Infrastructuring for the Long-Term: Ecological Information Management

Helena Karasti
Department of Information Processing Science, University of Oulu
helena.karasti@oulu.fi

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

Abstract

This paper foregrounds the long-term perspective and the role of information management in creating infrastructure to support collaborative ecological research. The case study of the Long-Term Ecological Research Network is an ongoing research collaboration that integrates ethnographic and action research approaches. We describe three interdependent elements of science, data and technology for which information management provides support, and the articulation work needed for balancing their inherent tensions and the requirements generated by short and long term timeframes. We further describe information managers' learning community and collaboration-in-design, two mechanisms created within the LTER for continuing technology development over the long-term. The notion of infrastructuring is related to ecological information management as an ongoing design process that highlights participation and co-construction, as well as the complex relationships between the long-term, data, participants, collaborations, information systems, and infrastructure. The understudied area that entails issues of long-term, care/maintenance, and infrastructure presents challenges for the design of large-scale collaborative information systems.

1. Introduction

This paper presents the long-term perspective and the role of information management in creating an infrastructure for large scale scientific collaboration. It is based on an ongoing research collaboration with the Long-Term Ecological Research (LTER) Network (http://lternet.edu) integrating ethnography and action research. Though information management (IM) is a recognized part of LTER activities, its practices and practicalities remain invisible to a large extent as described by an information manager in the following quote:

"We don't do things that are in the metrics that the PI [principal investigator] community value. We don't write multi-million dollar grants. We don't publish a bazillion papers every year. We are too busy getting the work out the door. So based on the metric that most of the traditional scientific community uses, we are pretty invisible."

The invisibility of information managers' work is partly due to the nature of their work of providing *support* for ecological science, "busy getting the work out the door" [cf. 35]. We have performed what Bowker calls an 'infrastructural inversion' [8] by focusing on information managers' support work and foregrounding the backstage elements of their everyday work practice, such as the taken-for-granted functioning of data management and database infrastructure maintenance that by definition are part of the background [31]. In so doing this paper joins emerging areas of research that discuss the importance of understanding the typically invisible maintenance and infrastructure work [e.g. 32, 34, 36, 37, 38] and address the need to broaden our understanding of information systems design to account for the variety that exists 'in the wild' [e.g. 4, 11, 40].

We start by giving some background on the LTER network and explaining our research approach and methods including our analytical assumptions about work practice and knowledge. Then we describe taken-for-granted yet invisible elements of information managers' work, i.e. the support they provide for science, data and technology, and the *articulation work* [36] through which they engage in balancing the tensions between the often-contradictory prerequisites inherent to long-term ecological information management. We continue the empirical case by describing two specific collaborative social mechanisms that the LTER information managers have created during two decades for dealing with ongoing technology development over the long-term. Then we discuss ecological information management as an enduring, collaborative design process of *infrastructuring* [32]. To illustrate the importance of this understudied area and transferability of the issues highlighted in this paper, we conclude by relating the LTER case with studies in other fields that have addressed long-term care, maintenance and infrastructure work, and by proposing a number of challenges for the design of large-scale collaborative information systems.

2. Background

2.1. Long-Term Ecological Research (LTER) Network

Ecology is a branch of science concerned with the interrelationship of organisms and their habitats; hence it studies the interaction of organisms and their physical, chemical, and biological environment. Though much of ecological research addresses time scales of less than a month, the importance of long-term phenomena in ecology is well-documented [22]. In ecology, historical change is the key to understanding the present and anticipating the future. For instance, long-term research is essential to revealing and understanding protracted phenomena, such as slow processes or transients, episodic or infrequent events, trends, and processes with major time lags, as well as to the formulation and testing of ecological theory [13].

The Long-Term Ecological Research (LTER) program was initiated in 1980 by the National Science Foundation (NSF) as six cooperating teams of researchers, each team conducting field, laboratory and theoretical investigations focused on a particular research site's geographic study area [10]. Each team at a site addresses patterns and processes that operate on year to decade to century time scales and extend over local to global spatial scales. Each LTER site has an opportunity to choose its own research focus so that relates to the five LTER core research areas, including primary production, decomposition and disturbance [17].

Today LTER is a federation of twenty-four independent research sites and one network office site in the United States [17], and there is a developing International LTER [15]. The U.S. LTER involves more than 1200 scientists and students from a diversity of disciplines conducting multidisciplinary scientific investigations of ecological phenomena in a variety of biomes. LTER sites range from cold polar to hot desert regions and from tropical rainforest to suburban watersheds.

Research proceeds at each site independently while participants also join together for cross-site work and to contribute to the LTER Network. Cross-site research is encouraged through the adoption of themes that span ecosystems, through support of multi-site participation and through promoting diverse multidisciplinary partnerships.

Sites are selected to become part of the LTER program through a competition held by NSF. Continuation of each site is judged every six years by a panel whose criteria include scientific progress and degree of cooperation with other sites. After the initial competition, sites no longer compete against one another for continuation. Rather, the intellectual integrity and coherence of a site's development is considered through the assessment of renewal proposals every six years as well as intensive three year site reviews. The longer-than-usual funding cycles provide a harbor for activities such as multidisciplinary studies, network participation, and community change discussion.

From the outset, LTER placed an emphasis on preserving data for the long-term. The LTER Network initial vision [13] and continuing mission [17] include the concept of data management, requiring it be a part of each research site's science plan. Furthermore, since the mid 1990's, each site is required to have data available on the Internet two years after its collection. Such activities have prompted development of a data management group consisting of a representative from each LTER site. This group initiated annual meetings early on and has evolved into a Community of Practice [21] that provides a forum for cross-site conversations, collaborative project development, and joint network technology strategizing. An elected executive committee has been formed as the need arose to maintain communication and subcommittee coordination between the annual meetings. Today the Information Manager committee focuses on information management [24, 26, 28] and ecoinformatics (http://www.ecoinformatics.org/).

In information management LTER is faced with the specific challenge of how to maintain datasets over the long-term [24]. The need for data stewardship is motivated by an awareness of an ongoing loss in informational content for data that results in the loss of usefulness of data over the long-term. This is captured in an often-referenced graph portraying 'information entropy' (Figure 1) that refers to the loss of information about the data collected to address a particular scientific question by a particular individual researcher subject both to 'retirement' and to 'death'. The extended temporal dimension of preserving data for decades to centuries poses challenges for the design of metadata and long-term memory, of large-scale databases and archives, and of technologies that support distributed collaboration.

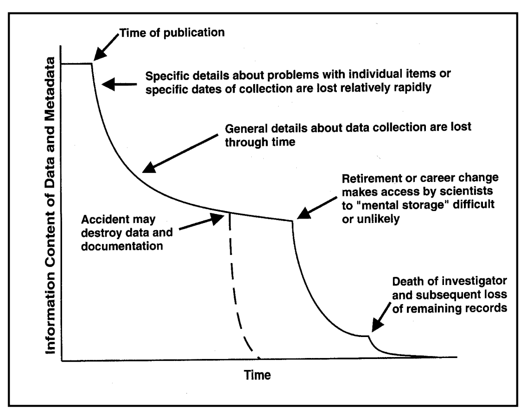

Figure 1. Graph depicting typical degradation of information content associated with data and metadata over time [25].

2.2. Research Project, Approach and Methods

The research reported in this paper has been carried out in an NFS funded Biodiversity and EcoInformatics project entitled 'Designing an Infrastructure for Heterogeneity in Ecosystem Data, Collaborators and Organizations' [3, http://pal.lternet.edu/projects/02dgo/]. The composition of our research team, both outsider and insider LTER investigators with varied backgrounds, allowed for the integration of ethnographic [5] and action research approaches [14, 19, 29]. While extensive ethnographic fieldwork was carried out during year 2002, analyses by the research team and dialogue with the LTER community continue.

Our research is motivated by a profound theoretical interest in *work practices* and *information ecologies*. It is useful to note that we use the term 'ecology' in two meanings, referring both to the science of ecology as in the case of LTER, and to the analytic approach we apply. Our interest in 'ecologies' is not only metaphorical, i.e. intended to evoke an image of biological ecologies with their complex dynamics and diverse species [cf. 27], but pertains to the very observations and analyses of data, practices, collaborations and infrastructures denoting concrete everyday work practices and situated knowledges [6]. Such an analytic approach requires an interest in the mundane, even 'boring' or 'singularly unexciting' things [31] as well as seeing knowledge and meaning as socially constructed within ongoing communities of practice [6] instead of, for instance, relying on a typical distinction in division of labor into 'routine/knowledge work' [7]. We have elicited and articulated significant elements of information and work practice 'ecologies' by focusing on the ways in which data from across a range of media is used; by scrutinizing the relationships between data and their multiple environments; by distinguishing relations between data, participants and their networks; and by relating participants, their organized data and collaboration practices with the social and technical infrastructures [34].

Due to the relatively long history, widely distributed nature and complex organization of the network, ethnographic fieldwork consisted of participant observation and interviews in varied locations and occasions supported by working both with paper and digital archives as well as following ongoing e-communications. An important part of empirical work was the continuous collaboration and dialogue between the fieldworker and the other research team members.

Participant observation [1] at one particular LTER location - that accommodated a number of members of one distributed site - continued for the entire year. This was supplemented by observation of both virtual and co-located meetings of the site. Shorter, more focused visits to a number of other LTER sites comprising observation of selected local work practices together with interviews with site personnel provided for understanding local contingencies and practicalities, as well as differences between the sites. To gain insight to network level activities participant observation was also carried out in various meetings organized during a period of nine months, including network's science, coordinating committee and executive committee meetings, as well as information managers' committee and executive committee meetings. Extensive debriefings and discussions in the research team guided the fieldwork.

Interviews with LTER participants and associates were carried out throughout the fieldwork. Themes were prepared by the research team in advance but sessions were conducted in an open manner that allowed for serendipitous and *in situ* topics and elaborations [18]. Interviewees were selected to cover a variety of LTER sites and all major roles, e.g. scientists, information managers, research assistants, field personnel, technicians, administrators, graduate students, and to provide – as much as possible – for diversity in views and perspectives. Discussions in the research team provided for more informed choices in the selection of interviewees.

Gradually, as the fieldwork progressed, more elements of action research were intertwined. We have created opportunities for participant reflection by sharing with the community our observations and developing understandings. We have presented initial findings for comment and dialogue, for instance, in information managers' executive and annual committee meetings, via postings on our project website and in writing in the community newsletters. Currently we are finishing a report on long-term information management trajectory to continue the dialogue. In comparison to more traditional stance of (participatory) system design, our interventions have been moderate. In choosing to work through 'modest interventions' we acknowledge and account for the multiple politically engaged and epistemically situated perspectives involved in the learning and change processes [16, 23].

The fieldwork resulted in rich material corpus totaling over 50 interviews that average approximately 2 hours in length, 10 notebooks of field notes, various kinds of paper-based and electronic materials, and digital pictures. All interviews have been transcribed, adding up to more than a thousand single-spaced pages. Qualitative, both individual and collaborative, analyses of interview materials [30] started during year 2002. For collaborative analyses each member of the research group would gather excerpts from the interview materials relating to the selected theme. In sessions members of the research team would offer their interpretations on these passages and they would be discussed. The emerging themes included scientific research and collaboration, information management, changes, boundaries, tensions and stories [20].

This paper builds on these analyses plus additional readings of the interview materials. The quotes for this paper are from the interviews, with the exception of two excerpts marked with an asterisk that are from an email conversation. Information managers' interview quotes are marked with (IM) and scientists' not focused on information management with (S).

3. LTER Information Management

This section describes the support information managers provide for science, data and technology. These elements are interdependent and need to be integrated as part and parcel of everyday work. They are also conflicting in many ways, hence information managers need to engage in the *articulation work* of balancing tensions. Strauss explains that articulation work "must be done to assure that the staff's collective efforts add up to more than discrete and conflicting bits of accomplished work" [36]. Star & Strauss explicate further that articulation work is contingent in nature: it is not part of routine, rational sequences of events, rather it accounts for getting things back 'on track' in the face of the unexpected and modifies action to accommodate unanticipated contingencies. [35]

3.1. Providing Support for Science

LTER information management focuses on providing support for 'site science', that is for a research team united by an ecosystem and a common field site. In the words of an LTER information manager:

"One of the things that I see as important is that information management is driven by the research. Information managers continue to come back to assessing whatever projects they want to develop to whether it is really going to support the research at the site." (IM)

As part of a community concerned with long-term issues, LTER scientists are engaged in ongoing discussions about data management and have developed expectations with respect to data issues. In the following quote a senior scientist explains his views about the relationship between science and information management:

"The expectation by others of information management….to take all the messy data and get clean and make it available to them [scientists]. (laughter). But in a sense I think that there is that naiveté, what's all this money going for and what do you get for it ... And they simply don't appreciate the time and the energy and the effort required just to do the nuts and bolts maintenance. Never mind any grandiose new stuff … I think the scientists come in a range of flavors, there is one flavor of scientists who would like all the information management to be totally transparent, the less they have to worry about it and the more that they can get from it the happier they will be. And there are some who take an interest in it and are party to proposals and efforts to both get money into the game and to make advancements in it per se. There is a full range of expectations." (S)

In response to these expectations, an information manager sees that:

"You [information manager] have to … be willing to, to some extent, accept a support role to the main scientific function of the LTER." (IM)

3.2. Providing Support for Data

Ecological research typically deals with heterogeneous data and poses for information management the challenge of dealing with data diversity:

"We have a lot of varied types of datasets. Some studies may have a ton of records, a 'deep database', not a lot of diversity, but huge volumes (like remote sensing). In ecological data in general you get much smaller databases that cover a much wider variety, 'wide databases'. In general you are struggling with the diversity of different types of data, therefore generic modes of maintenance are a challenge. In genetics, for example, in comparison, databases are deep but not as complex." (IM)

Creating a legacy of well-designed and documented long-term experiments and observations for use by future generations requires scientific data be accompanied by contextual information that describes the data collections. These descriptions are called metadata (data about data).

"It's [metadata] really unlike anything that has been done in ecology, and it does preserve datasets over time. Ecological Society of America has tried to identify datasets at risk, important ones to the discipline as a whole and to get them documented. That has been based on the work done in the LTER network, as far as establishing what needs to be documented, the practices…. The network has had a great influence, pushing forward a standardized approach to collecting metadata." (IM)

Long-term data concerns extend the temporal horizon: to the future as well as to the past. Information managers address the varied concerns in their everyday 'data care' work. An information manager tells about recovering forgotten data sets and recording additional historical context:

"I was trying to document a lot of historic stuff … and just asked the PI questions… he was coming on with Alzheimer's and I knew that he was going to retire … and I had a series of interviews with him and I got INCREDIBLE docu, I mean, I got all the documentation for these early corporate [data], like stream chemistry and things, all from just doing interviews with him." (IM)

Another information manager elaborates on taking care of the current, ongoing data capture and archive procedures:

"getting their [scientists'] data into our system from the very beginning to, whether it is to help them with data entry forms, setting up data entry programs, all the way from you know QA/QC programs to getting it archived into our system and accessible on the internet." (IM)

Yet another information manager explains the need for metadata standards as one element of designing data infrastructures for the future:

"as we envision it also that we'll also be adding the EML [Ecological Metadata Language] … And sort of often go back and forth between whether we want to do that from the ASCII files or the database. … but at any rate we'll somehow make EML available dynamically on the Internet to the group at large, to support EML in that effort for having a standard exchange format for metadata." (IM)

3.3. Providing Support for Technology

As technologies are developed at increasing speeds, staying technologically informed is an important aspect of an information managers' work:

"the need for people [information managers] to remain current in technology" (IM)

"technology keep changing, original tape library and mainframe system, it was really kludgey but cool at the time. It is a constant battle to keep up with things". (IM)

Although staying technologically current is a major driver, other factors that relate to the long-term perspective underscore the merits of modest and unadventurous approaches in site information management systems. Judicious decisions about technology procurement are influenced by the features of high reliability, easy maintainability, and low risk for long-term data management and science support. An information manager's foremost concern in aligning developing technologies with existing technologies and practices (with infrastructure) is to minimize disturbance of ongoing data archival and use followed by interest in optimizing long-term data re-use.

"that experience we have had with several of our things… that the issue isn't how you do it, it's how do you maintain it and how do you make it so that it is easily maintainable." (IM)

On one hand, there is the concern for having in place a data-safe, functional system for maintaining the integrity and availability of the long-term datasets. On the other hand, incorporation of new capabilities to enhance data capture, use and preservation always holds the potential for extra facilitation of science.

In addition to balancing the tension between the speed of technological change and the work of 'data care', an information manager is required:

"to do long range planning when new technologies can be placed in, look for the windows of opportunity for proposals for major upgrades for technological infrastructure". (IM)

The evaluation process that places research sites under scrutiny every three years sets a timeframe for some technological updates:

"We manage to update it [web pages] every three years, for review and proposal. We are on this cycle, and we end up putting a lot of energy into updating." (IM)

However, transitions of a larger magnitude occur less often, and the persistence of technological change prompts cautious thinking and careful balancing of options.

"we are transitioning our whole design, we are really facing a lot ... then it stabilizes again. Every so often things need to migrate, the technology changes so much." (IM)

"having the investment in [current technology], it is not so bad yet that I would want to go and rewrite all my interfaces." (IM)

These ongoing and judicious technology procurement and implementation processes produce "a kind of archaeological layering of artifacts acquired, in bits and pieces, over time" [39]. Infrastructures are embedded into and inside other structures, social arrangements and technologies [34]. In LTER an infrastructure is subject to change particularly due to the push for site science, the pull of network expectations, and the emergence of long-term perspectives.

3.4. Balancing Tensions

This subsection puts forward two sets of tensions that are suggested to be systemic and inherent to the field and work practice of ecological information management [cf. inner contradictions, 12]. From information managers' accounts arise three interdependent but also deeply conflicting areas of work and expertise. We present one view of the articulation work they need to engage in by identifying and connecting these elements of support work and relating them to the backdrop of their setting in information management for ecological science that deals with multiple time scales.

3.4.1. Science, Data, Technology. The three research elements of science, data, and technology (depicted in Figure 2) that long-term information management supports come to play in ways that inflict tensions:

"[It's] important to recognize that technology is a tool, and should not be used as an end itself. What does the technology provide for the data you are securing? Potential danger of having just technocrats as information managers, without proper coordination and interaction with the science base. … So that the service that has been provided serves the needs of science as well as providing the protecting cocoon and the ability to service that data to others outside the community." (IM)

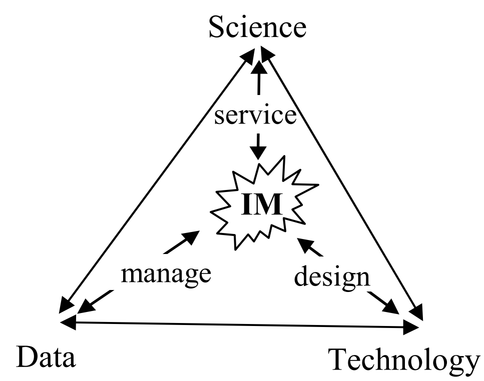

Figure 2. Information management providing support for and mediating between science, data and technology.

Information managers engage in continuous *articulation work*, striking a balance between the intertwined elements. In this work they draw on complex expertises, local knowledge and working experience. The skillful balancing of tensions requires ongoing triage and prioritization while immersed in everyday work activities as described below by a senior information manager:

"when I first started my job, I found … very difficult … there would always be some things that I thought that needed to be done that I could never get to, because I kept having to do triage everyday, and decide what was the most important thing to focus on, and set priorities. Eventually I came to some kind of peace with that, because I felt that was part of my job, to prioritize and decide what was going to get attention and was not going to get attention and occasionally to require more resources." (IM)

Having to deal with the tensions on a day-to-day level has created a particular position of a mediator for the information managers.

"Information manager acting as a communication node between getting the science done: the scientists and the technology" (IM)

"there is a delicate balance there of how you participate…I do think that the LTER… information management community, because of where it sits. See most of the people who are doing, the specialists that are you know that are doing the big projects …They are embedded in an environment of computer science and information technology. On the flipside the LTER information managers, and LTER as an information manager embedded in a matrix of ecologists. And that gives them I think some special insights into what will work in their community and what won't." (IM)

3.4.2. Multiple Time Frames. Rapidly developing technology, data requiring continuous 'slow time' care and science having to cope with short-term funding and long-term motive suggest very different time scales. Long-term science is concerned with the research need to collect and keep records of the same measurements over long periods of time. At the same time it is necessary to attend to the short-term concerns of innovative site research and publications that are assessed at three year intervals and critical to success in securing the next increment of six year funding. LTER scientists are engaged in ongoing discussions about information management. A senior scientist observes the tensions between short-term and long-term issues and the implications for information management in providing support for science:

"some of the tension came from the difference between people wanting to use the resources for short-term business as usual, process oriented studies, versus maintaining a long-term program with a legacy of a database." (S)

Bringing the long-term view into an organization's vision and thus into its research plans introduces a complex set of long and short-term considerations. Figure 3 presents four distinct timeframes: the immediate (I), the short-term (II), the long-term (III), and a combination of short-term and long-term (IV). This quadrant approach is a simple heuristic for representing and understanding more fully research activities that occur simultaneously yet contribute to and hold value in differing timeframes. Empirical science with field and laboratory data collection entails work that often cannot be delayed so falls within the 'immediate' quadrant. The discovery work of a project is placed within the 'short-term' second quadrant since the way data is used will change as it is being taken and analyzed. The third quadrant represents the 'individual career' where a scientist integrates information/knowledge gained over multiple projects. The LTER appears in the fourth quadrant since research sites (and their datasets) are expected to integrate over multiple individual careers that have shared a common ecosystem focus, creating an infrastructure relevant to both short and long term work.

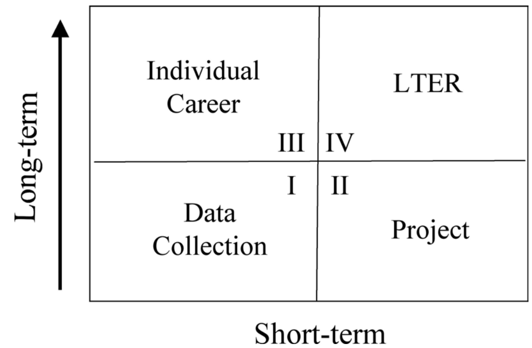

Figure 3. The interplay of short-term and long-term in science timeframes.

The founding vision of LTER places (or perhaps pushes) LTER participants into the fourth quadrant. Individuals from sites are often heard discussing strategies for balancing work that falls within the different quadrants. For instance, the mandate to share field data within two years of collection increases the pressures to provide support for short-term science (and also addresses a historical need to dedicate time in the short-term to make data useful in the long-term). Note, the LTER scientist's quote immediately above reflects the complexity of integrating long-term and short-term science: "short-term business as usual" refers to 'immediate' data collection and to 'short-term' project activities; "maintaining a long-term program with a legacy of a database" refers to fourth quadrant work.

4. Endogenous Social Mechanisms of Long-Term Information Management

There are two central mechanisms within LTER information management that particularly well reflect adaptation to the long-term way of thinking and federated way of operating: learning community and collaboration-in-design. They offer collective forums where a variety of changes: ecological, technological and organizational are addressed and incorporated into information management.

4.1. Learning Community

Although anchored by the realities and needs of their sites, information managers have created a network level forum, the Information Managers' Committee that forms their Community of Practice [21]. They bring with them an appreciation for local settings and diversity of local infrastructures, yet are united by the vision of long-term data for long-term science, sharing interests in technology and data issues that cross geographical and ecosystem bounds. Information managers learn together through dialogue and joint data and technology design projects as well as through attention to informing and training new members.

"Over the last two decades, the LTER information managers have taken the time that fosters an integrative, sustainable approach with technology, ensuring that we learn together." (IM*)

"and it's all like being mentored really by the overall group. So I see we are training folks." (IM)

Through annual meetings and community support for workshops, through a newsletter and listserves, through a common struggle with diversity and consensus, the group defines its community. Awareness of the long-term provides an opportunity to develop a community with continuity. It provides a safe place that is both a sounding board for ideas and an arena where information managers' voices can develop:

"It is also true that we have a group of people who have been doing this for a while. Seeing familiar faces when going to meetings and not having to rebuild every time, we have created trust." (IM)

"it's a place where people can let their hair down, be themselves, be natural. And it is safe to say things that demonstrate, I think failure is too strong, but where people have not been as successful, or disappointments. As soon as you are able to do that in a group, there is a bonding that occurs." (IM)

Participation in the LTER IM committee gives information managers a special point of view to gain an understanding of the LTER network level activities that 'rank and file' scientists may even lack.

"We have made a greater impact as a group. The network is not that cohesive as far as science goes, every site is very independent. Network is just that we have been funded by the same group and have figured out to do cross-site science, but it is the information managers of the network groups that have really created a network framework. We're the ones that get along better; scientists do not that well, there are big egos there. We make things happen… We are an incredible asset to the whole LTER program." (IM)

However, it also creates yet another tension for LTER information management between site and network level activities:

"where should the information managers' time be going, should it be only to support site activities, or should some of it be going to support network activities?" (IM)

4.2. Collaboration-in-Design

Collaborative design processes take place at site and network levels. The LTER organizational structure with site-based information management positions and venues for network level activities, such as an all-site information management committee and various temporary small focus groups, facilitate collaboration in design.

4.2.1. Site Level. Technology at a site may be introduced by site investigators or by information manager(s). The technologies that survive the test of time ultimately are integrated into a site's common information management framework by the information manager in coordination with the site scientists:

"There has to be that two way street between science and the techie." (IM)

As technology is ultimately evaluated against its value for ecological research, information managers are accustomed to thinking about and designing technologies embedded within their social and organizational contexts of scientific work and collaboration. Their practices reveal awareness that technologies need to be designed together with social mechanisms. Such socio-technical sensitivities are exemplified here by excerpts from an information manager's description of policies, forms and their flexible uses:

"they [scientists] need to complete a Notification of Research form. … Is [the data] something that LTER is going to want to archive? … they need to complete a Project Metadata Documentation form. … then you maintain communications with the researcher, and then … as they are generated, metadata for the individual datasets … Then the question is are those [data and metadata]] going to be entered through the control of information management's data entry system, or is it going to be handled by the researchers themselves." (IM)

Collaborative evaluation of technologies and technology development strategies go on formally and informally, as part of site reviews and everyday work:

"Each information manager and PI at each site is faced with articulating over time the 'why' or 'why not' of their particular development strategy at annual meetings, site reviews and proposal renewals. This 'discuss it or adopt it' phenomena becomes effective in contributing to a learning process both for those who think they know the answers as well as for those that think they don't know the answers. This … has created an opportunity to bridge technology with science in a manner unique to LTER." (IM*)

"The many facets of technical issues are revealed in the dialogues carried on between information managers and scientists." (IM)

Local approaches may differ in their emphasis and methods of technology development: some sites prefer to "keep it simple", some emphasize "data availability, data accessibility and possibilities for exploration" and some go after "automating systems and experimenting with new technologies".

"Our approach on the IT infrastructure is to keep things as simple as we can and still provide the services that we can." (IM)

"We put a lot of effort into making data accessible. My goal has been to have it be easily accessible. When the project started in the early years before we had a lot of the pieces in place, a typical scenario would be that a researcher would want a data set, they would have to come to the information manager, and then a day later have the dataset that they wanted, and I think that a system like that inhibits the exploration of data, to be able to ask questions as they occur, a really exploratory relationship with the database. We at our site have put a lot of effort to the accessibility and making sure that people could access the data directly and could do it in ways that they could get just the data they wanted to get." (IM)

"Philosophy for information management at our site is that the website is the filing cabinet for the site. … people are widely distributed. … My philosophy with regard to information management is to make it so that the investigators and students do as much as possible, so I spend more time setting up systems that actually maintaining them." (IM)

4.2.2. Network Level. Coming together in technology design at the network level, information managers are faced with the diversity of sites' approaches:

"A lot of the bottom-up characteristics are important for LTER information management. Ability to deal with heterogeneity not by limiting it but by dealing with it." (IM)

The tradition of collaborative undertakings in information systems development was initiated from the beginning by an LTER working group composed of scientists and information managers working together to outline the specifications for a minimum installation of technologies at each site, followed by, for instance, establishment and organization locally of Geographic Information System centers and of web sites. The development of a queriable all-site climate database as a cross-site activity by a subset of information managers in turn led to a broader conceptual design of a Network Information System [2]. Recently, the critical need for scientist-information manager dialogue on adoption of the concept of the ecological metadata language (EML) standard and metadata design and implementation prompted creation of a Network Information System (NIS) standing committee.

The LTER tradition of 'prototyping into consensus' is based on the idea of each module effort being led by an interested information manager who coordinates design, presentation, and communications with the LTER community throughout development and implementation. Interested sites are frequently recruited to serve as test users and ultimately code signers as the module becomes a *boundary object* [33] that is shared and discussed, redesigned and modified. Although only a few sites may participate originally, discussions during presentations or break out groups at annual meetings elicit the voices of the larger community.

In addition to posing a challenge for consensus building, the diversity engendered by the 24 research sites also provides an arena for sharing of ideas and learning from each other's experiences. Technological heterogeneity is not only allowed, it is also seen as one of the strengths of the LTER IM network. One information manager describes this phenomenon as a 'cherry picking octopus':

"one of the advantages with 24 sites is that there is always someone doing a major upgrading, they're out there looking for the solution that would work the best, they might find the solution through IM meetings, word of mouth, Databits, and it may also solve my problem (Boom!) … looking around what is going on within the network: 'do not spend so much time looking at your own stuff, that you never look at other's stuff'. I learn more by looking at other LTER sites, if I see they are doing something neat, I'll try to find out how they did it. Good things, bad things. …. There is always some site looking for something new, cherries are the good pieces of software ... 24 opportunities to find good ways, it needs to be an octopus as they need to be connected." (IM)

'Prototyping into consensus' and 'cherry picking octopus' both describe learning through collaboration-in-design in the course of network wide selection processes where each site is a 'laboratory' with its local specificities.

5. Ecological Information Management as 'Infrastructuring'

The LTER community infrastructure comprises organizational, technical and social elements. This paper has described the work of LTER information management that aims to ensure the longevity of the network's infrastructure by providing support for science, data and technology, by balancing between their systemic tensions, and by creating ways of learning and designing collaborative infrastructures that draw on the inherent characteristics of the networked organization.

We began by presenting the challenge of predictable decay of data over time and presented the graph that is widely used in LTER to motivate long-term data preservation (Figure 1). We put forward another heuristic to depict the innate tensions between short-term and long-term perspectives in ecological science and information management (Figure 3). The four quadrants shift our attention away from the individual scientists' datasets and careers and allow us to refocus on heterogeneous data ecologies and diversity of scientific collaborations together with the variety of challenges they pose for information management and infrastructure work that have to deal with the multiple time frames of today's long-term ecological science.

We have identified interrelated elements of support in an LTER information managers' infrastructure work and the inherent, systemic tensions between science, data and technology (Figure 2). The day-to-day articulation work of balancing tensions portrays how essential it is to integrate these elements in the infrastructure for long-term ecological science in an ongoing, continuous manner. Furthermore, Star & Bowker nudge our perception of infrastructure beyond an image of complex interconnected series of hidden physical objects by presenting us with a verb, i.e. 'to infrastructure' [32]. This changes the notion of infrastructure into an ongoing design process and creates a metaphor suggesting participation and co-construction as an integral constituent of ecological information management. We have described LTER's grass-roots approaches to 'collaboration-in-design' through the examples of 'prototyping into consensus' and 'cherry picking octopus' that account for joint learning and designing by drawing on and utilizing the character of the network of federated sites.

The LTER role of information management, emerging within the shelter of a long-term science community but subject to ongoing technological and organizational change, forms itself and is being formed between and in relation to the research elements it supports (as we have tried to depict in Figure 1). The information manager serves as a mediator between the elements of science, data and technology as well as between local practices within sites, the network level Community of Practice, and the larger world of technology development. In the field of ecological information management they are exceptionally well-positioned relative to the eight salient features of infrastructure defined by Star & Ruhleder [34] based on their rich understandings of the local (and network) conditions and the domains of ecological science and expertise 'from within' the LTER. For instance, they are able to account for the embeddedness of infrastructures in other social and technological structures; the transparency in invisibly supporting tasks; both the spatial and temporal reach or scope; and the taken-for-grantedness of artifacts and organizational arrangements as learned as part of membership; and they shape and are shaped by the conventions of practice [34]. From the practice-based standpoint of 'infrastructure betweenness', information managers recognize "[i]nfrastructures subtend complex ecologies: their design process should always be tentative, flexible and open" [32] as they wrestle with the inertia of the installed based in aligning new technologies with existing ones; and optimize for designing and fixing infrastructure in modular increments, not all at once or globally [34].

Concluding remarks

Long-term perspective and issues of infrastructure work remain largely unexplored in the field of information systems design. Despite this, and based on our case study of the LTER information management, we maintain that the issues discussed above are important and relevant to a surprising degree and extent, yet requiring 'infrastructural inversion' [8] to make them visible. Two vastly different studies in clearly distinctive fields of practice are introduced to illustrate resemblances and analogies in the concerns for long-term, maintenance and infrastructure work. Strauss et al. have studied the social organization of medical work. Similar to nurses tending to chronically ill patients [36, 37], information managers provide long-term 'data care'. In the field of architecture, Brand has written about the change processes that occur with buildings after they are designed and constructed. In accordance with what he observes about the extensive modification of buildings that takes place over time after architects' design [9], information managers engage in ongoing, longitudinal processes of infrastructure design, use, evaluation and maintenance.

Bringing the long-term perspective and infrastructure work to the forefront of priorities creates a need for change in approaches to information systems design and technology development, particularly large-scale collaboration infrastructures. The challenges are both theoretical pertaining to the very assumptions of central concepts as well as methodological reflecting the spatial and temporal expansion of what needs to be studied.

Theoretical challenges suggest moves from technologies as high-tech devices towards more inclusive conceptualizations of thickly interwoven socio-technical infrastructures encompassing mundane technologies and practices and information systems design from one-time technology development towards ongoing processes of infrastructuring. Together these openings challenge us to explore *designing for infrastructuring*, i.e. how to design for the blurring of borders between use and design, for ongoing changes, ease of maintenance, and tailoring of flexible and adaptable systems.

Methodological challenges that open are manifold, including, for example: understanding the relationships between 'routine work' and knowledge work; interest in 'boring infrastructure things' in addition to high-tech devices; studying extensive distributed networks in a way that accounts both for situated practices as well as large-scale collaborations; developing methods to study long-term collaborations and their developments over long periods of time; nurturing a research relationship with the studied community, for instance, sharing and learning with participants and creating appreciative interventions; and overcoming problematics with longitudinal studies such as raising funding and sustaining working relations within a multi-domain research team.

Short-term perspectives and practices along with recent developments - not all triumphant - in collaborative infrastructures, such as large partnerships and grid technologies, demonstrate the need for research directed at understanding and designing for the long-term perspective and infrastructure work.

Acknowledgements

We offer special recognition to the third member of our research team Geoffrey C. Bowker and to the LTER community. NSF Grants EIA-01-31958, DBI-01-11544 and OPP-02-17282 as well as Academy of Finland, Finnish Cultural Foundation and Scripps Institution of Oceanography support this work.

References

[1] P. Atkinson & M. Hammersley, Ethnography and Participant Observation. Handbook of Qualitative Research. N. K. Denzin & Y. S. Lincoln. Thousand Oaks, Ca., Sage Publications, Inc., 1994, pp. 248-261.

[2] K. S. Baker, B. J. Benson, D. L. Henshaw, D. Blodgett, J. H. Porter, & S. G. Stafford, Evolution of a Multisite Network Information System: The LTER Information Management Paradigm, BioScience, 50(11), 2000, pp. 963-978.

[3] K. S. Baker, G. C. Bowker & H. Karasti, Designing an Infrastructure for Heterogeneity in Ecosystem Data, Collaborators, and Organizations. Procs of the 2nd National Conference on Digital Government Research, 20-22 May 2002, Los Angeles, CA., Digital Government Research Center, pp. 141-144.

[4] E. Balka, Participatory Design in Women's Organizations: The Social Structure of Organizational Structure and the Gen-

dered nature of Expertise. Gender, Work and Organization 4(2), 1997, pp. 99-115.

[5] H. S. Becker, Tricks of the trade: How to think about your research while you're doing it. Chicago: University of Chicago Press, 1998.

[6] J. Blomberg., J. Giagomi, A. Mosher & P. Swenton-Wall, Ethnographic Field Methods and Their Relation to Design. Participatory Design: Principles and Practices. Schuler, D. & A. Namioka (eds.). Hillsdale, NJ, Lawrence Erlbaum Associates, 1993, pp. 123-155.

[7] J. Blomberg, L. Suchman, & R. Trigg, Reflections on a Work-Oriented Design project. Human-Computer Interaction 11, 1996, pp. 237-265.

[8] G. C. Bowker, Science on the Run: Information Management and Industrial Geophysics at Schlumberger, 1920-1940. Cambridge, MA: MIT Press, 1994.

[9] S. Brand, The Clock of the Long Now: Time and Responsibility: The Ideas Behind the World's Slowest Computer. Basic Books, New York, 1997.

[10] J. T. Callahan, Long-term Ecological Research. BioScience, 34, 1984, pp. 363-367.

[11] Y. Dittrich, S. Eriksén & C. Hansson, PD in the Wild: Evolving Practices of Design in Use. In Binder, T., J. Gregory & I. Wagner (Eds.) Procs of the Participatory Design Conference. Malmö, Sweden, 23-25 June 2002. CPSR, Palo Alto, CA., 2002.

[12] Y. Engeström, Developmental Work Research: Reconstructing Expertise through Expansive Learning. In Nurminen, M. I. & G. R. S. Weir (eds): Human Jobs and Computer Interfaces. Amsterdam, North-Holland, 1991.

[13] J. F. Franklin, C. S. Bledsoe, & J. T. Callahan, Contributions of the Long-term Ecological Research Program: An Expanded Network of Scientists, Sites, and Programs Can Provide Crucial Comparative Analyses. BioScience, 40(7), 1990, pp. 509-523.

[14] J. Greenbaum & M. Kyng, Eds. Design at work. Lawrence Erlbaum, Hillsdale, 1991.

[15] J. Gosz, Ecology Challenged? Who? Why? Where is This Headed? Ecosystems (2), 1999, pp. 475-481.

[16] D. Heath, Bodies, Antibodies, and Modest Interventions. In Downey, G. L. & J. Dumit (eds.) Cyborgs and Citadels: Anthropological Interventions in Emerging Sciences and Technologies. School of American Research, Advanced Seminar Series, SAR press. pp 67-82, 1997.

[17] J. E. Hobbie, S. R. Carpenter, N. B. Grimm, J. R. Gosz, & T. R. Seastedt, The US Long Term Ecological Research Program. BioScience 53(1), 2003, pp. 21-32.

[18] J. A. Holstein, & J. F. Gubrium, The Active Interview. Thousand Oaks, Sage Publications Inc., 1995.

[19] H. Karasti, Increasing Sensitivity Towards Everyday Work Practice in System Design. PhD Thesis. Department of Information Processing Science, University of Oulu, 2001.

[20] H. Karasti, K. S. Baker & G. C. Bowker, Ecological Storytelling and Collaborative Scientific Activities. SIGGROUP Bulletin, 2003, in press.

[21] J. Lave & E. Wenger, Situated Learning: Legitimate Peripheral Participation, Cambridge University Press, 1991.

[22] G. E. Likens, Long-Term Studies in Ecology: Approaches and Alternatives. Springer-Verlag, New York, 1989.

[23] R. Markussen, Politics of Intervention in Design: Feminist Reflections on the Scandinavian Tradition. AI & Society, 10(2), 1996, pp. 127-141.

[24] W. K. Michener & J. W. Brunt, Ecological Data: Design Management and Processing. Blackwell Science, 2000.

[25] W. K. Michener, J. W. Brunt, J. J. Helly, T. B. Kirchner, & S. G. Stafford, Nongeospatial Metadata for the Ecological Sciences. Ecological Applications, 7 (1), 1997, pp. 330-342.

[26] W. K. Michener, J. W. Brunt, & S. G. Stafford, Environmental information management and analysis: Ecosystem to global scales. Taylor & Francis, Bristol, PA, 1994.

[27] B. A. Nardi & V. L. O'Day, Information ecologies : Using technology with heart. Cambridge, MA: MIT Press, 1999.

[28] B. Rogers, Ed. Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information, 4-7 November 1996, Lake Buena Vista, FL. Ann Arbor, MI: Environmental Research Institute of Michigan (ERIM), 1996.

[29] D. Schuler & A. Namioka Eds. Participatory Design: Principles and Practices. Hillsdale, NJ, Lawrence Erlbaum Associates, 1993.

[30] D. Silverman, Ed. Qualitative Research: Theory, Method and Practice. London, Thousand Oaks, New Delhi, Sage Publications, 1997.

[31] S. L. Star, The Ethnography of Infrastructure. American Behavioral Scientist, 43(3), 1999, pp. 377-391.

[32] S. L. Star & G. C. Bowker, How to Infrastructure in The Handbook of New Media. L.A. Lievrouw & S.L. Livingstone (eds). SAGE Publications, London, 2002, pp. 151-162.

[33] S. L. Star & J. R. Griesemer, Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science, 19, 1989, pp. 387-420.

[34] S. L. Star & K. Ruhleder, Steps Towards an Ecology of Infrastructure: Design and Access for Large Information Spaces. Information Systems Research, 7(1), 1996, pp. 111-134.

[35] S. L. Star & A. Strauss, Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. Computer-Supported Cooperative Work: The Journal of Collaborative Computing, 8, 1999, pp. 9-30.

[36] A. Strauss, S. Fagerhaugh, B. Suczek & C. Wiener, Social Organization of Medical Work. University of Chicago Press, Chicago, 1985.

[37] A. Strauss & G. Glaser, Chronic Illness and the Quality of Life. Saint Louis, The C. V. Mosby Company, 1975.

[38] L. Suchman, Supporting articulation work: Aspects of a feminist practice of technology production. In A. Adam et al. (eds.) Women, Work and Computerization: Breaking old boundaries - building new forms. New York: IFIP/North-Holland, 1994, pp. 7-22.

[39] L. Suchman, J. Blomberg, J. Orr & R. Trigg, Reconstructing technologies as social practice. American Behavioral Scientist, 43(3), 1999, pp. 392-408.

[40] M. Vehviläinen, Women defining their information technology - Struggles for textual subjectivity in office workers' study circle. European Journal of Women Studies, Vol. 1, 1994, pp. 71-91.