suitable for everyday management. The same goes for other visualization algorithms of huge hierarchies such as Botanical Trees, PhylloTrees, Fractal trees, PolyPlane trees, Circular trees and others [156, 109, 69].

Content Map (2009) prototype provides an overview of "file hierarchy" in a mind map like view. It is populated from a file system but later changes (moving, multi-categorization, etc.) remain limited to the tool which forces users to manage two different structures [52].

Authors of already mentioned Contact map (see subsection 2.4 about email contact based visualization) had planed to associate files with contacts, so files could be navigated to through a social network. Such navigation would probably be good for only a small number of files that would be connected with people in contact list.

File hierarchy visualization has not received such attention as email (see subsection 2.4). This might be because files do not have consistent user friendly attributes such as email (sender, receiver(s), time, subject, body), slow growing rate of a file hierarchy and because file managers are not easily extendible. Some file types such as photos and music provide more opportunities to visualization as they have a consistent set of user friendly attributes (such as time, GPS location, author, genre, etc.). For this reason, several tools specialized for managing only photos, music and also video emerged.

Visualization of publicly available semi-structured documents and their retrieval received more attention of a research community. We already mentioned some such as VIBE and Dataland. However there are many more examples such as Document Atlas (http://docatlas.ijs.si/), Envision [111], Nirve [138], BEAD [31], ThemeScapes [169], Galaxy of news [123], etc. To our knowledge such systems were used mainly in environments where documents were not familiar to users - thus not in personal information space and will not be covered here.

4 Web bookmarks and web history

Most present web browsers have a web bookmarks manager and web history embedded. Regardless of this, people employ several other web bookmark managing techniques to retrieve previously visited web sites such as: sending emails with URLs to oneself, copying web addresses to personal web pages or documents, printing out web pages, saving web pages in a file hierarchy or writing URLs on a paper [74]. Even so, web bookmarks managing tools and web browsers' history have not significantly changed to help people to return to pages they consider important. But there has been several research prototypes developed to address this issues. This section first describes bookmarking prototypes, then history prototypes and at the end prototypes that integrated bookmarking and history.

4.1 Web bookmarks

Some early bookmarking prototypes tried to impose metaphors from the physical world. Such examples are WebBook and WebForager (1996). The first is a 3D book metaphor where pages of a book present web pages, while the second tool provides a 3D room-like environment to store and manage WebBooks [28]. Data Mountain (1998) also used a 3D inclined plane to spatially manage thumbnails of bookmarks. This helped users to retrieve their bookmarked pages based on a position and visual clues of thumbnails [128]. The problem was again the limited space and scalability. A study also showed that such 3D spatial information space used in Data Mountain leads to slower retrieval times compared to a 2D system [36]. Another study for NIRVE document system (see section 6) has also shown similar results: only under the right combination of task, user, and interface did 3D visualization result in performance comparable to functionally matched 2D and textual tools [138].

Web spy (2001) augmented bookmarks with a reminding system that monitored for changes on selected bookmarked pages and alert users if changes occurred [146]. TopicShop (2001) helped users evaluate and organize collections of web sites by providing web sites profiles (links to other sites in the collection, number of pages, images, multimedia, etc.), spatial management of thumbnails, tags (annotation), and clumsy piling [4]. Tags (2003) were also used in a multiple classification interface for managing bookmarks based on a flat set of attributes [118]. Participants in evaluation study welcomed the multiple classification functionality, which lead to the improvement in retrieval speed. However, users were tested on a preselected set of web pages, tagging often resulted in keyword listing, and some tags were duplicated due to disambiguation and misspelling. Providing "tags only system" also makes it harder to classify information that is naturally hierarchical.

One of the tools, supporting findings that users sometimes want to bookmark only a chunk (snip) of a web page and not all of it, is Snip!t (2003) (http://www.snipit.org/). It lets users store and multi-categorize notes or snips of web pages together with titles and URLs, it allows managing snips in categories and sharing snips with other users. It even goes further and suggests future actions if it recognises a type of data in a bookmarked text (such as a post code, dates, names) [30]. Similarly was a notion of integrity of information items [77, p. 10] questioned in Haystack (see subsection 5.3) and Personal Project Planner (see subsection 5.4).

4.2 Web history

Some prototypes focused on a visualization of web browsing history. As with email, acquisition of web history is automatically done by logging browsing sessions. Although management of web browsing histories lacks organization, it still allows retrieval and to some level maintenance (e.g. deleting items from a history list) of visited web pages. In recent years several browsers even allow simple one click disabling and enabling logging of the browsing session.

WebMap (1995) visualized a map of web browsing and relationships between web pages with several graphical representations where each page was represented by a small circle and types of links between pages indicated: (1) whether the destination document is located on a same or different server (green and black solid lines), (2) direct jumps to other pages on different servers (dashed blue lines) or (3) cross-jumps between nodes on the same server (dashed orange lines). The graphs produced by WebMap could be saved, for use by others [42]. Similarly MosaicG (1995) depicted a graphical left-to-right hierarchical history of web pages' thumbnails visited during a browsing session [6] Portions of this tree could also be collapsed (zoomed out). Padprints (1998) similarly introduced a zoomable tree view of web browsing histories with web page thumbnails as nodes. Users could zoom in to a particular excerpt of their browsing from an overall tree view [70]. While all these tools used a 2D representation of web history, WebPath (1998) used a 3D representations of web navigation, where pages were visualized as cubes (with thumbnail, title and other information on its surface) and connected with coloured arrows distinguishing domains [58].

Domain Tree browser (2000) visually presented trees of a browsing history thumbnails adjacent to the currently viewed page, organized by domains [61]. A user study between Domain Tree, GlobalTree (a graph based history developed by Sun) and a regular history list interface indicated that the use of visual aids is more powerful than using just text or the current history methods [87].

Footprints (1999) made use of real-world metaphors such maps (a graph of web sites), paths (a tree view history), signposts (popularity of documents by colour), etc., to expose history of use by a group of users. A controlled study showed that users were able to get the same amount of work done with less effort [165]. Browsing icons (2001) displayed animated graphs of users' browsing paths similar to Footprints' maps. Every web session built an individual graph which could be organized in a hierarchy of user-defined tasks. Users could interrupt tasks and continue later using graphs which could also be reused for similar or recurrent tasks. By organizing history hierarchically according to tasks, authors tried to cope with a scale of browsing history and provided a concept for revisitation of web pages, which had been reported as useful in a study [104].

History-Centric Browsing (2006) prototype associated pages from web history to the currently displayed page through three types of relevancies: temporal sequence, URL/location-based proximity, and content similarity. It could show these related pages during the browsing session through 3 different graphical visualizations: a statistical summary of a web history, thumbnails of relevant pages on the top of the current web page, and a temporal exploration of history based on the relevance to the currently visited page [141]. Contextual Web History (CWH) (2009) integrated several above ideas and introduced history search field in a browser allowing to search and browse thumbnails and metadata of previously visited web pages. It also integrated thumbnails of relevant web pages from the history on a search engine web page together with results returned by the search engine [171].

A visual and consistent representation web pages during the browsing session and later in a web history, has led to development of Visual snippets (2009) which where aimed to replace thumbnails of web sites. Snippets were built by extracted web site logo, title and image to easily evaluate and later recognize already visited web pages [152].

4.3 Integration of web bookmarks and web history

BookMap (2000) was a prototype that integrated bookmarks, web history (last few pages), most visited pages and pages that were marked to read later. It visualized all web references in a 2D graph which could be freely managed [66]. WebView (1999) allowed users to browse their web browsing histories in several modes: (1) a tree view based on nesting relationships between the storage location of the pages, (2) a tree view based on cross site navigational links and (3) temporal view. It also allowed visually annotating thumbnails with dogears (bookmarking) based on visiting frequency and provided a shortcut menu from any thumbnail to navigate to any subordinate child page [35].

Integration of bookmarks and web history was also used in Bookmark Pruner (2001). This prototype let users score bookmarks on an importance scale and used the web history (amount of time spent at each web page, frequency of revisits, and length of period between revisits) to automatically order and gradually change the list [146]. Similarly History Harvester (2001) [146] created a list of bookmarks from the whole browsing history as input and produced a list of web bookmarks based on user's predefined areas of interest (keywords).

Integrated System (2001) integrated the browsers back button, a browsing history, and bookmarks in one interface. Users could use a back button to access recently visited web pages presented as thumbnails, and mark (tag with dog-ears) these as bookmarking action. Users could search and filter the list using dynamic queries [78].

3D bookmark (2004) was a a bookmark system based on a 3D interface. It allowed users to browse bookmarks and history (although separately but in a consistent interface) in three listing modes: a book mode, a circular mode, and a cube mode. A slide bar and buttons were used for moving between cached web pages and all three modes previewed previous and following pages, displaying several pages simultaneously [172].

Present web browsers made bookmarking easier with one mouse click, automatically piling web bookmarks in a root folder, better search tools over bookmarks and synchronizing bookmarks between different devices (web browsers). Some browsers also allow tagging in conjunction with a hierarchy (e.g. Firefox). Tagging is also used in some standalone applications such as miTaggedMarks and on many social bookmarking sites such as Delicious (2003) which provide an online personal and collaborative classification and taxonomy of web bookmarks.

A time ordered history list of previously visited web pages also gained a search feature and real time results when typing in browsers' address bar. Other than this it did not get any other visualization even if studies suggest that users revisit anywhere between 60% to 80% of web pages [151, 37]. It is also surprising that users use web histories in only a small percentage (in some studies even less than 1%) of revisits [164, 73, 74, 26, 5]. there are also some commercial products such as Browseback (http://smilesoftware.com/BrowseBack/), that let users visually scan thumbnails of web pages, or MindRetrieve (http://www.mindretrieve.net/) which lets users to search their web history, tag bookmarks and even files (although the development seams to be stopped).

The new Firefox feature called Panorama (previously known as Tab Candy) brings a new visual way to manage web sites in groups on a spatial surface. It allows to have opened only web sites of one group and navigate between groups of web sites. Future versions promise to support collaborative web browsing. Panorama is a kind of mashup between bookmarking and (collaborative) browsing (or performing tasks), using several ideas form many mentioned prototypes in this paper (virtual desktops, zooming interface, piling, etc.).

5 Information integration

Information fragmentation is considered one of the main obstacles in cross-tool and cross-device information management [22, 120] and has since been a focus of the PIM research community. Some of the so far mentioned prototypes provided some level of integration. Such prototypes are Raton (email, calendar, to-do list and notes – see subsection 2.3) and TaskMaster (emails, URL's and attached files – see subsection 2.3). But their focus was on one information type while providing the support for managing non native information types as well. They use a standard user interface and embed the support for other information types. Contact Map's idea was to integrat contacts with several types of communication (email, instant messaging, telephone) and files; although only email was implemented (see contact based visualization in subsection 2.4). On the contrary, this section provides an overview of information integration through other means. First are presented applications that provide integration through searching (5.1), followed by applications that integrate through context, time and tags (5.2), semantinc desktops (5.3), task support (5.4), display (5.5), and integration of physical and digital worlds (5.6).

5.1 Integration through search and associative browsing

Probably most known and used cross-tool PIM applications are desktop search engines. One of the first such prototypes was Stuff I've seen (2003) (SIS) which provided a keyword search over several kinds of information types such as files, email, calendar entries, contacts and visited web pages. It also showed users rich contextual clues with returned results such as parts of the text and attributes and it also provided a faceted search over the list of returned results [48]. It was later upgraded by Milestones in Time (2003) with showing results on a timeline together with public (important events, holidays) and personal milestones (events from calendar, photos) to support episodic memory [127]. SIS also lead to a new desktop search engine Phlat (2006) which merged search and browsing through various associative and contextual clues (people, tags, time, hierarchy path) [39].

Semex (SEMantic EXplorer) (2005) enabled browsing personal information by automatically created semantically meaningful associations [44]. Retrieving information by association has lead to development of Quill (2008) [64] and Feldspar (2008) [32]. Quill is a so called narrative-based interface where user fills in predefined sentences (tells a story about a document) and the system returns a list of thumbnails of related documents. Feldspar is very similar but its interface is not based on a text story as Quill's. Feldspar lets users interactively link predefined sets of information types presented in columns with associations.

Although powerful desktop search engines are a part of our computer lives only small percentage of people use them regularly and as such can not be an answer to a cross-tool information management [18]. It is also hard or impossible to retrieve for example all information items related to one particular task with one search query and such queries would probably require high cognitive load. These search engine are good at retrieving a particular item given a right set of searched keywords. Associative retrieval can take a lot of time with present interfaces and it probably has a prominent future integrated with a better voice recognition.

5.2 Context, time, tags

Time proximity can help as a reminder in information retrieval as we mentioned above with Milestones. Although its focus was more on search rather than information management (see subsection 5.1). Lifestreams (1996) was a time-centred information integration application. It showed several types of information items (such as files, email and calendar inputs) in a timely ordered document stream that replaced the conventional file hierarchy. Stream could be managed, organized and summarized by applying different filters to it. Old documents moved towards the tail of the stream providing simple archiving function (note that archiving is one of the activities users are not good at) similarly to what TimeScape did for files (see subsection 3.4). Present documents and new email were in front of a stream while calendar events and in future placed documents served as reminders [59]. One drawback of Lifestreams is that it replaced one superordinate aspect of the document (its location in the hierarchy) with another (its location in the timeline). If we compare it to the above mentioned TimeScape, the later demands from the user to remember both the time and the hierarchy location to find a desired document.

Some prototypes provided context over several information types showing related information. Webtop (2002) showed a list of documents related by various types of relationships (the same folder, bookmarks, direct link (inward and outward), and similar content) to a currently opened document [170]. Haystack also provides related information to a current object in view (contacts, email, documents or other information objects) but its focus is on semantic desktop and will be presented in next subsection 5.3 . Project Infospace (2003) replaced desktop metaphor and implemented unified handling of all types of documents (text, pictures, emails, etc.), multiple classification of documents (i.e. in various contexts), bi-directional links between documents and groups of documents (to manage semantic associations and access related information), visualizing temporal evolution of documents, etc. [121]. Implicit query (2004) also generated a list of related information items to the users' current activities by context-sensitive search [50].

Attributes or tags were a major feature in a prototype named Presto (1999). It was a part of the Placeless Documents project and allowed spatial management of information and their (automatically or user defined) attributes. Dynamic groups could be formed by search queries or manually by users and could be further refined by dragging additional attributes on them which would change their content accordingly. Presto used piles as graphical presentation of information [45]. KnownSpace (1999) extended a before mentioned file limited DomainView (see subsection 3.1) with allowing users to set tags on other information types as well (files, emails, contacts, web pages, etc.). A step forward in dynamic documents' structures are desktop applications such as Haystack, Gnowsis and its successor NEPOMUK which will be presented later (see subsection 5.3).

Poyozo (2010) is also based on tags or keywords which are automatically aggregated from personal information of one's online activities (such as email, browsing history, social networks, blogs, etc.) [106]. It provides an overview or reflection of users online life on a calendar interface and does not provide any management functionalities. Poyozo is available as a web browser's extension.

Unification through time, tags or context is not present in mainstream information managers. Tagging is supported in several applications separately, which use different approaches as there is no standard to follow. Lifestreams was released as commercial application called Scopeware, but it never became widely used as a desktop application. Although lifestreamig is used by several other tools such as: social networking web services and blogs. Social networks aggregators also unify information from different social networks and present them in a time ordered stream of documents. Similar aggregation is done in Vodafone 360 for all communication streams of contacts from a unified contact list (email, social networks, phone).

5.3 Semantic desktop

Haystack (1999-2005) was a semantic desktop application that replaced all existing PIM tools with its own interface to overcome all drawbacks of the former. It provided unified data environment with a uniform namespace for referring to all individual information objects. These objects could be grouped (in collections with drag and drop), annotated, and linked in any possible way, giving users total control over the organization. RDF was used for naming objects and relationships between them. Objects could be smaller that actual information items, as Haystack shredded typical information items (files, emails, web pages, contacts, etc.) into many individual information objects that were connected through application specific relationship. The granularity of information was also addressed in previously mentioned Snip!t[30] (see section 4.1) and later discussed PPP[75] (see subsection 5.4).

IRIS (2005) also provided an integrated environment of different information types. But instead of building new tools for managing data it used existing tools and integrated them in to a coherent interface, allowing to annotate, classify and display related information [33]. DeepaMehta (2005) aim was to integrate information from separated desktop applications into an integrated workspace with a graph-based user interface (Topic Map). It enabled the user to organise, describe and relate information objects such as text notes, documents and media, browse the web and create semantic networks. It provided yet another tool and another organization(s) to manage [126].
.

Gnowsis (2003-2007) also brought semantic web features to the desktop with a unified RDF graph of all personal information [135]. It aimed to complement established desktop applications and the desktop operating system. In addition to providing an interface for managing personal information, it also provided interfaces for other applications to access this information. The central part of gnowsis was the personal ontology handled in Personal Information Model PIMO that also allows to annotate or tag information. The same ontology system (data format for describing ontologies) was used in gnowsis' successor NEPOMUK (2007-2009). This semantic desktop system introduced collaborative work and communications without introducing new interface (such as Haystack), but rather exploit existing applications.

The use of personal ontologies was also implemented in other prototypes besides semantic desktops. One such prototype is DELOS Task – a task-centred information manager [30]. Its aim was to automate user's most frequent activities, by exploiting the collection of personal information and compute task inferences. Task based information integration prototypes are presented in the next subsection.

To foster standardization and interoperability between different implementations, the community around the NEPOMUK project founded the OSCA Foundation whose standards are now used by KDE, gnome, freedesktop and Nokia's Maemo. NEPOMUK is getting integrated in KDE - K desktop environment (which exist for all major operating systems including Linux, Windows and OS X). Although the development is very slow, it is already possible to tag, rate, and comment files in a KDE file manager Konqueror and search this information with a search engine. The future versions plan to extend semantic desktop features to all KDE applications. Another NEPOMUK based application is a Gnome project called Tracker which is a search engine based on an ontology created using NEPOMUK (http://projects.gnome.org/tracker/). DeepaMehta project is also still developed (http://www.deepamehta.de/) while the development of Open IRIS prototype has ceased (http://www.openiris.org/).

A KDE project worth mentioning here is also Akonadi. Although not related to semantic desktop, it provides a unifying storage and retrieval services of data and metadata to every PIM application that traditionally have different data storage and handling methods.

5.4 information integration through task support

Task management has long been a focus of researchers. One of the first prototypes that addressed this with a new interface was the Personal Role Manager (1994) [142]. PRM allowed users to manage multiple technological formats under the dominant organizational dimension of roles (and their associated tasks). Roles could be broken down into a task hierarchy represented as windows in which relevant information could be stored. The role and task descriptions were also made available in a calendar. However, like many other PIM-unification systems, no evaluation has been reported.

WorspaceMirror (2003) helped users maintain similar hierarchies across applications with mirroring folders across hierarchies. Although the prototype was not meant to be specifically used for tasks, researchers found out that mirrored folders tend to be task related [23]. It was not used much as authors predicted, and they argued that users employ different management techniques in each hierarchy (e.g. not filing emails because of its amount).

Another prototype that used task as a central paradigm was UMEA (2003) [81]. UMEA let users define high level tasks. At a certain time only one task could be selected and all information items that were opened/viewed/edited during this time were automatically stored under selected task. UMEA also automatically ordered information of each task based on the time spent working on each information item. In a very similar fashion, TaskTracer (2005) monitored users' activities and documents usage and with user's assistance associated these activities to tasks. Users could later access records of past activities and restore task context [46]. The prototype gained automatic task prediction with a TaskPredictor [140]. It attempted to automatically detect task shifts which worked only with lower-level, specific tasks done in support of larger tasks.

TaskVista (2006) was a to-do list which could be created with typing in to-dos or by dragging-and-dropping emails or files to it. It is a part of a bigger application called ACTA (Activity-Centered Task Assistant) which creates a task centred environment with separated folders for documents, attendees, correspondence, contacts etc. [15].

A different approach to task management was introduced with Personal Project Planner (2005) or PPP (also called Universal Labeler which was an environment for supporting and testing PPP [76]). PPP let users write a project plan in a word processor with headings, subheading, etc. which were automatically turned into a file hierarchy in the background. Text from web pages and emails could be draged-and-droped in a document and links to original sources were automatically created making the possibility to return to original content and context of information. It even let users write emails from the document and automatically link them.

Ideas from UMEA and TaskTracer were later used in several time tracking applications (e.g. Chrometa). Some task support was included in Windows XP Explorer via task pane on the left side of the window containing a list of common actions and destinations that were relevant to the current directory or file(s) selected. It did not integrate information, however it provided some online services in a file manager. This was removed in later version in Windows 7.

5.5 Display integration

Display is also one mean of information integration. Users can have information of different types together on a screen managed by separate applications to support their tasks. Management of application windows on a display screen was always a challenge. With several applications opened while working on different tasks, computer screen easily becomes cluttered.

Rooms (1986) introduced virtual desktops [68]. It supported sharing the same information on different desktops, an overview of desktops, saving and restoring desktops. The Task gallery (2000) prototype was a 3D window manager. Application windows could be grouped by tasks and hanged on walls, ceiling and floor. User could move freely in a room where the stage in this room was used for windows needed in the current task [130]. Another task supporting application is GroupBar (2003) which allows users to group applications related to a task in groups on a windows taskbar. As a result, users can also switch between tasks' groups and not separated applications only [144]. Scalable fabric (2004) let users group and minimize application windows as they were dragged to the screen edge [129]. Wincuts (2004) was also developed to fight windows clutter but rather then ”extending” the available screen space, it focused on cropping application windows to show only relevant information for a task at hand [150].

The main problem with all these applications is their temporal nature as the position of windows, groups and excerpts are lost after a computer restart. They proved helpful in studies and might be an answer to information management on a display and display clutter of shot lived tasks. Several window managers and desktop environments are available especially in the open source domain to help uses manage their information on the screen in relation to tasks, projects, activities. Virtual desktops exist in some of the present operating systems by default (OS X, Gnome, KDE) and some even allow to save the state of these desktops for later reuse. Some window managers and desktop environments also let users group application windows title bars as tabs (e.g. Fluxbox, PWM, KDE, Ion, Enlightenment) so windows related to a task can be tabbed together. But such solution visualizes only one task related window at a time while other windows are obscured. Some also let users use tiling and tabbing to maximize screen estate usage and tabbing windows (e.g. Ion, KDE). Some applications let stick windows as tabs on the edge of the screen (such as Sticky Windows). Above mentioned GroupBar was released as an addon for Windows XP allowing to have ”multiple taskbars” (similar to tabbed windows), however the new Windows 7 taskbar uses grouping based on applications and not tasks.

5.6 Integration of digital and physical information space

It was intriguing to bring capabilities of digital world to managing documents in a physical space and vice versa, ever since the personal computers became popular. There are mainly three approaches that address this issue: (1) copy a physical state in a digital world and run both copies simultaneously, (2) bring digital documents to physical space with augmented reality or (3) bring digital management capabilities to physical documents. Several techniques were used to achieve this from imitating physical file cabinets in the electronic environment, monitoring physical and/or digital activities, augmented reality, RFID, etc.

Portofoil (1994) was and electronic file cabinet to manage scanned physical documents in an digital environment to enhance their manipulation with digital capabilities [119]. Similar was a self-organized file cabinet (1999) that enhanced a physical file cabinet with a digital version in where documents were scanned, OCRed and suggested a position on behalf of a user [92]. The user could query the system based on any combination of text and colour using a natural language.

One of the first prototypes that monitored users' activities was DigitalDesk (1992). It was a system that interpreted the activities on an office desk (monitored with a camera) to provide digital facilities to physical paper management [110]. Magic Touch (2001) monitored a kind of wearable tags attached to user's hands and all physical documents. It tracked all movements in the physical space, and it recorded them in a database [114]. The number of tags, allowance of moving only one document at a time and the fact that only users with tags on their hands could move these documents were major drawbacks of this system. Kimura (2002) monitored activities in digital and physical environments (through sensors), interpreted working contexts and automatically assigned user's activities and documents to different working contexts. Working contexts were displayed as a montages of images on a wall. These montages helped users to quickly return to a working context they want to work on [163].

A wish to manage digitally enhanced documents in a physical form led to several ideas. Most common solution is through tagging [113, 7]. But there were also other technologies used. PaperSpace (2006) adds a code and a command bar on each printed document from a digital system [145]. This allows users to manipulate digital documents (annotate, open, email, link, get information) using a command bar on a printed physical version with a computer vision which recognizes gestures of a user over a document.

WebStickers (1999) allowed users to manage web bookmarks in a physical world with printing them on a paper as barcodes. These could be read back and opened in a browser using a barcode reader [95]. BubbleFish (2001) on the other hand tried to bring digital documents in a physical space with augmented reality where digital documents represented as bubbles, could be dragged out of the virtual world and managed together with physical documents [53].

Some of the mentioned solutions exist as commercial products (managing documents with RFID tags). However they are not used for personal documents management and we are still far away from the "Internet of things" vision.

6 Other prototypes

There are several prototypes that do not fall in to so far mentioned categories because they are not associated with files, email, bookmarks or their unification. Here are divided into two categories: category of prototypes that allowed to manage arbitrary information formats, notes or information scraps, and a category of life-logging prototypes that automatically log activities and surroundings of a user. We will mention only a few of these prototypes.

The problem of taking arbitrary information notes and retrieving them has lead to some interesting ideas and interfaces. LotusAgenda (1990) was one of the first prototypes that allowed entry of free form text information without imposing a structure in advance [80]. It was also commercially available and it is now available as an open source graphical application Chandler. MessyDesk (2002) was a tool for spatially managing photos and text excerpts dragged and dropped from other applications on a free form bulletin board and which could be collaboratively shared on a MessyBoard (a large projected bulletin board) between a group of users. Jimminy (2003) was a note taking device that suggested and showed old notes based on user's physical context such as location and people being around [124]. It has been found that content of recently and currently entered notes are better indicators of which archived notes might be relevant to the user. Jurknow (2007) allowed users to enter arbitrary text entries with automatic context capture to support rich presentation and retrieval [155]. ChittyChatty (2007) was a PDA prototype that recorded speech while notes were written and enabled users to navigate the recorded voice by selecting locations in written notes [79]. List.it (2009) took an interesting approach as an extension for a web browser which lets users writing notes, dragging text from web sites and other applications, URLs , etc. although it does not preserve links to information source (which is planned in the future versions) [88].

Lifeloging is a process to log one's life with computer based devices. This field also produced few interesting interfaces to deal with big amounts of information. Forget-me-not (1994) was an memory aid prototype (running on a mobile device) that logged users' physical context to facilitate retrieval of information exploiting human episodic memory. LifeLines (1996) was a prototype for visualizing personal history of bibliographic data showing multiple facets of the records as individual time lines together with discrete events [117]. This tool was meant to visualize personal information such as health or court records (persoanl information managed by others). MyLifeBits (2002) was a system for logging user's activities, a storage and a multiple visualization interface of one's digital life. It supported collections, full-text search, text and audio annotations, and hyperlinks which replace traditional information management interfaces. It worked together with a SenseCam which took photos of the environment and few other data in time intervals or with changing a surrounding environment. A simple screen capturing prototypes PersonalStreams and PersonalStills (2002) helped augment memory for remembering past tasks performed on computers. They logged and could later show one's activities as short video excerpts or still images. Storing one's life worth of digital information is very challenging and is a focus of research projects such as Digital Lives (http://www.bl.uk/digital-lives/) and Memories for life (http://www.memoriesforlife.org).

There are certainly other prototypes that also focused on personal information of other types such as calendar ([105]), instant messaging ([51, 160]), etc. But are beyond the scope of this review.

7 Conclusion

This paper tries to provide an overview of research prototypes for managing personal information. In particular we wanted to see which ideas were transferred to mainstream everyday applications. We described several research prototypes that have been proposed to upgrade or replace present desktop metaphor and hierarchies, prototypes that brought some automation to the management, prototypes that focused on task support and other ideas that could make our information management easier. Yet these ideas very slowly trickle to our everyday life.

In the meanwhile in the last two decades the amount of available information has grown so quickly that is hard to keep track of and find relevant information at hand when needed. We are assessing information differently than we did just a few years ago because of its amount and our limited time. In the future, the storage to keep this information in our desirable and most efficient ways and forms, will have to adapt to our needs to successfully perform different roles we have in our lives.

It is hard to predict where future research and development of personal information management will direct itself. There are certainly many routes. We will probably still store a part of our information on our personal digital devices while a part of this information will be also accessible in the cloud. With the advances of web 2.0 (social networks, blogs, collaborative knowledge building, etc.) a lot of our personal information has already moved on-line and with potential of the web certainly more information will move there in the future (on-line archiving, office applications, email, etc. and mash-ups of these tools). This brings new possibilities (live introduction and testing of new features) to the research of real life information management and some prototypes already take advantage of it. A part of personal information will be shared with others and some kept private as it is today.

What about ubiquity of information, wearable information logging, embedded data and augmented reality? All these are possible paths. But PIM will probably become more personal, focused and more computer aided [136, 17] maybe with agents, contextual adaptation, task inference or in any other possible way. No matter where and how our information will be stored it will be linked and intervened with our lives more as it is today.

References

[1] D. Abrams and R. Baecker. How people use www bookmarks. In *CHI '97: Extended abstracts on Human factors in computing systems*, pages 341–342, New York, NY, USA, 1997. ACM. 1

[2] A. Agarawala and R. Balakrishnan. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1283–1292, New York, NY, USA, 2006. ACM. 11

[3] R. Agrawal, J. Roberto J. Bayardo, and R. Srikant. Athena: Mining-based interactive management of text database. In *EDBT '00: Proceedings of the 7th International Conference on Extending Database Technology*, pages 365–379, London, UK, 2000. Springer-Verlag. 12

[4] B. Amento, L. Terveen, W. Hill, and D. Hix. Topicshop: Enhanced support for evaluating and organizing collections of web sites. In *UIST '00: Proceedings of the ACM symposium on User interface software and technology*, pages 201–209. ACM Press, 2000. 14

[5] A. Aula, N. Jhaveri, and M. Käki. Information search and re-access strategies of experienced web users. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 583–592, New York, NY, USA, 2005. ACM. 16

[6] E. Z. Ayers and J. T. Stasko. Using graphic history in browsing the world wide web. Technical report, Georgia Institute of Technology, 1995. 15

[7] M. Back and J. Cohen. Page detection using embedded tags. In *UIST '00: Proceedings of the ACM symposium on User interface software and technology*, pages 159–160. ACM, 2000. 22

[8] R. Baeza-Yates, T. Jones, and G. Rawlins. A new data model: Persistent attribute-centric objects. Technical report, University of Chile, 1996. 10

[9] O. Bälter and C. L. Sidner. Bifrost inbox organizer: giving users control over the inbox. In *NordiCHI '02: Proceedings of the second Nordic conference on Human-computer interaction*, pages 111–118, New York, NY, USA, 2002. ACM. 3

[10] D. Barreau. The persistence of behavior and form in the organization of personal information. *Journal of the American Society for Information Science and Technology (JASIST)*, 59(2):307–317, 2008. 10

[11] D. Barreau and B. A. Nardi. Finding and reminding: file organization from the desktop. *SIGCHI Bulletin*, 27(3):39–43, 1995. 1, 10

[12] D. Bauer, P. Fastrez, and J. Hollan. Computationally-Enriched 'Piles' for Managing Digital Photo Collections. In *VL/HCC '04: IEEE Symposium on Visual Languages and Human Centric Computing*, pages 193–195, 2004. 11

[13] V. Bellotti, N. Ducheneaut, M. Howard, and I. Smith. Taking email to task: the design and evaluation of a task management centered email tool. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 345–352, New York, NY, USA, 2003. ACM. 2, 4

[14] V. Bellotti and I. Smith. Informing the design of an information management system with iterative fieldwork. In *DIS '00: Proceedings of the 3rd conference on Designing interactive systems*, pages 227–237, New York, NY, USA, 2000. ACM. 4

[15] V. Bellotti and J. Thornton. Managing activities with TV-ACTA: Taskvista and activity-centered task assistant. In *SIGIR '06: Personal Information Management Workshop*, volume 4, 2006. 20

[16] O. Bergman, R. Beyth-Marom, and R. Nachmias. The project fragmentation problem in personal information management. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 271–274, New York, NY, USA, 2006. ACM. 1

[17] O. Bergman, R. Beyth-Marom, and R. Nachmias. The user-subjective approach to personal information management systems design: Evidence and implementations. *Journal of the American Society for Information Science and Technology (JASIST)*, 59(2):235–246, 2008. 23

[18] O. Bergman, R. Beyth-Marom, R. Nachmias, N. Gradovitch, and S. Whittaker. Improved Search Engines and Navigation Preference in Personal Information Management. *ACM Transactions on Information Systems (TOIS)*, 26(4), SEP 2008. 2nd International Workshop on Personal Information Management, Seattle, WA, AUG 10-11, 2006. 10, 12, 17

[19] T. Bladh, D. A. Carr, and J. Scholl. Extending tree-maps to three dimensions: A comparative study. Technical report, Luleå University of Technology, 04 2004. 13

[20] S. Bloehdorn, O. Görlitz, S. Schenk, M. Völkel, and F. I. Karlsruhe. Tagfs - tag semantics for hierarchical file systems. In *I-KNOW 06: Proceedings of the 6th International Conference on Knowledge Management*, pages 6–8, 2006. 10

[21] R. Boardman. *Improving tool support for personal information management*. PhD thesis, Imperial college London, University of London, 2004. 2

[22] R. Boardman and M. A. Sasse. "stuff goes into the computer and doesn't come out": a cross-tool study of personal information management. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 583–590, New York, NY, USA, 2004. ACM. 10, 17

[23] R. Boardman, R. Spence, and M. A. Sasse. Too many hierarchies? The daily struggle for control of the workspace. In *HCI International '03: Proceedings of Interantional Conference on Human-Computer Interaction*, 2003. 1, 9, 20

[24] O. Bondarenko and R. Janssen. Documents at hand: Learning from paper to improve digital technologies. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 121–130, New York, NY, USA, 2005. ACM. 1, 9

[25] G. Boone. Concept features in re:agent, an intelligent email agent. In *AGENTS '98: Proceedings of the second annual conference on Autonomous Agents*, pages 141–148. ACM Press, 1998. 3

[26] M. D. Byrne, B. E. John, N. S. Wehrle, and D. C. Crow. The tangled web we wove: a taskonomy of www use. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 544–551, New York, NY, USA, 1999. ACM. 16

[27] S. Card, G. Robertson, and J. Mackinlay. The information visualizer, an information workspace. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 181–186. ACM New York, NY, USA, 1991. 11, 12, 13

[28] S. K. Card, G. G. Robertson, and W. York. The webbook and the web forager: an information workspace for the world-wide web. In *CHI '96: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 111–ff., New York, NY, USA, 1996. ACM. 14

[29] P. Castro and A. Lopes. Magnet mail: A visualization system for email information retrieval. In A. Butz, B. Fisher, M. Christie, A. Kruger, P. Olivier, and R. Theron, editors, *Smart Graphics*, volume 5531 of *Lecture Notes in Computer Science*, pages 213–222. Springer Berlin - Heidelberg, 2009. 5, 8

[30] T. Catarci, A. Dix, A. Katifori, G. Lepouras, and A. Poggi. Task-Centred Information Management. *Digital Libraries: Research and Development*, 4877/2007:197–206, November 25 2007. 14, 19

[31] M. Chalmers and P. Chitson. Bead: explorations in information visualization. In *SIGIR '92: Proceedings of the ACM SIGIR conference on Research and development in information retrieval*, pages 330–337. ACM, 1992. 13

[32] D. H. Chau, B. Myers, and A. Faulring. What to do when search fails: finding information by association. In *CHI '08: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 999–1008, New York, NY, USA, 2008. ACM. 10, 17

[33] A. Cheyer, J. Park, and R. Giuli. Iris: Integrate. relate. infer. share. In S. Decker, J. Park, D. Quan, and L. Sauermann, editors, *ISWC '05: Workshop on The Semantic Desktop - Next Generation Personal Information Management and Collaboration Infrastructure at the International Semantic Web Conference*, volume 175, November 6th 2005. 19

[34] R. Chin. Three-dimensional file system browser. *Crossroads*, 9(1):16–18, 2002. 13

[35] A. Cockburn, S. Greenberg, B. McKenzie, M. Jasonsmith, and S. Kaasten. WebView: A graphical aid for revisiting Web pages. In *OZCHI '99: Proceedings of the Australia conference on Human-Computer Interaction*, volume 99, pages 15–22, 1999. 16

[36] A. Cockburn and B. McKenzie. 3D or not 3D?: evaluating the effect of the third dimension in a document management system. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, page 441. ACM, 2001. 14

[37] A. Cockburn and B. McKenzie. What do web users do? an empirical analysis of web use. *International Journal of Human Computer Studies*, 54(6):903–922, 2001. 16

[38] A. Cockburn and H. Thimbleby. Reducing user effort in collaboration support. In *IUI '93: Proceedings of the 1st international conference on Intelligent user interfaces*, pages 215–218, New York, NY, USA, 1993. ACM. 3

[39] E. Cutrell, D. Robbins, S. Dumais, and R. Sarin. Fast, flexible filtering with phlat. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 261–270, New York, NY, USA, 2006. ACM. 17

[40] P. J. Denning. Acm president's letter: electronic junk. *Communications of the ACM*, 25(3):163–165, 1982. 2

[41] E. Diep and R. Jacob. Visualizing E-mail with a Semantically Zoomable Interface. In *INFOVIS '04: Proceedings of the IEEE Symposium on Information Visualization*, 2004. 8

[42] P. Dömel. Webmap: a graphical hypertext navigation tool. *Computer Networks and ISDN Systems*, 28(1-2):85–97, 1995. 15

[43] J. S. Donath. Visual who: animating the affinities and activities of an electronic community. In *MULTIMEDIA '95: Proceedings of the ACM international conference on Multimedia*, pages 99–107, New York, NY, USA, 1995. ACM. 8

[44] X. Dong. A platform for personal information management and integration. In *VLDB '05: Workshop at the International Conference on Very Large Databases*, page 26, 2005. 17

[45] P. Dourish, W. K. Edwards, A. LaMarca, and M. Salisbury. Presto: an experimental architecture for fluid interactive document spaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 6(2):133–161, 1999. 10, 18

[46] A. Dragunov, T. Dietterich, K. Johnsrude, M. McLaughlin, L. Li, and J. Herlocker. TaskTracer: a desktop environment to support multi-tasking knowledge workers. In *IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces*, pages 75–82. ACM, 2005. 20

[47] N. Ducheneaut and V. Bellotti. E-mail as habitat: an exploration of embedded personal information management. *Interactions Magazine*, 8(5):30–38, 2001. 2, 4

[48] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins. Stuff i've seen: a system for personal information retrieval and re-use. In *SIGIR '03: Proceedings of the ACM SIGIR conference on Research and development in information retrieval*, pages 72–79. ACM Press, 2003. 17

[49] S. Dumais, E. Cutrell, and H. Chen. Optimizing search by showing results in context. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, page 284. ACM, 2001. 3

[50] S. Dumais, E. Cutrell, R. Sarin, and E. Horvitz. Implicit queries (IQ) for contextualized search. In *SIGIR '04: Proceedings of the ACM SIGIR conference on Research and development in information retrieval*, page 594. ACM, 2004. 11, 18

[51] M. Eisenstadt and M. Dzbor. Buddyspace: Enhanced presence management for collaborative learning, working, gaming and beyond. *Procedings of JabberConf Europe*, 2002.

[52] K. Espenkrona and M. Svensson. Composing an evaluation framework for personal information management tools. Bachelor Thesis Report No. 2009:043, University of Gothenburg, Department of Applied Information Technology, 2009. 13

[53] D. Fallman. The bubblefish: Digital documents available on hand. In *INTERACT '01: Proceedings of IFIP TC13 Conference on Human-Computer Interaction*, Waseda University Conference Center, Shinjuku, Tokyo, Japan, July 9-13 2001. 22

[54] S. Farnham, W. Portnoy, A. Turski, L. Cheng, and D. Vronay. Personal map: Automatically modeling the user's online social network. In *INTERACT '03: Proceedings of IFIP TC13 Conference on Human-Computer Interaction*, pages 567–574, 2003. 7

[55] S. Fertig, E. Freeman, and D. Gelernter. "'Finding and reminding"' reconsidered. *SIGCHI Bulletin*, 28(1):66–69, 1996. 10

[56] D. Fisher, A. J. Brush, E. Gleave, and M. A. Smith. Revisiting Whittaker & Sidner's "'email overload"' ten years later. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 309–312, New York, NY, USA, 2006. ACM. 2

[57] S. Frau, J. C. Roberts, and N. Boukhelifa. Dynamic coordinated email visualization. In V. Skala, editor, *WSCG '05: International Conference on Computer Graphics, Visualization and Computer Vision*, pages 187–193, Plzen, Czech Republic, January 2005. (Jan 31 - Feb 4). 6

[58] E. Frecon and G. Smith. Webpath - a three dimensional web history. *INFOVIS '88: Proceedings of the IEEE Symposium on Information Visualization*, pages 3 –10, 148, October 1998. 15

[59] E. Freeman and D. Gelernter. Lifestreams: a storage model for personal data. *SIGMOD Record*, 25(1):80–86, 1996. 18

[60] X. Fu, S. Hong, N. Nikolov, X. Shen, Y. Wu, K. Xu, and N. Australia. Visualization and analysis of email networks. In *APVIS '07: International Asia-Pacific Symposium on Visualization*, pages 1–8, 2007. 8

[61] R. Gandhi, G. Kumar, B. Bederson, and B. Shneiderman. Domain name based visualization of web histories in a zoomable user interface. *Proceedings of 11th International Workshop on Database and Expert Systems Applications*, pages 591 –598, 2000. 15

[62] D. Gifford, P. Jouvelot, M. Sheldon, et al. Semantic file systems. *ACM SIGOPS Operating Systems Review*, 25(5):16–25, 1991. 10

[63] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70, 1992. 3

[64] D. Gonçalves and J. A. Jorge. In search of personal information: narrative-based interfaces. In *IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 179–188, New York, NY, USA, 2008. ACM. 10, 17

[65] J. Gwizdka. Taskview: design and evaluation of a task-based email interface. In *CASCON '02: Proceedings of the 2002 conference of the Centre for Advanced Studies on Collaborative research*, page 4. IBM Press, 2002. 5

[66] M. Hascoët. A user interface combining navigation aids. In *HYPERTEXT '00: Proceedings of the eleventh ACM on Hypertext and hypermedia*, pages 224–225, New York, NY, USA, 2000. ACM. 16

[67] J. Heer. Exploring Enron: Visual data mining of e-mail. http://jheer. org/enron/, 2005. 8

[68] J. D. A. Henderson and S. Card. Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. *ACM Transactions on Graphics (TOG)*, 5(3):211–243, 1986. 21

[69] I. Herman, G. Melançon, and M. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, pages 24–43, 2000. 13

[70] R. R. Hightower, L. T. Ring, J. I. Helfman, B. B. Bederson, and J. D. Hollan. Graphical multiscale web histories: a study of padprints. In *HYPERTEXT '98: Proceedings of the ninth ACM conference on Hypertext and hypermedia*, pages 58–65, New York, NY, USA, 1998. ACM. 15

[71] C. Horn. Natural metaphor for information visuzalization. Master's thesis, Massachusetts College of Art in Boston, 2007. 8

[72] B. Johnson and B. Shneiderman. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *Visualization '91: Proceedings of the 2nd IEEE conference on Visualization*, pages 284–291. IEEE Computer Society Press, 1991. 6, 13

[73] W. Jones, H. Bruce, and S. Dumais. Keeping found things found on the web. In *CIKM '01: Proceedings of the tenth international conference on Information and knowledge management*, pages 119–126, New York, NY, USA, 2001. ACM. 16

[74] W. Jones, S. Dumais, and H. Bruce. Once found, what then? A study of "keeping" behaviors in the personal use of Web information. *ASIST '02: Proceedings of Annual Meeting of the American Society for Information Science and Technology*, 39(1):391–402, 2002. 1, 14, 16

[75] W. Jones, P. Klasnja, A. Civan, and M. L. Adcock. The personal project planner: planning to organize personal information. In *CHI '08: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 681–684, New York, NY, USA, 2008. ACM. 19

[76] W. Jones, C. Munat, and H. Bruce. The universal labeler: Plan the project and let your information follow. In *ASIST '05: Proceedings of Annual Meeting of the American Society for Information Science and Technology*. American Society for Information Science and Technology, 2005. 20

[77] W. Jones and J. Teevan, editors. *Personal Information Management*. University of Washington Press, 2007. 14

[78] S. Kaasten and S. Greenberg. Integrating back, history and bookmarks in web browsers. In *CHI '01: Extended abstracts on Human factors in computing systems*, pages 379–380, New York, NY, USA, 2001. ACM. 16

[79] V. Kalnikaité and S. Whittaker. Software or wetware?: discovering when and why people use digital prosthetic memory. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 71–80, New York, NY, USA, 2007. ACM. 22

[80] S. J. Kaplan, M. D. Kapor, E. J. Belove, R. A. Landsman, and T. R. Drake. Agenda: a personal information manager. *Communications of the ACM*, 33(7):105–116, 1990. 22

[81] V. Kaptelinin. Umea: translating interaction histories into project contexts. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 353–360, New York, NY, USA, 2003. ACM. 20

[82] D. Karger. *Unify everything: it's all the same to me*, chapter 8, pages 127–152. Univ of Washington Pr, 2007. 2

[83] D. R. Karger, K. Bakshi, D. Huynh, D. Quan, and V. Sinha. Haystack: A customizable general-purpose information management tool for end users of semistructured data. In *CIDR '05: Proceedings of the Conference on Innovative Database Research*, pages 13–26, 2005. 10

[84] D. R. Karger and W. Jones. Data unification in personal information management. *Communications of the ACM*, 49(1):77–82, 2006. 2

[85] B. Kerr. Thread arcs: An email thread visualization. *INFOVIS '03: Proceedings of the IEEE Symposium on Information Visualization*, 0:27, 2003. 6

[86] B. Kerr and E. Wilcox. Designing remail: reinventing the email client through innovation and integration. In *CHI '04: Extended abstracts on Human factors in computing systems*, pages 837–852, New York, NY, USA, 2004. ACM. 4

[87] B. Killam. A study of three browser history mechanisms for web navigation. In *IV '01: Proceedings of the Fifth International Conference on Information Visualisation*, Washington, DC, USA, 2001. IEEE Computer Society. 15

[88] M. Kleek et al. Note to self: examining personal information keeping in a lightweight note-taking tool. In *CHI '09: Proceedings of the international conference on Human factors in computing systems*, pages 1477–1480. Association for Computing Machinery, 2009. 22

[89] B. Kwasnik. *The influence of context on classification behavior*. PhD thesis, Rutgers University, East Rutherford, NY, 1989. 3

[90] B. Lam. Mailstacker: Applying a pile metaphor for email management. Bachelors Theses, School of Information Technology and Electrical Engineering, The University of Queensland, 2005. 8

[91] M. Lansdale and E. Edmonds. Using memory for events in the design of personal filing systems. *International Journal of Man-Machine Studies*, 36(1):97–126, 1992. 12

[92] D. Lawrie and D. Rus. A self-organized file cabinet. In *CIKM '99: Proceedings of the eighth international conference on Information and knowledge management*, pages 499–506, New York, NY, USA, 1999. ACM. 21

[93] G. Lepouras, A. Papatriantafyllou, A. Katifori, and A. Dix. Time2hide: spatial searches and clutter alleviation for the desktop. In *AVI '08: Proceedings of the working conference on Advanced visual interfaces*, pages 355–358, New York, NY, USA, 2008. ACM. 12

[94] H. Liao, M. Osada, and B. Shneiderman. Browsing Unix directories with Dynamic Queries: An evaluation of three information display techniques. Technical Report CS-TR-2841, UM Computer Science Department, 1998. 11

[95] P. Ljungstrand and L. E. Holmquist. Webstickers: using physical objects as www bookmarks. In *CHI '99: Extended abstracts on Human factors in computing systems*, pages 332–333, New York, NY, USA, 1999. ACM. 22

[96] W. E. Mackay. Diversity in the use of electronic mail: a preliminary inquiry. *ACM Transactions on Information Systems (TOIS)*, 6(4):380–397, 1988. 2, 4

[97] J. Mackinlay, G. Robertson, and S. Card. The perspective wall: Detail and context smoothly integrated. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 173–176. ACM, 1991. 12, 13

[98] P. Maes. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):30–40, 1994. 3

[99] T. W. Malone. How do people organize their desks?: Implications for the design of office information systems. *ACM Transactions on Information Systems (TOIS)*, 1(1):99–112, 1983. 11

[100] T. W. Malone, K. R. Grant, F. A. Turbak, S. A. Brobst, and M. D. Cohen. Intelligent information-sharing systems. *Communications of the ACM*, 30(5):390–402, 1987. 2

[101] U. Manber and S. Wu. Glimpse: a tool to search through entire file systems. In *WTEC'94: Proceedings of the USENIX Winter 1994 Technical Conference*, pages 4–4, Berkeley, CA, USA, 1994. USENIX Association. 11

[102] R. Mander, G. Salomon, and Y. Y. Wong. A "pile" metaphor for supporting casual organization of information. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 627–634, New York, NY, USA, 1992. ACM. 11

[103] M. Mandic and A. Kerne. famailiar & intimacy-based email visualization. *INFOVIS '04: Proceedings of the IEEE Symposium on Information Visualization*, pages p14 – p14, oct. 2004. 6

[104] M. Mayer and B. Bederson. Browsing Icons: A Task-Based Approach for a Visual Web History. Technical report, UM Computer Science Department, 2003. 15

[105] T. M. Mitchell, R. Caruana, D. Freitag, J. McDermott, and D. Zabowski. Experience with a learning personal assistant. *Communications of the ACM*, 37(7):80–91, 1994. 23

[106] B. Moore, M. Van Kleek, D. R. Karger, and M. Schraefel. Assisted self reflection: Combining lifetracking, sensemaking, & personal information management. In *CHI 2010: Workshop - Know Thyself: Monitoring and Reflecting on Facets of One's Life*, April 2010. 2, 18

[107] B. Nardi, J. Miller, and D. Wright. Collaborative, programmable intelligent agents. *Communications of the ACM*, 41(3):96–104, 1998. 3

[108] N. Negroponte. Books without pages. *ICC '79: Proceedings of the IEEE International Conference on Communications*, 156(1):56.1–8, 1979. 11

[109] P. Neumann, M. S. T. Carpendale, and A. Agarawala. Phyllotrees: Phyllotactic patterns for tree layout. In *EuroVis '06: Proceedings of Eurographics / IEEE VGTC Symposium on Visualization*, pages 59–66. Eurographics, 2006. 13

[110] W. Newman and P. Wellner. A desk supporting computer-based interaction with paper documents. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 587–592, New York, NY, USA, 1992. ACM. 21

[111] L. T. Nowell, R. K. France, D. Hix, L. S. Heath, and E. A. Fox. Visualizing search results: some alternatives to query-document similarity. In *SIGIR '96: Proceedings of the ACM SIGIR conference on Research and development in information retrieval*, pages 67–75, New York, NY, USA, 1996. ACM. 12, 13

[112] K. A. Olsen, R. R. Korfhage, K. M. Sochats, M. B. Spring, and J. G. Williams. Visualization of a document collection: The vibe system. *Information Processing & Management*, 29(1):69 – 81, 1993. 11

[113] J. O'Neill, A. Grasso, J. Willamowski, B. Chidlovskii, J. Fuselier, L. Lecerf, A. Kempe, N. Cancedda, A. Sandor, A. Kaplan, et al. Rfid: enhancing paper documents with electronic properties. In *CSCW '06: Workshop Collaborating Over Paper and Digital Documents on Computer Supported Cooperative Work*, 2006. 22

[114] T. Pederson. Magic touch: A simple object location tracking system enabling the development of physical-virtual artefacts in office environments. *Personal and Ubiquitous Computing*, 5(1):54–57, 2001. 21

[115] A. Perer, B. Shneiderman, and D. W. Oard. Using rhythms of relationships to understand e-mail archives. *Journal of the American Society for Information Science and Technology (JASIST)*, 57(14):1936–1948, 2006. 6

[116] A. Perer and M. A. Smith. Contrasting portraits of email practices: visual approaches to reflection and analysis. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 389–395, New York, NY, USA, 2006. ACM. 7

[117] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. Lifelines: visualizing personal histories. In *CHI '96: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 221–227, New York, NY, USA, 1996. ACM. 23

[118] D. Quan, K. Bakshi, D. Huynh, and D. Karger. User interfaces for supporting multiple categorization. In *INTERACT '03: Proceedings of the Ninth IFIP TC13 International Conference on Human-Computer Interaction*, pages 228 – 236, 2003. 14

[119] R. Rao, S. Card, W. Johnson, L. Klotz, and R. Trigg. Protofoil: storing and finding the information worker's paper documents in an electronic file cabinet. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 180–185. ACM, 1994. 21

[120] P. Ravasio, S. G. Schr, and H. Krueger. In pursuit of desktop evolution: User problems and practices with modern desktop systems. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 11(2):156–180, 2004. 1, 10, 11, 17

[121] P. Ravasio, L. Vukelja, G. Rivera, and M. Norrie. Project infospace: From information managing to information representation. In *INTERACT '03: Proceedings of IFIP TC13 Conference on Human-Computer Interaction*, 2003. 10, 18

[122] J. Rekimoto. Time-machine computing: a time-centric approach for the information environment. In *UIST '99: Proceedings of the ACM symposium on User interface software and technology*, pages 45–54, New York, NY, USA, 1999. ACM. 12

[123] E. Rennison. Galaxy of news: an approach to visualizing and understanding expansive news landscapes. In *UIST '94: Proceedings of the ACM symposium on User interface software and technology*, pages 3–12, New York, NY, USA, 1994. ACM. 13

[124] B. Rhodes. Using physical context for just-in-time information retrieval. *IEEE Transactions on Computers*, 52:1011–1014, 2003. 22

[125] B. Rhodes and T. Starner. Remembrance Agent: A continuously running automated information retrieval system. In *PAAMS '96: Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi Agent Technology*, pages 487–495, 1996. 11

[126] J. Richter, M. V
"olkel, and H. Haller. Deepamehta - a semantic desktop. In S. Decker, J. Park, D. Quan, and L. Sauermann, editors, *ISWC '05: Workshop on The Semantic Desktop - Next Generation Personal Information Management and Collaboration Infrastructure at the International Semantic Web Conference*, volume 175, November 6th 2005. 19

[127] M. Ringel, E. Cutrell, S. Dumais, and E. Horvitz. Milestones in time: The value of landmarks in retrieving information from personal stores. In *INTERACT '03: Proceedings of IFIP TC13 Conference on Human-Computer Interaction*, pages 184–191, 2003. 17

[128] G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel, and M. van Dantzich. Data mountain: using spatial memory for document management. In *UIST '98: Proceedings of the ACM symposium on User interface software and technology*, pages 153–162, New York, NY, USA, 1998. ACM. 14

[129] G. Robertson, E. Horvitz, M. Czerwinski, P. Baudisch, D. Hutchings, B. Meyers, D. Robbins, and G. Smith. Scalable Fabric: flexible task management. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 85–89. ACM, 2004. 21

[130] G. Robertson, M. Van Dantzich, D. Robbins, M. Czerwinski, K. Hinckley, K. Risden, D. Thiel, and V. Gorokhovsky. The Task Gallery: a 3D window manager. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, page 501. ACM, 2000. 21

[131] S. Rohall and D. Gruen. Remail: A reinvented email prototype. In *CSCW '02: Demonstration on Computer Supported Cooperative Work*, New Orleans, LA, November 16-20 2002. 4

[132] S. Rohall, D. Gruen, P. Moody, and S. Kellerman. Email visualizations to aid communications. In *INFOVIS '01: Proceedings of the IEEE Symposium on Information Visualization*, 2001. 4, 5, 6

[133] W. Sack. Conversation map: a content-based usenet newsgroup browser. In *IUI '00: Proceedings of the 5th international conference on Intelligent user interfaces*, pages 233–240, New York, NY, USA, 2000. ACM. 8

[134] M. Samiei, J. Dill, and A. Kirkpatrick. Ezmail: Using information visualization techniques to help manage email. *IV '04: Proceedings of International Conference on Information Visualisation*, 0:477–482, 2004. 4

[135] L. Sauermann, G. Grimnes, M. Kiesel, C. Fluit, H. Maus, D. Heim, D. Nadeem, B. Horak, and A. Dengel. Semantic desktop 2.0: The gnowsis experience. *ISWC '06: Proceedings of International Semantic Web Conference*, pages 887–900, 2006. 19

[136] M. Schraefel, P. André, and M. Van Kleek. This Time it's Personal: from PIM to the Perfect Digital Assistant. In *CHI '08: Workshop on Personal Information Management*, volume 26, 2008. 23

[137] M. Schraefel, M. Karam, and S. Zhao. mSpace: interaction design for user-determined, adaptable domain exploration in hypermedia. In *AH '03: Proceedings of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems*, 2003. 12

[138] M. M. Sebrechts, J. V. Cugini, S. J. Laskowski, J. Vasilakis, and M. S. Miller. Visualization of search results: a comparative evaluation of text, 2d, and 3d interfaces. In *SIGIR '99: Proceedings of the ACM SIGIR conference on Research and development in information retrieval*, pages 3–10, New York, NY, USA, 1999. ACM. 13, 14

[139] R. B. Segal and J. O. Kephart. Mailcat: an intelligent assistant for organizing e-mail. In *AGENTS '99: Proceedings of the third annual conference on Autonomous Agents*, pages 276–282, New York, NY, USA, 1999. ACM. 3

[140] J. Shen, L. Li, T. G. Dietterich, and J. L. Herlocker. A hybrid learning system for recognizing user tasks from desktop activities and email messages. In *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces*, pages 86–92, New York, NY, USA, 2006. ACM. 20

[141] Y. Shirai, Y. Yamamoto, and K. Nakakoji. A history-centric approach for enhancing web browsing experiences. In *CHI '06: Extended abstracts on Human factors in computing systems*, pages 1319–1324, New York, NY, USA, 2006. ACM. 15

[142] B. Shneiderman and C. Plaisant. The future of graphic user interfaces: personal role managers. *People and Computers*, pages 3–9, 1994. 20

[143] V. Sinha and D. R. Karger. Magnet: supporting navigation in semistructured data environments. In *SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 97–106, New York, NY, USA, 2005. ACM. 12

[144] G. Smith, P. Baudisch, G. Robertson, M. Czerwinski, B. Meyers, D. Robbins, and D. Andrews. Groupbar: The taskbar evolved. In *OZCHI '03: Proceedings of the Australia conference on Human-Computer Interaction*, volume 3, 2003. 21

[145] J. Smith, J. Long, T. Lung, M. M. Anwar, and S. Subramanian. Paperspace: a system for managing digital and paper documents. In *CHI '06: Extended abstracts on Human factors in computing systems*, pages 1343–1348, New York, NY, USA, 2006. ACM. 22

[146] C. Sorensen, D. Macklin, and T. Beaumont. Navigating the world wide web: bookmark maintenance architectures. *Interacting with Computers*, 13(3):375 – 400, 2001. 14, 16

[147] J. Stasko, R. Catrambone, M. Guzdial, and K. McDonald. An evaluation of space-filling information visualizations for depicting hierarchical structures. *International Journal of Human-Computer Studies*, 53(5):663–694, 2000. 13

[148] S. Sudarsky and R. Hjelsvold. Visualizing electronic mail. In *IV '02: Proceedings of Sixth International Conference on Information Visualisation*, pages 3 – 9, 2002. 3

[149] J. Takkinen and N. Shahmehri. Cafe: A conceptual model for managing information in electronic mail. In *HICSS '98: Proceedings of the Thirty-First Annual Hawaii International Conference on System Sciences*, volume 5, page 44, Washington, DC, USA, 1998. IEEE Computer Society. 3

[150] D. S. Tan, B. Meyers, and M. Czerwinski. Wincuts: manipulating arbitrary window regions for more effective use of screen space. In *CHI '04: Extended abstracts on Human factors in computing systems*, pages 1525–1528, New York, NY, USA, 2004. ACM. 21

[151] L. Tauscher and S. Greenberg. How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies*, 47:97–138, 1997. 16

[152] J. Teevan, E. Cutrell, D. Fisher, S. Drucker, G. Ramos, P. André, and C. Hu. Visual snippets: summarizing web pages for search and revisitation. In *CHI '09: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 2023–2032. ACM New York, NY, USA, 2009. 16

[153] E. Tulving and Z. Pearlstone. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5(4):381 – 391, 1966. 3

[154] M. Tungare and M. A. Pérez-Qui nones. Mental workload in multi-device personal information management. In *ASIST '09: Personal Information Management (PIM) Workshop*, November 7-9 2009. 1, 2

[155] M. Van Kleek, M. Bernstein, and D. Karger. Gui—phooey!: the case for text input. In *UIST '07: Proceedings of the ACM symposium on User interface software and technology*, page 202. ACM, 2007. 22

[156] J. van Wijk, F. van Ham, and H. van de Wetering. Rendering hierarchical data. *Communications of the ACM*, 46(9):263, 2003. 13

[157] G. Venolia, L. Dabbish, J. Cadiz, and A. Gupta. Supporting email workflow. *Microsoft Research*, pages 2001–88, 2001. 3

[158] F. Viegas. Mountain. http://alumni.media.mit.edu/~fviegas/projects/mountain/index.htm, 2005. 7

[159] F. Viégas, D. Boyd, D. Nguyen, J. Potter, and J. Donath. Digital artifacts for remembering and storytelling: Posthistory and social network fragments. In *HICSS '04: Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, page 10, 2004. 6, 7

[160] F. Viégas and J. Donath. Chat circles. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 9–16. ACM, 1999.

[161] F. Viegas and M. Smith. Newsgroup crowds and authorlines: visualizing the activity of individuals in conversational cyberspaces. page 10 pp., jan. 2004. 7

[162] F. B. Viégas, S. Golder, and J. Donath. Visualizing email content: portraying relationships from conversational histories. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 979–988, New York, NY, USA, 2006. ACM. 7

[163] S. Voida, E. D, B. Macintyre, and G. M. Corso. Integrating virtual and physical context to support knowledge workers. *IEEE Pervasive Computing*, 1:73–79, 2002. 22

[164] H. Weinreich, H. Obendorf, E. Herder, and M. Mayer. Off the beaten tracks: exploring three aspects of web navigation. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 133–142, New York, NY, USA, 2006. ACM. 16

[165] A. Wexelblat and P. Maes. Footprints: history-rich tools for information foraging. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 270–277, New York, NY, USA, 1999. ACM. 15

[166] S. Whittaker, Q. Jones, B. Nardi, M. Creech, L. Terveen, E. Isaacs, and J. Hainsworth. Contactmap: Organizing communication in a social desktop. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 11(4):445–471, 2004. 7

[167] S. Whittaker and C. Sidner. Email overload: exploring personal information management of email. In *CHI '96: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 276–283, New York, NY, USA, 1996. ACM. 1, 2

[168] Wikipedia. Archy - the humane environment project. http://en.wikipedia.org/wiki/Archy, 2005. 11

[169] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: spatial analysis and interaction with information from text documents. In *INFOVIS '95: Proceedings of the IEEE Symposium on Information Visualization*, page 51, Washington, DC, USA, 1995. IEEE Computer Society. 13

[170] D. Wolber, M. Kepe, and I. Ranitovic. Exposing document context in the personal web. In *IUI '02: Proceedings of the 7th international conference on Intelligent user interfaces*, pages 151–158, New York, NY, USA, 2002. ACM. 11, 18

[171] S. S. Won, J. Jin, and J. I. Hong. Contextual web history: using visual and contextual cues to improve web browser history. In *CHI '09: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1457–1466, New York, NY, USA, 2009. ACM. 15

[172] T. Yamaguchi, H. Hattori, T. Ito, and T. Shintani. On a web browsing support system with 3d visualization. In *WWW Alt. '04: Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, pages 316–317, New York, NY, USA, 2004. ACM. 16

[173] K. Yee. Zest: discussion mapping for mailing lists. In *CSCW '02: Proceedings on the ACM Conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, USA, November 16-20 2002. ACM. 6

[174] K. P. Yee and M. Hearst. A visualization to facilitate productive discussions. In *CHI '05: Beyond Threaded Conversation Workshop on SIGCHI conference on Human factors in computing systems*, 2005. 6

[175] K. S. Yiu, R. Baecker, N. Silver, and B. Long. A time-based interface for electronic mail and task management. *Advances in human factors/ergonomics*, pages 19–22, 1997. 5