# Social Theoretical Issues in the Design of Collaboratories: Customized Software for Community Support versus Large-Scale Infrastructure

Geoffrey C. Bowker

Susan Leigh Star

University of California at San Diego

*The same factors which have thus coalesced into the exactness and minute precision of the form of life have coalesced into a structure of the highest impersonality; on the other hand, they have promoted a highly personal subjectivity.* Simmel, 1960, p. 413

## 1. Introduction: The Problem of Electronic Community

- ☐ I live on the net.
- ☐ The Internet is a new nation.
- ☐ The laboratory of the 21st century has no walls and no boundaries, but is a virtual community.
- ☐ We are all Netizens now.
- ☐ Electronic communication has revolutionized the way science is done.

Those of us studying the use of electronic media are often faced with statements such as these. The popular media often confound daily life and routine work and practice, electronic communication such as email, and that which transpires "over the net," with the concept of "community." There is a substantial elision of experience, material conditions, structural positions in particular social networks, communication, and location in discussions of "the net," and "the web."

As social scientists, this worries and intrigues us. We have been involved in the development and evaluation of several software systems for use by scientists.

Whatever is happening with the scientists for whom the software was designed, it is clear that neither custom software efforts, nor the larger electronic environment are *literally* places where people live full time, nor even work full time. It does not provide either face-to-face local community nor full working infrastructure, as those concepts might plausibly be applied to close-knit collegial networks or to occupational communities and labor pools.

Early sociologists struggled along similar lines with the concept of community as it was affected by another large-scale technological shift, the industrial revolution. They used the term *gemeinschaft* to refer to tribal and village life, the sort one is born into, where everyone knows who you are by virtue of family and station. At first glance, it seemed that this sort of life was being threatened or even destroyed by cities and factory work. *Gesellschaft* referred to its opposite: the world of the industrialized city, marked by commerce, where who you are reflects your place in a more distanced, formally commercial and secular order. There were many arguments about whether urbanization had destroyed *gemeinschaft*, both among social scientists and in the general public sphere. Since their use during this time, the terms have been generalized more loosely to refer to intimate versus distanced relationships. (As well, the early sense of the disappearance of community has been modified and complexified.)[1]

During the infrastructural shift that is occurring with the building and integration of global electronic networks, some of the past lessons from sociological studies of communities and networks are useful. With the interest in terms such as *electronic community* today, the intimate versus distanced distinctions are again important. We seek here to clarify some senses in which *both* apply to the development of collaboratories. We are learning, as did early sociologists, that online community is not a choice between familiar locale and alienated metropolis, but that elements of both are important for design and analysis (Bishop, et al., in press).

The work of doing science still takes place substantially off line, even though electronic communication and data sharing aspects are becoming increasingly more important. Even the parts of the work that may be isolated as "information work" are not necessarily conducted via electronic media: People talk to each other, run down the hallway, write things by hand in notebooks and on labels, and also FAX and telephone each other constantly. Further, one is not "born" into the scientific "community, " but one's sponsorship and apprenticeship occur over a long period of time "offline" for the most part, in graduate school and via long-standing collegial and friendship networks. At the same time, electronic tools are of growing importance, and elements of nature are being tested, taught, and modified in virtual space.

Bowker (1994a, 1994b) has used the concept "infrastructural inversion" to describe a conceptual shift in the social studies of technology, especially in history of technology. He implies that a figure-ground gestalt shift has occurred: decentering individuals, single artifacts, or even social movements as causal factors in large-scale scientific change. Instead, when infrastructural change is treated as the primary phenomenon, collective processes of transformation are more richly explained. Changes in infrastructural networks such as transportation, information, and domestic technologies explain a great deal about other forms of social change and social relationships-they are not simply substrate, but substance (Bowker & Star, 1999; Star, 1999). If this is true, then the substantive changes effected by (among many others) the National Information Infrastructure Initiative, the Collaboratory efforts, and the Digital Library initiatives have significance both for science and far beyond it. We now have a chance to observe this phenomenon as it unfolds. It is a moment that will not recur, and which requires extraordinary effect from social, information and computer scientists. One democratic concern here is that the social inequities and distributions of information resources will become somehow frozen or reified in the infrastructural changes. This begins with the simple notion of "information rich" and "information poor," but extends more complexly as more of life's business is conducted electronically. The digital divide has become omnipresent.

## 2. On the Concept of Community: Lessons from History

Infrastructural shifts on the scale of global electronic networking are not all that common in human history. They tend to generate extensive discussion of "basics": values, concepts, moral directions. During the great shifts of the 18th and 19th centuries from rural to urban, agricultural to industrial, and to large-scale organized capitalism, some basic questions were raised about community that echo those found today in *Science* magazine's articles on the global electronic laboratory and very large databases.

Distinctions between types of human bonding-and the troublesome notion of community-indeed informed the founding of the discipline of sociology. Nineteenth-century sociologists worried a great deal about the forces, or nature of the links that hold people together (e.g., Durkheim, 1984; Marx, 1970; Simmel, 1950; Tönnies, 1957), which seemed to be undergoing massive transformation throughout the Industrial

Revolution. Did movement from the village, farm, and tribe to city, factory and contracts change us, body and soul? Did what held us together in small, rural groupings differ radically from that which held us together in large urban conglomerates? If so, how?

Volumes, if not libraries, have been written on the contrast between *gemeinschaft* and *gesellschaft*, including the authors who find no essential change in human organization and bonding between country and city, farm and factory. The concept of *community* and all it might entail is one of the most contested in the history of social sciences. Some social scientists, such as Simmel (1950), held with confidence that the metropolis has spawned a new "mental life." In his classic essay on the topic, he claims that: "There is perhaps no psychic phenomenon, which has been so unconditionally reserved to the metropolis as has the blasé attitude."

Others went on to investigate the notion in rural societies in less-developed countries, such as Redfield (1947). As time went on, their claims were modified and challenged by other social scientists such as Lewis (1949) who found plenty of evidence for conflict and arms'-length relationships in the village, and Willmott and Young (1960) who found plenty of *gemeinschaft* in the close-knit working class neighborhoods of urban London's East End. It seems that community is not such an easy notion to place, nor such an easy one to dispense with.

Social scientists need concepts to describe how it is that people "stick together" within groups. But as with other fundamental concepts, such as "species" in biology or "culture" in anthropology (Clifford & Marcus, 1986), ironically it becomes both the glue that binds and the thing *least* agreed upon. (Perhaps it is just that, as Simmel noted long ago, conflict itself is "socializing," and the agreement to disagree is what binds these disciplines together.) Stacey (1969), in a magisterial review of the decades of arguments about the term *community*, notes that "It is doubtful whether the concept 'community' refers to a useful abstraction. Certainly, confusion continues to reign over the uses of the term" (p. 134).

Like "power" or "profession," the term community clearly points to a phenomenon of key interest for understanding large-scale change-but exactly what, or why, remains elusive. Stacey goes on to note that while one author deplores the romanticism of the notion, another claims it as the framework in which humans are introduced to civilization itself:

> This is so vague as to be nonsense: there is no such thing as community which does this. Various agencies are involved in this process of introduction, perhaps neighbors, almost certainly parents' kind and friends (which may live next door or miles away). These institutions may, or may not, be locally based. They may, or may not, be inter-related. If they are locality-based *and* inter-related then there may well be a local social system worth studying, but one would hesitate to call this a community. Nor is there any less lack of confusion in earlier usages of the term. (p. 135)

The conflation and generalization of these terms continue to this day. *Communities* seem to be everywhere in the media: the diplomatic community, the international community, the women's communication, the African-American community, the high-tech community, the bird-watching community. Perhaps our favorite of the season has been the "heterosexual community." What is the concept that joins such disparate collections of allegiances and heterogeneities?

In the world of science, the notion of "scientific community" (and its cousin, "invisible college") has been equaled only by the term *paradigm* in the amount of disputed terrain contained within it. Methodological divergences in how one would measure a community, where its boundaries are, whether it's a meaningful unit of analysis, and so on, have enriched and confused the world of social scientific work about science (STS, or science, technology and society). Alternative notions have been proposed to emphasize one dimension or another of the social relations under scrutiny: communities of practice (Lave & Wenger, 1991); social worlds (Clarke, 1991; Clarke & Montini, 1993; Strauss, 1978); relevant social groups (Bijker, 1995), and actor networks (Callon, Law & Rip, 1986), to name just a few.

The development of large-scale electronic infrastructures have given a new life to the battered notion of community, and have direct relevance to collaboratory research and its descendents, such as studies of cooperative work on the web and related CSCW (computer-supported cooperative work) projects. Early on, these events had two basic foci: *custom software projects* (such as the NSF collaboratories) which remotely link research already joined by common interests or heritage; and *tools for browsing* the extant internet resources in order to discover useful information or possible colleagues.

The collaboratory concept emerged in the late 1980s from a top-down initiative from the National Science Foundation in Washington. Dr. William Wulf, then director of the NSF Directorate for Computer and Information Science and Engineering wrote a foundational white paper: "The proposal, then, is to undertake a major, coordinated program of research and development leading to an electronic "collaboratory" a "center without walls", in which the nation's researchers can perform their research without regard to geographical location, interacting with colleagues, accessing instrumentation, sharing data and computational resources, accessing information in digital libraries" (Lederberg & Uncapher, 1989, p. 19). Wulf's paper recalls Vannevar Bush's canonical *Science-The Endless Frontier*, which led to the foundation of the NSF. Bush had also called for improvements to scientific information flow. Bush's logic had been government through big science will give citizens health, wealth and security -his clarion call opened with the words: "The Government should accept new responsibilities for promoting the flow of new scientific knowledge and the development of scientific talent in our youth. These responsibilities are the proper concern of Government, for they vitally affect our health, our jobs, and our national security" (Bush, 1990 [1945] p. 8).

Wulf's logic, along similar lines, was that government through distributed science will give the nation technology, which leads to economic and military stature and competitiveness-in the opening words of his paper: "The health of the United States, economically and militarily, depends on its technology base. The technology base depends on the number, quality, and productivity of the nation's research scientists and engineers" (Lederberg & Uncapher, 1989, p.19). Several important historical developments have occurred between the time of Bush and that of Wulf, of course. Most notably, has been inserted as a major mediator between science and nature (through simulations and advanced instrumentation). A new kind of information technology has become an essential support for scientific work, which has important implications for the practice of science as well as the notion of community.

Wulf's paper led to a workshop held at Rockefeller University in March 1989 (Lederberg & Uncapher, 1989, p. i). The workshop report found several needs for a collaboratory. Instruments, it was noted, were often to be found in environments hostile or inaccessible to people. Further, people themselves were inconveniently distributed, so that remote interaction with colleagues was necessary "whenever the appropriate mix of talents to address an interdisciplinary problem is not collocated anywhere." Finally, not only people and instruments, but also data were distributed-remote interaction was needed when: "the data are too vast to be replicated and managed at a single location," for example with the global seismic database. Accordingly, the goal was to build: "no less than a distributed intelligence, fully and seamlessly networked, with fully supported computational assistance designed to accelerate the pace and quality of discourse, and a broadening of the awareness of discovery: in a word, a Collaboratory." Here then is one defining reading of the word: collaboratories are about distribution in every sense of the word-distribution of things, people (thus Gallo formed a distributed college without walls-a "dream team" to fight AIDS) and information. This gives a great deal of scientific and theoretical power to computer and information scientists-who have developed theories of distributed databases and long-range communication networks; it allows a translation between government imperatives and new tools being developed within computer and information science: "The Collaboratory is much more than just a set of tools. It is a national computer-based infrastructure for scientific research."

The phenomenal growth of the web scarcely needs any review in this venue; the number of users grows exponentially. At the same time, there is considerable concern about the lack of good indexing and sorting tools, as well as a need for better means of ascertaining data quality. Thus the distinctions between custom projects and the development of the web are not absolute, bur rather foci that may blur into each other over time.

Even custom projects interface with the larger networks through email and links with databases (and increasingly this is the case for multimedia as well); browsing tools when adopted and extensively used in local sites come to form a *de facto* custom package over time.

With the advent of the Web, the repertoire of social science tools for its analysis has also vastly expanded, to include social studies of structural and nascent communities; web-based social networks (Kiesler and Sproull, 1996); online/offline "ecologies"; digital libraries (Bishop, et al., in press; Weedman, 1992), and the use of advanced modeling techniques including visualization and VR.

### 3. **How Can Electronic Networks Support a "Community"?**

In spite of all the controversy associated with exact definitions of community, there is general agreement that the sense of community rests on nontrivial, ongoing relations among people; some degree of shared knowledge, understandings, material objects, or conventional practices; and the idea that these two are not independent. Initial research on computer-mediated collaboration, or electronic communication systems showed some interesting effects in terms of the social relations, or role differentiation among members (Sproull and Kiesler, 1996). This in turn affected shared understandings, or consensus formation in decision making. A common finding in this research was that technologically mediated communication creates less role differentiation among group members that did face-to-face communication. This was attributed to the fact of less visible differentiation between group members when communication occurs electronically (you can't see someone's race or gender, for example; you couldn't hear their voice; you can't impute social cues about power or position by looking at their dress or posture). Consequently, concluded these researchers, there tends to be more uninhibited communication, more (vertical) communication between members of different status, and more equal participation in groups communicating via electronic mail (Kiesler, Siegel, & McGuire, 1984; Rice and Rogers, 1984; Siegel, Dubrovsky, Kiesler, & McGuire, 1986).

Researchers argued that the nature of decisions and agreements in such computer mediated groups may be affected, in part, by the decreased role differentiation that occurs in come computer mediated groups. Some researchers have found that less role differentiation allows group members to generate more ideas in problem-solving situations (Connolly, Jessup & Valacich, 1990; Jessup & Tansik, 1991). When there are more ideas to choose from, group members may have a harder time deciding on a solution creating a longer time to reach consensus, or a point of action. This is obviously a relative notion, depending on small, identifiable groups with co-generated histories. Whether these ideas scale up to larger electronic groups is highly debated both about whether the distinction is too simplistic, or about whether other forms of social cues emerge over time. The larger and more distributed the groups, the more difficult status-role differentiation is to analyze. Baym (1995; 2000) argued that emergent groups on the net, such as Usenet groups, develop strong cues within the email messages themselves, including differentiation, leadership, and novel semantic conventions based on practice.

A linked finding was that groups within organizations communicating with computer mediation seemed to take longer to reach consensus than groups communicating face-to-face. Without traditional social cues to inform decision making, the process become more open, less reified or controlled by traditional lines of authority. Siegel et al. (1986) reached the same conclusion after observing greater choice shift and inefficient communication among group members communicating electronically than among group members trying to reach consensus in face-to-face settings.

On the other hand, Galegher and Kraut (1990) suggested that decreased role differentiation has the opposite effect on consensus formation. They focus on the diffuse responsibility and joint ownership that results from electronically mediated joint production. With less individual identification with the product, they argue, there will be a greater tendency toward conformity thus speeding up the consensus formation process.

In science, technological mediation-or electronic community systems- are becoming routine parts of scientific teamwork, both locally and at a distance. Modern, "big" science consists of scientists addressing problems so large they cannot be solved by a lone scientist. Such large problems require the combined efforts of scientific team members. Historically, one of the primary determinants of scientific collaboration has been physical proximity. However as the scope of scientific endeavors grows and members of the scientific community are more mobile, scientists working on the same problem often are not co-located. Physicists have a particularly well-developed infrastructure to support international collaboration of this sort (Traweek, 1988). Therefore, the benefits of productive collaborations afforded by proximity must be achieved by other means. It was claimed and hoped that electronic community systems will provide the means for frequent interactions, and joint access to tools and information for scientists collaborating long-distance (Kraut, Egido & Galegher, 1990; Lederberg & Uncapher, 1989; Schatz, 1992). Since the mid-1990s, most of these tools operate via or alongside the Web.

One feature often integrated with electronic community systems is, of course, email. Early studies, partly in dialogue with the organization decision support studies of electronic "rooms" and local electronic messaging, tried to show how email would affect role differentiation and consensus formation in scientific collaborations. The findings were somewhat inconclusive. Tombaugh (1984) argued that greater role differentiation would have facilitated the international communication in an international asynchronous messaging, or conference system. Scientists in his study felt the need for more leadership. Hiltz (1983) studied 103 scientists communicating asynchronously over the net. This technologically mediated communication resulted in greater perceived understanding of other scientists' interests or theories, thereby affecting problem solving and decision making. Although this early research did not take place in anything like the Internet environment, it is useful in identifying problems in the design of electronic community systems for large, geographically dispersed scientific communities, and for illuminating some of the conceptual problems addressed earlier in this chapter.

Electronic infrastructural developments have made it possible to include in electronic community systems a variety of functions in addition to electronic mail (e.g., information sharing, document editing and collaborative writing tools, and data visualization techniques and collaborative tools). Although the emergent functionalities of a community system will depend on how community members use it, there are two fundamental functionalities for which such systems are designed:

1. An electronic system is designed as a quicker means for existing communications and social relations that compose a community. For instance, in a scientific community, journal articles can be viewed as a mode of communication of ideas in science-and electronic journals and pre-prints can provide the same function, more quickly.
2. The electronic medium can be used to support new activities. Different things are possible on electronic systems and the notion of scientific community can be changed. For instance, traditionally by keeping all information about an area available online, with annotations by and dialogues among experienced professionals, the training of newcomers can become an activity supported by the electronic system.

The degree to which these capacities of electronic community systems may lead to changes in the scientific community itself are only beginning to be analyzed. The first capability may lead to quicker spread of information or diffusion among a relatively well-defined group of scientists. The second approach may make it easier for new people to come into a specialty, and more difficult to maintain boundaries. In other words, electronic community systems may have the effect of reducing the gatekeeping roles of interpersonal relations (this is one hope of public policy advocates of such systems). However, if large resource differences continue to stratify electronic infrastructures (e.g., with respect to advanced multimedia) they may reinforce extant disparities.

These are some of the things that might be considered by designers of electronic community systems for scientists. An electronic system can be developed for a specific scientific community by community members or by outsiders. In addition, a "generic" electronic community system can be developed to support all types of general activity. When a scientific community adopts a generic electronic community system, its members can adapt some of the systems features for their specific use. Schatz, for example, proposed migrating the *c.elegans* WCS (discussed next) collaboratory system to other scientific communities, and this has been implemented for drosophila and yeast, among others[2].

## 4. **The Worm Community System (WCS)**

Scientists have, of course, been remotely operating scientific equipment with computer mediation for more than 20 years (e.g., the space program-see Olson & Atkins, 1990 for a discussion of this point). New to the collaboratory was a suite of communication and collaboration tools that allow distributed scientists to work together on data. This novelty can be seen in the Worm Community System (WCS), a customized piece of software designed to support the collaborative work of biologists sequencing the gene structure, and studying other aspects of the genetics, behavior and biology of *c.elegans*, a tiny nematode (Schatz, 1991)[3]. It is a distributed "hyperlibrary," affording informal and formal communication and data access across many sites. It incorporates graphical representations of the physical structure of the organism; a periodically updated genetic map; formal and informal research annotations (and in this way constitutes an electronic publishing medium); directories of scientists in the community; a thesaurus of terms linked with a directory of those interested in the particular subtopic, and a quarterly newsletter-the *Worm Breeder's Gazette.* It also incorporates a database independently developed in Europe designed for the community, *acedb*. Many parts of the system are hypertext-linked with each other.

WCS was developed in 1990-93 by Bruce Schatz and Sam Ward at the University of Arizona (after which it moved to the University of Illinois, and parts of its core programming were adapted in the Illinois Digital Library Project, 1994-98, and in the ongoing Interspace Project). WCS was designed for a particular group with the idea of eventually migrating the structures of the software to other groups. Like the Sequoia 2000 project which deals with global change, this project was seen as being interesting both for its domain specific support and the nature of the computer science involved- meeting the information needs of scientists as well as providing a challenge for basic computer science research. (Some of the complexities and difficulties of this relationship are explored in Star and Ruhleder, 1996; and Weedman, 1992.)

Star and Ruhleder (1996) worked as ethnographers on the project, considering potential sociological effects and dynamics within the system as a whole. They traveled to worm labs across the United States and Canada, interviewed and observed the use of computing and other features of "worm work," including aspects of routine work and communication. As well, they asked questions about other features of the work, such as scientific careers in biology, competition, routine information-sharing tasks, and how computing infrastructure is managed. They visited more than 30 labs and interviewed more than 100 biologists during a 3-year period.

The worm community (their term for themselves, by the way!) consists of more than 1400 scientists distributed around the world in more than 120 laboratories. They are a close-knit community and are very friendly. Until recently, most people were first or second generation of the field's founders. After the choice of *c.elegans* as the "model organism" for the Human Genome Initiative, some of this been changing (with increased resources, visibility, and competitiveness). Model organism means both that the actual findings from doing the worm biology and genetics will be directly of interest to human geneticists-for example, when homologues are found between oncogenes (cancer-causing genes) in the worm and in the human (although worms do not get cancer as such, there are developmental homologues) and that tools and techniques developed in the *c.elegans* mapping effort are put to use in the human project.

The worm itself is remarkable both as an organism, and as a component of a complex pattern of information transfer integral to the biologists' work. It is microscopic and transparent, with the surprising and convenient capability of being able to be frozen, mailed to other labs via parcel service, thawed out and retrieved live for observation. Worms and parts of worms travel from one lab to another as researchers share specimens. Worm strains with particular characteristics, such as mutation, may be mailed from a central Stock Center to labs requesting specimens. Tracking the location and characteristics of organisms thus is an important part of record-keeping and information retrieval. Two points emerge immediately here. First, both the worm and the scientist become part of a single distributed community. The worm biologists travel a virtual network and their community only exists at their set of nodes; their subjects travel a mail network, and live only at the nodes. Second, and most significantly, the worms and the scientists can only travel (really and virtually in turn) if there is a common set of standards shared across the network. Protocols must be in place to ensure that a given set of electronic signals means a contribution to a distributed database, and will be understood as such at each point along the way. Equally, there must be protocols to ensure that a given package of worm tissue means a standard mutation in each lab in the node. Berg (1997) calls this the need to discipline local practice-and he notes that his is a paradoxical feature of the attempt to create a transparent, flexible infrastructure.

Usage patterns in WCS reveal the sometimes competing nature of custom versus global emphases in information systems design. Many potential users of WCS moved to simpler, less functional web tools such as Gopher, Usenet, or simple email (and after the Web, migrated there). Star and Ruhleder (1995) analyze this in part as an unfamiliarity with some of the infrastructural tools, such as the Unix operating system, as well as other aspects of local infrastructure and support systems.

## 5. **The Collaboratory and the Nature of Work**

Ruhleder (1995) has written about the ways in which classical Greek scholarship changed with the introduction of the Thesaurus Linguae Graecae. This complete canon of classical Greek literature on database was made available online, during the 1970s with updates since. Tracking down the occurrence of a single word throughout that canon, with a view to uncovering its modalities, used to be a lobar involving sensuality (the feel and smell of the book so beautifully rendered by Charles Lamb), prodigious memory and a goodly set of notes. It was not something one did at the start of one's career. One learned, though apprenticeship over the years, how and what and where. Now this work can be done with the touch of a button by a graduate student embarking on a PhD. Is it the same work in those two instances? Or-like Borgès' two Don Quixotes, one written by Cervantes and one by a later academic-does the work itself change with its context, in this case its infrastructural support? We have so far looked at theories of work practice informing the development of collaborative infrastructure; in this section we explore how what it is to be a scientist is being affected by the development of high speed networked information infrastructures.

Steven Hawking speculated that by the turn of the century theoretical physics would be province of the computer, the role of the human being would be to attempt to understand and appreciate discoveries made elsewhere. Although this seems unlikely, there are two basic ways in which the new information infrastructure is radically changing the nature of scientific work: the nature of representation and the nature of the scientific product.

The *nature of representation* is changing in the sense that theoretical work is increasingly being delegated to the intelligent instrument-which works through terabits of data streaming in and decides (by one of a number of algorithms) which data is interesting and which is not; and then represents the interesting data graphically according to another set of algorithms. The interpretive work is deliberately partially delegated to the machine, in order to cope with information retrieval. It is of course true that the appropriate algorithm can be changed, but one suspects that once this act of delegation is made and ramified (as one infrastructure submerges into another) then the attached algorithms will be naturalized (in the anthropological sense). (A relevant robust finding from library science is that patrons will use a convenient electronic source they know to be incomplete in preference over a card catalog that they know to be complete.) There will literally be no other way to see the world.

The nature of the scientific product is therefore changing in several ways. The scientific paper is arguably no longer the *terminus ad quem*. It is an archival document of use to people in other disciplines, argued one researchers (*Science*, 3/24/95, p. 1764). It is certainly increasingly the case the publication is proceeding online. This means more than the mere transposition of linear texts onto the screen. The new information infrastructure is just as significant as Eisenstein (1979) has persuasively argued that the book was for the presentation of scientific data. Information is not being presented here in linear form, with the word as the center of attention-rather the representation becomes the thing, with linear argument as secondary. At the limit, the scientific product becomes itself a library-the human genome for example-which is to be consulted as a huge interactive database created collaboratively by an array of henceforth anonymous authors.

## 6. **The Unusual Role of Theory in Electronic Collaboration**

In computer and information science there is a unique relationship between theory and practice, fact and artifact, word and thing. The word (in the form of a computer program) *is* the thing (the instrument, the tool, the communication medium, the simulator); just as in a fully distributed, computer-mediated work environment the map *is* the territory. So many things that we as social scientists have been trying to prove for years-the inscription of theories in technical objects, the theory-ladenness of observation and so forth-are suddenly literal truths. We need to develop new theoretical sensibilities in order to move around in this brash new world.

There is a strand in science and technology studies that argues that social theory and values play a central role in engineering design (Winner, 1980; Star, 1999). In a classic article on the electric car, Callon (1989) argued that the "socio-engineers" designing the vehicle were at the same time betting on a theory of society in which the car itself became inevitable. Bowker (1994a) looking at *Schlumberger*'s scientific work argued that the geophysical company engineered work practices and social life around the oil well in such a way that their own science became first possible then inevitable. With the collaboratory-conceived here as the deliberate creation of a new information infrastructure for science-there is a multiple interpenetration of social theory and work practice.

At the broadest level there is the network architecture of the national information infrastructure itself. As is well-known, this architecture originated in the United States in the late 1960s in the need felt by the military for linking ARPA (the Advanced Research Projects Agency) projects-at the same time France was developing the CYCLADES network. This early connection with the military meant that much scientific development in the new information infrastructure has been physics-led. Thus physicists developed the electronic preprint service that for many scientists has replaced the reading of journals (currently to be found at http://xxx.lanl.gov); the World Wide Web was first developed at the CERN laboratory.

In a brilliant paper, Abbate (1994) recently discussed the theoretical understanding by the main players of the nature of information networking-and how this played out in a series of relatively irreversible technical choices. There has been over the past few decades a conflict between two major protocols (X.25 and TCP/IP) describing the ways in which computers can talk to each other. The ARPA model (TCP/IP) assumed that the network itself should have low-level structure-this allowed greater heterogeneity among the laboratories being interconnected; although at the same time it entailed a greater degree of computing sophistication on the part of the labs themselves. The alternative protocol (X.25) grew out of a CCITT (Comité Consultatif International Télégraphique et Téléphonique) initiative. Drawing on the model of the telephone, the federated PTT's involved chose instead to make the network itself well-structured-so that at the other end there could be unsophisticated dumb users, just like clients of telephone companies. ARPA offered complexity with the advantage of control (Abbate, 1994, p. 202-203); X.25 usability but little flexibility. ARPA modeled communication as being between well-equipped laboratories with their being a premium on speed and redundancy of communication - traditional military values; CCIT put the premium on volume and usability.

These differences have had very practical consequences for the practice of scientific work and for scientific communication. It was the flexibility offered by ARPANET that allowed the Internet to develop in such a distributed, anarchistic fashion-the efflorescence and rapid propagation of new programs and standards from a huge array of sources, since each laboratory in the military web was assumed to have an interest in maximal cooperation. The success of this model is most clearly demonstrated by the development of NCSA Mosaic in the early 1990s. Mosaic revolutionized the World Wide Web (developed at CERN); its HTML standard created a new kind of usability and transparency. Within this distributed model, the interactive sharing of information has been difficult technically (so many possible standards and protocols need to be taken into account simultaneously in order to create a transparent system). Recognizing this, ARPA concentrated work in the early 1980s on ways of translating email messages between incompatible systems: "The result has been that electronic mail is the only form of connection that has been fully achieved between diverse networks. Thus the communication aspect of networking has been emphasized, encouraging new uses of computers that stress interaction rather than calculation (Abbate, 1994, p. 207). Typical of the complexities of the information infrastructure environment ahs been that this battle between the systems of X.25 and TCP/IP has not resulted in the death of one in favor of the other. Rather, X.25 is generally used, but only that subset of it that supports TCP/IP (Abbate, 1994, p. 206). In computer-mediated work, these forms of level shift are frequent. (For a fuller account of these transformations, see Abbate, 1999)

The ARPA theory of a computer network as a set of connections between "centers of calculation" (Latour, 1987) has, then, led to a degree of local autonomy unusual in the history of large technical systems (e.g., compare the story of electricity networks as told by Thomas Hughes [1983]); and has led to an email-driven version of scientific community. There is no necessary paradox between military centralism and the democratization of Internet development. The key vision is that of the center of calculation being the node; and the question then is only who controls the nodes (and for a long period in post-World War II America, the military effectively controlled most significant high-tech laboratories).

However, it is not only at the level of network architecture that the theory of community with which one is working comes into play; on the contrary, in general one can say that with information infrastructure development of the order of scientific collaboratories and now digital libraries that it is theory all the way down. Another level of theory comes when one is trying to design an infrastructural tool that will actually get used by scientific researchers. To do this, one needs a model of what it is that scientists do, and then a suite of programs that will permit them to do the same better and faster.

Picture as a thought experiment an information systems designer assuming the truth of Merton's norms for the scientific community (cf. McClure et al., 1991, p. 93) for a discussion of what might happen to scientific norms with the introduction of electronic research networks). In the past, these norms may well have been a very useful part of the discourse of scientists: A positive self-image that demarcates science from business and projects an image of the scientist as the fair-minded arbiter of truth. But when it comes to developing an information infrastructure such as a collaboratory with which scientists will work to share and shape their experiments, then whether or not the norms represent working reality will make a difference. Thus, when Star as sociologist on the design team looked at the community of worm biologists, she found that these latter certainly did not want to share all information with rival labs on an ongoing basis, contra Merton's norm of communism: indeed, particularly early in their careers (as ambitious postdoctoral students carving out their own part of the genome), these biologists placed a great premium on secrecy. Scientists would not use a system that permitted no room for privacy: working practice had to be made explicit in order to design a usable product, and that practice had to be informed by sociological theory. Star and Ruhleder (1996) note that the WCS was designed to support tight long distance collaborations among the many researchers involved in the project; co-ordinate this scientific work; and allow the rapid involvement of new scientists by way of on-line recruitment and training.

In order to meet these goals, the infrastructural collaboratory had to reasonably reflect working practice in order to gain initial acceptance-though one assumes, of course, that a series of workarounds developed over time will increasingly allow for affordances between theory of practice inscribed in the infrastructure and practice itself.

### 7. **Hands-On Versus Automated Work**

In the Internet, properties are constantly being swapped between humans and non-humans - such that what constitutes mind, memory, intelligence and scientific work are being redefined by the new infrastructure. It also impacts efficiency. Hiltz and Turoff's classic work *The Network Nation*, first published in 1978, makes the claim that: "Although a crucial endeavor to the maintenance of our society as we know it, research is a highly inefficient process when compared to other institutional functions" (p. 212). This passage reminds us that efficiency has long been a problem and a goal in electronic communication. This goal has been there, indeed, since the origin of computing: Babbages's original 1832 design for his calculating engine was based in part on his admiration for Prony's method of calculating logarithms--itself an application of the factory principle of the division of labor (Bowker, 1994b).

Hiltz and Turoff (1993) described a working computer conferencing system (CCS) that operated within an electronic information exchange system (EIES) as: "Another capability being incorporated into EIES indicates the role a CCS can play in the area of resource sharing. A fairly sophisticated microprocessor with its own computer-controlled telephone dialer has been programmed to engage in the conference system as a full-fledged member, with the same powers of interaction as any human member (Hal Zilog, as it/she/he is referred to). This entity may perform any of the following tasks:

1. It may enter EIES and receive or send messages or retrieve and enter items into the other components of the system.

2. It may exercise certain analysis routines or generate display graphics from data provided by other EIES members, and return the results to them.
3. It may phone other computers and select data from existing data bases or obtain the results of a model to send back to any designated group of EIES users.
4. It may drop off and pick up communication items from other conference and message systems." (pps. 25-26).

This new member of the network was a full equal of any human members. Further, human members of the CCS were themselves transformed by the process in the sense that their disembodied intelligence would be free to act to its fullest capabilities: "computer based teleconferencing acts as a filter, filtering out irrelevant and irrational interpersonal 'noise' and enhances the communication of highly-informed 'pure reason'-a quest of philosophers since ancient times" (p. 28-citing Johansen, Vallee & Collins, 1977).[4] Sociological studies of computer conferencing systems have tended to disconfirm these claims of enhanced filtering (Baym, 1995; Yates & Orlikowski, 1992); and indeed work in organization theory-notably inspired by Janis' classic article on Groupthink-has opened the question of the value of cooperative work (Kraut et al., 1990, p. 152).

However, the dream remains alive, as foreshadowed in the report of the 1993 workshop on collaboratories cited earlier, which aimed at creating "a distributed intelligence, fully and seamlessly networked." A scientific instrument in this view no longer a passive intermediary between the mind of the scientist and nature: "Incorporation of intelligence into the instruments allows the possibility of 'self-directed' data gathering, with the instrument itself deciding when data is significant and should be transmitted, setting parameters based on local feedback, and doing preliminary data reduction. This can lead to both reduction in communications and archiving requirements and better scientific data." (pps. 13-14) A formal model of information flow here can no longer have the structure subject+verb+object = scientist+instrumental action+nature; just as likely is instrument-instrumental action-nature, with the information systems designer and not the scientist as the *deux ex machina*.

So we have on the one hand the pure human mind and on the other the artificial mind of the intelligent information retrieval system: each made possible by the development of a seamless infrastructure. The human and the intelligent agent meet in the infrastructural database-thus the famous Los Alamos e-print archive referred to above (http://xxx.lanl.gov) has a welcome to humans but a warning to nonhumans to be on good behavior: "ROBOTS BEWARE: indiscriminate automated downloads from this site are not permitted." Further exploration recovers the following message, topped by a picture of a no entry sign superimposed on a photograph of Commander Data from Star Trek the Next Generation:

> This www server has been under all-too-frequent attack from "intelligent agents" (a.k.a. "robots") that mindlessly download every link encountered, ultimately trying to access the entire database through the listings links. In most cases, these processes are run by well-intentioned but thoughtless neophytes, ignorant of common sense guidelines.
>
> (Very few of these same robotrunners would ever dream of downloading entire databases via anonymous ftp, but for some reason conceptualize www sites as somehow associated only to small and limited databases. This mentality must

> change --- large databases such as this one [which has over 100,000 distinct URL's that lead to gigabytes of data] are likely to grow ever more commonly exported via www.)
>
> Following a proposed standard for robot exclusion, this site maintains a file /robots.txt that specifies those URL's that are off-limits to robots (note that the title/author listings for all archives remain available for remote indexing). Continued rapid-fire requests from any site after access has been denied will be interpreted as a network attack, and the automated response will be decidedly unfriendly. (Click here to initiate automated "seek-and-destroy" against your site.)

The point here is that as far as the *computer database* is concerned, there is very little difference between humans and nonhumans; just as for the scientific instrument there is very little difference between its own directive intelligence and that of the scientist.

The collaborative infrastructure is being designed precisely in the context of collapsing the human-nonhuman divide. The work of creating a seamless intelligent network is at present posited squarely on the possibility of rendering humans and nonhumans interchangeable. Impelling the collaboratory effort is a central focus on information as product and commodity. Notoriously, in any information-centered vision of the world, it matters little formally where the information is being held or processed (its container): The only thing that is important is that it circulate flawlessly and be analyzed thoroughly. Early work on the nature of information science after World War II instantiate this claim. One example is Wiener's early, outrageous conjecture that people might be sent down a variety of telephone wires as code, and restored physically at the other end: after all, we were just information (Wiener, 1957, p. 23). Another is Turing's famous demonstration that his abstract logic machine (an infinite roll of tape, and a mechanism for moving and marking) was equivalent to any real logic machine-whether this be lodged in the human brain or in a scientific instrument (see Bowker, 1994b for a further exploration of these issues).

Sociological and philosophical theory is getting literally encoded into the infrastructure of collaborative scientific work on the net. It becomes an active resource drawn on by engineers and designers as they create a new system. Thus any given theory-if encoded into successful software-can be truer now than it was before. Winograd and Flores' Coordinator, for example, was one of the early and most influential groupwork programs; it is one of several such programs that utilizes Searle's speech-act theory (Rodden, 1992, p. 5; Winograd & Flores, 1987). The program structures office communication into assertives, directives, commissives, expressives, and declarations; and any future researcher analyzing collaborative work using this system will find these naturalized categories structuring all computer communication and having ramifications for non-computer mediated work. One could say that the infrastructure can only work if the theory that everything is information is true; a preferable statement would be that it can only work if it can make the theory that everything is information true-change the world in such a way that this is a fair description of the nature of things (see Suchman, 1994 for a critique of Winograd and Flores on these grounds, for example).

This is a somewhat pessimistic view of the collaboratory as architect of humans as vessels of pure reason, computers as same, and computer networks as reason distribution engines. There is of course another way of reading the collapse of the human/non-human divide-as represented by the works of Haraway (1985; 1997), Star (1991; 1999) and Latour (1993). According to this reading, categories that were assumed to inhere in the head (such as "memory", "cognition" and "learning" can themselves be understood as spatially and temporally distributed). Cicourel (1990, p. 223), for example, pointed to a unity between social analysis and infrastructural development: "The idea of socially distributed cognition refers to the fact that participants in collaborative work relationships are likely to vary in the knowledge they possess, and must therefore engage each other in dialogues that allow them to pool resources and negotiate their differences to accomplish their tasks. The notion of socially distributed cognition is analogous to the idea of distributed computing"(p. 223). Or again, following historically on the lead of Maurice Halbwachs, an influential group of workers within the world of computer supported cooperative work (CSCW) have argued that memory is invariably a social phenomenon, which is both spatially and temporally distributed (Middleton and Edwards, 1990). Work in distributed artificial intelligence also speaks of "composite systems" of humans and machines.

## 8. **Conclusions**

Rather than indulge in an intellectually vacuous exercise of determining "where is the boundary" between online work and offline work, or the exact distinction between "electronic communities" and "scientific communities," we offer some findings and cautions about the visions of collaboratories. The community metaphor is a powerful one, but one whose heritage is so fraught that it is almost useless to try to retrieve it intact (see Jones, 1998 for excellent discussions of the matter in cyberspace). (Yet, the sheer level of use of the term demonstrates that we are using it for *something* important to us.) Rather, we would like to make the following observations.

Collaboratories contain inherently competing goals (as do all large systems). This is why they are communicative systems. With the growth of big science, for instance with *c.elegans* a designated model for the human genome initiative, an electronic community system is designed with the following goals in mind: to support long-distance collaborations among the many researchers in the large scientific effort; to help coordinate the large scientific effort; to allow the rapid involvement of new scientists by way of online recruitment and training. The worm researchers, however, were explicit in that they did not want to lose that informal, close-knit community feeling. A more structured scientific effort may impose formals ties and weaken the informal ties upon which the community has been built. The availability of online training, and reduced need for mentorship MAY indeed decrease the boundary definition and intimacy of the community, although this is by no means clear as yet. Therefore, the goals of the system are antithetical to the feelings of the existing scientific community. Therefore, the goals of the system are antithetical to the feelings of the existing scientific community. Indeed the goals of the system are in conflict with each other in precisely this way. Simultaneous attempts to build *gemeinschaft* and *gesellschaft*; to blue locale and global reach; to preserve intimacy *and* extend community infinitely will not work without careful attention to the distribution between hands-on and automated work, between *gemeinschaft* and *gesellschaft.*

In scientific communities like *c.elegans*, the scientists valued the close-knit collaborative working environment and were open to an electronic community system that supported this. An electronic system without a formal structure of information and communication links might best support this closeness. However, users of an electronic system get frustrated without some structured access to information, and the desire to use the system for training suggests a system with a more arms' length set of relationships and structures. These two requirements of an electronic scientific community system create an inherent design conflict. The solution to this conflict may come from a more "organized" system, that evolves in response to the community evolution, or perhaps from newly evolving forms and conventions that we cannot yet imagine.

The notion of collaboratory itself is a moving target. However, key elements are an orientation to information flow-between instruments, people, and documents-embedded in an integrated information infrastructure. It is assumed that within this infrastructure the map will become the territory. Just like Huysmans (1981) in *A Rebours* deciding not to visit London because he had already in a quayside restaurant in Calais experienced all the sensations that the visit itself would produce; so do the new scientists not need to see each other, their instruments, or the world in order to do their valuable work of theory production. At the same time, theories of situated action and workplace studies (Hutchins 1995; Lave, 1988; Star, 1995; Suchman, 1987) show us that without attention to the local contingencies and differences in hand-on, craft aspects of even formal work, systems and knowledge risk irrelevance and rigidity.

We see the important of vase, dense electronic networks that scientists use as an opportunity not to engage in boundary disputes, but rather to use the conceptual tools from different parts of the social sciences to understand the phenomenon empirically and theoretically. One thing is clear: Media hype is not helpful in this enterprise, nor is extending the inherited clutter and entropy associated with the concept of community.

However, we hope that perhaps a real opportunity to combine the empirical work from ethnography and situated studies in analyzing new forms of communication, media for work practices, and affiliations, if not communities, may help shed some light on this old problem.

There is nothing new about scientific collaboration being distributed in time and place-the correspondence circles of 17th-century scientists; Darwin's enormous range of correspondents; the large-scale collaborations during World War 2 leading to the development of the atomic bomb. What is new and interesting is the embedding of specific forms of collaboration into an information infrastructure that impacts the very nature of scientific work. We have argued that as the information infrastructure becomes ever more deeply engrained, then the successful theories inscribed into it will be naturalized. They will come to seem true and unproblematic. However, it should be noted that this is not equivalent to saying that the best theory will win. The history of standards and infrastructures makes it very plain that a series of contingencies mean that often the second best (or even the worst) will triumph. The Lotus 123 spreadsheet, the DOS operating system, and VHS format are (in)famous examples. Nor does it mean that the theory which becomes naturalized will give the best possible description of scientific work and practice. After all, Merton's norms were long naturalized in the scientific community; but they certainly did not describe the ways in which scientists acted. But this naturalization does count for something. The new truth becomes the problematic in terms of which theory-critical and otherwise-is defined.

Social theorists have taken on a very active role in the development of collaboratories. Much good sociological analysis has been written by authors involved in the design of systems. Where the field of science studies has in the past called attention to the sophistication of the actors' own perspective and served as spokespeople for the actors, the shoe is now firmly on the other foot. The designers of the new infrastructure are calling attention to the sophistication of sociology's views of information and work, and are serving to represent these views within their programs (Bowker, et al., 1997). A fearful, but quite pleasing, symmetry.

## Acknowledgments

WCS was partially funded by NSF under grants IRI-90-15047, IRI-92-57252, and BIR-93-19844. The Interspace project is currently housed within the Community Systems Laboratory (CSL), headed by Bruce Schatz, affiliated with the Graduate School of Library and Information Science and the National Center for Supercomputing Applications at the University of Illinois, Urbana-Champaign. Additional support was provided by the University of Arizona and the University of Illinois. Co-PIs Bruce Schatz and Sam Ward, and developers Terry Friedman and Ed Grossman were extremely generous with their time, comments, and access to data and meetings; we also thank our anonymous respondents for their time and insight. Additional thanks to Karen Ruhleder and Bill Turner.

## References

Abbate, Janet. 1994. “The Internet Challenge: Conflict and Compromise in Computer Networking,” In J. Summerton, ed. *Changing Large Technical Systems* (pp. 193-210). Boulder: Westview Press.

Abbate, Janet. 1999, *Inventing the Internet*. Cambridge, MA: MIT Press.

Ackoff, Russell L., Thomas A. Cowan, Peter Davis, Martin C. J. Elton, James C. Emergy, Maybeth L. Meditz and Wladimer M. Sachs. 1976. *Designing a National Scientific and Technological Communications System*. Philadelphia, PA: University of Pennsylvania Press.

Baym, Nancy K. 1995. "From Practice to Culture on Usenet," In S. L. Star, ed. *The Cultures of Computing* (pp. 29-52). Oxford, UK: Blackwell.

Baym, Nancy K. 2000. *Tune In, Log On : Soaps, Fandom, and Online Community.* Thousand Oaks, CA: SAGE.

Berg, Marc. 1997. *Rationalizing Medical Work - Decision Support Techniques and Medical Problems.* Cambridge, MA: MIT Press.

Bijker, Wiebe E.. *Of Bicycles, Bakelite, and Bulbs : Toward a Theory of Sociotechnical Change***.** Cambridge, MA: MIT Press.

Bijker, Wiebe E., Thomas P. Hughes, and Trevor Pinch, eds. 1987. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology.* Cambridge, MA: MIT Press.

Bishop, Ann, Laura Neumann, Susan Leigh Star, Cecelia Merkel, Emily Ignacio and Robert Sandusky. In press. “Digital Libraries: Situating Use In Changing Information Infrastructure.” *Journal of the American Society for Information Science, 51.*

Bowker, Geoffrey C. 1993. ‘How to be Universal: Some Cybernetic Strategies,’ *Social Studies of Science, 23,* 107-127.

Bowker, Geoffrey C. 1994a. *Science on the Run: Information Management and Industrial Geophysics at Schlumberger, 1920-1940.* Cambridge, MA: MIT Press.

Bowker, Geoffrey. 1994b. “Information Mythology and Infrastructure.” Pp. 231-247 In L. Bud-Frierman, ed. *Information Acumen: The Understanding and Use of Knowledge in Modern Business* (pp. 231-247). London: Routledge.

Bowker, Geoffrey, and Susan Leigh Star. 1994. "Knowledge and Infrastructure in International Information Management: Problems of Classification and Coding." In L. Bud, ed. *Information Acumen: the Understanding and Use of Knowledge in Modern Business* (pp. 187-213). London: Routledge.

Bowker, Geoffrey C., Les Gasser, Susan Leigh Star and William Turner (eds.). *Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*. Princeton, NJ: L. Erlbaum Associates.

Bowker, Geoffrey, and Star, Susan Leigh. 1999. *Sorting Things Out: Classification and its Consequences*. Cambridge, MA: MIT Press.

Bush, Vannevar. 1990/1945. *Science-The Endless Frontier: A Report To The President on a Program for Postwar Scientific Research*, Washington, DC: National Science Foundation.

Callon, Michel. 1989. “Society in the Making: The Study of Technology as a Tool for Sociological Analysis.” In W. E. Bijker, T. P. Hughes and T. Pinch, eds. *The Social Construction of Technological System* (pp. 83-105). Cambridge, MA: MIT Press.

Callon, Michel, John Law and Arie Rip, eds. 1986. *Mapping the Dynamics of Science and Technology*. London: Macmillan.

Cicourel, Aaron V. 1990. "The Integration of Distributed Knowledge in Collaborative Medical Diagnosis." In J. Galegher, Robert E. Kraut and Carmen Egido, eds. *Intellectual Teamwork: Social And Technological Foundations Of Cooperative Work* (pp. 221-242). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clarke, Adele E. 1991. Social Worlds/Arenas Theory as Organizational Theory. In D. Maines, ed. *Social Organization and Social Process: Essays in Honor of Anselm Strauss* (pp. 119-158). NY: Aldine de Gruyter.

Clarke, Adele E. and Theresa Montini. 1993. "The Many Faces of RU486: Tales of Situated Knowledges and Technological Contestations" *Science, Technology and Human Values*, 18, 42-78.

Clifford, James, and George E. Marcus. 1986. *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley: University of California Press.

Conolly, T., Jessup, L. M., & Valacich, J. S. 1990. "Effects of Anonymity and Evaluative Tone on Idea Generation in Computer-Mediated Groups. *Management Science*, 36, 689-703.

Durkheim, Emile. 1984. *The Division of Labor in Society*. Translated by W.D. Halls. New York: Free Press.

Eisenstein, Elizabeth L. 1979. *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early Modern Europe*. Cambridge: Cambridge University Press.

Galegher, Jolene, Robert E. Kraut and Carmen Egido, eds. 1990. *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Groupe Géode. 1995. *Planifier l'émergence: la formation d'une stratégie dans une économie des compétences et des savoirs*, Paris: CERESI/CNRS, May.

Haraway, Donna. 1985. "A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s." *Socialist Review, 80*, 65-107.

Haraway, Donna. 1997. *Modest-Witness@Second-Millennium. FemaleMan-Meets-OncoMouse: Feminism and Technoscience*. New York: Routledge.

Hiltz, Starr Roxanne, and Murray Turoff. 1993/1978. *The Network Nation: Human Communication Via Computer* (rev.ed.). Cambridge, MA: MIT Press.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.

Huysmans, J. K. 1981. *A rebours*. Paris: Imprimerie Nationale.

Jessup, L.M., and D. A. Tansik. 1991. "Decision Making in an Automated Environment: The effects of Anonymity and Proximity with a Group Decision Support System." *Decision Sciences*, *22*, 226-279.

Jones, Steven G., ed 1998.. *CyberSociety 2.0 : Revisiting Computer-Mediated Communication and Community***.** Thousand Oaks, Calif.: SAGE.

Kiesler, S., J. Siegel, and T. McGuire. 1984. Social Psychological Aspects of Computer-Mediated Communication. *American Psychologist*, *39*, 1123-1134.

Kiesler, Sara, ed. 1998. *Cultures of the Internet***.** Mahwah, N.J : Lawrence Erlbaum Associates.

Kohler, Robert E. 1994. *Lords of the Fly: Drosophila Genetics and the Experimental Life*. Chicago: University of Chicago Press.

Kraut, Robert, Carmen Egido, and Jolene Galegher. 1990. "Patterns of Contact and Communication in Scientific Collaborations." In J. Galegher, Robert E. Kraut, and Carmen Egido eds., *Intellectual Teamwork: Social And Technological Foundations of Cooperative Work* (pp. 149-172). Hillsdale, NJ: Lawrence Erlbaum Associates.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Latour, Bruno. 1993. *We Have Never Been Modern*. Translated by Catherine Porter. Cambridge, MA: Harvard University Press.

Lave, Jean and Etienne Wenger. 1991. *Situated Learning : Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lave, Jean. 1988. *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life.* New York : Cambridge University Press.

Lederberg, Joshua, and Keith Uncapher. 1989. *Towards a National Collaboratory: Report of an Invitational Workshop at the Rockefeller University*, March 17-18.

Lewis, Oscar, 1951. *Life in a Mexican Village: Tepoztlán Restudied***.** Urbana, University of Illinois Press.

Marx, Karl. 1970. *Capital, Vol. 1*. London: Lawrence and Wishart.

McClure, Charles, et al. 1991. *The National Research and Education Network (NREN): Research and Policy Perspectives*. Norwood, NJ: Ablex.

Middleton, D.S. and D. Edwards, eds. *Collective Remembering: Memory in Society.* London: SAGE.

Olson, Gary M., and Daniel E. Atkins. 1990. "Supporting Collaboration with Advanced Multimedia Electronic Mail: the NSF EXPRESS Project" In J. Galegher, Robert E. Kraut, and Carmen Egido eds. *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work* (pp. 429-451). Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, Gary, Ray Levitt, and Laurence Rosenberg. 1993. Introduction, Coordination Theory and Collaboration Technology Workshop, NSF, July 8-10.

Rader, Karen. 1994. "The Production, Distribution and Uses of Genetically-Standardized Mice for Research, 1909-1962. In D. Barkin, ed. *Ideas, Instruments and Innovation. Pasadena, CA: CalTech Workshop Conference.*

Redfield, Robert. 1947. *Tepoztlán: A Mexican Village*. Chicago: University of Chicago Press.

Rice, R. E. and Rogers, E. M. 1984. "New Methods and Data for the Study of New Media." In R. E. Rise & Associates, eds. *The New Media: Communication, Research, and Technology*. Beverly Hills: SAGE.

Rodden, Tom. 1993. “Technological Support for Cooperation." In D. Diaper and Colston Sanger eds., *CSCW in Practice: An Introduction And Case Studies* (pp. 1-22). Berlin: Springer-Verlag.

Ruhleder, Karen. 1995. "Reconstructing Artifacts, Reconstructing Work: From Textual Edition to On-line Databank" *Science, Technology and Human Values, 20,* 39-64.

Schatz, B. 1991. Building an electronic community system. *Journal of Management Information Systems*, 8, 87-107.

Science Policy Study, Background Report No.5. 1986. *The Impact of Information Technology on Science*. Report prepared by the Congressional Research Service, Library of Congress, transmitted to the Task Force on Science Policy; Committee on Science and Technology, U.S. House of Representatives, 99th Congress, Second Session, Serial T, September.

Siegel, J., Dubrovsky, V., Kiesler, S. and McGuire, T.W. 1986. Group Processes In Computer-Mediated Communication. *Organizational Behavior and Human Decision-Processes*, 37, 157-187.

Simmel, Georg. 1950. 1908. 'The Stranger.' In Kurt Wolff (ed.). *The Sociology of George Simmel*. Glencoe, IL: Free Press, 402-408.

Kiesler, Sara and Lee Sproull. 1996. *Connections : New Ways of Working in the Networked Organization***.** Cambridge, Ma : MIT Press.

Stacey, Margaret. 1969. "The Myth of Community Studies," *British Journal of Sociology*, *29*, 134-147.

Star, Susan Leigh. 1989. *Regions of the Mind: Brain Research and the Quest for Scientific Certainty*. Stanford, CA: Stanford University Press.

Star, Susan Leigh. 1991. 'Power, Technologies and the Phenomenology of Standards: On Being Allergic to Onions.' In J, Law, ed. *A Sociology of Monsters? Power, Technology and the Modern World: Sociological Review Monograph 38*. Oxford: Basil Blackwell, 27-57.

Star, Susan Leigh, ed. 1995. *The Cultures of Computing. Sociological Review Monograph Series*. Oxford: Basil Blackwell.

Susan Leigh Star. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist*, *43*, 377-391.

Star, Susan Leigh and Karen Ruhleder. 1996. ‘Steps toward an Ecology of Infrastructure: Design and Access for Large Information Spaces.’ *Information Systems Research, 7,* 111-134.

Star, Susan Leigh and Alaina Kanfer. 1993. "Virtual Gemeinschaft or Electronic Gesellschaft?: Analyzing an Electronic Community System for Scientists." Paper presented at the Annual Meetings of the Society for the Social Study of Science (4S), Purdue University, West Lafayette, Indiana.

Strauss, Anselm. 1978. "A Social World Perspective." *Studies in Symbolic Interaction, 1,* 119-128.

Suchman, Lucy. 1987. *Plans and Situated Action.* Cambridge, UK: Cambridge University Press.

Tombaugh, J.W. 1984. "Evaluation of an International Scientific Computer-Based Conference. *Journal of Social Issues, 40*, 129-144.

Tönnies, Ferdinand. 1957. *Community and Society (Gemeinschaft und Gesellschaft).* Translated and edited by Charles P. Loomis. East Lansing, MI: Michigan State University Press.

Traweek, Sharon. 1988. *Beamtimes and Lifetimes: the World of High Energy Physicists.* Cambridge, MA: Harvard University Press.

Weedman, Judy, 1992. "Origins of Multi-Sector Scientific Collaboration: A Report on Research in Progress," Paper presented at the Annual Conference of the International Network for Social Network Analysis, Dallas, TX, June.

Wiener, Norbert. 1948. *Cybernetics; or, Control and Communication in the Animal and the Machine.* New York: J Wiley.

Young, Michael Dunlop, and Peter Willmott. 1957. *Family and Kinship in East London. London: Routledge and Kegan Paul.*

Winner, Langdon. 1980. Do Artifacts Have Politics? *Daedalus, 19\09,* 121-136.

Winograd, Terry, and Fernando Flores. 1987. *Understanding Computers and Cognition.* Norwood, NJ: Ablex.

Yates, JoAnne and Wanda J. Orlikowski. 1992. "Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media." *Academy of Management Review, 17*, 299-326.

---

1 This section owes a great deal to discussions with Alaina Kanfer, and draws extensively on an earlier paper coauthored by Star and Kanfer, "Virtual Gemeinschaft or Electronic Gesellschaft? : Analyzing an Electronic Community System for Scientists," presented at the Society for the Social study of Science (4S), Purdue University, November, 1994. Her consent to use this analysis is gratefully acknowledged.

2. Compare Rader (1994) and Kohler (1995) for the lives of rats and drosophila in scientific networks. Kohler in particular, discusses the survival strategies the drosophila employ to protect this highly rarefied niche.

3. A description and representation of WCS can be found at http://www.canis.uiuc.edu, along with a description of the Illinois Digital Library Project and the ongoing Interspace Project.

4. It is only fair to note here that Hiltz and Turoff also point out that the authors themselves question the operation of this filtering.