

01001100 01010100 01000101 01010010

LTER DATABITS

01001100 01010100 01000101 01010010

Fall 1992, Scientific Information Management Newsletter of Long-term Ecological Research

Welcome to a new issue of DATABITS! This issue has a good balance of technical and non-technical features. There are reports on the use of interactive data access and data management policies within the LTER network. There are also plenty of hints on how to better use the computers at your site, including software reviews and technical hints. Finally, there are articles on conducting GPS campaigns and some solutions to common problems in scientific information management. There have also been some superficial format changes to DATABITS. Its now created using Framemaker software on a Sun Workstation. As always, DATABITS depends on you for contributions. Thanks!

CONTENTS

Site Flashes	1
Interactive Data Access	3
LTER Information System - Simplified Access and new Capabilities	5
H.J.Andrews GPS Campaign	6
Challenges in Data Management	6
Data Access Policies -- A Preliminary Report:	7
LTER Guidelines for Site Data Management Policies	8
EUDORA - Mail for the Mac!	9
Using a Site-wide Email Calendar	9
Software tips from Antarctica	10
Publication Information	10

Site Flashes

CDR-- We had a site review during the week of July 13th 1992. Dr. Susan Stafford was on the review panel. This is a very positive step since it highlighted the increasingly critical role of data management in LTER projects.

Our data is all on line now in two sites: The University of Minnesota and Idaho State University (Pocatello). All updates will be done from the University of Minnesota. We have also moved from SAS 6.03 to SAS 6.07 thus giving us access to SQL at the same time to the analytical power of SAS.

In the Hardware arena, MSI/TSR funds from this spring were used to acquire an additional SparcStation IPX and a Sparc-10, model 54 with 4 CPU's and 128 MBytes of physical memory expandable up to 640 MBytes. We will have symmetric processing capabilities as soon as the SS10 and

the Solaris 2.0 are delivered in the fourth quarter of 1992. This will allow us to do analyses that were not possible before. Also we will be moving to a new building on the St. Paul Campus in December of 1992. This will enable most of our researchers and hardware to be in one location (Ecology, Soil Science, Forestry, etc.) and allow greater sharing and efficient utilization of resources. Another new surprise this summer! This has been the coldest summer since 1904. This had effects on both species richness and biomass, as did the drought of 1988. Our data would not have been as useful if we did not have 11 years of experimental data with which to compare it. -- *Abderrahman Elhaddi, Cedar Creek LTER, - AEIHaddi@LTERnet.edu*

CWT-- During the past year, the Coweeta DM & GIS Lab has doubled its size. Every hardware and software package used in the lab has been upgraded, and all the UGA Co-Pis are now connected to the Coweeta Local Area Network. In addition, a high-speed 9600 baud modem has been installed to facilitate the communication and data transfer between the Coweeta Hydrologic Laboratory field station and the Institute of Ecology facilities.

The Data Management Personnel has been involved in three separate projects. Considerable energy has been spent in upgrading and giving system support to the GIS facilities. One major upgrade has been the installation of a new Image Processing workstation currently being used to process both SPOT and LANDSAT scenes, and the scanning and digitization of low altitude aerial photographs. Another project has been the documentation and archival of older data sets into the Coweeta data bank, and the formatting of data sets of on-going research. Thirdly, the "Annotated Bibliography of Publications on Watershed Management and Ecological Studies at Coweeta Hydrologic Laboratory" have been entered into "Citation" format to aid us on reference upgrading and the publication of such

a bibliography. An On-line Bibliographic Catalog is being implemented using the BibIX bibliographic software. This on-line catalog will be available to all the LTER sites via e-mail

We have recently installed the INGRES relational database package. It will be used to replace the current data archival-data receival system, and to implement a more robust database. Using the ARC-INGRES and the SAS-SQL modules to interface with INGRES relation databases, we plan to develop Scientific RDBMS environment which will automate most aspects of data management, and give the Copi's a more user-friendly interface to the Coweeta data bank.

Geographic Information Systems at Coweeta are progressing in two directions. First, the basin wide database is available for researchers interested in modeling processes across the site. These layers include geology, soils, topography (slope, elevation, and aspect), and some vegetation information. A spatial site management system is being established to allow researchers to select research sites by querying the database: e.g., find all the north facing slopes with wet soils at an elevation > 900m. It will also provide the ability to display past or current research sites by querying the Coweeta Databank. The system will utilize Arc/View to provide a user friendly interface to the data.

The basin wide database is also being expanded to the regional scale. Landsat and SPOT satellite images will be classified using Erdas image processing software to provide landuse/landcover information. The USGS 1:250,000 Digital Line Graph and Census Bureau TIGER data will provide base information including roads, rivers, basins, and civil divisions.

The second direction is mapping the gradient, gap, riparian and stream studies at a larger scale to afford more detail. A Total Station transit is being used to acquire high precision position and topographic information. These data are integrated into location information collected by global positioning systems. The descriptive data collected will be stored in the Ingres database and modeled using Arc/Info. -- *Gil Calabria, Coweeta LTER, GCalabria@LTERnet.edu*

JRN -- In late spring we hosted the second annual 1992 Jornada Symposium that proved to be a successful coming together of range scientists and ecologists who have or are planning on conducting research within the Jornada study area. We also enjoyed a visit with a delegation from the Tarim River Basin in western China who are establishing a long term monitoring station there. Foxpro2 is up and running for our database software on the PC along with ARC/INFO and ERDAS for our GIS software. We also have a new software package RDL-SIMS for downloading, error checking, graphing, and report generation of data from our weather station Campbell CR10 datalogger. We are embarking on a GPS field project that will include expansion of the GPS grid for photogrammetrics and the mapping of research sites, rain gages, wells, fencelines and roads. This

effort is in collaboration with the USDA, BLM and Forest Service. -- *Barbara Nolen, Jornada LTER, BNolen@LTERnet.edu*

KBS-- KBS LTER is in the process of upgrading its weather station. New purchases include a 10m tower for the wind instruments, soil and air temperature probes, Vaisala temperature and RH probe and a quantum sensor. Efforts are being made to assure that data from this station are compatible with the KBS Pond Lab weather station. The Pond Lab Station is also part of the inter-site climate database project and currently managed by John Gorentz and Rich Losee.

Management of the LTER weather station include daily distribution of automated weather reports via email to interested persons. A hard copy graphic summary of daily, monthly and cumulative precipitation is distributed at the beginning of each month. These summaries also include 30 year means.

On the GIS front, our VAXstation 3100 has been running with ARC/INFO version 5.0.2 using DEC windows. Baseline data from the 1988 LTER spatial variability study have been incorporated in the ARC/INFO database and applications are now available which allow for automated mapping of kriged data sets. A menuing system allows the investigator to examine any one of 28 variates currently on-line for patterns of soil variability. Geostatistical models used for kriging are also available for viewing. Work is being done to link frequency distributions, TIN's and individual plot data to this menuing system. -- *Lolita Krievs, Kellogg Biological Station LTER, LKrievs@LTERnet.edu*

LUQ-- Data Management at the Luquillo Experimental Forest has been involved in completing the details of what we understand are the basics of Data Mng. in a site: developing the Site Data Management Policies, guidelines to complete forms, and facilitate the documentation of the data sets. All the forms, guidelines and policies are now available in a manual. Also, the data sets catalog for the site was updated and a printout of the 57 LTER data sets following the LTER Net format (with some additional information added) produced. This catalog is now available. Every year we update our bibliography which now includes all publishable activities. Together with our PI we have finally developed a format to assign keywords to facilitate the retrieval of the desired citations. Now we have started to obtain long-term core data sets, namely, streamflow, meteorological and water chemistry data from USGS, Dept. of Commerce and other agencies. We are manipulating those data sets we have received and changed their format to one which can be easily plotted. The aim is to include them in our next LTER proposal. Our future plans are more oriented toward the manipulation of data, although we are not forgetting about the documentation of the data sets. An on-line catalog for reference is also in our short term plans for development. -- *Eda Melendez, Luquillo LTER, EMelendez@LTERnet.edu*

NWT-- We have made a number of important software acquisitions during the past 12 months. We have obtained the PC version of GEO-EAS in order to facilitate geostatisti-

cal analyses. Applied Information Systems' PC version of EasyEntry is being used to standardize quality control during manual entry of data. Climate data will be accessible from an INGRES database in the near future; subsequently all metadata and key data sets will reside in INGRES databases. New hardware acquisitions include a 386 PC dedicated to information management, as well as peripherals (1.3-GB desktop storage pack, Exabyte tape drive, laser printer) attached to our SPARC 2. In addition, TDR (time domain reflectometry) hardware and software will enable us to quantify spatial and temporal distribution of soil moisture at our site and 2 gas chromatographs provide the basis for a fairly sophisticated trace gas laboratory at our field research station. Conversion from thin-wire to twisted-pair ethernet connections for microcomputers and workstations has virtually eliminated maintenance costs for those connections. We have recently completed a manual that facilitates access to and documents methods for maintenance of our climate data set. -- *Rick Ingersoll, Niwot Ridge LTER, RIngersoll@LTERnet.edu*

PAL-- The first field season from September 91 to March 92 was completed. LTER large-scale and small-scale sampling grids were established along with GPS location of the grids. Sampling included initiation of an oceanographic program, continuation and expansion of krill and bird survey programs, and installation of the first automatic weather station. At the Palmer post-field season all-scientist meeting held in May 92 at UCSanta Barbara, data issues included the investigation of networking of remote pi's via internet to the Santa Barbara CRSEO Computer Center, establishment of a central 600Mbyte archive disk, demonstration of a metadata browse program, as well as consideration of data forms and data exchange. There is a synergism between the new Sequoia2000 project which is a Univ. of California collaboration with DEC Computer Company since their objectives include working to develop Earth Observing System (EOS) era technology including data management and data visualization. --*Karen Baker, Palmer LTER, KBaker@LTERnet.edu*

SEV-- SEV Sevilleta crews recently completed the field work for a high-resolution GPS survey. 10 high resolution monuments were established in a network design across the Sevilleta National Wildlife Refuge tying into Highway department GPS points north and south of the Sevilleta. Data from this survey have not been processed pending the arrival of new processing software.

Work has just been completed on a T-1 bandwidth connection from the UNM campus to the Sevilleta field research station. The T-1 provides voice and data service to the field station with data speeds of about 1.1 Mb/s. Many field season data processing chores have now been relocated at the field station which houses 2 workstations, 2 PC's, and 8 terminals. The voice lines are UNM campus extensions. Real cost for this connection is about 8K per year.

Installation is continuing on an RF network connecting eight sevilleta meteorological stations with the field station

network and making it possible to monitor any station from workstation in the campus lab.

Because of increased requests to the data management group, a request tracking system has been instituted. The simple software database system provides a way to monitor the number of requests and what stage of processing they are in.

Two new Sun workstations have been installed and a new DEC 5000/240 server. These workstations have 3rd party super-high resolution graphics capabilities for working with planned low altitude aerial survey ortho photos.

Currently, testing is underway on the use of SAS AF and SCL modules for data entry. Preliminary results look good - all the functionality is there, although the learning curve is a bit steep at first it levels out quick. -- *James Brunt, Sevilleta LTER, JBrunt@LTERnet.edu*

VCR-- In February we completed installation of 38 high-resolution GPS markers spread throughout the lower 40 km of the Delmarva Peninsula. Data processing of so many baselines (over 150) and network analysis required an upgrade of the LTER data management PC to a 486-based machine. In the past year we have worked on establishing a research site database in our GIS system, along with a standardized set of species codes for both the fauna and flora of the Virginia Coast Reserve. In cooperation with the University of Virginia Library, Academic Computing Center and Department of Environmental Sciences, our ARC/INFO capabilities have been greatly expanded. An RS6000/340 workstation capable of supporting up to 12 simultaneous ARC/INFO users was installed, along with 6 X-terminal and two additional digitizing tablets. -- *John Porter, Virginia Coast Reserve LTER, JPorter@LTERnet.edu*

Interactive Data Access

At the 1991 Data Managers' Meeting in San Antonio, there was a working group on Interactive Data Access (IDA). The group focused attention on a very important aspect of data management -- how easy is it for a researcher to get the data in a useful form. We recognized the connection between the research productivity of a site and the system for accessing data. The multidisciplinary nature of LTER research adds additional impetus to the need for easy access. IDA supports a productive workshop/working group environment. An additional argument for the development of IDA is the increasing work load on the data management staff in handling data requests.

Given the interest in and importance of IDA, the working group decided to conduct a survey of the LTER sites to assess the current state of IDA. At the 1992 Data Managers' Meeting, Barbara Benson summarized the responses to this questionnaire and data managers from all sites were given the opportunity to edit summary sheets of responses. This

article will present the updated summary tables and overall conclusions of the survey. The survey helped identify sites which have already made substantial progress in this area and it is hoped this summary will help us share our technological experience.

The first question in the survey dealt with who performs data access. Does the individual researcher or data management staff perform the retrieval? If the individual researcher has direct access, is this because individuals control their own data or there is automated access to a centralized database? The answers to these questions varied across the LTER sites and also across data sets within a site.

TABLE 1. Who performs data access

<i>Site</i>	<i>Individual researcher</i>	<i>Some Automated Access</i>	<i>Data Management</i>
AND		X	X
ARC	X		X
BNZ	X		X
CDR			X
CPR		X	
CWT			X
HBR		X	X
HFR	X		X
JRN			X
KBS	X	X	X
KNZ		X	X
LUQ			X
NIN			X
NTL			X
NWT	X		X
PAL		X	X
SEV		X	X
VCR			X

Several sites currently have a user interface to LTER databases. KBS has IDA for weather, herbaria and some spatial data. HBR has access to data through their bulletin board. KNZ has IDA for weather, soil moisture, data catalog and publications list. NWT makes climate data available through IDA. PAL makes subsets of more extensive data sets available. AND has an IDA under development which is not yet publicly available. CPR and CDR both have IDA to their LTER data.

The next group of questions dealt with what software is being used for various data management tasks relating to data access. The overall picture here is again quite heterogeneous. The following table summarizes which Data Base

Management Systems (DBMS) are being used both for the data and the data set catalog.

TABLE 2. DBMS Software

<i>A. Data DBMS software</i>	<i>Sites:</i>
dBase III+	HFR, PAL
dBaseIV	BNZ
Foxpro	AND, JRN
Informix	CDR
Ingres	BNZ, CWT, KBS, NTL, NWT, VCR
Oracle	KNZ
Paradox	ARC, LUQ
SAS	CDR, NIN
<i>B. Metadata DBMS software</i>	<i>Sites</i>
dBaseIV	VCR
Foxpro	AND, JRN
Gopher	CWT
Ingres	BNZ, NWT
Notebook II	HFR
Oracle	KNZ
Paradox	LUQ
SAS	CDR
Symphony	NWT
Unix structured	PAL

The next question asked what software is used for manipulation of data in the data retrieval process. Some sites also have developed their own inhouse software for data manipulation.

TABLE 3. Software for data manipulation within the retrieval process

<i>Software</i>	<i>Sites</i>
dBaseIII+	HFR, PAL
dBaseIV	BNZ
Excel	NTL, NWT
Foxpro	AND, JRN
Ingres	BNZ, KBS, NTL, NWT
Lotus	HFR, LUQ
Paradox	ARC, LUQ
Qpro	ARC, LUQ
SAS	AND, BNZ, CDR, CWT, KBS, KNZ, LUQ, NIN, NWT, PAL, SEV, VCR

TABLE 3. Software for data manipulation within the retrieval process

<i>Software</i>	<i>Sites</i>
SPSS	PAL, VCR
Unix tools	BNZ, CDR, NWT, PAL, SEV

The last software question asked what software is used for data entry. At some sites the software used is quite diverse because individual researchers perform data entry (BNZ, HBR, KBS, NWT, PAL, VCR). Several people complained that they were dissatisfied with all the available software options.

TABLE 4. Software for data entry

<i>Software</i>	<i>Sites</i>
dBaseIII+	HFR
Foxpro	AND
Informix	CDR
Ingres	KBS
Paradox	LUQ
Easy Entry	NIN, NWT
Knowledge Manager	CDR
Excel	ARC, NTL, NWT, PAL
Lotus 123	BNZ, LUQ, NWT
with D.A.V.E.	JRN, SEV
Lotus 2.2	HFR
Quattro Pro	ARC, BNZ, JRN, LUQ
20/20	KBS
SAS-FSP/AF	CDR, CWT, LUQ, SEV
FORTTRAN program	HBR, JRN
inhouse software	CDR, CWT, KNZ, NWT, SEV
PC program	CPR
text editor	KBS
autoload via command procedure (from datalogger)	LUQ, KBS

IDA is an expanding area of interest at LTER sites. Six sites are now in the process of implementing an IDA system (AND, BNZ, CDR, CPR, KNZ, SEV). Another six sites currently have plans to implement IDA (CWT, HFR, NIN, NTL, NWT, VCR). Three other sites indicated they would like to implement an IDA system (JRN, KBS, LUQ).

Sites were asked to identify what the barriers were to the implementation of an IDA system at their site. Almost all sites mentioned some combination of time, personnel or money as limiting. Connectivity issues were important for several sites because interactive access from remote site locations required network connections. Proprietary issues would need to be addressed in any automated access system. Many sites felt that the need to operate on multiple platforms would present a barrier. Several sites anticipated problems with computer comfort of some end- users. Other problems mentioned by individual sites included system performance, ongoing changes in hardware and software, tracking who has access to data and when, integration of software, and the wide variety of spatial and temporal sampling among data sets.--*Barbara Benson, North Temperate Lakes LTER, BBenson@LTERnet.edu*

LTER Information System - Simplified Access and new Capabilities

Access to the LTERnet Information System has been simplified by shortening it's Internet name to LTERnet.edu (from LTERnet.Washington.edu). **LTERnet.edu** now also has an additional name **ESnet.edu** (for Ecological Sciences Network), chosen to emphasize the LTER network's efforts to reach out to the wider ecological research community, including the ESA Long-term Study Section members, the Land Margin Ecosystems Research Program and numerous other ecological research groups. The previous LTERnet addresses will remain valid and synonymous with LTERnet.edu (or simply LTERnet on Bitnet. New capabilities have been added to the LTERnet Information system, including query functions for the LTER data set catalog (by e-mail and direct access), additional bulletin board readers (nn, mrn), a menu-driven information system access. Please send any message to Overview@LTERnet.edu to get the most recent copy of our LTERnet system overview document.

The public information directory at LTERnet can now be browsed through the **Gopher information server**, which provides menu-driven access to the files in the anonymous FTP directory. Anonymous FTP can still be used to retrieve the files in this directory, but menu-driven Gopher access is intuitive and has additional capabilities. Gopher lets you view files, mail them or print them. Viewing includes the **capability to see graphics images**, such as satellite images of sites (GIF files of Landsat scenes, for example). Below is a short description on how to use Gopher. Gopher has just been installed at LTERnet, so there may still be a few

glitches. If you encounter any problems, please let us know, so we can fix them. Please send your messages, comments and suggestions to Helper@LTERnet.edu.

There are three main ways of Gopher access:

1) On any system that has an executable copy of the Gopher client program, type

```
gopher LTERnet.edu <RETURN>
```

(If you have a SUN workstation running SUN OS, you can get a copy of the 'gopher' client from LTERnet, directory `~ftp/programs/unix/sun`).

Follow the menu to browse through the various information directories. If you have an X windows terminal, have a look at an example of image display. Go to "LTER Sites file exchange space", "H.J. Andrews Experimental Forest (AND)", then press <RETURN> on any of the image files there (Patience! Depending on network traffic, the display may take a couple of minutes to come up).

2) Log on to LTERnet. If you have an X windows terminal and if you want to display image files, enable LTERnet to use your display by typing 'xhost LTERnet <RETURN>' on your local host computer. On LTERnet, type 'setenv DISPLAY yourHostName:0.0 <RETURN>'. Then, on LTERnet type

```
gopher <RETURN>
```

>From there, browse the directories as in 1).

3) This option will be functional soon: From a Gopher on any Internet host, select "Other Gopher Information Servers", "North America", "Long-Term Ecological Research Network (LTER)". From there, browse the directories as in 1) and 2). -- *Rudolf Nottrott, LTER Network Office, RNottrott@LTERnet.edu*

H.J. Andrews GPS Campaign

The GPS campaign is off to a successful start at the Andrews Forest. HJA has the high precision Trimble GPS units for the months of September and October, and Fred Bierlmaier, Hazel Hammond and Art McKee are directing our efforts. Jeff Jefferson from North Inlet LTER was on hand for the first week in September, and was invaluable in preparing and training our folks for the campaign.

The Andrews faces some unique problems in accomplishing this survey. Distances from two known globally positioned stations outside of the Andrews to one primary control point in the Andrews are 26 and 38 km. These are considered very long baselines for successful high resolution accuracy. A second problem involves having direct line

access to four satellites simultaneously with the many high ridges and tall mature/old-growth forest comprising the landscape. These obstructions can block large portions of the horizon, severely limiting satellite reception.

The baseline problem has not caused significant measurement errors. The crew set up the antennas and receivers for 5 hours (and repeated on two or three days) for the HJA primary control points. Maximum replication errors were less than one-half cm (in lat-long) over one 26 km baseline, and were always less than 6 cm over the longer baselines. Now that we have established these control points, we will have considerably shorter baselines, and are receiving satellite signals for shorter collection times.

The more significant problems involve the limited horizons available and placement of permanent monuments. While this is not easily solved, Hazel Hammond has developed an innovative computer program which greatly enhances our planning and chances for success. Hazel's program uses color graphics on a polar coordinate system to display the tracking of satellites over time. Input files (available from the satellites) with satellite azimuths and elevations by time allow us to display the satellite tracks color-coded by hour. The program then overlays the specific site's horizon (obtained from field measurements), and we can quickly and easily visually determine which hours are optimal for receiving adequate satellite data. This program has immensely improved our planning and provides necessary information for determining the feasibility of various site locations for placing permanent monuments. It also allows us to know precise times of optimal reception, which will shorten our satellite reception times to two and one-half hours.

What advice can we offer for other sites? You cannot be too prepared. Choose wisely the placement of permanent monuments, and install them before the equipment arrives. If horizons are limited around certain monuments, evaluate optimal satellite times before the campaign begins (we will be happy to loan out the satellite program we have developed). Have a field crew of at least three people prepared for a month of long work days. Send Bill Michener letters of support for the NSF grant to buy the new GPS software (it would be nice). And one equipment note: keep the antennas inside over night since they don't function if too cold. -- *Don Henshaw, Andrews LTER, DHenshaw@LTERnet.edu*

Challenges in Data Management

(Editors note: This is the first in what we hope will be a series of articles where individual data managers will talk about challenges they have faced.)

I think everybody will agree that being a Data Manager is a rewarding, interesting, but not always easy way to earn a living. I have been at the Luquillo Experimental Forest, LTER in Puerto Rico for more than three years now. Only recently have I started to understand the "Big Picture", not

only at my site but for the whole network. It was not until this last Data Management meeting in Hawaii that I realized how much further I still have to go as Data Manager.

I wonder how many other Data Managers ever felt the same. I want to share with you the most important problems I was able to resolve. I do not intend, by any means, to say I found the magic formula for these problems, but maybe my examples will assist some of you finding solutions to your data management issues. After all, that is why we call ourselves a network, because we share common tasks... and problems.

Issue: Facilitate documentation of old as well as current data sets.

Solution Reached: I asked the investigators at my site to collaborate with me in the process of designing the documentation forms. This was after going through a process of looking at the other sites documentation forms and designing a basic form with the help of the guidelines in "Data Management at Biological Field Station", Report of a Workshop, May 17-20, 1982, W. K. Kellogg Biological Station at Michigan State University (there is a 1992 issue). The final outcome were documentation forms that were directed to obtain the metadata in a form that is suitable for ecological data containing entries that have been proved in other sites to be useful and necessary, and that the collaborators at my site were familiar with. The fact that the investigators assisted in the design of these forms showed to be very helpful later on. Being acquainted with the forms made it easier for them to comply with this requirement. They all showed interest and recognize the importance of this time-consuming process.

Issue: Obtain support to solve hardware and software problems one encounters every day.

Solution Reached: After finding out there are certain problems that hardware or software documentation does not help you to troubleshoot, I discovered the DMAN group using the E-Mail facility (my boss talked me into it, and Michener finally convinced me at our last meeting). It is not to say that I am now in love with the process and that it has given me a solution every time I searched for one but it does help. Finally, I was able to understand that it is sometimes easier to get the attention from a group of computer lovers' like us through a computer rather than through a plain phone or written communication.

Issue: Give a structure to Data Management at my site such that I can accomplish the common tasks of Data Management (i.e. serve the scientific community) and play an active role in the research activity done at the site.

Solution Reached: First of all, since the beginning of my job as LTER Data Manager I was able to see the importance of LTER program as well as of the essential role that Data Management plays in it. The problem is one of "feeling" and communication. Luckily enough I did not have to go

through the process of making my principal investigator aware of this because he had to take care of some of tasks I am taking care of now before he hired me. But the need of the other investigators' collaboration was obvious to me since the beginning. So I requested their input, either by to me meeting with them or by writing, for issues such as their need of assistance to manipulate data, what software programs they used to manipulate their data, what was the role they thought a Data Manager should have within the LTER program, etc. I established a direct contact with the investigator and discussed with them my opinion on the role that Data Management has in the LTER program. The ice was then broken, and the rest was a matter of maintaining the communication channels open. My next step was to prepare a work plan converting a year approval and feedback from my principal investigators. The benefits of this go to beyond making myself noticed by the one who manages the budget to; it organizes my ideas and improves the efficiency of my job as well. It also serve as a guideline to write my annual report that is included in our annual report to NSF. I do this every 12 to 15 months, and the feedback. I receive from my two PI'S is always positive and constructive.

Future foreseen problems: Getting myself in the mode of publishing. This is my first collaboration to DATABITS, maybe a good place to start. -- *Eda Melendez, Luquillo LTER, EMelendez@LTERnet.edu*

Data Access Policies -- A Preliminary Report:

History:

Data access policies were discussed at length during the 1990 LTER Data Managers' Workshop. Following that discussion, the LTER Coordinating Committee formed an ad hoc committee on data access. The committee was composed of three Principal Investigators (Hobbie, Magnuson, and Myers) and three Data Managers (Michener, Porter and Stafford) and was charged with formulating guidelines for LTER Site Data Management Policies. The committee formulated guidelines based on principles listed by data managers and investigators during a joint workshop at the LTER All-Scientists' Meeting. The resulting guidelines were then distributed to individual sites (see below).

The guidelines included 10 provisions that should be addressed by an individual site's data management policy and a sample policy which embodied those provisions regarding data access. However, it left the investigators at each site to decide what specific policies should be implemented at their site. It is the purpose of this article to provide a preliminary look at the policies based on the guidelines.

A request for copies of site data management policies was issued in the spring of 1992. Ten sites responded (this includes two sites that did not have formal data management policies). It is not possible to provide details of each of the policies in the limited space here. Instead emphasis will be placed on examining the scope of the policies, the areas they covered and the approaches they took.

Many policies exceeded the scope of the original guidelines. Several addressed administrative issues, formalizing control of data management at a site. Five sites included information on quality assurance, quality control and archival procedures. At all five sites individual investigators submitting data to the site database were responsible for the quality of the data. Four sites included standards for when data should be submitted, which varied from six months after the end of the calendar year in which data was collected to following publication.

Specifics on archival storage of data were included in five policies. Four of the five used a centralized scheme, where archival storage was the responsibility of the site's data manager. At one site individual investigators were responsible for maintaining their own data and documentation. One site included detailed protocols for submitting and accessing physical samples as well as electronic data.

Data access was an important part of all the policies. At seven sites, release of data required the permission of the investigator who collected the data for at least some period of time. At one site data was only available following publication. At two of the seven sites where investigator permission was required, the site PI could also release data.

Time limits for restricting data access were in place at three sites, with limits ranging from 1 to 5 years. Some sites recognized classes of data (typically baseline or meteorological data) which were available upon request to all LTER investigators. Four sites limited the uses to which data could be put to purposes specified on a data request form or restricted the sale of data. Four sites specified procedures for charging for data to recover processing costs. Six sites specified that copies of publications using LTER-derived data should be provided to the site and five specified that data sources should be adequately cited.

This review indicates that there are a number of policy elements which have been incorporated into the majority of site policies. This emerging standard may be useful to other sites in creating or amending their own policies. More information will be required to evaluate the effectiveness of individual policies in achieving the goals of database development and database use. The upcoming All-Scientists Meeting may provide an opportunity for further refining the LTER guidelines for site data management policies. - John Porter, Virginia Coast Reserve LTER, JPorter@LTERnet.edu

LTER Guidelines for Site Data Management Policies

(Editors note: This is a reprint of the guidelines distributed to each site during the fall of 1990)

Each site should develop its own policy in consultation with their investigators and higher administrative units. The following provides general guidelines and rationale, but each site should be prepared to defend its own policy for data management through the site and peer review process. Bear in mind that the general policy of the Division of Biotic Systems and Resources (NSF) is that the data are public property one year after termination of the grant.

General Guidelines

The management policy should include provisions that assure:

- A. the timely availability of data to the scientific community,
- B. that researchers and LTER sites contributing data to LTER databases receive adequate acknowledgment for the use of their data by other researchers and sites receive copies of any publication using that data,
- C. that documentation and transformation of data is adequate to permit data to be used by researchers not involved in its original collection,
- D. that data must continue to be available even though an investigator were to leave the project through transfer or death.
- E. that standards of quality assurance and quality control are adhered to,
- F. that long-term archival storage of data is maintained,
- G. that researchers have an obligation both to contribute data collected with LTER funding to the LTER site database and to publish the data in the open literature in a timely fashion.
- H. that costs of making data available should be recovered directly or by reciprocal sharing and collaborative research,
- I. that LTER data sets not be resold or distributed by the recipient, and
- J. that investigators have a reasonable opportunity to have first use of data they collected.

Example Policy

The following is an example of a policy that will meet these guidelines with respect to data sharing.

Data Type I. Published data and meta data (i.e., data about data). Policy: data are available upon request without review

Data Type II. Collective data of the LTER site (usually routine measurements generated by technical staff). Policy: data are available for specific scientific purposes one year after data are generated.

Data Type III. Original measurements by individual researchers. Policy: data are available for specific scientific purposes after two years of generation. Data can be released earlier with permission of the researcher.

Data Type IV. Unusual long-term data collected by individual researchers. Policy: The principal investigator of the LTER site can designate that such data can be withheld for longer periods. Such action should be rare and justified in writing.

EUDORA - Mail for the Mac!

I'd like to extol the virtues of Eudora (Popmail server for the Macintosh). It's without doubt one of the most convenient and slick packages for e-mail that I know of, and to boot it's free!

Features:

A. Simple logon procedure with manual or automated password.

B. Checks for e-mail at user designated times (I use every hour, but you can choose once a day, every 5 minutes, or whatever pleases your desire to know that someone wants to talk with you.

C. Optional autodelete files from the server (they're stored on your mac in compact form.

D. Hierarchical "mail bins" to organize your messages (or save to text files for word processing, etc.)

E. Compose any number of messages, queue them up, and send them all at once, minimizing time for the application to logon to the server, and transfer messages.

F. Use of nicknames, meaning that I could effectively alias anyone's name and use the alias for the e-mail address. (all this...@LTERnet.edu stuff is what the COMPUTER needs for routing, not you and I.

G. This is the best one: You can attach documents, software, or whatever to your message. It automatically binhex codes the attached document. If on the other end you also have eudora you will receive the message and (your choice) the attached message would automatically be decoded to show

up on the desktop as its original. A good example of its use would be if Bruce Hayden didn't have to use all of those old methods of graphing "functions" in his news letter, CED.

H. The signature appended to this message is automatically sent with each e-mail message (optional)

I. Put an "alias" of the application in your start-up folder, and setup the auto-mated password option and your mail will be automatically delivered to your mac to be read first thing in the morning.

Where to get it:

FTP.QUALCOMM.COM

-- Carl Bowser, North Temperate Lakes LTER,

CBowser@LTERnet.edu

Using a Site-wide Email Calendar

Similar to most LTER sites, we have problems keeping things coordinated between widely-spaced investigators. One tool we've developed to help with this is a reminder service that uses electronic mail to send daily messages to each investigator or graduate student detailing LTER activities for the next two days. Here is an example listing of an LTER-Calendar generated message:

To: bph@Virginia.EDU, lkb2e@Virginia.EDU,....
Subject: VCR-LTER Calendar
Status: OR

This is automated entry from the VCR-LTER Calendar. Entries for the calendar and changes to the mailing list should be sent to jhp7e@virginia.edu. They will be automatically posted one day prior to the date listed and on the date itself.

11/03/92 VCR-LTER PI meeting 3:30
11/03/92 L. Blum returns from shore

Fortunately, this service is extremely easy to implement on a UNIX computer. A below is a simple shell script used to produce our calendar:

```
#!/bin/sh
cd /home/jhp7e/ltercal
\rm -f entry
calendar >entry
if [ -s entry ] ;
then
cat calendarhead entry| \
/usr/bin/mail -s "VCR-LTER Calen." ltercal
fi
```

It uses the UNIX "calendar" program to scan a text file named 'calendar' containing lines of announcements. The calendar program passes lines containing dates for the current day and the next day to an "entry" file. If the "entry" file is not empty, a standard "header" file containing an informative message is prepended to the calendar entries

and the whole business is piped into a mail command and sent. The user 'ltercal' is a mailing group created using aliases in the .mailrc file. The program is executed at 6:05 AM every weekday morning using a crontab entry:

```
5 6 * * 1,2,3,4,5 /home/jhp7e/ltercal/ltercal
```

We are working on expanding the service to include a weekly calendar to be sent on Monday of each week and a monthly calendar to be sent on the first of each month.

-- John Porter, Virginia Coast Reserve LTER,
JPorter@LTERnet.edu

Software tips from Antarctica

1) BULLETIN BOARD POINTER

SYSTEM: Unix

If your local unix machine is using RN or XRN to read bulletin boards, you can set setenv NNTPSERVER LTERnet.edu to read the lternet bulletin boards over internet. You need to deal with bulletin board history information if you want to point to more than one set of boards regularly.

2) BIBLIOGRAPHIC SOFTWARE PRODUCT: BibIX

SYSTEM: Unix

NOTES: Bibliographic system that can be used simply or with great power. We use it as both a stand-alone record of references in our files, of our own group references as well as with nroff/troff documents for report/manuscript inclusion of references. It's free, well documented and the records remain viewable as ascii so the user never feels too far removed from the information.

REFERENCE: BibIX-A Bibliographic Database & Text Formatting System for UNIX R. P.C.Rodgers, Kenneth Gardels, Anat Findelshain, April 1989 Release 1.4 UCSF Pharmaceutical Chemistry CALM/MedIX Technical Report 89-1-4:

INFO/ORDER: From the BibIX documentation manual: BibIX is a licensed software system, distributed by the Campus Software Office, 295 Evans Hall, University of California, Berkeley, CA 94720, (415)643-7201. Information can also be obtained from the authors:

R.P.C.Rodgers, MD
Dept of Pharmaceutical Chemistry,
School of Pharmacy
Laurel Heights Campus,
Box 1204 3333 Calif Street,
Room 102,
San Francisco, CA 94118
(ucbvax.berkeley.edu!cca.ucsf.edu!rodgers)
or
Kenneth Gardels
Landscape Architecture Computer Applications Lab
202 Wurster Hall
Univ of Calif Berkeley, CA 94720

(ucbvax.berkeley.edu!violet!gardels)

3) GENERIC MAPPING TOOLS/Unix

PRODUCT: GMT

SYSTEM: Unix

NOTES: Generic mapping tools (GMT) is a software package to map and display. It's available over Internet at no charge. It's a well-structured unix set of programs that have a very short learning curve to produce land maps with latitude, longitude site overlays.

REFERENCE: Free Software Helps Map and Display Data, P.Wessel and W.H.F.Smith, EOS Transactions, American Geophysical Union, Vol72(41) October 8, 1991, p441+.

ORDER: Connect to kiawe.soest.hawaii.edu (128.171.151.16) using anonymous ftp and copy the binary compressed tar archive "gmtv2.0.tar.Z" from the pub/gmt directory. Uncompress the file and extract its contents at your local site:

```
my_box% ftp 128.171.151.16
Name: anonymous
password: <use your email address>
ftp>cd pub/gmt
ftp>binary
ftp>get gmtv2.0.tar.Z
ftp>quit
my_box% zcat gmtv2.0.tar.Z | tar -xBf
```

More information and demonstration of software use in the article.

4) SPREADSHEET, TEXTPROCESSOR

PRODUCT: AsterX

SYSTEM: Unix

NOTES: We've been using AsterX (spreadsheet, textprocessor, draw) unix xwindow set of programs. It would be interesting to know if anyone has done a comparison with Lotus or Framemaker.

ORDER: commercial product-- Karen Baker, Palmer LTER, KBaker@LTERnet.edu

Publication Information

<i>Editor:</i>	<i>Circulation</i>
John Porter	Rudolf Nottrott
Virginia Coast Reserve LTER	LTER Network Office
Dept. of Environmental Sciences	College of Forest Resources AR-10
Clark Hall, University of Virginia	University of Washington
Charlottesville, VA 22903	Seattle, WA 98195
Phone: (804) 924-8999	(206) 543-8492
Email: jporter@LTERnet.edu	rnotrott@LTERnet.edu
jporter@LTERnet (BitNet)	rnotrott@LTERnet (BitNet)