Science on the Net

Nathan Bos, Judith S. Olson, Gary M. Olson
Collaboratory for Research on Electronic Work (CREW)
School of Information
University of Michigan

Since 1989 when Bill Wulf coined the term "collaboratory" as a laboratory without walls, there have been over a hundred collaboratory projects in science and engineering. Most were built "one off," with different sets of either customized or off-the-shelf collaboration technology fashioned for each according to their perceived needs. The social aspects of these collaboratories differed as well. For example, some were strictly managed, some more adhoc; some had formal convenants for communication and decision making, others were more amorphous. Some succeeded and some failed. Lessons learned in early projects were not well transmitted to the new adopters. Since it is very expensive to fail on these projects, it is time to analyze these past efforts, to glean the best practices in both technology and social arenas, to insure lower development costs and higher rates of success.

At Michigan and with funding from NSF, we are doing this analysis in a two-pronged approach. First, we are gathering some standard information about all the collaboratories we can identify, to date numbering 77. We collect information about the domain, location, facilities, funding sources, start-stop dates, and a description of what they did. In the course of doing this we noted some clusters of types, which we will describe later in this presentation. Second, we are going into more depth on a more modest number of collaboratory projects, detailing more about their environment, development method, management structure, social conventions, etc. as well as any measures of use and success available. From these we expect to be able to find some lessons learned and practical prescriptions about how to form a successful collaboratory in science and engineering. We have been doing this for a year and a half now, and have some preliminary findings.

First, there appear to be a variety of types of collaboratories. Lessons learned will likely differ depending on which type they are. The seven types are:

- **Distributed Research Center**
- **Shared Instrument**
- **Community Data System**
- **Virtual Community of Practice**
- **Virtual Learning Community**
- **Expert Consultation**
- **Open Community Contribution Syste**m

We will both define these and give several examples of each to illustrate their variety. Based on some of our previous field and laboratory work, we have confirmed some general aspects of these collaboratories that seem to affect their success.

1. How the work is coupled is very important: the more easily partitionable the work, the more likely it can be conducted long distance. Working through technology is not good for work that is tightly coupled or highly ambiguous.
2. People have to have common understandings of what they are doing, what is called "common ground," both about the nature of the work and how and when they will communicate.
3. The technical infrastructure has to be sophisticated enough to accommodate the new technology. The more uniform the infrastructure, the better.
4. The community supported has to have a spirit of collaboration for the collaboratory to be successful. You cannot *make* people collaborate by implementing collaboration technology.
5. Incentives for participation must be carefully designed to encourage sustained participation in some types of collaboratories.

One issue that we are currently investigating more deeply is the role of incentive systems in sustaining collaboratories. CSCW systems are often hampered by a mismatch between which participants are expected to do most of the work, and which participants are expected to reap most of the benefits (Grudin, Orlikowski). This problematic configuration is an example of a social dilemma, which are common in social systems of all types, including collaboratories. 'Solving' social dilemmas requires careful attention to distribution of incentives and rewards.

Community Data Systems are a type of collaboratory where well-designed incentive systems are particularly important. A CDS is an information resource that is created, maintained, or continuously improved by a geographically distributed community. We will describe basic features and policies of four such systems: The Zebrafish Information Network, GenBank, The Alliance for Cell Signalling Molecule pages, and the Slashdot weblog.

	Data	Participants	Incentive/ sanction
GenBank	Genetic sequence data on thousands of organisms	Genetics researchers	Would-be journal authors must show proof of GenBank contribution before publication
ZFIN	Genetic, anatomic and other data on Zebrafish model organism	Researchers using zebrafish	ZFIN relies on goodwill of a close-knit but growing community for contributions
AfCS molecule pages	Each molecule page aggregate everything that is known about signaling activity of one chemical	Researchers in cell signaling	Contributions are co-published by the journal Nature as a publication of sorts
Slashdot	Informative postings on current topics	Software developers, 'nerds'	Postings to Slashdot discussions are rated by moderators and highly-rated contributors accrue 'karma points'

This set of CDS's pose interesting contrasts. While ZFIN has managed to survive and thrive calling only upon the goodwill and interest of participants, GenBank some years ago adopted a model that is very effective, but requires cooperation from a large number of academic journals. Slashdot also relies upon goodwill and interest, but supplements this with a surprisingly successful 'karma point' system for acknowledging contributions. And the molecule pages are forging a new model, whereby contributions to a model cell will be acknowledges as equivalent to journal publications.

This talk will (time allowing) describe and contrast these community data systems, and attempt to generalize which of these incentive system models will be useful for which types of collaboratories.