

01001100 01001100
01010100 01010100
01000101 01000101
01010010 01010010

LTER DATABITS

01001100 01001100
01010100 01010100
01000101 01000101
01010010 01010010

Fall 1990

Data Management Newsletter of Long-term Ecological Research

Another full issue of Databits! We have a feature articles on archival storage at LTER sites, using PC-NFS and a review of two optical disks. Additionally we have reports on the LTER Data Managers Meeting(s) and the ESA Long-Term Ecological Research Section, news from the sites and two new "columns:" "GIS corner" and "I learned it the hard way".

Developing Protocols for the Storage/Archival of LTER Core Data Sets

This article attempts to summarize the experience of data managers across the LTER network with issues of storage and archival of LTER core data sets. I hope such an exercise will help us learn from each other and alert the newer sites to aspects of storage/archival they may not have addressed yet. I want to thank the data managers at the 10 sites who responded to the questionnaire which I sent out; the sites were: AND, NWT, KNZ, HVF, JRN, VCR, BNZ, SEV, CWT, NTL.

There are several different storage media being used at the sites. Reporting on which media are used for storage is complicated by the differences among storage, archival and backup functions and the ways these three overlap at the various sites. If you consider from which medium you would have the most direct and usual access to electronic data, five of the reporting sites use hard disks (sometimes in a network environment), one uses 9-track magnetic tape, one uses floppy disks and magnetic tape, and four have some data on hard disks and some on magnetic tape. A number of the sites are moving or planning to move from a tape based storage system to optical disks. Some sites are moving to online storage rather than using tapes on a mainframe.

A different question is that of the storage medium sites storing core data on hard disks use for backup. The responses included: diskettes, cartridge tapes, 9-track tape, 8mm tape. One site is backing up copies of data in process on Bernoulli disk cartridges. Frequency of backup ranged from daily to weekly to monthly. Responsibility for backup ranged from being contracted out to a Computer Science Department to being the task of individual investigators. Other media in addition to those already mentioned being used for archival of data include: WORM optical disks, RW optical disks, microfiche and, of course, paper.

The number of electronic copies of the core data (in addition to any copy on a hard disk) ranged from two (at six sites), three (at three sites) and three to six (at one site). Almost all the sites maintain copies in at least two different locations. Two sites reported special concern for the safety of the location, with one site archiving diskettes in a fire resistant room and one site shopping for a specialized heavy-duty safe to provide protection of digital media from

fire, etc.

The protocols for replacing tapes ranged from every three to every five years. One site reads and rewrites archive tapes annually and replaces them every four years or by inspection. Advice I have received from our campus computing center and seen in trade magazines suggests ten years as the life span of a magnetic tape, so it would appear that three to five years would be a safe replacement interval. However, several people with replacement protocols in place for tapes verbalized a distrust of magnetic media. Some example quotes are: "[it is] unwise to rely on the integrity of magnetic media for more than a couple of years"; "I am a non-believer with regard to magnetic tape". One site reported a 5 year replacement protocol for diskettes. Three of the newer sites had not yet established a replacement protocol.

All sites except one have experienced media failure. The media mentioned included: magnetic tapes (even relatively new ones), tape cartridges, diskettes, hard drives, Bernoulli cartridges. In almost all cases, a system of multiple backups prevented disaster (i.e., data loss or major recovery costs). One site reported having a major recovery problem early in their LTER experience before the redundant backup system was created.

A number of other concerns were raised by the people responding to the questionnaire. I will summarize these issues below. Perhaps someone would like to address them in DATABITS or at the next Data Managers' Meeting. The issues included:

- accessibility of a storage/archival system. It would be valuable to contrast different media on their ease of use and access.
- automation of the backup system.
- the effect of technical advances, the example being given of the loss of information which resides on computer cards, once a reasonable archival medium. Media redundancy was suggested as insurance against changes. In addition to media changes, data can become inaccessible through changes in data management packages. A cost incurred by several sites was the major effort involved in transferring data when the central campus computing center changed its machine.
- the storage and archival of non-traditional data sets

Although not mentioned in the questionnaire specifically, we are all facing the task of developing protocols for remote sensing images and GIS coverages. How do we back them up, what meta data belongs with them, which ones do we archive, how do we maintain a log of the source of derived images? In addition, there are other non-traditional data sets (e.g., photographs, soil cores, plant and animal collections). How do we keep them in a usable condition?

It was apparent from many of the responses that we feel a need for more information and advice on the issue of storage media. Specific questions mentioned were: what and when to purchase, what are reliable figures on reasonable replacement times for common storage media, what is the latest consensus on optical versus magnetic storage.

The scope of this article does not include a review of the technical literature on these issues though such a review would be timely. I did come across an interesting article contrasting optical disks and magnetic tape when I did a search in Computer Library Periodicals on CD-ROM. The article is entitled "Optical Disks Save Space, Time, Maintenance Costs" and appeared in Digital Review, July 23, 1990. It deals with cost per megabyte, maintenance costs, life span, storage capacity, access times and other pertinent issues.

I hope we can find some mechanisms to share technical information which don't involve a lot of time for anyone. Perhaps we can maintain a list of relevant, current articles in the trade magazines on the LTER bulletin board and any of us could add to this list as we came across good articles. Perhaps we could maintain a list of "experts" (i.e., those with some experience) among the data managers on selected topics. I have appreciated John Vande Castle's information on optical drives sent to us via email.

Hopefully this summary of our existing protocols has helped us become more knowledgeable about what other sites are doing and has organized the many issues involved in establishing a protocol for storage/archival. Let me close by quoting from James Brunt's responses.

THE BOTTOM LINE: "The most important thing is 'having' a protocol that adequately covers your assumable risk and adhering to it. There's not a lot of room for slack here..." --Barbara Benson, North Temperate Lakes LTER

1990 LTER Personnel Directory

It is about time to make our year-end effort to get the latest information for the 1991 edition of LTER Personnel Directory. I will send out the information forms by electronic mail in November. Since most of you have been sending regular personnel updates to the Network Office, probably all you have to do is send back a few corrected entries. Thank you for your help with this. -- Rudolf Nottrott, LTER Network Office

From the Sites

BNZ--Our field season has ended, and everyone is in analyzing data and generally trying to catch up after the summer. We are anxiously awaiting the arrival of Mark

Klingensmith, our new Data Manager. He will be on board sometime in December. The campus networking is progressing, but ever so slowly. It still appears that we might be networked any day now. Our SUN 4 is still living in temporary quarters and has not moved to its final resting place. But it is happily functioning where it is as our GIS machine. The correct digitizer finally arrived and is functioning now as a part of our GIS system. Our three SUN 386i machines will soon be traded out for three SUN SPARC 1's.

That is the news from here, but I have one more thing I would like to mention. The LTER Climate Committee plans to begin using e-mail to circulate information of current weather and climatic interest. It will be sent to those on the "climate" mailing list. Please pass the word to each site, so that anyone who is interested can request to have their name added to the "climate" mailing list.

-- Phyllis Adams, Bonanza Creek LTER

HFR-- Two major projects for the upcoming year are installing a LAN at the Harvard Forest and connecting it to the campus network and the Internet, and developing a long term plan for data management. We also hope to continue work on computer modeling of hurricane winds, and implementing statistical tools for studying correlation of spatial data. --Emory Boose, Harvard Forest LTER

KNZ--1) Our Local Area Network has now been functional for over a year. In addition, our Novell network has been modified such that telnet can be run from any PC workstation. For the first time, most of our investigators are using electronic mail on a daily basis. I would like to thank Rudolf for a job well done on the electronic mail forwarding system. Our investigators enjoy it!

2) Unfortunately, our full time network manager Cornell J. Kinderknecht has been hired away by the private sector. Thus, come the first of the year we will be looking for an individual to take his place. Hopefully, I will send out a job notice within a month or two. Let me know if you know of any good people.

3) With our last supplement, we bought some network software, increase our GIS/ Remote sensing capabilities and purchased some workstations and in this next year we will purchase an optical disk drive for our data archive.

4) Has anyone installed PC ARC/INFO on a Novell network? I will be "trying" to set up a system within the next three months and would like to know of any and all pitfalls. -- John Briggs, Konza LTER

NTL-- Recently an interesting research context has presented useful challenges for data management. Now that our site has collected ten years of data on many components of our lakes, questions are being posed which were not feasible before. For example, there now exists enough fish data to see two different pulses of strong year classes of yellow perch in Crystal Lake. A team of NTL researchers has been examining "top-down" versus "bottom-up" control of water clarity. This research has required the data management staff to access a variety of different databases: precipitation, groundwater, lake chemistry, chlorophyll, zooplankton, and fish.

As we amassed the data in suitable formats for analysis, we found areas where our existing data management worked well and areas that needed attention. Importantly, individ-

ual researchers also came to focus on data management and feel commitment to improvements. As a result, one data set which had not been computerized has been entered and earlier data for a different data set are being integrated into the database.

In the design of our sampling scheme and data management system, much thought was given to facilitating research which would cut across disciplines. We have found both in the current research effort described above and in the VARNAE workshop that these broader research endeavors provide a constructive test of our data management system.-- *Barbara Benson, North Temperate Lakes LTER*

VCR-- The size of our data management staff has doubled with the addition a graduate assistant data manager. Ramesh Caushik is a graduate engineering student who will be helping out with the day-to-day tasks of data management. Our Remote Sensing an Geographical Information System Laboratory recently moved into the largest office in the building (formerly the Department Chairman's office) and our SUN systems have been upgraded with two 24-bit graphics cards --*John Porter, Virginia Coast Reserve LTER*

Using PC-NFS

To enhance the Internet connectivity of our IBM PC compatible equipment, we have begun installing PC-NFS from Sun Microsystems. Previously, our PCs were connected using the NCSA Telnet software and an Ethernet card, enabling them to act as virtual terminals on any Internet host, to transfer files to or from any Internet host, and to act as FTP file servers on the Internet. Even though this mode of operation has been useful, the need to interrupt any other work on the PC in order to get a file from (or put a file onto) another host has been perceived disruptive.

PC-NFS has solved this problem (within the limits of what the single-task DOS operating system can achieve) by implementing the Network File System (NFS) protocol originally defined by Sun and now common on most Unix and DEC VMS machines. In addition to the Telnet and FTP programs, PC-NFS provides transparent access to "remotely mounted" files and printers on other Internet hosts. Given proper authorization, any directory available on an Internet host can be mounted on a virtual drive of the PC (usually about 15 virtual drives are available on a PC). For example, directory LTERnet:~ftp/public could be mounted on a PC anywhere on the Internet in such a manner that it would appear as drive E: on that PC. File access permissions are given on the basis of passwords (those on the host exporting the files). Mounting a printer remotely is just as transparent, even though this may not make sense for printers that are further away as your building. PC NFS provides an easy way to share printer resources.

We have installed PC-NFS on two PCs now, with three more to follow. Both PCs access files via Ethernet from a VAX (LTERnet), a Sparc (GIS server) and a generic 386 PC (general file server). I am also planning to install PC-NFS over a modem connection (Serial Line IP, SLIP). Contact me if you are interested to hear more about our experience with this networking setup. I am also interested to hear from anybody who has looked at similar products (e.g. FTP Software, Inc.) --*Rudolf Nottrott, LTER Network Office*

1990 LTER Data Managers' Meetings

The LTER Data Managers met at the All Scientists Meeting (ASM) in Estes Park, September 28, 1990. Twelve LTER sites and 5 agencies were represented. Collectively, we made efficient use of our two-hour allotted time slot! The "Orchid Report" was distributed, summarizing the results from the 1990 LTER Data Managers Meeting.

The ASM Data Managers Workshop meeting was divided into 4 categories; status reports and updates on current projects, 3 working group sessions, a presentation on DOE PARKNET, and a final planning session for the 1991 Data Managers Workshop. A brief overview of the highlights of each session are given below.

STATUS REPORTS AND UPDATES ON CURRENT PROJECTS:

1. The LTER Core Data Set Catalog is ready to go! The original estimate of 200 pages expanded to 330 pages! It includes site descriptions and site maps from all 17 sites and a color cover. As a result of the increased size and fixed budget no more than 300 catalogues could be printed. The planned distribution will be 100 for NSF, 100 for the Network office, and 5 per site. The printed version will be available in mid-November to early December. There was some discussion about periodic updates to the catalog. Rudolf Nottrott (NET) will be maintaining an electronic copy of the catalog on the network office computer.

2. Plans for the Data Management Symposium are underway. The emphasis will be on global and regional efforts of managing large, environmental data sets. There will be 4 core areas: conceptual basis/needs assessment for global and regional scientific databases (John Pfaltz is a potential speaker); a "Nuts & Bolts" section on QA/QC, etc.; Archival storage; and Spatial Data Management and GIS. Abstracts will be needed in fall, 1990. Invited speakers will be asked to bring a copy of the paper to the meeting. Springer Verlag has expressed interest in publishing the Symposium volume.

3. The Site Bibliography project was discussed. A major proposal (~\$100K) would be needed to create a keyworded bibliography. It was recommended that this proposal have a strong science basis. Each site would need someone to oversee keywording. A question was raised whether a smaller proposal had been suggested. It had been, but the LTER Coordinating Committee (LTER/CC) suggested a larger, comprehensive approach. One site commented that their institute had a keyworded bibliography but it was a lot of trouble to maintain. We discussed whether it was worth starting with a small set of keywords indexed by site. The group consensus was no, it needed to be well indexed to be useful. An action item was to try writing another proposal. Jim Halfpenny (NWT) will take the lead on this.

4. The Climate Distributed Database Project is underway. This project, the outcome of a working group at the Snowbird meeting, is a distributed climate database to be accessible via e-mail. Plans for implementation require a gatherer server to translate e-mail requests into requests from individual sites, handle merging and manipulations, and return the requested data. Money (about \$100 K) is needed for the gatherer server. More money will be eventually needed for individual site capabilities (although not all sites will need this capability right away). There is a possibility

of developing software for formulating data requests (for sites with UNIX). It is paramount that the scientific value of this endeavor be emphasized in the proposal. It was suggested that John Gorentz (KBS) marshal his arguments and get letters of support by prospective users. The identification of prospective users is critical to show its potential usefulness. Dave Greenland (NWT), Climate Committee, said that his group liked the concept and added their formal endorsement of the project at the LTER/CC meeting. The most exciting potential use of the distributed database is by modelers. We discussed the need for people to know errors, the nature of the data, etc. There was some discussion about checking into integrating with historical climate databases.

5. The "China Connection" was discussed. Initial comments from Dr. Zhao Shidong, Chinese Ecological Research Network (CERN) and Beryl Leach, National Academy of Science, have been very positive. Susan Stafford (AND) and Barbara Benson (NTL) are writing a preproposal outlining our interest in giving a training workshop for CERN data managers in the US. Jim Gosz (SEV) reported that a high ranking delegation of CERN scientists were planning to come to the US in May, 1991. Their itinerary includes visits to several LTER sites; specifically KBS, AND, SEV, KNZ, and CPR. The current plans calls for 2 days per site with special emphasis on GIS activities.

6. Plans for the Monitoring of Natural Resources Conference, May 6- 8, 1991, at the Andrews Site are underway. The final agenda will be set by early November. Anyone with specific suggestions, questions, and/or contributions, should contact Susan Stafford (AND) or Art McKee (AND) as soon as possible.

7. Proprietary Issues was the topic of a separate data sharing workshop at the ASM. The seed for this topic originated at the spring KBS workshop and was further developed at the Snowbird meeting. There was additional elaboration of the pros, cons, and legal issues at this meeting. Judy Meyer (CWT) made a presentation of a set of Network Guidelines to LTER/CC. It states that each site develop site specific policies which address timely availability of data, adequate documentation, public access to core data, death or departure of PIs, quality assurance and quality control, and assuring adequate acknowledgment. Sites should review this set of guidelines and return their suggestions to Judy Meyer (CWT) before the January LTER Executive Committee Meeting.

8. In terms of Outreach Activities, Rudolf Nottrott (NET) reported that he has received letters from ESA (and specifically the long-term studies section) requesting extension of our mail forwarding system. Rudolf has replied that he would be pleased to make this available if additional resources are available. Ideally, this type of arrangement should extend beyond simply mail capabilities to scientific collaboration. It was suggested that the larger group of ESA (6,000 members) may need to implement a Bulletin Board. The consensus was that the highest and best use of Rudolf's time and expertise is on innovative technological advances of network capabilities. The custodial role of maintaining other organizations' mailing lists should be assigned to someone else using new, outside funds.

WORKING GROUPS

1. META-DATA

A draft copy of the minimum standards for a Data Set Abstract, prepared by Don Henshaw, Susan Stafford, and Gody Spycher (AND) was distributed and discussed. This was designed for PI's to let them know about the minimum standards. Specific comments on revising the draft document included giving a data set title instead of study title; adding a contact person, date, and georeferenced site location; and recording the location of the raw physical copy of the data as well as the location and form of the electronic data. Some discussion followed about whether it was practical to track the publications resulting from a particular data set. It would be difficult to keep it updated and it may not be appropriate for a minimum set of documentation. Some sites combine experimental design with methods. At virtually all sites, there are higher levels of documentation. This document was merely proposed as a guideline for a minimum set for the Data Set Abstract, i.e. a checklist, that could be distributed to data managers to use to leverage additional documentation from reluctant PIs. It was decided that this list would not be distributed to PIs as these are MINIMUM criteria for a data set abstract and in some cases, could be used to the detriment of the data management activities when additional detailed documentation is requested. There was some discussion about whether or not this form should be required for the data set abstract. It was determined that this should serve as a guideline for the network. Data managers are encouraged to review the text and convey any additional suggestions to Don Henshaw or Susan Stafford (AND). Tom Siccamo (HBR) made an interesting suggestion about doing an intersite QA/QC analysis on our data management procedures. This may be something we would like to pursue.

2. SITE REVIEW CRITERIA

A working group discussed criteria for site review of data management. Because a major emphasis of the LTER program is the accumulation of long-term databases, we felt that it was crucial that data management be evaluated during a site review. The neglect of this evaluation could be misinterpreted by investigators as an indication that data management is not considered important by NSF. In addition, such a review encouraged adherence to high standards. The working group compiled a list of areas to be evaluated. These included such topics as documentation, safe archival, policy for release, accessibility (electronic form, ease of query, etc.), formal statement of data management protocols, and specimen/sample preservation. It was also suggested data managers be included on site review teams to make cogent data management suggestions. Barbara Benson (NTL) will edit this list and email a copy to all data managers for comments. The resulting draft will be presented to the LTER/CC and following their approval, it will be sent to NSF as a set of recommendations for the Site Review of Data Management.

3. NEW DIRECTIONS FOR THE 1990'S

Long term, data publication is an issue we will be facing. There are already several models within the network; NWT publishes data at cost and NIN provides data summary books. In the future these could be available on CD's. Peer review is very important. Having problem/question driven databases is also an option. The LTER Network would be in a unique position to facilitate this.

Eventually the Catalog should be extended to all data sets - not only the core LTER data sets. This would allow a wider user base. This could be accomplished possibly through a secondary catalog. This will also promote data sharing. There had been some opposition to making data available during the IBP but today there is much enthusiasm for making problem/question driven databases more accessible.

There is a need for support for UNIX system administration across the network. To begin with, it would be helpful to have a list of regional experts to provide system support. We need to find a mechanism to provide this level of UNIX support. One possibility would be to add a day to the data management meeting for UNIX administration.

A Proposal for Biological Databases needs to be made. Some are already in progress, i.e. developing online databases with the distributed climate database. The Symposium will also address this. We need to check with the Committee on Earth and Environmental Studies (CEES) to forge links to the greater ecological and scientific community. We need to continue to attempt to jointly identify places where the LTER data managers can participate with research workshops; perhaps by establishing more formal protocols and workshops in ecology which will need programming and computational support. We also need to recognize that some scientists may want to do everything themselves.

PARKNET

Dale Bruns, from the DOE National Ecological Research Parks (PARKNET), gave a brief overview to the PARKNET program. (Recall Walt Conley briefed us on PARKNET at the Snowbird meeting). John Heurer (Savannah River Research Park) was introduced and is playing a key, leading role in the development of PARKNET data management activities. There are several DOE sites and a series of computational workshops have been planned. A network (email) query system has been developed with the assistance of Walt Conley and Rudolf Nottrott (NET). PARKNET has a low-tech online database system. Next year's LTER Data Managers meeting may include a joint DOE PARKNET component.

1991 DATA MANAGERS AGENDA

On the list for possible agenda items for the 1991 meeting are the following:

- * Protocol development of Remote Sensing/GIS archival
- * Quality Assurance
- * Network Security
- * Meta-Data
- * Data Publication of Question-driven Data Sets

- * Expansion of Data Catalogue to Non-Core LTER Data Sets

- * Pressing Issues in Data Management: NSF Proposals to Robbins Pgm

- * Facilitation of On-line Database: Electronic Climate Database

- * Establishment of Network UNIX Support Group

As you can see, once again we have set ourselves an ambitious agenda! The list is longer than the time available. The LTER Data Managers Task Force (Barbara Benson (NTL), James Brunt (SEV), John Porter (VCR), Bill Michener (NIN) and Susan Stafford (AND) welcome your suggestions as we begin planning for the 1991 Data Managers meeting. As of this writing, the 1991 Data Managers Meeting will be held next summer preceding the ESA meeting in San Antonio, Texas. --*Susan Stafford, Andrews Forest LTER*

GIS Corner

Activities of the GIS Working Group-- Last summer a survey was conducted of GIS and remote sensing capabilities across the LTER network. In September, David Foster and Emory Boose (HFR) organized a workshop at the All Scientists Meeting that focused on the results of the survey and on research applications of GIS at each site. Participants agreed on the need to share information on GIS developments, perhaps through the LTER bulletin board. A report on the current status of MSI technology in LTER is being prepared, with assistance from the Network Office and the Connectivity Committee, and will be distributed later this fall to the sites and to NSF. -- *Emory Boose, Harvard Forest LTER*

ERDAS MIA -- This is a short little memo about the new ERDAS module called "MIA" Multivariate Image Analysis. MIA is a hybrid from of image analysis which allows users to interactively delineate interesting pixel groupings in a feature space. Principal component scatterplots provide the key and access to the data structure. I am finding it an excellent way to explore principal component analysis in a very graphic way and I hope to use it to "tease" out some minor components of the vegetation that get "swamped" by the dominant classes. So far it appears to have some promise. --*John Briggs, Konza LTER*

Simplifying Window Management with ERDAS -- When running ERDAS on a SUN computer it is necessary to frequently switch between the menu and display windows. One way to avoid "switch fatigue" is to hook a terminal or PC emulating a VT100 terminal to the serial port of your SUN. You then login on both the main console and the terminal. The console can be used entirely for ERDAS displays while the terminal displays the menu. You can also add a second SUN display -- but that is much more expensive than adding a terminal! --*John Porter, Virginia Coast Reserve LTER*

Getting PC more memory-- PC ARC/INFO users can use all the memory they can get (I'll say!) because ARC is a real hog. There are several extended memory managers that can ease the crunch. We've been using QEMM-386 (Version 5.0) from Quarterdeck to force everything we can into high memory. This process is not for the meek, the easy install took about a week (Greg, Tom, and I know a lot more about pc memory allocation now - aahhh UNIX!). We have successfully put most of the system drivers and PC-NFS into high ram. We have not been able to put the VDI/CGI drivers into high ram. However, ARC News Summer 1990 reports that one memory management system has been successful at this - Hi386 AMS Upper Memory Manager available through Global Systems Modelling, Ltd. (they wrote the article, isn't capitalism great). It is also reported that up to five simultaneous ARC/INFO or sub-module sessions have been run for many hours without problems using Quarterdeck's DESQview (Version 2.26). It is not reported how well DESQview works with Hi386. QEMM-386 requires at least 1.5 Mb of memory (kind of strange since all our machines will only run on 1, 2, 4, etc). --James Brunt, *Sevilleta LTER*

Cleaning Tektronix Printers -- Our Tektronix 4696 ink-jet printer was nearly useless. Clogged jets were becoming a way of life. No sooner would one clog abate than another would appear. After consulting with a Tektronix technician we had the key to cleaning -- distilled water!

It turns out that the ink is water soluble. Over long periods of time crystals can form through evaporation of the ink. The crystals then clog the print heads. Adding distilled water dissolves the crystals and clears up the clogs. There is an additional hint: turning OFF switches 2 and 3 on the DIP switch on the back of the printer then holding down the 'select' and 'feed' buttons and turning ON the printer will put it into a clean and test mode. It will go through a cleaning cycle, print some test lines, then go into another cleaning cycle. This continues for 100 iterations -- though only the most stubborn clog takes more than a few. --John Porter, *Virginia Coast LTER*

If you have a hint or information for the GIS Corner, send a copy to Emory Boose. His e-mail addresses are: EBOOSE@LTERNET.WASHINGTON.EDU (Internet) or EBOOSE@LTERNET (Bitnet)

Data Access Guidelines Formulated

During the LTER Coordinating Committee meeting at Estes Park an ad hoc committee on data access was appointed. It was comprised of three principal investigators (Myer, Hobbie and Magnuson) and two data managers (Michener and Porter). The committee met in Estes Park and compiled a draft set of guidelines. These guidelines address, in a general way, objectives for data access. Each site will be expected to formulate a data management policy that contains specific policies to implement the general objectives stated in the guidelines. *For more details on the contents of the guidelines, see Susan Stafford's report on the Data Managers' meetings.*

Currently the draft guidelines are being circulated for comments to the PI's at individual LTER sites and should be finalized at the Executive Committee meeting in January. -- John Porter, *Virginia Coast Reserve LTER*

Review of Optical Disks

At the LTER Data managers meeting in Snowbird this year the subject of the status of optical disks came up once again. Discussions from sites suggested that a number of options still exist. The optical disks from Pinnacle Micro and Alpatronics appear most useful. Both companies have versions for most types of computer systems. Disks MAY be interchangeable between the same operating systems. Pinnacle-15265 Alton Parkway, Irvine CA 92718 (800) 553-7070 Alpatronics-2300 Englert Drive PO 13687 Research Triangle Park NC 27709 (919) 544-0001.

Both are rewritable using Sony drives and 600Mb (300/side - 1 side at a time) 512byte/sector disks for most machines.

Bill Michener at NIN-LTER has been using the Alpatronics (3) on a Sun. We have been using a Pinnacle on a Sun SPARCstation and 80386 PC under DOS. We have had problems using it under Interactive Unix on the 80386-PC.

I had Alpatronics stop in for a demo to compare it to the Pinnacle. Both drives worked the same. They are fairly slow for writing and average disk access is about 65ms - fast enough, but not fast enough for your main or only disk drive. The most important part - both systems can read each other's disk. I popped a disk written on the Pinnacle into the Alpatronics, it worked fine, and so did the other way around. This is probably only true for similar file systems. This does mean that files could be sent across the network when needed...

Alpatronics does market software to be able to use disks from other file systems (like DOS or DEC-VMS) - This is still a to-be-released product (12/90) on the Sun but exists for the VAX. On the Sun, Alpatronics looks just like a standard "sd" disk drive. They have special (copy protected) formatting and setup software. The Pinnacle has its own drivers and formatting software. I have, however used the Pinnacle as an "sd" disk - but I don't recommend it. (Both drives however, can be used as a boot device.)

One word of warning - 300Mb/side is not all that much with images and GIS files. I would expect larger capacities in the future... Maxtor has a drive that is much faster (it uses a Tahiti subsystem) and can write non-standard larger format disks and perhaps (I have never used one) the standard ISO/ANSI 512 or 1024 byte/sector disks.

Some sites (like NTL) have been using WORM drives like the 200mb IBM for the PC. These are also great, but once written, the files are unchangeable (they can be delete, but the disk space remains used). Being able to move 200-300Mb of random access disk on and off systems is better than trying to use tape... For most purposes the disk can be used just a another hard disk...

A word of warning however. One person in Forestry has an early version IBM WORM drive. It cannot read disk written on new versions of the same drive... Hopefully the ISO/ANSI standards have solved this type of problem, but I wouldn't count on it.

The Remote Sensing database from the satellite acquisition will be archived on the Pinnacle Optical and to-be-purchased secondary Alpatronics drive.

I would appreciate further information/experiences or comments. --John Vande Castle, LTER Network Office

Editor's Note: Subsequent to John's initial article his Pinnacle disk drive failed (still under warranty) and the warranty service was much slower than expected. Following circulation of an e-mail message describing the problem, a new drive arrived. Contact John directly for news of any late developments.

I Learned it the Hard Way

Mounting Bad File Systems-- Unix users and superusers remember to ALWAYS do a file system check fsck(8) on a partition you are intending to mount from the command line with mount(8). Mounting a corrupted file system WILL crash THE system. If possible do not fsck a file system that is mounted, if you have to, address the character device (eg. /dev/xy0a) and not the raw device (eg. /dev/rxy0a). The raw device is usually checked because it much faster. --James Brunt, Seville LTER

Taking the Sting Out of rm * -- On our UNIX system we had repeated problems with users accidentally deleting entire directories of files because they used the 'rm *' command carelessly or were in the wrong directory. With backups we were able to save all but a few hours work. However, adding the line 'alias rm rm -i' to users' .cshrc files has eliminated the problem. A user is then prompted for each individual file before it is deleted when the rm command is used. --John Porter, Virginia Coast LTER

If you would like to contribute to a future "I learned it the hard way" column, send contributions to James Brunt (JBRUNT@LTERNET.WASHINGTON.EDU). It's a place for confessions by would be guru's and will feature short quips and tips that involve "struggles." These "struggles" can be expounded up on or simply alluded to.

1990 Long-Term Ecological Research Section Meeting

The Long-Term Ecological Research Section of ESA met during the annual Ecological Society of America meeting in Snowbird. It is reviewed here because it touched on many issues (notably adoption of standards) with potential impacts on LTER research data management.

The meeting began with selection of new officers. A slate consisting of Bob Parmenter (Chair), Nalini Nadkarni (Vice-chair), Bob Waide (Sec.), J. Franklin and T. Stohlgren (Counselors) was proposed by the sitting leadership and elected without revision.

Symposium topics for the next couple years ESA meetings were suggested. A one-half day symposium on plant mortality was proposed by Mark Harmon and endorsed by the section for 1991. Tentative proposals for 1992 included models of long-term processes, standards and land use change. Additional symposia can be proposed directly to the section leaders. Finally, a questionnaire will be sent out to all section members, along with an LTER newsletter.

The discussion of standards began with a brief introduction by Jerry Franklin and formal presentations on decomposi-

tion, tree gaps and data management. Jerry noted the advantages of doing things in the same way in the field and laboratory and managing data in a common database for comparative analysis. The disadvantage of standards was the freedom lost by individual ecologists to do things "their way." The decomposition presentation focused on the LTER-wide decomposition experiment coordinated by Mark Harmon. The tree gap presentation discussed the establishment of protocols for measuring and classifying tree gaps. In two workshops spanning four years, standards were worked out for measuring gap size, criteria for determining when a gap was closed and other important characteristics. The data management presentation focused on the LTER MSI (either "Minimum Standard Installation" or "Minimum Station Installation" depending on who was talking) consisting of a GIS system, network connections and high-capacity storage.

Bill Michener was also given time to mention standards within LTER data management. He focused on the upcoming report from the data management workshop at Kellogg Biological Station as being a good place for learning about emergent standards in data management, both within LTER and the field station and marine lab community as a whole. John Vande Castle also spoke briefly about standards for GIS file formats. Jerry Franklin concluded this portion of the discussion by noting that many standards are neutral and we just need to pick one and stick to it.

A wider discussion on what parameters and standards were needed ensued. There was also a discussion of what actions the section could take to advance standardization. The final resolution of the discussion was that the section would take two actions. First, ESA would be approached about endorsing a standards manual. Secondly, the section would adopt as a project the compilation of a catalog of existing standards. --John Porter, Virginia Coast Reserve LTER

DATABITS Publication Information

Editor:

John Porter
Virginia Coast Reserve LTER
Dept. of Environmental Sciences
Clark Hall, University of Virginia
Charlottesville, VA 22903

Phone: (804) 924-8999

E-mail: jporter@lternet.washington.edu (Internet)
jporter@lternet (BITNET)

Circulation:

Rudolf Nottrott
LTER Network Office
College of Forest Resources AR-10
University of Washington
Seattle, WA 98195

Phone: (206) 543-8492

E-mail: rnottrott@lternet.washington.edu (Internet)
rnottrott@lternet (BITNET)

LTER DATABITS is produced by integrating text from electronic mail messages and direct file transfers using WordStar version 5.5 to produce a Postscript file at the University of Virginia. The Postscript file is then transferred electronically to the Network Office at the University of Washington for printing and distribution. Electronic copies of DATABITS are available from the Network Office upon request.