Phase Two
Toward A Subjective Reality Of The collaboratory

CHAPTER FIVE
Criteria for Inclusion as collaboratory

Phase One of this study constructs an objective reality of the collaboratory based on the publications made available through the world's libraries. The assumptions of relative equality of contribution to, and interdisciplinarity of, collaboratory research are proved as practiced principles reflected by the literature, and an emergent theory that the collaboratory an ungendered intellectual environment is put forth. Practically speaking, the collaboratory is represented as a technologically-enabled shared "mind space" (Schrage 1995) where people, ideas, instruments, and information come together to make "new types" of knowledge. Objectively, the collaboratory *seems* to exist. But the practical question remains: does the collaboratory exists outside the publications of its researchers, and is it as the library represents?

A distinct boundary between objective reality and subjective reality has never been successfully resolved. Haraway (1991) prefers a hybrid object/subject construct, maintaining that the two can not be considered separately, while Wilson (1999) contends that all things, including human subjectivity, are ultimately reducible to objective physics.

Objective reality is generally understood as having an embodiment outside the mind in objects or actions that can be "seen" and therefore epistemologically proved by empirical investigation, with negation the motivation in the quest for truth. Subjective reality is based on the feelings, emotions, experiences, assumptions, and considerations internal to the human mind. Subjectivity cannot be disproved, even when it is factually wrong. Objective reality is connected to things external to self, while subjectivity is internal and separate from things outside of self (Goldberger 1996).

Phase Two seeks a subjective reality of the collaboratory based on immersion in the online environment. Chapter Five introduces and examines the third key document of the collaboratory literature, the National Research Council's (1993) *National Collaboratories: Applying Information Technologies for Scientific Research,* which sets the instrumental foundation for the collaboratory. NRC (1993) is synthesized with Robbin's (1995) evaluative criteria for collaboratory implementation, and the **CIRAL** matrix of necessary but not sufficient criteria for inclusion as a collaboratory is developed. Chapter Six presents the collaboratory testsite experience. Chapter Seven presents three additional case studies developed during immersion in the larger collaboratory environment.

National Collaboratories (NRC 1993)

On the philosophical foundation of Wulf's (1988*) White Paper,* and the intellectual foundation of the Lederberg and Uncapher (1989) *Report*, the National Science

Foundation, in December 1991, convened the Committee on a National collaboratory and charged it with establishing a user-developer partnership to actualize the collaboratory. The Committee's task was to study and report (NRC 1993) on the need for and potential of information technology to support collaboration in the conduct of scientific research (vii). The yearlong effort included frequent face-to-face meetings and three two-day disciplinary workshops (molecular biology and specifically genome research, oceanography, and space physics.) As the NRC study evolved,

> the idea of developing a single national collaboratory was replaced by the idea of developing multiple scientific collaboratories which would share network and computing resources, software, and infrastructure but would have unique features dictated by the needs of particular scientific disciplines. (viii)

The committee's motivation was to inspire tool development and leverage the efforts of computer scientists. Its goal was to "develop a more explicit partnership between scientists in general and computer scientists in particular" (1). Science's traditional culture of competition and individuality, and the traditional system of rewards and recognition among scientists, were recognized as major impediments to achieving effective collaboration. Bottom-up motivation was recognized as an essential success factor (2). Education and research were given equal priority (3). A resolution of the perception of technology and tools development as being somehow "less than" research needed to be reached (4) in order to achieve

> an environment in which a scientist’s instruments and information are virtually local, regardless of their actual locations... [and the] ...virtual presence of an individual in someone else’s laboratory [can be achieved]. (2)

The 105-page *National Collaboratories* includes an Executive Summary, six sections, and four appendixes. Three of the sections focus on specific requirements for building collaboratories for genome research, oceanography, and space physics. Together, these criteria provide the generic model for the collaboratory. The final section of *National Collaboratories* focuses on the logistics for implementation of a national program.

"Building and Using Collaboratories"

Section Five of *National Collaboratories,* "Building and Using Collaboratories" distills the discipline-specific needs identified in its preceding chapters. It enumerates the basic capabilities a collaboratory should support, and analyzes the social, organizational, technical, and practical considerations for individuals and institutions.

There are four general classes of information-related problems and specific needs within those areas that the collaboratory must address technically for large-scale functionality to achieved:

1. **Data Sharing**

- Electronic libraries combining databases, literature, and software
- Easily accessible data archives

A comprehensive system supporting retrieval of data from any and all sources,

between items in this federated data

2.

-
-
-

3. **Controlling Remote Instruments**

Reliable networking with sufficient speed to support interactive responses

Method for capturing and transmitting output

Standardized telemetry protocols

Communicating with remote colleagues

- Including voice, video, text, data, images
- Synchronous and asynchronous modes. (56)

Functions between and among these information problems must be transparent

agents, might be developed to achieve this transparency.

Social and Institutional Factors

identifies social and institutional factors relevant to the

individual scientist that must be tackled:

Motivation for participation

- Confidence and security in data
- Rewards and recognition for contributions
- Resolution between the perceived value of tools and technology construction, and research (64).

The NRC workgroup recognized that a sense of community was necessary in order to establish and sustain collaborative communication among scientists, and recommended starting with scientific communities predisposed to communication while research developed. The workgroup recognized that the costs of using computer-based collaboration technology for individual scientists include:

    - Incompatibility/critical mass of tasks and people (having to switch between media)
    - Economic costs
    - Dependence on network and vulnerability (of data and functionality)
    - Need for ongoing education and training

- Development of effective, equilateral partnerships (65).
- Issues for individual institutions supporting collaborating scientists include:
    - Providing local infrastructure support (including new administrative, control, and accountability models, and increased support for technicians and support personnel)
    - Managing the results of increased interaction, including increased travel and managing shared results (68).

one agency is charged with infrastructure development, it is generally perceived that allocating funds implies taking money away from individual or domain-specific

scientists have in collaborating in any circumstances led the NRC to recommend that collaboratory projects be selected from the bottom up,

> scientists who recognize a need to collaborate and who manifest an interest in applying more and better information technology. (70)

The technological base for the generic model of the collaboratory includes five criteria:

> Interoperability
>
> Transparency
>
> Customizability
>
> Integrity
>
> Extensibility (28-29)

Robbin (1995, 40) synthesized the collaboratory environment described in *National Collaboratories,* collaboratory for evaluation of her collaboratory project. The NRC-Robbins generic model includes:

- A distributed computer system
- Networked laboratory instruments and data-gathering platforms
- Tools to enable a variety of collaborative activities
- Financial and human resources for maintaining, evolving, coordinating and assisting in the use of computer-based facilities
- Digital archives and libraries that include tools for organizing, describing and managing data, including images, to enable large-scale sharing of data
- Digital libraries which include tools for organizing, describing, and managing the information derived from analysis of the data.

The CIRAL Matrix for Evaluation of the Online collaboratory

Phase Two of this study seeks a subjective reality of the collaboratory based on immersion in the online environment. The subjective reality seeks to determine whether, and to what extent, the collaboratory actually exists, and whether the subjective experience aligns with the objective reality determined in Phase One. To establish preliminary criteria for evaluating the existence of the collaboratory, Robbin's NRC-based generic model for evaluating the collaboratory is synthesized and reconstructed. The fundamental components of the Robbin/NRC model are: distributed **C**omputing, networked **I**nstruments and tools, support **R**esources, data **A**rchives, and digital **L**ibraries, or **CIRAL**.

The matrix rendered as Appendix D considers the needs of expansive investigation into the other key constituents of an information ecology (people,

and institutional levels as identified in *National Collaboratories.*

scope of this study, that is, a synoptic exploration of the environment, the **CIRAL**

and to answer the fundamental question: Does the collaboratory exist?

Discussion of Criteria for Inclusion

necessity the resources to support them, are a given. Without computerized networks and the resources to support them, no collaboratory be online and so would not exist. Several articles from the collaboratory literature offer online

Psychological

Science

includes most of the collaboratories that are given individual attention in or are otherwise mentioned by the remaining literature. Simple searches for the single word

Yahoo, etc.) produce thousands of hits to sites that call themselves "collaboratory." These results advance the assumption that computerized networks and the

collaboratory online.

Data archives must also exist if the **CIRAL** criteria for inclusion are met. Data archives may simply be the collection of computer files and programs that facilitate USE of the computerized networks for the collaboratory, and so are also a given if indeed an online collaboratory exists. But, data archives may be much more than a collection of enabling software, and, indeed, the archives may be of an entirely new sort. They may include archives of data created by collaboratory participants or instruments during the course of collaboratory activity, such as with virtual "notebook" systems that allow the logging of text, images, messages, and other digital information in central or decentralized systems available to collaboratory participants. The question of data archives raises the question of digital libraries, and the difference between them, which will be discussed later. But first, the question of remote instrumentation needs attention.

It is clear from Wulf (1988), Lederberg and Uncapher (1989) and NRC (1993) that the intended definition of remote instrumentation is a mechanical device or contrivance which serves as a tool for specific types of work, is rare or not widely held, and which may be manipulated by someone remotely located using computerized networks. It may be argued that the computer network is itself a collaboratory instrument. Thus, interactive online environments that rely on the instrument of the network (as with virtual reality systems or interactive web sties), or those that rely on the human mind (as with MUDS and MOOS, which construct text and object-based interactive online environments), may be included as collaboratories. While it is easy to imagine these online environments are at least of the same family as the collaboratory, for this study the more strict definition is adopted. Accordingly, any computerized network supported by resources that has data archives and digital library resources, but which does not include access to and the ability to manipulate remotely located instrumentation, is excluded as a collaboratory.

Likewise, any site that does not include access to and support of the digital library must be excluded. The obvious distinction between data archives and digital library is the application of librarianship, a professional activity (Abbott 1988) concerned with the intellectual activity of collecting and disseminating information resources to support specific human activities or goals. Whether this service is digital, as with a webpage of relevant links, or is an intelligent agent, (a computer program that searches for, harvests, formats, and delivers relevant data), or is provided in real time by a human being, is irrelevant to this study. An exact and rigid definition of "digital library" is also not necessary since the concept is still very much in its own research and development stage (Spink 1999). For the purposes of this study, however, digital library is defined as a specifically built and managed collection of relevant internal and external data and information intended to support the knowledge activities of collaboratory participants. As such, digital library resources are distinct from, but may include, data archives, which are defined as files that support the collaboratory's technological functions and/or record its activities. Certainly, one collaboratory's data archives may be another collaboratory's digital library, so further distinction must be made. For the purpose of this study, data archives are defined as the accumulated and saved records that support the collaboratory's existence, and digital library resources are defined as value-added data, or information saved, accumulated, and made accessible in service to collaboratory activities.

Conclusion

Phase One's objective reality of the collaboratory as reflected in publications held by the library proves as practiced principles Wulf's assumptions of relative equality of contribution to, and interdisciplinary of, collaboratory research, and theorizes that the collaboratory is an ungendered intellectual environment. Chapter Four begins Phase Two's subjective reality with an investigation of the collaboratory's third foundation document (NRC 1993), which sets the instrumental foundation for the collaboratory and from which an evaluative instrument, the **CIRAL** matrix for inclusion as a collaboratory, is constructed. The necessary but not sufficient criteria for inclusion as a collaboratory are identified and discussed. Chapter Five reports the collaboratory testsite experience, and Chapter Six presents the empirical evidence of the study's immersive phase.

CHAPTER SIX

Phase Two of this study seeks a subjective, experiential reality of the collaboratory based on a prolonged immersion in the online environment. The

and a derivative evaluative instrument, the **CIRAL**

relates the experience of a prolonged immersion in a collaboratory testsite, and tests the usefulness of the Matrix for guiding exploration of the larger collaboratory environment, which is preseented as a series of case studies in Chapter Seven.

through July 1999.

In February 1998, a web search using the term "collaboratory" on Digital's Alta

® search engine and index at http://www.altavista.digital.com produced 489

an almost ten-fold increase. The same search of the indices at http://www.hotbot.com, http://www.excite.com, and http://www.yahoo.com also

any number of factors (including index age, algorithm configuration, index effectiveness, index growth, or increase in sites using the word). The results are taken here only as an indicator there is sufficient reason to assume a collaboratory exists and that it is publicly accessible.

Five Collaboratories are referred to frequently in the library literature, and are also represented in the web search results herein described. They are:

- The Materials MicroCharacterization (MMC or M2C) Collaboratory of the U.S. National Laboratories;
- The Upper Atmospheric Research Collaboratory (UARC) at the University of Michigan;
- The Collaboratory for Research on Electronic Work (CREW), also at the University of Michigan;
- The Management Information Science (MIS) Collaboratory at the University of Texas; and
- The U.S. Department of Energy's Diesel Combustion Collaboratory (DCC).

These five Collaboratories were selected for first contact because of their longevity and institutional affiliation as likely to be fully functioning. A simple web search for the full name of each collaboratory quickly produced web portals and the email address of the most logical gatekeeper (usually the project coordinator or Principle Investigator). An email message requesting consent to enter the collaboratory for environmental dissertation research was sent to each gatekeeper. An interactive consent form was made available on the web at http://www.intertwining.org/dissertation/consent/p2consen.htm. M2C returned first consent, so was selected as the testsite for this study. Following is the account of the researcher's immersive experience in the M2C Collaboratory.

> Dire chimera's conquest was enjoined;
> The mingled monster, of no mortal kind;
>
> A goat's rough body bore a lion's head;
> Her pinchy nostrils flaky flames expire;
>
> --Homer, Iliad

Its difficult to tell the head from the tail of the Materials MicroCharacterization

coverage and saturation of online search indices are a true indicator, there can be no doubt M2C is the both the loudest and proudest collaboratory in the world.

Collaboratories, a sort of mega-collaboratory or collaboratory of collaboratories. The M2C Collaboratory includes three US government laboratories: Argonne National

California, and Oak Ridge National Laboratory (ORNL) in Tennessee. Also included are the National Advanced Materials Testbed of the National Institute of Technology

Microanalysis of Materials. The M2C collaboratory is supported, in part, by the U.S.

Department of Energy's DOE2000 initiative, and by DOE's prior initiative, the "Distributed Computing Electronic Environment" or DCEE.

All five labs are concerned with microscopy: they look at very small things up close, and with materials science: the interconnections of materials at their point of contact. Materials Science is a

> blend of a multitude of disciplines ranging from basic science to applied engineering, from physics and chemistry through metallurgy and ceramics…which rely on techniques employing electrons, ions, photons, x-rays, neutrons, and mechanical and/or electromagnetic radiation to elucidate the microstructure of matter. (http://tpm.amc.anl.gov/MMC/HomePage.html)

Each of the M2C's TPM sites sends and receive web-based video and provides web access to and remote manipulation of unique microscopy instruments. Each maintains a unique online address with a distinct URL. Each has its own set of webpages.

While several of the TPM sites look alike, several use distinctly different webpage designs. Movement between the look-alike sites is virtually seamless and, unless close attention is paid to the URL on the browser status bar, it is easy to move from one site to another without knowing. But, each TPM site also uses a different primary name for their larger online presence, making it difficult during preliminary searching to determine that while they are, at the larger level, disconnected and distinct, they are at the functional level, are very interconnected and interdependent. The concept is caught elegantly in the M2C's animated online logo (Figure 24), which slowly highlights one acronym at a time in a sort of peacefully rotating disconnected connection.

Figure 24. M2C collaboratory logo

ANL — LBNL
U. of I. M2C NIST
ORNL

At least seven Industrial Partners participate technologically and financially in the M2C collaboratory. Industrial Partners include Gatan Inc., R.J.Lee Group, EMiSPEC Systems Inc., Philips Electronic Instruments, Hitachi Scientific Instruments, Inc., JEOL USA Inc., and Graham Technology Solutions. None of these partners have a strong graphical or informational presence at the M2C Collaboratory's five TPM sites.

Together, the M2C Collaboratory's five TPM sites have $50 million in rare and expensive microscopy instruments, with key pieces hardwired to the network, and hundreds of highly trained scientists and technicians with tens of thousands of hours of specialized scientific experience at each others' disposal. None of the labs do exactly the same thing, and each has specialized and unique instruments that, when brought together as M2C, do the larger job at each individual lab.

The "hand" of the M2C collaboratory seems to be at the Argonne National Laboratory, http://tpm.amc.anl.gov (Zaluzec 1999), from where the coordinated web video telepresence and an aggressive public relations program originates, and the M2C's most public personality, Nestor Zaluzec, is located. The unofficial "head" is at Oak Ridge National Laboratory (Wright 1999) where the information science aspect of the project is centered.

Zaluzec arranged for two video cassettes (Argonne 1998, Dept. of Commerce 1998a), and a CD-ROM (Dept. of Commerce 1998b) about M2C and DOE 2000 to be sent to the researcher as soon as consent to enter the Argonne TPM site was granted. The researcher was also advised of a recently published "hot" article about the M2C collaboratory (Kling 1998), which provides a nice introduction to the project, and which was not discovered during Phase One's library research. The multimedia materials are intended for educational use, are distributed freely to the public, and are available on request to the U.S. Department of Commerce NIST Public & Business Affairs Office in Gaithersberg, MD.

The professionally produced CD-ROM and video cassettes present the story of an exciting collaboration between scientists at NIST, Argonne, and Texas Instruments (TI) during which a microscopic problem that developed during silicon chip production at the TI manufacturing plant in Texas is resolved online in a M2C Collaboratory session. The TI scientist prepared and sent the problem chip to Argonne. Argonne positioned the chip in one of its microscopes, and opened a collaboratory session with TI. The TI scientist could see through the eyes of and manipulate the microscope, and collaborated with the Argonne microscopist to identify the problem and devise the solution.

But, full color, multimedia glitz and promising public relations aside, the actual experience of M2C Collaboratory's virtual facility is a hard-on-the-brain cacophony of overcrowded information presented on a battleship gray webpage awash in a sea of acronyms. The Argonne TPM web page is chunked into five or more scrollable frames, all crowded onto one computer screen. One quick click on the wrong acronym switches the participant, without fanfare, from one lab to the next, one instrument to another, from one side of the country to the other. The whole arrangement makes it difficult to maintain orientation. Nevertheless, the promise that anyone, anywhere in the world, can log on to the public Internet, go to the Argonne TPM web site, and fine focus million dollar microscopes variously located at five National Laboratories around the country, makes complaints about the cacophony of the interface design petty. Such potential is a powerful feeling, and provides a taste of what the future may hold for virtual scientists.

Consent to enter the ORNL, UofI, and NIST TPM sites was also received. Consent to enter the LBNL collaboratory was denied because the site would be inactive during the term of this research, and the primary contact at LBNL was out of the country. The testsite experience, then, became an exploration of four Collaboratories, as one.

TPM : TelePresence Microscopy

The Argonne TPM site is the control center of the M2C's telepresence. It coordinates constant video feed between and among the five TPM sites using "push" technology. It does not support audio transmission, so the experience is all eyes, no ears. The video transmission rate over a simple dialup connection to the web is three frames per second and begins as soon as the website is accessed. Three frames per second delivers jerky video: more a slow steady series of independent snapshots than a flowing stream. The connection, however, is constant and consistent if the client system is not running superfluous other programs, particularly anything from the Microsoft family of software, the AutoSave functions of which consistently cause crashes and session interruptions. Argonne recommends a high speed dedicated connection for full effect.

Whether through a dedicated connection or by Internet dialup, maintaining a connection to M2C's video feed requires lots and lots of RAM, or Random Access Memory, the "desktop" space available during any computer session. For ease of this research, a personal computer dedicated exclusively to the collaboratory session was necessary (additional systems configuration is discussed later in this chapter.)

Argonne's TPM website provides multiple, simultaneous live or recorded video transmissions from itself and any ONE of the other four TPM sites, with as many as six separate video transmissions available on a single webpage. Switching between the video views and between TPM sites is as quick as the click. The web telepresence uses a point-to-point video feed rather than a multi-point videocast, as it would if all five sites were sending and receiving video simultaneously (which TPM facilitates on special occasion using CuSeeMe teleconferencing software.) Participants are able to switch between and control the views from the microscopes and the macroscopes (room view video cameras) at any of the sites. Split screen "collaboratory sessions" allow the participant to view microscope and macroscope video transmissions from two separate cameras at each of two sites simultaneously for a sort of "four-eyed" experience. Everyone accessing the TPM website sees the same video feed, and anyone can change which video is seen. The host-site may override and disable a remote controller, and any change in video view changes the view for everyone, so a warning is posted that if the view suddenly "changes" its because someone else who's accessing the site has switched views. There is no information provided about who else, or how many others, are accessing the TPM sites, so it feels as if only you, and the people you can see on the macroscope view from the lab (if there are any), are present, when in fact thousands of others might be.

To further confound the confusion about who, or how many, are seeing what view from where, the number of possible simultaneous video feeds on a single screen is six, making a sort of six-eyed experience with each eye seeing something different. Getting oriented to what is on the screen is difficult and confusing, and often, as soon as orientation is gained, the "view" suddenly changes.

The Argonne lab is the video "headquarters" of the M2C collaboratory, and must be included as a partner in any two-lab collaboratory session. The Argonne setup does not allow direct connection between two remote labs, only between itself and one other lab, or any other lab and Argonne. The same is true of the other sites, which allow a view from themselves and Argonne, or from Argonne and any one other site. Accessing any one of the labs by itself is also possible, either from a connection at Argonne, or by logging on to that TPM's separate URL.

For instance, a collaboratory session between the Oak Ridge Lab and the Berkeley Lab is not supported, however two sessions can be set up simultaneously, as between Argonne and Oak Ridge, and Argonne and Berkeley, and participants can switch between these sessions, requiring orientation to as many as twelve video feeds. Collaboratory sessions are preprogrammed and executable with a one click switch to "collaboratory mode." Remote users may "join" an ongoing collaboratory session, and will receive microscopic and macroscopic video from two labs on a single screen (see "collaboratory mode" later in this chapter). In other words, participants may be visually "in" one of the labs alone, or, in collaboratory mode be "in" two labs at once, or even be in multiple collaboratory sessions, conceivably four at once.

For a scientist at one of the labs, or for one of the M2C Industrial Partners, or someone from the wider Internet, to use instruments at any of the five TPM sites, the materials sample has to be prepared and sent to the lab and a technician or scientist has to receive the sample and load it into the instrument. This setup, then, allows the remote scientists to have their eyes at the remote site, one on the lens of the microscope, and the other viewing the room, with their hands on the remote control buttons in a separate frame on the M2C webpage. They can see through and manipulate the microscope, but they cannot move the sample or talk to the remote scientist without a subsidiary telephone connection. A "snapshot" of the microscope's results can be viewed either as real-time online video, or the results can be captured and posted to a collaboratory notebook (the notebook configuration is discussed in detail later in this Chapter.)

It is possible and practicable for anyone to connect to and enjoy the full function of the M2C collaboratory and participate in TPM collaboratory activity with a 14.4mps dialup ISP connection to the Internet from a 486-66mhz PC with 16 MB RAM and a standard video monitor. It's slow, but possible. A frames-capable version of Netscape Navigator® (2.0 or higher) is required to receive video, otherwise no special software or unusual configuration is required. Microsoft's Internet Explorer® browser does not support the file type used for video transmission. This research used a 28.8 Mbps dialup Internet connection from a 200mhz Pentium II with 128mg

RAM and 2mg Video Ram, as well as a 166 MHz Pentium laptop with 16 MB RAM and 1 MB video Ram connected by dialup via a 56.6 Mbps modem.

Only one area of each TPM site has restricted access: the "Online Control" of the hardwired instrumentation. Online control of the instruments must be justified and arranged ahead of time. As U.S. National Laboratories, however, each of the TPM sites is required to make their instrumentation available to qualified outside researchers. A form for requesting instrument use is online at each site. Users may view others' use of the instruments, or read the contents of the public notebooks, (password protected private notebooks are also set up) and may remotely manipulate the site's macroscopes without password access however.

Video from and manipulation of the room-view cameras, or macroscopes, is provided around the clock, whether anyone is in the lab, or not. However, video feed is often not "live" in that is may be a prerecorded transmission intended for demonstration purposes. It is often difficult to distinguish whether the video is live, or prerecorded. Of course, macroscope and microscope views that are prerecorded have the remote control options disabled.

Each TPM site hosts shared public and private notebooks. The notebook is a shared, online document space in which participants may post pages containing text or still graphic shots captured from the microscopes. Only one person can have realtime control of any given notebook page, so chat-like communication is not possible. Private notebooks are password protected and reserved for specific experiments, which may be of a proprietary nature, as with a Collaboration with Texas Instruments or one of the other Industrial Partners. Public notebooks allow real-time addition of pages via web forms technology. Each notebook page has an email-like header identifying the author, subject, etc. While there is a provision for searching the subject line of notebook pages, there is no provision for full text search, or for annotating or linking among notebook pages, or among notebooks.

The Argonne TPM site hosts a data archive where programs, software patches, and activity logs are stored and shared via file transfer protocol. The site hosts an "Ask-the-Microscopist" service, and provides information about setting up middle/high school experiments. There is a questionnaire and comments form, and access to join or read the archives of a six-year old distributed email list. The researcher was granted access to conduct research on the archives of the M2C collaboratory email list, an opportunity which will provide data for expansive studies of other aspects of the collaboratory ecology, but which was not undertaken in this study.

The site design relies heavily on web frames (dividing the screen into multiple, smaller windows) to organize and present these many options. The interface is cacophonous and confusing. The framed sections quickly become quite small and are packed with options, most of which involve arcane that may be confusing to the novice user. The confusion is compounded because each of the five laboratories has its own specific set of acronyms along with an individual name and domain, and one click can mean a switch between labs and instruments without noticeable change in layout or design of the web interface. At certain times, the M2C Collaboratory's TPM sites feels like five labs in one, and at others, five individual labs.

This study monitored the M2C collaboratory through the Argonne TPM site at odd times and for varying durations and intensities, on at least a weekly basis, for ten months from October 1998 until July 1999. But, it wasn't until three months into the experience that the "who's on first" feeling of disorientation was relieved, and comfortable orientation to the larger environment was achieved. The final confounding variable is that contact information from each of the sites leads back to Argonne, making the Argonne TPM site "feel" as if it is the center of the Materials MicroCharacterization collaboratory. But M2C may also feels "centered" at any of the other TPM sites when primary access is made through that lab's distinct URL. M2C is actually a megacollaboratory encompassing five individual but intertwined collaboratory sites: ANL, LBNL, NIST, ORNL, and U of I.

During the ten months of this study's immersion, the main microscope at TPM Argonne was frequently offline, and two of its seven microscope lenses were under repair. No public experiments were conducted. Many times a prerecorded video of collaboratory activities was broadcast and it looked "as if" there was live lab activity, when in fact it was not "real-time" activity at all, but prerecorded broadcast.

Because the macroscope controls were accessible even when no one seemed to be in the lab, users were free to "poke around," panning the cameras, and focusing in and out on pieces of equipment and researcher desk space to get a sense of the place. Because the transmission rate is three frames per second, however, and because the views could be changed without notice by some unknown, other visitor, sometimes someone would just "appear" in a chair in the laboratory, and just as quickly "disappear" between frames. It was not immediately possible to tell whether this appearance was real, or recorded, and often from which lab it originated. In fact, it was generally difficult to tell what was "real" and what was recorded, what was live, and what was contrived, during most of the M2C sessions. At no time during the ten month immersion did any collaboratory participants make contact with the researcher while in the collaboratory. No response was posted to the researcher's entry in the public notebook.

Because of the specialized, scientific focus of the M2C, an understanding of instrument type and function would help with orientation, and for educational purposes an elementary overview of the equipment and the work they used for would be helpful. Much of the researcher time was spent learning which piece of equipment did what sort of job. Without this specialized knowledge it is virtually impossible to get a comfortable sense of the work that takes place in the M2C.

The Argonne TPM public notebook contains 125 entries, dating from March 1997. Most of the posts are brief messages concerned with startup technicalities and configuration tests for various graphic settings. Again, this was no place for beginners as none of the messages made any attempt to explain either technicalities or acronyms. There are a few series of messages (indicated by the subject line of the notebook entry), which are mostly single frame graphic files that appear to have been produced on a microscope during remote collaboration with a microscopist. None of the notebook pages that containing graphics provided more than a brief, acronym-laded description of the shot; there was no discussion of the contents of the graphic file, nor any sort of formal diagnosis of the materials problem.

The M2C's digital library (although the word "library" is not used) is a separate collection of internal and external web links maintained by the microscopist/webmaster at Argonne. The link to the digital library (called "General Info about TelePresence Microscopy") is not prominantly displayed but, once located, transports the user to a separate web address at http://www.amc.anl.gov, home of the Microscopy and Microanalysis WWW Server. The server offers a collection of project papers, proceedings, and reports aimed at general education about the M2C concept and the science of microscopy. None of the documents were for specific project support nor were they about ongoing experiments. They seemed intended for visitor information rather than to support Collaboratory activity. The "General Information" site also points to data archives that include the site's distributed email list archive, and links to related microscopy societies and other external WWW resources. There were no preprint papers, and no peer review postings. The M2C collaboratory felt as if its main purpose was the exchange of video and captured shots from the instruments. Very little attention paid is paid to the intercommunication of participants.

With all the **CIRAL** criteria in place and functioning, the Argonne TPM site of the M2C collaboratory proves the collaboratory exists as it was philosophically, intellectually, and instrumentally described by Wulf (1988) Lederberg and Uncapher (1989) and NRC (1993). It is not possible to say M2C was the first or is the only full implementation of the collaboratory, but it is possible to say that the Argonne TPM site of the M2C collaboratory appears to have been fully functional in March 1997, when its public notebook went online.

NIST TelePresence Microscopy Site

At first, there is no apparent difference in the look and feel of the Argonne TPM and the National Institute of Standards and Technology (NIST) TPM site, and for a while the researcher was not aware they are two distinct TPM sites. Even though the online address, or URL, is different, the interface and design are nearly identical, as if the Argonne TPM webpages were mirrored on the NIST site at http://scanner.cme.nist.gov.   Only the acronyms seemed to change. But, on closer inspection, the Argonne TPM site showed 41,566 visitors logged on since 1997, while the NIST site logged 1801 visitors. The microscope video feed from each site is different (although this might only be apparent to a microscopist.) All other options seemed identical.

Achieving collaboratory mode between the Argonne and NIST TPL sites is a one-click operation that brings both sites together onto a single screen (Figure 25). Mico- and macroscopic video views from each site are framed above frames for video control, comments, and instrument controls frames, for a total of seven frames on a single webpage. Similar sessions can be configured between Argonne and ORNL, Argonne and U of I, or Argonne and LBNL.

Figure 26 is a screenshot of a TelePresence Microscopy session between Argonne National Laboratory (ANL) and National Institute of Standards and Technology (NIST) on May 16, 1999, as seen from an Internet connection to the website at http://tpm.amc.anl.gov/TPMLVCollabNIST.html.

Figure 25. Screenshot of M2C TelePresence Microscopy session

The upper left frame is a lens shot *from* the Advanced Analytical Electron Microscope at Argone in Illionois. The upper right frame is a shot *of* the NammMetrology SEM Microscope at NIST in Maryland. The top center shot is a macroscope (room) view of the Argonne lab, and the middle center shot is the macroscope (room) view of the NIST lab. The lower left frame has buttons to change between the instrument or cameras at either site. The lower center frame contains links to switch between collaboratory sessions and to other web-based resources. The lower right frame is the online instrument controls, which, in this shot, are offline.

Consent from ORNL's TPM gatekeeper suggested the researcher access the M2C from http:tpm.amc.anl.gov/MMC/, yet another portal to the site. This access point provides a thorough overview of the M2C collaboratory and includes an introduction and background information, goals, tools, status, opportunities for involvement, expansion plans, partners, visions, publications, workshops, images, a form to request use of resources or information at/from any of the TPM sites, email access to TPM gatekeepers, an overview of research programs, and connection to the project's steering committee. While all this text-based information helps make sense of the science of microscopy, and the notion of the collaboratory generally, it doesn't aid visual orientation to the TPM experience.

The M2C experience has a sort of "virtual" feeling of being there, somewhere, within and among these five labs, but in eyes only, and with a lingering sense that others MAY be around, but are hidden. Because there was no human interaction, either via audio transmission or real-time chat, and because it was difficult to determine whether the video feed was live, or prerecorded, and also difficult to orient to which lab one was connected, the M2C experience was generally disorienting and produced a feeling of disconnected connection, of being alone online with unknown virtual others.

Conclusion

The Materials MicroCharacterization collaboratory, M2C, meets the **CIRAL** criteria for inclusion as a fully functioning collaboratory as put forth in Wulf (1988), Lederberg and Uncapher (1989), and NRC (1993). It includes computerized networks, remote instrumentation, resources to support, data archives, and digital library-like resources. The collaboratory exists, even if it doesn't include direct human-to-human communication (waves and two-fingered peace signs notwithstanding).

CHAPTER SEVEN
Empirical Findings

In Chapter Six the existence of at least one collaboratory that meets all the criteria for inclusion of the **CIRAL** matrix is established. The Department of Energy-funded Materials MicroCharacterization (M2C) collaboratory, with its five remotely-located TelePresence Microscopy (TPM) sites uses resource-supported computerized networks to provide access to and manipulation of remote, rare, and expensive instrumentation, and hosts digital archives and digital library resources to support the work of collaborating scientists and researchers. Such a system, a "collaboratory" was first visualized and described by Wulf in 1988, and further developed by Lederberg and Uncapher (1989) and NRC (1993). Chapter Seven seeks to discover whether other full collaboratory implementations are in place, to determine the nature of their environment, and also explores the nature of sites that use the word collaboratory, but do not meet the **CIRAL** criteria for inclusion.

There are virtually thousands of websites that use the word collaboratory but which obviously do not meet the **CIRAL** criteria for inclusion. Most of the thousands of websites clearly miss the definition because they do not incorporate access to and manipulation of remote instrumentation. Many, in fact, are simply personal or institutional websites that incorporate some level of interactivity using web server functions that are now generally and widely available to the public. In fact, a simple website with a link to the Library of Congress and set up to facilitate a Microsoft NetMeeting session (which allows point-to-point real time audio/video transmission, remote access to and control of either personal computer, file sharing, and whiteboard capabilities), can claim to meet all the criteria for inclusion as a collaboratory *except* for providing access to and manipulation of rare and expensive remote instrumentation.

Online environments that use the word "collaboratory" but that do not meet the **CIRAL** criteria for inclusion are mostly academic or educational websites. Most of the sites incorporate various collaborative or communication functions, including links to email, chat, or forum functions, and provide access to various types of data archives or links to internal or external digital library facilities. Consequently, access to and manipulation of remote instrumentation is either the determining criteria for proper use of the word "collaboratory," or the definition of the word "collaboratory" needs revision. If redefinition is adopted, however, the word "website" and "collaboratory" become virtually interchangeable, the Internet has become the collaboratory (or vice versa), and the distinct notion of a collaboratory is meaningless.

Derivative "collaboratories"

Many publicly accessible collaboratories are K-12 oriented web-based information clearinghouses that present some sort of interaction between students and remote scientists, educators, or datasets. For example, the CoVis collaboratory, which is funded in part by the National Science Foundation, is

> the website of the *Learning Through Collaborative Visualization Project* at Northwestern University, comprising…a community of thousands of students, over one hundred teachers, and dozens of researchers, all working together to find new ways to think about and practice science in the classroom. (http://www.covis.nwu.edu/)

Many collaboratories, such as the University of Michigan's Learning Collaboratory at http://databases.si.umich.edu/Learningcollaboratory/ are intercollegiate university computer labs that can connect to each other using point-to-point or multi-point audio or video transmissions, but do not host access to remote instrumentation.

Some collaboratories, such as the nonprofit Research collaboratory for Structural Bioinformatics (RCSB) at http://rcsb.nist.gov/ are shared data repositories but do not incorporate human-to-human communication functions or access to remote instrumentation.

Some collaboratories, such as the University of Michigan's Collaboratory for Research on Electronic Work (CREW) at http://crew.umich.edu/index.htm, and the MIS Collaboratory at the University of Texas at http://cism.bus.utexas.edu/collab.html, are metacollaboratories, or web sites for university-based centers of research about collaboratories or collaboratory work which offer extensive data archives and digital library functions, but (in the case of CREW) no remote communication function, and (in the case of both) no access to remote instrumentation.

Undoubtedly, some government agencies, such as NASA and branches of the U.S. military, have fully operational Collaboratories that facilitate manipulation of remote instrumentation (such as launch rockets, and space station instruments.) But these setups are "hidden" from public access and, for security purposes, are more likely to co-locate participants in command centers than they are to co-locate instrumentation virtually. Likewise, there are undoubtedly private, proprietary implementations of the collaboratory environment, as in large corporations such as at IBM (1997) and Bellcore (Cooper 1993). Such sites are not pursued by this study.

SPARC : The Space Physics and Aeronomy collaboratory

The Upper Atmospheric Research Collaboratory (UARC) at the University of Michigan was undergoing major revision during the term of this research, and has been renamed the Space Physics and Aeronomy Research collaboratory (SPARC). SPARC is funded, in part, by the National Science Foundation under the Knowledge and Distributed Intelligence (KDI) initiative as part of NSF's Upper Atmospheric Facilities.[1] SPARC maintains a primary public web interface on the University of Michigan's Collaboratory for Research in Electronic Work (CREW) server at http://www.crew.umich.edu/UARC/.

Buried deep in the SPARC site is a form for requesting guest access or a logon account for access to the password-protected "working" portion of the

[1] The NSF's Upper Atmospheric Facilities include four large incoherent-scatter radar and optical facilities located along a longitudinal chain from Greenland to Peru. Along with the Sondrenstrom Facility which is accessed via SPARC, they include Millstone Hill Radar near Boston, Massachusetts, Arecibo Observatory in Puerto Rico, Jicamarca Radio Observatory in Peru (http://www.geo.nsf.gov/atm/uaf.htm#desc).

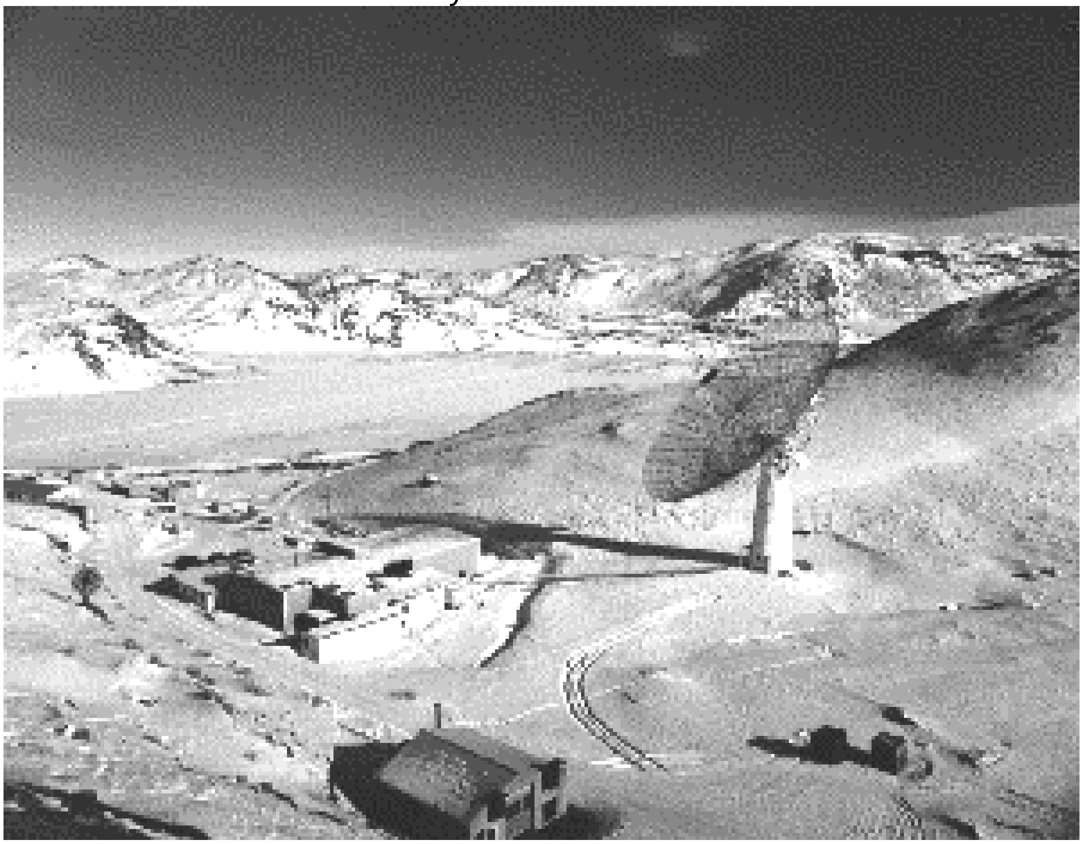
collaboratory. This access point is at http://sparc-1.si.umich.edu/sparc/central/edit/Login/REGISTER.

SPARC also maintains a separate K-12 education-oriented web site with games, experiments, and other instructional material at http://www.windows.umich.edu/sparc/. The education site uses Java technology and requires the most recent version of any cookie-enabled graphical web browser software for full functionality.

Upper atmospheric space physics is concerned with the effects of the sun on activities in the Earth's upper atmosphere. SPARC's purpose is to design, develop, deploy, and evaluate Internet-based technology to space scientists work together in collaborative studies of space and upper atmospheric science. The SPARC team includes an international community of space, computer, and behavioral scientists (http://www.windows.umich.edu/sparc/).

The primary instruments to which SPARC provides access are located at the Sondrestrom Facility, part of a global network of incoherent scatter radars. The Sondrestrom Facility is located north of the Arctic Circle in western Greenland. Figure 27 is an aerial photo of the Sondrenstrom Facility, and shows the transmitter and instrument building, power plant, and crew housing. SRI International, one of the world's largest not-for-profit independent research enterprises (http://explorer.csc.com/C3TEI/sri.html), operates Sondrestrom for a variety of universities and government labs.

Figure 26. The Sondrestrom Facility

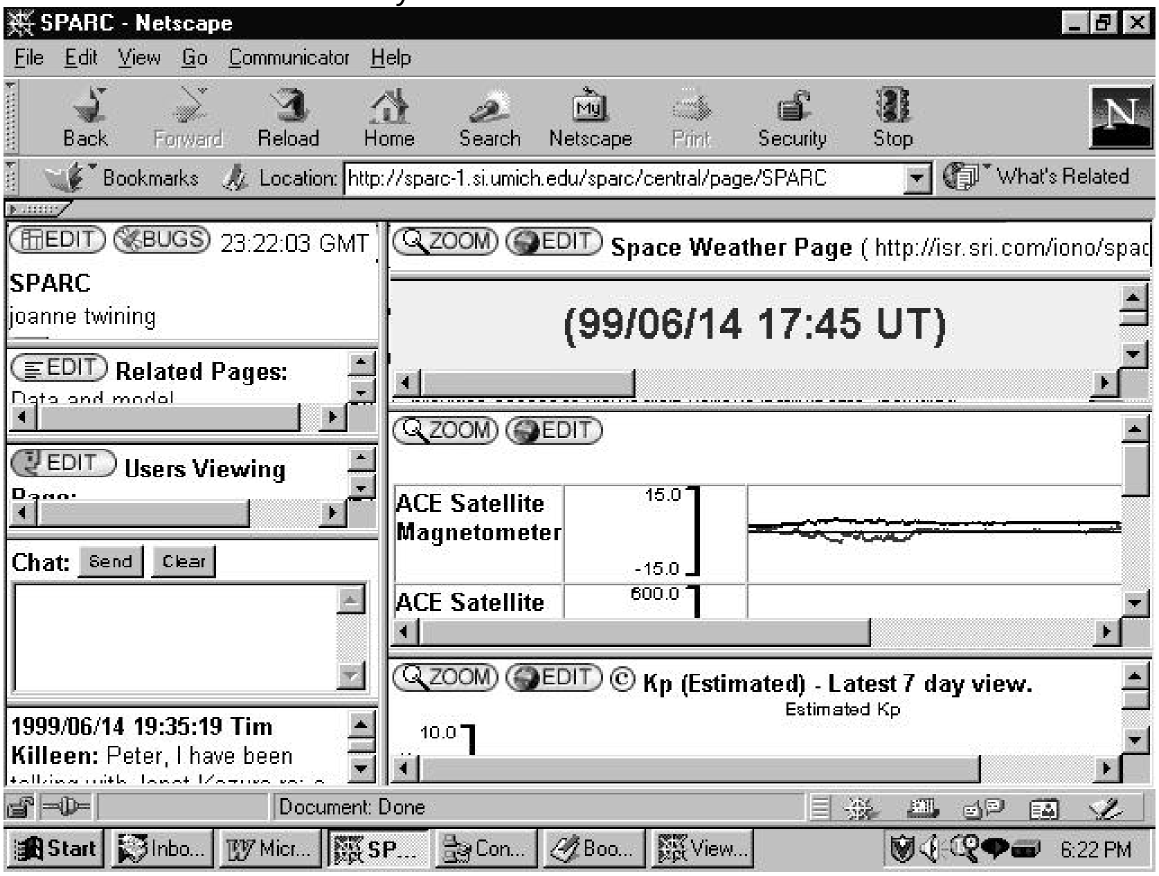


The Sondrestrom Facility's incoherent scatter radar has a fully steerable 32-meter parabolic antenna to which a wide range of optical and radiowave instruments are attached. Prior to the UARC/SPARC collaboratory, scientists had to travel to the site to use the instruments (http://128.18.44.75/iono/issfsond.html).

While SPARC provides access to various archived and live datafeeds from the various Sondrestrom instruments, manipulation of the remote instrumentation is available only through specially configured workstations at five locations. Real-time manipulation of the instruments is not available via public Internet and SPARC gatekeepers refused researcher access to the instrument controls.

As with the M2C Collaboratory's TPM web interface, the SPARC web interface is a sea of frames, a technique for dividing a web page into multiple, scrollable sections. The SPARC's main web page (Figure 28) includes nine frames, each with different options, and each less than an inch deep, making it necessary to scroll through the individual frames a line at a time, or open each frames in a new window. Even though the SPARC site doesn't transmit video, it does have a RAM-hogging refresh default, and quickly maximizes the desktop. Like the M2C Collaboratory's TPM web design, the site is awash in a cacophony of acronym-laden choices that may be difficult for users without a background in space science physics to understand. Unlike the M2C interface, SPARC offers a chat window and a list of others who are online, providing a real sense of co-presence that is absent from the M3C Collaboratory. However, a SPARC session feels more like reading than it does like seeing, as with the M2C experience.

Figure 27. SPARC collaboratory website



DOE2000

The U.S. Department of Energy, though its DOE 2000 initiative, and its prior initiative, the Distributed Computing Experimental Environment, or DCEE,[2] takes the lead in implementation of the collaboratory environment. Besides DOE's M2C TPM Collaboratory examined in the previous Chapter, and NSF's SPARC as described above, this research found two other functioning Collaboratories in the

[2] see http://www-itg.lbl.gov/DCEE/Overview.fm.html for DCEE Program overview and final report.

public domain that meet all the **CIRAL** criteria for inclusion, and both are funded by DOE.

Granted, successful collaboratory implementation depends on a great deal of basic and applied research, much of which cannot be easily identified for reasons discussed in Phase One. Any successful implementation rests on the very broad, if unrecognized shoulders of others. Such research and development areas include telecommunications, computer science, multimedia transfer protocols, and hardware and software development, and comes from virtually every quarter of the government as well as from industry and commerce. And, just as there has been no direct tracking of collaboratory-specific research, there has been no direct tracking of federal expenditure on research and development leading to implementation of the collaboratory.[3] Nevertheless, for bringing the research together and establishing a functioning collaboratory environment that is open to the public realm, DOE owns the prize.

Besides the M2C collaboratory, two other DOE collaboratory environments have been successfully implemented since 1996. The first is the Remote Experimental Environment (REE) collaboratory built around the DIII-D Tokamak fusion reactor at General Atomics Corporation in San Diego, California. The second is the Environmental Molecular Science collaboratory (EMSL) at the Pacific Northwest National Laboratory (PNNL) in Washington State. In addition to the technical and architectural research they provide, the REE project and EMSL

[3] for detailed historical tables of federal funding for research and development 1959-1999, see http://www.nsf.gov/sbe/srs/nsf99347/htmstart.htm

also each includes a sociological research component. The REE social research was contracted to Sara Bly, of Sara Bly consulting, but results of these studies are intended for internal use and are not published in the scholarly literature. They are, however, available online. This chapter investigates the REE DIII-D Tokamak Collaboratory and the Bly Reports, as well as the ESML Collaboratory, and its sociological and psycho-social studies, and concludes with a general synthesis of the environment of the functioning collaboratory.

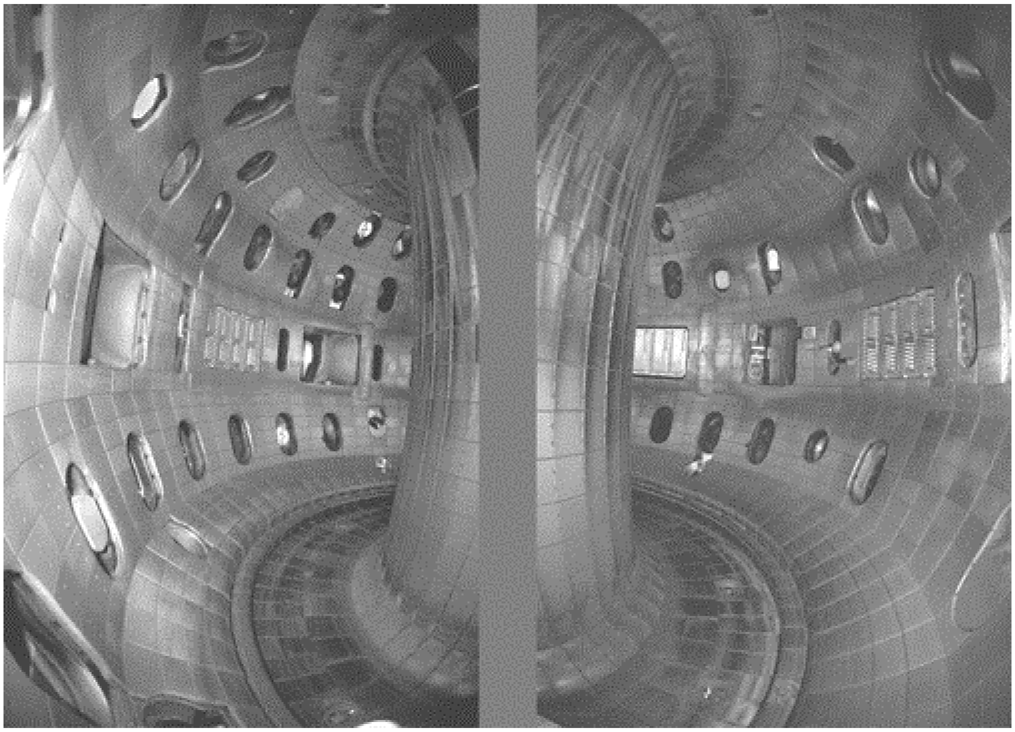
REE collaboratory Testbed at the DIII-D Tokamak

The Remote Experimental Environment (REE) collaboratory is a DOE-funded testbed built around the use of the DIII-D Tokamak at General Atomics Corporation. A tokamak is a

> toroidal plasma confinement device invented in the 1950s by the Russians Tamm and Sakharov …The word "tokamak" is a contraction of the Russian words: "toroidalnaya", "kamera", and "magnitnaya", meaning "toroidal chamber-magnetic.
> (http://fusioned.gat.com/webstuff/Tokamak.html)

More simply, the DII-D Tokamak it is a very large (about three quarters the size of a football field) donut-shaped (toroidal) chamber in which plasma, the fourth most common and the most abundant state of matter on Earth (along with solid, liquid, and gas) is magnetically suspended at very high temperatures, and into which hydrogen atoms (from water) are introduced, where they fuse, creating energy. Figure 29 shows the inside of the DIII-D Tokamak at General Atomics.

Figure 28. Inside the DII-D Tokamak



Hydrogen atoms are injected into the high temperature magnetically suspended plasma inside the tokamak, where they fuse, or come together, in a controlled thermonuclear reaction. The reaction creates helium energy the same way the sun and stars create energy. Unlike traditional fission nuclear reactors, which split apart heavy atoms such as uranium, and produce troublesome waste and other side effects, tokamak fusion reactors bring two light hydrogen atoms together. Unlike fission, the physics of fusion make it inherently safer: a fusion reactor cannot go through a meltdown. The waste generated by fusion is expected to be less radioactive and to have a much shorter half-life, and is thus easier to dispose of than fission waste. Fusion energy is also much cheaper to produce: 50 cups of sea water contain the same amount of energy as two tons of coal, or a thimble of the element created by the fusion process is equivalent to 20 tons of coal. If advancements continue at the present rate, it is expected that energy break-even with fusion reaction could be accomplished by the year 2010. Commercial power plants could then come on line just as the Earth's oil gauge becomes critically low, about the years 2050-2060.[4]

The DIII-D Tokamak at General Atomics is one of several major fusion programs worldwide. It is funded by the U.S. Department of Energy (DOE) and operated by General Atomics (GA). Collaborating organizations include Princeton Plasma Physics Laboratory (PPPL), Lawrence Livermore National Laboratory (LLNL), and Oak Ridge National Laboratory (ORNL), some of the very same labs involved in the M2C TelePresence Microscopy Collaboratory discussed in the previous chapter. As an international resource for plasma physics and fusion energy science research, scientists and researchers from around the world are also involved in REE Collaboratory sessions.

The complexity and cost of firing up a tokamak fusion reactor require a high degree of international collaboration. The energy released by the fusion reaction is abundant, and harvesting multiple very large datasets from multiple, simultaneous, sequential experiments is possible, and usual. There are a variety of instruments attached to the tokamak, and each is used for a different type of experiment, which take place simultaneously at various stations around the tokamak. Instruments can be added to, and removed from the tokamak as needed.

[4] The preceeding description is synthesized from an introductory slideshow available at http://fusioned.gat.com/Teachers/SlideShow/SlideMenu.html

Unlike the TelePresence Microscopy Collaboratory, which involves two or more remotely-located scientists working together at a single microscope with a single sample, DIII-D Tokamak collaboratory sessions involve hundreds of scientists and researchers. These scientists and researchers are either stationed at various instruments around the tokamak, in other buildings nearby, in labs elsewhere in the United States, or logged on from positions around the globe. Scientist may be conducting a shared experiment, or may be working on individual experiments, all at the same time, but everyone must nevertheless, and to varying degrees, stay in touch with each other and with the local or remote control team, which coordinates the overall "firing" of the fusion reactor. Unlike the M2C collaboratory, the tokamak equipment is not accessible to the public via the Internet. Rather, it is wired to an internal network to which password-protected remote logon must be arranged.

Coordinating communication between and among the hundreds of researchers, scientists, and technicians who are present at or logged on to any given tokamak sessions was the prime concern of the first REE collaboratory experiments in 1997. Rapid adjustments of instrument settings between the rapid series of shots that make up a tokamak session are often necessary. These fine-tuning adjustments rely on input from any of the people involved (Bly 1997). Communication is achieved via audio/video transmission from the DIII-D control room, and also includes the broadcast of meetings, use of inter-process communication software to post events to the network during a tokamak shot, the creation of a DCE (Distributed Computing Environment) cell for creating a common collaboratory environment, and distributed use of computer cycles, remote data access, and remote display of results (McHarg et. al. 1997).

Not everyone involved in a tokamak session needs to stay in touch with everyone else all the time, however, so communication channels are customized. Any given scientist might need to be communicating on more than one channel, to more than one subgroup of operators or scientists, at any one time during the experiment. Communication involves a very complicated, cacophonous, and continuous connection. Simultaneous communication between and operation of the instrument by as many as fifty scientists around the world is described by one participant as being very much like playing a musical machine:

> …you're playing it. And you're there all the time. You'd doing a dynamic experiment. You're continuously changing the experiment. You're watching the data as it comes out. You're adjusting things, you're measuring things, you're changing the experiment as you go along. (Bly 1997, 3)

In the preliminary report on the ways collaboratory technology was used and the effects of those technologies on the work of scientists and technologists during the 1997 experiments, Bly et. al. (1997) find that

> Cultural evolution in creating collaboratories may be as critical as computing technology development. The introduction of collaboratory technologies will influence the design and use of the technologies in ways that are different from many groupware situations today. Understanding the work activity as well as the technology is critical to the long-term success of these projects. (1)

At the beginning of her study, Bly (1997a) identified six aspects of scientific work common across the laboratories involved in the DCEE initiative:

1. Expensive and hard-to-duplicate equipment for data collection

2. Complex planning and coordination of the experimental design
3. Multiple person, multiple specialties needed for carrying out experiments
4. Rapidly iterating and changing experimental parameters
5. Collaborators who are geographically distributed
6. Shared analysis and results.

At the conclusion of her study, Bly (1997b) identifies seven characteristics of activity that took place during the 1997 DII-D Tokamak collaboratory session:

1. Nothing is fixed.
2. Everything is interdependent.
3. There are multiple paths to and from multiple sources.
4. Activity is constant and busy, but not frantic.
5. There is strong dependence on recognized, habitual signals and activities.
6. The frustration level varies.
7. The pre-operations meetings are valuable and relatively straight forward.

The REE DIII-D Tokamak Collaboratory experiment was much more a communication- and media-oriented event than are either the M2C or the SPARC projects, which maintain ever-ready and permanent presence on the Internet. An RRE Collaboratory session is a one-time occurrence involving hundreds of people with special configurations that rely on the coordinated firing of a very large instrument.

While collaboratory session in M2C and SPARC are also events, they occur within a stream. The M2C instruments are pretty much always available and require minimal preparation and coordination for special use, and the SPARC instruments are always online, supplying a steady stream of data. To continue the musical metaphor, M2C is like a CD player that can be turned on to listen to a specific piece of music. SPARC is like muzak in the background of a public elevator: its always on unless something's gone wrong. RRE is like assembling an orchestra for a command performance at the White House. Thus, we see, the type of instrument has much to do with the configuration of any given collaboratory, and so is necessarily the major influence on the other factors of a collaboratory ecology.

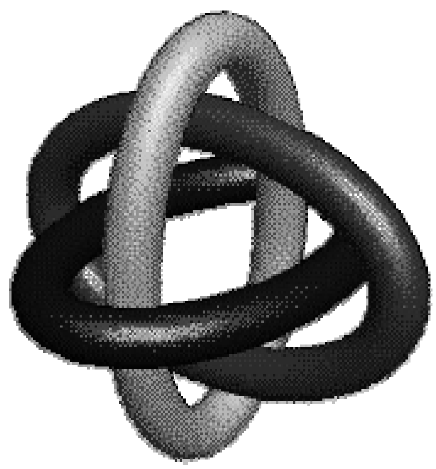
EMSL : The Environmental Molecular Science collaboratory

The Environmental Molecular Science collaboratory (EMSL) at http://www.emsl.pnl.gov:2080/docs/collab/ is concerned with the Department of Energy's mission to develop new technologies to clean up the nation's hazardous waste sites. EMSL is specifically but not exclusively concerned with the Hanford Nuclear Reactor site in southeastern Washington State. The Hanford Site has approximately 1.4 cubic kilometers of hazardous and radioactive wastes, 150 square miles of contaminated aquifer, and 60 millions gallons of radioactive wastes (260 MCi) stored in underground storage tanks (of which more than one-third are believed to be leaking). The Hanford Site also has 270 tons of spent fuel, 9 inactive reactors, and 7 major inactive reprocessing plants. The site is the equivalent of nearly 1400 Superfund sites divided into 78 distinct groups sharing common traits and geographies (Kouzes, n.d.).

The William R. Wiley Environmental Molecular Sciences Laboratory hosts EMSL at Pacific Northwest National Laboratory (PNNL) in Richland, Washington.

The EMSL "Virtual Scientific Facility" is operated by PNNL for the DOE Office of Biological and Environmental Research. The EMSL's philosophy of the synergy of the collaboratory is represented by its logo (Figure 30), which is Borromean Rings: three symmetric interwoven loops that cannot be pulled apart although no two rings are interlinked, and removing (breaking) one allows the other two to slide apart.

Figure 29. EMSL Logo



EMSL houses many unique facilities for basic scientific research, including the world's first commercial gigahertz Nuclear Magnetic Resonance (NMR) spectrometer, a scanning near field optical microscope, and the most powerful IBM parallel supercomputer yet built. Overall, EMSL is home to some 300 researchers with unique expertise, equipment, and software.

As part of its work, EMSL is developing software tools to support collaboratory operations, specifically CORE, the Collaborative Research Environment, a suite of software tools that provide general collaboration capabilities ranging from distributed file systems to videoconferencing. CORE supports real time computer display sharing, a Web based Electronic Notebook, secure direct acquisition control of the EMSL spectrometers, and the ability to consult with EMSL staff for training, as well as during setup to optimize the experiment, or at any time during analysis or other activities.

EMSL does not provide the public direct web access to any of its instruments, nor to active, ongoing collaboratory sessions. To participate in any of the EMSL activities, the CORE software must be downloaded and configured on the users' computer. CORE software may be downloaded from the EMSL main webpage at http://www.emsl.pnl.gov:2080/docs/collab/.

EMSL Collaboratory sessions are achieved using X-Windows technology, a cross-platform interface which allows remote scientists, regardless of the type of computer they're using, to telnet (or remote logon to the computer in the lab), open a graphical window on their computer, and participate fully in ongoing activities. Collaboratory sessions must be arranged ahead of time by submitting a proposal, negotiating the time for the experiment, acquiring passwords, and arranging for any necessary training. EMSL collaboratory sessions include both short term, project specific work, and long term ongoing collaborations. An example of an ongoing EMSL Collaboratory project is the one between Kelly Keating at EMSL (who does nothing but collaboratory science) and a scientist at Lawrence Berkeley Lab, which involves investigating the breakdown of DNA after exposure to radioactive waste. (Keating 1999).

Payne and Myers (1996) provide technical specifications for the architecture of the CORE suite. How a typical EMSL collaboratory session is set up between two remotely located scientists using the CORE suite is described in words and pictures at http://www.emsl.pnl.gov:2080/docs/collab/virtual/VNMRFScenario.html.

The CORE suite of software is designed to support four basic types of collaborations:

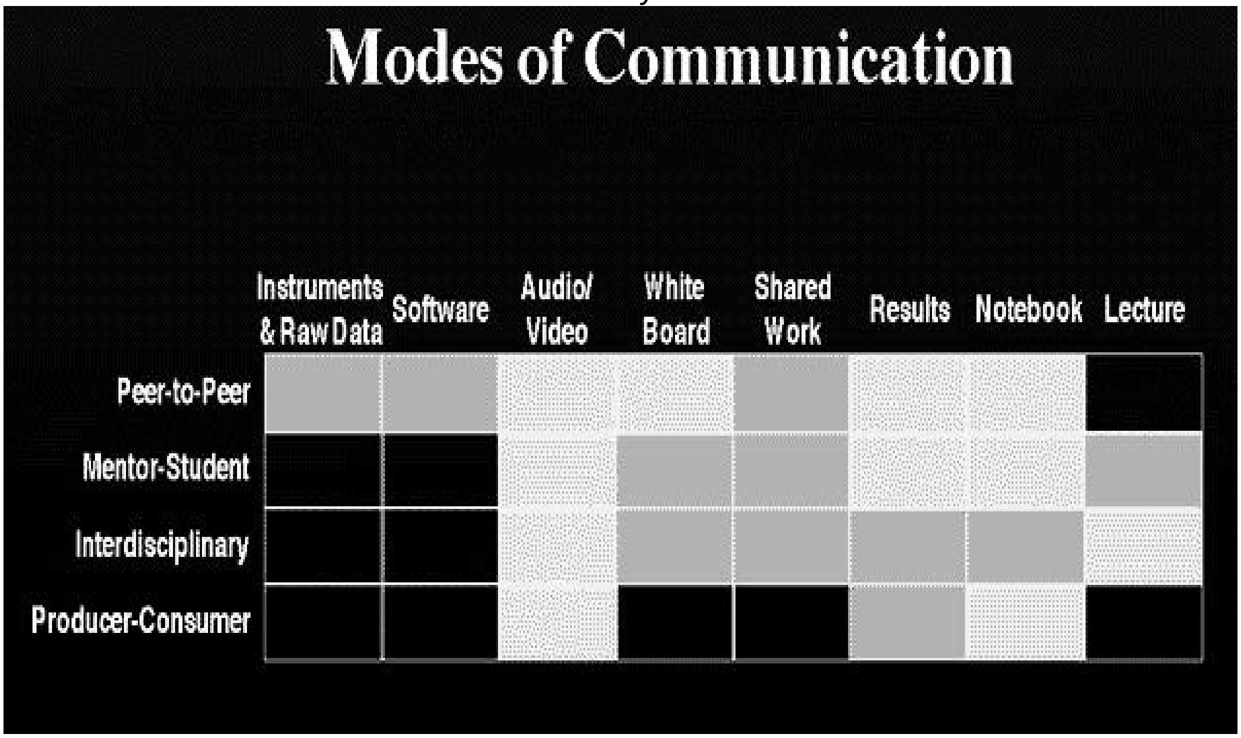
> **Peer-to-peer** collaborations between scientists from the same discipline who share the same vocabulary and focus, who need to share realtime manipulation of the instruments and have access to unanalyzed data, and who may be in competition;
>
> **Student-Mentor** collaborations where there is unequal knowledge, the introduction of new vocabulary, the need for reference materials and lectures, and the need to observe student efforts;
>
> **Interdisciplinary** collaborations between scientists from different fields, which may be bi-directional student-mentor projects, in which there are shared concepts, goals, and samples, but limited common vocabulary, where access to analyzed data and data summaries (rather than raw data) is needed, and communicate about the meaning of results is necessary; and
>
> **Producer-consumer** collaborations where resident scientists perform experiments and interpret the data on behalf of others who prepare and send samples to the lab, and in which there is no competition, few shared

concepts, and limited common vocabulary (http://www.emsl.pnl.gov:2080/docs/collab/research/CollabSociology.html).

EMSL research maintains that these four types of collaborations are universal and apply to education, medicine, and business environments, and that when one type of collaboration is undertaken, it often spawns one of the other types (Payne and Myers 1996).

EMSL's sociological research further identifies the communication needs of each of these four types of collaboratory sessions. Figure 32 is an EMSL-produced slide that shows the different modes of communication required for the four types of collaboratory sessions. (http://www.emsl.pnl.gov:2080/docs/collab/presentations/talk3.96/pg11.gif). The black squares represent modes that are not necessary, the medium gray squares represent modes of communication that are most necessary, and the light gray squares represent modes of communication that are moderately necessary.

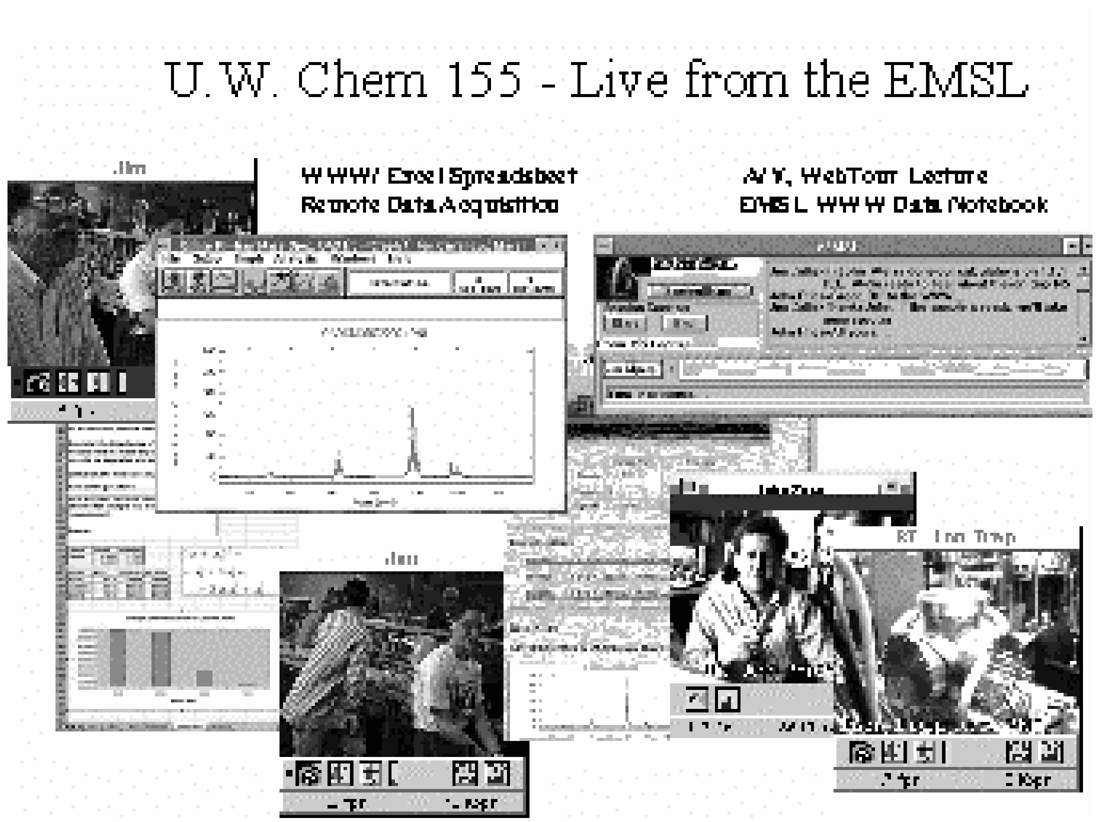
Figure 30. Communication needs of collaboratory Modes



The EMSL Collaboratory includes access to digital library services in the form of collected published and unpublished articles, including preprints. Digital archives include scientist and experiment Notebooks (password protected), software files, and system logs. EMSL also has an aggressive program of both public and formal education, and is developing interrelationships with several regional colleges and universities using its CORE technologies to improve undergraduate science education.

EMSL's collaboratory for Undergraduate Research Education (CURE) connects regional colleges/universities with each other, and with the lab. Again, downloading and configuring the CORE proprietary software is necessary to participate in the program, as is training, and arranging for access to password-protected portions of the site. Figure 31 is a screenshot of a CURE web session during an introductory Chemistry course at University of Washington. The CURE program is dedicated to the proposition that learning science can be done by doing science: interacting with real data and professional scientists, working and learning collaboratively project oriented and interdisciplinary work (http://www.emsl.pnl.gov:2080/docs/collab/projects/CURE/index.html).

Figure 31. EMSL CURE Session



EMSL's preliminary psycho-social research into the collaboratory is concerned with how the introduction of electronic communications might change and/or enable new types of collaborations, and how the synthetic environment of the collaboratory might be sustained. Findings of this psycho-social research are:

- **A collaboratory is a Social System**…nothing less than the village square and campfire juxtaposed to the Information Age
- It relies on **Social Mechanisms of Cooperation** in which collaborators must gain value from a relationship in order for it to be sustained
- It uses **Props and Controls for Social Discourse**
- **Either People or Work are the Focal Object**
- It relies on **Precedence and Ritual** organized by spatial syntax and by ritualized norms of acting
- **A Sense of Place is Essential.** (Wise n.d.) (emphasis added)

Conclusion

Chapter Six explores the Materials MicroCharacterization (M2C) Collaboratory's five TelePresence Microscopy (TPM) sites, and tests the usefulness of the **CIRAL** matrix of criteria for inclusion that was developed in Chapter Five. Chapter Seven explores "derivative collaboratories" that do not meet the **CIRAL** criteria for inclusion along with three additional Collaboratories that meet the **CIRAL** criteria. Case studies of the Space Physics and Aeronomy Research collaboratory (SPARC), the Remote Experimental Environment (REE) collaboratory experiment at the DIII-D Tokamak, and the Environmental Molecular Science Laboratory (EMSL) collaboratory are developed.

Most of the thousands of online environments that use the word "collaboratory" fail to meet all the criteria for inclusion because they do not provide access to and manipulation of remote instrumentation. There are undoubtedly other functioning collaboratories operated by business or the government which are outside the reach of this research because of privacy or security reasons.

Each of the four collaboratories explored in Phase Two is a unique information environment. The instrumentation to which each provides access is a major determining factor of the nature of those environments.

In preliminary sociological studies, EMSL identified four types of collaboratory: peer-to-peer, student-mentor, interdisciplinary, and producer-consumer, and identified the communication modes required for each. Each of the four Collaboratories serves a different scientific purpose, and each deploys different combinations of communication functions.

Comparison of collaboratory architectures

The M2C Collaboratory focuses on the simultaneous transmission of microscope and macroscope video views to and from solo scientists working together online with one of the five TPM sites. Participants have access to and can manipulate remote instrumentation, are supplied raw data in the form of microscope lens views, or lens view snapshots captured in collaboratory notebooks. Digital archives, and limited digital library resources are available. There is no need for special software. M2C does not support audio, whiteboard, chat, or forumfunctions. M2C functions as each type of collaboratory, but does not provide the full range of communication functions described by ESML.

The SPARC collaboratory is centered on the delivery of text and chart-type data constantly transmitted from remote instruments, and text-based communication among unlimited participants logged on via the Internet using publicly-available software. Participants have access to the data flow from but cannot control the remote instrumentation. Control of the instrumentation is limited to five specific sites that use specialized workstations. SPARC does not offer collaboratory notebook, whiteboard, audio, or shared work space functions, nor does it support lectures. SPARC functions as all but the Producer-Consumer type of collaboratory, but does not provide all the communication modes identified by EMSL.

The REE-DIII-D Tokamak Collaboratory experiment is centered around the audio and video intercommunication among hundreds of scientists, researchers, and technicians who either have their hands directly on some aspect of a very large piece of equipment, or are logged on remotely to participate in experiments or control the tokamak. REE-DII-D Tokamak Collaboratory sessions are events that require considerable preparation and coordination. Special software is required to participate, and all the communication modes identified by EMSL are offered. The REE-DIII-D Tokamak Collaboratory experiment functions as all four types of collaboratory.

The EMSL collaboratory is focused on facilitating shared, project-oriented experiments in either one-time mode or on a continuing basis. Special software is required to participate. EMSL does not make its instruments available via public access to its website. EMSL functions as all types of collaboratory, and includes all the communication modes identified.

Comparison of collaboratory data types

The primary data produced by the M2C collaboratory are single shots taken

transmission from the labs, the archives of the site's distributed email list, and the content of researcher notebooks. The primary data produced by SPARC are

of the text-based chat sessions. The primary data produced by the REE DIII-D Tokamak Collaboratory experiment are multiple, huge data sets created by

records of the intercommunication between experiment participants. The primary data produced by EMSL are raw data from the instruments, and the contents of

session logs.

CONCLUSION OF PHASE TWO

has many manifestations. The common element of these environments is the humans who participate in collaboratory activities. The differentiating factor is the

and uses seems to determine the level and sophistication of communication functions built into each collaboratory.

understanding the sociological aspects of the collaboratory, specifically communication needs of collaboratory participants and the nature of collaboratory work activity. Preliminary sociological studies are underway. From the REE DIII-D Tokamak collaboratory Experiment, Bly (1997) identified six characteristics of collaboratory activity and six aspects of collaboratory work. EMSL researchers identified six psycho-social conditions of the collaboratory, and clearly identifies the collaboratory as a social environment. This study finds that the social environment is instrumentally-determined. Whether and how the synthetic environment of the collaboratory can be sustained will be determined by improvements in communication specific to each environment. Bly believes the answers to the problem of collaboratory communication will be found in understanding the cultural evolution of the collaboratory.

It is clear that future collaboratory research must expand to include and concentrate on the social aspects of the environment, and that Nardi and O'Day's (1999) key constituents of information ecologies offer a guide to that research. It is also clear there can be no one "formula" for collaboratory configuration. Each collaboratory is a unique environment determined by its instrumentation, and so also is a unique ecology.

With this overview of the environment of the collaboratory, this research may now turn its attention to the people, and the practices and the values of the collaboratory. Ecological studies must examine how collaboratory work is conducted, how the subtleties of human communication might be better mediated technologically, and how collaboratory skills might be taught to prospective participants. These types of studies require the researcher to have access to the collaboratory's physical location. Phase Two's subjective examination of the

Phase Three of this study takes a preliminary step toward such expansive research by engaging collaboratory Pioneers identified during Phase One and

for the collaboratory and identify the skills collaboratory Pioneers value in prospective participants.