

From Group Memory to Project Awareness Through Use of the Knowledge Depot

Michael Kantor

University of California, Irvine
Information & Computer Sci.
Irvine CA, 92697
1 (714) 824-2703
mkantor@ics.uci.edu

Beatrix Zimmermann

Bell Atlantic S&T
500 Westchester Ave.
White Plains, NY 10604
1 (914) 644 2503
bz@nynexst.com

David Redmiles

University of California, Irvine
Information & Computer Sci.
Irvine CA, 92697
1 (714) 824-3823
redmiles@ics.uci.edu

ABSTRACT

In the CSCW literature, systems labeled as group awareness tools are generally designed to provide relatively instantaneous awareness among a group of individuals. Such a tool is provided to enhance the ability of the members of the group to communicate and coordinate with one another. This paper presents a variant of group awareness which this paper refers to as Project Awareness. This type of awareness provides individuals with awareness of a set of groups (rather than a set of individuals), over days or weeks (rather than fractions of seconds or minutes). Project awareness provides a parallel benefit to group awareness: it enhances the ability of groups to communicate and coordinate with other groups. This paper discusses an existing group memory, the Knowledge Depot, and illustrates how a group memory can be enhanced into a project awareness tool.

Keywords

Group awareness, Project Awareness, Awareness Tools, Group Memory Tools

1 INTRODUCTION

In the CSCW (Computer Supported Cooperative Work) literature, systems labeled as group awareness tools are generally designed to provide relatively instantaneous awareness among a group of individuals. This type of awareness is intended to enhance people's ability to communicate and coordinate with one another. This paper describes a new type of awareness tool that enables a different kind of communication and coordination. This tool focuses on providing individuals with awareness of a set of groups (rather than a set of individuals), over days or weeks (rather than fractions of seconds to minutes). Such a system provides a parallel benefit to group awareness: it enhances the ability of a group to communicate and coordinate with other groups. Table 1 illustrates a categorization of existing awareness tools using a categorization based on the difference between group awareness and this new type of awareness. The two dimensions are the frequency with which users receive information and the unit of observation. A tool whose unit of observation is an individual, helps users maintain awareness of a set of individuals [5, 9, 7].

A tool whose unit of observation is a group helps users maintain awareness of a set of groups [1, 3]. Information that users need to be aware of over days rather than over minutes tends to be much more conceptual and task oriented than physical. Instead of providing awareness of some physical fact through graphical or audio means (e.g. the user is in room X, sitting, talking on the phone and browsing the web), we instead provide information that is much more abstract (e.g. a person plans to work on certain tasks, the tasks have certain priorities, and certain problems are delaying them) which type of information is best represented textually, as is done in most calendars. Whereas the unit of information in a typical awareness system over minutes is an individual's current low level activity, the unit of awareness in a system aimed at groups over days is the task or set of tasks of the group. Thus, awareness of the task is what makes this type of awareness so important.

In sum, this paper is about an approach being developed to fill the Group/Days quadrant of the diagram. This paper will refer to tools that fit the Group/Days quadrant as Project Awareness tools.

1.1 Requirements For Awareness

There are three key aspects to providing an effective awareness tool:

1. information capture
2. information distribution
3. information presentation

The effectiveness of an awareness tool will largely depend upon how well it provides these three features. It must capture useful information, distribute it quickly to the appropriate people, and display the information in a meaningful manner. The Portholes system, for example, captures information by taking periodic photographs of people in offices, distributes the information to people who have told the system that they want to remain aware of certain people, and presents the information as a set of photographs on the users background.

A newsgroup system could be argued to be an awareness system, as it captures information, distributes it by making it available to newsgroup

TABLE 1. Characterizing different types of Awareness

Frequency	Unit of Observation	
	Individual	Group
Seconds to Minutes	What is a person's location and current activity? (example tools are Portholes, Piazza)	Is there a group meeting? Where? What types of tasks is the group working on? Who is in the group? (example tools are Video Windows, wOrlds).
Days to Months	What is a person trying to accomplish this week? What are a person's plans for this week? What problems is a person working on solving? (example tools are various calendar applications and distribution lists).	What is a group working on this week? What kinds of problems is the group encountering? What changes has the group made in the task the group is working on? When will the task be complete? (we seek to fill this block with our tool, Knowledge Depot).

readers, and presents it as a list of messages. However, the information capture requires users with information to deliberately start the news reader, navigate to a topic and submit the information, and distribution requires users to go to appropriate software and search for the information. While a newsgroup can be used effectively to help people be aware of various kinds of knowledge, it requires explicit work on the part of the user to go out and retrieve the information, to enter the information, and to figure out where to go to add and retrieve information. Some newsgroups are set up to automatically forward all messages to a list of people, and to receive mail. These tend to behave more like the distribution lists discussed next.

A distribution list can be considered an awareness tool over time where the unit of observation is an individual. Studies of distribution lists [10] have reinforced this perspective by showing that they make peripheral group members more aware of their group. Information capture takes place when a member writes a message. The message gets distributed to everyone on the list, and the recipients view the message with a mail reader. This mechanism has a poor information capture mechanism because it is only allowed information that people explicitly design to be distributed to the entire group. It also has a poor distribution mechanism because it forwards the information to everyone in the group instead of the subset of the group that is interested in the specific item of information. If used with restraint, this can be an effective way to make people aware of what's happening in other people's work. It is common though for such mailing lists to get flooded with discussions that people have to ignore in order to get work done.

The rest of this paper will focus on the benefits of project awareness, the Knowledge Depot group memory tool, and how the Knowledge Depot was enhanced from a group memory into a project awareness tool.

2 BENEFITS OF PROJECT AWARENESS

One reason awareness of groups is important is that it allows people to keep track of the status of the tasks being worked on by other groups. When people join together into a project group to engage in a complex task, (e.g. build a software system, organize an event, write a report, etc.) it is common for the project group to divide into a number of smaller groups, each working on a different aspect of the task (e.g. construct a software component, obtain equipment for the event, research and write a section of a report). The types of information that group members want to know about another group concerns the status of the other group's task. If the work is behind schedule, the groups whose work depends upon the delayed task need to be aware of the delay. If a group changes their design in a way that might affect the work of other groups, other groups may need to adjust their own designs to compensate. It is failures of communication that are one of the three main problems faced by Software Engineering today according to Curtis, Krasner and Iscoe [4].

Using the terminology of Curtis, Krasner and Iscoe, a powerful awareness system would provide a link between groups, where the link connects groups within an organization as well as connecting groups across organizational boundaries. Furthermore, the types of awareness information that would flow through these links would depend on the nature of the relationship between the groups. Teams working on interacting components need a flow of information about changes to the component and to the schedule for completing the component. A link between a development team and a group of users would provide developers with an awareness of the future user's concerns about the system, and carry awareness of some alternatives being considered by the developers back to the users (e.g. a link that can be especially important for developing the user interface for a system). A link between developers and designers can carry issues under consideration in

the architecture (e.g. allowing developers to give some input under certain circumstances), and carry problems encountered implementing the architecture back to the designers.

2.1 Approaches to Project Awareness

Three common strategies for distributing project awareness information are:

1. Mail information to everyone the sender can think of who might be affected, and trust that they will forward it to anyone the message missed,
2. Send information to everyone, and assume that the right people read the right part of the information, rather than treat it as junk email, or memos to be filed in a trash can.
3. Generate formal status reports (automatically or by hand), reporting only on issues that the automated system is able to understand, or that the person generating the report by hand is willing to commit to a formal document, and feels is important enough (from the author's perspective, which will differ from the readers perspective) to be worth the time to write.

The first two approaches fail as an effective distribution mechanism, because if either method is used frequently, either people will regularly fail to receive needed information, or people will ignore important information hidden within too much irrelevant mail.

The third approach requires a lot of careful work to make reports contain relevant information that can be scanned to get an awareness of what is going on, and read in depth to understand issues of interest. Such reports will contain less information than the first two strategies (which if done right is good in that it prevents information overload). In general, Reports will provide information that is better organized and easier to search and skim than the same information after being sent in bits and pieces over email. However they provide less information on the day-to-day discussions and issues and often will only report on them after they have been resolved. Furthermore, such reports would take intense amounts of forethought in order to provide the required awareness to all of the different types of groups with different types of interest. No single report can maintain all of the links to all of the groups that each have an interest in their own specific type of information about a group (though of course there are cases where all anyone cares about is a completion date).

A system that better matches the definition presented above for a powerful project awareness tool is the Information Lens [8]. The Information Lens has a mechanism for capturing information (e.g. people send email to a user account "Anyone"), a means of distribution (e.g. users specify rules that the system uses to determine which messages will be of interest to them), and a means of presentation (e.g. a person's electronic mail box).

While the creators of the Information Lens did not describe it as an awareness tool, they described many of the same benefits from their system as are defined above for a project awareness system. One example they use is distributing engineering change notices. If a plan or design changes, someone on the group making the change will send mail to "Anyone" describing the change. All people whose work will be affected by changes to this aspect of the project will have set preferences to insure that they will receive information about such changes.

This still leaves users with a limited view of the change. They know that the change was made, but will have no idea why and receive no warning that it might happen. If they set their rules to capture all of the information leading up to the change, they will end up with too much information in their mailbox.

An ideal presentation method then would be a method that presents a fish-eye view of the project [6]. At the focus of the eye is the mail sent within a subgroup directly to the subgroup's members. These people receive the entire message and are assumed to have read most of the message. Just outside of the focus of the eye are other groups whose work is connected. People on the project who are interested will receive summaries of these related group's emailed discussions. This will allow them to stay aware of changes being discussed and decided upon, and of problems delaying the group's task. The summaries would include a link allowing them to access the entire message if they need more than an overview. The proximity to the focus of a person's work determines how frequently these summaries are received. Farthest from the focus are groups working on unrelated tasks. From them a user will only receive an occasional message that is broadcast to the entire project.

The rest of this paper describes a system that meets the desired traits of an project awareness tool: information is captured about many groups over time, it is distributed daily and weekly to those who specify that the information is relevant to them, and the information presentation attempts to approximate a fish-eye view of the project allowing each user to focus on issues relevant to their own task.

3 KNOWLEDGE DEPOT

Groups that use email to announce decisions, discuss problems and their solutions, and to carry on other work related discussions produce a rich pool of information as a side effect of these conversations. The Knowledge Depot, an enhanced version of GIMMe [11], captures and organizes this information, allowing users to browse through the information to rediscover, or learn for the first time, why different decisions were made, what problems were encountered as a result of those decisions, and allowing the user to regain some of the context in which those decisions were made.



Figure 1: Knowledge Depot displaying a title frame, a controls frame, a hierarchical topic browser frame and a message list frame.

The system organizes its knowledge around a set of topics defined over time by all users of the system. The topics frame is shown in figure 1 as a hierarchical list of discussion items. A topic is four things

1. A *phrase* describing a concept, task or activity representing aspects of the group's work
2. A *place* where people go to find information
3. A *definition* of the type of information the system looks for to determine whether something belongs in the topic
4. A *destination* that people will aim their messages at in order to have the message stored correctly for later retrieval

A topic then might have a name like "Portholes." This indicates that any message that has "Portholes" in the subject line will be captured in this topic, and people browsing for messages will know to look for email discussing Portholes within this topic. People having an email conversation about Portholes then need only put "Portholes" in the subject, and CC the Knowledge Depot system for the information to be captured and put in an appropriate place. Furthermore, the hierarchical organization of topics allows a message that has "Portholes Installation" in the subject to enter the "Portholes" topic, and then enter the "Instal, Config, Distrib, Setup" subtopic.

If a new type of discussion begins, any group member can create a new topic or subtopic to capture the new type of discussion. If the terminology changes, users can change the definition of the topics. For example, if there is a topic “NYNEX, collaboration” for discussing work done with NYNEX, and NYNEX changes its name to Bell Atlantic, the topic would simply be updated to “NYNEX, Bell Atlantic, collaboration”. This allows the organization of knowledge to evolve over time as the group itself evolves.

This is very similar to TeamInfo [2], which is also an email based group memory that automatically organizes its messages. However, use of the subject line to determine a messages location rather than a keywords field in the header of an email message’s header (as is done in both TeamInfo and the Information Lens) frees users from constraints on the type of email software they can use.

The Knowledge Depot uses the same type of privacy mechanism as the Information Lens: information does not become publicly available unless explicitly emailed to a special user name. The Information Lens used the account “Anyone” to determine which knowledge should be publicly distributed, while Knowledge Depot uses a mail account named after the group to capture mail that is to be archived in the group memory.

While the Depot focuses on email, it can also process electronic newsgroups, and users can upload formatted documents (i.e. PDF, HTML, GIF, etc...). These documents can’t be searched by their contents the way mail and newsgroup messages can, but can still be

accessed by browsing topics, and by giving the system queries by date, author and/or subject. Furthermore, because users can email information and use a web page to upload documents into the depot, so can software. Software can mail the depot information such as errors, usage, status, etc... Software can also use the URL users use to upload documents to automatically upload its own documents such as charts (such as might be produced by project management software), images, and other non-textual information.

The Knowledge Depot has many similarities to a newsgroup, if the newsgroup is used as a place where people Cc mail so that it can be made available to non-recipients of the message, and so that it can be stored for a time. The Knowledge Depot though is not constrained to a single newsgroup with a linear list of messages, nor is it equivalent to a large number of newsgroups where users have to determine which group or groups are most appropriate for their needs. The depot uses a flexible set of topics where any user can at any time add a new topic, and based on the definition of the topic, all existing messages will be checked to see if they belong in the topic (in addition to any other topics that the message is already in). Potentially, the entire topic structure could be removed, and replaced with a new topic structure, with messages automatically reorganizing themselves to account for the new hierarchy. While a newsgroup acts as an oversized shared email box, Knowledge Depot acts as a shared database. The database is used to organize all of a groups discussions and documents, and provides

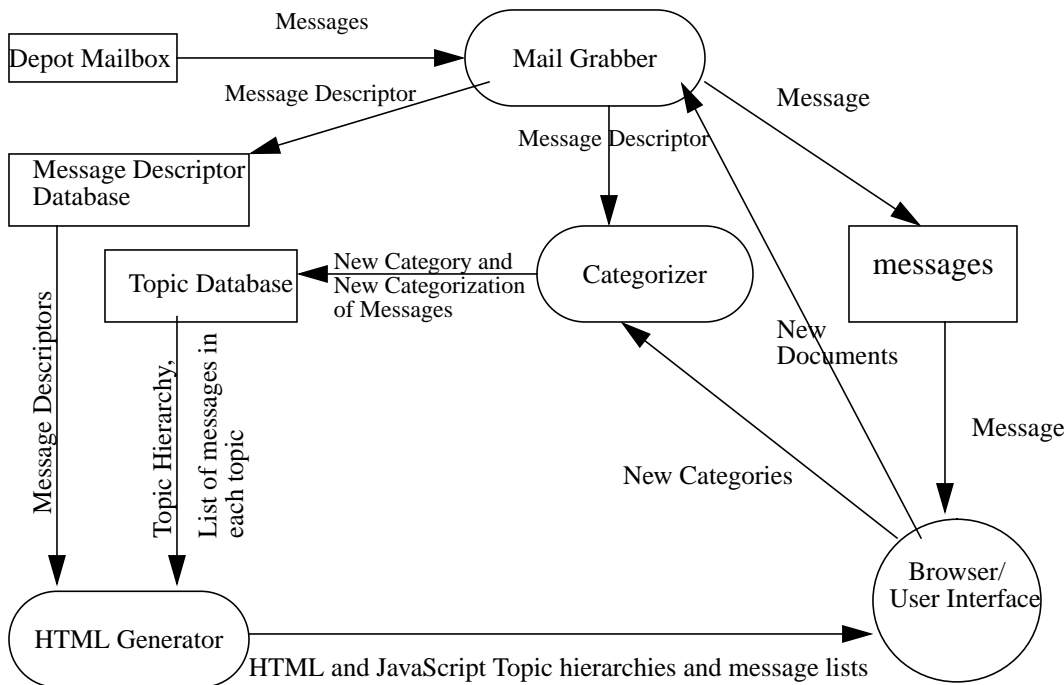


Figure 2: Knowledge Depot Architecture. Arrows indicate flow of data from one component to the next.

standard types of database queries and permits dynamic reorganizations.

3.1 Implementation

Knowledge Depot has four main components implemented in PERL as CGI scripts and in Javascript as features of the user interface: Message Grabber, Categorizer, HTML Generator and Browser.

The Message Grabber takes messages from the Depot's mail box once every hour and takes documents sent by the user through the Browser. For each message or document, it creates a file, makes an entry into a database describing the document and then sends the description to the categorizer.

The Categorizer is in charge of categorizing messages into the topic hierarchy. It takes new messages captured by the Message Grabber and inserts the path to a message into each topic that it matches. In cases where the user adds or modifies a topic, the categorizer recategorizes all messages in the Depot to account for the new topic hierarchy.

The HTML Generator takes descriptions of messages and of the topic hierarchy and sends it to the Browser as HTML or javascript. If the user is using an appropriate web browser, the topic hierarchy is sent as a javascript. What is a group working on this week? What kinds of problems is the group encountering? What changes has the group made in the task the group is working on? When will the task be complete? (we seek to fill this block with our tool, Knowledge Depot).d to allow users to browse through the topic hierarchy on their local machine. Other web browsers receive the topic hierarchy as plain HTML. This component provides the Browser with all of the information and javascript functions needed to make the system usable.

The Browser uses a web browser to display all information and uses javascript functions to allow the user to execute various browsing and editing commands. Javascript allows for more complex relations between various frames and windows of the system than HTML, and allows for significantly more local interaction before accessing the web server.

There are three main data components:

1. A folder containing all of the messages and documents, allowing easy access to the documents via the web
2. A database containing descriptors of all messages, used by the categorizer to recategorize messages, and by the HTML generator to sort and query messages by author, date or subject, and to get the information that needs to be displayed in the Browser about each message.
3. A database containing the topic hierarchy and which messages were in each topic. This is used by the HTML generator to quickly find out which messages are in the topic it is preparing to print.

3.2 System Usage

Versions of Knowledge Depot have been in use at NYNEX Science & Technology and the University of Colorado for three years, and at UC Irvine for one year. The system has had varying levels of success depending upon the situation in which it was used. Its main successes have been in geographically distributed groups and in groups where usage is mandated. For distributed groups, email is generally the primary means of communication, and a central archive of messages provides a single location for people at all locations to go for information. The one group of users where usage was mandated was a group that used the system to help become ISO 9001 compliant. The group memory was used to capture and categorize messages containing meeting minutes, plans, reviews, reports and other information that can be used to document a process for an ISO auditor. Other user groups have arranged for the system to capture all mail broadcast within the group to guarantee that information is not lost, but rarely use the information, and are generally no longer aware that the information is being captured.

Other potential uses of the system are

1. a group memory to capture and organize all communications between customers and developers
2. a group memory to capture automatically generated information such as back-up logs, error messages, and data sent by systems to their designers to report on status and usage
3. a newsgroup

These uses will become more interesting when combined with the project awareness discussed next.

4 FROM GROUP MEMORY TO PROJECT AWARENESS

Turning the Knowledge Depot into an awareness tool is made easy by the fact that it already has a means of capturing and organizing information. All that is required is for people to indicate that a topic is of interest to them, and of how much interest. The greater their interest, the more frequently they will receive summaries on new information that arrives in that topic. Future versions may also provide more detailed information about the contents of a topic for users who are especially interested in a topic. Direct email is still used to discuss a topic with group members, arriving immediately in full detail in the mail-boxes of the group members, but summaries of information generated by other groups can now arrive in users mailboxes at an interval reflecting the level of their interest.

Figure 3 shows the current subscription method. Users select a topic or subtopic in the topic browser, and select "Add Subscriber" from the operations pull-down menu. This presents them with a form for describing the subscription. People choose between receiving header information (e.g. author, date, subject and URL to the message), or the entire message. They also

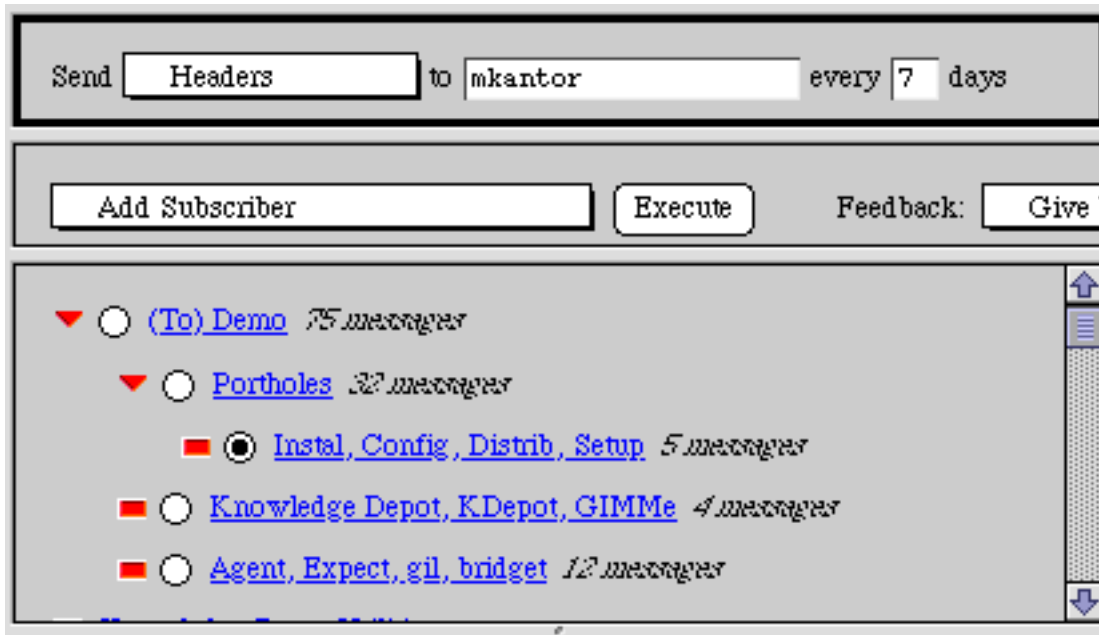


Figure 3: Subscribing mkantor to receive weekly updates of the “*Instal, Config, Distrib, Setup*” topic

choose who the updates will be sent to (e.g. usually themselves, unless they are subscribing someone who does not have access to the system such as a customer who wants to remain aware of the status of some part of the project).

The key advantage that a system like this has over any Information Lens type of mail distribution system is that it is built upon a group memory. As a group memory, this system gains two advantages in awareness:

1. Users can get a better approximation of a fish-eye view of what is going on, but can still get the full information on items of interest (e.g. by following a hyperlink into the group memory).
2. Users can note things that might be relevant later, and when it becomes relevant, they can browse through the group memory to find the item and all related information and context.

Add to that advantage a strong flexibility in the types of information that can be displayed (any information viewable over the web), flexibility of environment (users can use any email software and most web browsers) and flexibility of information sources (email sent by users and by automated systems, formatted documents uploaded by users and by automated systems), and this technology can be seen to fit in a wide variety of work environments.

The updates keep people aware of issues, and of the knowledge stored in the Knowledge Depot itself. A programmer may receive a warning that a component the person depends upon may be about to change, or has changed. A manager may receive updates describing various bug reports, giving the manager a

general impression of how many bug reports are coming in. A developer may receive summaries of email passed between marketing and the customer helping the developer keep track of what kinds of promises are being made or discussed. Because these summaries depend on email subject lines, they may often be uninformative, but if the topic itself is descriptive, and there is a lot of activity taking place in the past week in a topic, then the quantity of activity itself is informative. It tells the recipient of the summary that a topic of interest is currently under heavy discussion, or there are a lot of bug reports coming in, or that some change concerning this relevant topic is about to take place.

5 RESEARCH DIRECTIONS

This research looks at a number of issues. One of the issues is to determine if the concept that is the basis for this paper is valid; is project awareness a meaningful and useful classification for awareness tools, and are the believed benefits presented above actually derived from such a tool? The other main issue is to see how closely Knowledge Depot approximates the main concepts of an project awareness system. An aspect of that question is to determine if one can generalize the changes made to the Knowledge Depot group memory to group memories in general.

An aspect of the system that needs to improve to achieve these research goals is the summary mechanism. While message subjects can provide useful information, it unlikely to be enough to provide consistently meaningful summaries of activity. Better summaries can be produced by displaying the first few lines of text (e.g. as is done in web searches), or analyzing the message to extract summary information.

We are currently testing this awareness system at NYNEX and at UC Irvine. It is our hope that the system will enhance the ability of the user groups to coordinate their efforts by enhancing awareness of the activities that might cause breakdowns in intergroup coordination.

ACKNOWLEDGEMENTS

Various aspects of this effort are supported by the National Science Foundation under Contract Number CCR-9624846 and by NYNEX/Bell Atlantic.

REFERENCES

1. Abel, M. (1990). Experiences in an Exploratory Distributed Organization. In e. a. Galegher (Ed.), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, Lawrence Erlbaum Associates, pages 489-510.
2. Berlin, L. M., Jeffries, R., O'Day, V. L., Paepcke, A., Wharton, C. (April 1993). Where Did You Put It? Issues in the Design and Use of a Group Memory. CHI '93. Conference proceedings on Human factors in computing systems, pages 23-30.
3. Bogia, D. P., & Kaplan, S. M. (1995). Flexibility and Control for Dynamic Workflows in the wOrlds Environment. Proceedings of conference on Organizational computing systems, pages 148-159.
4. Curtis, Krasner, and Iscoe (1988). A field study of the software design process for large systems. *Communications of the ACM*, 31(11):1268-1287, November 1988.
5. Dourish, P., & Bly, S. (1992). Portholes: Supporting Awareness in a Distributed Work Group. CHI '92. Conference proceedings on Human factors in computing systems, pages 541-547.
6. Furnas, G. W. (1986, April 1986). Generalized Fish-eye Views. CHI '86. Conference proceedings on Human factors in computing systems, pages 16-23.
7. Isaacs, E., Tang, J., Morris, T. (1996). Piazza: a desktop environment supporting impromptu and planned interactions. In CSCW '96. Proceedings of the ACM 1996 conference on Computer supported cooperative work, pages 315-324.
8. Malone, T. W., Grant, K. R., Lai, K.-Y., Rao, R., & Rosenblitt, D. A. (1989). The Information Lens: An Intelligent System For Information Sharing And Coordination. In M. H. Olson (Ed.), *Technological Support for Work Group Collaboration*, Hillsdale NJ: Lawrence Erlbaum, pages 65-88.
9. Mantei, M. M., Baecker, R. M., & Sellen, A. J. (1991). Experiences in the Use of a Media Space. CHI '91. Human factors in computing systems conference proceedings on Reaching through technology, pages 203-208.
10. Sproull, L., & Kiesler, S. (1991). Increasing Personal Connections, *Connections: New Ways of Working in the Networked Organization*, MIT Press, pages 79-101.
11. Zimmermann, B., Lindstaedt, S., & Girgensohn, A. (1997). Growing Group Memories in the Workplace: A Field Study of GIMMe. Technical Report, White Plains, NY: NYNEX Science & Technology.