Representing Community: Knowing Users in the Face of Changing Constituencies

David Ribes
School of Information, University of Michigan
1075 Beal Avenue, Ann Arbor, MI 48109-2112
dribes@umich.edu

Thomas A. Finholt
School of Information, University of Michigan
1075 Beal Avenue, Ann Arbor, MI 48109-2112
finholt@umich.edu

ABSTRACT
This paper traces the use of the concept 'community' by drawing attention to the ways in which it serves as an organizing principle within systems development. The data come from an ethnographic study of participants and their activities in the Water and Environmental Research Systems Network (WATERS). WATERS is a US National Science Foundation-funded observatory and cyberinfrastructure project intended to serve the heterogeneous scientific disciplines studying the water environment. We identify four vehicles by which WATERS participants sought to know the needs, conflicts and goals of their diverse communities: engaging in vernacular discussions; organizing community forums; implementing surveys; and requirements gathering. The paper concludes that the use of community in IT development projects is substantially divorced from its traditional meanings which emphasize collective moral orientations or shared affective ties; instead, within systems development, community has a closer meaning to a 'political constituency,' and is used as a short-hand for issues of inquiry, representation, inclusion and mandate.

ACM Classification Keywords
K.4.3 Organizational Impacts: Computer-supported cooperative work J.4 Social and Behavioral Sciences: Sociology

AUTHOR KEYWORDS
Community, Ethnography, Cyberinfrastructure, Infrastructure, Science, eScience, Studies of user-centered design and requirements engineering

1. INTRODUCTION
Infrastructure is intended to provide stable, reliable and accessible resources, supporting the work activities of a *community*. This intuitive formulation has become a staple in studies of infrastructure[19], but this definition also hides enormous complexities, stabilizing the meaning of precisely those elements which are fluid and often contentious. In this paper we focus in particular on *participants' work in defining community* – that is, the users of an infrastructure – and ask how these activities come to inform a trajectory of development.

CSCW'08, November 8–12, 2008, San Diego, California, USA.

To explore this question empirically, we focused on the planning process for the Water and Environmental Research Systems Network (WATERS), a US National Science Foundation (NSF) project intended to serve the heterogeneous scientific disciplines studying the water environment. WATERS is a cyberinfrastructure (CI) in that it seeks to provide access to data, services and computing resources to a broad group of scientists, educators and policy advisors. It is the identity of those groups, and what kinds of resources to provide, that are *at stake* in defining the WATERS community. Thus, our goal was to reveal the problematization of 'community' by participants themselves. Rather than taking it as a given, as obvious, or even a pre-existing entity, *we explore how 'community' is debated, researched and ultimately constituted by representatives who seek to mobilize its identity as they go about the work of planning technological development trajectories.*

Mirroring our actors' categories we shall use the term community to refer to the general body of the scientific domain or discipline. Communities are amorphous and abstract entities that come to be known through inquiry in the form of, for example, surveys or community forums. In turn, through their representatives, communities come to shape the technologies, services and data which will be made available through information infrastructure.

We focus on a particular event in the planning of the WATERS Network in which its target communities were substantially overturned. The WATERS Network began in 2005 as CLEANER: Collaborative Large-scale Engineering Analysis Network for Environmental Research. CLEANER's remote sensors and IT resources were to be directed primarily at the environmental engineering community. However, goals for data integration and cost-sharing led to an alliance between environmental engineers and hydrological scientists. In 2006 the Consortium for Advancement of Hydrologic Science (CUAHSI) was added to the CLEANER team and the project was renamed WATERS to reflect its plans to serve both communities. There is a blurry boundary between these groups of researchers: nominally both are engaged in the study of the water environment but often they do so with diverging purposes. Put simply, environmental engineers tend to identify themselves as conducting research with an orientation to application (e.g., mitigation of human impact on water quality), while scientists tend to emphasize contributions to a base of knowledge (e.g., influences on the flow and course of rivers). More complexly, *the boundary between the scientific and engineering communities stands-in for debates over the kind of data that will be collected, what research the infrastructure will support, and what organizations will receive funding awards*.

In order to understand the event which led to the creation of WATERS we must delve into the scientific and technical ambitions of the project and we must trace the methods by which participants come to know their communities. This event became an occasion for us to investigate 'the purpose of infrastructure': it rendered visible to the ethnographic researcher the work of participants as they scrambled to redefine their communities, scientific research agendas and the technical development trajectories of a major research equipment and facilities project.

At a broad level, this is a study of 'user-centered design in action' and of how participants give voice to a community in the development of technical systems. We demonstrate that in a large-scale technology development project, its 'purpose' is a moving target. We focus in particular on how the constitution of a 'community' is significantly at play in defining its purpose. In development, 'community' serves a similar role to that of 'constituency' in the political sphere: it identifies a target population in order to conduct inquiries on the interests, goals and purposes of future users; and in turn, representatives articulate this knowledge in the planning process in the hopes of shaping a technological trajectory of development.

2. THE USE OF COMMUNITY

In an analysis of infrastructure development we must be cautious not to hypostatize 'community,' since in many senses this entity is both the object and outcome of the endeavor. Participants' use of the term community within systems design has a distinctly different orientation than those debates endogenous to social science. In building sensing and information systems for WATERS the term community has had at least two general uses: i) it has referred to an *already existing* but amorphous body of the domain, but ii) has also acted to identify a *future body* of users who will be linked by multidisciplinary ties and the networked resources of the infrastructure. The two meanings are used somewhat interchangeably and reflect the mandates and purposes of the WATERS project. WATERS seeks to serve hydrologists and environmental engineers, but it also seeks to develop a technical platform enabling collaboration and to form an organizational entity to coordinate across these fields. The infrastructure will support existing scientific and engineering activities, but it is also intended to transform those activities, inaugurating novel interdisciplinary relations – thus, it is a 'community building endeavor'.

The use of community has four primary components comparable to constituency: (i) identifying and circumscribing a collective on which to conduct *inquiry* (ii) so as to produce a form of authoritative (political) *representation*, thus; (iii) generating *accountability by* demonstrating the relevance of the future resources or services back to that collective, and; (iv) to show a *mandate* to funding institutions in order to secure funds for building and sustaining those resources and services.

The categories 'hydrologists' and 'environmental engineers' include thousands of academic and industrial researchers distributed in universities and research institutes across the US, in departments as varied as geology to resource management. The totality of American environmental engineering researchers and hydrologists who could benefit from or use WATERS (its constituency) are never present at a planning meeting for WATERS. Rather, it is the representatives of that community which speak in its name. Michel Callon [4] has described such representatives, or spokespersons, as "punctualized" organizational actors: individual participants who collectively stand-in for the whole. These actors always appear specifically, for example, in our cases as principal investigators, committee members, executive directors of institutions or irate scientists. We can think of that small selection of participants in WATERS who come together to build the Network as a sample of the larger hydrology and environmental engineering community. *The work of 'knowing the community' is partially that of sustaining those relationships which ensure that this sample remains representative of the larger body.* In this sense, the use of the term 'community' within systems development comes closer to that of 'constituency' in the political sphere. Like constituencies, the communities of science have their demographics compiled, are polled for opinion and surveyed to identify their needs. In the planning and design meetings of WATERS, the interests of communities are represented by (usually) senior scientists, who speak in their name. In turn these community representatives describe a certain responsibility to their constituencies; they bring community issues 'to the table', and feel accountable back to them.

We took as our research object *participants' work* as they went about knowing, defining and demarcating their communities. Just 'what the community wants' can come to be a significant resource as representatives debate amongst themselves future technological development trajectories, funding allocations, and divisions of labor. Put briefly, we treat 'community' as an actor's category the content of which is at stake in the process of infrastructure development.

2.1 The traditions of community: affective and moral or practical outcome?

In the social sciences 'community' is a notoriously prickly concept. For example, historian Thomas Bender has traced a one-hundred year debate in sociology over the classical formulation of 'Gemeinschaft and Gessellschaft' [2]. In 1887 the German scholar Ferdinand Tönnies published his now canonical piece defining Gemeinschaft as human associations characterized by affective ties and shared moral orientations, and Gessellschaft as a society characterized by self-interest and a high division of labor. Bender argues that an irresolvable debate has emerged over the relationship between the two: Is traditional community declining in the face of modern anonymous society or is community continuously reinvented at the interstices of society's endlessly evolving structure?

IT research is by no means immune to this debate. These two interpretations of Tönnies' thesis (decline of community vs. its reinvention) have trickled into research on digital and virtual community. Analysts describe the fall of civil society in the face of the internet or see its reformulation in blogs and social networks In many senses the debate over the status of community has become intractable; for the purposes of this paper, we would like to turn our attention to a slightly different phenomenon: *the use of community*.

More specifically, we propose to leave the definition of community to the field. How have the interests, goals and wishes of a community come to be problematized for participants in infrastructure building ventures?

It is difficult to reconcile the use of ‘community’ in WATERS with traditional social science understandings where community is “a network of social relations marked by mutuality and emotional bonds,” in which a “sense of self and community may be hard to distinguish” and where “a community is an end in itself” [2:8p]. 'Water researchers' only rarely know each other directly, most often do not interact at all, and are usually tied only by loosely framed interests or through common organizational and funding agencies; yet 'community' is regularly evoked by participants in the WATERS Network. Community is an actor's category; moreover, it is a keyword and organizing principle in the development of information infrastructure, often discussed, debated and deployed.

In this sense we are following an alternate scholarly tradition to those described above which take it as the analyst's prerogative to evaluate the rise and decline of community. Instead, we draw on a repertoire of thought which takes community as an *outcome* of participants' activities; we will briefly encapsulate this approach through the concept of 'boundary work.’

In studies of boundary work the object of study becomes the activities of actors as they go about defining and defending demarcations of who is or is not a member of a group. Sociologist of science Thomas Gieryn defines boundary work as the “attribution of selected characteristics to the institution of science (i.e., to its practitioners, methods, stock of knowledge, values and work organization) for purposes of constructing a social boundary” [6:782p] We focus on the boundary work around the 'hydrology and environmental engineering communities' as participants seek to define each disciplinary field, their diverging methods, goals, and data. In WATERS, boundaries are defined in order to understand disciplinary difference, but then also in order to find commonalities and develop a single 'umbrella infrastructure' [16]. In system development, part of the task is to identify the heterogeneity of users' needs and requirements. This is analogous to knowing and representing the interests of a constituency in the political sphere. The diverging interests of hydrologists and environmental engineers must first be identified before they can be reconciled through the construction of a single infrastructure.

2.2 A study of user-centered design in action

In this paper we do not use ethnography to inform design, as a form of requirements elicitation or usability testing [8]. Rather, we turn the ethnographic eye upon the work of elicitation and testing itself [27]: how is community known? User-centered design techniques are intended to transform how work is done while still respecting existing practices – technology products are informed by the present but seek to inaugurate novel futures:

> The project of constructing a prototype on this view affords possibilities for respecifying a relevant form of work in and through the act of building a new artefact; one that simultaneously reconfigures the work’s practice while maintaining its accountability to relevant professional and organizational constituencies. [22:166p]

It seems high time for such an analysis, as it promises to help us disentangle the use of community within information systems design from its traditional moorings which emphasize culture, affect or moral orientation.

We do not take 'the concept' or 'meaning' of community as the seminal site for investigation. Rather, community's importance emerges as it comes to be a category of action in techniques of inquiry, deployed in surveys and then summarized in reports. We seek to respecify these central epistemological efforts as "commonplace discursive and practical activities," [11:5] and demonstrate how they come to inform planning and technological design.

As we will see the common or distinct 'scientific practices' and 'needs' of environmental engineers and hydrologists are canvassed in vernacular conversation, community workshops, surveys and requirements gathering. The results of such inquiry outline the contours and boundaries of what come to be the 'heterogeneous communities' of environmental engineering and hydrology. That data is then mobilized by representatives in the WATERS planning process as they design a single 'umbrella infrastructure' [16] for water environment researchers. Thus our question is not ‘are communities the relevant unit of analysis for scientific activity?’ but rather, *how do those practices and techniques which make it known come to render community an actor [3] in the development of information infrastructure?*

3. STUDY RATIONALE AND CASE

Communities of the scale and distribution of those studied in this paper are not directly observable entities for the ethnographer. Even at the largest conferences such as the Geological Society of America (GSA) 'environmental engineers and hydrologists' cannot be known by participant observation. Well targeted interviews can reveal a great deal, but the vast majority of people remain an undifferentiated mass of streaming conference participants. In this sense, the ethnographic observer is no more privileged than any other participant in how they know 'community'.

Our study rationale is to follow the full range of participants' activities – whether they use ‘vernacular’ forms of knowing such as conversations and community forums, or more ‘formal’ means such as surveys – in order to trace the process of how participants go about knowing community. We use the term participant to broadly designate those involved directly in the planning, design, implementation of WATERS. We use the term representative to designate those participants who seek to stand-in for a community.

Our main field sites have been the meetings of principal investigators (PIs), designers and implementers of WATERS. These are often senior scientists in their fields and leaders in computer science and IT development. Such meetings are excellent sites in which participants regularly discuss the evaluation, strategizing and planning of their projects. Because WATERS is distributed and often organized at a distance, field research has been highly multimodal, conducted, for example, through video-based communication (e.g., Access Grid) and conference phone calls. We supplemented our ethnographic work with semi-structured interviews. In addition we have been granted access to portions of the WATERS email listservs. Because these projects are geographically distributed, email discussions often include everyday planning and decision making, providing a continuous stream of rich data.

Ethnography does not only describe our method (participant observation), it also describes our goals: we seek to capture the orientation of participants as shaped through their own activities. In particular we focus on the constitution of WATERS' communities: What is the work that goes into knowing community, and how do the results of such work inform planning and funding decisions? These activities are the domain of the surveyors and requirements engineers, and so in this paper we focus in particular on the activities of these participants, as well as their instruments (questionnaires) and results (reports). Thus, our method is closer to the theoretical sampling of grounded theory than the thick description of cultural anthropology [20]. We focus on our actors' actions and work, around their usage of the term 'community'.

Finally, in each project we have also been *participants,* contributing to aspects of planning, proposal writing, social dimensions feedback, user-studies or requirements elicitation [14].

3.1 The WATERS Network

The WATERS Network is an ambitious and complex endeavour involving the participation of hundreds of scientists, engineers, administrators and technical staff. It is geographically distributed across the US and its timeframe for planning and construction stretches to decades. The most optimistic participants do not anticipate that the actual construction of WATERS will begin until 2012, with an expected award of hundreds of millions of dollars. But today twelve planning grants have been awarded and completed, a project office has been established, and eleven testbed sites are in operation, totalling an approximate investment of three million dollars.

WATERS will include remote automated sensor networks; tools for archiving, sharing and analyzing data; and support for multidisciplinary and distanced collaboration: "[T]he WATERS Network plans to deploy an integrated and distributed system of environmental observatories at sites across the country." [12:6642p]. In addition to hydrologists and environmental engineers, the Network will also provide services to social scientists, geochemists, geomorphologists, educators and policy advisors[1]. The larger goal is to transform the scope and topics of scientific research. To do so WATERS intervenes at the level of data collection, modelling and computational resources:

> Research at these sites would be aided by tools for collection, storage, and dissemination of environmental data; interactive models that could be tested in real or near-real time; and an integrative cyberenvironment that would help multidisciplinary, geographically dispersed teams of researchers work together effectively. [12:6645p].

The Network is an infrastructure in the sense that it is intended to support 'community research' in a general fashion, rather than serving to answer a single set of scientific questions or a particular team.

WATERS is seeking to secure funds from NSF's Major Research Equipment and Facilities (MREFC) account. This is an agency-wide special account set up to pay for the acquisition, construction, and commissioning of major scientific infrastructure and equipment.

As the WATERS Network is described on its homepage:

> The WATERS Network is actually a combination of two national environmental observatory planning initiatives: CLEANER […] which has been supported by NSF's Engineering Directorate and the CUAHSI […] initiative for Hydrologic Observatories, which has been supported by NSF's Geosciences Directorate. [26]

It is beyond the scope of this paper to fully explore why CLEANER and CUAHSI were *brokered* into a single project, however, participants and NSF officers generally account for the merger through four explanations: i) that both environmental engineers and hydrologists are studying the water environment, ii) that CI should encourage multidisciplinary collaboration, ii) that CI should facilitate data and resource interoperability, and as such a single infrastructure would facilitate the creation of shared metadata standards and; iv) that the significant investment necessary to develop observatories and CI, demands that projects be leveraged. We title the drive to broker CI projects 'the logic of leveraging.'

The larger phenomena of brokering within CI, and the logic of leveraging which largely informs it, deserves fuller attention and will be the topic of our future research. However, for the purposes of this paper we focus on the result of that brokering – the flurry of activity which ensued to reshape CLEANER's environmental engineering endeavour by incorporating the 'interests, goals and needs' of the hydrological science community. As an NSF officer described the logic of the CLEANER/CUAHSI merger: "When it comes to environmental observatory construction, particularly at this scale [MREFC], and particularly at this level of investment and complexity, two is too many," (NSF, interview).

4. INQUIRY ON COMMUNITY

The recent past of computer science is dotted with acclaimed and then abandoned high-end computing projects and applications. From such endeavors, those people designing and implementing new systems have learned a great deal about 'best practices' for planning. For example, the Atkins Report [1] (a programmatic CI document based on the work of a blue ribbon panel convened by NSF) collects insights from assessments and scholarly reviews of these programs, as well as from sociological research in informatics documenting both successful and fatally flawed results. Many of these studies emphasize the importance of communication, coordination and the difficulties of distanced collaboration; these findings have trickled into the policy make-up of CI and observatory

[1] While there are multiple constituencies which WATERS seeks to serve, in this paper we focus on the 'event' which brought together hydrology and environmental engineering. CUAHSI/hydrology was the first constituency to be 'leveraged' into the initially environmental engineering endeavor.

development. For instance, the Atkins Report tackles the challenge by encouraging new management structures and implementation strategies in collaboration with social scientists:

> Much of the effort under way to use cyberinfrastructure for collaborative research is not giving adequate attention to sociological and cultural barriers to technology adoption that may cause failure, even after large investments. [1:13p]

These ‘social’ barriers are to be understood as mismatches between technical development and existing routines and conventions of scientific work within the community. Within CI a general consensus has emerged that the solution is to be found in user-centered design practices or requirements engineering: iterative planning and development which is closely informed by the system's intended user base.

The number of methods and approaches to user centered design is quite large, they range from ethnography, surveys and participatory design to cultural probes and user-modeling [18]. Overall, the goal of these methodologies is to lead the development of 'domain-oriented systems' [5], tailored to specific kinds of users or tasks while supplying new tools. While most of the participants in WATERS are not trained in such methods, 'knowing community' and having that knowledge inform – even 'drive' – design has come to be a well accepted tenet of contemporary scientific information infrastructure development. Today, social and information scientists, their methods and their conceptualizations are *always already* infused into large-scale technology ventures [15, 17]. In this section we will focus on four mechanisms for knowing community as enacted within WATERS: vernacular engagements, community forums, user surveys, and requirements elicitation.

Participants' primary orientation towards the merger between CLEANER and CUAHSI was an anxiety about how to develop resources relevant to their respective communities. While the merge presented logistical and organizational problems, the overarching concern for participants has been 'how should an infrastructure serve both groups?' As one environmental engineer characterized the sentiment:

> the problem is […] that instead of just satisfying the environmental engineering community […] we now have to also try to meet the needs of the hydrologic research community, and they can be a little bit tough. […] Our community ( ) is more problem-focused, and the hydrologic community tends to look at things [as] more scientific and research based; on the things that they’re interested in as opposed to “here’s the problem. I want to try to find a solution to it”.

Speaking in sweeping terms, the merger of CLEANER and CUAHSI is about 'engineering vs. scientific' research; however, this phrasing is merely a short-hand for a myriad of concerns which when inspected more carefully also reveal substantial historical, organizational and technical content. Participants in WATERS themselves make a point to distinguish disciplinary histories of environmental engineering and hydrology; they note differing scientific norms and traditions, varying 'science questions' and demarcate criteria for the kind and quality of data that will be necessary to answer those questions.

In order to understand the boundary 'science/engineering' and how it may affect the development of the new WATERS collaboration, participants have embarked on multiple ventures to 'know their communities'. They do so in order to use that knowledge to inform planning and design, but also to demonstrate *accountability* to that community. That is, to show that the WATERS Network will reflect the needs and research goals of hydrologists and environmental engineers.

4.1 Vernacular Discussions

Vernacular methods of inquiry are usually excluded from formal discussions of requirements engineering; however, we found that participants often emphasized that their sense of 'community opinion' was substantially shaped through i) everyday ('watercooler') conversations with colleagues, ii) in question and answer periods following presentations at conferences and iii) through face-to-face meetings specifically organized to test new ideas with resistant or vocal representatives ('canaries in the coal mine'). Casual conversations can later serve to inform the development of more formal techniques of inquiry, i.e., vernacular impressions require the force of survey findings to gain rhetorical force. For example, during an interview, in asking a respondent ('A') how she knew what the hydrological community wanted from WATERS, she pointed to vernacular means for knowing, and then when pressed, to more formal approaches:

> DR: How do you know what hydrologists will want from WATERS?
> A: Well, I go to a lot of meetings and conferences, I read all our journals, I talk to people. So that’s how I know…
> DR: So you know because you are a participant in the community?
> A: Well … when it comes right down to it though, if we want some hard numbers, we've also done those surveys, the CUAHSI HIS survey ( ) and then the CLEANER one. Those surveys give us specific details we can share with others.

We have found that the reverse is also true. Reporting of formal findings can become an occasion to discuss what has been learned through vernacular means.

We will use the term *vernacular* to refer to information gathering that is primarily structured in interaction (e.g., conversations or question and answer periods following presentations). We will use the term *formal* to refer to social and information science methods, such as interviews, surveys, user-studies and requirements engineering. [2]

Conversational interactions are often the richest means for gathering community opinion. Structured *in situ,* conversations allow the participants to range freely, unlike, for example, surveys which substantially constrain topic and length of responses. Often such discussions occur with colleagues who are also friends, collaborators or old school-mates. The conversations can be casual, frank, opinionated and need not

[2] This said, we do not wish to belabor the distinction vernacular/formal. In practice they interlock so tightly as to merit disavowing the terminology altogether. However, the distinction serves to remind us of the heterogeneity of methods participants draw on to know community, and that information gathering is both an 'everyday' occasioned activity and one that is planned and implemented.

be rendered accountable by reference to extrinsic 'facts' (i.e., as is expected of an official community representative, see below). Because such discussions are 'not on the record' there is a greatly lessened concern with repercussions; the converse of this, though, is that while such conversations can serve to inform an understanding of community opinion, the information so gathered is not particularly accountable. They are simply "impressions I've gathered from talking to people," or "the general sense of what the community is thinking" (Env.Eng.Interview). With vernacular discussions, it is difficult to prove to the community that their opinions are being taken into account.

Vernacular discussions are often used as the 'testing grounds' for design decisions and new planning documents, what during an interview one WATERS PI referred to as the "canaries in the coal mine," those individuals who serve as the vocal vanguard of community discontent:

> B: There are certain people that are more sensitive on this topic than others, that I use as kind of the canaries in the coal mines. Where, if you keep those people happy, then we're doing okay.
> DR: Are they particular individuals or a group?
> B: No, they're individuals. One of the individuals that we (have) from the CLEANER division is ['Bob'], our engineering canary in the coal mine. He's very worried about this (merger), very worried…. And we have to work hard to make sure Bob is happy, because that's important. Because if Bob's unhappy, the core of our community is also going to be unhappy. So his role is actually, I think, is important. (Env.Eng.Interview)

These individuals, who are 'sensitive' to the issues of design relative to, for example, the brokering of WATERS, may not have any formal authority in the community but are "respected, connected or just happen to say what everyone is already thinking" (Env.Eng.Interview). By consulting first with these representatives (for example, in reviewing the draft of a forthcoming planning document) the hope is that later dissent can be tempered by modifying language or ceding concessions for design in advance.

4.2 Community Forums

Vernacular opinion is often collected opportunistically at the professional meetings of scientists or in specially arranged forums for the community to 'provide feedback'. In both cases such community opinions are usually expressed following presentations intended to educate or inform the community about the infrastructure development endeavor.

The scope of resource investment in WATERS – financial, human, and temporal – is larger than any comparable previous single endeavour in the history of environmental engineering and hydrologic research. Its plans for large-scale instrumentation are ambitious, and the goals for data interoperability and tool provision can seem opaque to domain scientists: "I think ( ) part of the issue is that the community doesn't / hasn't worked with the volume of data that these observatories are going to generate. A few people have, but the bulk of the community hasn't. So I don't think they understand what they're going to be (faced with)," (interview, WATERS). Unsurprisingly, practicing scientists have occasionally responded with ambivalence and at times with stated opposition to "putting all our eggs in one basket" (comment at WNCW, see below). It is for this reason that when opinions are collected face-to-face they often occur in forums which first establish the future goals and services of the Network.

These elaborate presentations introduce the 'science goals' and informational services of WATERS while contextualizing it as a public good, or, at least, a good for the scientific community:

> I think there's the issue of infrastructure versus research. And first of all, helping the community to understand *this is an infrastructure project* that *enables* the research, and that it does fund some research, test-bedding, so that you [the community] can understand what the infrastructure is that you need. (emphasis in original, interview, WATERS)

Community forums are formal in the sense that they are events specifically organized to educate the community on the existence of projects and then collect feedback on their responses, evaluations and opinions. However, these events are also vernacular in that, while they are organized and orderly, the character of the discussions and how 'community feedback' is expressed is often casual and conversational.

Community forums provide a venue for scientists to convey dissenting opinions but they also serve to provide status reports and demonstrate the emerging results of the planning process. For example, the two-day 2008 WATERS Network Community Workshop (WNCW) was organized to collect feedback on a new programmatic document. The workshop drew approximately 100 participants, which included (prominent) practicing scientists, NSF officers, education and outreach specialists, and representatives of various state agencies (EPA, USGS etc...). The presentations moved 'top down,' beginning with NSF's disciplinary interests and funding commitments; followed by discussions of vision, goals and community needs; then impact and outreach; and finally a (still very high order) discussion of design, outlining details such as scales of measurement, kinds of available instrumentation, data standards and tools for collaboration. Eliciting community opinion is often intertwined with forming it.

Lucy Suchman [21] has entitled such performances *public persuasions* and described how they seek to create structured venues to engender 'community debate,' while also serving to shape that debate. For instance, in her study of the planning process for the construction of a bridge, engineers had created a 'community viewpoint' image. This is a fictional model of the view of the bridge from the nearby residential zone: a sunny skyline highlighting the modern-marvel bridge as children cheerfully played soccer by the shores. The image 'advised' the community about what the bridge would look like, but it also served to soften expressions of discontent about a potential eyesore. Similarly WATERS has created cartoonish diagrams populated by instrumented rivers, basins, and cornfields while contended scientists lie in grassy fields and interoperated data are funneled to their laptops. Such imagery, and the programmatic speeches that accompany it, convey the 'transformative vision' of WATERS to scientists attending these meetings; but they also quell opinions that the infrastructure may not serve the needs of actual scientists.

The question and answer periods at the interstices of formal presentations become moments for 'community feedback,' (or dissent) as scientists (not directly involved in the organization of WATERS) step up to microphones scattered across the room and direct their questions to seated panelists, often challenging science, planning and implementation efforts.

The overall tone in the single room of the WNCW was earnest. Attendees expressed a sense that what would come to be referred to as ‘community opinion’ was being forged by the many prominent scientists, leaders of the research institutions and representatives of state agencies. For example during an open forum one environmental engineer asked a question around the distinction scientist/engineer, that is, that environmental engineers are focused on water quality while hydrologists tend to orient to questions of water quantity:

> I think your presentation was very interesting. And I noted – and I think you have it right – that we should worry about both water quality and water quantity. But I also noted that the water quantity piece was very well-developed and the water quality piece was pretty primitive. […] If this were a hydrological MREFC, I think we’re well along. If this is a combined water quality and water quantity proposal, I just don’t see that it’s clear enough just yet. So in my mind we need to do a fair amount of work to make sure that the two disciplines are really working together. (comment at WNCW)

Nominally, environmental engineers are concerned with quality: pollution, contamination, sewage, hypoxia, potability. Hydrologists are concerned with quantity: drainage, erosion, and so on. The speaker continued by specifying details that would shift emphasis to engineering concerns, such as focusing on water budgets and nutrient fluxes, and ensuring that sensors captured their primary markers such as nitrogen and phosphorus levels. This discussion is 'about science' in the sense that it focuses on research questions and data, but it is also a negotiation about the trajectory of planning and development, and which communities will be best served by the future infrastructure.

A common complaint in such community workshops was to note the lack of clear mechanisms for incorporating the 'feedback' expressed by attendees: "how is what I'm saying here going to change what you are doing?" (question posed by attendee at breakout group of WNCW). Inevitably in such workshops, the suggestion was made that such community forums are 'all show,' merely designed to create an impression of community engagement. Workshop organizers found themselves at pains to demonstrate otherwise:

> We're here listening. We want to know what the community thinks. And most of the material will be placed online, including videos of the presentations and a report that we'll write from this workshop. (WNCW presentation to attending audience)

The claim that community workshops are 'simply all show' appears extreme. More reasonably, they are *simultaneously* rhetorical ventures to shape community opinion, performative displays to demonstrate community engagement, *and* genuine testing grounds (cf. canary in the coal mine) to vet plans and documents. That 'videos online' and 'reports' are the only signs that the community 'is being listened to,' is simply the nature of the available means for summarizing and communicating events such as workshops.

Put briefly, it is difficult to demonstrate accountability to the community's feedback as stated within public forums. However, both casual conversations and workshops also inform the production of surveys and requirements elicitation instruments[3], with which participants hope to substantiate vernacular opinion with 'hard data': "we *know* that environmental engineers are more enthusiastic about WATERS [today] than they were a year ago, now we just need to be able to communicate it," (interview, emphasis in original, WATERS).

4.3 Surveys and Requirements

Surveys and requirements gathering produces the most tangible results about community opinion and needs. In the planning process, findings from such studies are deployed by representatives seeking to shape technical development in ways which render themselves accountable to their communities. Results are accountable in that they can be publicly displayed in reports and presentations using numerical values, graphical representations (such as pie charts and bar graphs) and direct quotes. Surveys and requirements generate findings that are objective, that is, visible and inspectable, and thus carry a rhetorical weight that knowledge created through vernacular means simply does not.

Communicating 'community opinion' in authoritative and unambiguous ways is one of the primary purposes for conducting surveys and requirements gathering. The findings of vernacular discussions and community workshops can be rich and complex; however, they are also difficult to express succinctly or to interpret them as directly informing design. In the excerpt below, an NSF program officer (responsible for the continued funding of WATERS) for environmental engineering is paraphrasing his demands from CUAHSI:

> You guys are hydrologists. So I can’t speak for hydrologists. You’re the experts in those fields. *The evidence to me that you actually have a community is when I see that you guys come up with a science plan.* […] We want you to have 3 (science goals). That’s your problem. If you want this to be successful, you’ve got to show me that you can pick 3 things and show that your community stands behind it. […] And I don’t want to hear you [expletive which means 'complaining'] about how diverse your community is. Every community has differences of opinions about the way they should go. (NSF, interview, emphasis added)

A great deal of ink has been spilt demonstrating how surveys or interviews produce findings that are 'tidier' than the opinions of their respondents [7, 25]; shades of grey are captured in black and white terms. In constructing questionnaires and surveys requirements, participants often express "the enduring assumption," identified by sociologist of computing Steve Woolgar, that community opinion and "actual requirements pre-exist our efforts to 'capture' them," [28:203]. It is more accurate to say that much opinion is formed at the junction of respondent and survey instrument. However, this is not to say that in the production of data, surveyors do not have an orientation towards accuracy and reliability. Rather, it must be

[3] Community forums also serve as the sites for *administering* questionnaires, thus generating formal results.

acknowledged that there is also a desire to clarify, simplify and more generally represent 'community opinion' in such a way that it can be easily codified and communicated. Similarly, with requirements engineering, needs must be captured in forms which can be "operationalized for design" (interview, CUAHSI, computer scientist).

The surveys conducted in CUAHSI, CLEANER and then WATERS did just that. These studies revealed a field of relatively clear differences and commonalities across the communities of environmental engineering and hydrology, defining the boundaries of each field while also pointing to promising avenues for cooperation. For instance, the WATERS survey differentially identified the databases from which hydrologists and environmental engineers commonly draw: "each group has very different needs for data sources. For example, CUAHSI members are substantially more reliant on USGS streamflow data," [9:3p] while engineers "had a substantially greater interest than CUAHSI members in the USGS and EPA[4] water quality databases," [9:3p].

Such results quantified and authorized findings from vernacular and community workshop discussions that had distinguished engineers as interested in quality of water and hydrologists with quantity (ibid.). The survey results are both more detailed (identifying particular datasets) and also more authoritative when communicated in percentages and pie-charts. The ephemeral boundaries of engineering and hydrology are buttressed by such findings, constituting a detailed reality from what was formerly impressionistic disciplinary difference.

This said, these fields of differences are complemented by revealing a wealth of commonalities. For example, while the two groups prioritize datasets differently, in aggregate they substantially agree on what they would like to do with data: "Both groups agreed that the most important features of new data sources were finding specific measures or variables (57%) and the accuracy of the data (52%)" [9:2p]. Surveys reveal disciplinary difference as ordered preference and heterogeneous practice but they also point to shared goals, widespread phenomena of interest and commonly desired computer support. These are the raw materials for demonstrating the existence of a single 'water environment community' and claiming a mandate for the construction of WATERS.

Since the community should 'drive' development, surveys and requirements can serve to demonstrate that the community is being given voice in design decisions. The participants in WATERS are called to demonstrate accountability to individual members of the community (e.g., in forums) but also (as the extended excerpt above shows) to the NSF officers funding decisions. Funding of WATERS is contingent on demonstrating a community mandate, i.e., showing how the design reflects the current research goals and needs of hydrologists and environmental engineers. Survey findings are accountable in that participants in WATERS can point to them as capturing the 'desires, needs and goals of the community.'

[4] United States Geological Survey and Environmental Protection Agency.

For example, in an interview, a computer scientist described how he had been challenged in the past with regards to his data integration efforts. However, by pointing to the results of the survey described above, he can silence dissent by locally deploying survey findings:

> The most important result from our HIS survey was understanding which datasets are most important to their [hydrological] community. Now, when we provide services, we know which databases to prioritize, which to integrate. [Hydrologists] care about hundreds of databases, and we only have time to work on a few at a time. And before, no matter which one we worked on, someone was accusing us of ignoring them. Now we just point to the table. You can't argue with the numbers.(interview, CS)

In order to receive funding from NSF, the leadership of WATERS will have to demonstrate a 'community mandate' indicating strong support for the extant science questions and particularities of design. For example, the National Academy of Science report, *Setting Priorities for Large Research Facility Projects* [13] states that its sixth criteria for determining which MREFC project to fund, is:

> *Which projects have the greatest degree of community support?*

The measure of 'community support' is an actor in deciding the future funding of WATERS. The concept of community is in the 'hard numbers' and 'soft impressions' of proper conduct for infrastructure design; and it is entrenched in funding mechanisms and evaluative criteria of the highest state agencies.

4.4 Representing community: Accountability and mandate

In what James Taylor has called the autonomous theory of organization [23] participants are actively engaged in making declarative statements in the name of 'their community'. In this model 'the community' is significant only to the extent that individuals can successfully mobilize it as their representatives, thus effecting the outcome 'community opinion and needs' [24:87 ft.17]. Because communities cannot speak on their own, requiring active representation, they are analogous to constituencies. The leading figures in WATERS and CUAHSI are one sort of community representatives; not in a simple sense of a hierarchical relation, since "authenticity" must be preserved, but in their special capacity to stand in for their communities [24:87 ft.17].

To be clear then, in the formulation in this paper, are communities actors [3]? Yes, they affect the outcome of infrastructure design, planning and implementation. However, communities require representation through enactment of their needs, desires, goals, extant work practices and so on. Does this mean that communities are 'socially constructed'? No. At least, not in any sense of the term which implies that the 'needs and opinions' of a community are fully malleable to its representatives. Knowledge of communities is established by surveys and other forms of inquiry, or more simply as a product of vernacular membership. While (we agree that) knowledge of community is shaped by the means of inquiry, it does not follow that representatives may cast communities as they please. Why? Because representatives must render themselves accountable, and thus, as 'authentic' representatives the community. In this sense, they are similar to political representatives and accountable to their constituencies. In order to faithfully stand-in for their communities, representatives must demonstrate a mandate on the part of that community.

For example, the requirements elicitation process in WATERS consistently showed that both environmental engineers and hydrologists wanted stable and usable support for their basic software suites, such as spreadsheets (specifically mentioned is MS Excel). Such tools are the bread and butter of today's practicing scientists. However, within CI such applications are considered to be 'stove pipes', making it difficult to share data or use them with the high end computational tools WATERS seeks to develop. Notably, this finding was in direct opposition to WATERS' goals for open, interoperable and sustainable data archives. But the finding could not simply be dismissed: it is 'what the community wants'. Representatives who veer too far from a sense of what the community wants can be, and have been, 'called on it.' Scientists with grand techno-utopian visions for data access but no relation to existing data practice can be placed in question as having lost perspective or having ‘gone native’ to computer science. In such circumstances just how representative they actually are of their community may be placed in question.

For spokespersons ‘authenticity’ is not firmly grasped in an iron glove; rather, they must perform a sense of responsibility towards their community, or perhaps more accurately, towards their constituency.

5. CONCLUSION

This has not simply been a study of actors' categories. We have shown community to be more than a matter of 'meaning.' Rather, we have traced how this category is set in motion in vernacular conversation, in community forums, in organizing surveys or enacting requirements, and in the streams of presentations and documents which report findings. Within cyberinfrastructure development, community has been "operationalized" in vernacular and formal techniques of inquiry. Moreover, community has been encoded into the very processes which lead to the funding, planning, design and evaluation of infrastructure development. To transition from its planning and construction phase, WATERS participants are developing detailed documents with plans for development and particular science questions. To transition to the next step in receiving MREFC funds they will also have to show a mandate, demonstrating community support for these plans and research goals. 'Community' is concept organizing inquiry, driving design and outlining criteria for evaluation.

Both amongst our informants, but also within the academic literature, there is a concern that ‘community’ is used as a rhetorical term: an inherently positive descriptor evoking ‘warm and fuzzy’ feelings and potentially obfuscating internal heterogeneity and discord. This concern clearly stems from a more ‘traditional’ usage of the term community as formulated, for example, in the work of Tönnies and more recent sociologists [15, 17, 18, 32]. In such formulations community has a clear positive valence, and its decline is considered to be a social problem. However in the ‘techniques and practice’ of CI participants themselves, we have not seen community take on such a valence. In our studies of large-scale technology planning, the use of community has:

- little or nothing to do with affective bonds or shared moral orientations (except to the extent that they speak of doing 'better science,' 'good research,' or ensuring an 'intelligent investment of resources'); rather, *community often serves as a concept that organizes understandings of difference and similarity*;
- only a glancing relationship to collocation (the 'neighborhood understanding' of community: same place, same time), but substantial relevance to collaboration and coordination;
- *served as an organizing principle for multiple forms of inquiry* which seek to 'know,' 'elicit,' or 'capture' the 'needs,' 'existing work practices,' 'requirements,' or 'goals' of future users;

and perhaps most importantly:

- *community is used as a short-hand for issues of representation, for example, ensuring inclusion and establishing a mandate – in other words, as a near synonym for what in the sphere of politics we would call a constituency*.

A primary contribution of this work to CSCW research and practice is the concept of constituency as a way to understand the planning and development aspects of technology projects where not all the future users can directly bring their concerns to the table. This is particularly significant as CSCW researchers turn their attention from the group level of analysis, where it is practical to assess the opinions of all participants, to the community or organizational (or larger) level – where previous requirement-gathering strategies will not scale. In a project intending to serve 'community interests' (such as infrastructure) these future users require representation. In our case, the concept of community organized a series of inquiries which came to define both their differences and commonalities; in turn, representatives mobilized these findings to inform technology design and demonstrate a mandate. Demonstrating a mandate meant accountably showing that planning decisions faithfully represented the wishes of the community.

The formation of a single community ('water environment researchers') from heterogeneous beginnings ('environmental engineering' and 'hydrology') was an exercise in (political) representation and then action. We traced participants' work as they sought to prove that a single infrastructure could serve the interests of heterogeneous communities. While WATERS may be an extreme case of technology planning in which the configuration of constituencies was completely overturned, it served to reveal the work of knowing community and how that knowledge is then mobilized in design. More generally, we have shown that in such projects the definition of community is at stake in shaping the trajectory of technological development, which is a key insight as CSCW researchers seek to understand and influence the creation of large-scale cyberinfrastructure (vis [10]).

A final important lesson learned from this case for CSCW researchers is that 'needs are in motion' and that knowing those needs is an iterative process of multimodal inquiry. From the 'perspective of user-centered design', the issue is not to deplore the indeterminacy which prevents us from fixing targeted communities, their needs or requirements, but to see this as a condition of the possibility for increasing future research opportunities through the planning and development of information systems. Another lesson is that 'opinion and needs' are not transparent representations of the community, they must be mobilized as findings, and often such mobilization is not only about good design, it also about the allocation of resources, and debates about 'who and how' technology will serve its users.

6. ACKNOWLEDGMENTS

We would like to thank Eric Cook, Libby Hemphill and Jude Yew for their feedback on drafts; Katherine Lawrence's past work with WATERS and conversations with Ann Zimmerman have also proven invaluable. We would also like to recognize our respondents who have been immensely forthcoming about their activities. This work was supported by the National Center for Supercomputing Applications (NCSA/NSF/OCI) Partner Project grant #04-38712.

7. REFERENCES

1. Atkins, D.E.C. Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure, National Science Foundation, 2003.
2. Bender, T. *Community and Social Change in America*. Rutgers University Press, 1978.
3. Callon, M. Some elements of a sociology of translation: Domestication of the scallops and the fisherman. in Law, J. and Kegan, P. eds. *Power, Action and Belief*, Routledge, 1986.
4. Callon, M. Techno-economic Networks and Irreversibility. in Law, J. ed. *A Sociology of Monsters? Essays on Power, Technology and Domination Sociological Review Monogrpah 38*, Routledge, London, 1991, 132-161.
5. Fischer, G. Domain-oriented design environments. *Automated Software Engineering*, *1* (2). 177-203.
6. Gieryn, T.F. Boundary-Work and the Demarcation of Science From Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, *48* (6). 781-795.
7. Goguen, J. Requirements Engineering as the Reconciliation of Social and Technical Issues. in Jirotka, M. and Goguen, J. eds. *Requirements Engineering: Social and Technical Issues*, Academic Press, London, 1994.
8. Goguen, J. and Linde, C. Techniques for Requirements Elicitation. in Fickas, S. and Finkelstein, A. eds. *Proceedings, Requirements Engineering '93*, Computer Society, IEEE, 1993, 152-164.
9. Lawrence, K.A., Finholt, T.A. and Kim, I.-h. Cyberinfrastructure for the WATERS Network: a Survey of AEESP and CUAHSI Members, 2006.
10. Lee, C.P., Dourish, P. and Mark, G. The human infrastructure of cyberinfrastructure. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, *NY, NY*. 483-492.
11. Lynch, M. *Scientific practice and ordinary action: Ethnomethodology and social studies of science*. Cambridge University Press, New York, NY, 1993.
12. Montgomery, J.L., Harmon, T., Hass, C.N., Hooper, R., Clesceri, N.L., Graham, W., Kaiser, W., Sanderson, A., Minsker, B., Schnoor, J. and Brezonik, P. The WATERS Network: An Integrated Environmental Observatory Network for Water Research. *Environmental Science and Technology*, *41* (19). 6642-6647.
13. NAS, T.N.A.o.S. *Setting Priorities for Large Research Facility Projects Supported by the National Science Foundation*. The National Academies Press, Washington DC, 2004.
14. Ribes, D. and Baker, K.S. Elements of Social Science Engagement in Information Infrastructure Design. *ACM: The Proceedings of the 7th Annual International Conference on Digital Government Research, San Diego, CA, May 21-24*.
15. Ribes, D. and Baker, K.S. Modes of Social Science Engagement in Community Infrastructure Design. in Steinfield, Pentland, B.T., Ackerman, M. and Contractor, N. eds. *Proceedings of Third International Conference on Communities and Technology*, Springer, London, 2007, 107-130.
16. Ribes, D. and Bowker, G.C. Organizing for Multidisciplinary Collaboration: The Case of the Geosciences Network. in Olson, G.M., Olson, J.S. and Zimmerman, A. eds. *Scientific Collaboration on the Internet*, MIT Press, Cambridge, 2008.
17. Sommerville, I., Rodden, T., Sawyer, P. and Bentley, R. Sociologists Can Be Surprisingly Useful in Interactive Systems Design *Proceedings of the HCI'92 Conference on People and Computers VII*, 1992, 341-353.
18. Sommerville, I. and Sawyer, P. *Requirements Engineering: A Good Practice Guide*. John Wiley & Sons. Inc., New York, NY, 1997.
19. Star, S.L. and Ruhleder, K., Steps Towards and Ecology of Infrastructure: Complex Problems in Design and Access for Large-Scale Collaborative Systems. in *Proceedings of the Conference on Computer Supported Cooperative Work*, (Chapel Hill, 1994), ACM Press, 253-264.
20. Strauss, A. *Continual Permutations of Action*. Aldine de Gruyter, New York, 1993.
21. Suchman, L. Organizing Alignment: A Case of Bridge Building. *Organization*, *7* (2). 311-327.
22. Suchman, L., Trigg, R. and Blomberg, J. Working artefacts: ethnometods of the prototype. *British Journal of Sociology*, *53* (2). 163-179.
23. Taylor, F. Shifting from a heteronomous to an autonomous worldview of organizational communication: Communication theory on the cusp. *Communication Theory*, *15* (3). 1-35.
24. Taylor, J.R. and Van Every, E.J. *The Emergent Organization: Communication as Its Site and Surface*. Lawrence Erlbaum Associates, Mahwah, NJ, 2000.
25. Throgmorton, J.A. *Planning as Persuasive Storytelling*. University of Chicago Press, Chicago, 1996.
26. WATERS (ed.), *<http://www.watersnet.org/history.html>*, accessed Dec.12.2008, 2008.
27. Woolgar, S. Configuring the user: The case of usability trials. in Law, J. ed. *A Sociology of Monsters*, Routledge, 1991.
28. Woolgar, S. Rethinking requirements analysis: Some implication of recent research into producer consumer relationships in IT development. in *Requirements Engineering: Social and Technical Issues*, Academic Press, New York, 1994, 201-216.