Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment

Geoffrey C. Bowker, Karen Baker, Florence Millerand, and David Ribes

Introduction

It is rare enough in human history that fundamentally new ways of working are developed, which changed the way in which information was gathered, processed, stored, and reused. When we developed speech, writing, and printing, new forms of action at a distance over time became possible (Clanchy, 1993; Goody, 1987). It now seems natural to update this triad by adding the internet. It is clear that each new work mode has accompanied – rather than preceded or followed – social, economic, and organizational upheavals.

In this article we explore the current change accompanying the development of the internet in terms of its relationship with the nature and production of knowledge. We move from a definition of infrastructure to the exploration of a historical context for its development. We continue with the organizational, political, and finally ontological dimensions of its development. Many of our examples will be drawn from scientific cyberinfrastructure; however, where appropriate, links to knowledge work in the arts and humanities and in business will be made.

PART 1: THE WHAT OF INFORMATION INFRASTRUCTURE

What Is Infrastructure?

The term “infrastructure” evokes vast sets of collective equipment necessary to human activities, such as buildings, roads, bridges, rail tracks, channels, ports, and communications networks. Beyond bricks, mortar, pipes or wires, infrastructure also encompasses more abstract entities, such as protocols (human and computer), standards, and memory.

G.C. Bowker (✉)
School of Information Sciences, University of Pittsburg, Pittsburgh, Pennsylvania, USA
e-mail: gbowker@pitt.edu

J. Hunsinger et al. (eds.), *International Handbook of Internet Research*,
DOI 10.1007/978-1-4020-9789-8_5,

Superadded to the term "information," infrastructure refers loosely to digital facilities and services usually associated with the internet: computational services, help desks, and data repositories to name a few. In the same vein but in a broader sweep, the Global Information Infrastructure (GII) refers to worldwide information and communication systems that process and transport data inside and outside national boundaries.

The field of science and technology studies (STS) explores the phenomenon of "infrastructuring" (Hughes, 1983, 1989; Scott, 1998). Infrastructure typically exists in the background, it is invisible, and it is frequently taken for granted (Star & Ruhleder, 1994). The work of infrastructure and its maintenance is itself often that of undervalued or invisible workers (Shapin, 1989; Star, 1991). In such a marginalized state its consequences become difficult to trace and politics are easily buried in technical encodings (Hanseth, Monteiro, & Hatling, 1996; Monteiro & Hanseth, 1997). The design of infrastructure itself can make its effects more or less visible (Bowker, Timmermans, & Star, 1995). Calls to study infrastructure in STS have engendered methods for making it and associated, emergent roles visible (Edwards, 2003; Karasti & Baker, 2004; Ribes & Baker, 2007): practical methods such as observing during moments of breakdown (Star, 1999) or conceptual methods such as "infrastructural inversion" (Bowker, 1994).

Here we take infrastructure as a broad category referring to *pervasive enabling resources in network form*, and we argue that a theoretical understanding of infrastructure is crucial to its design, use, and maintenance. This understanding plays a critical role in associated fields such as informatics, library science, and new media – all fields that underpin communication in large-scale and long-term collaborative science. In our analysis we extend conventional understandings of infrastructure as "tubes and wires" to the technologies and organizations which enable knowledge work: supercolliders, orbiting telescopes, supercomputer centers, polar research stations, national laboratories, and other research instruments of "big" science. In addition our image would be incomplete without the variety of scientific organizations, such as funding agencies, professional societies, libraries and databases, scientific publishing houses, review systems, and so on, that are inherent to the functioning of science. As Leigh Star has noted, infrastructure is relational: the daily work of one person is the infrastructure of another (Star & Ruhleder, 1996). Finally, we further open the conceptual umbrella of infrastructure to include the individuals – designers and developers, users and mediators, managers and administrators – in existing and emergent roles associated with information infrastructure.

Large-scale information infrastructures (or in today's language of large-scale scientific projects, cyberinfrastructure and e-Science) aim at supporting research practices through a vast array of community digital services and resources (collaboratories and centers, data and code repositories, best practices and standards development, visualization tools and high performance computing, and so on). Two main issues are associated with such projects: first, the idea of a shared infrastructure in the sense of a public good; second, the idea of sustainability, of supporting research over the long term. What we understand by the concept of infrastructure has significant consequences in terms of how we design the support environments for scientific research. For instance, concerns about infrastructure in everyday scientific practices tend to follow a reactive pattern such as when scientific imperatives require updated infrastructure components. When a new instrument comes available, there is frequently a lot more data or a new type of data to process and manage than existing resources can afford (basically "more resources are needed for the work to get done"). Traditional vision tends to favor immediate responses, usually in terms of additional human resources or new technological tools.

An alternative vision of infrastructure may better take into account the social and organizational dimensions of infrastructure. This vision requires adopting a long term rather than immediate timeframe and thinking about infrastructure not only in terms of human versus technological components but in terms of a set of interrelated social, organizational, and technical components or systems (whether the data will be shared, systems interoperable, standards proprietary, or maintenance and redesign factored in).

Thinking of infrastructure this way requires a major shift in thinking. It involves changing common views and metaphors on infrastructure: from transparency to visibility, from substrate to substance, from short term to long term (Karasti, Baker, & Halkola, 2006; Star & Ruhleder, 1996). Usually perceived as something "just there," ready-at-hand, completely transparent, something upon which something else "runs" or "operates" (a system of railroad tracks upon which rail cars run; a computer network upon which a research lab operates or disseminates data like the WWW), any infrastructure that has been the target topic of activities has probably also been the object of passionate debates – for the engineers in charge of building the railroad system or for the scientists and technologists in charge of developing the network. Related to this taken-for-granted aspect of infrastructure, STS speaks of invisible work, complex problems, and the challenges of alignment in the face of breakdowns. Understanding the nature of infrastructural work involves unfolding the political, ethical, and social choices that have been made throughout its development. Analytically, this exercise consists in "going backstage" (Goffman, 1956; Star, 1999), looking for the infrastructure in the making and practicing "infrastructural inversion" (Bowker, 1994) that is shifting the emphasis from changes in infrastructural components to changes in infrastructural relations. Infrastructure is indeed a fundamentally relational concept; it emerges for people in practice, connected to activities and structures. It consists of both static and dynamic elements, each equally important to ensure a functional system.

Defining Information Infrastructure

Most often, information infrastructure is defined by jotting down a laundry list of characteristics. Instead of discussing the technologies, or even the particular affordances, of infrastructure, we focus on capturing the underlying concept and placing this within a historical lineage. We ask, beyond the immediacy of introducing high-end technologies to the sciences, what is information infrastructure really about? Early National Science Foundation reports focused on digital libraries and global environmental system science (Futrell & the AC-ERE, 2003; Hedstrom et al., 2002). The Atkins NSF report (Atkins et al., 2003) focuses on cyberinfrastructure, defining cyberinfrastructure as those layers that sit between base technology (a computer science concern) and discipline-specific science. The central concentration is on value-added systems and services that can be widely shared across scientific domains, both supporting and enabling large increases in multi-disciplinary science while reducing duplication of effort and resources. According to the Atkins Report, cyberinfrastructure consists of "...hardware, software, personnel, services and organizations" (Atkins et al., 2003, p.13). This list recognizes from the outset that information infrastructure is about more than just wires and machines. A more recent cyberinfrastructure vision document is similarly diffuse, though it regrettably somewhat sidelines the social and organizational in the definition[1]:

> Cyberinfrastructure integrates hardware for computing, data and networks, digitally enabled sensors, observatories and experimental facilities, and an interoperable suite of software and middleware services and tools. Investments in interdisciplinary teams and cyberinfrastructure professionals with expertise in algorithm development, system operations, and applications development are also essential to exploit the full power of cyberinfrastructure to create, disseminate, and preserve scientific data, information, and knowledge. (*NSF CI Council, 2006*, p. 6)

Both these definitions do, however, draw attention to the complex if not dynamic nature of cyberinfrastructure development. A recent report to NSF (Edwards, Jackson, Bowker, & Knobel, 2007) presents the history and theory of infrastructure as one approach to understanding the dynamics, tension, and the design of infrastructure.

Cyberinfrastructure projects come in many forms but they often seek to bring together, under a single umbrella, various domain sciences with novel information technologies. For example, GEON "the geosciences network" seeks to provide computing resources such as visualization and data integration to the broader earth sciences. Here, a single information infrastructure is to become the clearinghouse for data in paleobotany, geophysics, and metamorphic petrology (http://geongrid.org; Ribes, 2006). Similarly, LEAD "Linked Environments for Atmospheric Discovery" seeks to enable real-time analysis of weather data for meteorologists and atmospheric scientists (http://leadproject.org). Both projects are ambitious and have encountered problems in working across disciplinary boundaries (Lawrence, 2006; Ribes & Bowker, 2008). Challenges exist across temporal boundaries as well. With concern for long-term data and information rather than information infrastructure per se, the LTER since 1980 has focused on creating long-term scientific understandings of biomes (http://lternet.edu). In building information infrastructures,

[1] This reminds us that while social and organizational theory have made inroads into the technically dominated fields of information infrastructure design, maintaining such accomplishments will require continuous and active engagement by practicing social scientists in the field of cyberinfrastructure.

always a long-term venture, programs and communities are faced with developing technologies that look farther ahead than immediate research concerns.

While accepting this broad characterization, the long-term perspective such as that of the Atkin's report and of long-term scientific endeavors incorporated today as part of a temporal continuum (i.e., the "long-now," Brand, 1999; the "invisible present," Magnuson, 1990) invites a discussion of first principles. For this we return to Star & Ruhleder's now classic definition of infrastructure (Star & Ruhleder, 1996) originally composed for a paper on one of the early scientific collaboratories, the Worm Community System. We show in Fig. 1 how their definitions can be ordered along two axes, one explicitly non-spatial and the other spatial: the social/technical and the individual/group or local/global. As Star and Ruhleder put it, the "configuration of these dimensions forms "an infrastructure", which is without absolute boundary or *a priori* definition" (p. 113). We think of infrastructures and their construction as distributions along these axes rather than as tensions between polar opposites. Our argument is an ecological one; it calls for investigating infrastructure building a set of distributed activities – technical, social and institutional.

Everyday decisions in terms of infrastructure design, development, or enactment involve such distributions. What needs a technical fix? What needs a social

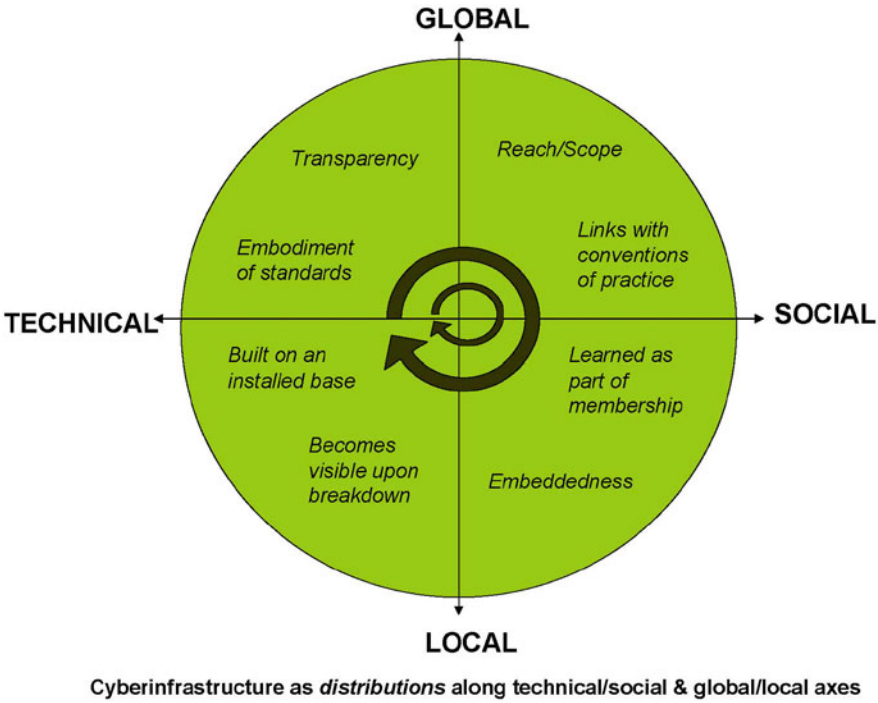

Fig. 1 Information infrastructure as distributions along technical/social and global/local axes

fix? Critical technical, social, and organizational path dependencies established in the present will have long-term consequences – with each new lock-in an aura of inevitability about technical and organizational choices builds (Arthur, 1994). Looking at the distribution of qualities between the social and the technical allows us to investigate the discontinuity within the (apparent) continuity of technological development that usually tends to mask decision points, thus allowing for writing stories in new forms (MacKenzie, 1993). We cannot do the history of software without doing the history of their surrounding organizations.

In building cyberinfrastructure, the key question is not whether a problem is a "social" problem or a "technical" one. That is putting it the wrong way around. The question is whether we choose, for any given problem, a primarily social or a technical solution, or some combination. It is the *distribution* of solutions that is of concern as the object of study and as a series of elements that support infrastructure in different ways at different moments.

An everyday example comes from the problem of e-mail security. How do I distribute my trust? I can delegate it to my machine, and deploy firewalls, password protection, and version controls. Or I can work socially and organizationally to make certain that members of my scientific community understand and police norms about data usage.

Similarly within cyberinfrastructures, metadata generation can be automatic (frequently problematic), the role of an *in situ* data manager or the role of a receiving archive. Generally it sits unhappily between the three. In general, perspectives, standards and conventions get distributed.

An example in scientific efforts comes from the challenge of data issues in the sciences, i.e., data access and sharing across various disciplines, institutions, technical platforms, and across long periods of time. What interoperability strategy best suits the design of shared databases with highly distributed and heterogeneous data sets? How to distribute work and responsibilities between databases, users, and institutions? One can think of an ontology that would allow requests in natural language from a variety of interdisciplinary scientist users – in this case the work of drawing equivalences between discipline terminologies is delegated to the machine (who's in charge of crafting the ontology remains at stake). One can also think of a standardized metadata language that would leave the competence to the scientists – who would have to appropriate first the syntax of the metadata language to be able to then discover the data (rewarding mechanisms for that remain to be defined).

Let us suppose the metadata standard approach was selected. Now how to distribute between the local and the global, or between individual versus community needs? "An infrastructure occurs when the tension between local and global is resolved" (Star & Ruhleder, 1996, p. 114). While occurring, how are an infrastructure's qualities being distributed between the local and the global? For instance, to what extent is a metadata standard designed generic enough to represent a domain ("reach or scope") while aiming at fitting local structures, social arrangements, and technologies ("embeddedness")?

For our purposes, cyberinfrastructure is the set of organizational practices, technical infrastructure, and social norms that collectively provide for the smooth operation of scientific work at a distance (Latour, 1992; NSF CI Council, 2006).

Being all three subjects to contingent distributive actions, these are objects of design and engineering; a cyberinfrastructure will fail if any one is ignored. Key to any new infrastructure is its ability to permit the distribution of action over space and time.

The Long Now of Information Infrastructure

Stewart Brand's "clock of the long now" will chime once every millennium: a cuckoo will pop out (Brand, 1999). Accustomed as we are to the "information revolution," the accelerating pace of the 24/7 lifestyle, and the multi-connectivity provided by the world wide web, we rarely step back and ask what changes have been occurring at a slower pace, in the background. For the development of cyberinfrastructure, the long now is about 200 years. This is when two sets of changes began to occur in the organization of knowledge and the academies that have accompanied – slowly – the rise of an information infrastructure to support them. On the one hand is the exponential increase in information gathering activities by the state (statistics) and on the other is the emergence of knowledge workers (the encyclopedists) and the accompanying development of technologies and organizational practices to sort, sift, and store information.

When dealing with information infrastructures, we need to look to the whole array of organizational forms, practices, and institutions that accompany, make possible, and inflect the development of new technology, their related practices, and their distributions. JoAnne Yates made this point beautifully in describing the first commercial use of punch card data tabulators in the insurance industry. She demonstrates that use became possible because of *organizational* changes within the industry. Without new forms of information management, heralded by such low status technologies as the manila folder or carbon paper and accompanied by new organizational forms, there would have been no niche for punch card readers to occupy (Yates, 1989). Similarly, Manuel Castells argued that the roots of contemporary "network society" are new organizational forms created in support of large corporate organizations, which long predate the arrival of computerization (Castells, 1996). James Beniger described the entire period from the first Industrial Revolution to the present as an ongoing "control revolution" in which societies responded to mass production, distribution, and consumption with both technological and organizational changes, designed to manage ever-increasing flows of goods, services, and information (Beniger, 1986). In general there is more continuity than cleavage in the relationship of contemporary "information society" to the past (Chandler & Cortada, 2003).

The lesson of all these studies is that organizations are (in part) information processors (Stinchcombe, 1990). People, routines, forms, and classification systems are as integral to information handling as computers, ethernet cables, and web protocols. The boundary between technological and organizational means of information processing is both diffuse and mobile. It can be shifted in either direction and technological mechanisms can only substitute for human and organizational ones when the latter are prepared to support such a substitution.

The impacts of contemporary infrastructure are pervasive. In the “long now,” two key facets of scientific information infrastructures stand out. One clusters around the nature of work in the social and natural sciences. Scientific disciplines were formed in the early 1800s, a time Michel Serres describes as the era of x-ology, where “x” was “geo,” “socio,” “bio,” and so forth (Serres, 1990). Auguste Comte classified the division of labor in the sciences, placing mathematics and physics as the most developed and best models and sociology as the most complex and least developed – more or less where Norbert Wiener placed them 130 years later in *Cybernetics and Society* (Weiner, 1951). This was also the period during which the object we now call the database came to be the lynchpin of the natural and social sciences. Statistics etymologically refers to “state-istics,” or the quantitative study of societies (states); it arose along with censuses, medical records, climatology, and other increasingly powerful techniques for monitoring population composition and health (Porter, 1986). Equally, the natural sciences – moved by the spirit of the encyclopedists – began creating vast repositories of data. Such repositories were housed in individual laboratories or in institutions, such as botanical gardens and museums of natural history. Today they are increasingly held in electronic form, and this is fast becoming the norm rather than the exception. Indeed, a researcher publishing a protein sequence must also publish his or her data in the (now worldwide) Protein Data Bank (http://www.wwpdb.org). A somewhat analogous publishing effort by the Ecological Society of America was initiated in 1998 for digital supplements, including databases and source code for simulation models (http://www.esapubs.org/archive/). The use of this forum has developed differently as one might expect, given the significant cultural and data differences between the fields of ecology and bioinformatics.

The second facet clusters around scientists’ communication patterns. In the 17th and 18th centuries scientists were largely “men of letters” created an information infrastructure to exchange both public and private correspondence, such as the famous Leibniz/Clarke exchange (Desmond & Moore, 1991; Leibniz, Clarke, & Alexander, 1998). From the early nineteenth century a complex structure of national and international conferences and publishing practices developed, including in particular the peer-reviewed scientific journal – at the same time as personal communication networks continued to develop (Desmond & Moore, 1991). Communication among an ever-broader scientific community was no longer two-way, but *n*-way. New forms of transportation further undergirded the development of a truly international scientific community aided also by *linguae francae*, principally English and French.

Similarly, today’s scientific infrastructures must be understood as an outgrowth of these developments rather than ahistorically as “revolutionary” or “radical.”[2]

[2]The claims of revolutionary change about our information infrastructure are somewhat akin to the millennialism of every generation since the industrial revolution. We are living the epoch of the database founded in the era of governmentality (late eighteenth century) – and all the claims that we see today about speed, time, and distribution have been with us since that epoch.

Databases and *n*-way communication among scientists have developed as organizationally and institutionally embedded practices and norms. There is far more continuity than many recognize precisely due to the 'always invented here' claims of much computer and information science. However, as scientific infrastructure goes digital, we can also identify a genuine discontinuity. The social and natural sciences evolved together with communication and data-processing technology. Changes in data handling and communications have profound ripple effects throughout the complex web of relations that constitutes natural and social scientific activity; infrastructures thus grow slowly.

PART 2: DESIGNING COMMUNITIES, TECHNOLOGIES AND KNOWLEDGE

In the context of the internet, Infrastructure Studies spans the set of information and communication technology and services that secure and underlie much of modern work and play. It explores the ways in which

- *new forms of sociality* are being enabled/shaped by and shaping Information and Communication Technologies (ICT). Objects of study here will typically be communities of scientists (e.g., cyberinfrastructure, e-science, collaboratories); new kinds of intentional community (e.g., support groups for sufferers of a particular rare disease), and studies of distributed collective practice (e.g., scientific networks, communities of practice, transnational businesses). This is the social dimension of the new infrastructure.
- *social, ethical and political values* are being built into these self-effacing, self-perpetuating infrastructures.
- the nature of *knowledge work* is changing with the introduction of new information technologies, modes of representation, and the accompanying shifts in work practice and systems for the accreditation of knowledge. This is the ontological dimension of the new infrastructure.

We shall examine some core concepts from each of these themes in turn, referring in each to aspects of the definition of infrastructure.

New Forms of Sociality: Organizational and Community Issues

Mediation

In contrast to large-scale physical infrastructures that may be viewed as the responsibility of a business, city, or government, the internet transitioned from a research environment into an open-bazaar, committee-run entity and today is being re-defined in a marketplace of policies and regulations. The internet and its associated information infrastructure is reaching beyond the immediacy of the physical and technical, extending to individual and community realms as well as into organizational and cultural structures and expectations. The scope and ramifications of "information" – from access, quality, and distribution to exchange – is vast, bringing with it responsibilities and exposing needs. Mediation appears an appropriately ambiguous term to suggest the multiple new dimensions to be developed and interfaces to be tended. With its whole-culture, whole-earth reach, there is aligning, weaving, and arbitrating to be carried out between systems and networks, people and organizations, habits and cultures. Infrastructure studies is best considered as "process building," involving both community building and system building. An infrastructure studies "process-building" framework must take account of both the use of technology for automation and effective disintermediation as well as the increased need for support, for tending by growing numbers and varieties of mediation professionals. New types of mediation are required for managing information, new forms such as that represented by activity theory or by ethnographic work (Engestrom, 1990; Nardi & O'Day, 1996).

Process Building

Because of the broad implications for information infrastructures, the sweep of information systems compares to that of governance systems. The pooling of information within infrastructures provides a distribution of relations such as group identity, makes available community findings, and provides a mechanism for staying in touch by sharing information and files – the coin of the realm. The need for participation holds the same import for information infrastructures as for democratic systems. Engaged participation of diverse participants ensures that issues such as standards formation and continued maintenance and update are addressed (Hanseth et al., 1996). If participants have been active in the formation of infrastructure elements, they are more likely to have a deeper awareness of alternatives and have had a voice in mediating choices inherent to issues such as standards formation and community goals (Millerand & Bowker, 2009). This then calls for a forum where multiple perspectives are considered and where the timeless tensions between local and general, between short term and long term are addressed. Once in place, information procedures and standards become a general requirement or driver that facilitate development for some, that misrepresent or perhaps ignore some, and that potentially alienate others (Berg & Timmermans, 2000). One aspect of infrastructure studies inquiry is consideration of new types of roles evolving with the process of building information infrastructure – roles such as digital librarians, information managers, and network specialists. These represent new strategies – and new attitudes – that are organizationally situated to support an internet generation of participants.

System Sustaining

Just as power and energy today underpin the functioning of automated services, so the information infrastructure has evolved to be regarded as an essential, ubiquitous service for delivery, access, and exchange of information. With technology-related resources, issues of change, redesign, and support arise with respect to hardware, software, networks, and human resources and relations as we gain experience with technology, its mediation, and its interfaces. Here, mediation encompasses relations between people, between machines, or between communities as well as a vast variety of human–technology interfaces. Through new roles explicitly incorporating mediation, there are opportunities for sensitivity to silences and absences that may be organizationally instantiated – a layer of professionals trained to notice who is not represented or who is misrepresented. Articulation work is well recognized in STS (Strauss, 1988, 1993), in the literature of design (Bratteteig, 2003; Schmidt & Bannon, 1992) and in ethnographic studies (Baker & Millerand, 2007; Sawyer & Tapia, 2006) though not always as an integral part of the information infrastructure process building– local, community, national, or international. We note that infrastructure work frequently entails frequent and ongoing articulation work in order to enable continued functionality.

Infrastructure in Time

Infrastructure building involves alignment work involving different time scales. The Long-Term Ecological Research program provides an interesting example with its mission to further understanding of environmental change through interdisciplinary collaboration and long-term research projects (LTER: http://lternet.edu; Hobbie, Carpenter, Grimm, Gosz, & Seastedt, 2003). In addition to the LTER moving beyond the "plot" of traditional ecoscience to analyze change at the scale of a continent, one of the chief challenges is to move beyond a 1–6 year funding cycle of projects or a 30-year career cycle of the scientist to create baselines of data spanning multiple decades. While the preservation of data over time, and their storage in conditions appropriate to their present and future use, has always been a priority within the LTER network, a new urgency arises with the development of scientific cyberinfrastructure projects aiming to support long-term and large-scale ecological research. Aligning what is naturally misaligned (funding cycles, scientists career trajectories, ecosystem cycles) is fundamentally an issue of distribution between technologies, communities, organizations, institutions, and participating individuals.

Lemke presents the important principle of heterochrony, variations in the parameters of temporal change associated with different parts of a system (Lemke, 2000). Understanding cyberinfrastructure building requires understanding the timescales for operation of its different components and how they are articulated so that processes have an apparent continuity across time. Ecological systems consist of different components with various change rates and scales of size. Indeed these differences are why they can absorb shocks and survive:

> Some parts respond quickly to the shock, allowing slower parts to ignore the shock and maintain their steady duties of system continuity. The combination of fast and slow components makes the system resilient, along with the way the differently paced parts affected each other. [. . .] All durable dynamic systems have this sort of structure; it is what makes them adaptable and robust. (Brand, 1999, p. 34)

Heterochrony in cyberinfrastructure development is a major issue. It is a statement of the obvious to say that given the extremely rapid developments in information technology, parts of a technological system are frequently outdated before the whole system can be assembled, thus requiring development of ad hoc, last minute arrangements. It is also common to find that when a new technology is released and becomes of high interest for user communities, it ceases to be of interest to its developers who are already working on another state-of-the-art technological project – thus leaving critical issues such as maintenance and redevelopment over time largely unaddressed. Other examples are drawn-out collaboration processes when disparate groups try to collaborate but work within differing time frames, resulting in different rates of interest, uptake, and/or learning.

A system – whatever the blend of technical, organizational, social – consists of multiple layers and dimensions at differing stages of maturity. In addition, the designers, developers, deployers, enactors, and users of data and information are at different phases of technological interest, awareness, and/or skills (Millerand and Bowker, forthcoming). Thus an individual's interface with others about technology or with the system itself at any one time presents a scenario of differing local arrangements, conceptual development, and individual understandings. Even with here-and-now, small-scale expectations and interactions, participants confront what can be called differing "readiness factors" (Olson & Olson, 2000). The introduction of technology is discussed in the champion literature with leadership depending upon different types of goals, champions (e.g., project, technical, and user champions), and innovators (e.g., gatekeepers focused on products in particular and conceptualists focused on approaches) (Howell & Higgins, 1990; Schon, 1963). In planning for the internet, there is a shifting baseline of understanding as the user/designer base of experience grows. In terms of the growing user/design literature, two points were developed early on: (1) for planning and learning to be effective, it needs to be based on the learner's experience (Dewey, 1902); (2) individuals have perceptual as well as information processing limits (Weick, 1979). As an example, consider the pervasiveness of electronic spreadsheets as an organizing tool. Experience with numbers stored in columns or matrix form was a territory and an approach over which accountants reigned for decades. Even after initial development of the electronic spreadsheet concept, the tool and the public matured together over time. In a sense, for those who are part of the digital culture, availability and familiarity with spreadsheet use has likely raised their readiness for work with say Data Grids.

Social and Political Values

What kind of a thing is the internet? We do not do it justice when we see it as wires and modems, bells and whistles. Conceptually, let us for a moment imagine it as a very large database, an outcome of the late eighteenth-century encyclopedist impulse to record all of the world's knowledge and make it freely accessible. It was rendered possible by not only the development of electricity but also by the development across the nineteenth century of large-scale classification systems in any of a number of domains, of library classifications above all (the MARC record turns electronic and then morphs into the Dublin Core). Today, one can find both widely held standards such as the ISO standards (International Organization for Standardization, http://www.iso.org) as well as an emergent set of working standards in community arenas. Now it is clear that how we arrange information in encyclopedias has social and political dimensions (do we look under "t" for "terrorist," "f" for freedom fighter, or "i" for "insurgent," for example?) cite. The art and science of classification plays an underappreciated but recurring role in the organization of information (Bowker & Star, 1999; Epstein, 2007).

It is a database that constantly needs updating. Indeed, looking back to the 1911 edition of the Encyclopedia Britannica – considered by many the finest edition – one finds a number of categories which did not stand the test of time and do not hold today – aether, phrenology, and poltergeist, for example (http://www.1911encyclopedia.org/). So how do you design a database, for the ages, which reflects the values that we as a society want to develop – be these the very abstract value of pluralism or specific ones, such as carbon neutrality or producing green technology. Looking forward from 2000, there are new topics to include – from blogging to dataspaces and infrastructure – as well as new data types to incorporate – from metadata to pervasive sensors and streamed data. Indeed, to date, the internet is inclusive. It is a "dataspace" into which we can add a growing variety of artifacts (Franklin, Halevy, & Maier, 2005).

These design issues are playing out across the board in information infrastructure development as the problem of deciding what kind of objects to people your world with, how to describe data with metadata and how to build ontologies (Sowa, 2000). Building ontologies involves gathering domain knowledge, formalizing this knowledge into a machine computable format, and encoding it into machine language. This is the stripped down technical understanding of knowledge acquisition. But in the work of building ontologies as undertaken today, ontology-building specialists typically find that domain practitioners are not readily prepared for ontology building. Ontology work is a quintessential act of distribution – taking knowledge out of a closed community of practice and allowing for its reuse and reshaping by others in different fields. First there is the enrollment of the domain. To bring experts on board is to inform them of the technology of ontology, its strengths in the face of other interoperability strategies, and the particular work it will require. Enrolling practitioners is securing an investment in technological direction by a domain community. Second is the work of knowledge acquisition. Written sources such as textbooks and technical treatises are often not precise enough for transformation into description logics: there may be competing accounts of the same phenomena, overlapping taxonomies and standards, or outright contradictions (Bowker, 2000). Similarly in consulting authorities in a domain, a programmer may find that these experts are not immediately able to state domain knowledge in the terms necessary for ontology building. In short, the ontology specialist often finds that what participants in a domain consider their validated and structured knowledge is not readily compatible with ontology building. So third is understanding that ontologies are not static. Though frequently viewed as a product to finish, it is dynamic ontologies with associated process-building activities designed, developed, and deployed locally that will allow ontologies to grow and to change. And finally, the technical activity of ontology building is always coupled with the background work of identifying and informing a broader community of future ontology users.

Social and political choices are evidently being made in the building of large interoperable databases and ontologies in the social sciences and the humanities – one needs only look at controversies around ethnic classification in the US census, debates around Wikipedia content and editing, or the classification system of the Library of Congress. However, they are also present across the range of natural sciences. As drug companies create databases of plants that may have medicinal value, they frequently relegate indigenous understanding of those plants to free text fields that are essentially useless – the political decision not to use local knowledge is enforced by a simple technical fix (Hayden, 2003; Latour, 1993). If a dominant subgroup within a given discipline gets to define the ontology of the database, they can reinforce their power – a simple example here is the neurophysiological approach to mental illness at the heart of DSM, the Diagnostic and Statistical Manual (Kirk & Kutchins, 1992): this classification system, used in both research and medical reimbursement, renders it extremely difficult for Freudian therapists to present their diagnoses.

Knowledge Work

Let us now anathematize techno hubris. Much technological infrastructure development has been motivated by the belief that if we build it they will come. Little was learned from the experience of supercomputer centers (built, they didn't come) or collaboratories (built, they left early). It is not enough to put out a new technical infrastructure – it needs to be woven into the daily practices of knowledge workers. It has emerged from the last ten years of information infrastructure development that a wide range of cultural and organizational changes need to be made if the new infrastructure is going to bear fruit.

We are dealing with a massively entrenched set of institutions, built around the last information age and fighting for its life. The rationale for publishing journals shifts (just as music producers are finding less purchase for the album) when single papers can be issued to a mass audience at the push of a button. This then leads to questioning the reasons to work with the publishing industry, since the work of peer review is done on a volunteer basis by the scientific community, with the journals contributing only to the expense and hence unavailability of the final product. However, one does not just click one's heels together and make a multi-billion dollar enterprise go away – as Chandler has shown there is remarkable historical continuity among major corporations (Chandler & Hikino, 1994). And yet the very nature of publishing is changing and will potentially change more. Similarly, universities grew up in previous eras of communication technology – one traveled to campuses in order to study with particular learned professors and to become immersed in a learning environment. There seems little reason to restrict universities to campuses (or libraries to buildings) when you have the ability to share information easily across the internet.

And yet there is inevitably a cultural dimension as well. People love libraries and campuses; they don't tend to feel so fond of their computer screens. Take the issue of publishing for example. Over the past 30 years, the Missouri Botanical Gardens has twice failed to move the Flora of North America into the electronic realm. The problem has not been technical capability. To view it in broadest perspective, it is that botany grew up with the printed book. Some of first incunabalae were botanical field guides. Works in botany stand the test of time – it is one of the few scientific fields where work from the eighteenth and nineteenth centuries is still regarded as valid and interesting. It has proven very difficult to get botanists trained in this tradition to move to the sharing of uncertain results in ephemeral form: their whole scientific tradition is devoted to provenance and finality.

Within the field of science studies, ethnographic work has described scientific knowledge production as a specialized form of work. By focusing on the hands-on processes of how knowledge, or at an even finer granularity, how "a fact" is constructed these ethnographic "laboratory studies" have contributed greatly to a body of empirical evidence complicating notions of a scientific method; knowledge and fact production; scientific rationalities; data sharing and interpretation (Knorr-Cetina, 1981; Latour & Woolgar, 1979; Lynch, 1985). In conjunction with a "turn to practice," lab studies were also crucial in a "turn to the material" by shifting analytic focus to instrumentation, data (including "second- and third-order" representations such as derived data sets, visualizations, photographs, charts, and graphs), and the movements of these "traces" across the physical site of the laboratory itself. By making a combination of the "practical and material" accessible to the sociological analyst another avenue was opened for the study of "scientific content." The methodological imperative, then, becomes to treat "technical" dimensions of activity no differently than any other sociological object of interest: with appropriate access and determination of the researcher, human activity is fundamentally observable whether "scientific," "lay," or "technical."

Thus rather than the rhetoric of revolutionary fervor that permeates cyberinfrastructure circles, infrastructure studies take as its object change at a much more mundane scale: as forms of practice, routine, or distributed cognition associated with knowledge work. Is this position against a possibility of "revolution?" Not at all. It is, rather, a research sensibility which seeks to make transformations of infrastructure visible relative to the everyday work of scientists, information technologists, or information managers.

As we have argued, the new information infrastructure is fundamentally about distribution. Consider the long history of European development over the past 1500 years. When trade was difficult and scarce, community centers were built up close to resource reserves: coal, water, arable land. With the first industrial revolution of the eleventh century, the use of water power (windmills) and the plough created conditions where cities could grow away from this cluster of resources. With the second industrial revolution (the eighteenth and nineteenth centuries), global resource distribution became possible through new communication and transportation infrastructures. Now, over the past 200 years, our emergent information society has (gradually) moved to a further form of distribution where complex social, organizational, and knowledge work can be practiced on a global scale, which leads to the question of how we study it.

PART 3: TOWARDS A SCIENCE OF INFORMATION INFRASTRUCTURE

The internet is changing what we know and how we know it; accordingly the study of information infrastructure studies is a field with an emerging research agenda. We need new images and mindsets. The classic argument here is from Paul David (David, 1985): computers were very bad typewriters; electrical dynamos were very bad steam generators – leading to the classic productivity paradox where their respective introductions led to 30 year drops in productivity. Only when we truly think the new technology can we use it to its fullest. The internet opens an array of information infrastructure issues whose resolution will frame the future of the digital realm, issues of data sharing and database query, community standards and data spaces, domain repositories and ontologies, grid portals and resource sharing: “Infrastructure is an idea, a vision or an ideal, but it is also a practice, a commitment and a long term endeavor” (Ribes, 2006, p. 299).

The fields which have contributed to this new science include, but are not limited to, computer science, information science, communication, organization theory, cognitive science, and science and technology studies. Teams of researchers have emerged from their disciplinary silos to claim special expertise on information infrastructure – often in the process denying the expertise of others, as is normal in attempts to professionalize (Abbott, 1988). We have argued in this article that we cannot simply fragment the parts of such studies into separate fields. As with transdisciplinary science in general (Gibbons et al., 1994), there is a fundamental scope and unity to the field that reflects the very nature of the new infrastructure – one in which the global and the local, the social and the technical are in flux in new and interesting ways for the first time in 500 years. The decisions we make over the next 50 years will have very long-term consequences.

We have looked so far at the nature of information infrastructure and some emergent themes in its study. We conclude with a few themes native to this emerging field that must be addressed in order to move the field to the next level.

The first theme is research organization. Following an infrastructure is not an easy task for the qualitative social scientist, who typically has been immersed in a particular tribe, project, group and has been able to interview key players and observe them at work. As some have complained in Hollywood, there is little to catch the eye in watching someone typing on a keyboard – and those who are spending a lot of time online are immersed in a community to which the researcher does not have physical access. There are two responses here. First, we need new models for scalable qualitative research: is the appropriate sized research team more of the order of 15–20 rather than the 0–2 typical of many programs? One way to look at this is in terms of Ross Ashby's Law of Requisite Variety (Ashby, 1956): we need as much span in our research teams as there is in the phenomena we are studying.

Further, we need better forms of multi-modal research – there is no one methodology which is going to prevail – the phenomena themselves evoke a range each with strengths and weaknesses. For example, from social network analysis we have a semi-quantitative approach that brings together a great deal of data but which typically ignores domain content, meanings, and practice. We also have the hermeneutic tradition of readings of scientific texts and more recently databases; this is a method that provides rich situated detail but it typically ignores organization. Infrastructure is "large" spanning time and space, but it is also "small" coming in contact with routine and everyday practice. Thus, infrastructure studies require drawing together methods that are equal to the ambitions of its phenomenon.

This leads directly into our second theme, an integrative view. In order to understand the new information infrastructure we need to move beyond seeing the social, organizational, and cognitive sitting somehow on top of or beside the wires and gateways of the physical infrastructure. Each layer is riven through with each of these dimensions – and we need to train social scientists and information scientists to move freely between all of them. This is not just a good idea – it is something of a law if we want to fully understand emergent phenomena in the development of new ways of knowing.

Our third theme is a direct consequence: we need to be sensitive to the development and spread of new ways of knowing across information infrastructures. Consider the issues of classification systems, ontologies, and metadata that subtend so much work on the web. Patrick Tort has written a marvelous book about the development of classification systems in the nineteenth century (Tort, 1989). In this book he charts the rise and eventual dominance of genetic classification systems, which sort things into piles by their origin points (the origin of species, for example) in a whole range of different fields. Now there was at this point no "science" of classification, no set of conferences that everyone was going to – the new systems spread virally among many different populations. The development of information infrastructures facilitates the viral spread of ontologies and classifications – indeed when you look under the hood of various cyberinfrastructure projects you often find the same underlying architecture and design philosophy being promoted by a small set of actors. To the extent that we train social scientists to look *within* particular fields or projects, we are missing one of the most interesting features of information infrastructures: their ability to promote combined (if uneven) development across a very broad range of work practices.

We are convinced that we are in the midst of developing fundamentally new ways of knowing – though we would put this in a 200 year span rather than a machine-centered 50 year frame. Understanding these in order to rethink the nature of knowledge on the one hand and improve design on the other entails developing fundamental new ways of working in natural and social science. We need to move beyond the "endless frontier" model of building information infrastructure and start to look at just what homesteading means in this new landscape. Information infrastructure is a great tool for distribution of knowledge, culture, and practice. Homesteading the space it has slowly opened out over the past two centuries involves building new kinds of community, new kinds of disciplinary homes, and new understandings of ourselves.

References

Abbott, A. (1988). *The system of professions: An essay on the division of expert labor.* Chicago: University of Chicago Press.

Arthur, B. (1994). *Increasing returns and path dependence in the economy.* Ann Arbor: University of Michigan Press.

Ashby, W. R. (1956). *Introduction to cybernetics*. London: Chapman & Hall.

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., Messerschmitt, D. G., et al. (2003). *Revolutionizing science and engineering through cyber-infrastructure*. Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure.

Baker, K. S., & Millerand, F. (2007, January 3–6). *Articulation work supporting information infrastructure design: Coordination, categorization, and assessment in practice*. Proceedings of the 40th Hawaii International Conference on System Sciences (HICSS), Big Island, Hawaii, IEEE Computer Society, New Brunswick, NJ.

Beniger, J. R. (1986). *The control revolution: Technological and economic origins of the information society*. Cambridge, MA: Harvard University Press.

Berg, M., & Timmermans, S. (2000). Orders and their others: On the constitution of universalities in medical work. *Configurations, 8*, 31–61.

Bowker, G. C. (1994). *Science on the run: Information management and industrial geophysics at schlumberger*, 1920–1940. Cambridge, MA: MIT Press.

Bowker, G. C. (2000). Biodiversity datadiversity. *Social Studies of Science, 30*(5), 643–683.

Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Bowker, G. C., Timmermans, S., & Star, S. L. (1995). Infrastructure and organizational transformation: Classifying nurses' work. In W. Orlikowski, G. Walsham, M. R. Jones, & J. I. DeGross (Eds.), *Information technology and changes in organizational work* (pp. 344–370). London: Chapman and Hall.

Brand, S. (1999). *The clock of the long now: Time and responsibility: The ideas behind the world's slowest computer.* New York: Basic Books.

Bratteteig, T. (2003). *Making change: Dealing with relations between design and use*. PhD Thesis, University of Oslo, Norway.

Castells, M. (1996). *The rise of the network society*. Cambridge, MA: Blackwell Publishers.

Chandler, A. D., Jr., & Cortada, J. W. (2003). *A nation transformed by information: How information has shaped the United States from Colonial times to the present* (pp. 185–226). New York, NY: Oxford University Press.

Chandler, A. D., & Hikino, T. (1994). *Scale and scope: The dynamics of industrial capitalism.* Cambridge, MA: Belknap Press.

Clanchy, M. T. (1993). *From memory to written record, England 1066–1307* (2nd ed.). Oxford, Cambridge: Blackwell.

David, P. (1985). *Computer and dynamo: The modern productivity paradox in a not-too-distant mirror*. Stanford: Center for Economic Policy Research, Stanford University.

Desmond, A. J., & Moore, J. R. (1991). *Darwin.* New York: Warner Books.

Dewey, J. (1902). *The child and the curriculum*. Chicago: University of Chicago Press.

Edwards, P. (2003). Infrastructure and modernity: Force, time, and social organization. In T. J. Misa, P. Brey, & A. Feenberg (Eds.), *The history of sociotechnical systems. Modernity and technology* (pp. 185–226). Cambridge, MA: The MIT Press.

Edwards, P. N., Jackson, S. J., Bowker, G. C., & Knobel, C. (2007). *Understanding infrastructure: Dynamics, tensions, and design*. NSF Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures". Retrieved from http://hdl.handle.net/2027.42/49353.

Engestrom, Y. (1990). *Learning, working, and imagining*. Helsinki: Orienta-Konsutit Og.

Epstein, S. (2007). *Inclusion: The politics of difference in medical research*. Chicago: University of Chicago Press.

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: A new abstraction for information management. *SIGMOD Record, 34*(4), 27–33.

Futrell, J., & the AC-ERE (2003, May). *Environmental cyberinfrastructure (ECI): Tools for the study of complex environmental systems*. http://catalog.rcls.org/ipac20/ipac.jsp?session=12746364037TU.158230&profile=mid&uri=link=3100007∼!974625∼!3100001∼!3100002&aspect=subtab111&menu=search&ri=1&source=∼!anser&term=Environmental+cyberinfrastructure+%28ECI%29+tools+for+the+study+of+complex+environmental+systems.&index=PALLTI

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Thousand Oaks, CA: Sage Publications.

Goffman, E. (1956). *The presentation of self in everyday life*. New York: Doubleday.

Goody, J. (1987). *The interface between the written and the oral*. Cambridge: Cambridge University Press.

Hanseth, O., Monteiro, E., & Hatling, M. (1996). Developing information infrastructure: The tension between standardization and flexibility. *Science, Technology & Human Values, 21*(4), 407–426.

Hayden, C. (2003). *When nature goes public: The making and unmaking of bioprospecting in Mexico*. Princeton: Princeton University Press.

Hedstrom, M., Dawes, S., Fleischhauer, C., Gray, J., Lynch, C., McCrary, V., et al. (2002). It's about time: Research challenges in digital archiving and long-term preservation committee: NSF-DGO Report on NSF Workshop on research challenges in digital archiving. Digital Gov Program, NSF Div of Info and Intelligent Systems.

Hobbie, J. E., Carpenter, S. R., Grimm, N. B., Gosz, J. R., & Seastedt, T. R. (2003). The US long term ecological research program. *BioScience, 53*(2), 21–32.

Howell, J. M., & Higgins, C. A. (1990). Champions of technological innovation. *Administrative Science Quarterly, 35*(2), 317–341.

Hughes, T. P. (1983). *Networks of power: Electrification in western society, 1880–1930*. Baltimore: John Hopkins University Press.

Hughes, T. P. (1989). Evolution of large technological systems. In W. E. Bijker, T. P. Hughes, & T. J. Pinch (Eds.), *The social construction of technological systems: New directions in the sociology and history of technology* (pp. 51–82). Cambridge, MA; London: The MIT Press.

Karasti, H., & Baker, K. (2004, January). *Infrastructuring for the long-term: Ecological information management*. Proceedings of the 37th Hawaii International Conference on System Sciences. HICSS38, IEEE Computer Society, Big Island, Hawaii.

Karasti, H., Baker, K. S., & Halkola, E. (2006). Enriching the notion of data curation in eScience: Data managing and information infrastructuring in the long term ecological research (LTER) network. In M. Jirotka, R. Procter, T. Rodden, & G. Bowker (Eds.), *Computer Supported*

Cooperative Work: An International Journal. Special Issue: Collaboration in e-Research, *15*(4), 321–358.
Kirk, S. A., & Kutchins, H. (1992). *The selling of the DSM: The rhetoric of science in psychiatry*. New York: Aldine de Gruyter.
Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford: Pergamon Press.
Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In W. E. Bijker & J. Law (Eds.), *Shaping technology, building society: Studies in sociotechnical change* (pp. 225–258). Cambridge: MIT Press.
Latour, B. (1993). *La Clef de Berlin et Autres Leçons d'un Amateur de Sciences*. Paris: La Decouverte.
Latour, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Beverly Hills: Sage Publications.
Lawrence, K. A. (2006). Walking the tightrope: The balancing acts of a large e-Research project. *Computer Supported Cooperative Work, 15*(4), 385–411.
Leibniz, G. W., Clarke, S., & Alexander, H. G. (1998). *The Leibniz-Clarke correspondence*. New York: Manchester University Press.
Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity, 7*(4), 273–290.
Lynch, M. (1985). *Art and artifact in laboratory science*. Boston, MA: Routledge and Kegan Paul.
MacKenzie, D. (1993). *Inventing accuracy: A historical sociology of nuclear missile guidance*. Cambridge, MA: The MIT Press.
Magnuson, J. J. (1990). Long-term ecological research and the invisible present. *BioScience, 40*(7), 495–501.
Millerand, F., & Bowker, G. C. (2009). Metadata standards, trajectories and enactment in the life of an ontology. In S. L. Star & M. Lampland (Eds.), *Formalizing practices: Reckoning with standards, numbers and models in science and everyday life* (pp. 149–167). Ithaca, NY: Cornell University Press.
Monteiro, E., & Hanseth, O. (1997). Inscribing behaviour in information infrastructure standards. *Science, Technology & Human Values, 21*(4), 407–426.
Nardi, B., & O'Day, V. (1996). Intelligent agents: What we learned at the library. *Libri, 46*, 59–88.
NSF Cyberinfrastructure Council (2006). NSF's cyberinfrastructure vision for 21st century discovery (Version 7.1). Retrieved from http://www.nsf.gov/od/oci/ci-v7.pdf
Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction, 15*, 139–178.
Porter, T. M. (1986). *The rise of statistical thinking, 1820–1900*. Princeton, NJ: Princeton University Press.
Ribes, D. (2006). *Universal informatics: Building cyberinfrastructure, interoperating the geosciences*. Unpublished Ph.D. dissertation, Department of Sociology (Science Studies), University of California, San Diego.
Ribes, D., & Baker, K. S. (2007). *Modes of social science engagement in community infrastructure design*. Proceedings of Third International Conference on Communities and Technology, Springer, London, pp. 107–130.
Ribes, D., & Bowker, G. C. (2008). GEON: Organizing for interdisciplinary collaboration. In G. M. Olson, J. S. Olson, & A. Zimmerman (Eds.), *Science on the internet* (pp. 311–330). Cambridge: MIT Press.
Sawyer, S., & Tapia, A. (2006). Always articulating: Theorizing on mobile and wireless technologies. *The Information Society, 22*(5), 311–323.
Schmidt, K., & Bannon, L. (1992). Taking CSCW seriously: Supporting articulation work. *Computer Supported Cooperative Work, 1*(1), 7–40.
Schon, D. A. (1963). Champions for radical new inventions. *Harvard Business Review,* 41, 77–86.
Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. New Haven: Yale University Press.

Serres, M. (1990). *Le Contrat Naturel.* Paris: Editions Françous Bourin.
Shapin, A. P. (1989). The invisible technician. *American Scientist, 77*, 554–563.
Sowa, J. (2000). *Knowledge representation: Logical, hilosophical, and computational foundations.* Pacific Grove, CA: Brooks/Cole.
Star, S. L. (1991). The sociology of the invisible: The primacy of work in the writings of Anselm Strauss. In D. Maines (Ed.), *Social organization and social process: Essays in honor of Anselm Strauss* (pp. 265–283). Hawthorne, NY: Aldine de Gruyter.
Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist, 43*(3), 377–391.
Star, S. L., & Ruhleder, K. (1994, October 22–26). Steps towards an ecology of infrastructure: Complex problems in design and access for large-scale collaborative systems. Proceedings of the Conference on Computer Supported Cooperative Work (CSCW 94 – Transcending Boundaries), Chapel Hill, NC. ACM Press, New York, pp. 253–264.
Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research, 7*(1), 111–134.
Stinchcombe, A. (1990). *Information and organizations*. Berkeley: University of California Press.
Strauss, A. (1988). The articulation of project work: An organizational process. *The Sociological Quarterly, 29*(2), 163–178.
Strauss, A. L. (1993). *Continual permutations of action*. New York: Aldine De Gruyter.
Tort, P. (1989). *La Raison Classificatoire: Les complexes Discursifs – Quinze Etudes*. Paris: Aubier.
Weick, K. E. (1979). *The social psychology of organizing*. Reading, MA: Addison-Wesley Publishing Company.
Weiner, N. (1951). *Cybernetics and society*. New York, NY: Executive Techniques.
Yates, J. (1989). *Control through communication: The rise of system in American management*. Baltimore, MD: Johns Hopkins University Press.