Proceedings of the 1995 LTER Data Management Workshop

Participants and Contributors: P. Adams, K. Baker, B. Benson, D. Blodgett, J. Briggs, C. Bristow, J. Brunt, G. Calabria, H. Chinn, D. Chang, M. Easter, J. Gosz, M. Hartman, J.Hastings, B. Hayden, D. Henshaw, R. Ingersoll, T. Kirchner, K. LaFleur, R. Lent, E. Melendez, B. Nolen, R. Nottrott, J. Porter, S. Stafford, M. Tomecek, J. Vande Castle, C. Veen

Table of Contents

---

1. Working Groups

1.1 Strategic Vision for LTER Information Management (back to TOC)

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**

Pursue information systems development and implementation from the context of ecological research needs. We approach information management from an ecological research perspective, guiding our development by the research uses of the information system and evaluating the system implementation in the light of specific research projects. Information management is an integral part of the LTER research platform, and provides crucial infrastructure for the LTER research enterprise.

Conduct information management in the nested context of site, research network, national, and international levels. Site-oriented agendas must be balanced with broader (network, national, and international) level goals. Network solutions should consider and facilitate local site solutions while local site solutions should take into account the strategic plans developed at the network level.

Emphasize the timely and effective transformation of data into information and the general availability of that information. The LTER community expects us to continually find better ways to access and use existing data to answer increasingly complex questions with minimum difficulty. Meeting these expectations will require new data products and information interfaces. Developers of environmental databases must address many issues including the storage and integration of a wide variety of data types, a large range of both temporal and spatial scales and sizes, and increasingly sophisticated analytical requirements. Design of any information system must assure data quality,

protection, and availability while considering cost and efficiency. While sites independently build information systems to meet their own needs, an information system at the network level may require the development and implementation of standards for LTER.

Ensure the long-term preservation and continual accessibility of information. Long-term research requires a data management environment which provides for the long-term availability of data and metadata. Preservation implies that appropriate quality control checking has been performed on the dataset, and that the associated metadata is sufficient and complete for interpretation of information in the future. Potential users of this information include scientists, academicians, managers, policy makers, as well as the public. Data publication of selected high-quality datasets should be pursued.

Ensure appropriate information system development through information management research. The information management system of the LTER must facilitate current needs, anticipate new technical horizons, and be extensible to new technology and research requirements. Research into new information system technologies must be conducted to ensure appropriate system development for LTER, with external resources cultivated when necessary. The LTER network should provide a testbed for new technological solutions, providing a means for implementation and improvement of the research platform. Publication of innovative implementations of new technology should be explored. Historically, we have explicated the necessary components each site should develop to provide acceptable data management at a site level in support of an operational network of sites, and this role will continue.

Develop human resources necessary for the continuing evolution of LTER information systems. As we move into the 21st century, we must ensure that the intellectual capital in this area exists for the next generation of long-term ecological research. We see a need to continue developing training materials and curricula in the area of information management. The Data Management Committee meetings are essential to maintaining our collective expertise, and other workshops will be necessary as we move toward a network development focus.

This document is a working document of the LTER Information Managers and as such will be dynamic.

2.2 LTER Network Information System (back to TOC)

In the past six years the LTER information managers have led a concerted effort to utilize and improve Internet connectivity among the research sites within the network. The resulting electronic infrastructure has revolutionized the ways in which sites exchange data and information. This infrastructure also provides the basis for the existing LTERnet information system, which has helped to transform a collection of individual sites into a coherent network. In formulating and implementing the vision of the LTER Network Information System for the coming decade and beyond, the DMC ensures that we will maintain the momentum of our past success.

The information managers are aware that, with the rapid changes in technology, this document on the LTER Network Information System is a dynamic working document. Because of the limited time available at this workshop, many planning and implementation details remain to be worked out. An ad hoc working group for developing an LTER Network Information System, the NIS Working Group, was established at the conclusion of this workshop. The group consists of Barbara Benson(NTL), Darrell Blodgett(BNZ), James Brunt(SEV), John Briggs(KNZ), Gil Calabria(CWT), Jordan Hastings(MCM), Don Henshaw(AND), Richard Lent(HFR), Eda Melendez(LUQ), Rudolf Nottrott(NET), John Porter(VCR), and associated research members Caroline Bledsoe, Harvey Chinn and Robert MacArthur. The group, in close communication with DataTask, will meet and communicate regularly to refine this planning document and initiate action to move along the implementation process.

Four working groups in several sessions focused on the following general questions:

- what is the "LTER Network Information System" - define the general functions at the site and network level of an "ideal" system as close as possible to the system that we realistically can expect to implement
- who are the users of the system and participants in its development
- what are the steps to implement that system, with alternative approaches to system design

The conclusions of the individual group discussions, as presented by rapporteurs, were then synthesized by the entire group with emphasis on system functions and characteristics, implementation considerations, likely system users participants and collaborators in the System's design and development.

Recognizing the importance of workshops in all phases of the design and implementation process, the working groups also addressed questions of what kinds of workshops and pre-workshop activities would be needed, what type of potential users and collaborators would be most interested and suitable participating, and the criteria by which to judge the success of those workshops.

The conclusions on these topics are detailed below.

System Functions and Characteristics

Participants agreed that the main purpose of the present and future LTER Network Information System is to support basic ecological research and science, both at the site and network levels. Several system characteristics follow from this premise. Participants recognized early on, however, that input from potential users and collaborators will be essential in completing and refining the list of system functions and characteristics, and that the best way to accomplish this task would be in a series of future workshops.

From an economic point of view, the System should be **capable of answering "standard" information requests** (those that occur >= 80% of the time), and should provide a **economically feasible mechanisms for building query systems for specific projects and special information requests.**

The System is a **distributed system** so that site data and information can be maintained locally. A **scaleable system** will enable users to expand site specific studies into intersite studies, and a prototype system or module that has been developed for a small group at one site or study should also be useful for larger groups and intersite use. **Queries and browsing of datasets and metadata, located at multiple sites**, are essential characteristic for the system to function as an intersite research tool. The system should have the capacity for **data input at t**he site level, as well as at the level of multiple sites, to accommodate the needs generated by specific research groups. **Platform independence** will make it possible to use the system in the environment of heterogeneous hardware and software in use by different sites.

To make the system useful to the users of widely ranging areas of research interest and levels of expertise in computing and communication systems, the characteristics above must be accompanied by a set of software tools that will **present data and information in a consistent fashion**, independent of their original format and location. The distributed system should have a **consistent appearance**. The location of the data and much of the work performed to retrieve data will be **transparent** to the users. Initially this may be achieved with relatively simple mechanisms, such as an expanded data catalog that provides transparent links to data and metadata, whereby the consistent "look and feel" would be generated by filter programs transforming information in variable site formats into an common LTER meta-format. The use of filters has been utilized in the existing All-Site Bibliography. Other mechanisms could employ DBMS standards such as ODBC in a client/server environment to achieve platform-independent consistency. **On-line help** will be essential in making the system friendly for beginning and casual users. A number of tools will be available to assist in such functions as analysis or complex queries, analysis and graphic display, and handle diverse project-specific data and information output (datasets, graphs, pictures, arbitrarily rich - combinations of many different types of data and information)

System developers and maintainers will also need their own tools, (e.g., tools to provide statistics on use, to archive queries (unsuccessful queries would be a valuable system research base), tools for evaluation of network functions, user feedback, monitoring logs, etc.).

System Implementation

Workshop participants emphasized the need to carefully balance planning, design and implementation considerations. A more detailed specification of system functions and characteristics is still needed. The precise implementation details

will depend on input from potential users and collaborators which will be provided through a series of future **workshops**.

There was agreement that the System should be **develop**ed **in a modular**, s**tep-by-step approach**, while **maintaining and building on present functionality**. Components of the existing system will be a good start. This naturally includes the existing LTERnet information server, catalogs (Core Data Sets, Images), the All-Site Bibliography, the LTER Personnel Directory, electronic connectivity, and other components of the LTERnet Network Support System.

An important issue is the consistent use of existing **content standards** for online data and metadata at all sites, as well as establishment of new content standards where they do not exist. These standards are a prerequisite for the distributed system's capability to present data and information in a consistent fashion, independent of their original format and location. Standardization is also critical for any meaningful synthesis of data across multiple sites. In contrast, network-wide **format standards**, which will simplify implementation of the System, are less critical than the content standards, since a uniform appearance could be achieved by using filter programs to create standard views from files in different site formats. Nevertheless, format decisions will still affect a variety of tasks, such as development of system parts like a distributed cross-site catalog. This will require consistency between sites, standardized keywords, development of the catalog query system (minimum metadata standards, structure/format/access method standard within site, writing filters to produce desired displays, etc.)
Ad hoc working groups for development of standards will play a vital role during system implementation.

We recognize the importance of **prototypes** for testing, demonstrating and **proving system concepts** as well as **analysis and integration tools**. Prototypes can be developed and used in an iterative fashion, e.g., a prototype could be defined as a result of a workshop. The prototype would then be developed by a team of specialists, and subsequently used (perhaps in a follow-up workshop) as a research tool and to refine system functionality.
NIS task groups composed of researchers, data managers, and occasional consultants will provide much of the personnel power necessary to design, fund, and implement various components of the system.

Workshops and Pre-Workshop activities

Workshops offer the most effective mechanism, both for completing a detailed initial definition of system functions and characteristics, and for planning and clarifying further implementation details. The **DMC, the NIS Working Group and Network Office personnel will be core groups** in this process. They provide consistent and complete site representation, expertise in practical information management and expertise in computer and communication systems, as well as knowledge of application of these disciplines to ecological research. The members of these groups also provide an important element of long-term "LTER institutional memory".

By the very nature of their research, **groups of investigators with a focus on intersite research** will be in the best position to help define detailed needs and characteristics of the Network Information System. However, because the system must also be useful for individual sites, the somewhat different requirements of research conducted at single sites is no less important. It is critical, therefore, that the needs and views of **site PIs and other researchers, including graduate students**, are properly accommodated in the System.

We also need to ensure that we draw upon the tremendous **pool of computer and information science experts that exists in organizations and agencies outside LTER**, in the wider ecological research community and in computing and commuincation research. If we do our work right, the LTER NIS will be of use not only to LTER, but also to this wider community. Interactions in the past with such organizations as the San Diego Supercomputer Center (SDSC), the National Center for Supercomputer Applications (NCAS), NASA, DOE, and others, have demonstrated this potential. Consultants with special skills and expertise may come from these organizations, as well as others, and could be instrumental in completing specific system implementation tasks or build various system components.

The **workshops will therefore include members from all these diverse groups**, as well as **ad hoc NIS task groups.**

On a regular basis, the DMC will dedicate a significant portion of future annual LTER information managers meetings to issues related to NIS planning and development.

The workshops will focus on topics including:

- detailed **definition of needs**, system functionality and characteristics; which types of scientific questions could be answered with the existing system, what are the most important missing pieces and obstacles; what are the opportunities; identify the types of questions the system must address (intersite projects as well as site researchers are asking these questions)
- **initial mini-workshops or sessions at individual sites** would focus on current capabilities; could fill out questionnaire after; develop common outline across sites for format of workshop, including range of recommended demos; should be very simple; possibly done at subset of sites or with varying formats to meet local needs
- detailed **definition of implementation tasks** and system modules (the latter would lead to development of prototypes
- system **architectural design** and **modules definition**
- design of **prototype modules** to support specific existing research efforts (intersite or site-specific); in subsequent workshop use the prototype for real work and make modifications and improvements

Pre-workshop activities

Given the relatively short time frame under which most workshops will have to operate, and the enormous size and complexity of the task, it is critical that a large amount of preparation is done prior to workshops, that the work continues between workshops and that work begun there will be completed successfully. Most of the mechanisms proposed rely on the use of the electronic network infrastructure we have developed in LTER and on the experience and availability of strong research programs at the sites. Effective means for the collaboration preparing for LTER NIS workshops, and generally collaboration outside workshops that will lead to creation of the envisioned NIS include:

- At the **sites**, get input from visionary PIs that ask envelope pushing questions; Ask "How is our existing system limiting?" or "What improvements in functionality do they see?" or "What would they like in its place?", site data managers can conduct one-to-one interviews; Outreach to all LTER scientists through **local site meetings**; approach depends on site (different levels of expertise); PI-grad student teams (take questions; apply to own and other sites)
- review the existing individual site iformation information systems (difficulties, blocks to syntheses, prototype system interactions - they know what they like)
- review existing network information systems that are publicly accessible
- information managers put together framework and PIs respond
- survey PIs -- information manager at each site can do informal query, would likely get constructive response in one-to-one interactions
- bring information managers, or subgroups, together frequently to compare notes, e.g could do as small regional working groups could meet regularly; use phone and video conferencing
- use electronic mail or LTERnet information server to circulate questionaire (what are site's current needs and capabilities? what services are most important and useful now? what have researchers tried to do and been frustrated? what functional additions are most critical?

Funding

Just as it is clear that an undertaking of the size and complexity of the LTER NIS cannot succeed with the support of a few individuals or sites alone, it is also clear that we will have to rely on multiple sources of funding and support, in addition to our present limited resources available specifically for information management. The Network Office will be a major player in NIS development, but there will be a delay in the availability of newly targeted funds until the beginning of the new funding cycle in 1997. We nevertheless need to make sure that our vision is properly reflected and supported in the upcoming proposal for the Network Office, as well as in any future site proposals. In addition, we must seek other resources aggressively, such as funding from the Database Activities in the Biological Sciences and other agencies (NASA, DOE, etc.).

2.3 Data Management Committee Structure and Implementation (back to TOC)

The data management committee (DMC) is a standing committee of the Long-Term Ecological Research (LTER) network consisting of at least one representative from each LTER site. It's function is to actively develop initiatives and address pertinent issues with regard to the accessibility, management, exchange, and analysis of ecological data. Decisions are made by consensus. Issues requiring decision are brought before the group as a whole either electronically or in a formal meeting.

The DMC has two standing subcommittees with shared membership: the Data Management Task Force (DMTF) and the Data Management Meeting Organization Committee (DDMOC).

The functions of DMTF are to 1) provide representation for the committee to the LTER Coordinating Committee, 2) ensure continued funding for future meetings, and 3) provide guidance and continuity to the DMC. The chair of the DMTF (which is elected by the members of DMTF) has the additional directive to represent the DMC to the Coordinating Committee. In addition, DMTF may appoint working groups/standing committees to ensure the continuation/completion of projects/meeting/workshops etc., when necessary, beyond the annual meeting. These groups shall serve as long as directed by DMTF. Any products/decisions that are produced by these groups will be presented to DMTF who will notify DMC.

DMTF will consist of six members, 2 of which are capable of serving as signatory Pis (should the need arise for proposal submission), with the Network Office data manager who will serve ex-officio. As of now, DMTF consists of Karen Baker (PAL), Barbara Benson (NTL), James Brunt (SEV), John Briggs (KNZ), John Porter (VCR), Rudolf Nottrott (NET-ex-offico) and Susan Stafford (AND). [John Porter is ex-offico due to his tenure at NSF], with James Brunt serving as chair. To encourage stability and maintain "institutional" memory in DMTF, the following rotation plan is proposed and with the current DMTF members having the following terms:

- Briggs and Brunt (January 01, 1994 to December 31, 1995)

- Porter and Benson (January 01, 1994 to December 31, 1996)

- Baker and Stafford (January 01, 1994 to December 31, 1997). (Note Susan Stafford was ex-offico during calendar year 1994 while at NSF).

Starting at the 1995 DMC meeting, two DMTF will be selected from the DMC. Terms for future DMTF will be for 3 years and will run from Jan. 011996 to Dec. 31 1999. At this time, no term limits are proposed. From the the DMTF a chair will be selected. DMTF will then inform the DMC. If possible, the chair should serve as long as possible during his/her DMTF appointment.

If more than 2 individuals are nominated, a vote of the DMC will be conducted either at the annual meeting or by electronic mail, with a simple majority ruling. The chair of the DMTF will oversee this election. If the chair of the DMTF is involved in the elections (i.e. a candidate), the data manager of the network office will conduct the elections.

RESULTS FROM THIS YEAR ELECTIONS:

Gil Galabria and Don Henshaw were elected and will serve on the DMTF from 01 January 1996 to December 31, 1999.

4. Site Bytes (back to TOC)

AND (Don Henshaw, Susan Stafford, Gody Spycher)

Information Access: We expanded information access by establishing a World Wide Web homepage (http://www.fsl.orst.edu/lterhome.html) which provides selected metadata and data as well as an extensive collection of GIS coverages. The homepage also offers an overview of the the LTER program at the H.J. Andrews Forest, a personnel roster, and an overview of the site facilities. We view this as a dynamic system whose use will continually expand in the future.

Data Production: We refined our quality control procedures for our major longterm, nonspatial databases by expanding rulebased tests. A typical example for vegetation databases is ensuring that status codes (ingrowth, live, dead) are properly sequenced over time. Quality control for metadata covers automated checks for completeness and consistency of all metadata information. Quality control of data now includes checks for ranges, codes, duplicates (where applicable), referential integrity (where applicable), integrity of temporal sequences (where applicable), and specific database rules .

The Andrews P.I.'s are undertaking a "data assessment initiative" to review what spatial and non-spatial datasets, associated meteorological data, and models are available for the Andrews and surrounding region, and to determine unfulfilled data and modeling needs. This project is intended to review what models are being used, what are the limitations of these models, what are the data requirements for these models, and what are the data products from these models. This initiative is considered an important component in the preparation of the LTER 4 proposal.

Field Station: Andrews field technicians recently completed installation of two new Meteorological Stations, completing the planned expansion to four stations. This expansion was based on an overall review of our long-term climate measurement program. The four stations are intended to capitalize on the existing hydrometeorological network, and to represent the environmental gradients across the Andrews.

Biodiversity: We are identifying all species lists for the Andrews, and plan to establish an on-line database of all species lists. We hope to have several on-line for the biodiversity workshop at the October Coordinating Committee meeting. We have lists for about 10 taxa, some of which include information on habitat, trophic status, and other information.

Future Plans: Developer Roen Hogg and the OSU College of Oceanography conducted an indepth evaluation of their computing environment, and recently completed the development of an information system based on that evaluation. We hope to update our information management architecture by adopting some of the basic results from this effort, which is based on the clientserver model and "low" cost shrinkwrap technologies. Central to the system is the use of ODBC compliant software.

We are planning a tandem approach in determining the appropriate server to house all core LTER datasets: 1) evaluate Microsoft's SQL Server on a scalable NT Server platform, and 2) evaluate Oracle which will be available on the new Forest Service UNIX-based computing system. Regardless of the server, we will continue to use Visual Foxpro as the front-end tool for quality control. Other front-end or "productivity" tools (ie., spreadsheets, graphics, word processors, statistics, or any ODBC compliant software) that are commonly used will be tested for connectivity and interoperability with the server. Over the next year we will begin to move LTER core data onto the new server and interface the system with the Internet and the productivity tools. Our mainfocus will be on resolving network access to a heterogeneous database environment over the Internet and we will work with other LTER Data managers in finding common solution strategies.

ARC (Jim Laundre)

The Arctic LTER had their NFS review in July . Mike Allen lead the team of reviewers which included John Porter as a member. The review report was positive and offered some helpful suggestions.

This spring we created World Wide Web pages for our site and it has allowed us to improve our on-line data access. The various types of data collected by the ARC LTER are described in web pages with links to the metadata files and the ASCII delimited data files on the gopher server. The people responsible for organizing and creating the web pages are : Neil Bettez - Lake data; Chris Harvey - Streams data; Jim Laundre - Terrestrial and Weather data; and Wil Wollheim -

Bibliography. More files are in the process of being added along with other sources of information about the arctic. A searchable index is to be added as soon as WAIS is running on the web server.

The field season is progressing well. The new field labs should be ready late summer. They had to be modified once at

the station to come up to the specifications of the original plans. They will greatly alleviate the crowded lab space as well as provide for a safer and a more convenient lab environment.

I will not be able to attend the Data Management Workshop this year since in the arctic peak biomass is the end of July. Beginning August is the start of fall weather and of finishing several experiments for the season. We will be doing a biomass harvest on a set of experimental plots that have be maintained for 15 years. This combined with the fact that most PIs were at the station in late June to early July for the review prevented us from sending any other assistant data manager during this important part of the season.

BNZ (Darrell Blodgett)

Information Flow and Access: This has been a busy year at Bonanza Creek as well. On-line HTML forms for creating and editing metadata (study descriptions and dataset descriptions) are in place and have been used by PI's to submit information into our World Wide Web server system. A large portion of our PI's are finally networked from their desktops as opposed to having a single networked workstation shared by 6 or seven PI's. Most PI's can now transfer data to our online system by simply copying the file to a network drive mounted on their personal computer. We installed PCNFS, Netscape and Eudora for our newly networked PI's. I think this has been a critical step to improve the flow of information into our online system. I will be putting some heavy pressure on the PI's this year after field season to get documentation and datasets on-line and archived. We have switched from Mosaic to Netscape as our primary browser/interface to our information system. Netscape had more complete support for on-line forms across all three platforms we are using here (Mac System 7,MS-Windows, SunOS/Solaris).

NSF Site Review: We had our Mid-Term NSF site review July 10-12. The weather was great, and the visit went pretty smoothly. Our information management system drew a positive review from the NSF panel. They did present us with several suggestions for improvements, which are being addressed.

System Upgrade?: We had plans to upgrade our main Server to Solaris 2.4 beginning July 19th. We found out that a critical component will not upgrade to the new OS. Another important issue is the cost of system maintenance for the server versus system maintenance and upgrade costs for a newer model machine. Currently weshare the system with another University of Alaska unit which both houses the server and pays for it's maintenance. One of our systems is already running 2.4, and I assisted another affiliated unit in upgrading thier network to 2.4 and NIS+. As some of you know, there are major differences between SunOS 4.1.3 and Solaris 2.4.

Security became a big concern throughout our University. Several Hackers were caught in the act of breaking into one ofUAF's systems. With the release of Satan, there was even more concern. The difficulty lies in our responsibility to make our data available to the public, and at the same time protect it from unauthorized modification/elimination. We have installed tcpwrappers as a first step in preventing break-ins, but our system security is still not as strong as I would like it to be.

Testing: I spent some time experimenting with the idea of a distributed information system based on accessing WWW servers at multiple sites from a single script. I found CERNS WWW line mode browser and developed a simple script to go to several sites and pull off what most resembled a data catalog. The script generated a single HTML document with links back to the source of the information.

CDR (Clarence Lehman, Charles Bristow)

This year, observations at Cedar Creek gathered over the last decade helped resolve a long-standing ecological controversy about biodiversity -- the question of whether complexity begets stability, as Elton (1958) concluded, or the opposite, as May (1973) concluded. Remarkably, both views were correct, but both had been simplifying a more complex question in different ways. Details are in press.

We continue to emphasize biodiversity at Cedar Creek and now have a new full-time research director, Johannes Knops, to assist. Dr. Knops oversees field work, data management, and intersite coordination, as well as conducting ecosystem research on carbon and nitrogen cycling.

We have installed a central Sun computer at the Cedar Creek field station. Formerly only satellite computers were there, with the central computers on campus. We have trained additional students as computer assistants to help with maintenance and development of all our computers.

Using the multiprocessor software on our central Sun Sparc-10 for computer simulations, we further explored conditions under which habitat destruction drives the top competitor extinct first (as reported last year), and found such conditions to be quite broad. We have expanded these simulations to an investigation of percolation effects on species recovery after habitat destruction.

The simple software we established last year for online laboratory notebooks has proved itself in keeping track of computer simulations, statistical analyses, and experimental designs, especially for a new large-scale experiment which examines effects of biodiversity on population, community, and ecosystem processes.

Finally, we are completing a cross-site study concerning photosynthetic functioning among seven different sites.

CPR (Tom Kirchner and Mark Easter)

Over the last year a committee of scientists associated with the LTER program reviewed and amended the data management plan for the project. The committee developed a statement of the objectives for data management and policies for implementation of those objectives. The data management policies list the responsibilities of the scientists and the data management staff, and also give guidelines to ensure that data collected for the project become publicly accessible on a timely basis.

The CPER site also completed a technical review of the data management and geographic information systems, and we are in the process of implementing many of the recommendations of that panel. Our goal for this year was to get much of our data installed in the data base system, and to provide the project scientists with the tools and interfaces necessary to access the data. The tools for accessing the data are essentially completed, with the exception of updating of the manuals to reflect recent changes in the software.

We have completed converting our system for handling bibliographic data from Fortran to C++, and have added an X-windows interface (Motif) for entering and keywording references. The bibliography was recently updated and we have also added an interactive tool for searching the bibliographic data base. The system enables one to save results of searches as ASCII files, WordPerfect files or FrameMaker files.

Our current system enables scientists from anywhere on the internet to access our publicly available data using the LTERMENU program that we developed. We have also developed a Mosaic "page" to provide access to our data. The data and data description files can be browsed with Mosaic. The advantage of LTERMENU over Mosaic is that it enables one to plot or to extract in tabular form specific columns of data from the data files. The Mosaic interface provides easy access to the files, but does not provide a way of easily extracting data. The data files are often hard to browse with Mosaic because most of our tabular data are stored in fixed field, card image formats. LTERMENU has been updated to support IEEE big-endian binary files (native SPARC format), IEEE little-endian binary files (Intel and M68000 family processors), External Data Representation (XDR) binary files, comma delimited ASCII files, blank and/or tab delimited ASCII files, and files using any arbitrary delimiter. Support for the display of other types of files, such as image files, through external programs or scripts has also been added to LTERMENU.

Tom Kirchner is being replace by Mark Easter as Data Manager for the CPER site. Mark is also serving as the Project Manager for the CPER LTER project. Our programmer position will be terminated at the end of October.

CWT (Gildo Calabria)

Introduction: This past year has been very exciting for Coweeta Data Management / GIS personnel. The proposed idea of using Web Pages as a front-end to the Coweeta LTER Information Server has been a success.

Accessibility: By combining the process strength of our Sparc-1000 server with the intuitive feel of WWW's

hypermedia environment, Coweeta LTER data sets and other site information are now at the fingertips of all Coweeta PI's; and at a lesser extent to the rest of the web world. This has been possible by developing gateway interfaces between our databases and the web server, and by writing numerous lines of HTML code. We are very excited about the possibilities of this information distribution system, and we are currently very busy expanding these services.

Automation: The Coweeta LTER Site collects very large microclimate data sets on a continuous basis. Hence, it has become imperative that an automated archival procedure be implemented for these data sets. This automation has proven to be an invaluable tool, and it has now been expanded to include several other data sets. In fact, the Coweeta WWW project has been a byproduct of the "extra" time produced by these scripts.

Training: To continue our training efforts, two new workstations have been added to our DM/GIS facilities. These machines have greatly improved our lab accessibily to the graduate students, and it has also upgraded our 24-bit graphic capabilities.

Connectivity: Since no workstations were available on site, a Sun workstation has been purchased for the Coweeta Hydrologic Laboratory field station. We are eager to improve the connectivity between the field station and our facilities at UGA, beyond our current high speed (28.8K) modem connection. We are anxious to learn about alternatives already implemented at other LTER sites.

Biodiversity: We are currently developing a catalog of all species lists for the Coweeta LTER site. This catalog will be distributed on-line via our Web server, and should be compiled in time for the biodiversity workshop at the October Coordinating Committe meeting.

GIS: We are also currently building the on-line Coweeta GIS Atlas. This atlas will depict all Coweeta base layers, and geographic infomation about our ongoing research. Finally, we are extremely excited about two new products developted this past year: (1) a notation of human and non-human usage and disturbance for the Coweeta basin, dating back 6000 years; and (2) the first draft for the Coweeta Basin Vegetation Map.

HBR (Cindy Veen and Paul Boon)A dedication ceremony was held to rename the field lab. There is a new conference room and dorms - so no more meetings in a hot, stuffy barn! The conference room can hold 160 people at a meeting.

Efforts have been expanded in archiving physical samples. Some of the samples date back to 1962. Others have been subsampled up to 9 times. Efforts have focused on documentation of previous samples, resampling sites and comparing with past samples, and rerunning some old samples to test for old analytical errors. Sample datasets have been archived.

The EDEX system has been on the EcoGopher at Yale. There are now 40 public EDEX files occupying 6.5 Mb. Several scientists have volunteered more data. The Hubbard Brook Publication list now has more than 1300 citations.

A location database in GIS is being created. The database will aid the advisory committee in evaluating requests for new studies. Some sites will be noted as off limits to new research. The research history of some sites will also be examined.

HFR (Richard Lent)

Harvard Forest's LTER site grant was funded for six more years. As part of this new phase of LTER, Harvard Forest has recently made significant advances in data management, computing, and networking. The Forest now has a full-time data manager (R. Lent), with responsibility for strategic planning and oversight of all aspects of research information management and support. A major part of this planning effort has been production of a data

management design document. This document provides an overview of the Harvard Forest information system and describes detailed procedures for generation, documentation, and archiving of data produced by researchers working at the Harvard Forest.

As of July 1995 we established a World Wide Web home page on LTERnet (http://LTERnet.edu/~rlent/hfrhome.htm). This Web site will provide an introduction to the Harvard Forest; information on fellowships, educational programs, and research opportunities;

abstracts of current research from the annual Harvard Forest Ecology Symposium; guidelines for prospective HFR researchers; bibliographies of Harvard Forest publications; links to Web sites of research collaborators and Harvard University; information on the Fisher Museum of Forestry; and HFR on-line datasets. This information system will evolve continually as additional datasets, graphics, and other information sources become available in on-line form.

Harvard Forest is now fully connected to Harvard University's computer network in Cambridge via a high-speed leased telephone line. All of our offices and laboratories in all three HFR buildings have been wired with network cable. Approximately 35 IBM-compatible PC's are now directly connected to Harvard and the Internet. All HFR personnel have Harvard University electronic mail accounts and have switched over to using Eudora e-mail software. In addition, a new telephone system (Nynex Meridian Norstar), with voice mail, was installed in December 1994.

In further support of Harvard Forest data management, a new archival facility for data, samples, and other materials is under construction with funding from Harvard University and an NSF Facilities Improvement Grant. Currently, the Archives occupy a fireproof vault in the main headquarters building (Shaler Hall) and contain a 90-year record of research in forest ecology. Much of this archive contains irreplaceable original data sheets, maps, photographs, and other records accumulated since 1906. Managing this archive is a monumental task because most of it, save for data collected beginning in 1988 with the start of the HFR LTER, is not in electronic format. The new facility will greatly expand the space available for data archiving, providing three rooms for storage and work space. A fourth room, also under construction, will provide space for a separate archive facility for samples such as soils, tree, and sediment cores.

Using LTER supplementary grant funds, we have significantly upgraded our mass storage and data archiving capabilities. New hardware acquisitions include six CD-ROM drives, plus a CD-recordable drive for making our own CD's; 35 Connor tape drives; five internal and one external Bernoulli cartridge drives; six gigabyte-range hard drives; and a Microtek ScanMaker III color scanner. This equipment will enable each of our 35 PC's to be backed up onto tape by individual researchers, and will also facilitate analysis and management of very large data files and remote sensing imagery. The HFR electronic data archive will be maintained on several of these mass storage devices (e.g., hard disk, tape, CD-ROM). In addition to the new hardware, optical character recognition software now enables scanning and machine-reading of old typescripts stored in the Archives. We are also evaluating neural-network software that translates scanned paper maps into GIS layers. These acquisitions will aid in the conversion of archival materials into machine-readable data.

JRN (Barbara Nolen)

This is the first year of our third funding cycle. This includes strong collaborative ties with USDA Jornada Experimental Range (JER) with Research Director, Kris Havstad joining the Jornada LTER investigator ranks.

Our main push at this time is to get our site on-line with respect to off-site availability of data and site information. All of the office computer systems are connected to a Novell Netware Network system (V4.1). The Netware network is at this point "in-house" only, though we continue to be linked via e-mail and able to ftp and telnet as required. We are evaluating several different Internet suite and World Wide Web server packages designed for Netware. We intend to be on-line with our own WWW Home Page by the end of the year.

Our site was featured in the February issue of Discover magazine. In May we co-sponsored the Ninth Wildland Shrub Symposium with focus on "Shrub Ecosystem Dynamics in a Changing Environment" with 200+ in attendance.

With regard to climate, the past two years have been very dry with corresponding reduction in grass cover. After 10 years of average or above rainfall, it will be interesting to see if this is a blip in time or the beginning of an extended dry period for the Jornada.

USDA JER will be archiving long-term data sets with Jornada LTER. An example is their current effort to rescue an

old date set of plot pantographs of vegetation cover that date back to 1915 with measurements continuing to date. These will be used to create a computer visualization of reconstructed landscapes over time using GIS.

An intersite, long-term small mammal research project that will incorporate work at Jornada LTER, Sevilleta LTER, and Mapimi (UNESCO Man and Biosphere site in Mexico) has been initiated with baseline measurements being completed this winter. "The purpose of this study is to determine whether or not the activities of small mammals regulate plant community structure, plant species diversity, and spatial vegetation patterns in Chihuahuan Desert shrublands and grasslands." Additionally, the "study will provide long-term experimental tests of the roles of consumers on ecosystem pattern and process across a latitudinal climate gradient."

KBS (Martha Tomecek)

We have spent most of the year making information about the Kellogg Biological Station LTER site and its datasets available over the WWW. Presently the following data sets are on the web: NPP, soil inorganic nitrogen, spatial variability, agronomic yields, gps, and climate (current and historical). We intend to also have accessible the datasets for insect population dynamics, biodiversity (plant and insect species lists), and soil total carbon and nitrogen. In addition we have provided background information about our site such as our research overview, experimental design, maps on soils, elevation, vegetation, etc., listing of investigators, bibliography of our publications, graduate student opportunities, and agronomic logs. A

recent publication by Robertson and Freckman (Ecology 76:1425) listed the URL citation for the dataset, possibly the first to have done so. We are also working to network co-Pi's who currently are not able to access the Internet.

KNZ (John Briggs)

Once again, weather (rainfall) was the dominant factor at our site this year. May of 1995 was the wettest May on record and as a result of the excessive rainfall, the road to our site was damaged by a mud slide. Thus, we are now forced to use a another road to reach our site.

Our goal for this year is getting all of our data on-line and this will be done before our renewal is written this winter! As of today, we have almost 30% of our data on line via WWW server (climate.konza.ksu.edu). In the past, we have focused on only getting data on-line that was requested (i.e publication list, fire dates, etc.) but now we are encouraging all of our PI(s) to get their data on-line. Data on-line without metadata is useless, so we are placing supporting data including documentation and our entire "Methods Manual" on-line complete with maps and figures. Having this information on-line was very helpful this spring especially to non-Konza scientists who were writing proposals for intersite work.

Due to a variety of reasons (mostly money), the Konza Prairie Data Management staff is back to one full-time person (me!). This is the first time since 1987 that only one full-time person has been involved with data managment at Konza. Thus, I have been somewhat busy.

MCM (Jordan Hastings)

Admin/Events: Data and Information Manager, Jordan Hastings, spent 5 weeks at the McMurdo field site in Taylor Valley, Antarctica. Associate Data and Information Manager, Dr. Dah-Wei Chang, was recruited and recently hired (1 July 1995). Assistant Data and Information Manager, Ms. Julie Matheson, was hired and trained during the year. Julie took full responsibility for Data & Information during the Manager's deployment in Antarctica.

- Two paid students are working on computer systems and networking, and database design issues.
- Eight student volunteers (4 in each of two semesters) were inducted to assist with the Taylor Valley GIS.

SOLA [Science On-Line Antarctica] proposal was submitted to NSF/DBA competition in Fall, 1994; recently, it was partially funded as a bootstrap collaborative project involving the larger LTER Network community. (Topic for discussion at this year's DMan meeting.)

- Technical: High-level documentation was prepared for all officially submitted datasets (~30 total), in the style of a bibliography. A few synthetic datasets (~10) were compiled from submittals, and documented similarly. A suite of software was developed for processing meteorological data from the meso-network in Taylor Valley, and the first two years' datasets from 9 stations are currently being worked up. Taylor Valley GIS significantly advanced, with major assistance from ESRI (both labor and product). Dialup connections were established from Taylor Valley field site back to Continental US and will be upgraded to SLIP/PPP this year.

Major computer equipment purchases were delayed a year - but are now done! (So no ftp, gopher, or www presence yet.)

- Publications: Four journal articles, three re: Taylor Valley GIS, and one re: Taylor Valley meteorlogical data; two workshop presentations, both re: Taylor Valley GIS; and one workshop proceedings, joint with Dr Wharton: Antarctic GIS were published.

Manager was appointed to Antarctic Computer and Communications Committee. Negotiations were opened with Smithsonian Institution Museum of Natural History, for Antarctic participation in the upcoming 'Forces of Change' exhibit.

NTL (Barbara Benson and Paul Hanson)

In the fall of 1994, the WWW home page for NTL-LTER was made public. This home page now contains a site description, the text and figures from two funded proposals to NSF in 1994, the personnel directory, bibliography, and a data catalog with links to on-line data sets. In 1995, a home page for the field station, Trout Lake Station, was developed by station staff and made publicly available.

In the past year, we have completed the migration of the databases (physical and chemical limnology, fish, and meteorology) which were in Ingres on the central campus computer into the Oracle database on our local Sun workstation. In addition, we have put meteorological and water temperature measurements from the raft on Sparkling Lake and ice-on/ice-off data into the Oracle database.

The NTL-LTER has chosen to use a relational database (Oracle) as the structure for its core data because this technology provides powerful query functionality and allows researchers to make linkages among diverse data tables. Many researchers at NTL-LTER have now become adept at using the client browsing tool (Oracle Data Browser) to directly access data from the Oracle database. Two training workshops on SQL and Oracle Data Browser were offered during the past year.

The Internet connection for the field station was completed in early 1995. The Internet connection is a major step in allowing researchers at the field station to access data and services directly for both intrasite and intersite research. It is facilitated through a 56Kb dedicated phone line. Both the file servers and Gatorboxes (a Gatorbox is an Appletalk router) at Trout Lake and the Madison laboratory are able to communicate directly using IP tunneling of Netware's native protocol, IPX, and IP tunneling of Apple's native protocol, Appletalk. The end results are that a PC user in either Trout Lake or Madison can log-in directly to either or both of the file servers; a Mac user in either Trout Lake or Madison can see both LANs as zones and login to computers at both locations using Apple's Appleshare; administrators can service either file server remotely.

NWT (Rick Ingersoll and Michael Hartman)

Our field headquarters, the Mountain Research Station, was ethernetted in late autumn 1994. FTP, not floppy diskette, is now the primary means of data transfer between MRS and the NWT LTER data management laboratory. A permanent (underground) power line to the Niwot Ridge Saddle (our main research site) is scheduled for late summer/autumn of 1995. The improvements will greatly enhance the scientific capabilities at the Saddle.

We designed and constructed our WWW server in late February and early March 1995. It provides access to virtually any type of information (including data) about our site that anyone might desire. We have received quite a bit of positive feedback, including designation as a "cool" (pun intended) site by the Boulder Daily Camera newspaper.

During late April over 250 cm of snow were recorded at the Mountain Research Station. May was the wettest on record with nearly 3.5 times greater precipitation than average. The result is that the typically short growing season will be even shorter this year --- snowpack conditions in the tundra in early July were representative of early June conditions in most years!

With the exception of an encounter with a vehicle driving in the wrong lane in Boulder Canyon, our site review in late June was a decidedly positive experience. Appropriately enough, the weather for the field visit was beautiful, while the weather for the indoor presentations was nasty! The review team provided overwhelmingly 'high marks' to the NWT LTER information management system. In fact, it would be impossible to provide the relevant comments from the report, without giving the appearance of shameless self promotion! Those comments will be provided to other LTER information managers upon request. This validation of our information management program came at an ideal time with Michael Hartman being hired full-time on 1 July, so that he can replace Rick Ingersoll, who is leaving the program at the end of August.

NET (Rudolf Nottrott, John Vande Castle)

Last year has seen substantial additions to the Network Office facilities and the LTERnet Support System. The University of Washington made substantial additional space available to accommodate the needs of the Network Office. In addition to new offices for individual NET staff, we now have a large room suitable for workshops, training sessions and other types of group meetings. Most of the equipment comprising the LTERnet system has been moved to this new space (LTERnet Center).

Collaborations with NASA have resulted in new database activities at the Network Office. Data from sun photometers at HJA, BNZ, NTL, MCM and SEV are part of a larger sun photometer network generating data for information on atmospheric aerosols and water vapor which are critical for current and planned remote sensing systems. Data from additional LTER sites will be included in the future. The sun photometers acquire data every 15 minutes and relay them hourly via GOES satellite to Wallops Island. The data are then processed at NASA Goddard. The database (now about 1gb) and the analysis software ('demonstrat') are mirrored on LTERnet. Additional Landsat-TM data will be acquired for the project and added to new on-line data storage (81Gb). This data storage will be used to keep key LTER remote sensing data on-line, in addition to the near-line Unitree archive currently maintained. This project is supported by NASA funds which will also increase the computational capabilites of the LTERnet GIS systems.

Our switch from X.25-based Sprintnet to SLIP/PPP dial-up access has proven so popular that it was substantially exceeding the funds allocated for this service. Following a decision at the VCR CC meeting that sites should pay for network access out of their own budgets, several sites have acquired or are in the process of acquiring their own 800 number for LTERnet access. The data manager at NET is assisting sites in this process.

Another measure of the popularity of the network system is reflected by the continuing growth of LTER personnel with valid e-mail addresses, now 93% of the nearly 800 people now listed in the LTER personnel directory (90% percent last year).

A new addition to our connectivity arsenal is the LTERnet Field Connectivity Kit, which has just arrived and is being configured for a first real life test at the ERECO in August in Budapest. The Kit consists of a portable Voyager Sparc Station with Ethernet and serial PPP interfaces. The Ethernet forms the basis of a LAN which is connected to the Internet through a dial-up connection to LTERnet. This type setup will prove useful in supporting information management at field station LANs where a high-speed T1 or similar options are not economical.

Now that we have put the basic network infrastructure for communication and information exchange in place, the NET information manager is shifting emphasis to the systems that can work on top of this infrastructure.

To prepare for the task of creating a network-wide distributed information system, including the capability to query relational databases, I have performed tests using ODBC network access tools that are now a de facto standard for DBMS client/server applications involving PCs and Macs. Preliminary tests have shown that off-the-shelf hardware and software, including shareware and public domain software, can be used to simultaneously query tables maintained

at in database servers at BNZ (climate table) and NET. A similar test including NTL is planned for next month.

I am presently exploring possibilities to get a shared license for some of these products (Oracle, Ingres).

Prompted by the directions from NSF, the Network Office has made a strong effort to promote international collaboration among lter researchers world-wide.

In March, the NET information manager has organized a two-week training course in which participants from CERN, TERN and Malaysia learned how to establish an Information server on the Internet.

As part of the First Regional Conference on International LTER, East Asia & Pacific Region, in Taipei in April, Rudolf Nottrott and John Vande Castle organized a Workshop (co-organizers Teng-Chiu Lin and Jen-Jiun Perng of TERN) on "International LTER Information Servers on the Internet - Global Communications and Connectivity Facilitating Long-Term Ecological Research World-wide". 21 scientists and from Australia, Japan, Mongolia, Taiwan and the U.S. participated.

To facilitate data and information exchange for ILTER, we have set up a new Internet domain, ILTERnet.edu, and we are now configuring an information server for the growing ILTER Network on this domain.

PAL (Karen Baker)

In General: The fourth field season from September94 to March95 continued the weekly sampling program at Palmer Station. The lterjan95 cruise was the third annual January cruise providing an interannual regional view that complements the seasonal work done in 1993. A Palmer LTER PI Meeting in May began planning for our renewal proposal.

Information Availability: More consideration is being given to online presentation of information as this becomes an established and expected medium of communication. The web homepage established in February of 1994 is being augmented with links to related historical sites. Weather information has been specifically made available online. Data online for Palmer LTER investigators is increasing as issues of documentation, quality assessment and publication are addressed.

Weather: Because of the importance to specific current site research questions and because of the importance in general for ecology at a long-term site, a number of issues pertaining to Palmer station weather and climate have been addressed over the last year. These efforts are ongoing. Historical weather records have been collected and analyzed for presentation in a chapter for the site's synthesis volume 'Foundations for Ecological Research West of the Antarctic Peninsula' to be published before the end of this year. A second automatic weather station (AWS) has been installed 90 km off shore from the first AWS installed near Palmer station. Both battery malfunctions as well as connector problems were identified and corrected due to a weekly monitoring program established to insure data quality. The value of overlapping weather measurements as a method of quality assurance using the Palmer operations and the AWS weather measurements has been documented in an Antarctic Journal Article.

Field: Co-ordination with the Antarctic Support Associates for field logistics, communications needs and electronic technical issues continues through participation with both the McMurdo Users Working Group and the Antarctic Communications and Computers Working Group. The possibility of developing a new satellite link to Palmer exists and will have tremendous impact on our field efforts if successful. Emphasis on display of data while in the field continues as does consideration of the use of GIS software for field data visualization.

Documentation: An increasing amount of time is dedicated to online participant lists, event sampling logs, sampling maps, bibliographic database, calendar and contributions to the Palmer monthly electronic agenda. Species list compilation has begun in response the intersite work. Work on procedure manuals continues. Some draft versions on field site computers exist and development is ongoing in preparation for the fifth field season and this season's lterjan96 annual cruise.

SEV (James Brunt)

The Sevilleta LTER has begun its second funding cycle. SEV II involves some major new science initiatives that build on the SEV I science. The first of two major components of this are a vegetation mapping project that is being done in using 13 TM images and verified

through a series of extensive hypothesis testing ground campaigns. The second is a modeling project to predict actual evapotranspiration. There are two new post-docs involved in this project.

Three new Sparc 5 workstations with video conferencing capabilities were added in the spring. Sun's ShowMe software has now been successfully tested over the network and is a reasonable product. The Sevilleta WWW server development continues with a collaborative grant from Sandia National Laboratories. The project is focusing on database queries of meteorological data. Oracle SQL and other database products will be added in 1996. In addition, a faunal treatment for Sevilleta NWR is almost complete. Recent proposals from Sevilleta investigators to NSF have included URLs.

New proposals have been funded to further parasite work at the Sevilleta and to enhance the computational capabilities at the Sevilleta Field Station. The latter includes new Sun Ultra-Sparc workstations and a Server to make the field station computing environment autonomous. A cross-site proposal was funded to extend the rodent exclosure work to the Mapimi Biosphere Reserve in Mexico. The proposal provides a satellite communication system to provide data transfer and voice communication.

(back to TOC)