Evolution of a Multisite Network Information System: The LTER Information Management Paradigm

KAREN S. BAKER, BARBARA J. BENSON, DON L. HENSHAW, DARRELL BLODGETT, JOHN H. PORTER, AND SUSAN G. STAFFORD

THE LTER NETWORK INFORMATION SYSTEM (NIS) IS A COOPERATIVE, FEDERATED DATABASE SYSTEM SUPPORTING MULTI-INVESTIGATOR, MULTISITE ECOSYSTEM STUDIES CREATED IN COLLABORATION WITH LOCAL INFORMATION MANAGEMENT PERSONNEL

Pelagic polar marine, temperate coniferous forest, urban watershed, coastal estuary, eastern deciduous forest, tropical rain forest, tallgrass prairie—these are just a few of the ecosystems represented in the 24 sites of the Long-Term Ecological Research (LTER) Network (Franklin et al. 1990). By combining information from the diverse ecosystems represented in the LTER network, participants have a unique opportunity for large-scale investigations of complex phenomena like climate change, biodiversity, soil dynamics, and environmental policy.

In 1996, to facilitate data exchange and synthesis from its multiple sites, LTER launched the LTER Network Information System (NIS), based on an independent site and central office organizational infrastructure. Other organizational partnerships provide examples of earlier efforts also focused on communications and data sharing: the Worm Community System, the Flora of North America Project (FNAP), and the Organization of Biological Field Stations (OBFS). The Worm Community System was developed—before Internet connectivity became available—as a collaborative software environment through which its 1400 widely dispersed researchers could share information on the genetics, behavior, and biology of the soil nematode species *Caenorhabditis elegans*. Insight into the complexity of a network structure was gained through attention to the design and analysis of both the system's structure and usability (Star and Ruhleder 1996). The FNAP system, in contrast, was developed with Internet technology. The FNAP, with a goal of identifying and cataloging all plant species, uses online technology to create

Karen S. Baker (e-mail: kbaker@ucsd.edu) is Palmer LTER Data Manager, University of California, San Diego (UCSD), Scripps Institution of Oceanography, La Jolla, CA 92093-0218. Barbara J. Benson (e-mail: bjbenson@facstaff.wisc.edu) is Associate Scientist and North Temperate Lakes LTER Data Manager, University of Wisconsin–Madison, Center for Limnology, Madison, WI 53706-1492. Donald L. Henshaw (e-mail: dhenshaw@fs.fed.us) is Andrews Forest LTER Data Manager, USDA Forest Service, Pacific NW Research Station, Corvallis, OR 97331. Darrell Blodgett (e-mail: blodgett@taiga.lter.alaska.edu) is Bonanza Creek LTER Data Manager, University of Alaska, Forest Soils Laboratory, Fairbanks, AK 99775-0740. John H. Porter (e-mail: jhp7e@virginia.edu) is Research Assistant Professor and Virginia Coast Reserve LTER Data Manager, University of Virginia, Department of Environmental Sciences, Charlottesville, VA 22904. Susan G. Stafford (e-mail: stafford@cnr.colostate.edu) is Professor and Department Head, Department of Forest Sciences, College of Natural Resources, Colorado State University, Ft. Collins, CO 80523-1470.

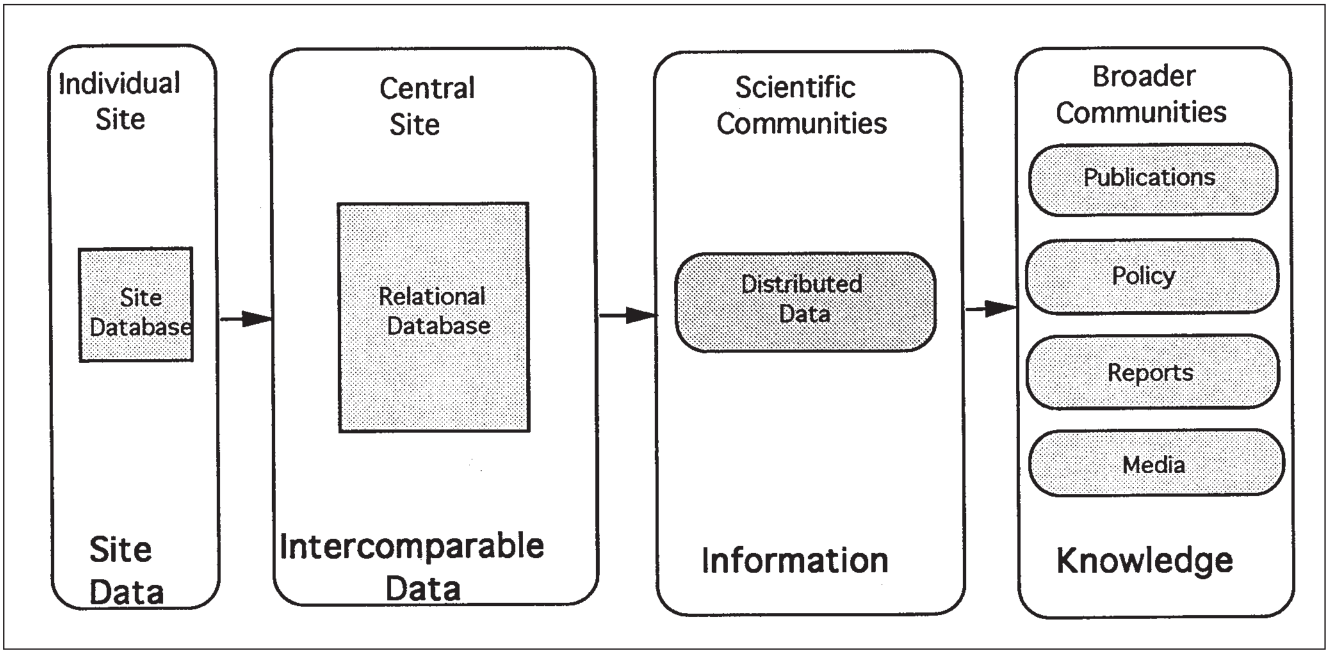

Figure 1. Information flow. Raw data travel through several transformations before providing useful knowledge for a wide range of communities. This information flow begins with the collection of data at an individual site. When data sets are structured to be intercomparable, a central (often relational) database can be established. Scientific communities then draw data from this center in formats suited to their analyses. Broader communities turn to interpretations of such analyses to obtain knowledge on pertinent topics. Such synthesis may take various forms: publications for scientists, policy statements for government needs, reports for business needs, and media presentations for the general public.

an electronic community of authors who can write, edit, review, and publish cooperatively (Tomlinson et al. 1998). OBFS, representing a community of more than 150 biological field station members with diverse environmental concerns, has introduced plans for an electronic medium to support community partnerships and pilot projects (Stanford and McKee 1999).

This article describes the evolution of the LTER NIS, highlighting important components and presenting specific examples of software modules that permit the integration of research data from different sites. We also describe how the LTER NIS work style fosters intersite communication, technology transfer, and an interactive, participatory approach to information management. The LTER NIS is a cooperative effort that allows for local site independence and a flexible modular design.

Today's biologists and ecologists are reaping a bounty of new data from their research projects, thanks to new instrumentation and data storage capabilities. But organizing these data and related information poses a challenge: The data must be easily available on a long-term basis; they must be stored, along with their supporting documentation, so that quality is assured and difficulties in interpretation are minimal; and the data sets must be structured so that they are comparable (Gurtz 1986, Porter and Callahan 1994, NRC 1995, Stafford et al. 1996). Meeting these goals is essential for the success of long-term research, multidisciplinary projects, and cross-site studies—leading funding agencies to focus increasingly on how data are managed. The quality of LTER sites' data management techniques, for example, is among the criteria the National Science Foundation (NSF) uses in reviewing individual sites.

How data are managed

LTER data flow from individual sites through several stages to their ultimate destination (e.g., publication, policy report, or media outlet; Figure 1). Management of the data stresses their accessibility and integration (Stafford et al. 1994).

The data manager is the facilitator, translator, and converter (Stonebraker 1994, Bowker et al. 1997, Kies et al. 1998) of these data, and must know both the science and the information technology. A data manager must effectively integrate data and metadata (documentation about the data) across studies (see box page 965). The data manager is also the attendant for a new vocabulary, which can either launch or limits efforts (see box page 966), describing changes that are introduced by technology (Kay 1998, Kies et al. 1998).

The site data manager's second focus involves developing a balance or flexibility that lets the system respond to priorities. Balance built into the system must be able to address specific scientist versus general-user requirements, short-term versus long-term data handling, and local versus generic design methods. The data manager's participation in the design of the data handling procedures offers a distinct advantage because he or she is familiar with the specific needs of the local scientific community. The manager's understanding of explicit and implicit site information helps establish the balances among competing needs when resources are limited. Without flexibility and balance, data sets and information systems can go unused.

The partnership between science and data management has influenced studies in areas such as ecosystem variability (Kratz et al. 1995), prairie climatology (Strebel et al. 1994), vegetation (Riera et al. 1998), climate (Greenland and Swift 1991), and ice (Magnuson et al. in press). Development of an ice phenology database by the Lake Ice Analysis Group (LIAG) for over 750 lakes and rivers of the Northern Hemisphere provides a specific example of how data management contributes to data structuring and intersite research. This database, used to explore the usefulness of ice data as a paleoclimate indicator, was developed for a 1996 workshop with 28 international participants organized by the North Temperate Lakes (NTL) LTER site. Before the workshop, project scientists worked closely with the NTL information management staff to identify variables of scientific interest as well as formats for construction of standard data sets. These often time-consuming activities, requiring close collaboration between scientists and data managers, were critical in database design and workshop preparation. Discussions defining the data set stimulated development of important descriptive information about the data, which increased the value of individual data sets by providing a broader context. Thus, a data set contribution, which might consist of only lake name and annual date of freeze, was augmented for submission to the ice phenology database with standard codes specifying the continent, water body type, latitude, longitude, elevation, mean depth, shore line length, population of largest city on lake (or river), and ancillary weather data.

Whether "ice-covered" meant total or partial cover was discussed, as well as how thawing and refreezing affected the definition of duration of cover. Instructions were developed to distinguish between an observation that the lake did not freeze and a case in which no observation was made.

In the 1980s, gathering of data by the lead scientists often required visits to locations involved in LTER intersite efforts. For the ice phenology workshop, specifications were electronically distributed to participants beforehand. Data submissions usually involved communication between contributors and information management personnel. Not all research groups could submit data electronically (e.g., some faxed it) or in the specified standard format (e.g., the month written as "Jan" or roman "I" rather than numeric "01"). In such cases, information management staff members transferred the data into standardized formats, and data managers also checked facts to aid scientists' quality assurance efforts.

By the time of the workshop, a significant portion of the data was available for on-site analysis. In earlier LTER intersite efforts, scientists spent a significant amount of time integrating the diverse data sets into Excel spreadsheets, which then supported subsequent analyses. In contrast, the ice phenology workshop data sets were incorporated into a relational database. Data views provided an initial point of departure, and the availability of the database at the workshop supported the generation of important research questions as the scientists interacted.

For instance, researchers were able to readily select subsets of water bodies with lengths of record or spatial locations best suited to their particular questions, thus maximizing individual productivity. A policy limiting data access addressed researcher concerns on data sharing, and after the workshop the data were accessible to LIAG members via the World Wide Web, promoting continued updates and communication. The diverse studies possible from this single database resulted in a series of papers presented at the 27th Congress of the International Association of Theoretical and Applied Limnology in 1998 (Magnuson et al. in press). The preparations required to make data comparable, deal with multiple formats, structure data in a database form, and provide access to it resulted in a rich body of scientific publication and a robust database available for future study. Without an adequate system and personnel for managing data, the magnitude of the effort needed to deal with large and complex data sets can be a substantial barrier to intersite research.

Data Manager or Information Manager?

The term "data manager" developed to describe an individual dealing with specific data sets. A data manager may prepare, calibrate, document, and assure the quality of raw data. Additionally, the data manager may develop techniques and formats for exchanging data with a central site and for eventual distribution and/or preservation. Current research often requires data sets to be integrated from multiple projects and sources into intercomparable groups of data sets. In fact, the term "information manager" may better describe the individual dealing with the broader aspects of data. Scientists, responsible for collecting and interpreting data, benefit from working in partnership with data and information managers, to ensure optimal data and information availability.

Data management developments. Scientific progress and efficiency increase when data from colleagues, libraries, and archives are available electronically because individual research units need less time to access the data and thus have more time for analysis and synthesis (Ingersoll et al. 1997). However, availability does not necessarily imply utility. Useful data must be of known quality, well described by metadata. With organizations spending increasing resources on data collection, there is growing attention to the development and implementation of archival metadata standards (Gross et al. 1995). For instance, the Federal Geographic Data Committee (FGDC 1995) created metadata standards as part of the National Information Infrastructure efforts. National and international master directory efforts are beginning to catalog the location of data sets. The Global Change Master Directory, supported by the National Aeronautics and Space Administration (NASA), is one example of an earth science database directory.

Traditionally, peer-reviewed publications have been the science community's data archive, but publications preserve only a subset of the authors' data. Electronically available complete data sets allow future researchers to address alternative hypotheses as well as unanticipated questions. Both the high cost of properly documenting data for electronic archives and the lack of a reward structure for supporting these efforts are significant deterrents to making data available (Michener et al. 1997). Incentives for openly exchanging data among ecologists, such as venues for publishing data sets and peer-reviewed papers about data sets, will help reduce this resistance (Olson et al. 1996).

In addition to publication and traditional short-term data storage by individuals, there are a variety of other data archival methods (Olson and McCord 1998). National scientific data archives emphasize data set identification, lineage, storage, and redistribution. NASA's Earth Observing System Data Information System, which includes interconnected Distributed Active Archive Centers, is one example of a national archive system. The National Oceanic and Atmospheric Administration's National Environmental Satellite, Data, and Information Services is another example with interconnected centers including the National Oceanographic Data Center and the National Climatic Data Center.

In addition, individual research sites with information systems are becoming stewards of long-term data repositories. Local site storage has the advantage of keeping the data resources close to the scientists who use them. Data stored locally are readily available for updates as analysis and synthesis proceed. A local repository promotes continuing dialogue between researcher and data manager about database entry, quality assurance, and data analysis. It also may promote inclusion, in addition to data sets, of full data lineage such as literature references and published texts as well as informal documents such as project plans, proposals, newsletters, brochures, and preliminary results.

Coming to Terms with Vocabulary	
Metadata:	text describing data, e.g. location, methods, units
Raw data:	data as recorded directly in the laboratory or field
Calibrated data:	data transformed to have relevant units
Intercomparable data:	data put into an agreed upon format in order to be compatible with other related data sets
Data management:	system for keeping, handling and making available data
Information management:	system making available data pertinent to a selected topic, often derived from intercomparable data
Knowledge management:	system making available interrelated information
Module:	an independent piece of an information system with a specific function
Prototype:	test module
Extensible:	able to be modified or augmented
Filter:	program code that reads data and converts to data which has a different form
In-reach:	process of providing information and knowledge within your immediate scientific community
Out-reach:	process of providing information and knowledge to a broader community

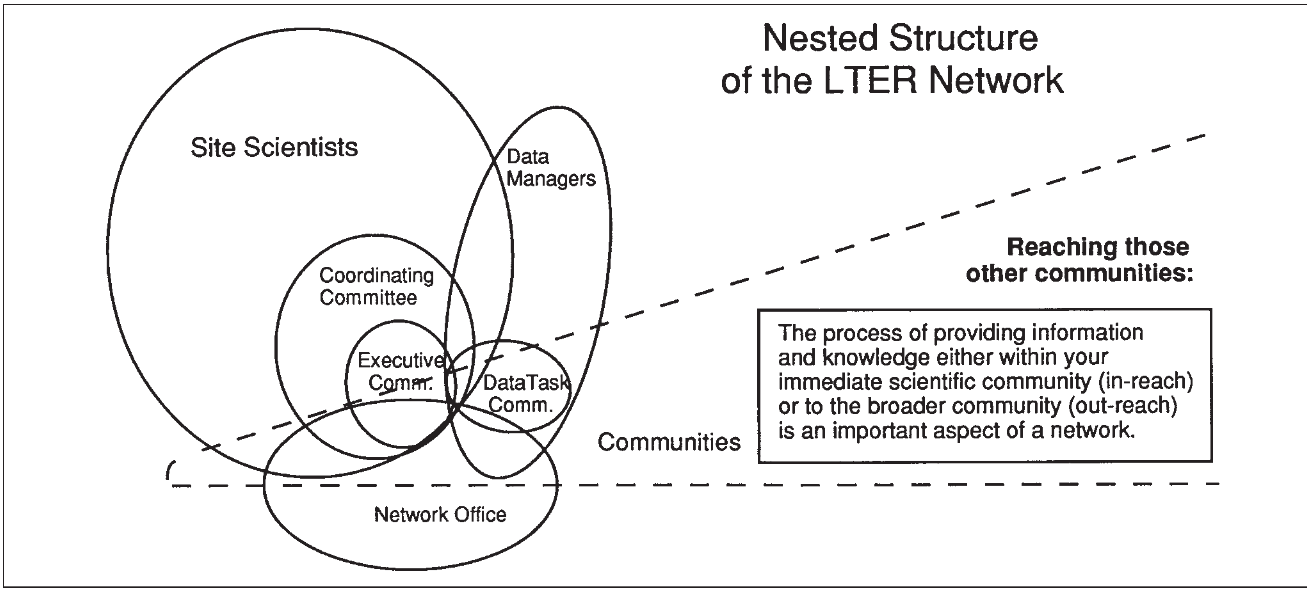

Figure 2. LTER Network—Background and nested structure. The Long-Term Ecological Research Network was established in 1981 by the National Science Foundation as a collaborative research program bringing together individual sites mandated to create a data legacy for future researchers. The original vision—coordination into a network of a group of sites addressing broad ecosystem issues—remains in place today (Callahan 1984). Each site consists of a multi-investigator research team exploring a specific ecosystem. These sites include forest, prairie, desert, tundra, agricultural, lake, river, wetland, estuary, and marine environments—all linked into a network through core study areas, cross-site studies, and ongoing collaborations (Franklin 1994, Franklin et al. 1990) as well as through electronic connectivity and information management (Brunt et al. 1990).

The LTER Network organization may be visualized as nested activities represented by ovals: individual site research with a data management component and the network with cross-site research coordination. Each site is represented on a coordinating committee, from which an executive committee is elected, and on a data manager committee, which also elects a coordinating committee. The central driver for these nested activities remains the site-funded ecosystem research, while cross-site activities have been supported by establishment of an LTER network office that maintains support personnel and facilities such as a computing center. In-reach activities may take the form of both intrasite and intersite communications. The central support mechanism facilitates in-reach of scientists within the LTER community and out-reach to communities such as other scientific networks, technical experts, government agencies, and the public. These broader communities (enclosed by dotted lines) represent another dimension of overlap. The unions of the loops represent joint LTER activities. For instance, there is coordination of LTER data management with LTER scientific pursuits through data manager representation at both the annual scientist coordinating committee meetings and cross-site committee meetings. An annual data manager committee meeting provides an essential forum for information exchange among the data managers themselves.

Other countries also have begun to focus on data management and to establish national data sharing networks. Two examples are the Environmental Change Network in the United Kingdom (Cuthbertson 1993, Lane 1997) and the Ecological Monitoring and Assessment Network in Canada (Canadian Global Change Program 1995). With a growing recognition of the need to coordinate networks on a larger scale (IGBP 1990), the International LTER network promotes development of national ecological networks and provides an international forum for global networking as well (Franklin 1994).

Data management paradigms. A variety of paradigms address issues associated with the processing and synthesis of data. Regardless of approach, they are all faced with the challenges of scaling up (Brown 1994, Stafford et al. 1994, Robbins 1995, Olson 1999) and dealing with both broader spatial data scales (regional and global) as well as expanded temporal ones (from short-term time series to paleo records). Scaling up in a network arena means interfacing successfully with other sites and networks. A number of efforts have addressed the interface of field science with computer science (Gurtz 1986, Michener et al. 1994, Stonebraker 1994, Strebel et al. 1994, Thorley and Trathan 1994, Robbins 1995). The challenges associated with data synthesis have been addressed by partnerships focusing on discipline or theme (e.g., ecology, Gosz 1994; oceanography, Flierl et al. 1992; worm community, Kouzes et al. 1996, Schatz 1993, Star and Ruhleder 1996; remote sensing in modeling, Vande Castle 1991; prairie climatology, Strebel et al. 1998); on region (e.g., Antarctica; ICAIR 1993); on task (e.g., management; Bannon 1996); and on public policy

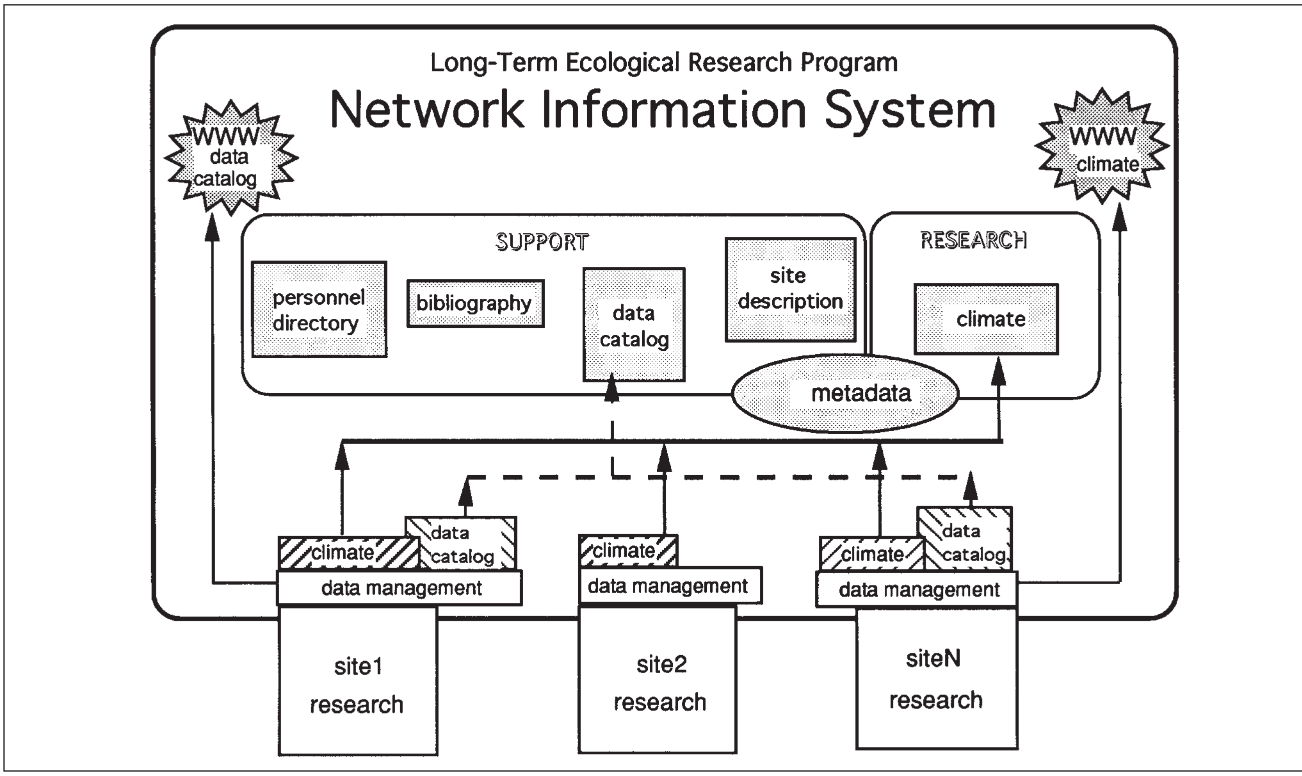

Figure 3. NIS. The Long-Term Ecological Research Network Information System facilitates information flow from individual research databases through composite relational databases into appropriately formatted out-reach uses by broader communities. NIS rests on the foundation of research done at individual sites. This example shows how data management interfaces and assembles data into individual data modules such as the climate database or the data catalog. Using appropriate filters, data managers arrange for these data to be transmitted to the NIS central climate database and the network data catalog. The Web presentation may migrate eventually from its original development location at a research site to the network office.

(e.g., Antarctic ocean resources, Commission for the Conservation of Antarctic Marine Living Resources; Knox 1988).

Another pivotal issue in multicontributor databases is the balance of local versus centralized management since data management functions may be designed and implemented at either local or central sites. Some multinational efforts, such as the Global Terrestrial Observing System (Heal et al. 1993) and the Antarctic Data Centers in support of the Antarctic Treaty (SCAR/COMNAP 1996, Agnew 1997), have proposed a networked structure of sites with relatively centralized management to coordinate and distribute data. In this context, an overview of the multinational Antarctic Biological Investigation of Marine Antarctic Systems and Stocks (BIOMASS) data program (Thorley and Trathan 1994) stressed, in hindsight, the need for better integration of science and data management efforts. Such an integration is the focus of the National Atmospheric Deposition Program (Aubertin et al. 1990), a US multiagency group funded to oversee the National Trends Network. The program's individual sites are responsible for data collection, using well-defined procedures and field form completion, while a central laboratory is responsible for quality assurance and data analysis.

The general organizational structure and the specific data management schema of biology projects are often influenced by a combination of historical development, user community factors, and scientific goals. In a top-down approach to data management protocols, an initial unifying element can be specified. For example, the Human Genome Project (HGP; Pearson and Soll 1991, Robbins 1992, Robbins et al. 1995) unified part of the molecular biology community structurally through the adoption of a single relational database software application to address the well-defined gene sequencing databases. Because it could project the type and amount of data involved before it started, the HGP carefully planned an electronic information system using or creating the necessary technology. The Sequoia Project (Dozier 1992) is another example of initial project unification in which coordination centered on the design, extension, and adoption of an interactive working environment for earth science projects using specific object-oriented module concepts. Elements of the common work environment flourished past the project end, in contrast with the earth

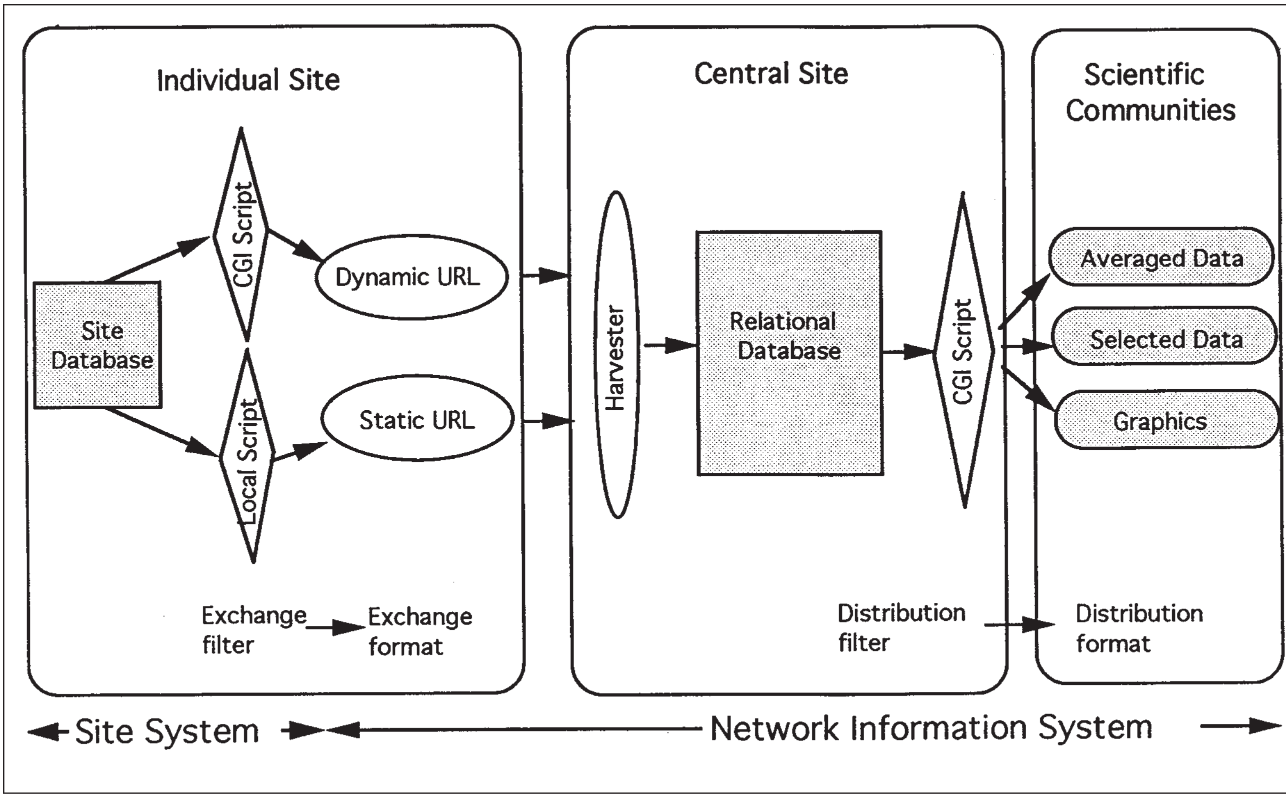

Figure 4. Module schema. A generic module schema illustrates how data originating at a local site can be posted to either a dynamic or static URL using exchange filters in an agreed-upon format. Given a set of site-identified URLs, a harvester from a central site can gather data and place it in a relational database. The data can then be displayed for the general scientific community using distribution filters in a variety of formats.

science collaborations that remained dispersed (Stonebraker 1994).

In many situations, an abundance of research questions are not well defined at the onset. Platt (1988) suggests that in the life sciences, the individuality of living things responding continuously to past and current events creates an abundance of research questions that simply cannot be well posed. For example, the LTER encompasses ecological databases in which the variable definitions themselves are under development both within the differing site system environments and at the network level. The definition of as basic a variable as productivity requires discussion when considering how productivity is measured in a rain forest as compared with a desert or an ocean. Group or project organizational structures will differ from the top-down model when the research variables and questions are under development.

Structure influences the behavior and expectations of research participants by defining how decisions are made, in a continuum ranging from autocratic or top-down to

Strategic Vision for the Long-Term Ecological Research Network Information System

Statement of Purpose

Our goal is to promote ecological science by fostering the synergy of information systems and scientific research.

Vision Statements

- Pursue information systems development and implementation from the context of ecological research needs
- Conduct information management in a nested context of site, research network, national, and international levels
- Emphasize the timely and effective transformation of data into information and the ease of access to that information
- Ensure the long-term preservation and availability of information
- Ensure appropriate information system development through information management research
- Develop human resources necessary for the continuing evolution of LTER information systems

Table 1. Components of the LTER NIS. The NIS includes operational modules located at the LTER network office in addition to prototype modules that may be hosted at either individual sites or the network office.

Modules	Description	Status
Support		
Personnel Directory	Telephone, address, and e-mail for LTER researchers	Operational
Electronic Mailing Lists	Supports topic-based electronic mailing lists for LTER researchers	Operational
Bibliography	Free-text searches of LTER site bibliographies	Under revision
Data Catalog	Free text and keyword access to site datasets	Operational
Site Description Directory	Basic site characteristics	Prototype
Research		
Climate	Consistent interface to meteorological data, with data selection and graphical display	Prototype
Species Diversity	Species data	Developing

democratic or bottom-up. LTER data management decisions are made at many levels, permitting a variety of decision-making strategies and promoting negotiation among interested or affected parties across sites, themes, and disciplines. The response to a policy query about productivity may be initiated at the LTER network level while the working definition of productivity for a research investigation may be prompted by both a site-specific question and a working group question at the multisite level. LTER's bottom-up organizational structure relies on the collaboration of sites to define research priorities as well as to initially unify data management protocols and develop specific network-level tools that synthesize information across sites.

Network Information System (NIS)

LTER data management focuses on the task of supporting local-site science efforts while providing network coordination to aid participants at all levels (Figure 2). The groundwork for LTER support of local-site data management was initiated during two workshops (Gorentz 1992) and further developed at LTER Data Manager Committee meetings held annually since 1988. Additional technology issues have been addressed by subcommittees at biannual Coordinating Committee Meetings for scientist representatives from all sites. By 1994, individual-site data management evolved to a point at which LTER data managers could begin discussing the need for an integrated network information system. However, the primary motivator for creation of the NIS was a 1994 mandate of the LTER Coordinating Committee (the governing body for the US LTER Network; Figure 2) that each site make available at least one online data set. As a result, data management efforts on a network level became imperative.

The LTER NIS focus is summarized in a "strategic vision" (see box page 969)—"to promote ecological science by fostering the synergy of information systems and scientific research" (LTER Data Management 1995). Using this vision as a guide, a 1996 implementation plan outlined development of the NIS through 2002 (Brunt and Nottrott 1996). Three important features have been the balance between responsibilities of the local sites and the centralized network, the modular design, and the process of prototype development.

The considerable diversity of data management systems across individual LTER sites reflects their distinct needs, resources, and organizational structures. This range derives from the critical need to facilitate local research and publication (Strebel et al. 1994), and results in part from the fact that some local data management systems predate site entry into the LTER network (Gurtz 1986, Briggs and Su 1994, Veen et al. 1994, Baker 1996, 1998, Benson 1996, Porter et al. 1996, Spycher et al. 1996, Ingersoll et al. 1997). Two common organizational structures are centralized models with a full-time site-designated data manager supported by staff and student help, and a part-time site-designated data manager who works with individual research team data managers associated with the site.

Initial LTER network-wide data management efforts focused on mail list services, a personnel directory, a data access policy, an all-site bibliography (Chinn and Bledsoe 1997), metadata standards (Michener et al. 1997), and data catalogs (Michener et al. 1990, Porter et al. 1997). Subsequent cross-site data synthesis efforts—all addressing data coordination from their inception—have included a remote-sensing sun photometer project (Vande Castle and Vermote 1996), a soil roots database (Caroline Bledsoe, University of California at Davis, personal communication), and a climate database (Henshaw et al. 1998). The growing availability of Internet tools (first Gopher and then the World Wide Web) has played an immediate and significant role in catalyzing the vision of an expanded, more integrated network-level information system.

NIS modules

A modular framework was agreed upon for the LTER NIS (Figure 3) so that individuals and working groups could develop prototypes independently. Modules may be divided into two content groups: support and research. The support modules, consisting solely of metadata databases, contain information about the research sites or the research data sets. These modules, such as a personnel directory, site description directory, bibliography, or data set catalog, may be designed and developed independently by data managers.

In contrast, research modules, such as intersite databases of climate, species lists, and net primary productivity, consist of research data and related metadata. Experience has shown that scientific questions are optimum drivers for development of research module prototypes, and that development is best accomplished by having lead discipline specialists working in collaboration with data managers (Stafford et al. 1986). This ongoing collaboration addresses design questions regarding module goals, variable definitions, and data structure. A generic module schema, developed originally for climate data (Figure 4), illustrates the interface of the individual site to the general scientific community with the mediating function of a central site. Table 1 lists prototype and operational components of the LTER NIS. Examples of both support and research modules are discussed below.

Climate database. The methodology for collecting climatic data at LTER sites is generally standardized, following guidelines provided by the LTER Climate Committee (Greenland 1997). The LTER climate database was one of the first proposed research modules for the LTER NIS because weather is an important parameter in many site and synthesis studies. The module objective is to provide current and comparable climate summaries for each site (Henshaw et al. 1998). Figure 4 illustrates how all data remain under local site control, with data involved in intersite exchange being provided in a standard exchange format. This approach is well suited to the diversity of LTER sites because it allows them to store data in any desired format. A uniform resource locator (URL) for each site is required to identify the location of the exchange-formatted data. The method used to produce the exchange format, known as the exchange filter, is determined by the site. The URL may be linked to a periodically updated static file or to a dynamic script or database program that generates the exchange data from the site's own database upon request.

Web harvesting techniques are used to automatically collect the exchange-formatted data from the site-supplied URL on a periodic basis, with the data deposited into a central relational database. Although quality assurance for the data is a local function, subsequent validation is performed at the central location to check for errors in data transmission and composite. A Web interface lets end users download reports and graphically view the database.

Distribution format views are independent of the database storage structure. Distribution filters, often implemented through programs known as common gateway interface scripts, create the distribution formats and graphic displays that the end user selects. Distribution filters are invoked through query forms within the Web interface. Implementing many parallel formats allows for meeting diverse end-user requirements and avoids discussions of which distribution format is most appropriate. For example, an initial climate committee forum recommended a single-variable matrix format while a subsequent climate workshop recommended a cross-site multivariable matrix format (Bledsoe et al. 1996). Both formats have been implemented, and new distribution filters may be added without affecting previously established formats.

The climate prototype was developed at the North Temperate Lakes LTER site using Oracle relational database software with access through a Web interface (Stubbs and Benson 1996). The final climate prototype was moved to the LTER network office where it was implemented in a Microsoft SQL Server relational database. Portability issues, considered during the design phase, affect the ease of prototype transfer from development site to full production site. The sites participating in the climate module benefit by not having to develop similar presentation report capabilities for their local climate data. Individual site investigators as well as outside community users can use this network climate database to access both a single site's data and the data from a group of sites in a common format. Such benefits to local sites provide incentives that promote participation in research module development efforts. There are 15 contributing sites whose minimum data set includes the years 1991–1995 while ongoing development focuses on climate metadata.

Data catalog. A network-level LTER data catalog creates a table of contents of the more than 2000 online data sets as a method of locating data at the LTER sites. Catalog entries are derived from site metadata documentation forms that were developed independently at individual sites throughout the 1980s (Michener 1986). Over the years, the LTER data managers have addressed metadata issues together in a sequence of working groups. The history of the LTER data catalog development illustrates how dramatically data access methods have changed. In 1990, site metadata descriptions of selected data sets were compiled and published in a hardcopy bound volume (Michener et al. 1990) through the network office. In 1993, these published data set descriptions were made available electronically at a network office Gopher Internet site while the actual data remained available at the local site. In 1997, the catalog was redesigned by the Virginia Coast Reserve LTER site as a module for the NIS, taking advantage of Web search engines and harvesting techniques (Porter 1997).

The centralized table of contents of the LTER data catalog, instead of harvesting the actual data sets, harvests metadata: individual site lists of local data sets including the data set name, the principal investigators, keywords, and local access number. This list is provided dynamically as a script or statically as a file at a local URL and is harvested regularly into a centrally located database. The LTER data set lists are then prepared as HTML files providing links to the actual local data sets and accompanying metadata from the central site. A Web search engine, WebGlimpse (Manber et al. 1997), generates indexes that allow site-specific and cross-site catalog searches for data sets. The engine provides keyword searching based on the locally provided keywords as well as free-text searching of local-site metadata descriptions. Though the issues of keywords as part of a controlled vocabulary search and of accession schemes have not been centrally addressed, the data catalog prototype has generated further discussions illustrating how the design, evaluation, and feedback steps are an inherent part of module development.

The data catalog was developed predominantly through expertise at a local site, but group discussion elicited product definition and consensus. Four prototypes, ranging in technical complexity from a simple catalog listing with links to local-site data sets to the more sophisticated creation of a new catalog archive, were presented to the larger group for consideration. Use of existing catalog archives was also considered, but update would be difficult. Each option was evaluated by data managers within the context of her or his site; the consensus choice was the prototype based on Web-crawling technologies; it is in production at the network office. The primary selection factors were ease of use and ease of update via automated processing. The data catalog includes site-specific as well as network-wide search capabilities, which benefit the site that does not want to invest in a locally developed search capability.

Bibliography. The LTER all-site bibliography (Chinn and Bledsoe 1997) is an example of an early NIS support module created using a variant of the harvest technique. In this case, individual site filters to convert local bibliographic files into a standardized, centrally located database were centrally developed rather than site developed as in Figure 4. However, this inhibited subsequent development because updating the database required manual steps, most of them performed at the central database, which distanced the site from the exchange filter used to convert its data. Such efforts provided a valuable opportunity for group education but also created a legacy of expectations since both module functionality and content may appear to be complete when this is actually true only momentarily. A working prototype can become incomplete as time passes unless a dynamic update keeps information current. The design is being modified to move the LTER bibliographic exchange filter to the local site and ensure that update capabilities can be handled more readily. Originally, the LTER database was made available online using the gopher protocol and a Wide Area Information Server (WAIS) search, but it was moved to a network office Microsoft SQL Server in 1998. The all-site bibliography contains more than 12,000 entries, with several index schemes being tested and a new interface strategy under development. The bibliography is used for both cross-site theme searches and documenting site publication.

Site description directory. The goal of the site description module is to provide a uniform presentation of site information about location, personnel, and research. In this prototype, information is entered through a Web form and saved to a central database. To ensure control at the site level, the entries can be edited there and are easily available for download. All users may view the site descriptions, with the local data manager overseeing each site's updates. For security, since passwords are not limited to local distribution, all input is saved to a temporary storage area and reviewed for integrity by the database manager before submission to the online database. A site description support module is critical for conducting cross-site research since it provides elemental site information such as location (latitude, longitude, elevation), plot size, and biome classification as well as overview information about climate, vegetation,

How Does One Establish a Network Information System?

One good approach to start an information system is to identify a common need centering on selected data sets that require integration and are pertinent to an ongoing research question. A working relationship between several scientists in addition to an ongoing data exchange creates an environment conducive to future exchanges. Thus, a beginning may involve just two people agreeing to put a year of temperature data into a common format. This is the start of a partnership.

Further developments might include the exchange of the data in several different formats with attendant metadata and a web posting to make the data more readily available. With web communication established, the participants might want to establish a mailing list to let others know about the project and offer comments. This small start can grow into a multi-investigator exchange of temperature data or into a collaboration large enough to require more specialized data handlers at the local site or a central location. Throughout the development of a data system, it is essential to stay focused on facilitating ongoing research.

soil, hydrology, geology, and education. Ultimately, it provides an index into cross-site LTER common science themes and a mechanism to link with other site description directories and other networks in the future. This module uses Web forms that are most useful for information requiring one-time entry or very occasional update.

Web form utility depends on how the input is stored and the tools available to work with the information. For the site description directory, Web forms interface to a relational database (MiniSQL) on a UNIX platform but is being transferred to another database (Microsoft SQL Server) on an NT platform in preparation for being moved to the production location. Some modules may use a hybrid mechanism, such as local-site file harvesting supplemented with network-level Web form update, until local and network efforts interface seamlessly. For instance, the site description directory and personnel directory could require harvesting to accommodate substantial changes for a local entry but might generally be served by network database Web form edits and additions.

Keys to success of a network information system

The first step in developing a network information system is understanding its intended use in depth. The concepts of information management may be addressed initially on a modest scale (see box page 972) as the requirements for data exchange and aggregation are defined. More comprehensive design elements used in the LTER NIS are highlighted in (see box on this page). Design elements may be incorporated gradually to improve the process as it expands from a few scientists exchanging data to a larger group establishing a procedure for ongoing exchanges. Elements include establishment of partnerships, definition of site responsibilities, creation of a modular structure, and development of independent prototypes. A support structure is needed to address the challenge of transforming individual site data into intercomparable data sets. The designation of a data manager and the development of a strategic vision at the local-site level have important influences on local data administration and database management. Similarly, a network information system requires data managers both at each site and at the network level. Agreement by sites on a strategic vision for the network-level information system that encompasses support of both local and cross-site ecological research is critical.

Communication is a key factor in making a large multidisciplinary project feasible (IGBP 1990). After all, the success of an enterprise is based on the exchange of information in support of its participants. There is a developing literature on organizational learning, computer-supported cooperative work, and communities of practice (Jordan 1996). Communication within the LTER network is promoted through a variety of informal means such as standardized lists, electronic mail lists, a network Web site, surveys, working groups, and symposia, in addition to the more formal committees, reports, proceedings, newsletters, and publications. Communication is essential to establishing and maintaining effective partnerships, which are another key design element for any information system. A nested structure (Figure 2) facilitates active partnerships among individual site scientists, data managers, a coordinating office, and the broader ecological community. Partnerships interlock at many levels, including within-site and cross-site as well as network-to-network. Communications specifically within the LTER data manager group have been encouraged through network office support and annual meetings. The concepts for online data handling are discussed frequently in working groups as well as plenary sessions of data management meetings. Surveys have been found to be effective tools for gathering information for group discussion; topics have included site overviews, electronic and weather instrumentation, bibliographic software, site software, data access, data policy, Web presentation elements, information system design, and information management support.

Design Elements for a Network Information System

Define focus and establish partnerships
- Encourage research-driven data management
- Include research, computer, and information sciences
- Design project structure to facilitate communication
- Consider local autonomy and heterogeneity
- Create a vision statement
- Balance in-reach and out-reach

Define site responsibilities
- Develop internet connectivity
- Maintain data electronically online
- Provide data documentation
- Address data quality issues
- Facilitate data update

Create a modular structure
- Ensure interoperability
- Build upon existing software including web tools
- Make use of filters for data exchange and distribution
- Establish ease of input/output as a priority
- Emphasize dynamic update capability

Develop independent prototypes
- Recognize small group development efficiencies
- Promote prototype assessments
- Evaluate scalability
- Ensure sustainability
- Enable site module migration to network location

Development of modules has been a crucial feature of the LTER NIS design. Within the framework of a modular structure, tasks can be considered independently as long as each is designed as a module that can interface with the NIS structure without disturbing other modules. An additional strength of modular design is that it is extensible: The information system can be extended or modified easily by adding, replacing, or deleting modules without compromising the NIS structure. This is made possible by having separate specifications for each part of the system (Brunt 1998).

Prototypes of modules are individual test designs that can be developed and implemented at any site but may migrate eventually to another location. Several prototypes may be proposed, tested, and modified in response to the need for a single module. A working prototype catalyzes development by providing a functioning product for immediate review and focuses discussion on concrete issues. Prototypes are developed by individual or collaborating sites and implemented by a few cooperating sites with the understanding that decisions made during development are subject to further modification. Evaluation within a small group strengthens the model before large inclusive group discussions. Still, the pace of development gives a site in disagreement with some aspect of an adopted model the time to create alternative suggestions. A variety of local-site initiatives demonstrating effective new approaches to data management are considered, but ultimately, only those that scale to the full network, as well as being generally robust and sustainable, are adopted. Since these initial explorations are limited in time investment and development scale, there is less reluctance to accept an alternate, more effective solution in response to feedback. Iterative module development is similar to the concepts of iterative software interface design (Kies et al. 1998). The local-site contributions to network prototypes expand the pool of technical expertise for the network. In the best of cases, there is a leveraging of resources that saves any one site from having to attempt each task individually.

Software tool developments have accelerated and created an environment of rapid change in data management techniques. Network tools have been a dominant integrative force in the development of the LTER Network Information System. Prototype implementations use tools such as Web pointers, which find Web addresses; Web harvesters, which provide automatic internet file retrieval; and Web forms, which allow database information to be added and updated. With site funding support focused primarily on science, local computing environments have relied extensively on existing software. This remains possible as long as visionary tool development in computer science is funded elsewhere. The LTER data managers have promoted collaborative exchanges with the computer science community to identify appropriate new tools and to serve occasionally as a test bed in their development. Adoption of new software is influenced by the consensus decision-making process employed at the network data management level, which acts as a conservative force in the adoption of new tools.

Different candidate hardware, software, and methods can undergo an evolutionary evaluation much like a natural selection process in which pieces that succeed are preserved and those that don't disappear over time (Kelly 1994, Spink 1997). The process is described by cybernetics in systems that have the capacity for learning and adapting. Although this feedback model does not promote linear progress, it offers flexibility because it simultaneously encompasses both design and evaluation as a part of development. Thus, instrumentation and software adopted at one site may be documented in surveys, reported upon, and discussed. Candidates that have not been discarded are refined and may spread through the network by example, and ultimately may be adopted at the network level. When a majority of sites have found a solution useful, it confirms the method's robustness. This feedback system enables sites to learn with and from each other. These communications foster replication of successful strategies and modification or avoidance of unsuccessful ones. Software choice will be influenced by the need to ensure site participation. Modules must be straightforward to implement as well as beneficial to the participating site. Their design can minimize interface difficulties and thus maximize the likelihood of site participation. Ensuring ease of input and output of data to the system is a basic priority. Providing a mechanism for data exchange and distribution is an obvious benefit to each site. The use of translation filters accommodates this exchange without dictating any changes in individual-site data management. Alleviating the difficulties of data update by the sites is also critical for long-term maintenance and should be a transparent part of the database process. Even the simple Web posting of prototype module participants provides a modest reward and encourages module implementation at other sites.

Looking ahead

As a community, the LTER may be considered a social system, or even a cognitive ecosystem, with its own unique infrastructure (Schatz 1993, Star and Ruhleder 1994, Tomlinson et al. 1998). The success of such a community system depends upon whether important issues can be recognized, communicated, and addressed (NRC 1995, Spasser 1997). In particular, it is critical for planners to keep in mind the differences between those who work to support the system and those who benefit from it (Grudin 1989). The LTER structure represents a complex interweaving of scientific, political, economic, technological, social, and educational issues. A broadening at selected sites to include social science and education components will bring new expectations and approaches. New developments will require a reexamination of both system size and organizational structure.

In looking ahead, it is important to consider the potential for LTER expansion in light of its history. The NIS has comfortably accommodated an increase in participants with initial sites unified by funding from the same National Science Foundation (NSF) division. The LTER network, which began as a group of six sites funded by NSF's Division of Environmental Biology (DEB), reached 16 sites in 1990. Several early LTER sites (H.J. Andrews Experimental Forest, Coweeta Hydrologic Laboratory, Jornada Experimental Range, Niwot Ridge, and Shortgrass Steppe) were initiated under the NSF International Biological Program. The network grew to 20 sites with the addition of two in Antarctica funded in 1991 and 1994 through the Office of Polar Programs and two urban sites funded in 1997 through a partnering of three NSF directorates: Biological Sciences, Social Behavior and Economics, and Education and Human Resources. In 1998, a 21st site was funded through DEB as a conversion of a former Land Margin Ecosystem Research (LMER) site. Consideration of additional coastal sites resulted in the identification of three new sites this year. Such a diversity of funding brings additional funds but with the added burden to network coordination given directorate differences in administration, support, and scheduling.

As the network continues to grow and diversify, participants must consider how the LTER model of nested structures (Figure 1) and consensus decision-making will be affected. We must be alert to whether additional communication mechanisms must be identified, whether the individual site will remain vested in LTER, and whether a consistent base of support can be maintained given an increasing reliance on funding sources with differing requirements and agendas. We must be aware of both the inefficiency inherent in large group dynamics (Grudin 1988) and the progress of computer-supported cooperative work (Baecker 1993, McCarthy 1994, Bowker et al. 1997) in addition to the balance of opportunistic network development against centralized control. The LTER NIS is a multidatabase system (Sheth and Larson 1990) that in the continuum of federated database systems described by Robbins (1995) is closer to the loosely coupled than the tightly coupled end of the spectrum. What changes will the NIS undergo? We will need to revisit the evaluation criteria themselves, which range from "Does the system support local science?" to "Does it promote cross-disciplinary synthesis?" to "Does it manage long-term data?" The growing need to address global ecological questions prompts us to ask how well the current LTER data management model will scale to more expansive networks and global database concerns, and leads to the question, "Will a paradigm shift occur?"

Conclusions

The Long-Term Ecological Research Network is a community of more than 1100 scientists and more than 700 students in approximately 140 institutions focused on ecological systems and common goals for the long-term data the community generates. It is a working model demonstrating how a data management structure can facilitate integrated science. Both the LTER program structure and the data management approach aim to integrate local and cross-site ecosystem research. Although early research methods emphasized the individual effort, contemporary paradigms are broadening to address community constructs (Jordan 1996). As data management becomes recognized as an integral part of a group effort and single-investigator personal data systems interface with broader data and metadata aggregations, issues such as documentation and data availability can be addressed in the larger context of integrated science. Unresolved issues such as module interoperability, data indexes, semantic capabilities, and data policy remain the subject of community discussion and research.

The LTER development of information management is an adaptive process, not a rigid prescription. It is an approach that facilitates science in the short term, fosters partnerships among data managers and scientists across the network, and creates technological interfaces. The structure of the LTER NIS allows exploration of rapidly changing computer science technology, avoiding the need to impose a single solution on an entire network. Prototype development, carried out by independent subsets of interested sites in an arena of group and subgroup discussion, has proved to be an efficient mechanism for exploring and evaluating new methods. Local LTER site initiatives, not constrained by network considerations, can also identify emerging technology that will benefit scientific progress. Yet the system is designed so that this experimentation does not impede the momentum of ongoing research.

With an emphasis on communication, ongoing feedback creates a synergy among LTER sites. This dynamic process permeates the spectrum of sites regardless of their stage of technological development. If one site cannot maintain a Web server initially, as long as it is connected electronically, communication is possible and the network can provide initial basic services such as Web page, storage, or remote-sensing support. Group connectivity means that computer-mediated communication is available to create an extended learning environment. Internet connectivity is not absolutely required for a network, but it has been a determining factor in the success of the LTER network and has been identified as a goal for all sites (Brunt et al. 1990).

Finding an appropriate balance of local development in concert with network development has guided the evolution of the LTER NIS. Local versus network dynamics can help maintain a healthy tension (Star and Ruhleder 1996, Susan Stafford, Colorado State University, personal communication) yet promote cohesiveness within the group of sites because of the participation by all in defining a system flexible enough to respond to today's rapid technological changes. From an organizational perspective, the decisions to maintain data management at a site level as well as to establish a network office that promotes and supports electronic connectivity on the Internet are significant contributors to the LTER program's success. Successfully identifying and allocating functions best maintained at the site level and at the network level is not a trivial assignment. Leadership and vision are necessary to ensure long-term viability of this complementary functioning.

The LTER NIS approach is appropriate for many groups of sites or countries faced with addressing both local needs and network cooperation. Such an approach also applies to a group faced with networking post hoc that wishes to build upon existing local data management efforts. Each LTER site has a unique research and administrative structure reflected in the local information system's balance of service, information management, and computer science. This balance is defined by the vision of the organizing body, the available technical infrastructure, and the expertise of the local data manager. The LTER data management approach, emphasizing partnerships, communications, and a modular structure developed through independent prototypes that are dynamically updated, has proven productive and has created a learning environment. The LTER Network Information System model permits site development to continue at the local level while drawing strength from the diversity of the sites in the network.

Acknowledgments

This work was funded by NSF Grants (OPP-96-32763 ksb/PAL, DEB-96-32853 bb/NTL, DEB-92-11769 db/BNZ, DEB-96-32921 dh&ss/AND, DEB-94-11974 jp/VCR, and DEB-96-34135 NET). The model presented builds upon contributions from the community of LTER site data managers as well as from the LTER scientists and administrators who have worked with vision and patience. Specific recognition is given to Robin Stubbs at the North Temperate Lakes site for development of the climate database, to Dawn Rawls for manuscript editing, and to James Thomas Callahan for undeterred support of data management.

References cited

Agnew DJ. 1997. Review: The CCAMLR ecosystem monitoring programme. Antarctic Science 9: 235–242.

Aubertin GM, Bigelow DS, Malo B. 1990. Quality Assurance Plan NADP/NTN Deposition Monitoring. Fort Collins (CO): Natural Resource Ecology Laboratory, Colorado State University.

Baecker RM. 1993. Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human–Human Collaboration. San Mateo (CA): Morgan Kaufmann Publishers.

Baker KS. 1996. Development of Palmer Long-Term Ecological Research Information Management. Pages 725–730 in Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

______. 1998. Palmer LTER information management. Pages 105–110 in Michener W, Porter J, Stafford S, eds. Data and Information Management in the Ecological Sciences: A Resource Guide. Albuquerque (NM): University of New Mexico.

Bannon LJ. 1996. Use, design and evaluation: Steps towards an integration. Pages 423–443 in Shapiro D, Tauber M, Traunmuller R, eds. The Design of Computer Supported Cooperative Work and Groupware Systems. Amsterdam: Elsevier.

Benson B. 1996. The North Temperate Lakes Research Information Management System. Pages 719–724 in Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Bledsoe C, Hastings J, Nottrott R. 1996. Xclimate workshop. Albuquerque (NM): LTER Network Office.

Bowker G, Star S, Turner W, Gasser L, eds. 1997. Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide. Mahwah (NJ): Lawrence Erlbaum Associates.

Briggs JM, Su H. 1994. Development and refinement of the Konza Prairie LTER research information management program. Pages 87–100 in Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Brown JH. 1994. Grand challenges in scaling up environmental research. Pages 21–26 in Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Brunt J, Porter J, Nottrott R. 1990. Internet Connectivity in the LTER: Assessment and Recommendation Report. Seattle (WA): LTER Network Office, University of Washington, College of Forest Resources AR-10.

Brunt JW. 1998. The LTER network information system: A framework for ecological information management. Pages 435–440 in Aguirre-Bravo C, Franco CR, eds. North American Science Symposium: Toward a Unified Framework for Inventorying and Monitoring Forest Ecosystem Resources; 2–6 Nov 1998; Guadalajara, Mexico. Fort Collins (CO): US Department of Agriculture, Forest Severice, Rocky Mountain Research Station. Proceedings RMRS-P-12.

Brunt JW, Nottrott R. 1996. The LTER network information system for the 21st century. Pages 104–104 in Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Callahan JT. 1984. Long-term ecological research. BioScience 34: 363–367.

Canadian Global Change Program. 1995. Long-term ecological research and monitoring in Canada. Final Report of the Long-Term Ecological Research and Monitoring Panel. Ottawa (Canada): The Royal Society of Canada.

Chinn H, Bledsoe C. 1997. Internet access to ecological information—The US LTER All-Site Bibliography Project. BioScience 47: 50–57.

Cuthbertson M. 1993. A database for environmental research programmes. Journal of Environmental Management 37: 291–300.

Dozier J. 1992. How Sequoia 2000 Addresses Issues in Data and Information Systems for Global Change: Berkeley (CA): University of California.

[FGDC] Federal Geographic Data Committee. 1995. Content Standards for Digital Geospatial Metatdata Workbook. Washington (DC): Federal Geographic Data Committee.

Flierl G, Bishop JKB, Glover DM, Paranjpe S. 1992. Data management for JGOFS: Theory and design. Pages 229–249 in Churgin J, ed. Proceedings of Ocean Climate Data Workshop; 18–21 Feb 1992; Greenbelt, MD; Springfield (VA): National Technical Information Service.

Franklin JF. 1994. United States—the U.S. Long-Term ecological research program: Present, future and international. Pages 22–62 in Nottrott RW, Franklin JF, Vande Castle JR, eds. International Networking in Long-Term Ecological Research. Seattle (WA): LTER Network Office, University of Washington, College of Forest Resources AR-10.

Franklin JF, Bledsoe CS, Callahan JT. 1990. Contributions of the Long-Term Ecological Research program. BioScience 40: 509–523.

Gorentz J. 1992. Data management at biological field stations and coastal marine laboratories, Report of an invitational workshop, W.K. Kellogg Biological Station; 22–26 Apr 1990; Lansing (MI): Michigan State University.

Gosz JR. 1994. Sustainable biosphere initiative: Data management challenges. Pages 27–39 in Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Greenland D. 1997. Standardized Meteorological Measurements for Long Term Ecological Research Sites. Seattle (WA): LTER Network Office, University of Washington, College of Forest Resources AR-10.

Greenland DE, Swift LW Jr. 1991. Climate variability and ecosystem response: Opportunities for the LTER network. Bulletin of the Ecological Society of America 72: 118–126.

Gross KL, et al. 1995. Future of Long-Term Ecological Data (FLED). Washington (DC): Ecological Society of America.

Grudin J. 1988. Why CSCW applications fail: Problems in the design and evaluation of organizational interfaces. Pages 85–93 in CSCW '88: Proceedings of the Conference on Computer-Supported Cooperative Work; 26–29 Sep 1988; Portland, Oregon. New York: ACM Press.

______. 1989. Why groupware applications fail: problems in design and evaluation. Office Technology and People 4: 245–264.

Gurtz ME. 1986. Development of a research data management system. Pages 23–38 in Michener WK, ed. Research Data Management in the Ecological Sciences. Columbia (SC): University of South Carolina Press.

Heal OW, Menaut J-C, Steffen WL, eds. 1993. Towards a global terrestrial observing system (GTOS): Detecting and monitoring change in terrestrial ecosystems. Report of a workshop, Fontainebleau, France; 27–31 Jul 1992.

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. Pages 123–127 in Michener W, Porter J, Stafford S, eds. Data and Information Management in the Ecological Sciences: A Resource Guide. Albuquerque (NM): LTER Network Office, University of New Mexico.

Ingersoll RC, Seastedt TR, Hartman M. 1997. A model information management system for ecological research. BioScience 47: 310–316.

[ICAIR] International Center for Antarctic Information and Research. 1993. Using information to further understanding of the Antarctic. Antarctic 13: 142–147.

[IGBP] International Geosphere–Biosphere Program. 1990. Data and information systems for the IGBP. Pages 1–16 in The International Geosphere–Biosphere Program: A Study of Global Change, IGBP: The Initial Core Projects. Stockholm: International Geosphere–Biosphere Programme.

Jordan B. 1996. Ethnographic workplace studies and CSCW. Pages 17–42 in Shapiro D, Tauber M, Traunmuller R, eds. The Design of Computer Supported Cooperative Work and Groupware Systems. Amsterdam: Elsevier.

Kay LE. 1998. A book of life? How the genome became an information system and DNA a language. Perspectives in Biology and Medicine 41: 504–528.

Kelly K. 1994. Out of Control: The New Biology of Machines, Social Systems, and the Economic World. Reading (MA): Addison-Wesley.

Kies JK, Williges RC, Rosson MB. 1998. Coordinating computer-supported cooperative work: a review of research issues and strategies. Journal of the American Society for Information Science 49: 776–791.

Knox GA. 1994. Management of the living resources. Pages 357–368 in Knox GA, ed. The Biology of the Southern Ocean. New York: Cambridge University Press.

Kouzes RT, Myers JD, Wulf WA. 1996. Collaboratories: Doing science on the internet. Computer 29: 40–46.

Kratz TK, et al. 1995. Temporal and spatial variability as neglected ecosystem properties: Lessons learned from 12 North American ecosystems. Pages 359–383 in Rapport DJ, Gaudet CL, Calow P, eds. Evaluating and Monitoring the Health of Large-Scale Ecosystems. New York: Springer-Verlag.

Lane AMJ. 1997. The U.K. environmental change network database: An integrated information resource for long-term monitoring and research. Journal of Environmental Management 51: 81–105.

LTER Data Management. 1995. Proceedings of the 1995 LTER data management workshop. Report of a workshop at Snowbird, Utah, 28–30 July. College of Forest Resources, University of Washington, Seattle, Washington, LTER Network Office.

Magnuson JJ, Wynne RH, Benson BJ, Robertson DM. In press. Lake and river ice as a powerful indicator of past and present climates. Verhandlungen der Internationalen Vereinigung fur Theoretische und Ange wandte Limnologie.

Manber U, Smith M, Gopal B. 1997. WebGlimpse—combining browsing and searching. Conference proceedings of the USENIX annual Technical Conference, Anaheim, CA; 6–10 Jan 1997. Berkeley (CA): USENIX Association.

McCarthy J. 1994. The state-of-the-art of CSCW: CSCW systems, cooperative work and organization. Journal of Information Technology 9: 73–83.

Michener W. 1986. Data management and long-term ecological research. Pages 1–8 in Michener WK, ed. Research Data Management in the Ecological Sciences. Columbia (SC): University of South Carolina Press.

Michener W, Miller A, Nottrott R, eds. 1990. Long-Term Ecological Research Network Core Data Set Catalog. Columbia (SC): Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina.

Michener W, Brunt W, Helly J, Kirchner T, Stafford SG. 1997. Nongeospatial metadata for the ecological sciences. Ecological Applications 7: 330–342.

Michener WK, Brunt JW, Stafford SG, eds. 1994. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

[NRC] National Research Council. 1995. Finding The Forest in the Trees: The Challenge of Combining Diverse Environmental Data—Selected Case Studies. Washington (DC): National Academy Press.

Olson R, McCord R. 1998. Data Archival, Data and Information Management in the Ecological Sciences: A Resource Guide. Albuquerque (NM): LTER Network Office, University of New Mexico.

Olson R, Voorhees L, Field J, Gentry M. 1996. Packaging and distributing ecological data from multisite studies. Pages 93–102. Proceedings of the Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Olson RJ, Briggs JM, Porter JH, Mah GR, Stafford SG. 1999. Managing data from multiple disciplines, scales, and sites to support synthesis and modeling. Remote Sensing Environment, Special Issue: Validating MODIS Terrestrial Ecology Products 79: 99–107.

Pearson ML, Soll D. 1991. The human genome project: A paradigm for information management in the life sciences. FASEB Journal 5: 35–39.

Platt T. 1988. Appendix D. Problems of biological oceanographic data. Pages 14–15 in Platt T, White G, eds. Joint Global Ocean Flux Study: Report of the JGOFS Working Group on Data Management, Bedford Institute of Oceanography, 27 & 28 September. Paris: International Council of Scientific Unions. Scientific Committee on Ocean Research.

Porter J. 1997. Prototyping Options for All-Site LTER Data Catalog, Data and Information Management in the Ecological Sciences: A Resource Guide. Albuquerque (NM): LTER Network Office, University of New Mexico.

Porter J, Hayden B, Richardson D. 1996. Data and information management at the Virginia Coast Reserve Long-Term Ecological Research Site. Pages 731–736. Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Porter J, Henshaw D, Stafford SG. 1997. Research metadata in Long-Term Ecological Research (LTER) IEEE Proceedings. Proceedings of the Second IEEE Metadata Conference; 16–17 Sep 1997; Silver Spring, MD.

Porter JH, Callahan JT. 1994. Circumventing a dilemma: Historical approaches to data sharing in ecological research. Pages 193–202 in

Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Riera JL, Magnuson JJ, Vande Castle JR, MacKenzie MD. 1998. Analysis of large-scale spatial heterogeneity in vegetation indices among North American landscapes. Ecosystems 1: 268–282.

Robbins RJ. 1992. Challenges in the human genome project. IEEE Engineering in Medicine and Biology 11: 25–34.

______. 1995. Information infrastructure for the human genome project. IEEE Engineering in Medicine and Biology Magazine 14: 746–759.

Robbins RJ, Benton D, Snoddy J. 1995. Informatics and the human genome project. IEEE Engineering in Medicine and Biology 14: 694–701.

Schatz BR. 1993. Building an electronic community system. Pages 550–560 in Baecker R, ed. Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human–Human Collaboration. San Mateo (CA): Morgan Kaufmann Publishers.

[SCAR/COMNAP] Scientific Committee on Antarctic Research/Council of Managers of National Antarctic Programs. 1996. Monitoring of environmental impacts from science and operations in Antarctica, Report for the Scientific Committee on Antarctic Research (SCAR) and the Council of Managers of National Antarctic Programs (COMNAP).

Sheth AP, Larson JA. 1990. Federated database systems for managing distributed, heterogeneous, and autonomous databases. ACM Computing Surveys 22: 183–236.

Spasser MA. 1997. The enacted fate of undiscovered public knowledge. Journal of the American Society for Information Science 48: 707–717.

Spink A. 1997. Information science: A third feedback framework. JASIS 48: 728–740.

Spycher G, Cushing J, Henshaw D, Stafford SG, Nadkarni N. 1996. Solving problems for validation federation and migration of ecological databases. Pages 695–700. Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Stafford SG, Klopsch MW, Waddell KL, Slagle RL, Alaback PB. 1986. Optimizing the computational environment for ecological research. Pages 73–91 in Michener WK, ed. Research Data Management in the Ecological Sciences. Columbia (SC): University of South Carolina Press.

Stafford SG, Brunt JW, Michener WK. 1994. Integration of scientific information management and environmental research. Pages 3–19 in Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Stafford SG, Brunt J, Benson BJ. 1996. Training environmental information managers of the future. Pages 111–122. Proceedings of the Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Stanford J, McKee A. 1999. Field Station 2000 initiative: Report of the OBFS Working Group on Networking. <www.obfs.org/Networking/OBFS_NCEASreport/OBFS_NCEAS.html> (10 Oct 2000).

Star SL, Ruhleder K. 1994. Steps toward an ecology of infrastructure: Complex problems in design and access for large-scale collaborative systems. Pages 253–264 in CSCW '94: Transcending boundaries: Proceedings of the conference on Computer Supported Cooperative Work; 22–26 Oct 1994; Chapel Hill, NC. New York: ACM Press.

______. 1996. Steps toward an ecology of infrastructure: design and access for large information spaces. Information Systems Research 7: 111–134.

Stonebraker M. 1994. Sequoia 2000: A reflection on the first three years. Pages 108–116 in French J, Hinterberger H, eds. Seventh International Working Conference on Scientific and Statistical Database Management; 28–30 Sep 1992; Charlottesvile, VA. Los Alamitos (CA): IEEE Computer Society Press.

Strebel D, Meeson B, Nelson A. 1994. Scientific information systems: A conceptual framework. Pages 59–85 in Michener WK, Brunt JW, Stafford SG, eds. Environmental Information Management and Analysis: Ecosystem to Global Scales. London: Taylor & Francis.

Strebel DE, Landis DR, Huemmrich KF, Newcomer JA, Meeson BW. 1998. The FIFE data publication experiment. Journal of the Atmospheric Sciences 55: 1277–1283.

Stubbs M, Benson B. 1996. Query access to relational databases via the World Wide Web. Pages 105–109. Proceedings of the Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Thorley M, Trathan P. 1994. The history of the BIOMASS data centre and lessons learned during its lifetime. Pages 313–322 in El-Sayed SZ, ed. Southern Ocean Ecology: The BIOMASS Perspective. New York: Cambridge University Press.

Tomlinson KL, Spasser MA, Sanchez JA, Schnase JL. 1998. Managing cognitive overload in the flora of North America Project. Pages 296–304 in Nunamaker JF, ed. Proceedings of the 31st Annual Hawaii International Conference on Systems Sciences, Vol. 2: Digital Documents Track; 6–9 Jan 1998; Kohala, HI. Kohala (HI): IEEE Computer Society Press.

Vande Castle JR. 1991. Remote sensing and modeling activities for long-term ecological research. Bethesda (MD): American Congress on Surveying and Mapping and American Society for Photogrammetry and Remote Sensing.

Vande Castle JR, Vermote E. 1996. Operational remote sensing data for comparative ecological research: Applications of atmospheric correction using automated sunphotometers. Pages 791–796. Proceedings of Eco-Informa Workshop, Global Networks for Environmental Information; 4–7 Nov 1996; Lake Buena Vista, FL. Ann Arbor (MI): Environmental Research Institute of Michigan.

Veen C, Federer CA, Buso D, Siccama T. 1994. Structure and function of the Hubbard Brook Data Management System. Bulletin of the Ecological Society of America 75: 45–48.