

Supporting Collaboration in Distributed Teams: Implications for e-Research

Wendy A. Kellogg

IBM T.J. Watson Research Center

P.O. Box 704, Yorktown Heights, NY 10598 USA

wkellogg@us.ibm.com

ABSTRACT

We argue that CSCW research, open source software development, the “web 2.0” and “mashup” paradigms, as well as experiences in supporting globally distributed teams within enterprises can and should inform the design of effective collaboration strategies and tools in e-Research. We describe some of the enterprise-based collaboration work at IBM that serves as part of the experience base from which to consider collaboration and CSCW issues in e-Research, and then briefly discuss issues and dimensions of distributed collaboration that have implications for supporting collaboration in e-Research.

INTRODUCTION

Our research interest in social computing and supporting collaboration and appropriation among diverse, distributed users [8,18] has led us to study a variety of collaboration applications. Arguably, many collaboration issues that have been studied in enterprise settings can contribute to a CSCW of e-Research. Enterprise collaboration continues to evolve at a rapid pace; for example, IBM recently announced the first enterprise ‘social networking’ product leveraging collaborative applications such as user profiles, communities, blogs, social bookmarking, and activities for global enterprises [16]. Blogs have proliferated inside companies, and collaboration and innovation have become a high priority. A survey of 750 CEO’s in 2006 found that 76% believed that their future success depended on collaboration and innovation [17].

The Social Computing and Cambridge-based Collaborative User Experience research groups have worked on a number of projects over the last several years related to these issues. These include understanding human productivity and scientific collaboration in high performance computing [3], task and social visualization in software development [12,7], team-based collaboration systems such as Babble [10], Loops [9], IBM’s Enhanced Audio Conferencing system [5], and games for virtual team-building [6], large-scale collaboration applications such as IBM World Jam [13,4], IBM Community Tools [27] and Many Eyes [14], and social networking applications such as Dogear [20], SmallBlue (social network visualization based on email communication patterns), Beehive (an enterprise-based Facebook-like application), and Sonar (an infrastructure for

capturing social network data and making it available to enterprise applications or mashups through APIs).

Although distributed collaboration within enterprises undoubtedly differs from open scientific collaboration in significant ways, we suspect it also exhibits many common issues. I hope to bring our experience with a wide variety of collaborative applications in enterprise settings to bear on the issues of concern to the workshop.

ISSUES IN DISTRIBUTED COLLABORATION

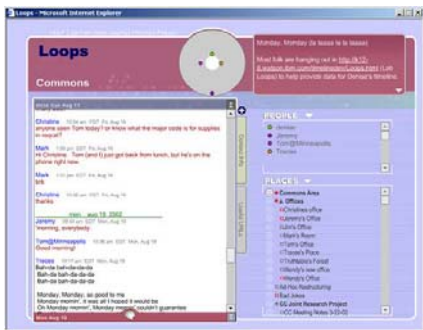
Collaboration in organizational settings has been studied for decades, but the growing importance of globally distributed teams in business and science (including amateur contributions) and the emergence of new collaboration paradigms such as large-scale collaboration have created a new impetus to understand the challenges and remedies of distributed collaboration. In this section we review some of the issues that have been articulated.

Distributed teams face a number of challenges that have been extensively documented by CSCW researchers, including developing and maintaining trust (both cognitive and affective), team cohesion, awareness of remote colleagues’ activity and project status, coordination of work, and accountability, to name a only a few [8,12,15,21,24].

One early issue that our work addressed was the need for distributed team members to interact informally, even playfully, over the course of an extended collaboration. Thomas and Kellogg [25] make a distinction between ‘instrumental’ (that which serves getting the job done) and ‘expressive’ (that which serves self-expression, identity, etc) communication. While organizations need both, software has tended to favor instrumental communication; indeed, one of the explicit goals of Babble [10] and Loops [9, Figure 1] was to support expressive communication. In general, systems that support conversation also support expressive communication.

Virtual teams have been shown to favor task-oriented interaction [11]. Paradoxically, an extreme task focus may lead to less effective communication, as it results in weaker relational links between team members [2]. A lack of social communication is also associated with lower trust and

cohesion in virtual teams [26], and with difficulties in establishing a shared knowledge base.



The Loops persistent conversation environment.

The fact that distributed colleagues used Babble and Loops to “hang out” together over periods of months and years enabled more than expressive communication: friendships emerged, barriers to asking for help among “familiar strangers” were lowered, and a sense of social identity among participants emerged. Babble’s style of interaction seemed to enhance what Austin [1] calls “Chance II,” beyond pure blind luck of “Chance I”:

In Chance II, something else has been added -- *motion*.
Years ago, when I was rushing around in the laboratory [conducting medical research], someone admonished me by asking, “Why all the busyness? One must distinguish between motion and progress”.
Yes, at some point this distinction must be made. But it cannot always be made first. And it is not always made consciously. True, *waste* motion should be avoided. But, if the researcher did not move until he was certain of progress he would accomplish very little...
A certain [basic] level of action “stirs up the pot”, brings in random ideas that will collide and stick together in fresh combinations, lets chance operate.
Motion yields a network of new experiences which, like a sieve, filter best when in constant up-and-down, side-to-side movement...
Unluck runs out if you keep stirring up things so that random elements can combine, by virtue of you and their inherent affinities.

Babble and Loops were designed to be socially translucent [8]: to use perceptually-based cues about participants and their activity to create mutual awareness and accountability. In particular, the use of persistent conversation, blended synchrony (seamless synchronous and asynchronous interactions), and social proxies (visual representations of people and their activity) created a distributed collaboration setting that was effective in supporting both task-oriented and social interactions.

Another complex of issues for virtual teams surrounds the coordination of work [12,15]: understanding who’s doing what, how much effort is being expended, the status of project deliverables, dealing with emergent issues, etc. Solutions have included structuring work to avoid unnecessary dependencies [21], and task and social visualizations [7] among many others.

Finally, a variety of logistical issues that are difficult, time-consuming or even impossible to solve exert their influence, particularly in cross-organizational distributed

collaboration. These can include time zone differences, cultural differences, incompatible software and infrastructure, privacy issues both technical (e.g., firewalls) and legal (e.g., intellectual property) [21,22].

DIMENSIONS OF DISTRIBUTED COLLABORATION

The prospects for collaboration of distributed colleagues, both technical and social, are a moving target, and have changed considerably since the advent of the internet. Nevertheless, there seem to be themes or dimensions that remain constant, though they play themselves out on ever-changing stages. In this section, we discuss a few of these dimensions of distributed collaboration.

Team vs. Large-Scale Collaboration

Distributed teams can be large or small, as can communities of practice or interest. But the emergence of significant open source software development [12,15], experiments in large-scale science collaborations such as the laboratories [22] or the MATLAB programming contests [11], the success of massively multiplayer online games (MMOGs), as well as social software applications such as del.icio.us, Flickr, Facebook, etc., represents a significantly different form of collaboration. Team-based collaboration is not likely to disappear, but the potential power of large-scale collaboration is enormous and must be embraced. E-Research will have to address a tension between team-based and larger-scale collaboration activities and CSCW research in this area can lead the way to understand how both ends of this collaborative spectrum can be leveraged.

The “Push-Pull of Competition and Cooperation”

Gulley [11] uses this phrase to describe how this tension played out in the MATLAB programming contests, calling it “strangely compelling.” He says

- We believe the contest was successful because it was
- competitive (*contestants are motivated*)
 - real-time (*contestants remain engaged*)
 - personal (*names are visible, discussion is encouraged*)
 - open-source (*all code is visible at all times*)

Crucial to the appeal is the fact that you can quickly modify and resubmit someone else’s entry. The winning entry in each contest represented the efforts of many people. Indeed, it’s fair to say that no single person could have written such an optimized algorithm. This push-pull of collaboration and competition is strangely compelling, and it echoes the popularity of open source programming, in that motivated people from all over the world contributed to create the best possible code.

The tension between cooperation and competition seems to play out differently in different arenas, but always seems to be present and an important element of what Olson and Olson [21] call “collaboration readiness.” From enterprise settings where consultants resisted sharing knowledge in the apparent belief that this would protect their unique value [23], to scientific settings where credit for initial discovery can have profound implications for career and financial success [22], to open internet collaborations like Wikipedia where significant, explicit mechanisms

governing contributions and conflict resolution have evolved with impressive rapidity [19], it is clear that the social and political realities of e-Research will frame how collaboration will succeed and fail.

Developing and Leveraging Resources

Another key determinant of collaboration and the cumulation of knowledge in e-Research will be the extent to which common resources can be developed and leveraged by individual projects. In social science, this may take the form of collecting and providing access to large data sources, particularly those that are proprietary and difficult for individual researchers to gain access. Indeed, this is an explicit goal of the newly-formed Consortium for the Science of Socio-Technical Systems (CSST), which grew out of a series of NSF workshops convened by Sara Kiesler of CMU, and being led forward by Tom Finholt of the University of Michigan. Creating new collaborative applications that allow a community of researchers to collectively “groom” or improve (e.g., by tagging or regularizing) significant data sources may allow them to be better leveraged and become more valuable over time. Here the models of open source development, which puts much of the data and process in the public domain, as well as the web 2.0 and mashup paradigms may serve e-Research well.

REFERENCES

1. Austin, J. *Chase, chance, and creativity*. Cambridge, MA: MIT Press, 2003. Cited in <http://blog.pmarca.com/>. Accessed August 14, 2007.
2. Burke, K. and Chidambaram, L. Do mediated contexts differ in information richness? A comparison of collocated and dispersed meetings. In *Proceedings of HICSS 1996*.
3. Danis, C. Forms of collaboration in high performance computing: Exploring implications for learning. In *Proceedings of CSCW 2006*.
4. Dave, K., Millen, D.R., and Muller, M. Flash forums and forumReader: Navigating a new kind of large-scale online discussion. In *Proceedings of CSCW 2004*.
5. Ding, X., Erickson, T., Kellogg, W.A., Levy, S., Christensen, J.E., Sussman, J., Vetting Wolf, T., and Bennett, W.E. An empirical study of the use of visually enhanced VOIP audio conferencing: The case of IEAC. In *Proceedings of CHI 2007*.
6. Ellis, J.B., Luther, K., Bessiere, K., and Kellogg, W.A. Games for virtual team building. Submitted to *DIS 2008*.
7. Ellis, J.B., Wahid, S., Danis, C., and Kellogg, W.A. Task and social visualization in software development: Evaluation of a prototype. In *Proceedings of CHI 2007*.
8. Erickson, T. and Kellogg, W.A. Social translucence: An approach to designing systems that mesh with social processes. *ACM ToCHI*, 7(1), 2000, 59-83.
9. Erickson, T., Kellogg, W.A., Laff, M., Sussman, J., Vetting Wolf, T., Halverson, C.A., and Edwards, D. A persistent chat space for work groups: The design, evaluation, and deployment of Loops. In *Proceedings of DIS 2006*.
10. Erickson, T., Smith, D.N., Kellogg, W.A., Laff, M., Richards, J.T., and Bradner, E. Socially translucent systems: Social proxies, persistent conversation, and the design of ‘Babble.’ In *Proceedings of CHI 1999*.
11. Gulley, N. Patterns of innovation: A web-based MATLAB programming contest. In *Extended Abstracts, CHI 2001*.
12. Halverson, C., Ellis, J.B., Danis, C., and Kellogg, W.A. Designing task visualizations to support the coordination of work in software development. In *Proceedings of CSCW 2006*.
13. Halverson, C., Newswanger, J., Erickson, T., Wolf, T., Kellogg, W.A., Laff, M., Malkin, P. "[World Jam: Supporting Talk Among 50,000+](#)" *ECSCW 2001* Poster.
14. Heer, J., Viegas, F.B., and Wattenberg, M. Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of CHI 2007*.
15. Herbsleb, J.D., Mockus, A., Finholt, T.A., and Grinter, R.E. Distance dependencies and delay in global collaboration. In *Proceedings of CSCW 2000*.
16. IBM Connections. <http://www-142.ibm.com/software/sw-lotus/products/product3.nsf/wdocs/connections>. Accessed August 14, 2007.
17. IBM Global CEO Study 2006. <http://www-03.ibm.com/press/us/en/pressrelease/19289.wss>. Accessed January 22, 2007.
18. Kellogg, W.A. and Erickson, T. (2005). Supporting appropriation work with social translucence, collective sensemaking, and social scaffolding. In Y. Dittrich, P. Dourish, A. Morch, V. Pipek, G. Stevens, and B. Torpel (eds.), *International Reports on Socio-Informatics*, 2, 2 (2005), 24-37. Available at <http://www.iisi.de>.
19. Kittur, A., Suh, B., Pendleton, B.A. and Chi, E.H. He says, she says: Conflict and coordination in Wikipedia. In *Proceedings of CHI 2007*.
20. Millen, D.R., Feinberg, J., and Kerr, B. Dogear: Social bookmarking in the enterprise. In *Proceedings of CHI 2006*.
21. Olson, G. and Olson, J.S. Distance matters. *Human Computer Interaction*, 15, 2002.
22. Olson, G., Teasley, S., Bietz, M.J., and Cogburn, D.L. Collaboratories to support distributed science: The example of international HIV/AIDS research. In *Proceedings of SAICSIT*, 2002.
23. Orlikowski, W.J. Learning from Notes: Organizational issues in groupware implementation. In *Proceedings of CSCW 1992*.
24. Powell, A., Piccoli, G., and Ives, B. Virtual teams: A review of current literature and directions for future research. *ACM SIGMIS Database*, 35(1), 2004.
25. Thomas, J.C. and Kellogg, W.A. The necessity of expressive communication in organizations. Unpublished manuscript, 2001. Available from authors
26. Walther, J.B. Relational aspects of computer-mediated communication: Experimental observations over time. *Organization Science*, 6(2), 1995.
27. Weisz, J., Erickson, T., and Kellogg, W.A. Synchronous broadcast messaging: The case of ICT. In *Proceedings of CHI 2006*.