CSCW 2008 Workshop Position Paper
"Designing Cyberinfrastructure for Science"

Mediation Roles, Local Scientific Data Communities, and Bottom-Up Infrastructure

Karen S. Baker
Scripps Institution of Oceanography
University of California, San Diego, US
kbaker@ucsd.edu

Helena Karasti
Dept of Information Processing Science
University of Oulu, Finland
Helena.Karasti@oulu.fi

Florence Millerand
Department of Social & Public Communication
University of Quebec at Montreal (UQAM), Canada
millerand.florence@uqam.ca

ABSTRACT
New roles are emerging to address new types of work associated with both local information infrastructure and cyberinfrastructure. Mediation – or work across boundaries of established arenas – is one of these types of work. We argue for the importance of recognizing and facilitating mediation roles in local-scale as well as global-scale infrastructure in order to ensure broad participation and continuing development in terms of infrastructure definition, design, and maintenance. We consider two common metaphors - building infrastructure (top-down) and growing infrastructure (bottom-up) – that together reveal and inform infrastructure development. The capacity to evolve and interface across multiple components is central to the work of advancing infrastructure.

INTRODUCTION
Information infrastructure is often referred to in the singular. The expression 'cyberinfrastructure' contributes to the notion of a singular entity when in practice there are a myriad of facets and components to infrastructure (e.g. Freeman 2007). Advancing infrastructure - the local and national, digital and human, historical and contemporary - comes from recognizing and acknowledging the multiples: multiple arenas of work, multiple facets in organizing understandings of interdependent aspects as well as multiple spatial and organizational scales associated with infrastructure. Considering cyberinfrastructure together with local infrastructure stimulates development of under-studied topics pertinent to development of infrastructure. Ethnographic work begins to reveal the magnitude of work and its distribution across different roles that is required for growing local infrastructures and creating interfaces between local-scale and larger or global-scale structures.

Position Paper

Workshop on Designing Cyberinfrastructure to Support Science

ACM Conference on Computer Supported Cooperative Work

Saturday, November 8, San Diego, CA

Understanding infrastructure requires crossing levels, categories, and scales which spurs development of much needed conceptual models and descriptive frameworks. Mediation roles ensure articulation of infrastructure issues and continuing dialogue across interfaces. These roles are key to developing language that informs ongoing efforts to infrastructure, revealing otherwise unrealized options and unarticulated choices.

We're all familiar with local socio-technical infrastructure and associated mediation roles – the electronics store clerk in a nearby shopping center and the colleague who suggests setting-up skype – and with digital infrastructure that blurs the line between local and global as storage devices preserving our files sit locally on our desks as well as within a community repository and/or remotely as Google docs. To move from these particular understandings of the digital realm to that of 'advanced' infrastructure, new approaches are needed that facilitate and promote articulation, dialogue and negotiation at planning meetings, at design tables, and at the boundaries of established arenas. New questions need asking: What and who is involved in designing infrastructure? Is infrastructure envisioned as an automated resource? Are new types or roles of mediation accounted for? Just how much standardization is needed? And what is envisioned for local infrastructures in relation to larger-scale cyberinfrastructures, and vice versa?

ARTICULATION WORK
Articulation work is a concept important to community-building and to long-term information infrastructure. Strauss (1988) devised the terms 'articulation work' and 'articulation process' in understanding how to describe the work within a project. The notion of articulation work and processes applies to technical and organizational arrangements but also to coordination of cooperative work involving interdependencies. Shapin's "Invisible Technician" (1989) emphasizes making visible activities that are not conceived as work but that still require skill and knowledge to perform. Such invisible work is associated with many roles emerging in association with creating, maintaining, and redesigning infrastructure.

In our work with an ecological scientific community that focuses on long-term data, advancements in the digital realm are welcomed. Yet even with an organizational structure that ensures technological mediation via the role of a local information manager on site, the need for dialogue, design strategies, and understandings of long-term ramifications associated with new technologies are under appreciated (Karasti and Baker, 2004). As an information manager juggles support for real-time analysis, data repository development, and technology integration, there is a need to consider the notion of information infrastructure as having multiple dimensions (Lee, Dourish et al, 2006; Baker and Millerand, 2007b; Ribes and Finholt, 2007). A strategy we have found useful in formulating issues and prompting dialogue is the use of multi-category frameworks that traverse traditional boundaries, for example to move from the perspective of infrastructure as 'technical', we have worked with the two category system of 'social' and 'technical' in addition to introducing a three category set - 'technical, social, and organizational' - or alternatively a pair of axes: local-global and social-technical. Once familiar with one framework, the ability to move between differing descriptive frameworks enhances the ability to understand and discuss digital infrastructure.

Aspects of Infrastructure

The three facets of infrastructure - technical, organizational and social –comprising one descriptive framework are interdependent. The value of these elements as a set becomes evident when one considers that in the realm of cyberinfrastructure, the term 'advanced' typically brings to mind first and foremost visions of the technical – computer platforms and digital arrays. So it is useful to have at hand a set of categories that prompts one to recall the multiple facets of 'advanced' including the organizational and the social in order that they are present when planning technological advancements.

To open up the work of infrastructure for data collecting scientific programs, we strive to develop understandings of and sensitivities to categories and vocabularies for community data tasks that constitute an information infrastructure. This infrastructure is a shifting blend of configurations of technology, organization and social. The interdependence of all three is inherent to work with data and involves the work of design, development, deployment and enactment (Millerand and Baker, 2005).

Within any one day a scientist may work with a new collaborative software application for communicating with a colleague, discuss the representation of data within an information system, and create a presentation about some recently collected field data. Creating an infrastructure that meets the heterogeneity of issues faced by such an individual is a challenge. But conceptualizing advanced organizational and social facets of infrastructure augments technical aspects that may be close at hand. For instance, in terms of organizational models, institutional and city library systems with their archives and public help desks provide an organizationally advanced example from another domain of both storage and mediation layers. In terms of "the social", a scientific network such as the Long-Term Ecological Research (LTER) that has grown from an initial six sites to twenty-six over a period of twenty-five years can provide insight into issues of sustainability and trust. In recent years LTER formalization of data sharing policies, governance practices, and a metadata standard in conjunction with collaborative data practices has 'advanced' our understanding of community infrastructure.

Mediation and Local Information Management

Mediation exists to facilitate the use of data by LTER scientists (Baker and Millerand, 2007a). In many communities, arrangements are made for technologists to work with scientists. Within the LTER, there exists another type of mediation carried out through the information management role established at each local site. From the domain science perspective, mediation work is required to interface at-sea with on-land data management as well as to create new combinations of datasets, data models, and information systems. From the information managers' perspective, mediation work is required to elaborate data processes, to articulate communication issues as well as options and constraints associated with technology in general and data organization in particular. Within LTER, the role of data management was renamed information management (Baker et al, 2000) in recognition of a growing complexity of work; the dialogue continues today in terms of considering roles associated not only with data and information management but also with informatics. As mediator, the information manager prompts dialogue, coordinates data activities, and elaborates on categories as well as on technical and organizational data issues. The role of information manager may be viewed as one part of a site's infrastructure, of 'growing one's own information infrastructure' (Karasti and Baker, 2008).

With the notion of re-using data, however, there arises the need to consider data exchange outside a particular community or arena. Today local information infrastructures are expected to join up or network with global or cyber infrastructures. When data flows between differing work arenas, however, there are new interfaces or boundaries that frequently require mediation. That is, the bottom-up local approach is required to meet the top-down as is described in standards' work by Millerand and Bowker (in press). In different circumstances, this same theme of meeting up in the middle is developed by Twidale and Floyd (2008).

CONCLUDING REMARKS

We distinguish local and global infrastructure arenas, highlighting the need for articulation work in both and attention to the joining of the two. The challenges of growing local infrastructures rivals the challenges faced in building global or cyberinfrastructure. Mediation work –

often invisible or unnoticed in everyday activities – involves the interrelation and alignment of components or parts. It is important that the infrastructure discourse reach across all facets and scales of infrastructure so that all are aware of the profound differences in possibilities and all are poised contribute to design and development of infrastructure. We argue for more efforts to investigate mediation roles in the context of local or site-based work and information management (bottom-up) as well as for their interface with global data center curators and developers (top-down). Recognition and involvement of diverse mediation roles will contribute to the understanding of what is at stake in development of infrastructure. How and with whom infrastructure is designed may be in the long-term equally as important as what is built and how it is used.

REFERENCES

1.Baker, K. S., B. J. Benson, et al. (2000). Evolution of a multisite network information system: the LTER information management paradigm. BioScience 50(11): 963–978.

2.Baker, K.S. and Millerand, F. (2007a). Articulation work supporting information infrastructure design: coordination, categorization, and assessment in practice. Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07), January 03-06, 2007, IEEE Computer Society, Washington, DC, 242a.

3.Baker, K. S. and F. Millerand (2007b). Scientific information infrastructure design: information environments and knowledge provinces. Proceedings of the American Society for Information Science and Technology (ASIST 2007).

4.Freeman, P. A. (2007). Is 'Designing' Cyberinfrastructure - or, Even, Defining It - Possible? First Monday **12**(6) (June 2007): URL: http://firstmonday.org/issues/issue12_6/freeman/index.html

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

6.Karasti, H. and K. Baker (2008). Community design – growing one's own information infrastructure. Participatory Design Conference, Bloomington, Illinois.

7.Lee, C.P., Dourish, P., and Mark, G. (2006). The human infrastructure of cyberinfrastructure. Proceedings of the Computer Supported Cooperative Work, November 4-8,Banff, Alberta, Canada.

8.Millerand, F., K.S.Baker, et al. (2005). Lessons learned from EML about the community process of standard implementation. LTER Databits Newsletter (Fall): http://intranet.lternet.edu/archives/documents/Newsletters/DataBits/05fall/#4fa.

9.Millerand, F., and Bowker, G.C. (in press). Metadata standard: trajectories and enactment in the life of an ontology. In Formalizing Practices: Reckoning with Standards, Numbers and Models in Science and Everyday Life. M.Lampland and S.L.Star (Eds), Cornell University Press.

10.Ribes, D. and Finholt, T.A. (2007). Tensions across the scales: planning infrastructure for the long-term. Proceedings of the 2007 international ACM SIGGROUP Conference on Supporting Group Work, Sanibel Island, Florida, USA.

11.Shapin, S. (1989). The invisible technician. American Scientist 77: 554-563.

12.Strauss, A. (1988). The articulation of project work: an organizational process. The Sociological Quarterly, 29(2), 163-178.

13.Twidale, M.B. and I.Floyd, 2008. Infrastructures from the bottom-up and the top-down: can they meed in the middle? Participatory Design Conference, Bloomington, Illinois.