

Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories

David Ribes¹ & Charlotte P. Lee²

¹*Communication, Culture & Technology Program (CCT), Georgetown University, 3520 Prospect St. NW, Suite 311, Washington, DC 20057, USA (E-mail: dr273@georgetown.edu);* ²*Department of Human Centered Design & Engineering (HCDE), University of Washington, 423 Sieg Hall, Box 352315, Seattle, WA 98195, USA (E-mail: cplee@uw.edu)*

1. Introduction

Cyberinfrastructure (CI), eScience and eInfrastructure are the current terms of art for the networked information technologies supporting scientific research activities such as collaboration, data sharing and dissemination of findings. These are the computational infrastructures that enable, for instance, global climate modelers to compile heterogeneous information sources in order to understand environmental change or the tools that make the massive quantitative data emerging from the Large Hadron Collider into tractable scientific visualizations. Within the US and Europe these ventures have garnered significant momentum in terms of funding and technological development. The greater funding of CI for the physical and biological sciences has led to a proliferation of CI studies in those areas (a bias reflected in our own special issue) with CI studies of humanities, arts, and social sciences growing more slowly. Cyberinfrastructure is heralded as a transformative force, enabling new forms of investigation and cross-disciplinary collaboration. The six articles in this special issue are dedicated to the detailed empirical exploration of CI systems, their development, and the changes that are emerging at the intersection of novel information technologies and everyday research practice.

The emphasis in the sciences on large-scale and long-term support for collaboration is gaining momentum. The field of Computer Supported Cooperative Work (CSCW), with its historical emphasis on workplace studies and

Dedication We dedicate this special issue to the memory of Susan Leigh Star. The influence of her work, and particularly her conceptualization of infrastructure, has been very influential in this growing research area. Almost all of the articles in this special issue cite her work or cite other works that were influenced by her own. Less apparent to others but strongly present in our lives was Leigh's role as a mentor, colleague and inspiration as she fostered the study of infrastructure. Her pioneering work demonstrated the important role of embedded social scientists that subsequently created openings to the field sites and funding sources that made much of this research possible. Leigh, you will be greatly missed.

collaborative system development (Grudin 1994), is well positioned to play an important role in the study of cyberinfrastructure enabled (and disrupted) scientific work. Cyberinfrastructure studies do not represent a wholly new wing of CSCW, but more detailed and systematic attention is required in order to better understand and facilitate this particular kind of computer supported cooperative work.

Three features characterize cyberinfrastructure's promised transformations in the sciences: community wide and cross-disciplinary collaboration, computationally driven collection, representation and analysis of data, and end-to-end integration.

Community wide and cross-disciplinary collaboration refers to the scope of CI projects. Rather than supporting teams or groups, CI practitioners speak in terms of communities, disciplines and domains. It is common to hear sweeping umbrella terms that redefine traditional disciplinary boundaries into vast domains of investigation, i.e., infrastructure for the earth-sciences, brain sciences, ocean sciences or environmental engineering. It is in this sense that CI is often attached to phrases such as 'the new big sciences'. While there are small scale projects that are characterized as CI, these are often wrapped up into larger assemblies of data standards, common services and shared computational infrastructures. Cyberinfrastructure departs from the classic big sciences (Galison and Hevly 1992; Price 1963) which were goal directed large-investment collaborations of scientists and engineers (as with the Manhattan Project, the Human Genome Project or the Large Hadron Collider). Instead, CI seeks to support research generally by identifying vast interdisciplinary swaths that could benefit from data and resource sharing, knowledge transfers, and support for collaboration across geographical, but also institutional and organizational divides.

These transformations are commonly initiated as collaborations of the natural and computational sciences, though CI ventures have now also extended to the social sciences and humanities. This generalization of disciplinary scope is heralded in changing terminology such as e-Research or e-Infrastructure (Edwards et al. 2009; Jankowski 2007; Olson et al. 2008). In this issue we refer broadly to these endeavors as cyberinfrastructure. Common to all these ventures is the leading role of computational, engineering and informational actors and the promised new technologies that are their research interest. These projects take into account the internet and the exponentially growing availability of computational resources, but they propose to take scientific infrastructure one step further through deployment of standardized sensor networks and ambitious sets of commonly available tools for storing, manipulating, sharing and analyzing the forthcoming 'data deluge' (Borgman 2007).

All these new features will be tied together in systems that are integrated 'end-to-end' (Gillespie 2006), arguably the most ambitious of the promises of CI. On one end of the system are the data collection and the distributed, standardized and often automated sensors and imaging technologies. At the other end are the

heterogeneous scientific users, working on richer and more diverse sources of data than ever before, using tools that seamlessly facilitate analysis, collaboration and dissemination of findings. Between data collection technologies and users lie the networks, computers, storage and a plethora of more subtle integration technologies that promise to facilitate communication and interoperation across all the boundaries that plague interdisciplinary collaboration: hard technologies such as fiberoptic cables and grid computing, soft technologies such as metadata standards and ontologies, and even softer on-paper agreements between institutions and agencies of science to facilitate the movement of 'siloed' data and findings. It is this massive assembly of computational tools, resources and collaborative support that together constitute cyberinfrastructure and promise to revolutionize the sciences (Atkins 2003).

In this special issue of JCSCW dedicated to the sociotechnical study of cyberinfrastructure we take these promises as our topic rather than a resource. We do not assume these outcomes but instead explore the work of their achievement as a matter of sociotechnical work, interdisciplinary collaboration and institutional formation. The authors of the six articles in this issue investigate the work of achieving these goals, the difficulties that practitioners encounter and how these tensions are resolved in order to enact the ambitions of CI. Below we present a theoretical summary by articulating seven common themes of cyberinfrastructure studies. We then follow with methodological approaches (or 'tricks of the trade') that are commonly employed to make these ambitious projects knowable to CSCW scholars. The themes and methodologies serve as an introduction to the articles within this special issue. Science is only a single context where large scale and long term information technologies are being deployed; for CSCW scholars, these themes and research approaches will also offer avenues to open a research terrain around a new generation of computational systems oriented to supporting work. We conclude with a discussion of the current limitations and future directions for the sociotechnical studies of cyberinfrastructure.

2. Themes of cyberinfrastructure studies

Used colloquially, the term infrastructure has come to prominence in many spheres of action: government, business, and, the focus of this issue, science. Infrastructure's initial meaning brought to mind the 'heavy steel' of rail and bridges or the 'tubes and wires' of sewage, electricity and the internet. It was within the ranks of CSCW itself that social scholars of infrastructure began to broaden its meaning and flesh out a theoretical understanding of infrastructure. Infrastructure came to be seen not as a 'what' so much as a sustained relation: "infrastructure means different things to different groups and it is part of the balance of action, tools, and the built environment, inseparable from them" (Star 1999, p.377).¹ One person's support structure is another person's everyday work of design, maintenance and upgrade. This Pragmatist

definition of infrastructure forestalls discussions of ‘what is really infrastructure’ and instead directs the researcher to ask whether there is analytic value in examining a phenomenon *as infrastructure*: what is supporting the work of another and who is sustaining those relationships? *Relationality* is the first of seven themes of cyberinfrastructure studies that have become thematic to the field and that are addressed in the articles of this special issue.

A second theme is the characterization of those relations, and in particular managing difference and the *integration of heterogeneity*. The focus is most commonly on the interdisciplinary relations that must be fostered in order to develop cyberinfrastructure and the novel science that it is hoped CI will thereafter bring about. A common framing of these differences points to the principal actors in CI: on the one hand the computer and information scientists that develop the technological systems and on the other hand the domain scientists that collaborate with them to develop technology, but who are ultimately interested in the final stable and functioning scientific resources (Lawrence 2006; Ribes and Bowker 2008). Upon closer inspection, these two groups splinter into many more actors with diverse skill sets and responsibilities, such as the data and information managers whose daily work is most closely tied to the sciences (Karasti et al. 2010, this issue); the program officers and representatives of the funding agencies that funnel financial and visionary support to CI (Kee and Browning 2010, this issue); and the everyday technicians that transition experimental technologies to production quality systems (Bietz et al. 2010, this issue). From these heterogeneous relations are spawned many of the tensions that characterize the development of large-scale, long-term and interdisciplinary infrastructure projects: aligning the end-goals of multiple participants, motivating contribution in the face of competing reward structures and designing novel technologies that are usable and stable (Ribes and Finholt 2009).

However integration of heterogeneity is not limited to human actors, CI’s diversity also rests in its technological systems and social institutions. CI includes a goal to seamlessly integrate the diverse technologies that make up these systems—sensors, databases, processors, operating systems, and visualization tools to name only a few components—and the ways that humans interact with them. The data that populate these systems are also heterogeneous, and each CI endeavor brings with it efforts to account for data’s context of production (Faniel and Jacobsen 2010, this issue), context of use (Bietz and Lee 2009), and its means of dissemination across multiple disciplinary communities and often also long periods of time (Karasti et al. this issue). For instance, Faniel and Jacobsen discuss the particularities of evaluating trustworthy data in a subfield of Earthquake Engineering, thus pointing at the difficulties of trying to achieve data that is trustworthy across disciplines and highlighting the difficulty of attempting to serve domains with different needs.

Sustainability is the third theme of cyberinfrastructure studies. Infrastructure is intended as a long term resource, certainly it has “reach beyond a single event or

one-site practice” (Star and Ruhleder 1996, p.115), but with cyberinfrastructure the goal is to look years and even decades ahead. This necessitates what Stuart Brand has called thinking in the Long Now (Brand 1999; Ribes and Finholt 2007), a form of design that considers immediate needs while also incorporating future requirements and planning for emerging changes in technology and scientific practice. In this issue Bietz et al. (2010) address this as a matter of thinking through the gap between intentional actions of design and the actuality of the emerging infrastructure. Thinking about sustainability necessitates the careful consideration of maintenance and breakdown. The technical staff that sustain everyday functioning over time and help to smooth transitions from established routines to novel technical configurations cannot act alone. Karasti et al. (2010, this issue) in particular address this as a matter of reconciling (short-term) project time with (long-term) infrastructure time. Their study of the implementation of a metadata standard highlights the integration of multiple kinds of heterogeneities: developers and information managers—but also different kinds of data—all with an eye to the long term sharing of resources across disciplinary boundaries. The response in their case study of Long-Term Ecological Research (LTER) is standardization of a data specification across the multiple disciplines that make up the ocean, atmospheric and geosciences. Data sharing too can easily be framed as an either-or technical question (How do we make data pipelines interoperate?) or a cooperation question (How do we get people to agree to share?). Karasti et al. (2010, this issue) show us the struggle of information managers, who often have a different temporality from the full-time developers they work with, and who strive to find a middle ground that will enable both stability and flexibility from both social and technical perspectives.

Standardization is both a goal and a method within many cyberinfrastructure ventures. It is one of the principle solutions to the integration of heterogeneity, whether of disciplinary difference (Ribes and Bowker 2009), data conventions (Millerand and Bowker 2009) or systems integration. Standardization is the foundation of interoperability, a promised technical solution that is underwritten by innumerable pacts and agreements. If a technical standard is ‘official’ (supported by a leading scientific organization, or funded to be a community standard) its development is often characterized by debates and difficulties in deployment. This is because adoption of a standard is a matter of changing the everyday practice for many kinds of actors. For example, in adopting a data standard to facilitate sharing (Karasti et al. 2010, this issue) or reuse (Faniel and Jacobsen 2010, this issue), scientists and their information managers must change their local practices of data collection and curation to match a common set of methods and styles of annotation (metadata). These changes percolate throughout entire scientific organizations in ways that are difficult to foresee in advance.

To study standardization is to inspect the ways that local divergence from a desired norm is managed, integrated and/or eradicated; how workarounds and gateways (David and Bunn 1988; Egyedi 2001) can add flexibility to rigid standard implementations; and, how boundary work (Star and Griesemer 1989;

Bietz and Lee 2009) can serve to translate and interlink difference while sustaining it. The papers in this issue show how adoption of standards allows participation in a more collective research practice. In the face of such global growth of infrastructure, a focus on standardization and its consequences can help us understand inherent tensions between local and global practice in cyberinfrastructure.

A key ambition of cyberinfrastructure is *scaling up or extension* of infrastructure: whether the number of collaborators, the quantity of data, availability of raw computing cycles or broader geographic reach. CSCW has traditionally focused on small group collaboration scenarios, but CI's goals are to support team activities while also keeping in mind entire communities of scientists. All of this is a form of growth, and thus a concern accompanying sustainability is extensibility. Extensibility is a design principle that takes into consideration future growth of the data system; as with Brand's Long Now, extensibility necessitates today's consideration of a distant future. Growth of the system can be through the addition of new functionalities or through modifications of existing functionalities. The central theme is understanding how to plan for change while minimizing impact to existing capacities. However, a concern with scaling up in these projects must be balanced with a concern with the everyday scientific practices these large systems are intended to support. In this issue we also draw attention to the often ignored technical minutiae of maintenance work, such as with de la Flor, Jirotko, Luff, Pybus and Kirkham's and Monteiro's studies of tool use within large CI enterprises.

All the above themes wrap together in a higher order theoretical issue of *the distribution between human work and technological delegation*. Cyberinfrastructure promises a lessening of the required effort for the accomplishment of labor intensive or tedious tasks. Data should flow across institutions without having to renegotiate cross-institutional agreements, computing cycles should be available on demand, and information visualization should facilitate scientific interpretation and analysis. All of this is a form of automation, or more accurately, a redistribution of labor between humans and technologies. As Lee et al. (2009) have noted, cyberinfrastructure is by no means devoid of a 'human infrastructure' that develops and maintains operations. In this issue Bietz et al. (2010) elaborate this concept by theorizing the day-to-day work of building infrastructural relationships. But just as technical workers can become invisible in the smooth operation of infrastructure, so too can its technological components: how are the novel tools of collaboration, of data collection or visualization encouraging or discouraging different kinds of scientific organization, investigation or dissemination? This overarching question is addressed through a detailed analysis of a novel technology in action in de la Flor et al. (2010, this issue). In their study of classicists interpreting an ancient manuscript, the Tolsum Tablet, they demonstrate that careful consideration of detailed ethnographic data can yield thoughtful iterative technical interventions that can support and amplify existing work practices.

Computer Supported Cooperative Work has from inception been a pioneering field for experimenting with novel intersections of the social, computational, information and design sciences and studies. Similarly, cyberinfrastructure itself has presented new opportunities and challenges to cross these great divides (Bowker et al. 1997). Arguably, it is the work of CSCW scholars and our like-minded kin in fields such as participatory design and science and technology studies (STS) that have inspired the leading figures in CI development and funding to pose a vision of the *always already social of cyberinfrastructure*. That is, both the highest order programmatic manifestos of cyberinfrastructure and the more mundane everyday activity of CI development are often conceived of as a social venture, often involving the direct participation of social scientists, and always seeking to imagine the impact on everyday scientific work. What follows from this conceptualization have been requests for the participation of ethnographers, social network analysts and other breeds of social scientists in various facets of CI development and enactment. At the same time that CI projects are pulling social scientists into project work, social scientists are also pushing themselves into the CI domain by negotiating entrée wherever there might be willing research participants. While self-funded scientists incur fewer obligations, conflicts, and distractions, the lack of a position on the critical path for work completion can result in occasionally inaccessible research sites, particularly during project upheaval.

3. ‘Tricks of the trade’: methodologies and challenges in CI studies

Below we outline some of the methodologies (or more revealingly, as Howard Becker would describe them, ‘tricks of the trade’) for tackling the challenges of studying cyberinfrastructure. The approaches we describe comprise a basic toolbox for the CI researcher.

Cyberinfrastructure is nascent and novel. The vast majority of studies to date are *investigations of ongoing planning, development and deployment efforts*. Such investigations are immensely rich, revealing the everyday goings on that come to constitute promised revolutions in technology. The hands-on methods represented in this special issue are as diverse as those found in the broader fields that make up infrastructure studies (historical, ethnographic, documentary, and interview-based to name a few), but the methodologies (or theories of inquiry) have coalesced around a handful of approaches that seek to make knowable the complexities of CI. Often conducted by ethnographers that combine participant observation and documentary analysis, these studies can be of the daily meetings of practitioners (as with Monteiro 2010, this issue; de la Flor et al. 2010, this issue; and Bietz et al. 2010, this issue), interview-based studies (as with Kee and Browning 2010, this issue; Faniel and Jacobsen 2010, this issue; and Bietz et al. 2010, this issue) or more virtually by tracking the vast streams of data created in these projects by emails, reports, published papers and online documentation (as

with Karasti et al. 2010, this issue). These are the moments of ‘social shaping of technology’ (Pinch and Bijker 1984) that later become black boxed, taken for granted and treated as inevitable. The careful documentation of the debates, choices and design decisions allow us to see how infrastructural resources come to be what they are, who they serve, and also allow us to imagine how they may have been otherwise.

A common technique in the study of infrastructure is to focus on activities of maintenance, upgrade and breakdown. A shorthand for this approach is *infra-structural inversion* (Bowker 1994), which refers to the focus on all the activities that hold together the functioning of infrastructure rather than those that it invisibly supports. For instance, the papers in this special issue that explore technical labor (whether of development or maintenance as in Karasti et al. (2010) or Bietz et al. (2010)) are inversions of the infrastructure, taking the enabling possibility as the focus rather than the activity that is supported (in this case, scientific work, as with de la Flor et al. (2010, this issue) and Monteiro (2010, this issue)).

Breakdowns themselves are a kind of natural infrastructural inversion. As Star and Ruhleder note, “The normally invisible quality of working infrastructure becomes visible when it breaks” (Star and Ruhleder 1996, p.115). It is at moments of failure that entire swaths of infrastructural activity (even those that didn’t fail such as properly functioning backups or routinized repair activities) are revealed not only to the analyst but to the everyday user of infrastructure. *Focusing on breakdown* is a variation of what are called ‘controversy studies’ in STS. For instance, debates over the etiology of AIDS (Fujimura and Chou 1994) or the existence of gravity waves (Collins 1981). In the sciences disagreement amongst practitioners can bring to the fore all the assumptions usually held tacitly by scientists about their fields, their methods and what they know about the world. Similarly, within infrastructure ventures (particularly those that promise to change the direction of an entire field) controversies can reveal the underlying expectations and turf battles that undergird seemingly ‘objective’ technology development. Karasti et al. (2010, this issue) explore the difficulties in developing and deploying a metadata language as technology developers and information managers’ interests go head to head. Kee and Browning (2010, this issue) explore tensions on institutional, individual, and ideological levels in funding infrastructure to show how an organization is built up over time by commitments across people and institutions.

Furthermore, controversies require explicit articulation by actor’s themselves. That is, the jargon and deep technical expression of normal scientific activity is often cleaned up as participants seek to clarify their position and convince others with more compelling evidence. *Tracing the adoption of expressions and their changing use* over time is a way to explore the development and negotiation of new practices and whole scientific communities.

The method of *tracing the adoption and use of a new technological artifact*, as a way to explore new sociotechnical arrangements (de la Flor et al. 2010, this

issue) will continue to be useful in CI just as it has long been for CSCW. In this issue, Monteiro's paper (2010, this issue) shows how an ethnographic study can give lie to the notion that digitization results in a disembodied practice divorced from the natural world. Advanced 4D computer modeling relies heavily on embodied understanding, and support for manipulation is of paramount importance; models are accepted as evidence because they refer to the natural world.

Having described some useful methodologies for the study of CI, we turn to a discussion of the challenges to which they are applied.

3.1. Challenging in studying cyberinfrastructure

The study of cyberinfrastructure presents formidable challenges and opportunities. Some of the research challenges are common to fields such as STS, the practitioners of which have developed multiple approaches to faithfully grasp the technical 'content' of a variety of scientific domains, or studies of large or distributed phenomena (Beaulieu 2010; Hine 2000). Cyberinfrastructure resides at the intersection of many of these methodological entanglements: it is technical (often doubly so, requiring a handle on information technologies and a domain science); it is distributed (nationally and often internationally); it is long term, at times stretching years into the past and prospecting into a future of decades; and brings together multiple heterogeneous expert actors.

While it is not an explicit theme in any of the articles in this special issue, behind each author's work reside multiple kinds of collaborative agreements that enable 'access' to the research site. Social scientists have acted as 'shrink wrapped' consultants within these projects, stepping in to provide expert advice, or have become committed members in give and take relationships that extend for years and possibly careers. Some of the authors in this issue come closer to being participants in CI projects than analysts of them. The result is wealth of perspectives and access points to these sites of research, but conversely this presents troubling issues that are not novel to social science: a challenge to traditional notions of objectivity and disinterestedness, limited openings for critical analysis and the pressure to neglect our own research in order to think and act as managers.

The stages of cyberinfrastructure development present a common set of difficulties whether in a nascent stage of development (Bietz et al. 2010, this issue; de la Flor et al. 2010, this issue), in practical deployment and operation (Faniel and Jacobsen 2010, this issue; Monteiro 2010, this issue), or stabilized as an institution (Karasti et al. 2010, this issue). These challenges are reflected in many of the themes above e.g., long term, large-scale, geographically distributed projects with high-level expert actors (both scientifically and technologically). The limitation of development phase studies (most designated cyberinfrastructures are nascent) is tied to the long cycles before CI can be considered as

successful, stable, or even usable infrastructure. Even in the most ambitious formulations, CI funding projects are 3–5 years, but from experience we can say that infrastructure is a matter of twice that number and probably in the range of decades before reaching maturity and anything approaching closure. Locating CI in a longer lineage of information systems development efforts (Edwards 1996) for the sciences (Abbate 1999; Bowker 2006) or as overlapping and overflowing as part of development efforts in other spheres, such as business (Haigh 2001), could alleviate some of the historical myopia of focusing on ongoing projects. Alternately, a focus on ongoing projects with informational components that have only recently adopted the moniker of CI (Karasti et al. 2010, this issue) can provide a strategy for studying more mature projects.

Cyberinfrastructure is relational, heterogeneous and extensible. What is the appropriate granularity with which to approach the study of cyberinfrastructure? There is no single answer, and the articles in this paper reflect multiple scales of analysis. de la Flor et al. (2010, this issue), for example, tackle a large scale project developing technologies to support the study of ancient texts and inscriptions, but then zooms in to the scale of scientific practice, employing the eye of an ethnomethodologist to track visualization tools in use.

The most common unit of analysis within CI studies is the project. GEON, LEAD, LTER, WATERS, BIRN, SEEK, NEES, CAMERA and an endless stream of trumpeted acronyms have become assumed shorthand to practitioners and analysts alike. However many such projects have been funded through common programs that encouraged large scale information system development (e.g., WATERS and NEES through the National Science Foundation's (NSF) MREFC, and GEON and LEAD through NSF ITR). Projects are not initiated or enacted in a vacuum. Moving *beyond the project scale* and treating CI as a holistic phenomenon is something sorely needed in the field. Kee and Browning (2010, this issue) take the arena of cyberinfrastructure development as the unit of analysis. Here CI is a national policy issue, a matter of agency level directives and funding incentives. Another approach is comparative, across multiple CI projects or across national funding regimes (for example, the field is ripe for a study comparing eScience, e-Research and Cyberinfrastructure in Europe, the UK and the US although we know of no such study forthcoming).

While projects are not conducted in isolation, this does not mean they are conducted in coordinated collaboration. Within any given field or subfield there are multiple competing computerization movements, standards, and scientific theories. While CI is described as singular ('the national cyberinfrastructure') seeking to unify resources under a common banner, in practice it also exists within an agonistic field of competing technical approaches, limited financial resources, and even more fundamental debates about the nature of the phenomenal world. The number of challenges facing CI studies is considerable, however the opportunities for CI and e-Research are considerably greater.

4. Conclusion: investigating design futures

This is the fourth special issue of a journal (the second of JCSCW) dedicated to the sociotechnical study of cyberinfrastructure (Jirotko et al. 2006; Jankowski 2007; Edwards et al. 2009). In addition there are four book-length treatments of the issue (Hine 2006; Jankowski 2009; Nentwich 2003; Olson et al. 2008). While we cannot yet point to a field, studies of CI are flourishing. This said, we remain at a stage of experimentation and exploration. With only an incipient set of ongoing themes and methods, we still require higher order synthesis and a deeper understanding of the long-term implications of CI, especially for the scientific establishment itself. The themes we present here of *relationality, integration of heterogeneity, sustainability, standardization, scaling up or extension, human work and technological delegation, and the always already social of CI*, represent the current landscape of CI studies—a sketch of a few notable landmarks and some approximate borders. We have only just begun charting this territory and there is much to discover.

Three investigative areas primed for rapid growth in CI studies include, in no particular order: how to support the particular practices of specific scientific sub-domains, how to facilitate data sharing, and how to support virtual organizations that have components of traditional organizations, but also have highly distributed and often short-term resources. It is important to note that these three areas are by no means mutually exclusive. One need only think of designing a web portal to upload and download data from a shared repository in order to find a research site that crosses all three areas of concern. However, as the CI literature matures, these areas provide a way to parse a research landscape that is as diverse as cyberinfrastructure itself. Numerous workshops relating to CI have taken place at the European Computer Supported Cooperative Work Conference (ECSCW), the ACM Computer Supported Cooperative Work Conference (CSCW), and the ACM Computer-Human Interaction (CHI) Conference to name just a few, and the diversity of attendee interests ranging from the micro to the macro and from the technological to the social to the scientific has been staggering. Given that each workshop by chance or by plan had a particular slant, no doubt some attendees at each wondered if they had wandered into the wrong tent. The fact of the matter is that CI, like so many phenomena, is not monolithic. As the creation of CI is itself incredibly heterogeneous and requires the collaboration of people of diverse interests and backgrounds, so too must be the researchers of CI.

It is tempting to engage with cyberinfrastructure phenomena through a narrow analytic lens (e.g., framing CI as fundamentally about hardware and software, or just at an organizational or team level), but the papers in this issue demonstrate the richness that is gained by considering the interplay of the social and technical from diverse viewpoints. The stories of cyberinfrastructure are revealed by looking across multiple levels of granularity, various facets of social life, and diverse technological actors. Much remains to be studied in the areas of

supporting domain specific practice, data sharing and curating, and infrastructural organizing. This is an exciting time for CI studies. Research is occurring in new and unexpected places, drawing on and bringing together the traditions of CSCW, information science, organizational studies, and science and technology studies. This cross-pollination, as exemplified by the papers in this issue, seems to be not only fruitful, but also very necessary.

Acknowledgements

We owe our thanks to many. This special issue would not have been possible without the insight and contributions of our co-editors Matthew Bietz, Helena Karasti, and Marina Jirotko; the support and patient guidance of Kjeld Schmidt; or the thoughtful consideration of all the volunteer reviewers. Stuart Geiger, Jessica Beth Polk and Matthew Bietz also provided valuable assistance with this Introduction.

Note

1. Studies of infrastructure stretch much further back, in particular with a lineage emerging from the study of large sociotechnical systems Coutard, O. (ed) (1999) *The Governance of Large Technical Systems* (London/New York: Routledge), Hughes, Thomas P. (1983) *Networks of Power: electrification in Western society, 1880–1930* (Baltimore: John Hopkins University Press), Hughes, Thomas P. (1989) ‘The Evolution of Large Technological Systems’, in Wiebe E. Bijker, Thomas P. Hughes and Trevor J. Pinch (eds), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, Massachusetts and London, England: The MIT Press): 51–82. However what Star gave us was a broadening of infrastructure beyond its technological meaning. Following from Becker’s work on Art Worlds Becker, Howard S. (1982) *Art Worlds* (Berkeley, CA: University of California Press) that explored all the professions and craft-workers tied to art production and dissemination, Star drew attention to the enabling resources and people that often remained invisible or unaccounted in many heterogeneous settings such as nursing, librarianship and technical labor.

References

- Abbate, J. (1999). *Inventing the internet*. Cambridge: MIT.
- Atkins, D. E. (Chair) (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*. Washington, DC: National Science Foundation.
- Beaulieu, A. (2010). From co-location to co-presence: shifts in the use of ethnography for the study of knowledge. *Social Studies of Science*, 40(3), 453–470.
- Becker, H. S. (1982). *Art worlds*. Berkeley: University of California Press.
- Bietz, M. J., & Lee, C. P. (2009). Collaboration in metagenomics: Sequence databases and the organization of scientific work. In E. Balka, L. Cioffi, C. Simone, H. Tellioglu, & I. Wagner (Eds.), *ECSCW 2009: Proceedings of the 11th European Conference on Computer Supported Cooperative Work*. London: Springer-Verlag.
- Bietz, M. J., Baumer, E. P. S., & Lee, C. P. (2010). Synergizing in cyberinfrastructure development. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9114-y.

- Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the internet*. Cambridge: MIT.
- Bowker, G. C. (1994). *Science on the run: Information management and industrial geophysics at Schlumberger, 1920–1940*. Cambridge: The MIT.
- Bowker, G. C. (2006). *Memory practices in the sciences*. Cambridge: MIT.
- Bowker, G. C., Star, S. L., Turner, W., & Gasser, L. (Eds.). (1997). *Social science, technical systems, and cooperative work: Beyond the great divide*. Mahwah: Lawrence Erlbaum Associates.
- Brand, S. (1999). *The clock of the long now: Time and responsibility*. New York: Basic Books.
- Collins, H. M. (1981). Son of seven sexes: the social destruction of a physical phenomenon. Knowledge and controversy: studies of modern natural science. *Special Issue of Social Studies of Science*, 11(1), 33–62.
- Coutard, O. (Ed.). (1999). *The governance of large technical systems*. London: Routledge.
- David, P., & Bunn, A. (1988). The economics of gateway technologies and network evolution: lessons from electricity supply history. *Information Economics and Policy*, 3, 165–202.
- de la Flor, G., Jirotko, M., Luff, M. P., Pybus, J., & Kirkham, R. (2010). Transforming scholarly practice: embedding technological interventions to support the collaborative analysis of ancient texts. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9111-1.
- Edwards, P. N. (1996). *The closed world: Computers and the politics of discourse in cold war America*. Cambridge: MIT.
- Edwards, P. N., Bowker, G. C., Jackson, S. J., & Williams, R. (2009). Introduction: an agenda for infrastructure studies. *Special Issue on e-Infrastructure of the Journal of the Association for Information Systems*, 10(5), 364–374.
- Egyedi, T. (2001). Infrastructure flexibility created by standardized gateways: the cases of Xml and the Iso container. *Knowledge, Technology & Policy*, 14(3), 41–54.
- Faniel, I. M., & Jacobsen, T. E. (2010). Reusing scientific data: how earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9117-8.
- Fujimura, J., & Chou, D. Y. (1994). Dissent in science: styles of scientific practice and the controversy over the cause of AIDS. *Social Science and Medicine*, 38, 1017–1036.
- Galison, P. L., & Hevly, B. (Eds.). (1992). *Big science: The growth of large-scale research*. Stanford: Stanford University Press.
- Gillespie, T. (2006). Engineering a principle: 'End-to-End' in the design of the internet. *Social Studies of Science*, 36(3), 427.
- Grudin, J. (1994). Computer-supported cooperative work: history and focus. *Computer*, 27(5), 19–26.
- Haigh, T. (2001). Inventing information systems: the systems men and the computer, 1950–1968. *The Business History Review*, 75(1), 15–61.
- Hine, C. (2000). *Virtual ethnography*. London, England, Thousand Oaks, California, and New Delhi, India: Sage Publications Ltd.
- Hine, C. (2006). *New infrastructures for knowledge production: Understanding e-science*. Hershey: Information Science Pub.
- Hughes, T. P. (1983). *Networks of power: Electrification in Western society, 1880–1930*. Baltimore: John Hopkins University Press.
- Hughes, T. P. (1989). The 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: The MIT.
- Jankowski, N. W. (2007). Exploring e-Science: an introduction. *Special theme: E-science, in the Journal of Computer-Mediated Communication*, 12(2), 549–562.
- Jankowski, N. W. (Ed.). (2009). *E-Research: Transformation in scholarly practice*. New York: Routledge.

- Jirotko, M., Procter, R., Rodden, T., & Bowker, G. C. (Eds.) (2006). Special issue: collaboration in e-research. *Computer Supported Cooperative Work (JCSCW)*, 15(4), 251–255.
- Karasti, H., Baker, K. S., & Millerand, F. (2010). Infrastructure time: long-term matters in collaborative development. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9113-z.
- Kee, K. F., & Browning, L. D. (2010). The dialectical tensions in the funding infrastructure of cyberinfrastructure. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9116-9.
- Lawrence, K. A. (2006). Walking the tightrope: the balancing acts of a large e-research project. *Computer Supported Cooperative Work*, 15(4), 385–411.
- Lee, C. P., Dourish, P., & Mark, G. (2009). The human infrastructure of cyberinfrastructure. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 483–492). New York: ACM Press.
- Millerand, F., & Bowker, G. C. (2009). Metadata standards: Trajectories and enactment in the life of an ontology. In M. Lampland & S. L. Star (Eds.), *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life* (pp. 149–167). New York: Cornell University Press.
- Monteiro, M. (2010). Reconfiguring evidence: interacting with digital objects in scientific practice. *Computer Supported Cooperative Work (JCSCW)*. doi:10.1007/s10606-010-9115-x.
- Nentwich, M. (2003). *Cyberscience: Research in the age of the internet*. Vienna: Austrian Academy of Science.
- Olson, G. M., Zimmerman, A., & Bos, N. (Eds.). (2008). *Scientific collaboration on the internet*. Cambridge: MIT.
- Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artifacts: or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, 14(1), 399–441.
- Price, D. J. (1963). *Little science, big science*. New York: Columbia University Press.
- Ribes, D., & Bowker, G. C. (2008). Organizing for multidisciplinary collaboration: The case of the geosciences network. In G. M. Olson, A. Zimmerman, & N. Bos (Eds.), *Scientific collaboration on the internet* (pp. 311–330). Cambridge: MIT.
- Ribes, D., & Bowker, G. C. (2009). Between meaning and machine: learning to represent the knowledge of communities. *Information and Organization*, 19(4), 199–217.
- Ribes, D., & Finholt, T. A. (2007). Tensions across the scales: Planning infrastructure for the long-term. *Proceedings of the 2007 international ACM conference on Supporting group work, Sanibel Island, Florida, USA* (pp. 229–238). New York: ACM Press.
- Ribes, D., & Finholt, T. A. (2009). The long now of infrastructure: Articulating tensions in development. In P. Edwards, G. C. Bowker, S. Jackson, & R. Williams (Eds.), *Special Issue on eInfrastructure in the Journal of the Association for Information Systems*, 10(5), 375–398.
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43, 377–391.
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, “translations,” and boundary objects: amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, 19, 387–420.
- Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: design and access for large information systems. *Information Systems Research*, 7(1), 111–134.