Towards an Infrastructure for Knowledge Communities

A position paper for the ECSCW '99 Workshop: "Beyond Knowledge Management: Managing Expertise"

Thomas Erickson and Wendy A. Kellogg
IBM. T.J. Watson Research Center
snowfall@acm.org; wkellogg@us.ibm.com

Abstract. In this position paper we summarize our efforts to design, implement, and deploy the infrastructure for conversationally-based knowledge communities. We believe that managing knowledge or expertise really means providing an on-line workplace within which users can engage socially with one another, and, in the process, discover, develop, evolve, and explicate knowledge relevant to shared projects and goals. Our basic approach is to design systems which, by making users and their activities visible to one another, can make 'knowledge work' visible, thus increasing its importance to coworkers and to the organization as a whole.

Introduction

Knowledge management is often seen as a problem of capturing, organizing, and retrieving information. Given this perspective, it isn't surprising that knowledge management evokes notions of data mining and text clustering and databases and documents. This is not wrong, but it is only part of the picture.

For example, one of us once interviewed accountants at a large accounting and consulting firm about how they might use a (proposed) database of their company's internal documents. A rather startling theme emerged from the interviews: the accountants said that they'd love to access the documents — so that they could find out who wrote them. As one explained, 'Well, if I'm putting together a proposal for

Exxon, I really want to talk to people who've already worked with them: they'll know the politics, and the history, and they can introduce me to their contacts. None of that gets into reports!'

There are two points of importance here. First, the accountants wanted to use a *data access* system to access *the people who produced the data.* It was only through the people — and the social networks they were part of — that the accountants could get the knowledge they really needed: often, the most important knowledge was too sensitive to be written down. The second point is that it was not just knowledge or expertise that the accountants were accessing: they were also getting access to *social resources* such as *contacts* and *referrals.* The accountants explained that the very worst way to approach a company with a proposal was by making a 'cold call'; it is much preferable to call someone and be able to say 'my colleague, Jil Smith, suggested I chat with you.' Being able to say that one was referred by a mutual acquaintance, to invoke a shared relationship (even a very tenuous one) is frequent and powerful facilitator for interpersonal interaction. Such social resources cannot, however, be embedded in databases. Calling someone and saying, 'Hello, I found your name in the corporate knowledge base' just isn't the same.

This sort of situation is not the exception, it is the rule. Knowledge — whether it be of bugs in the Java Virtual Machine or how to begin negotiations with an executive from another culture — is deeply social. Our position is that one can't isolate knowledge from its social context without denaturing it, without stripping it of the social resources which contribute to its utility. Instead, we believe a more productive approach is to create a context within which both knowledge and the social context within which it is created, modified, and disseminated, can flourish.

What's Next

In the remainder of this paper we describe a project whose aim is to design the infrastructure for knowledge communities, on-line workplace environments which can support long-running, deep, productive conversations. We begin with a vision of where we're headed, including its basic rationale. Next we summarize where we are in the project: we describe a prototype that we've implemented, called "Babble"; and we discuss the way in which we've come to use it as part of our daily work practice, as well as our experience in deploying it to other work groups. Finally, we summarize some of the questions and issues we've encountered and which we hope to pursue in the workshop.

Where We're Headed: Knowledge Communities

Imagine a knowledge management system which was designed from a social perspective, a system predicated on the assumption that knowledge is distributed throughout a network of people and that only a small proportion of it is captured in concrete form. As the above vignette suggests, such a system would need to provide a rich set of connections back to the social network of people who produced the information. In fact, we think such a system ought to encompass these social networks.

Given this view, additional possibilities suggest themselves. Imagine that the knowledge management system provided access not only to creators of knowledge, but to people who were accessing and using the knowledge. Suppose that — just as we look for busy restaurants, notice crowded trade show booths, or are drawn to engaging conversations — we could see similar traces left by those using information in a knowledge management system. We could notice popular knowledge sources, encounter other users with similar interests, and perhaps get glimpses of how knowledge was being re-purposed. That is, because users often must do a lot of work to adapt knowledge to their own ends, they develop an understanding of its shortcomings and particularities (as well as building on it) which would be very valuable to others engaged in similar efforts. Such a system would not be just a database from which workers retrieved knowledge, it would be a knowledge community, a place in which people discover, use, and manipulate knowledge, and encounter and interact with others who are doing likewise.

A knowledge community of this sort has a formidable social problem to overcome: *Why should those who produce and use knowledge take the time to engage in such interactions? Why should they want to? What personal benefits would they gain from sharing their knowledge?* The general solution that we are exploring in this work has to do with making people and their activities visible to one another. That is, by making knowledge work visible it becomes possible for it to be recognized and rewarded by the organization, and knowledge work can shift from something that takes time away from 'real work' to being 'real work' in and of itself.

Where We Are: The Babble Prototype

As a working system, the Babble prototype has been in existence for about two years, although it has evolved considerably over that time. Essentially (we will describe the system shortly), Babble is an on-line space which allows people to carry out multiple text-based conversations which may be either synchronous or asynchronous.

Our laboratory group has used Babble as part of our daily work practice over the last two years, and we have also deployed it to nine other groups with IBM. In all cases we've gathered logs of conversation, server activity, and (in some but not all cases) done field work involving surveys, interviews, and participant observation (a portion of this field work is described in our ECSCW paper [1]).

In terms of the implementation of the knowledge community infrastructure, there are three phases of work: conversation support; reuse of conversation; and creation of an organizational knowledge space. At present, we have implemented and deployed much of the first phase (described in [2]); that latter two phases -- which are crucial to supporting knowledge communities -- remain to be done.

What’s Been Implemented

In pursuit of our goal of creating a system in which knowledge work is more visible, Babble does two sorts of things. First, as a number of systems do, it results in the creation of a persistent trace of the textual conversation. These traces give the system the potential to function as a knowledge store, or what we prefer to call a ‘discourse base.’ Second, Babble makes the presence and activity of the participants visible by means of what we call the social proxy.

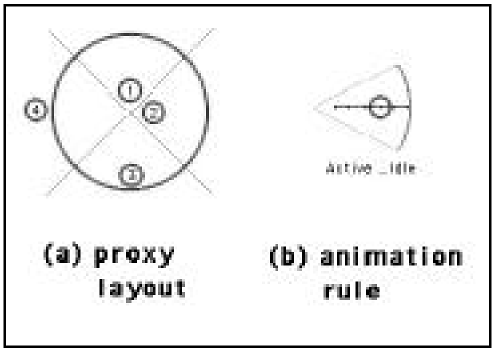
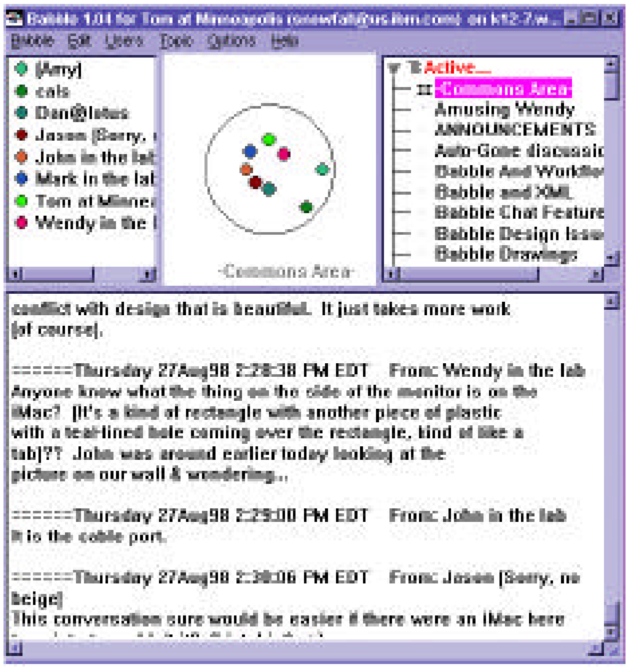


Figure 1. Social proxy schematic. In part (a), dots 1, 2 and 3, inside the circle, are part of the ‘current’ conversation; dot 4 is in another conversation. Part (b) shows how dots move: in this proxy they move abruptly to the center when they are active, and slowly drift to the periphery with inactivity.

The Social Proxy

The social proxy portrays the conversation as a large circle, and the participants as colored dots (shown as small numbered circles in the schematic in Figure 1), referred to, hereafter, as marbles. Marbles within the circle are involved in the conversation being viewed; marbles outside the circle represent those who are logged on but are in other conversations. The marbles of those who are active in the current conversation, either contributing (i.e. typing) or ‘listening’ (i.e., interacting with the conversation window) are shown near the circle’s center; with inactivity marbles drift out to the periphery. When people leave the current conversation their marbles move outside the circle; when they enter the conversation, their marbles move into the circle. When people log onto the system it creates virtual wedges for their marbles, adjusting the position of all the marbles in the social proxy; when they depart, the wedges are destroyed, and the remaining marbles adjust to uniformly occupy the space. All marble movements are shown with animation, thus making arrivals, movements, and departures visually salient. Although simple, this social proxy gives a sense of the size of the audience, the degree to which the audience is actively listening or contributing, as well as indicating whether people are gathering or dispersing, and who it is that is coming and going. We are currently experimenting with other forms of social proxies (see [3]) which show user activity over time, so that, for example, it is possible to see that a highly asynchronous conversation has a large audience, even though no two people are ever present simultaneously.

The Babble Interface



conflict with design that is beautiful. It just takes more work (of course).

======Thursday 27Aug98 2:28:38 PM EDT From: Wendy in the lab
Anyone know what the thing on the side of the monitor is on the iMac? (It's a kind of rectangle with another piece of plastic with a teal-lined hole coming over the rectangle, kind of like a tab)?? John was around earlier today looking at the picture on our wall & wondering...

======Thursday 27Aug98 2:29:00 PM EDT From: John in the lab
It is the cable port.

======Thursday 27Aug98 2:30:06 PM EDT From: Jason [Sorry, no beige]
This conversation sure would be easier if there were an iMac here

Figure 2: The Babble Interface. From the upper left, clockwise: the list of all users logged on; the social proxy; the conversation topics list; the conversation.

Figure 2 shows a screen shot of Babble's interface. In the upper left pane of the Babble window is a list of all users who are logged on, each name accompanied by its 'marble.' The upper middle pane contains the social proxy (usually called 'the cookie'); here it shows that all 8 participants are in the current conversation, two of whom are idle, and the other six having all 'spoken' or 'listened' recently. The upper right window pane shows a list of all conversations, with the current conversation — "Commons Area" highlighted. Clicking on a conversation moves the user to it, resulting in the conversation being displayed in the bottom pane of the window, and in the user's marble moving out of the circle (from the perspective of the other participants), and then appearing in a social proxy for the new conversation. Participants contribute to a conversation by typing into an entry window; each new comment is appended to the end of the conversation, with a name and time stamp. When a conversation has new material added to it (relative to a particular user), it's title in the topic list pane is shown in red. Because comments persist across sessions (indeed, they cannot be deleted), users do not need to be logged in at the same time to see the messages of other participants. Other functionality includes the ability to hold a private chat with another participant, and to find out status information about other users (including when they have last read a conversation).

Usage Experience within our Work Group

While we are aware of the shortcomings of studying the use of a system by its designers, we nevertheless believe that it is a reasonable point to begin. The group of users studied is centered around the software development group (AKA "the lab") that designed and implemented the system, and includes a mix of computer scientists and social scientists (including the authors). Over the period of time examined in this study, the Babble community ranged in number from nine to nineteen users. This growth is primarily due to members of the lab inviting colleagues with whom they had strong social or professional ties to join Babble.

Geographically, the group of Babble users is about half co-located, and half distributed. Most of the lab members are located within the same building, although offices tend to be distributed around the building — so actual adjacency is rare. To counteract this, four or five members of the lab spend most of the work time in a shared laboratory space. Two members of the lab are telecommuters, and spend the majority of their time tens to hundreds of miles away; other members of the lab frequently work at home. Four of the six associated colleagues (i.e. those not officially members of the lab, but users of Babble) are remotely located: three in the Boston area, and one in Austin.

Socially, the members of the lab are a cohesive group, with considerable camaraderie, as well as a significant amount of group work on software development projects. The associates vary in their ties to the lab members, some being well known to almost all lab members, and others being well known only two one or two lab members with whom they have shared interests or professional affiliations. Conversation in the Babble system moves fluidly between work and social talk; it is always civil, frequently informal, and joking, teasing, and other ludic behavior is not unusual.

Overall, the Babble system as used by this group can be characterized as a core of relatively synchronous social activity surrounded by a constellation of asynchronous work oriented conversations. At the center of activity is the "Commons Area," a place where collocated and remote members share news, engage in banter, get help, and 'hang out.' The Commons Area conversation tends to be more synchronous than other conversations, where hours or days often separate comments or bursts of comments. Over all, about 90% of the conversations in the lab Babble are work oriented, either involving group maintenance and management tasks, or being project or issue related.

Uses of Babble can be grouped into three general categories: social/ludic; group awareness; and instrumental. Social/ludic activities are those engaged in for entertainment and social purposes. Examples of such activities include a custom of exchanging morning greetings at the beginning of the day, a topic devoted to jokes, and the activity of 'playing with Archie', a dog who 'accompanies' one of the core users onto the system. Group awareness activities have to do with actions on the system that are addressed to the group as a whole, or to no one in particular, and generally are done without expectation of a reply or responsive action. These activities include posting announcements and other news believed to be of general interest, commenting on project activity, and keeping on-line notebooks or offices. The third type of activity is instrumental, that is, activities engaged in with a particular end in mind. These include starting or participating in focused discussions, posting bug reports, holding on-line meetings, and asking questions. These activities are often, though not always, addressed to a particular participant or group of participants.

Other Deployments

As noted above, we've deployed Babble to nine other workgroups within IBM, some of which we've studied (as are described in [1]). Although some of these deployments have been successful in the short term (two to six months), none have developed into the relatively productive and robust activity characterized by our own Babble. There are a number of possible reasons for these failures of long term adoption, ranging from software instability, to lack of critical mass, to the facilitation of communicative practices not welcomed by the majority of the group. We hope to discuss some of the issue that these deployment experiences raise.

Questions and Issues for the Workshop

At this point in our work we have more questions than answers. Our general approach, as discussed in [1, 2, and 3], clearly has some benefits. The Babble prototype supports a wide range of work and social interactions among its users. The minimalist visual representations of people and their activities (i.e., social proxies) do seem to support certain types of useful interactions such as the digital equivalent of hallway encounters, or the ability of one person to waylay another (e.g. [3]). But, although our efforts have had some promising results, we are left with more questions than answers.

1. Adoption. One of the biggest questions we are left with is what it means for a group to "adopt" a software system. In the vast majority of our deployments, we have had little difficulty getting groups to try out the system; instead, what we see is that after about six weeks, activity in the system begins to solidify (for the near future), or it drops off. It seems to us that, given the central importance of adoption in CSCW, that there are surprisingly few studies of it, or conceptual frameworks, which enable system designs to make sense of it, or design for it.
2. Critical Mass. Related to the issue of adoption is the concept of critical mass, which is often invoked when explaining the failure of CSCW. But what exactly is critical mass? As we note in [3], what constitutes a critical mass varies from one communicative practice to another: some communicative practices can exists quite well with a critical mass of two; others require a much larger number. From an intuitive standpoint, it seems to us that what is important is not critical mass, *per se*, but rather the number, variety, and interrelationships of communicative practices. We have explored some approaches to framing this [3, 4], but as yet are without a satisfactory answer.
3. Metrics. How does one measure interaction in an on-line, computer-mediated interaction system? How can one tell whether a system is succeeding or failing, or beginning to succeed or fail? Might there be indices of 'social health' or

‘interactive temperature’ or other metrics that might attest to the on-going viability of a system? Rather to our surprise, since the HCI field is so fixated on measurement, we can find no previous work on trying to quantify such things. We presume that there are tools and techniques from other fields that we can borrow, something we hope other workshop participants can point us to.

4. Negotiating the tradeoff between privacy and visibility. Our basic strategy is to make people and their activities *vis a vis* knowledge more visible to one another. Visibility means that others become aware of knowledge work, thus permitting it (and those who perform it) to be valued by the organization or institution within which it is conducted. And such visibility also provides useful cues for others, and make behavior more coherent and sociable. But at the same time, the increase in visibility means a decrease in privacy. While we do not think that personal privacy is an unmitigated good, at the same time we believe that the tradeoff between visibility and privacy is subtle, and best negotiated with extreme caution.

Next Steps in the Project

We’ve described our social approach to knowledge management, and our vision of knowledge communities; and we’ve described the Babble prototype we’ve implemented, and said a little bit about our experiences using it ourselves and deploying it to other groups. And we’ve summarized some of the most important questions we’ve facing at the moment. We’ll conclude this position paper with a look at the next steps we plan to take in evolving the Babble prototype towards the vision of knowledge communities.

The next phase of development is to make the knowledge embedded in conversations should be re-useable. That is, we want to move from today’s state, where conversation is of value primarily as it occurs, to a state in which conversation is a useful work product that can be browsed, mined, and restructured at a later time. Babble does not currently support such re-use, and the frustration of dealing with two years’ worth of conversation riddled with useful but hidden veins of information bears eloquent testimony to its value. Thus we intend to develop tools for searching, navigating, and visualizing conversations (note that conversations have considerable structure which can be exploited for these purposes — e.g., ‘find all dialogs between John and Amy that lasted more than a week and contain the word “DARPA”’. It is also important to provide *t*ools that permit participants to add structure to conversations, summarizing, glossing, highlighting, linking, and otherwise annotating them — this adds value for later browsers, and creates ‘social landmarks’ (e.g., ‘Show me comments linked to by more than 5 people’).

The final phase of the project involves the creation of an organizational knowledge space. That is, our experience with Babble suggests that knowledge creation, use, and re-purposing will proceed most easily in a relatively small group, one that is an informal, semi-private environment where knowledge workers feel ‘safe’ enough to venture tentative interpretations and conjectures. But this is at odds with the goal of making the knowledge visible to the larger organization. We imagine that a solution to this tension is to create a constellation of knowledge communities which can control the extent to which their knowledge work is outside the community. One possible approach is to allow ‘statistical information’ to be visible, so that a search engine might be used to reveal the locus and frequency of a particular concept being discussed within a knowledge community, without revealing the particulars of the discussions. Given clues that useful knowledge is present, interested parties could *request* summaries of the topic, *petition* for admission to the community, or simply *converse* with some of the community members. Notice how this strategy blends technical and social mechanisms: technology is used to locate hot spots, social mechanisms are used to control access.

Acknowledgments

The work described herein is highly collaborative, and we acknowledge the substantive contributions of our colleagues at IBM. Thanks to David N. Smith for creating the Babble prototype, Mark Laff, and Amy Katriel for implementations of the social proxy and Java versions of Babble, Cal Swart for critical assistance in the deployment of Babble, and Rachel Bellamy, Jonathan Brezin, Brent Hailpern, Charlie Hill, Amy Katriel, Wendy Kellogg, Mark Laff, Petar Makara, Peter Malkin, John Richards, David N. Smith, Cal Swart, John Thomas, and our summer interns Erin Bradner, Jason Ellis, Beth Venott, and Tracee Wolf, for productive, long-running conversations.

References

[1] Bradner, E., Kellogg, W, & Erickson, T. The Adoption and Use of Babble: A Field Study of Chat in the Workplace. To appear in the Proceedings of the European Computer Supported Cooperative Work (ECSCW '99) conference.

[2] Erickson, T. Smith, D. N., Kellogg, W. A., Laff, M. R., Richards, J. T., and Bradner, E. "Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of 'Babble.'" In Human Factors in Computing Systems: The Proceedings of CHI '99. ACM Press, 1999.

[3] Erickson, T. & Kellogg, W. "Computer-Mediated Communication and Collaboration: Designing Digital Systems that Support Social Processes." To appear in Transactions on Computer-Human Interaction. In preparation, 1999.

[4] Erickson, T. Making Sense of Computer-Mediated Communication: Conversations as Participatory Genres, CMC Systems as Genre Ecologies. Paper submitted to the Thirty-third Hawai’i International Conference on Systems Science. In preparation, 1999.