Building Environmental Information Systems: Myths and Interdisciplinary Lessons

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

Karen I. Stocks
San Diego Supercomputer Center
University of California, San Diego
kstocks@sdsc.edu

Abstract

With databases and information systems playing an increasing role in large scientific research projects, there is a growing stake in understanding how to design a useful information system and in broadening our understanding of what constitutes the scientific work involved in building these systems. Both experience and theory indicate that non-technical considerations, such as management and communication structures, are as important as technical decisions in system development. We examine four case examples of environmental information system development: the Ocean Biogeographic Information System, the Long Term Ecological Research Network, the California Cooperative Oceanic Fisheries Investigation, and SeamountsOnline. We then draw from a wide interdisciplinary literature, including science and technology studies and social informatics, to identify common myths and misconceptions about system development and consider alternatives. Our goal is both to provide a set of concrete models and a theoretical foundation useful to other projects.

1. Introduction

Some of the most critical science and management challenges facing marine ecology today are large scale, long-term, and/or interdisciplinary. Examples include predicting climate change impacts, describing and understanding hotspots of biological diversity, and managing ecosystems instead of individual fisheries stocks. By nature, these topics cannot be addressed within a single research project but instead must integrate data and knowledge across many studies and sources. Such broadly relevant topics can be supported by development of federated data systems.

Information systems (e.g. data access, web services, and communication technologies) are increasingly important tools for ocean and environmental research (e.g. Bisby 2000, Porter et al. 2004). Individual researchers and laboratories are making models and results available through websites. Diverse institutions such as museums and fisheries management organizations are creating systems for accessing their data resources remotely; National Ocean Data Centers and large-scale oceanographic programs such as the Joint Global Ocean Flux Study (JGOFS), Global Ocean Ecosystems Dynamics (GLOBEC), and Global Ocean Observing System (GOOS) are investing in information infrastructure.

As these efforts increase, there is a growing stake in understanding how to design and implement an effective information system. Developing these systems is neither easy nor straightforward (Kaplan & Seebeck 2001). They tend to be delivered late, over budget, or lack functionality. Rather than focusing solely on technological aspects, we consider the organization and social dynamics of these projects in order to identify characteristics that foster success, and also present working examples of mechanisms that projects are using to promote these characteristics.

Conclusions are presented not as abstract management generalities but as the lessons learned from existing environmental information system development projects that provide four different, concrete models of organization. In our combined experience in reviewing oceanographic, biological, and environmental information system development proposals from the US and Europe, it is common to see project management plans that include, for example, timelines, regular meetings, and steering committees. But we have rarely seen proposals that plan for design or enactment activities, cite social science or management literature, or make use of experience of past projects other than their own.

In the second part of the paper we present a series of information system design myths, along with alternative perspectives drawn from the literature. The goal is not to be comprehensive but rather provide an entryway into a few topics in this widely relevant field.

A much broader set of literature than that covered in the myths section can inform information system implementation. The developing field of information system research is beginning to identify relationships between information systems, methods, and values that reflect particular ethics and epistemiologies (Boland & Hirschheim 1987, Lyytinen 1987, Boland & Tenaski 1995, Hirshheim et al. 1996). Less common are discussions of system failures (Rocheleau 1997, Sauer et al. 1997, Lyytinen & Robey 1999). Many are emphasizing the information system as a process rather than as a product, a perspective further supported in science and technology studies and social informatics (Becker 1998, Greenbaum & Kyng 1991).

Organizational and management theory bring integrative approaches and resources to community science, with research on design characteristics and trade-offs, complex and adaptive systems, information flows and the nature of knowledge. Studies of collaboratories (Finholt & Olson 1997, Finholt 2004) are beginning to document the range and complexity of distributed collaboration, and system scientists are exploring holistic theories and practical approaches to information systems and communities (Senge 1990, Checkland & Holwell 1998). Research using the metaphors of an "information ecology" and the "social life of information" highlight the interdependence of the diverse system factors involved and foreground the emerging concepts of knowledge and how its management might unfold (Davenport 1997, Nardi & O'Day 1999, Brown & Duguid 1991).

In the few published case studies of environmental or oceanographic information system development, several have highlighted that adequate funding (at least in the short-term) and broad talent and technical expertise are far from sufficient to guarantee that a project meets its goals (e.g. Stonebraker 1994, Thorley & Trathan 1994, NRC 1995, Star & Ruhleder 1996, Tomlinson et al. 1998, Hale et al. 2003, Schnase et al. 2003). Within the US, reports from multiple National Science Foundation program areas (computer science, engineering, biology, environmental studies and social sciences) on planning cyberinfrastructure have all recognized the importance of "the social" (Hayes et al. 1995, Atkins 2003, Futrell 2003, Woolley 2005, Berman and Brady 2005). This is not to imply that technical factors are unimportant, but, to date, technical design and implementation have received substantially more planning and attention in project development than enactment and articulation work.

Table 1. Summary of Example Project Characteristics - as of 2005.

	OBIS	LTER	CalCOFI	SeamountsOnline
Age	5 yrs	25 yrs	56 yrs	4 yrs
Participation	10 regional nodes; 22 organizations providing data	26 sites; 1200 participants; 26 information managers; 95 organizational affiliations	3 institutions; 500 individuals on conference email list	1 investigator, 2 part time staff
Management Structure	International Committee (steering); Management Committee (regional node coordination); Technical Committee, Editorial Board (quality assurance)	Executive Committee; Information Management Executive Standing Committee; Network Information System Advisory Committee	CalCOFI Committee (steering)	single investigator
Main Data Type(s)	Biological (species occurrences)	Ecological interactions (species, communities, physical)	Biological (species occurrences), physical	Biological (species occurrences), habitat descriptions
Scientific Focus	Biogeography	Ecosystems	Fisheries Management	Seamount Ecology
Geographic Scope of Data	Global	US	California	Global
Geographic Scope of Management	International	National	State	Local
Serving Method	Distributed	Distributed & Centralized	Centralized & Distributed	Centralized

2. Case Examples and Lessons Learned

The examples below are necessarily brief so illustrative case examples rather than full case studies.

Ocean Biogeographic Information System. OBIS is an international federation of distributed data providers (http://www.iobis.org; Grassle & Stocks 2000, Zhang & Grassle 2003; Costello et al. 2005). The scope of the data (at least to date) is narrowly defined: records of particular species being observed or collected at particular locations. OBIS arose from a series of community workshops organized by the Census of Marine Life in the late 1990's to identify projects that would most advance the field of marine ecology. Organizationally, OBIS is managed by 1) an International Committee that acts as a steering committee to guide overall direction (and acted as the sole management body early in OBIS' development); 2) a Management Committee that runs the operation of approximately 10 regional nodes around the world; and 3) a Technical Committee that guides technical development. It also has a secretariat headed by the chair of the International Committee and portal development staff at Rutgers University. An editorial board is under development to provide quality assurance through expert taxonomic and technical advice and to assist with identifying quality datasets and tools.

Long Term Ecological Research program. LTER is a US federation of twenty-six independent research teams, each with dozens of researchers focused on understanding the structure and interactions of a particular terrestrial, coastal, or aquatic ecosystem (http://lternet.edu; Hobbie et al. 2003). LTER involves more than 1800 scientists, students, and staff collecting long-term data on ecological phenomena. Initiated in 1980 with funding from the National Science Foundation for six sites, it actively promotes regionalization of local efforts, cross-site collaborative and interdisciplinary research. LTER is a loosely coupled network with a coordinating committee and an affiliated international network (ILTER, http://www.ilternet.edu). An information manager working at each site facilitates documentation, preservation, and access to heterogeneous datasets (Karasti and Baker, 2004). They provide support for site science, local technology implementation, and network activities (Baker et al. 2000).

California Cooperative Oceanic Fisheries Investigation. CalCOFI is a more than 55 year-old study of the southern California coastal ocean, making it one of the longest-running ocean monitoring studies in existence (http://calcofi.org; NRC 1995, Ohman & Venrick 2003; CalCOFI, 1988). Its purpose is to gather data needed to improve regional fisheries management. Originally conceived in response to the collapse of the California sardine industry, the program now focuses on understanding and modeling the effects of biological and physical variability on the ecosystem. The core of the program is quarterly cruises during which a standard suite of physical, chemical, and biological samples are collected on a standard grid of stations. CalCOFI cruises and data management are carried out jointly by the Southwest Fisheries Science Center of the US National Marine Fisheries Service, Scripps Institution of Oceanography, and other regional governmental, scientific, and conservation organizations. Datasets are made available online.

SeamountsOnline. SeamountsOnline is a small, centralized database gathering information on species that have been recorded from seamounts (http://seamounts.sdsc.edu; Stocks 2004). It draws data from the literature and from electronic datasets provided directly by scientists after seamount expeditions. It is essentially a one-person project, with some part-time data entry and programming assistance. SeamountsOnline is currently expanding to become the information system for the Census of Marine Life on Seamounts, an international research program.

Table 2. Characteristics that facilitate information system development project outcomes. Although not exhaustive, the list represents elements for which specific mechanisms have been developed within the four case examples discussed.

Goal: Specific interdisciplinary communication
Mechanisms:
- create mixed technical and oceanographic/environmental expertise on decision-making bodies
- support cross-trained project personnel
- establish interdisciplinary "theme" working groups

Goal: Strong community and participant support
Mechanisms:
- give wide public recognition of people and groups who contribute in a variety of ways
- allocate seed funding to prompt local initiatives
- develop consensus, representational decision making
- develop a "learning environment" to foster ongoing professional development

Goal: A useful system
Mechanisms:
- create clear project goals in collaboration with user communities
- interact with user communities throughout design, development deployment, and enactment cycle
- plan explicitly for flexibility and change making use of articulation and enactment in system design
- consider community assumptions and myths

These four project examples span a range of complexity and age, and the authors have in-depth experience with each. Baker is information manager with two LTER sites. Stocks has been the vice-chair of the OBIS International Committee and leads the SeamountsOnline project. Both have worked with CalCOFI.

2.1 Characteristics of Success

Through comparing and contrasting these different information system development projects (Table 1), there emerge characteristics that have facilitated project outcomes and deliverables (Table 2). They are presented below, with examples of the specific mechanisms used to support each characteristic.

Interdisciplinary Communication

Successful information system development generally requires, at minimum, collaboration between the scientists in the domain for which the system is being built (such as oceanographers, ecologists, or taxonomists) and information technology personnel. Larger projects can have several scientific domains represented, several types of information technology experts (such as data managers, technologists, computer scientists, and web developers), and can expand to include experts such as education and outreach specialists, instrument developers, and science writers. We hypothesize that projects that recognize the interdependence of the different project facets devote more time (a valuable and limited resource) to establishing and maintaining good communication between domains. When each participant understands constraints, goals, and timelines of others, more informed decisions emerge.

Among the example projects, several mechanisms are employed to foster interdisciplinary communication. In OBIS, there is overlap between the committees: several technical experts sit on the overall International and Management Committees, and several scientists from oceanography and biology sit on the Technical Committee. This approach yields project members who can translate for their domain. For example, the biological scientists on the Technical Committee participated in the discussions on implementing standards and thus fully understood why a standard schema was critical to developing an integrated system and was worth the extra effort needed to implement it. These translators were then able to explain these needs and decisions to the other biologists who were developing datasets. Standardization efforts consequently met with less resistance. In the LTER program, each site has an information manager working as part of a team of investigators. From these local positions, data management practices and new technology are introduced gradually but persistently.

In LTER, OBIS, and SeamountsOnline, multi-perspective or “mediator” members perform critical roles. These are individuals whose primary training is, for example, in physics or biology, but who later undertook substantial formal or informal training in information technology. These roles provide translation and communication work so may not fit within traditional scientific measures of productivity (i.e. data collected, papers published, or code written). These roles illustrate a more general phenomenon of the emergence of new specialties – system architects, information managers, information officers, content managers – valuable to interdisciplinary projects.

Incentives for Participation

A critical element of project sustainability is the ability to retain personnel and engender support and investment from the larger community. This is particularly true as information system development projects scale to multi-site, national and international efforts: while some top-down coordination is needed, bottom-up support is essential. The case examples below demonstrate a variety of reward types.

Public Recognition. OBIS creates official titles for people who have a variety of ongoing roles with OBIS, and is developing an editorial board for those who have less consistent contributions but who nevertheless provide expertise in particular areas such as taxonomic nomenclature and identifying available datasets. Often, this level of participation is not a formal commitment, and is frequently overlooked by projects. In OBIS, the editorial board members will be prominently featured on the website. Similarly, when the European Register of Marine Species published a journal article describing its system, it devoted a full page to listing the many (unpaid) taxonomists who contributed to reviewing name lists, and a half-page to listing collaborating organizations (Costello 2000). Note that maintaining a reputation for scientific quality is essential for recognition to be a reward.

Seed funding. Small amounts of seed funding may provide critical focus and can efficiently leverage activities. Some seed projects promote local initiatives and hence encourage local leadership; some may shelter innovate high-risk prototypes at low cost. As the European Register of Marine Species project demonstrates, many academics have some discretionary time they choose to devote to projects considered important or interesting (Costello 2000). OBIS has found that a mini-grant can effectively create an individual priority. LTER has arranged seed funding at times, e.g. supporting small fellowships for site participation in prototype designs that could scale to network or community use. This approach produced the LTER climate, hydrological, and site description databases (Henshaw et al. 1998, Baker et al. 2002).

Open Decision Making. For projects where participation is voluntary, individuals are most willing to invest time and effort when they have a voice in decisions and directions, in contrast to projects that are led from “on high.” We define open decision making as allowing for the input of diverse voices. This input may take the form of continuing two-way communications and/or effective representation. Several of the example projects consider reaching broad consensus a goal for their decision-making process. OBIS has created a steering committee that is a microcosm of the communities with which they interact. Many of the projects that are the largest or oldest data providers to OBIS have seats on the committee, as do individuals who have devoted enormous (and mainly unfunded) time to OBIS development. To date, taking a formal vote has not been needed; topics are discussed until a consensus decision emerges. The LTER data management group similarly uses dialogue: products and procedures are adopted by the group after continuing discussion and modification. This is a slow process but often leads to the voluntary adoption of common tools and standards. If done otherwise, standardization can be either strongly resisted or quietly resented. The LTER data management group has further developed an elected executive committee with rotating members. Reports of monthly committee conference calls are distributed through a brief summary email to the data management group at large and reviewed at annual meetings of all site representatives.

As projects scale in size, not every participant can be on the decision-making body, but leadership can still be representative. None of the example projects have a steering committee solely made up of scientific experts in the domain (e.g. established taxonomists when building a taxonomic information system). A group with diversity, including, for example, technologists and information managers, brings new perspectives to leadership, avoids an "us and them" dichotomy, and holds the potential to prompt innovation (Brown and Duguid 1991, Lievrouw & Livingstone 2002). More time is needed, however, to consider heterogeneous perspectives and discordant views. Themed working groups are used within LTER as a mechanism for lateral communication crossing site and/or disciplinary boundaries. A working group on ecological disturbance or habitat heterogeneity, for example, may draw ecologists, students, information managers, educators, technicians, technologists and program representatives from multiple sites.

Learning Environment

Both oceanography and information technology are rapidly changing fields. Their practitioners need and want to learn about new methods, technologies and theories. Large projects can provide forums through which participants can keep each other up-to-date and seek advice. Such a "Community of Practice" (Lave & Wenger 1991) provides an efficient and collegial place for ongoing professional exchange and learning. The LTER program has successfully built a community of practice for data managers through support for annual meetings, by organizing an informal newsletter for opinions and discussion on technology (*Databits*), and by actively supporting participation in working groups. LTER data managers, who are geographically distributed across sites, develop a sense of camaraderie from annual meetings. A willingness to devote time to advising and helping fellow managers is created – essentially an informal capacity-building mechanism across sites that vary in their technical expertise and approaches. A concrete demonstration of this unofficial mentoring is the set of papers self-organized by the LTER information management community for two sessions at an international conference. Several papers had a co-author for whom this was the first experience with publication preparation, professional manuscript review, or public presentation – all unique planning, learning, and synthesis activities.

Conducting thoughtful, repeated joint assessments as a group is another method to build a shared understanding. Assessment involves gathering feedback in order to gain a measure of understanding of the current state of an endeavor and to adjust plans in relation to goals. Assessments take a variety of forms. An exercise as simple as diagrammatically capturing a discussion of community relations on a whiteboard makes larger contexts visible and elicits information on local semantics, worldviews, and unspoken priorities. It takes practice and a shift in perspective to recognize such work as a learning opportunity rather than another administrative burden. The field of education, facing analogous challenges to prompt and measure learning, has explored and developed concepts and tools of interest for effective assessment (Wiggins 1998).

System Usefulness

The CalCOFI program was designed to provide data that fisheries managers use to predict fish production. In essence, it gets the right data to the right people in the right way. SeamountsOnline has a small-scale approach to meeting a similar end: the primary developer is a seamount biogeographer using the system to do research. In addition, the project collaborates with a small set of other seamount ecologists also using data in the system, receiving frequent input on desired expansions and problems. Within LTER, site data systems have developed and transformed over time as use by scientists creates new ideas on system function. Originally tasked with creating a data repository, changes in technology, networking, and local scientists' expectations have created calls for data access and, more recently, for queriable data and, ultimately, interoperable datasets.

One lesson from these examples is the need for clearly stated, focused vision and goals with broad input from a defined user community. Input from the user community is critical – project managers may think they know the needs and interests of a particular community, but often there is unarticulated work and tacit knowledge (Choo 1995, Davenport & Prusak 1998, Whitley 2000). The development of an Ecological Metadata Language (EML) is a case in point (Jones et al. 2001). This project, led by a group of ecological and technical experts, employed an informal requirements specification phase to meet the needs of multiple communities, including LTER. The tools for working with EML fit the needs of some data providers though not all the diverse communities and did not easily accommodate legacy systems. Subsequently, development continues on EML, a 'working standard' with software fixes identified and informed in part by an ongoing multi-year, multi-tiered enactment period (Millerand and Bowker, forthcoming).

Flexibility in System Design

It is difficult to get everything right from the start when designing a new system so the ability to re-design and adapt is critical. In LTER, an informal system of multiple prototyping has emerged as a result of the independence of sites. Development at one location starts, for example, with a bibliographic citation application. This informs participants via regular information manager meetings, where in-progress reports are given. Efforts may coalesce, or other sites may suggest elements important for their local needs. Over time, one approach may become a standard across more sites. The strategy of multiple initial prototypes promotes dialogue, innovation, and community understanding. A focus on design, enactment, and articulation work promotes community development (Baker and Millerand, 2007).

In contrast, the SeamountsOnline project achieves design flexibility through being small and centralized. The developer is able to try different configurations sequentially on a development site. Because changes do not propagate through a larger system, there is less associated overhead. While this approach does not fully scale, the concept of modularizing system development, of initially growing modules on a smaller scale, can foster flexibility even in larger projects.

3. Interdisciplinary Perspectives: Myths and Assumptions

At first glance there appear to be a host of common myths regarding technology and technology development in the case examples, e.g. "a paperless office is imminent" and "the more data the better". Working to identify assumptions provides impetus to examine more closely issues contradicted by everyday experiences such as "once a system works it will continue to work" and "computers make your work easier". Information technology assumptions are less typically recognized perhaps because IT development experience is relatively new. Below, we expand on several of today's statements of mythic stature. Each is presented in our words, but we draw on the literature to explain concepts and suggest alternative views. In some cases, these statements and their alternatives may seem obvious and well known though difficult to change in practice.

In contemporary work arenas, are we able to recognize assumptions or to choose our myths in order to facilitate information system development? The power of myths has been described as self-renewing and transformative: "A myth is something that never happened but is always happening" (Campbell, 1988). Considered by some as unfulfilled expectations, myths may also be viewed as overarching themes.

3.1 Myth 1: System development is linear, and involves design, implementation, and then production. Star (1991) speaks to the heterogeneity of knowledge and its representation, which ultimately precludes "getting it right" in a linear fashion. In part, this is due to complex interdependencies and feedbacks that exist in a system with many, interacting components. These dependencies are such that no one piece can be finished before the next is started. This phenomenon is in part a result of the constant change: new technologies develop, funding resources shift, and collaborations coalesce. In part because design decisions involve negotiations between what is to be represented and how best to do so, there is no single 'right' path.

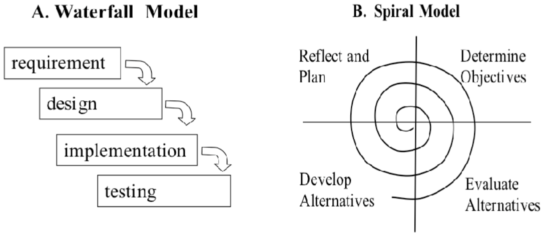


Figure 1. Two models of information system development: linear waterfall model and iterative spiral (Boehm 1986).

In response to linear development scenarios, such as the waterfall model (Fig 1a.), alternative models have been proposed. One of these is the iterative, spiral model of Boehm (1986) (Fig. 1b). In this approach, a project prepares for repeated cycles of planning, determining objectives, evaluating development options, and implementing selected options. A key difference between the two models is that a project using an iterative approach would be able to plan, create, and sustain mechanisms for periodic re-assessment. This carries an administrative cost but is more likely to achieve an effective outcome. Additional models of information system development and their trade-offs are offered by the fields of system design (Jirotka & Goguen 1994, Bowker et al. 1997), ecological design (Odum 1998, Orr 2002), and meta-design (Fischer 2003). Participatory design brings expected end users into the whole development process not just evaluating the end product, in order to benefit from their experience and understanding throughout

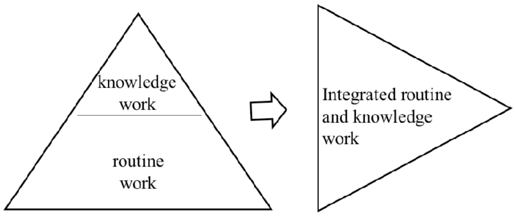
(Schuler & Namioka 1993; Letondal & Mackay 2004). As with spiral design, it requires the project be able to adapt to suggestions for change throughout the design process. It has the benefit of identifying gaps or problems early, when there is less cost to changing approach than if identified in a near-final, beta version. As LTER has found, having information managers closely connected to local user community of scientists – working jointly on long-term site science, and providing ongoing demonstrations of the system's use – fosters critical, often tacit, bidirectional communication about users' needs and technical considerations that guides local development.

Star and Ruhleder (1996) further underscore the need for designing flexible systems and maintaining ties with user communities. As users, in our case ecological and oceanographic researchers and program managers, learn to recognize and to articulate data needs, new system requirements and priorities are generated. For example, in LTER the original vision of a technical repository for field measurements grew over time to a delivery mechanism that can selectively order, group, and document data. In CalCOFI the vision of supporting fisheries management as well as local research has broadened to include providing data for a wider audience of interested users.

3.2 Myth 2: Data management is just technical work. In a traditional, hierarchical view of the workplace, high-level leadership performs the planning and management tasks that require conceptual development, contextual knowledge, and management skills. Staff then carry out the project production work. Blomberg et al. (1996) illustrate this approach with a triangular arrangement whereby a base of routine work supports knowledge work at the top (Fig. 2a). Technology increasingly has led to a "flattening" of workplaces, as a blend of skills and judgment are needed by all project personnel (Fig. 2b). Indeed, such changing conditions suggest shifts in how roles are viewed or develop over time (Boland & Tenkasi 1995, Zemke et al. 2000, Karasti & Baker 2004, Lamb & Davidson 2005, Friedman, 2006).

The lesson for information system development projects is that all members, from programmers to project administrators, are daily making a myriad of design requirement changes, management negotiations, contextual integrations, process optimizations, and alignment repairs that significantly impact a project's overall development and function. In flattening the hierarchy, not only is expertise throughout the system recognized, but those individuals from the upper levels traditionally divorced from the actual function of the project are brought into the development process previously considered production or ‘routine work’.

Another perspective is provided by action researchers and ethnographers who highlight "invisible work" and "silent voices" (Argyris et al. 1985, Star & Strauss 1999, Karasti 2001). Invisible work includes everyday work practices with their messy work-arounds, undocumented negotiations, and ongoing communication efforts. Silent voices – which often include technical mediators, information facilitators, and social or organizational managers – are participants whose efforts are invisible because of their very functionality. For example, few notice the effort involved in the routine work of system maintenance until the system goes down. The nature of these organizational operations is such that the perspectives and knowledge of maintenance workers are often not drawn upon until a breakdown occurs, whereupon putting out the fire is a reactive task rather than a proactive endeavor.



a)Traditional hierarchical structure b)Integrated lateral structure

Figure 2. Workplace organizational models (after Blomberg et al. 1996). a) A traditional hierarchical structure, with knowledge work isolated in a top strata; b) An integrated lateral structure showing blended routine and knowledge work. The lateral organizational model is consistent with the flattening of information arenas and the emergence of new workplace roles (e.g. system architects, information managers, chief information officers, content managers).

3.3 Myth 3: Technology is objective. Another myth is that technology is a distinct entity, an object that, once created and explained, can be adopted by an organization. To put it colloquially, this is the “you build it and then they use it” model. An alternative view recognizes that enacting technology is a dynamic and continuing process: not only do users adapt to new technologies (Fountain 2001), but new technologies can be informed by existing user practices. Thus, both the user's work practices and the technology itself co-evolve. So if being widely adopted and integrated into a user community is the goal of an information system project, planning a means to continue assessment, articulation, and adjustment (after traditional beta testing when the product is generally considered finished) can be critical.

3.4 Myth 4: Good communication involves just talking and listening carefully. In practice, good communication can be fostered with an awareness of factors outside of the message content itself: the communicator's background and frame of reference as well as the unstated expectations, tensions, and assumptions that underlie perspectives. For example, some participants may focus on details of systems and their characteristics, while others see the overall system as the frame for discussion. Mutual consensus is difficult when one person is speaking of the forest and the other discussing the trees. Could this be one reason the social sciences literature documenting direct observations of technology development case studies is largely untapped by projects actively developing information systems? System developers are largely focused on building the system in order to understand the natural world; cultural or sociological aspects are background concerns raised only when they must be considered to understand the natural ecosystem or the IT system. In contrast, with a figure-ground reversal, social scientists see cultural ramifications as the central focus with system development and the natural world as parts of the human world (Redman 1999). Note, the focus, interests, and even styles of writing differ between the two perspectives, creating barriers to communication.

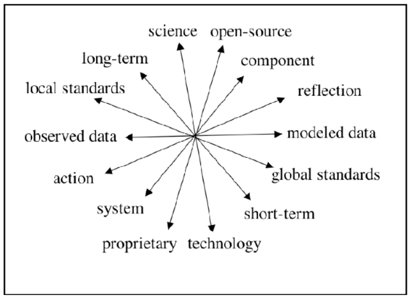


Figure 3. An example of a tension diagram, indicating axes along which perspectives and frames of reference vary.

Creating a simple diagram of the tensions within a group is one tool for understanding their impact on a project. Figure 3 provides an example, but the particular axes will vary by project. Each tension, for example between the need for short-term output and long-term planning, or between maintaining local flexibility and adopting community standards, represents a spectrum along which choices are made.

Identifying often unspoken tensions can prompt productive discussion, negotiation, and understanding, if not consensus. When global, more generic standards are then implemented locally, they require adaptation to fit them back to local data and needs (Star & Greismer 1989). This process of recontextualization can represent a local increase in work and loss of flexibility, thus contributing to a tension between local and global standards development.

4. Planning for the Long-Term: Environmental Information Systems and Community Design

The examples and myths presented above are limited but serve as reminders of the multitude of non-technical considerations in information system development. These issues become more important as information system projects, and science in general, scale up from individual or small group products to international and interdisciplinary collaborations. The breadth and variety of perspectives required for understanding how to create effective information ecologies requires a kind of interdisciplinary partnership not traditionally supported within existing organizational arrangements. The extent and complexity of collaborative work, whether interdisciplinary, local, regional, or global, is often taken for granted but, when recognized, can become part of a planning process for designing both environmental information systems and communities.

Planning and budgeting for collaboration and design is critical with scaled projects to ensure sufficient time and resources for the processes of exchange and assessment. We recognize that in an environment where time and money are inherently limiting elements, allocating these resources to goals as intangible as "fostering communication" and "developing a learning community" does not seem a priority. Yet the lesson has long been reported: "project management is a serious problem" (Stonebraker 1994).

In today's research arena, there is an awareness that ecosystem-scale questions require interdisciplinary collaboration from biologists, chemists, physical oceanographers, technologists, and other specialists. We suggest that in stepping back to view the information system and its community of users, we benefit from stretching partnerships and arranging new types of interdisciplinary collaborations involving a larger system perspective by involving, for instance, information, social and environmental scientists together with technology, organization and management experts. Rather than a burdensome requirement, this represents an opportunity for new research that holds the potential to embed information systems design approaches as an integral part of community development, contributing to both short and long-term community building processes.

Over the last decade the question "how to build a data system" has broadened to "how to build an information system that works" but emerges today as "how to build a community" – a community with diverse voices and proccesses organized to identify, address, and respond to the technical, social, and organizational issues arising in the ongoing process of information system design.

This work is supported by CalCOFI, OBIS, LTER, & SeamountsOnline (NSF/OPP#04-05069, NSF/OPP #02-17282, NSF/OCE #04-17616, NSF/OCE #03-40839) in addition to the Comparative Interoperability Project (CIP NSF/SBE/SES #04-33369) and Ocean Informatics at Scripps Institution of Oceanography. This work benefits from early conversations with H.Karasti and G.C.Bowker and helpful comments provided by P.Brewin, M.Costello, B.Hentschel, E. Venrick, F.Millerand, and D.Ribes.

References

Argyris C, Putnam R, Smith DM (1985) Action science: concepts, methods and skills for research, and intervention. Jossey-Bass, San Francisco

Atkins D and the NSF Blue-Ribbon Advisory Panel on Cyberinfrastructure (2003) NSF-AP Revolutionizing science and engineering through cyberinfrastructure. http://www.communitytechnology.org/nsf_ci_report

Baker KS, Benson BJ, Henshaw DL, Blodgett D, Porter JH, Stafford SG (2000) Evolution of a multisite network information system: the LTER information management paradigm. BioScience 50(11):963-978

Baker KS, Brunt J, Blankman D (2002) Organizational informatics: site description directories for research networks. Proceedings of the 6th World Multi-Conf on Systematics, Cybernetics, and Informatics, Orlando

Baker KS, Millerand F (2007) Articulation work supporting information infrastructure design: coordination, categorization, and assessment in practice. Proceedings of the 40th Hawaii International Conference on System Sciences (HICSS), 3-6 January, Big Island, Hawaii, IEEE Computer Society, New Brunswick, NJ

Berman F, Brady H (2005). Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences.

Bisby FA (2000) The quiet revolution: biodiversity and informatics on the internet. Science 289: 2309-2312

Becker HS (1998) Tricks of the trade: how to think about your research while you're doing it. Univ of Chicago Press

Blomberg J, Suchman L, Trigg RH (1996) Reflections on a work-oriented design project. HCI 11:237-265

Boehm A (1986) Spiral model of software development and enhancement. ACM SIGSOFT Software Engr Notes 11(4):14-24

Boland RJ, Tenkasi RV (1995) Perspective making and perspective taking in communities of knowing. Organizational Science 6(4): 350-372

Boland RJ, Hirschheim RA (1987) Critical issues in information systems research. John Wiley and Sons, NY

Bowker, GC, Star SL, Turner W, Gasser L (1997) Social science, technical systems, and cooperative work. Lawrence Erlbaum Associates, London

Brown JS, Duguid P (1991) Organizational learning and communities-of-practice: Toward a unified view of working, learning and innovation. Org Science 2:40-57

CalCOFI, 1988. Symposium of the CalCOFI Conference, 1987. California Cooperative Oceanic Fisheries Investigations, CalCOFI Reports, Volume 29.

Campbell, J 1988. The Power of Myth. Anchor Books, NY.

Checkland P, Holwell S (1998) Information, systems, and information system: making sense of the field. John Wiley & Sons, Chichester

Choo CW (1995) Information management for the intelligent organization: roles and implications for the information professions. Digital Libraries Conference, March 1995.

Costello MJ (2000) Developing species information systems: the European Register of Marine Species. Oceanography 13(3):48-55

Costello MJ, Grassle JF, Zhang Y, Stocks K, Vanden Berge E. 2005. Online mapping of marine species. MarBEF Newsletter No. 2. Ecological Consultancy Services (EcoServe) Ltd, Dublin.

Davenport TH (1997) Information ecology: mastering the information and knowledge environment. OxfordPress, NY

Davenport TH, Prusak L (1998) Working knowledge: how organizations manage what they know. Harvard Business School Press, Boston

Finholt JT, Olson GA (1997) From laboratories to collaboratories: a new organizational form for scientific collaboration. Psychological Science 8(1):28-35

Finholt TA (2004) Collaboratories. Annual Review of Information Science and Technology 36:73-107

Fischer G (2003) Meta-design: beyond user-centered and participatory design. Proceedings of the 10th International Conf on Human-Computer Interaction. Crete, Greece

Fountain J (2001) Building the virtual state: information technology and institutional change. Brooking Institution Press, Washington, DC

Friedman T (2006) The world is flat: a brief history of the twenty-first century. Farrar, Straus and Giroux, New York.

Futrell J and the AC-ERE (2003) Environmental cyberinfrastructure (ECI): Tools for the study of complex environmental systems. AC-ERE 1, May 2003

Grassle JF, Stocks K (2000) A Global Ocean Biogeographical Information System (OBIS) for the Census of Marine Life. Oceanography 12(3):12-14.

Greenbaum J, Kyng M (1991) Design at work. Lawrence Erlbaum, Hillsdale

Hayes EF, Bement AL, Hennessy J, Ingram J, Pitts N, Young PR, Kollman PA, Vernon MK, White AB, Wulf WA, Voigt R (1995) Report of the task force on the future of the NSF supercomputer centers program.

Hale SS, Miglarese AH, Bradley MP, Belton TJ, Cooper LD, Frame MT, Friel CA, Harwell LM, King RE, Michener WK, Nicolson DT, Peterjohn BG (2003) Managing troubled data: coastal data partnerships smooth data integration. Environ Monit Assess 81:133-148

Henshaw DL, Stubbs M, Benson BJ, Baker KS, Blodgett D, Porter JH (1998) Climate database project: a strategy for improving information access across research sites. In: Michener W, Porter J, Stafford S (eds) Data and Information Management in the ecological Sciences: A Resource Guide. Albuquerque: LTER Network Office, University of New Mexico. P 123-127

Hirschheim R, Klein HK, Lyytinen K (1996) Exploring the intellectual structure of information systems development: a social action theoretic analysis. Accounting, Management and Information Technology 6a:1-64

Hobbie JE, Carpenter SR, Grimm NB, Gosz JR, Seastedt TR (2003) The US long term ecological research program. BioScience 53(2):21-32

Jirotka M, Goguen J (1994) Requirements engineering: social and technical issues. Academic Press, London

Jones MB, Berkley C, Bojilova J, Schildhauer M (2001) Managing scientific metadata. IEEE Internet Computing 5(5):59-68

Kaplan S, Seebeck L (2001) Harnessing complexity in CSCW. Proceedings of the Seventh European Conference on Computer Supported Cooperative Work, Kluwer Academic Publishers, Dordrecht, p 359-397

Karasti H, Baker KS (2004) Infrastructuring for the long-term: ecological information management. Proceeding of the Hawaii International Conference for System Science, IEEE Computer Society Press, Big Island, Hawaii

Karasti H (2001) Increasing sensitivity towards everyday work practice in system design. PhD thesis, Department of Information Processing Science, Univ of Oulu, Finland

Lamb R, Davidson L (2005) Information and communication technology challenges to scientific professional identity. The Information Society 15:1-24

Lave L, Wenger E (1991) Situated learning: legitimate peripheral participation. Cambridge Univ Press, Cambridge

Letondal C, Mackay WE (2004) Participatory programming and the scope of mutual responsibility: balancing scientific, design and software commitment. Proceedings Participatory Design Conference 2004. Toronto, Canada

Lievrouw LA, Livingstone S (2002) The handbook of new media. SAGE Publications, London

Lyytinen K (1987) A taxonomic perspective of information systems development: theoretical constructs and recommendations. In: Boland RJ Jr, Hirschheim RA (eds) Critical issues in information systems research. John Wiley and Sons, New York, p 3-41

Lyyntinen K, Robey D (1999) Learning failure in information systems development. Info Sys J 9:85-101

Millerand, F, Bowker GC (forthcoming). Metadata Standards. Trajectories and Enactment in the Life of an Ontology. In SLStar and .Lampland (Eds), Formalizing Practices: Reckoning with Standards, Numbers and Models in Science and Everyday Life

Nardi BA, O'Day VL (1999) Information ecologies: using technology with heart. MIT Press, Cambridge

NRC (1995) Finding the forest in the trees: the challenge of combining diverse environmental data. Committee for a Pilot Study on Database Interfaces, National Academy Press, Washington DC

Odum E (1998) Ecological vignettes: ecological approaches to dealing with human predicaments. Harwood Academic Publishers, Amsterdam

Ohman MD, Venrick EL (2003) CalCOFI in a changing ocean. Oceanography 16(3):76-85.

Orr DW (2002) The nature of design: ecology, culture, and human intention. Oxford University Press, Oxford

Rocheleau B (1997) Governmental information system problems and failures: a preliminary review. Northern Illinois University, Division of Public Administration, Public Administration and Management: An Interactive Journal. ISSN 087-0091

Porter DE, Small T, White D, Fletcher M, Norman A, Swain D, Friedmann J (2004) Data management in support of environmental monitoring, research, and coastal management. Journal of Coastal Research 45: 9-16

Redman CL (1999) Human dimensions of ecosystem studies. Ecosystems 2: 296-198

Sauer C, Southon G, Dampney CNG (1997) Fit, failure and the house of horrors: toward a configuration theory of IS project failure. Proceedings of the 18th International Conference on Information Systems. Atlanta, p 349-366

Schnase JL, Cushing J, Frame M, Frondorf A, Landis E, Maier D, Silberschatz A (2003) Information technology challenges of biodiversity and ecosystems informatics. Information Systems 28(4): 339-345

Schuler D, Namioka A (1993) Participatory design: principles and practices. Lawrence Erlbaum, Hillsdale

Senge P (1990) The fifth discipline: the art and practice of the learning organization. Currency Doubleday, New York

Star SL (1991) Power, technology and the phenomenology of conventions: on being allergic to onions. In: Law J (ed) A sociology of monsters: essays on power, technology and domination. Routledge, London, p 26-56

Star SL, Griesemer JR (1989) Institutional ecology, translations, and boundary objects: amateurs and professionals in Berkeley's museum of vertebrate zoology. 1907-1939. Social Studies of Science 19: 387-420

Star SL, Ruhleder K (1996) Steps toward an ecology of infrastructure: design and access for large information spaces. Information Systems Research 7(1):111-134

Star SL, Strauss A (1999) Layers of silence, arenas of voice: the ecology of visible and invisible work. Computer Supported Cooperative Work 8:9-30

Stocks KI (2004) SeamountsOnline: An online resource for data on the biodiversity of seamounts. In: Morato T, Pauly D (eds) Seamounts: Biodiversity and Fisheries. Fisheries Centre Research Reports 12(5):13-16

Stonebraker M (1994) Sequoia 2000 – a reflection on the first three years. Seventh International Working Conference on Scientific and Statistical Database Management, 28-30 September 1994, Charlottesville, VA. IEEE Computer Society Press, p 108-116

Thorley M, Trathan P (1994) The history of the BIOMASS data centre and lessons learned during its lifetime. In: El Sayed SZ (ed) Southern Ocean Ecology: The BIOMASS Perspective. Cambridge University Press, NY, p 313-322

Tomlinson KL, Spasser MA, Sanchez JA, Schnase JL (1998) Managing cognitive overload in the flora of North America project. Proc of 31st Hawaii Int Conf on System Sciences, IEEE Computer Society Press, Big Island, Hawaii

Whitley EA (2000) Tacit and explicit knowledge: conceptual confusion around the commodification of knowledge. In: Brown L (ed) Knowledge management: concepts and controversies, p 62-64

Wiggins G (1998) Educative assessment: designing assessments to inform and improve student performance. John Wiley, San Francisco

Woolley J and Subcommittee on 21st Century Biology (2005) Building a cyberinfrastructure for the biological sciences (CIBIO): 2005 and beyond: a roadmap for consolidation and exponentiation

Zemke R, Raines C, Filipczak B (2000) Generations at work. American Management Association, New York

Zhang Y, Grassle JF (2003) A portal for the ocean biogeographic information system. Oceanologica Acta 25:193