Users as essential contributors to spatial cyberinfrastructures

Barbara S. Poore[1]
United States Geological Survey, St. Petersburg, FL 33705

Edited by Michael Goodchild, University of California, Santa Barbara, CA, and approved February 1, 2011 (received for review August 15, 2009)

Current accounts of spatial cyberinfrastructure development tend to overemphasize technologies to the neglect of critical social and cultural issues on which adoption depends. Spatial cyberinfrastructures will have a higher chance of success if users of many types, including nonprofessionals, are made central to the development process. Recent studies in the history of infrastructures reveal key turning points and issues that should be considered in the development of spatial cyberinfrastructure projects. These studies highlight the importance of adopting qualitative research methods to learn how users work with data and digital tools, and how user communities form. The author's empirical research on data sharing networks in the Pacific Northwest salmon crisis at the turn of the 21st century demonstrates that ordinary citizens can contribute critical local knowledge to global databases and should be considered in the design and construction of spatial cyberinfrastructures.

Cyberinfrastructure (CI) will simultaneously transform the technical tools and the social arrangements of contemporary scientific work according to the Atkins Report (1) of the National Science Foundation (NSF). Following the models of modern biological or climate change research, many sciences will become data-driven, relying on high-performance computing grids capable of processing, analyzing, storing, and indexing enormous datasets. Cross-disciplinary collaboration among teams of geographically dispersed researchers will become more prevalent. Spatial cyberinfrastructure (spatial CI) or CyberGIS, as the similar research trend in geographic information science (GIScience) is alternately referred to, combines the tools and computing technologies of CI with the power of spatial analysis to address complex environmental and social issues such as climate change, disaster response, transportation planning, and national security. Although the use of CI in GIScience is relatively recent, there is an important tradition of spatial data infrastructure research that emphasizes the social aspects of online data sharing and interoperability among data communities. In addition, the recent efflorescence of Internet mapping by private companies, mapping agencies, and ordinary citizens, commonly referred to as the geospatial web (2) can be a resource for GIscientists as they undertake the development of spatial CIs. Spatial CIs can contribute new insights to the more general CI effort by promoting research on the social and organizational aspects of infrastructure development and by demonstrating the important role that ordinary citizens might play in CI.

To date, writings on technical and social transformations projected to result from CI and spatial CI efforts have focused on professional scientists. Both the way scientists work and how they interact with others are expected to change. Cross-disciplinary collaborations will become the norm and they will require new technologies to support interactive virtual environments, shared analytical tools, and in silico models. The Atkins Report does not describe precisely how the technical and the social intermingle in CIs; however, the report asserts that this relationship is key to the success of CIs.

As with any deployment of information technology, it cannot be assumed that just because tools are provided they will be used, or that the affordances of new technologies for collaboration will inevitably lead to the dismantling of the isolated single investigator hypothesis-driven model of science. Although the Atkins Report recognizes that social and cultural issues may impede CI adoption, it proposes no solutions, and only glances at user issues. For example, the report calls for a new interdisciplinary workforce that would include social scientists. The role of social scientists would be to explore how information technology can be used to collaborate across domains, how to achieve broader participation in CI by minority-serving institutions, and how to provide access to CI resources by a wider audience including citizen scientists. However, in other passages, the report takes a passive view of the user, describing how the information needs of professional scientists will be "served" by computer scientists in new operational centers established by NSF. Appendix B of the Atkins Report contains an illuminating user survey. In contrast to the forward-looking tone of the report, which emphasizes the new affordances of such technological breakthroughs as tools for long-distance collaboration, the majority of survey respondents was unaware of these tools and did not see themselves as using them in the future. Of considerably more importance to survey respondents in terms of CI usability were the seemingly more mundane issues of removing impediments to accessing data in federated repositories or digital libraries. These user desires are not accounted for in the Atkins Report.

To be sure, the NSF has subsequently funded several CI projects, described in a subsequent section, in which social scientists are studying collaborations among domain scientists as CIs are being constructed, however, broad categories of potential CI users, such as teachers, students, citizen scientists, and the general public have not been the focus of much of the CI-related research since the report. This work will argue that as the GIScience community "goes cyber" and invests significant resources in the construction of a spatial CI, it has an opportunity to think critically about the ultimate aims of a CI. A focus on users should be paramount, and a spatial CI should look beyond professional domain scientists to the potential contributions of nonprofessionals.

There are several reasons for incorporating these types of users into a spatial CI. Spatial CIs will address questions of universal social importance such as climate change, global pandemics, hazard forecasting, and water availability. Ordinary citizens have a stake in these issues and should be participants in the decision making. These issues have a geographic basis, and a spatial CI will provide sophisticated analytical tools and visualizations, but much of the data that is needed to address these important questions are quite local in scale and not easily attainable by traditional methods. There are many

Author contributions: B.S.P. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

[1]E-mail: bspoore@usgs.gov.

efforts within CIs to support the use of sensor networks and mobile devices for data collection by professional fieldworkers (3), but some projects, for example, in phenology and ornithology, also rely on citizen scientists to collect these highly local data (4, 5). Spatial CIs might benefit from citizen scientists who make use of Web 2.0 tools to contribute data (6). Finally, although the geospatial community has strong quantitative traditions, alternative critical and qualitative research traditions have been developed in recent years that give it access to broader user communities.

Although the other papers in this issue focus on GIScience technologies for spatial CI, this work draws attention to these alternative social science-based, qualitative research traditions. These traditions include a long history of research on social aspects of data integration (7), spatial data infrastructures (8, 9), public participation GIS (10), spatial data decision support systems (11), and volunteered geographic information (12). How do we find and engage the ordinary nontechnical users who are potential citizen scientists (13)? This paper will argue that GIscience researchers already have the necessary tools to make these users visible. In the context of qualitative research, going cyber takes on additional meaning, because researchers use participatory techniques and social software tools such as Facebook and Twitter to interact directly with users and understand what infrastructure might look like from their standpoint. In effect, the researchers try to step inside the infrastructure, rather than serving the abstract "users" from outside the user experience as was proposed in the Atkins Report. This work is laid out as follows: a review of the literature on the histories of infrastructure and the emerging social science field of CI studies is followed by a discussion of related topics in the GIScience literature that can contribute to this new terrain. The work will draw on the author's ethnographic study of data sharing communities as they grapple with saving salmon habitat in the Pacific Northwest to speculate on how spatial CI's can be enriched by a focus on users as contributors of local data.

Studies of Infrastructure

The mix of technical and social issues that constitute an infrastructure has been studied for a number of years by historians of technology and by social scientists. Historical studies of infrastructure, referred to as large technical systems (LTS), open up previously black-boxed and invisible infrastructure to scrutiny to reveal that the deployment of these systems depends as much on such social intangibles as organizations, institutions, standards, laws, and markets as it does on the technologies. Edison did not just invent a substance called electricity, he envisioned an entire system from generator to transmission line to lamp (14, 15). The dams and electrical power plants in the Pacific Northwest are not pure technologies, they are tightly coupled to the region's history, to its economics, and to the natural systems of rivers and salmon (16). American cold-war computer defenses were, in part, shaped by metaphors of a closed, self-contained world, influencing the subsequent trajectory of computer development (17). The Internet works as it does today due, in part, to bureaucratic battles between the US Department of Defense and the hacker community over proposed machine communication standards (18). Sometimes highly visionary but technically feasible projects such as a French scheme to build a personalized mass transit infrastructure fail, and these failures cannot always be easily explained (19).

Histories of infrastructure expose the "long now" of infrastructure (20, 21)—the amount of time it takes, starting from the initial vision, for successful infrastructures to spread out in space and through time and become simultaneously ubiquitous and invisible. We do not often think in these extended terms and, furthermore, does studying an infrastructure as it is in the process of forming give clues to its future? A number of social science researchers are studying evolving CI projects to produce generalizations about the sustainability of CIs that can be widely applied (22–24).

However, studying infrastructures is a paradoxical venture because infrastructures are hardly meant to be noticed, much less studied in depth. According to Star and Ruhleder's (25) pioneering study, infrastructures have certain characteristics that contribute to their invisibility. Infrastructures are

i) embedded in other structures,
ii) built incrementally on an already installed base, and
iii) only become visible upon breakdown.

To penetrate the invisibility of infrastructure, Star and Ruhleder had to go cyber. They became participant observers in the creation of a virtual laboratory for worm biologists from both the top-down perspective of the designers (computer scientists) and the bottom-up perspective of the users (biologists). Although the previously noted characerisitics of infrastructures can be said to apply to the designer-eye view, Star and Ruhleder also uncovered characteristics of infrastructure from the users' perspectives. Infrastructures are

i) meant to reach many users across space or through time,
ii) to be transparent to the user,
iii) to be learned as part of membership in a community,
iv) to be linked with community conventions of practice, and
v) to embody community standards (25).

Star and Ruhleder used ethnographic methods in their study. Ethnography is a participatory research technique that had its origins in anthropology but which has since been adopted in many other fields, including information systems research (26). The ethnographer, whether she studies an exotic culture in a far away land or a mundane information system close to hand, renders what anthropologist Clifford Geertz (27) terms a "thick description," that is, not just an isolated description of a gesture, an artifact, or a statement, but the social and cultural contexts through which these items are manifested. Ethnography is a particularly suitable technique for studies of infrastructure because ethnographers are trained to find what is hidden behind outward appearances, because infrastructure is hidden. For example, designers of information systems often deploy master narratives that guide how the final system will look or behave (e.g., that the computer will transform office work), and these narratives can be exclusionary (e.g., for a housewife or an artist who might not organize their work into desktops, file folders, and trash cans) (28). In the biology community Star and Ruhleder studied, the scientists, motivated by professional boundary policing, insisted that system designers not solicit system requirements from the secretaries who helped produce the scientists' publications, although the secretaries were potential users of the system. The authors also uncovered the ways in which the representations of work in system design impose unfamiliar processes that must be learned, and require articulation on the part of users to fill in the gaps between what is specified in a requirements analysis and what the user has to do to make the system work in practice (25).

Social scientists have begun to apply insights from the historical studies of LTS and the ethnographic techniques of Star and Ruhleder to studies of CI as they are being developed. Beginning with the study of single CI projects such as the NSF-funded GEON project (www.geongrid.org), a CI for the earth sciences (29), and the National Institutes of Health's Biomedical Informatics Research Network (www.birncommunity.org) (30), a biomedical research CI, the field has progressed to comparative studies. For example, the Comparative Interoperability Project (31) is comparing three CI projects for the earth and environmental sciences, GEON, the Long Term Ecological Research Network (LTER) (www.lternet.edu), and Ocean Informatics (http://oceaninformatics.ucsd.edu), a data-sharing initiative for the ocean sciences. A recent workshop on the history and theory of infrastructure sums up this research strain (15). The history of LTS demonstrates that systems typically begin with a system builder, such as Edison. Technology transfer invariably leads to local differences and the emergence of competing systems. Transparent, reliable infrastructures are only formed when standardized gateways between local systems develop, but path dependence can mean that early moves toward standardization limit later possibilities. An analytic focus on the scientists who use the systems reveals tensions that are typical in infrastructure development including how standards for data sharing and preservation are negotiated, and how conflicts between local and global implementations are resolved. The report recommends designing processes that allow for flexibility in the face of technological change, easy navigation for users, and the inclusion of marginalized users. These CI studies describe users and developers as enmeshed in the same web (32). Thus, it is impossible for one to act without affecting the whole system.

Implications for Spatial CI

The history of infrastructure building and the studies of ongoing CI development discussed above will undoubtedly be valuable for the designers of spatial CIs. However, these designers must also grapple with lessons learned in their own discipline from spatial data infrastructures (SDI) at the local, regional, national, and global level that predate the current focus on CI. An SDI has been defined as "the "technology, policies, standards, human resources, and related activities necessary to acquire, process, distribute, use, maintain, and preserve spatial data. (33)." SDIs arose in the 1990s as a perceived response to the transition to a networked society (34). Geographic information systems (GIS) software that had operated on the large mainframes of government computers was becoming available on personal computers and spreading into state and local governments and the private sector. The Internet allowed people to find and share geographic data that had been prohibitively expensive to create. In response to calls in the United States for a national information infrastructure (35), federal agencies that used GIS technology conceived of a national spatial data infrastructure that would provide a reliable system for online production and sharing of data among many diverse data producers and users (36).

The idea of SDIs has spread globally (37), but despite >10 years of development and research, SDIs have had decidedly mixed results. To date, it is difficult to point to an unequivocally successful SDI that has all or even most of Star and Ruhleder's characteristics, listed above. A global survey of the state of SDIs finds data that are missing, poorly documented, incompatible, and hard to access and use. The infrastructures are not interoperable; there is a lack of global leadership; the cultural and organizational issues involved in infrastructure building have been underestimated; and SDIs provide insufficient training for participants and users (38). Several researchers have pointed to the lack of critical evaluation in SDI research in favor of an emphasis on the technology and the utopian promises of SDIs (39, 40). There is general agreement among critics of SDI research that it could benefit from more emphasis on such interpretive methods as Star and Ruhleder's ethnography (41, 42) and that users need to become a more central focus of SDI research (9).

The necessity to revise our ideas about the role of users in SDIs and in spatial CIs has become more urgent in the new world of the geospatial web. Since the declassification of the GPS in 2000, the publication of the Google mapping platforms for creating map mash ups in 2005, and the spread of geo-enabled cell phones, millions of nonprofessional users are contributing data to online participatory mapping systems (43). OpenStreetMap, an open-source volunteer effort to map the entire world, is beginning to rival SDIs in terms of coverage, completeness, and accuracy (44).

Users are no longer the passive recipients of data from national mapping agencies and the SDIs they have established—users have become the producers of their own data (45). It will benefit spatial CI and SDI researchers alike if they can recognize the users who are also data producers by bringing their hidden work to light. Participatory techniques are required. Researchers must go cyber to look at CI from both the top-down view of system designers and the bottom-up view of users at the local level. These techniques can uncover the metaphors and master narratives of CI builders as well as articulate who is excluded. Ethnography can help communities collaborate in cyberspace by uncovering common meanings in dissimilar concepts (46).

There has been a strong tradition of map-use research in cartography (47), but maps are only one potentially fleeting product of the underlying networks that are powered by data, software, hardware, standards, and people. Thus, spatial CI research must go beyond cartographic user studies to consider how the end user interacts with the entire system, including the interface, as well as how organizations use tools to generate, analyze, and provide access to data (48). Although some have argued that SDI research is too often mired at the level of data (39), data are precisely where a bottom-up study should begin (46). Studies of CI have shown that: Data, and the anxieties and tensions it occasions, represents the front line of CI development; its main site of operation, its most tangible output, and in some ways (as the NSF's Cyberinfrastructure Vision document lays out) the target of its highest ambition (ref. 15, p. 31).

CIs are inherently complex, involving many participants with different expertise, spread out over large geographies. CI studies have found that the tensions are different depending on the scale of observation (49), but the local scale, where data are created, has not been the focus of much research either in the CI community or the SDI community. The case study that follows describes tensions at a local level as watershed workers struggle to use technology to collect data on their watershed and merge them into a larger regional SDI, a shared network of data on the rivers and streams of the Pacific Northwest. These workers are seen as essential builders of infrastructure as they articulate local knowledge into a more universal system. Although this study was conducted before the global expansion of the geospatial web, it clearly demonstrates how membership in a technological community is established and how deployment of an infrastructure depends on community conventions of practice. The case study is an example of why the history of LTS is relevant to today's practice.

Case Study: A Regional Spatial Data Infrastructure

In 1999, nine salmon populations were listed as endangered or threatened under the Endangered Species Act by the National Marine Fisheries Service and the National Fish and Wildlife Service (50). Detailed regional maps of stream conditions in the watersheds of the Pacific Northwest were essential to improving water quality and habitat so that salmon could return to spawn successfully. As digital stream representations became widely available from national mapping agencies in the late 1980s, many organizations in the Pacific Northwest began using these data to reference the location of water quality, habitat conditions, and salmon spawning. The digital stream network thus became a framework for the creation and management of other types of data. However, each organization mapped only to its own borders and no further. There were differences in what each organization mapped and the techniques it used to map. Each database of streams had its own spatial reference system, scale of representation, data model, semantics, and attributes. The salmon listings helped spur efforts to create one single consistent representation of the rivers and streams throughout the region so that information on habitat conditions and salmon spawning could be easily visualized and shared. A project was initiated to create and maintain this common framework of stream data through a clearinghouse on the Internet (51).

Approximately 40 organizations including federal, state, and local governments, private companies, universities, and nonprofit organizations participated in this project. As a shared dataset, the Pacific Northwest hydrography framework (PNWHF) is an example of a spatial data infrastructure, linked to the National Spatial Data Infrastructure (NSDI) in the United States (33). Apart from the local data creation issues that are the focus of this paper, lessons learned from the PNWHF are very useful to developing a spatial CI. These issues are consonant with the findings of the CI studies discussed above. For example:

The importance of the long now: The hydrology data sharing project, which formally came into being in 1999, built on data sharing practices that went as far back as the late 1980s. These practices were entwined with the history of regional electrical and water power development since the 1930s. On one hand, there was a strong tradition of regional cooperation, but differences in how Oregon and Washington governed their water resources made data sharing difficult. An SDI is expected to endure into the future; such a process must be planned with past trajectories in mind.

How standards are negotiated: Competing ways of referencing stream locations existed between the two major sources of hydrology data, the US Geological Survey and the Environmental Protection Agency. These measurement frameworks (52) were related to longstanding differences in how scientists from different disciplines collect data to monitor stream pollutants versus salmon habitat. Evolving a common data model became a three-year process of regionwide and national negotiation.

Emergence of competing systems: Although negotiations among federal, state, and local agencies in Washington and Oregon were proceeding toward the development of a common data model, the Native American tribes were forging ahead independently, developing their own linked water framework encompassing the region. One key technologist was advising both groups but promoting different solutions.

Coordination of data communities across space and through time: A loosely coordinated governance structure that transcended official governmental boundaries and reached into local watersheds evolved over the course of the project, held together by a series of protocols that were both technical and social in nature.

Importance of gateways: A data clearinghouse with a standard method for data representation was established on the Internet through which data could be shared and updated by watershed workers. Only when this gateway emerged was there a way for local watershed workers to contribute data to the larger system.

Articulating Local Knowledge—Users as Contributors to Spatial Data Infrastructures

Producing data that can interoperate and be shared by different groups is critical to all SDIs and CIs. The PNWHF project was modeled after an idea proposed by the Federal Geographic Data Committee to develop basic framework datasets for the NSDI (53). This national framework was both technical and social in that a governance structure for managing the data was part of the initial vision. A hierarchy of producers and data managers with different roles, from the national level down to the local level, was imagined. Each level was responsible for managing the activities of those beneath it. It was assumed that local governments would feed their detailed data upward into the national system, but very little was prescribed in advance about how this would happen. The streams and rivers of the Pacific Northwest also are hierarchically organized into watersheds, but these do not follow neat political divisions. In its first phase, the PNWHF project built a regional database from existing data across state, county, municipal, and departmental lines. Reconciling those datasets was difficult enough, but the available data were at a relatively coarse scale. The project needed large-scale data on each small watershed to effectively monitor progress in salmon recovery. User/contributors in each watershed were also needed for long-term maintenance; the long now of infrastructure as it progresses into the future.

To see how local watershed users became essential contributors to the regional SDI, the author followed a worker from the Cedar River Watershed outside Seattle as she collected data on the watershed, used online tools to integrate the data into the regional database, and worked with colleagues to make corrections and additions. The observations that follow were derived from this participant observation and from interviews in the Cedar River watershed and elsewhere in the region in the early 2000s (8).

The GIS manager of the Cedar River watershed knows that the hydrography data the watershed received from the Washington Department of Natural Resources (WADNR) is at too small a scale to show all of the streams in the Cedar River watershed; they must be mapped locally. In addition, the WADNR data, originally derived from US Geological Survey topographic maps, are known to be inaccurate. He asks my informant, Elizabeth (not her real name), a GIS specialist, to begin ground truthing the watershed's stream network by using GPS technology. When the mapping of the stream network is complete, the watershed's stream maps will serve as a linear referencing point to link habitat data collected by biologists and hydrologists who work in the watershed. If the linear referencing system conforms to the standard that has been developed by the PHWHN project, it will make it easier to share information with managers in neighboring watersheds.

Elizabeth goes into the field with a biologist in the early spring to survey Walker Creek. They are mapping stream contours, habitat conditions, and the markers left by biologists who surveyed salmon nests the previous fall and winter. The fieldwork and subsequent work in the GIS laboratory consists of tensions and articulations. The tensions occur on many levels—between Elizabeth and the technologies, between Elizabeth and the biologist, between what is observed on the ground, what is on the maps, and what is in the machine. Articulations are connections Elizabeth makes between different types of local knowledges and the measurement technologies she uses to record them. The GPS is not easy to use, and Elizabeth and the biologist are neophytes. They are on a steep learning curve. As a technology, GPS articulates a whole network of rocket launches, tracking stations, and geodetic calculations, and all must function smoothly for Elizabeth to produce points of interest for her database. There is some trouble locating satellites in the dense woods. Elizabeth and the biologist must move the apparatus around to get the proper readings. Collecting the points is an articulation of the different worlds of the scientists in the watershed. Elizabeth and the biologist negotiate over which attributes—species name, appearance, etc.—should be captured about each patch of gravel where salmon have laid eggs. Not all of the attributes the biologist wants to record can fit into the limited fields of the GPS data logger. Negotiating in the absence of standards is articulation work and shows the importance of tacit knowledge in SDI and CI construction (54).

Back in the GIS laboratory, connected to the global network of the Internet, Elizabeth adds the points she collected in the field to the Cedar River hydrography database and begins to integrate the Cedar River data into the regional database. While looking at a much older technology, the topographic map, Elizabeth notices that her survey is apparently in conflict with the map. According to the survey she just made, Webster Creek flows into Hotel Creek, which, in turn, flows into Walsh Lake. The topographic map and the data from the DNR that are derived from the topographic map show Webster Creek and Hotel Creek flowing separately into Walsh Lake. These data are old and in the Pacific Northwest, it is not uncommon for streams to change course over the years.

Further fieldwork will confirm Elizabeth's observations, but this is the kind of local knowledge that can be brought into the database only because people are out in the field collecting it. Remote sensing methods for collecting stream courses underneath dense tree canopies have yet to be perfected.

Elizabeth's experiences in the laboratory demonstrate how an outsider learns to become an insider to the CI; how a user becomes a contributor and a designer. Her job is an entry-level position, typical of many jobs at the local-level city and county GIS shops. These workers are the shock troops of SDIs, but in their ability to add data and correct old errors, they become a creative force of major proportion. Elizabeth is not starting from scratch to construct the Cedar River data; she starts with the data the watershed inherited from the WADNR. She is improving the accuracy of the line work based on the GPS survey, but she is also adding something entirely new to the database that comes as a result of the regional project—stream routing. A route records the connections and flow among separate lines or arcs that represent streams in the database. This adds intelligence to the data and is useful for modeling the flow of contaminants downstream or the progress of salmon upstream.

Elizabeth is new to the stream-routing process, but luckily the PNWHF group has developed a set of methodologies and digital tools to help. She downloads smaller scale data from the regional clearinghouse. It is prerouted. She adds the larger scale data she has collected, and the system automatically merges these new data into the database and defines the routes. However, the merge is not entirely automatic. As with the articulations, Elizabeth had to get the GPS system to perform, human intervention is required. In the process of merging the Cedar River dataset with the regional data, the tool assigns some streams to what Elizabeth believes might be the wrong headwaters. She plots the data on a new map; she shows the map to the watershed hydrologists. When she returns to the laboratory, the map is richly covered with scribbles—local knowledge.

The programmers at the regional level who built the tools she is using designed a process by which the user could assert control over routing, correcting routes that went in the wrong direction, extending routes that did not cover the larger scale data. Using the new sketch map with the hydrologists' information, Elizabeth not only makes corrections, she decides how to name several streams. Her ability to do this empowers her as a user. Through her articulation work she is becoming more than a user; she is becoming a designer. By means of the feedback of the local data into the regional networks, she provides information about what exists on the ground in her watershed, information that would be difficult to obtain in any other way.

Even though technologies have changed and the geospatial web has made end user involvement much simpler, the problems of learning and membership in a community that are critical to infrastructure development remain. For example, the recent surge of volunteer mapping for the OpenStreetMap community after the Haitian earthquake required a special wiki for new volunteers with links to YouTube videos and discussion lists (55). The same qualitative methodologies that were valuable for examining inexperienced watershed workers could be applied in the case of CIs, and many of the same findings would hold. These are the same issues that have emerged from infrastructure studies of all sorts over the past few decades that cannot be addressed through technical means, but are artifacts of the social and cultural contexts of infrastructures. Only qualitative social science methodologies can capture these contexts.

Conclusion

The designers of spatial CIs should give serious consideration to involving critical human geographers and other social scientists in projects from the beginning. These researchers, using qualitative tools, can contribute a number of insights to a developing CI. Knowing the histories of infrastructures, and in particular of SDIs, can counteract the utopian visions that frequently accompany the rollout of new systems, making these systems more effective in the long run. The focus in user studies in GIScience has traditionally been on the individual user and his or her response to the map interface, but this emphasis may be misplaced. Usability must take account of previously unappreciated work practices and articulations that the user has to make, and the tacit knowledge required. Uncovering these knowledges can only be attained by ethnographic methods. Finally, because of the affordances of Web 2.0, common citizens now have the tools, the interest, and the ability to make enormous contributions to solving scientific problems (56). They find new galaxies, they count birds with enough accuracy that their results are used in scientific publications, they are on the watch for diseases as they snorkel coral reefs, and they make maps that rival those of professional mapmakers. The wicked problems (57) of long-standing environmental issues such as the decline of native salmon or global climate change have no "right" solution but must be addressed through a mutual learning process among all affected. These are precisely the types of problems that spatial CIs are meant to address, and seeking out coinvestigators who have the tools to address them can make spatial CIs truly user-centered and inclusive.

ACKNOWLEDGMENTS. I thank Dr. Dawn Wright for organizing this special issue and Dr. Michael Goodchild and three anonymous reviewers for suggestions that improved the paper.

1. Atkins DE, et al. (2003) *Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (Natl Sci Found, Arlington, VA).
2. Scharl A, Tochterman K (2007) *The Geospatial Web* (Springer, London).
3. Aanensen DM, Huntley DM, Feil EJ, al-Own F, Spratt BG (2009) EpiCollect: Linking smartphones to web applications for epidemiology, ecology and community data collection. *PLoS ONE* 4:e6968.
4. Morisette JT, et al. (2009) Tracking the rhythm of the seasons in the face of global change: Phenological research in the 21st century. *Front Ecol Environ* 7:253–260.
5. Cooper CB, Dickinson J, Phillips TB, Bonney R (2007) Citizen science as a tool for conservation in residential ecosystems. *Ecol Soc* 12:11.
6. O'Reilly T (2005) *What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software* (O'Reilly Publications, Sebastopol, CA).
7. Harvey F, Chrisman N (1998) Boundary objects and the social construction of GIS technology. *Environ Plan A* 30:31683–31694.
8. Poore B (2003) The open black box: The role of the end-user in GIS integration. *Can Geogr* 47:62–74.

9. Budhathoki NR, Bruce BC, Nedovic-Budic Z (2008) Reconceptualizing the role of the user of spatial data infrastructure. *GeoJ* 72:149–160.
10. Elwood S (2008) Grassroots groups as stakeholders in spatial data infrastructures: Challenges and opportunities. *Int J Geogr Inf Sci* 22:71–90.
11. Nyerges TL, Ramsey KS, Wilson MW (2006) *Collaborative Geographic Information Systems*, eds Balram S, Dragicevic S (Idea Group, Hershey, PA), pp 208–236.
12. Tulloch DL (2008) Is VGI participation? From vernal pools to video games. *GeoJ* 72:161–171.
13. Shneiderman B (2008) Computer science. Science 2.0. *Science* 319:1349–1350.
14. Hughes TP (1983) *Networks of Power: Electrification in Western Society, 1880-1930* (Johns Hopkins Univ Press, Baltimore).
15. Edwards PN, Jackson SJ, Bowker GC, Knobel CP (2007) *Understanding Infrastructure: Dynamics, Tensions, and Design* (Univ of Michigan, Ann Arbor, MI).
16. White R (1995) *The Organic Machine* (Hill and Wang, New York).
17. Edwards PN (1996) *The Closed World: Computers and the Politics of Discourse in Cold War America* (MIT Press, Cambridge, MA).
18. Abbate J (1999) *Inventing the Internet* (MIT Press, Cambridge, MA).
19. Latour B (1996) *Aramis: Or the Love of Technology* (Harvard Univ Press, Cambridge, MA).
20. Ribes D, Finholt TA (2009) The long now of technology infrastructure: Articulating tensions in development. *J Assoc Inf Syst* 20:375–398.
21. Brand S (2000) *Clock of the Long Now: Time and Responsibility* (Basic Books, New York).
22. Ribes D, Baker KS (2007) Modes of social science engagement in community infrastructure design. *Proceedings of the Third Communities and Technologies Conference, Michigan State University, East Lansing* (Springer, Berlin), pp 107–130.
23. Coutard O, Hanley RE, Zimmerman R (2005) *Sustaining Urban Networks: The Social Diffusion of Large Technical Systems* (Routledge, London).
24. Bos N, et al. (2007) From shared databases to communities of practice: A taxonomy of collaboratories. *J Comput Mediat Commun* 12:16.
25. Star SL, Ruhleder K (1996) Steps toward an ecology of infrastructure. *Inf Syst Res* 7:111–134.
26. Myers MD, Avison D (2002) *Qualitative Research in Information Systems* (Sage, Thousand Oaks, CA).
27. Geertz C (1973) *Towards an Interpretive Theory of Culture* (Basic Books, New York).
28. Star SL (1999) The ethnography of infrastructure. *Am Behav Sci* 43:377–391.
29. Ribes D, Bowker GC (2008) *Organizing for Multidisciplinary Collaboration Scientific Collaboration on the Internet*, eds Olson GM, Olson JS, Zimmerman A, Bos N (MIT Press, Cambridge, MA), pp 311–329.
30. Lee CP, Dourish P, Mark G (2006) The human infrastructure of cyberinfrastructure in Proceedings of ACM. *CSCW06 Conference on Computer-Supported Cooperative Work* (Assoc for Comput Machinery, New York) pp 483–492.
31. Ribes D, Baker KS, Millerand F, Bowker GC (2005) *5th ACMIIEEE-CS Joint Conference on Digital Libraries* (Assoc for Comput Machinery, Denver), pp 65–66.
32. Millerand F, Baker KS (2010) Who are the users? Who are the developers? Webs of users and developers in the development process of a technical standard. *Inf Syst J* 20:137–161.
33. US Office of Management and Budget (2002) *Circular No. A-16. Revised: Coordination of Surveying, Mapping, and Rrelated Spatial Data Activities* (Office of Manag and Budget, Washington, DC).
34. Castells M (1996) *The Rise of the Network Society* (Blackwell, Cambridge, MA).
35. Kahin B (1992) *Building Information Infrastructure* (McGraw-Hill, New York).
36. National Research Council, Mapping Science Committee (1993) *Toward a Coordinated Spatial Data Infrastructure* (Natl Acad Press, Washington, DC).
37. Onsrud H (2007) *Research and Theory in Advancing Spatial Data Infrastructures* (ESRI, Redlands, CA).
38. Craglia M, Johnston A (2004) Assessing the impacts of spatial data infrastructures: Method and gaps. *7th AGILE Conference on Geographic Information Science* (Heraklion, Greece).
39. Georgiadou Y, Puri SK, Sahay S (2005) Towards a potential research agenda to guide the implementation of spatial data infrastructures: A case study from India. *Int J Geog Inf Sci* 19:1113–1130.
40. Budhathoki NR, Nedovic-Budic Z (2007) Expanding the spatial data infrastructure knowledge base. *Research and Theory in Expanding Spatial Data Infrastructure Cconcepts*, ed Onsrud H (Esri, Redlands, CA), pp 7–32.
41. Georgiadou Y, Harvey F, Miscione G (2009) A bigger picture: Information systems and spatial data infrastructure research perspectives. In *GSDI 11*, eds Van Loenen B, Onsrud H, Rajabifard A, Stevens A, Rotterdam).
42. De Man EWH (2007) *Are Spatial Data Infrastructures Special? Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ed Onsrud H (Esri, Redlands, CA), pp 33–54.
43. Goodchild MF (2007) Citizens as sensors: The world of volunteered geography. *GeoJ* 69:211–221.
44. Haklay M (2010) How good is OpenStreetMap information? A comparative study of OpenStreetMap and Ordinance Survey datasets for London and the rest of England. *Environ Planning B* 37:682–703.
45. Coleman DJ, Georgiadou Y, Labonte J (2009) Volunteered geographic information: The nature and motivation of produsers. *Intl J Spatial Data Infrastructures* 4:332–358.
46. Schuurman N (2008) Database ethnographies using social science methodologies to enhance data analysis and interpretation. *Geography Compass* 2:1529–1548.
47. MacEachren AM (1995) *How Maps Work: Representation, Visualization, and design* (Guilford, New York).
48. Elzakker CP, Wealand K (2007) In *Use and Users of Multimedia Cartography Multimedia cartography*, eds Cartwright W, Peterson MP, Gartner G (Springer, Berlin), pp 487–504.
49. Ribes D, Finholt TA (2007) Tensions across the scales: Planning infrastructure for the long-term. *Proceedings of the 2007 International ACM Conference on Supporting Group Work* (Assoc Comput Machinery, Sanibel Island, FL), pp 229–238.
50. *Federal Register* 64 (1999), pp 14308–14328.
51. US Department of Interior, Bureau of Land Management (2007) *Pacific Northwest Hydrography Framework Memorandum of Understanding* (Bur of Land Manag, Portland, OR).
52. Chrisman NR (1997) *Exploring Geographic Iinformation Systems* (John Wiley and Sons, New York).
53. Federal Geographic Data Committee (1995) *Development of a National Digital Geospatial Data Framework* (Fed Geograph Data Comm, Washington, DC).
54. Gerson EM, Star SL (1986) Analyzing due process in the workplace. *ACM T Off. Inf Syst* 4:257–270.
55. OpenStreetMap (2010) WikiProject Haiti/New Mapper. Available from http://wiki.openstreetmap.org/wiki/WikiProject_Haiti/New_Mapper. Accessed August 8, 2009.
56. Hirsh AE (January 13, 2009) Guest column: A new kind of big science. *NY Times.* Available at http://opinionator.blogs.nytimes.com/2009/01/13/guest-column-a-new-kind-of-big-science/. Accessed February 11, 2011.
57. Rittel HW, Webber MM (1973) Dilemmas in a general theory of planning. *Policy Sci* 4:155–169.