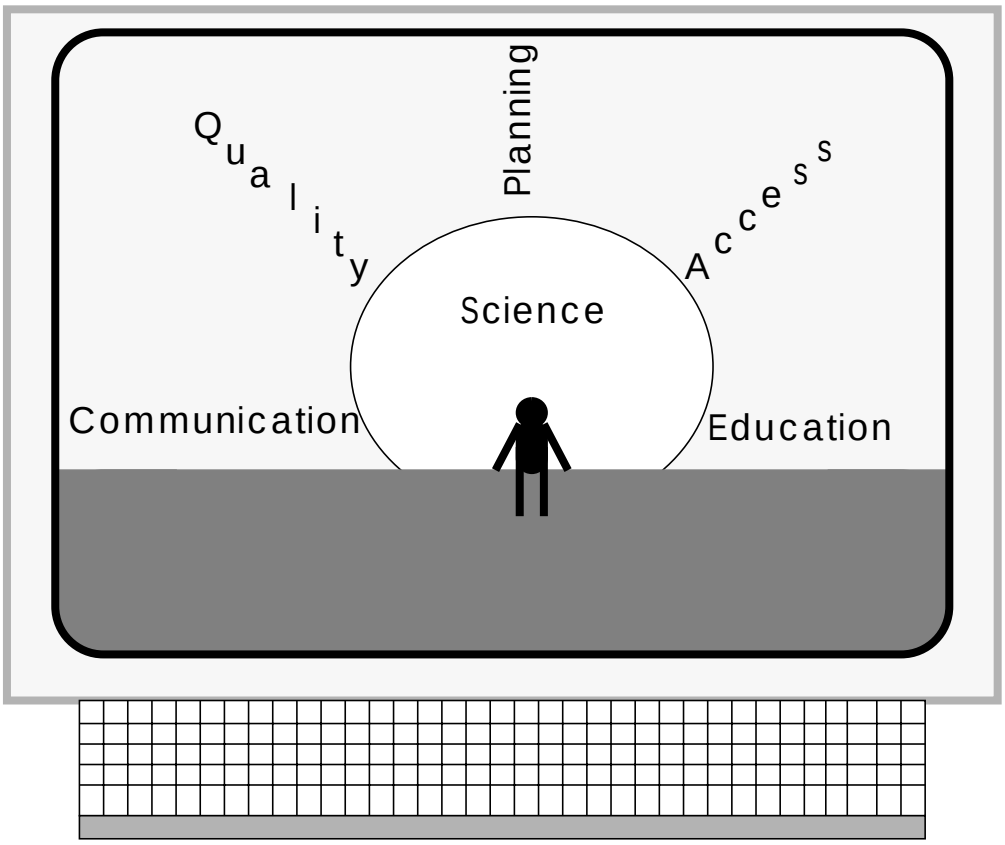
Proceedings of the 1992 LTER Data Management Workshop



EDITORS: R. Nottrott and J. Porter

CONTRIBUTORS: M. Andrews, K. Baker, B. Benson, E. Boose, J. Brunt, C. Bledsoe, J. Briggs, G. Calabria, A. El Haddi, D. Henshaw, R. Ingersoll, T. Kirchner, M. Klingensmith, L. Krievs, J. Laundre, R. Lent, E. Melendez, W. Michener, B. Nolen, R. Nottrott, J. Porter, S. Stafford, C. Veen

Long-Term Ecological Research Network (LTER)

Proceedings of the 1992 LTER Data Management Workshop

Editors: Rudolf W. Nottrott and John H. Porter **Contributors:** M. Andrews, K. Baker, B. Benson, E. Boose, J. Brunt, C. Bledsoe, J. Briggs, G. Calabria, A. El Haddi, D. Henshaw, R. Ingersoll, T. Kirchner, M. Klingensmith, L. Krievs, J. Laundre, R. Lent, E. Melendez, W. Michener, B. Nolen, R. Nottrott, J. Porter, S. Stafford, C. Veen

1.0 Executive Summary

The 1992 Long-term Ecological Research Data Management Workshop was held August 7-9 in Honolulu Hawaii. A total of 23 data managers and other interested scientists attended the workshop, including at least one representative from each LTER site. The workshop focused on two major areas. First, reports were presented on activities coming out of the 1991 workshop. Secondly, working groups were established to formulate plans for the coming year and to address issues of LTER-wide concern.

1.1 Data Management Activities since the 1992 Workshop

1. The DM symposium proposal **"Environmental Information Management and Analysis: Ecosystem to Biosphere Scales"** has been funded. The symposium will be held in May 1993.
2. LTER **Core Dataset Catalog access was broadened** by implementation of query functions via electronic mail and directly over the Internet.
3. **Connectivity within the LTER Network has been improved**. The Sevilleta field site received a T1 Internet connection (1.54 MBit/second). The LTERnet Network Support System was connected to Sprintnet. The LTERnet->Sprintnet, part of the recommendations of the Connectivity Report, provides a new way to connect to the Internet from remote locations anywhere in the U.S and abroad.
4. A **survey** of "**The Management of Electronically Collected Data Within the LTER Program**" was carried out and published as a report. The evaluation of the management of electronically collected data should be considered an ongoing process. The report can be brought up-to-date periodically with the format consisting of edited tables with a brief addendum comprising the text.
5. A **survey** on the **status of data accessibility by researchers at LTER sites** was carried out. A complete summary will be published in the Fall issue of LTER DataBits.
6. A grant has been received from NSF by Sevilleta researchers James Gosz and James Brunt to organize a **research data management training program for the Chinese Ecological Research Network (CERN)**. Four data managers will be in Beijing in September assisting CERN scientists with research data management issues.These efforts are part of the planned continuing collaboration between LTER and CERN.

7. Data managers at the Andrews Experimental Forest site submitted a **proposal to NSF requesting funds to create a conceptual design for a User Interface for the Forest Science Data Bank (FSDB**) at Oregon State University. The Interface Design will emphasize generic **support for basic data activities and access in a heterogeneous network**.
8. **GPS field campaigns** were conducted with strong participation by data managers at several sites. A network-wide proposal for a technical supplement to support the **purchase of state-of-the-art GPS data processing software** and workshop was funded by a NSF technical supplement. The **workshop on processing of GPS data** will be held at the NTL LTER site following the 1993 Ecological Society of America meeting.

1.2 1992 Workshop Accomplishments

1. A **10-year plan was developed outlining the goal for Scientific Data Management in ecological research**. Four broad categories are considered in the plan: (1) training and education, (2) communication and resourcing, (3) technology, and (4) developing a new mindset.
2. An outline was developed for a **presentation on data management at the 1993 LTER All Scientists meeting**. The presentation will have a network orientation emphasizing the evolution of data management as the LTER network developed from a collection of individual sites to a network of sites. The link between data management and research is crucial to this presentation.
3. **Data exchange tool development.** Specific plans were made to implement a standard inter-site data exchange format. When fully-developed, the format will facilitate the exchange of data and documentation between sites and the development of data management tools within the LTER network. This development should help to reconcile the need for individual approaches to data management within sites (to accommodate site-specific circumstances), with the need for an easy way to exchange data across sites. This approach is also complementary to the development of interactive data access systems which was extensively discussed at the 1991 workshop.
4. The **LTER all-site bibliographic database** discussed previously will be explored and developed as a **pilot project** using data from several LTER sites. The pilot project will serve to demonstrate the benefits of the database and to resolve technical details relating to access and maintenance methods. Envisioned uses of all-site bibliographic database for individual sites, the LTER network as a whole, and the wider ecological research community are wide-ranging. LTER scientists could more easily plan multi-site experiments and synthesize information if site research information were more accessible. Scientists from the wider ecological research community can compare their data or sites to similar LTER data or sites, developing more robust, general ecological principles. A bibliographic database can be cross-linked to the LTER Core Data Set Catalog, providing researchers with access not only to citations, but also to descriptions of specific data sets. In addition, agency administrators, decision makers and managers will have a very large body of ecological information identified and indexed for their use.

5. **Plan for the development of Network Resources**: Emphasizing the crucial role of the LTER network in sharing research tools and expertise, the data managers decided to develop a plan that outlines specific resources targeted for further development, or to be newly created and maintained in the network. These common resources will be useful in **supporting ecological research** and will be freely and easily available to researchers at all LTER sites, and beyond that to researchers in the wider ecological community (e.g., OBF/SAML, ESA LTSS). The list of network resources targeted for development has 3 general categories: (1) on-line documents, (2) a data management toolkit and (3) on-line databases (Catalog of Long-Term Datasets, intersite bibliographic database, LTER Personnel Directory, intetrsite datasets not specifically maintained or archived by any site). The existing on-line LTER Catalog of Core Datasets will be expanded beyond core data sets to contain all long-term datasets at the sites.

2.0 Activity Reports

2.1 Environmental Information Management Symposium (James Brunt, SEV)

The Environmental Information Management and Analysis Symposium has been funded by the National Science Foundation. Approximately 200 abstracts were received and 37 have been chosen for inclusion in the program. Symposium participants include leading scientists from a variety of disciplines, including ecology, geography, computer science, remote sensing, and statistics. Primary objectives of the symposium are to (1) identify state-of-the-art technologies for managing and analyzing environmental data collected at ecosystem to global scales and (2) to facilitate future long-term and broad-scale research efforts. The international symposium will include interdisciplinary discussions of long-term and broad-scale research efforts, relevant technologies, and future directions for information management and analysis. Technical sessions will focus on Geographic Information Systems and scientific databases (e.g., design and development, management, quality assurance/quality control, data transfer, and standardization) and analytical approaches for addressing questions at ecosystem to global scales (e.g., spatial pattern detection, scaling issues, expert systems, and linking GIS and remote sensing to ecological simulation models. The meeting will be held at the University of New Mexico in Albuquerque from Thursday, May 20 through Saturday, May 22, 1993.

2.2 GPS Network Proposal (Bill Michener, NIN)

The Technical Supplement for Network-wide GPS software was funded. This proposal will support purchase of 3 software packages during 1993 which facilitate post-processing of high precision data. In addition, a GPS data processing workshop is tentatively scheduled to occur at the North Temperate Lakes site following the ESA meetings.

2.3 LTER DM Site Reviews

John Briggs (KNZ), Bill Michener (NIN) and Susan Stafford (AND) participated in the site reviews of ARC, BNZ, JRN, SEV, HFR and found the experience worthwhile. They strongly recommend that LTER Data Managers avail themselves of opportunities to participate in future reviews. A common problem during the site reviews was that not enough time could be allocated for the data management to permit a comprehensive analysis. The participating data managers partially dealt with this problem by spending time on the issues after dinner or by arriving before the official beginning of the review. The large contrast between some sites in the amount of historical (paper) data can make cross-site comparisons difficult. Data managers on the review team can play a crucial role in identifying commonalities.

2.4 Coordinating Committee Meetings (Rudolf Nottrott, NET)

Two **Coordinating Committee meetings were held in 1992, at Trout Lake Station (NTL) Jan. 29-March 1, and in Fairbanks (BNZ) on Aug. 2**. Minutes of the Trout Lake meeting have been distributed by the Network Office.

The **agenda topics at Trout lake** included the 10-year LTER review, international activities (CERN interactions; IGBP Focus 2/GCTE), Climate Committee report, network remote sensing acquisition status, LTER/NASA collaboration, a synthesis session and data management issues.

James Brunt (SEV) presented the **report on the 1991 Data Management meeting** (San Antonio). His presentation, and the report handed out at Trout Lake, emphasized to the CC committee members the critical role the data managers play at their sites and for the network as a whole.

Jerry Franklin reported that he had planned to include the paper on the **Minimum Standard Installation (MSI)** in the technical report prepared by Foster/Boose. It became clear, however, that the MSI specifications dating to 1988 would first need updating with input from PI's and data managers. Prior to the CC meeting, Rudolf Nottrott sent a proposed draft for an MSI update to the PI's and data manager groups on LTERnet, but had received few responses. Responses suggested that an updated MSI would require at least one full-time data/systems manager and that it is premature to adopt Structured Query Language as a data access standard. The LTER Executive Committee recommended that any updated MSI should be published in a refereed journal, such as BioScience. Rudolf Nottrott requested additional input on the MSI update from the PIs.

The meeting **agenda at Fairbanks** was dense. Detailed minutes will be distributed by the Network Office. The list of agenda items included: Mary Clutter's (NSF) talk on NSF's perspective on funding, opportunities and long-range planning; the 10-year evaluation of the LTER program; 10-year plan for LTER's future; the collaboration effort between LTER and NASA and the agenda for the 1993 All Scientists Meeting.The agenda items directly related to data management were the Catalog of Core Datasets, the Minimum

Standard Installation (MSI) update, the intersite bibliographic database and the University of Alaska's GIS/Remote Sensing Lab.

In the session on data management issues, Rudolf Nottrott advised that now, with improved access to the LTER Core Dataset Catalog (query functions via electronic mail and directly over the Internet), the next logical step in developing the catalog is to expand the number of listed datasets. This will make the catalog much more useful to the LTER Network and the wider scientific community. With over 2,000 datasets network-wide, hard copy documentation would no longer be practical and always out of date. The current catalog only lists a small percentage of the total number of datasets. LTER needs to go beyond listing core datasets only and include all long-term datasets (not actual data). Polled by Jerry Franklin, all site representatives provided feedback on the proposed expansion of the Catalog, considering issues such as the definition of "long-term," cost, and dataset quality.

Most sites supported the proposal, at least conceptually. Most reservations expressed by the site representatives relate to (1) the magnitude and scope of the task (problems with the accessibility of old datasets in different forms; time and cost of fielding data requests resulting from catalog references; where to draw the lines on which datasets should be included - just LTER datasets or the whole site resource; difficulties with large sites and numerous datasets; effort can't come out of the basic site science budget; potential drain on other resources; need resources to develop critical standards/structures), (2) proprietary or legal issues (individual investigators may resist; catalog should include disclaimer), (3) the involvement of other agencies (in case of non-LTER datasets, concern that agency partners get due recognition; obligations to other agencies must be considered), (4) whether the would be a "consumer audience" for the catalog (people are often more interested in contacting scientists than in gaining access to a listing which is incomplete) and (5) whether intersite research is really limited because not all datasets are listed in the catalog. Datasets should not be a higher priority than publications (process studies which result in publications should be emphasized) and research should be question-driven.

Several people suggested that a possible implementation of the expanded catalog should include links to the a bibliographic database.

Jerry Franklin **concluded the discussion by asking the Data Managers for a staged proposal for the catalog expansion for the Coordinating Committee's consideration**. He noted that part of LTER's legacy, is creating "footprints in time." We have a tremendous responsibility in archiving for future generations. How will we be providing access to the wider scientific community? The future of LTER as a program and as a network depends on our developing whatever initiatives necessary to enable us to act responsibly in this area.

Jerry Franklin then described the importance of updating and publishing the **Minimum Standard Installation (MSI)** document on which he had asked Rudolf Nottrott to work with the sites prior to the Trout Lake meeting. Since the original MSI document was completed, the Network has revised the old concepts and developed new ones without documenting them. Over the next six months, the MSI will be made available to other sites, and

other research programs around the world. JF asked for site input, noting that after the Data Managers meeting, RN will circulate the revised version for further comment.

Concerns from initial circulation of the document prior to the Trout Lake meeting include: (1) cost of maintenance and personnel, (2) inclusion of global positioning systems, (3) the steep developmental curve with systematic Network-wide upgrades, and (4) the problems associated with "lockstep" upgrading. RN noted that the MSI will never be definitive, and that the revised document should state that costs, concepts and ideas are constantly changing.

Meeting participants made several comments: It would be desirable to change title of the MSI to Minimum Standard the LTER's instrumentation. An updated and broadened MSI would be useful if it is written up clearly and concisely, emphasizing that the MSI is one tool in the service of science, with publication and synthesis as the goals.

Caroline Bledsoe, reported on the development of an LTER **all-site bibliographic database**, a proposed project she is exploring under subcontract to the collaboration grant. The project goal is to combine all bibliographies from all sites into a single, on-line searchable database with annual updates. Such a resource would have many science uses.CB reported that she has been working with Niwot Ridge and Martha Andrews, a librarian at the University of Colorado, Boulder who will attend the Data Managers meeting to discuss software, commercial services, etc., and help to make recommendations, estimate costs, and plan a funding proposal.

A **demonstration University of Alaska's GIS/Remote Sensing Laboratory** utilized the Internet for remote data access (SEV Landsat-TM scene from the 1991remote sensing acquisition), processing and display. The demonstration was based on recent interactions between research groups at theLTER Network Office, SEV and The Khoros Group, and the San Diego Super Computer Center (SDSC).

Ken Steube (SDSC) demonstrated vegetation analyses from TM data stored at SDSC. The data were processed locally and on an SDSC Cray, and displayed on multiple displays in the lab. Many participants, including Mary Clutter and Jerry Franklin, noted the future possibilities for LTER in linking with such groups as the SDSC and the National Center for Geographic Information Analysis (NCGIA) for specific processing and data storage capabilities.

2.5 Views from the National Science Foundation

James T. Callahan of the National Science Foundation addressed the workshop. He stressed the importance of vision and imagination in addressing the problems of scientific information management. He also stressed the need to aim at objectives, with less focus on means. He pointed out that with the 3 to 5 year rotation in technology, technical means tend to catch up with goals.

Dr. Callahan discussed five major themes. The first was the emergence of scientific information management as a scientific discipline in its own right. He discussed the responsi-

bilities of a new discipline -- to develop curricula, a service ethic to the scientific community, but most of all he emphasized the critical role of research and publication. Secondly, he discussed how position augmentation of data managers could be achieved by legitimization of scientific information management as a valid discipline. Here he noted the important role of education, both of information managers and other scientists in this process and the critical role of performance. Scientific information management must deliver what is needed to other scientists. Thirdly, he discussed LTER as a centerpiece of long-term projects and environmental biology and emphasized the need to locate and utilize additional sources of funding. Fourthly, he discussed the 10-year review of the entire LTER program, with the organizing theme: "What has LTER accomplished that would not have been accomplished otherwise through individual funding?" Finally, he talked briefly about the proposed National Center for Ecological Synthesis and Analysis.

2.6 Chinese Ecological Research Network (James Brunt, SEV)

Following an LTER Delegation visit to China last year to interact with the leaders of Chinese Ecological Research Network (CERN), two developments have further advanced the collaboration with CERN. First, 4 data managers will be in Beijing in September assisting CERN scientists with research data management issues, and second, a grant has been received by James Gosz and James Brunt (SEV) from NSF to organize a research data management training program for CERN. These separate, but related efforts are part of the planned continuing collaboration between LTER and CERN.

2.6.1 LTER DM in China

In cooperation with the National Academy of Sciences' Committee for Scholarly Communication with the Peoples Republic of China (NAS/CSCPRC), the World Bank and the Chinese Academy of Sciences, LTER Data Managers will be in Beijing during the month of September to make contributions to the development of design plans for a data management system and data management training for CERN. James Brunt (SEV), Tom Kirchner (CPR), Bill Michener (NIN), and Susan Stafford (AND) will be working with CERN personnel at the invitation of Dr. Zhao Jianping and Dr. Sun Honglie, Chinese Academy of Sciences (CAS). The effort, lead by Bill Michener (NIN), represents a continuation of collaboration begun at the 1990 LTER All Scientists Meeting which led to a CERN delegation visit to LTER sites in May of 1991 and LTER delegation visit to research sites in China in September of 1991. During those visits, data management was identified as one of the key areas where CERN and LTER could begin immediate collaboration.

One of the highlights of this trip will the presentation of a two-day data management workshop in which LTER Data Managers will get to meet individuals responsible for CERN data management from the field stations, sub-centers (disciplinary centers responsible for key efforts), and the synthesis center (a national center that will provide resources for the synthesis of data and training of scientists). Topics to be discussed include data management system design, data management training, and the modeling-data management interface. Some of these scientists will be candidates for the 1993 data management training program at the University of New Mexico (SEV).

2.6.2 CERN Data Managers in Albuquerque

The National Science Foundation has awarded funds to LTER investigators James Gosz and James Brunt (SEV) for "A Proof-of-Concept Prototype" training course in research data management to be held at the University of New Mexico in 1993. The proposal was developed over the course of the CERN collaboration through the combined efforts of the investigators, Barbara Benson (NTL), Bill Michener (NIN), Susan Stafford (AND) and Beryl Leach, CSCPRC Program Manager at NAS. The funding of this proposal was a combined effort between NSF Divisions of International Programs and Environmental Biology. Besides those mentioned above, many other LTER data managers will be contributing their expertise to the development of this training course and furthering relationships between CERN and LTER.

The training will address basic concepts of managing long-term ecological research information across both the LTER and CERN networks. The primary audience are CERN data managers and data administrators responsible for design and development of a system of Research Data Management (RDM) within CERN. Data management personnel from the LTER network will team-teach modules focused on the needs of CERN, illustrate the diversity of techniques, and work out possible solutions. The training course will cover a wide range of data management issues from broad principles of design and organization to specific implementation issues. The format will include lectures, demonstrations, discussions and hands-on computer lab experience. The curriculum materials resulting from this course will be used to develop other courses.

This effort will establish groundwork for protocols and standards that will facilitate exchange of ecological research data internationally, and will set the stage for future cooperative efforts between CERN and LTER. Group discussions in the proposed training course will focus on the needs for international collaboration and sharing of long-term datasets, and address the establishment of clear and attainable international data management objectives. The workshop will culminate in facilitation, scheduling, and curriculum development for a follow-up intensive workshop in the Peoples Republic of China (PRC), in which LTER and trained CERN data managers will work side-by-side in training other CERN data management personnel. The trained participants will make training larger groups back in China possible and productive. Collaborative assistance of this kind will help cement the bonds between CERN, LTER, and the international ecological community, and thus, open the door to future cooperation in global research.

2.7 Inter-site Climate Database Project (John Gorentz, KBS, John Briggs, KNZ)

Work is in progress on the inter-site climate database project, under whose auspices KBS, BNZ, NTL, KNZ and NET are developing a prototype "gatherer-server" for electronic

mail access to climate data.

This database is intended to help researchers having access to Internet e-mail obtain climate data from multiple sites. These data are already comparable and (in many cases) accessible, but obstacles to their use exist due to differences in their form and organization, and in the methods of identifying them in time and space. The use of standardized database software does not necessarily do much to overcome these obstacles and in any case is often impractical because of historical or institutional reasons.

A prototype system is being developed by which climate databases at LTER sites can be accessed as if they were a single database. Users will be able to request data by sending an electronic mail message to an LTERnet address, which will then forward it to a central "gatherer-server" (located at KBS). The gatherer-server will parse the request into separate sub-requests, which will in turn be e-mailed as SQL queries to the relevant remote databases. SQL-servers at these sites will receive the requests, process them, and e-mail the requested data back to the gatherer-server. The gatherer-server will then aggregate and merge the data into a canonical format (reconciling the differences in form and organization, and in time-space identification) and e-mail them back to the requester.

At this time, Ingres SQL-servers at KBS, NTL BNZ and NET are functional. Although the gatherer-server uses Ingres to store requested data while manipulating them, it has been decided to use SAS to do most of the gatherer-server manipulations. Software has been written that takes users' requests for data and generates SAS procedures, using SAS's SQL where possible, to do the necessary data aggregation and merging to prepare the output to be sent back to the requester. SAS's rich repertoire of date-time functions, and the SAS/ACCESS product, which allows it read and write to an Ingres database, were factors in the choice of this method.

There will be further announcements when the system is available for testing.

2.8 Electronic Data Collection Report (Rick Ingersoll, NWT)

The data manager at each of the sites should have received a bound copy of the report, in addition to appendices (unbound copies of the individual responses to the survey). Data managers should insure that relevant technical personnel (i.e. those involved in the purchase, operation, and maintenance of the electronic equipment) and investigators are aware of, and have access to, these documents. This could be an initial step in reduction of the "information gap" among investigators, data managers, and technical personnel involved with electronically collected data. The report suggested that better communication among these individuals, at both the intra- and inter-site level, would be beneficial to LTER in terms of improved data quality, more efficient problem solution, and greater input to commercial manufacturers.

Much of the information contained in the report has become or soon will be outdated given the rapid evolution of the technology. Thus, evaluation of the management of electronically collected data within LTER should be considered an ongoing process. The report can be brought up-to- date periodically (annually?) fairly easily with the format consisting of edited tables with a brief addendum comprising the text.

Rick Ingersoll will communicate with the data managers by e-mail and ask that they mail him photocopied pages of text or tables with suggested revisions, which then will be incorporated into the updated report. The updated report will be publicly accessible on LTERnet.

2.9 Interactive Data Access (Barbara Benson, NTL)

The Interactive Data Access (IDA) working group at the 1991 LTER Data Management Meeting had decided to survey the LTER sites to collect information on the status of data accessibility by researchers at LTER sites. Barbara Benson summarized the results of this survey. A more complete summary will appear in the fall issue of Databits.

Currently at almost all sites, data management staff perform the retrievals for at least some of the LTER data sets (as opposed to direct access by the researchers). Eight sites have user interfaces to at least some of their data sets. Twelve sites have an IDA system either planned or in process. All sites reported some combination of time, personnel or money as a barrier to implementation. The need for network connections to remote sites was also mentioned by some sites as a problem. Other commonly cited barriers were proprietary issues and multiple computer platforms.

2.10 Databits (John Porter, VCR)

Databits, a newsletter serving the data management community, continues its publication. Edited by John Porter and distributed by the Network Office, it includes contributions from data managers at all LTER sites. It is available both in paper form and electronically from LTERNET. Numerous discussion participants suggested that an index be created of all the **Databits** issues so that relevant articles could more quickly be located and that electronic data should be in several standard forms (ASCII, Postscript) to facilitate downloading. Creation of on-line searching capabilities for **Databits** is being actively pursued.

2.11 Andrews Forest Database Activities Grant Proposal to NSF (Don Henshaw, AND)

Susan Stafford, Gody Spycher and Don Henshaw have written an NSF proposal requesting funds to create a conceptual design for a User Interface for the Forest Science Data Bank (FSDB) at Oregon State University. The Interface Design will emphasize generic support for basic data activities and access in a heterogeneous network. Production activities include assembly of descriptive data, key entry and data imports, program-supported data editing, and archiving. Data access will support searches of on-line database catalogs, queries to one or more autonomous databases, and export of data into analysis systems. Remote access from various platforms will be supported and access to data will be controlled through a permission framework established by the data owners. The User Inter-

face will lead users systematically through common data management activities without requiring them to master the underlying database management system. The proposed system will allow the researcher greater control and easier access to the FSDB.

2.12 All Site Bibliographic Database (Martha Andrews, NWT, Caroline Bledsoe, NET)

The proposal "Supplement to NSF-BSR 9011658: "Long-Term Ecological Research in the Colorado Alpine" contained support for an information scientist (Martha Andrews) to assist Caroline Bledsoe in development of an on-line, interactive LTER network bibliography. Martha Andrews was charged with providing input in software and hardware options, based on her experience with several combined-source databases within the polar regions. Prior to the LTER Data Managers meeting in Honolulu in August, 1992, Martha obtained answers to a questionnaire filled in by each LTER DM for his/her site. Some of the results of the questionnaire were compiled and handed out at the meeting. Other discussion points which had arisen from questionnaire answers were included in Martha's presentation, and there was active discussion relating to the issue of an LTER All Site Bibliographic Database. Some questions concerning the scope of the bibliography, how it would be kept current, whether a controlled vocabulary would be used, etc., were not clearly resolved at the meeting, and will be taken up by Caroline Bledsoe in the coming year. Taking the position that such a bibliography would be produced, options for production of such a database were discussed. All existing, machine-readable files would be combined, either with or without changes and/or additions.

This bibliography would contain over 11,000 references. It would have a wide range of uses for individual sites, the LTER network as a whole, and the wider ecological research community: LTER scientists could more easily plan multi-site experiments and synthesize information if site research information were more accessible. Scientists from the wider ecological research community can compare their data or sites to similar LTER data or sites, developing more robust, general ecological principles. A bibliographic database can be cross-linked to the LTER Core Data Set Catalog, providing researchers with access not only to citations, but also to specific data sets. Agency administrators, decision makers and managers will have a very large body of ecological information identified and indexed for their use. Other uses mentioned relate to the value of a bibliographic database for historic purposes, to demonstrate site productivity; to share information and improve the local bibliographies and for general public relations value.

Linking the bibliographic database with other existing databases (LTER Core Dataset Catalog and Personnel Directory) would be desirable. Considerations as to structure (of records, of the database), indexing (controlled vocabulary with appropriate but simple terminology), and the exact methods of access and updating were raised often in the discussion.

Additional discussions with Caroline Bledsoe and Tom Callahan resulted in Caroline Bledsoe taking the lead after the Honolulu meeting in continuing to develop the project of the all-site bibliographic database. Caroline Bledsoe is currently working with experts in computer and information science at UC Davis (Advanced Networked and Scientific

Applications (ANSA) Group in the newly-formed Information Technology Center at Davis) on a pilot project using data from several LTER sites. The ANSA Group includes both librarians and computer and networking specialists; the group designs tools to allow educators to access and use scientific information with computers and electronic networks. The ANSA group is interested in using the LTER bibliography project to test new software/hardware tools. Martha Andrews remains interested in the project and is kept advised of its progress by Caroline Bledsoe.

2.13 LTERnet (Rudolf Nottrott, NET)

Connectivity within the LTER Network has been improved by the connection of the LTERnet Network Support System to Sprintnet. The LTERnet->Sprintnet, part of the recommendations of the Connectivity Report, provides a new way to connect to the Internet from remote locations anywhere in the U.S and abroad. The connection charges are inexpensive compared to long-distance voice phone connections (from $6/hour off-hours in major metropolitan areas up to $14/hour for 1-800 connections from any location). Currently the Network Office is picking up the charges. A demonstration of access to LTERnet through Sprintnet is described in “Sprintnet->LTERnet Access” on page 39.

The LTER Core Dataset Catalog can now be queried by e-mail and directly over the Internet. E-mail query extends access to the catalog to users without Internet connections (such as Bitnet, U.S. Forest Service, OMNET, and others). Instructions on the use of the on-line queryable Catalog are also available by e-mail (any message to Catalog@LTERnet...).

A separate server function at LTERnet allows access to the LTER Personnel Database via e-mail and directly over the Internet (instructions for its use can be obtained by sending any message to Address@LTERnet...). Because of the similarity of the Catalog and Address server function, linking the two functions seems a logical and desirable next step. When searching for an investigator’s name (or part thereof), the combined server function could, for example, return all catalog entries of the investigator, plus the investigator’s entry from the personnel database (providing address, phone, e-mail, specialty, etc.).

3.0 Working Groups

3.1 Plan for Development of Network Resources

Committee: Rudolf Nottrott (NET), Eda Melendez (LUQ), Bill Michener (NIN), Martha Andrews (NWT), Rick Ingersoll (NWT), Mark Klingensmith (BNZ)

Emphasizing the crucial role of the LTER network in sharing research tools and expertise, the data managers decided to develop a plan that outlines specific resources targeted for further development, or to be newly created and maintained in the network. The plan is not meant to be exclusive, but should help to us focus on those resources that we have come to recognize of central importance. As we proceed in the developing those resources, new ones are likely to be added (or existing may be dropped), keeping in mind

that these common resources need to be useful in **supporting ecological research** and must be freely and easily available to researchers at all sites. In continuation of our efforts to reach out to the wider ecological research community, these resources should also be available beyond the LTER Network (e.g., OBF/SAML, ESA LTSS)

Specifically, the list of network resources targeted for development has 3 general categories: (1) on-line documents, (2) a data management toolkit and (3) on-line databases.

3.1.1 ON-LINE DOCUMENTS

All sites now have Internet connections and thus electronic document distribution between all sites is a realistic goal. Sets of documents suitable for electronic distribution, but presently filed and distributed mostly as hard copies, include the data management history, publications by the Network office, as well as many publications of network-wide relevance distributed by a variety of institutions (e.g. the US Forest Service report on "Climate Variability and Ecosystem Response"). A summary of site capabilities relevant to data management (expertise, hardware, software, etc.), proposed by Karen Baker (PAL) and Eda Melendez (LUQ), would also fit this category.

In addition to making these files widely available over the Internet, having them on-line will open the possibility of searching them for keywords or text strings.

To be able to fully benefit from electronic document storage and distribution, we need to agree on standards for on-line formats. Encapsulated PostScript (EPS), and text (ASCII) are obvious candidates, but emerging newer formats provide mechanisms for better user interfaces (hypertext, graphics visualization, WYSIWYG editing). This will remain an area of discussion, but to be able to share documentation on-line within the LTER network it is important for us to settle on a few formats. Plain text (ASCII), our least common denominator for sharing of documents in the past, may no longer be adequate. FrameMaker was proposed as an alternative format already in use by a number of sites.

3.1.2 DATA MANAGEMENT TOOLKIT

Many tools that have been developed at a particular site or elsewhere are suitable for solving problems at other sites. The data management and access routines developed at the CPR and CDR sites, as well as tools developed for the "Science Work bench" (/Intersite data exchange) are examples.

It is imperative that all tools that are part of the tool kit are thoroughly documented and indexed. A work group should be established to coordinate the compilation of tools and to assure that documentation standards are followed. Bill Michener has volunteered to chair the group.

The toolkit should be made available for public access (read only) over LTERnet's various access facilities (Gopher, WAIS, FTP).

3.1.3 ON-LINE DATABASES

On-line databases need to be further developed or newly created in the following categories: (1) Catalog of Core Datasets, (2) intersite bibliographic database, (3) LTER Personnel Directory, and (4) intersite datasets not specifically maintained or archived by any site.

Core Dataset Catalog;

The LTER Core Dataset Catalog should be expanded to ALL long-term data sets at the sites, not just LTER collected core datasets. The expanded Catalog would be maintained as an on-line version only. In addition to the present catalog search functions (investigator, core area, site, and text string in abstract or title), catalog entries should be searchable for keywords (controlled vocabulary). The structure of the present catalog entries may be modified to include the data dictionary for the dataset.

Mechanisms need to be explored to add long-term sets that are not maintained by LTER personnel at a site. This may be straightforward for datasets at LTER sites, but may require an editorial committee for datasets at the site. (An editorial board would be required to evaluate submissions from the larger ecological community if we open it up to others).

Rudolf Nottrott will prepare a proposal to expand the dataset catalog to all long-term datasets at a subset of LTER sites (5 to 10 sites). Based on the experience with this subset, a later proposal should be prepared to cover all long-term datasets at all LTER sites.

Inter-site Bibliographic Database

An intersite bibliographic database with the capability to access records from all sites should be developed. The bibliographic database should be linked to dataset catalog and the personnel directory so that cross-referencing is possible. For example, it should be possible to search simultaneously the data catalog, the bibliography and the personnel directory for a particular author or text string (key word).

The bibliographic database should be accessible over the Internet (TCP/IP), which excludes many of the common PC-based software packages. Martha Andrews and Caroline Bledsoe will be examining types of software that are available in network versions and the appropriate scope of records to be included (possibly start only with those that have ISBN numbers)

Personnel Directory (On-line and Hardcopy)

The effort to extend the on-line personnel directory to include researchers from the wider ecological research community beyond the LTER network should be continued as capacity permits.

A mechanism for sites to do their own directory updates will be in place soon. Initially this will involve a few sites that will update their site personnel’s entries using menu-driven access to the directory on LTERnet. The goal is that most sites will do their own updates

by the end of 1993. Ways to maintain personnel records at the sites and accessing them through a central access point, such as LTERnet, will be explored in the coming year.

The on-line personnel directory will always reflect any updates made by the sites immediately and is therefore more accurate than the hard copy, which is presently distributed annually to all sites. Despite this, workshop participants felt that the hardcopy has been a useful reference document and should still be distributed for the time being.

Future development of the on-line directory should consider the possibility of linkage with the core dataset catalog and the intersite bibliography.

Intersite Datasets

Long term datasets of interest to multiple sites, but not specifically archived or distributed by a site, should be made available on-line over the Internet. An example for such a dataset is 100-year cyclone frequency for the continental US compiled by Bruce Hayden (VCR). These datasets should have entries in the LTER Catalog of Core Datasets. Ways to link the data to the catalog such that browsing the catalog may include the possibility of retrieving actual data will be explored by Rudolf Nottrott (NET).

3.2 10-Year Plan for Development of Scientific Information Management in LTER

Committee: Barbara Benson (NTL), Barbara Nolen (JRN), John Briggs (KNZ), Emery Boose (HFR), James Brunt (SEV), Susan Stafford (AND)

At the1992 LTER Data Managers Meeting, in Hawaii, we challenged ourselves by brainstorming about the future and where we would collectively like to be in the next ten years. Recognizing that we are working with a technology that is changing every 6 months, we did not concern ourselves with technological specifics but rather concentrated on the “big picture” and what “big picture” capabilities we would like to have. Our thoughts were divided among four general categories: (1) training and education, (2) communication and resourcing, (3) technology, and (4) developing a new mindset.

3.2.1 TRAINING AND EDUCATION

Research Information Management (RIM) will be recognized as a scientific discipline with recognized curricula at the BS, MS and PH.D level combining biology, ecology with computer science. One implementation step will be to have developed Research Experiences for Undergraduates (REU) at every LTER site incorporating RIM training and education.

The data managers as we know them today will be obsolete. We will have moved to a role of facilitation for synthetic research efforts. A criterion to measure this by will be the frequency with which the Data Managers are Co-PIs on the LTER grant demonstrating a science partnership role.

There will be educational opportunities for “in-service” network training and cross-training among data managers. This will be measured by the number of collaborative papers, workshops, and research proposals.

3.2.2 COMMUNICATION AND RESOURCING

We will be using each other as resources and the competitive, high-stakes environment will be minimized.

Electronic mail (e-mail) will be fully exploited.

There will be standing committees across the network which will be will be in frequent communication utilizing state-of-the-art video conferencing technology. Leaders from each standing committee will be in regular contact with each other to disseminate the information from each standing committee. There will be travel resources to facilitate data manager travel among sites for cross-training and resourcing.

3.2.3 TECHNOLOGY

The creation of datasets for intersite workshops and analysis which is now in most cases complex and time-consuming, will be a “piece-of-cake”. We will have articulated our software needs to validate this goal. We will have fostered stronger ties with the software/hardware industry and have worked with the commercial software companies to recognize particular need of scientific databases in a way similar to NSF’s experience with the SAS Institute in North Carolina.

There will be increased speed over the Internet and full connectivity of all LTER sites and biological field stations.

There will be platform and software independence.

E-mail will be fully exploited and merged with teleconferencing.

3.2.4 VIEW AND MINDSET

We will have learned how to function as a network. We will have fostered a better sense of community within the LTER network. There will be well established mechanisms and widespread occurrences of cross-site synthesis work.

We will be a “Network of Networks” in which LTER will be a part of a very large ecological network. The LTER Network Office and its staff will be playing a continuing larger role in this development as the number of players continues to grow.

3.3 Data Exchange and Access

Committee: John Porter (VCR), Tom Kirchner (CPR), Richard Lent (HFR), Karen Baker

(PAL), Jim Laundre (ARC), Cindy Veen (HBR), Don Henshaw (AND), Abderrahman El Haddi (CDR), Gil Calabria (CWT)

Facilitation of scientific information exchange is a primary objective of LTER information management programs. The ultimate value of LTER data sets are determined the by the degree to which data are used, not by the degree to which they are archived. Workshop participants identified three basic areas where steps could be taken to facilitate exchanges of data. The first of these was the technical area which includes use of interactive data access (IDA) systems, wide-area networks and compatible computer hardware and data exchange formats. The second area is standardization. Optimum data exchange standards, creation of parallel datasets at different sites, development of data dictionary standards and development of global descriptors (such as standard formats for presenting dates) fall into this area. The final area affecting data access facilitation is administrative. Individual LTER sites need to implement policies that permit data exchanges.

3.3.1 PREVIOUS ACTIVITIES

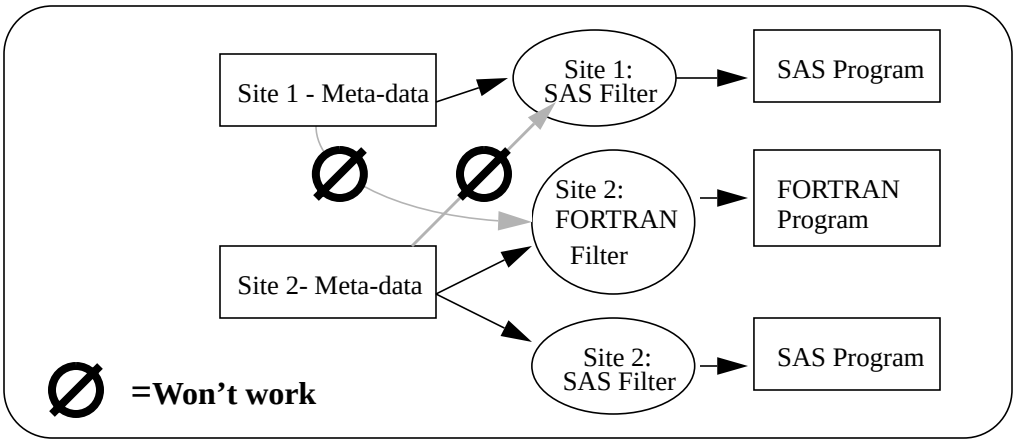
Past workshops and proposals and reports created by workshop participants have included detailed discussions of many of these topics. The 1990 workshop and subsequent sessions at the LTER All Scientists Meeting led to generation of network-wide guidelines for data management policies. The 1991 workshop had a working group which focused on the growing use of IDA systems within the LTER network. Some existing systems are well advanced, such as the intersite climate database and “gopher” servers at the network office and Hubbard Brook and Virginia Coast LTERs. The “Minimum Standard Implementation” has done much to standardize the computer equipment used at LTER sites and the “connectivity report” and subsequent technical supplements have done much to improve electronic network connectivity between LTER sites via NSFNET.

3.3.2 STANDARD EXCHANGE FORMAT

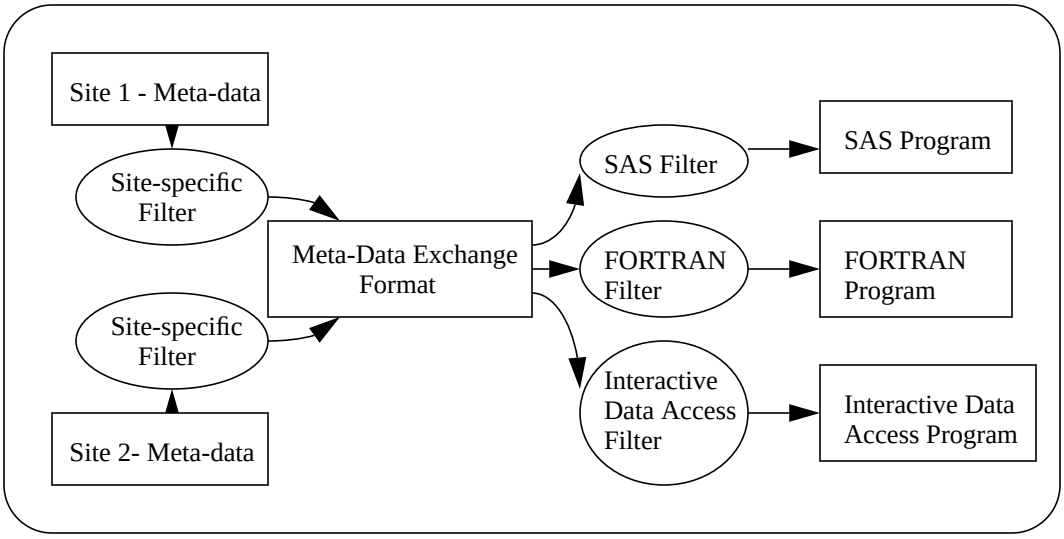
At this workshop, the data access facilitation working group focused on the development of a standard format for the exchange of meta-data (meta-data consists of the documentation and other materials which are required to utilize data, but are not themselves part of a data set). Current practice for data exchange is to provide meta-data in site-specific formats, many of which are not machine readable and no two of which are the same. Investigators receiving data sets from different sites must manually produce programs to read data into whatever analytical software they wish to use. A standard, machine-readable format for meta-data would make it possible to automate the process so that basic analytical programs could be automatically produced. Advantages of using a standard format include minimization of errors in exchanging data, increasing the ease of exchanging data and more efficient software development (through a reduction in redundancy). The disadvantages of adapting a standard are that individual sites would need to develop a translator to convert their locally maintained meta-data into the standard format and that care must be exercised in the development of such a format to assure that it is sufficiently general to accommodate individual site differences.

FIGURE 1. Comparison of processing using unique and common data exchange formats

No Common Exchange Format. Each site must develop its own tools.



Common Data Exchange Format: Sites can share program filters.



Desirable characteristics of a meta-data exchange format include both human and machine readability, minimization of redundancy and flexibility in its structure to accommodate a wide range of potential data formats. Importantly, the proposed format is for the exchange of data, not for the development and maintenance of meta-data within a site. Individual LTER information management programs need the flexibility to exploit unique local resources in the management meta-data. Although individual sites may choose to exploit the format on a strictly local basis, its primary focus is on intersite data exchange.

A major advantage of a common format is that it will facilitate the development and exchange of software which use the information in the meta-data to produce programs that manipulate the data themselves. Use of a common meta-data exchange format is complementary with use of interactive data access systems as well. Development of the common format is not cost-free. Participating sites will need to develop a translator from their internal format to the standard format. Additionally, care must be taken to assure that any format that is adopted is sufficiently flexible to accommodate different types of data and meta-data.

3.3.3 EXCHANGE FORMAT DEVELOPMENT

Development of the common data exchange format will be a lengthy process. There are four steps that must be taken:

1. We need to develop a consensus that such standardization will help and not hinder individual sites data management efforts. To a major degree, this was accomplished at the 1992 workshop. There was a high degree of consensus that, if developed, the data exchange format would be both useful and used.
2. Individual sites will need to be queried to develop a lists of “basic” meta-data items that must be included, plus optional items that can be included, but are not required. A starting point for this will be the “minimum documentation standards” compiled at the 1990 workshop.
3. The basic and optional meta-data items need to be incorporated into a flexible machine readable format which is capable of handling a wide variety of data types, from simple numbers to long blocks of text. Tom Kirchner (CPR) and Abderrahman El Haddi (CDR) have been working on machine readable formats within their own sites (See “Interactive Data Access Program Developed at CPR” on page 31. and “CDRLIB: A library of Functions for Scientific Data Manipulations” on page 36). The Kirchner format seemed to be more flexible in some aspects, so Tom Kirchner was asked to take a first stab at designing the machine-readable implementation of the common exchange format, in consultation with Abderrahman El Haddi.
4. Individual sites will need to develop software tools to facilitate translation of their local meta-data format into the data exchange format. Tools for translating from the data exchange format to formats used for specific analysis tools (statistical packages, spreadsheets, interactive data access programs) must also be developed.

The resulting format can be expanded as needed to accommodate additional types of data using a similar process.

3.4 Proposed National Center for Ecological Synthesis

Committee: All participants. Discussion leader: William Michener

Proposed National Center for Ecological Synthesis Group discussion focussed on providing input on the National Center in terms of scientific objectives and research needs.

Numerous valuable ideas and insights were contributed. A joint committee comprised of members of the Ecological Society of America and the Association of Ecosystem Research Centers met at an NSF-funded workshop in Albuquerque, NM on 25-27 October 1992. Ideas suggested at the LTER Data Management Workshop were brought to the attention of workshop attendees by William Michener (North Inlet). The workshop report, edited by James Brown (Sevilleta) and Steve Carpenter (North Temperate Lakes), will be completed by February 1993. Availability of the report will be publicized through bulletins of AERC, ESA, and the LTER Network.

3.5 All Scientists Meeting

Committee: All Participants. Discussion leaders: Susan Stafford (AND), John Briggs (KNZ)

The All Scientists Meeting Plenary Sessions, September 1993 provide the LTER Data Managers with an opportunity to showcase our accomplishments and set the stage for the coming decade. We decided we need a dynamic speaker who can emphasize the “network value-added functionality” of our research information activities. The presentation needs to have a network orientation emphasizing our evolution from a collection of individual sites to a network of sites. A vision of where we are going via a video or slide show would be appropriate and provide a time-line for the history of our collective data management efforts across the network. We should personalize this without offending anyone. A cartoon summarizing what data managers do would be clever.

The research link is crucial to this presentation. The plenary presentation should emphasize how Data Management has facilitated research with examples from synthetic and long-term studies. Accomplishments include CERN, Symposia, MSI. LTER DM has been at the forefront in the ecological community. We should outline our plans for future activities and our 10-year plan. RESEARCH should be a key word to help remove the mindset of data management as just a SERVICE component. We must give examples of how data management at the network level facilitates intersite research, i.e. the data management in DIRT and LITTERBAG studies. Also how data management at the network level affects general communications and connectivity - i.e. mail forwarding and the distributed climate database.

Another theme is information sharing that is facilitated by data management at a network level. For example, standardization of metadata, catalog and personnel directory, network resources (toolkits - LTERmenu and CDRlib), as well as technology transfer via DataBits. We have also been successful in attracting outside expertise and resources and have a critical mass of people in Data management. There have been “spin-offs” of LTER products and expertise to the outside scientific community - Parknet, China.

“Synthesis” and “products” were the important buzz words here. The consensus was that we should have a series of bulleted items highlighted in the Plenary Session as well as an ambitious schedule of demonstrations throughout the week of meetings. Ideas for each are as follows:

3.5.1 HIGHLIGHTS FOR THE PLENARY SESSION

1. The WAN meeting that produced as its first product LTERNET and the mailing alias system providing easy connectivity among all LTER PIs.

2. The development and installation of an MSI from a series of science and technology committees: the Shugart committee, the Gosz' Committee, and the Connectivity Report.

3. The Baruch Volume in '85 from the first symposium on Research Data Management

4. 1993 Symposium and the creation of white papers on various topics relating to research data management across all scales.

5. KBS Gorentz Blue Book

6. Establishment of the Network Office and the hiring of Rudolf Nottrott. Network DM activities, accomplishments, and resources.

7. Core Data Set Catalog and On-line Catalog

8. DataBits - John Porter

9. Remote-Sensing Activities - John Vande Castle

10. GPS - Bill Michener

11. The whole GIS Scene with hardware, software, and training

12. Early work of Walt Conley, Tom Carl, and Carl Bowser and others in the prototype data analysis workshop. Include other examples of collaborative data workshops that have evolved from the first Las Cruces workshop.

13. Provide a series of Benchmarks tracking our progress in hardware, software, personnel, and networking. This could include a list of the workshops we have held over the last few years.

14. China Connection

15. Catalog of Core Data Sets and Personnel Directory on-line

16. One suggestion was to show productivity of DM over time in terms of pubs in a graph. This could be as LTER data managers and as individuals.

17. Data-sharing and requests

18. Data management model - DM as a PI

19. DM on NSF review committees

20. Facilitation of research - synthetic, intersite, and long-term

21. Resurrect the Data manager Task List from the Toronto meeting - show how we have moved from one Data Manager to a DM staff. Identify the RIM model with DM on the same level as investigators.

3.5.2 DEMONSTRATIONS

1. Distributed Climate Database. John Briggs and John Gorentz's and Tom Kirchner's approach at the first attempt at an intersite synthesis.

2. Interactive Data Access

3. GIS demos

4. Network Office Facilities

5. Demo of white papers from symposium in May - perhaps using Frame Maker

6. Set up LAN on site

7. Khoros demo

8. GPS

9. Access to LTERnet on-line data

10. SprintNet access to LTERnet and INTERnet

11. LTERmenu - Tom Kirchner

12. Data Filters - El Haddi

13. Introduction to UNIX

14. E-mail software (e.g. pop-mail)

15. Analytical tools

4.0 Appendices

4.1 Site Flashes

4.1.1 AND

Data managers Susan Stafford, Gody Spycher, and Don Henshaw have submitted a proposal to the NSF Database Activities Grant Program. We are requesting funds to create a conceptual design for a User Interface for the Forest Science Data Bank at Oregon State University (OSU) (a more detailed description follows later in this report).

A program for viewing GIS coverages and metadata using a menu driven interface has been written by Hazel Hammond. The program allows viewing or printing of coverages or measurement sites by geographic area, data theme, or data type. The program is an ArcInfo aml which uses ArcPlot.

Data requests continue at a steady pace. Our Request Handling Program (demonstrated and described at last year's meeting) has been a tremendous asset for meeting these requests. This Foxpro program generates all necessary correspondence and metadata in the handling of requests for data, and provides a system of tracking data requests.

The Andrews is installing 26 monuments in preparation for the GPS campaign slated for September-October. We will have an instructor from the North Inlet site for the first week.

Ground breaking for a new office/conference center at the Andrews Field Site should occur this fall. We hope to have a T1 connection to the site in about one year.

Our Quantitative Sciences Group (QSG) has hired a new network assistant to work with Mark Klopsch.

The LTER program continues to benefit from close associations with OSU Forest Science and US Forest Service Research.

Lisa Carlson has done some development work on an Oracle interface to the FSDB for the Research Natural Area (RNA) database in our Region.

We hosted a regional monitoring workshop in March, "Improving Natural Resources Management through Monitoring". Over 200 participants attended and there will be a special issue of "Environmental Monitoring and Assessment" published on the proceedings.

We are preparing for our National Advisory Committee visit in November. This summer, Susan Stafford served on the NSF review committee of the Cedar Creek LTER site. All

data managers should be encouraged to serve in this capacity when the opportunity presents itself.

4.1.2 BNZ

We now have our 6 GB (10 media on line at one time) erasable optical jukebox up and running and on line after dealing with some frustrating problems, resulting from the hardware shipping with out-of-date firmware (chip) and other software bugs. The jukebox file system is attached to a SPARCstation 1+ and is mounted and available at bootup. As we and many others of you have surely discovered, image files and other GIS data, such as digital elevation models, eat up an incredible amount of file space. Jukeboxes can be an effective way of dealing with this problem. The Archival Management and Storage System (AMASS) software that drives the virtual file system on the jukebox allows erasable optical disk (EOD) volumes (up to 65,000) to be managed in the jukebox or on-the-shelf (out of the jukebox) in what is often called near-on-line status. Files larger than one EOD (up to 2 GB) can be striped across more than one side of a piece of media. Files in the AMASS system can be listed and searched for, even if they are not physically on-line, with normal everyday unix commands. On-the-shelf data can be moved immediately on-line with only a few commands.

To further streamline access to data we have set up the local SUN networked file system (NFS) to automount the AMASS file system on to any one of the five Sun workstations (in four different buildings!) that requires access to this data, eliminating the need to physically move data around the network as individual need arises. The jukebox is also being used for archiving data with its ability to write out individual standard Unix file system EOD's. ERMAPPER software is now on line for image processing. ERMAPPER uses savable algorithms, eliminating the need for multiple copies of an image as it is manipulated.

We have moved climate data into INGRES and I am working with John Gorentz to get the mailer interface for the climate database up and running.

4.1.3 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 as if we did not have 11 years of experimental data with which to compare it.

4.1.4 CPR

The CPER Bibliography was updated and distributed. It contains listings of publications relevant to the CPER classified by journal articles, technical reports, chapters in books and symposium proceedings, theses, dissertations, and abstracts. The entries are indexed by authors and keywords.

The interactive data access program, ltermenu, that we are developing has been converted to run under the X windowing system, and will soon run on personal computers under MS-DOS. The program previously ran only under Sunview. Ltermenu enables one to examine the CPER data sets by accessing the data and data descriptions using the Internet. The program uses menus to select and display data or data descriptions. The data can be displayed in tabular form, or as graphs. Data files can be downloaded to the local machine and saved, and data items selected by the investigator can be extracted from the data file and saved.

As part of a supplement through the University of Michigan we are developing a library of functions which can be used to transfer data between computers attached to the Internet. The routines can be compiled to use either BSD Unix stream sockets, or to use the X/Open Transport Interface (XTI). XTI is a transport layer independent interface compatible with ATT's Transport Layer Interface (TLI). These routines can be used to transfer entire files, to transfer records of data, and to pass messages between concurrently executing programs. The functions can be accessed from either Fortran or C, and provide a base on which distributed data management tools can be built easily.

4.1.5 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 inter- face with INGRES relation databases, we plan to develop Scientific RDBMS environment which will automate most aspects of data management, and give the Co-Pi'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 soft- ware 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.

4.1.6 HBR

We have now logged approximately 20,000 into our sample storage archive. Use of already stored samples has been somewhat limited, but several proposed studies include analysis or re-analysis of samples. The bulletin board 'Source of the Brook' continues to operate, and we are slowly adding datasets to our archive system. EDEX'es (Ecological Data Exchange) are now available over the Internet via Gopher. Our eventual goal is to duplicate the data on the bulletin board on Internet so that a greater number of people have access to the data.

4.1.7 HFR

A new data manager/ecologist (Lent) was hired this summer at Harvard Forest. We recently completed a first draft of our data catalog, currently with 40 documented data sets. This will be continually revised as we add more data sets and refine our documentation protocol. A catalog of GIS map coverages, containing 184 digital maps and overlays, was completed by Emery Boose. HFR's extensive archives are being prioritized for computerization; these date from over 80 years of ecological research and are mostly in paper form (stored in a fireproof vault). We are also working on getting better network access; our current access to LTERnet is via Sprintnet. Two 486-based PC's will be purchased this fall to aid networking and data management activities. We are considering purchasing a site license for a bibliographic software package (Papyrus) for use by individual researchers and to manage our library's collections.

4.1.8 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 fro 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, fence lines and roads. This effort is in collaboration with the USDA, BLM and Forest Service.

4.1.9 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 e-mail 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.

4.1.10 KNZ

The data management staff at Konza Prairie Research Natural Area was involved with putting together a document entitled "Long Term Ecological Research at the Konza Prairie Research Natural Area: Site Description and Research Summary (1981-1991). This document summarizes the LTER research that has been conducted at Konza Prairie since 1981. This was a good exercise for the data management program and it allowed us to evaluate many of our data sets and documentation. Other important developments included 1) Use of supplement this year to fund acquisition of TDR (Time-Domain Reflectometry) to measure soil moisture. 2) Upgrade of our remote sensing lab 3) The second volume of the CR-ROM series on the FIFE experiment was released. This CR-ROM includes most of the satellite images that was used in the FIFE experiment. 4) Purchase of an OPTICAL drive for our Sun machine and a CD-ROM for our Novell network. 5) Due to the large number of data requests that Konza Prairie LTER is receiving (for the period Jan. 01 1992 to Dec. 31 1992, we received 142 formal request for data from outside scientists), we have installed a IDA (interactive data access) on a UNIX machine. To interact with this system, log on to the machine (bison.konza.ksu.edu) as lterknz with the password as infosys. Simply follow the on-line instructions and remember that UNIX is case-sensitive. This system is modelled after FIS (FIFE Information System).

4.1.11 LUQ

Data Management at the Luquillo Experimental Forest has been involved in completing the details of what we understand are the basics of Data Management in a site: developing the Site Data Management Policies, guidelines to complete the forms, and facilitate the documentation of the data sets. All these 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 (and some additional) information 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 obtained 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, not

forgetting about the documentation of the data sets. An on-line catalog for reference in also in our short term plans for development.

4.1.12 NET

Numerous software additions to the LTERnet Network Support System (NSS) have resulted in extended connectivity and new functions for improved access to the databases at LTERnet (Core Dataset Catalog and the LTER Personnel Directory). LTER researchers without direct Internet connections now have an additional access path to the LTER NSS. We have installed at LTERnet a gateway to the Sprintnet data communications system (formerly Telenet). This connection allows researchers away from their home institutions, or otherwise unable to obtain a network connection or computer account, to achieve access by local phone call to the LTERnet Information System and the Internet, independent of their location. Regular users of this option include researchers at such geographically dispersed LTER sites as Toolik Lake in northern Alaska, Harvard Forest in Massachusetts, and Sevilleta in New Mexico.

The number of users of the LTERnet mail forwarding system has increased steadily, to 80% of the 555 LTER members listed in the on-line personnel directory. To cope with increased workload for computer systems administration, the LTER Personnel database has been moved from a personal computer to the LTERnet communications server, and integrated with the electronic mail forwarding system. Thus, information on the 600-plus researchers in the Network can now be accessed directly on-line and by e-mail. To encourage further interactions within the ecological research community, we have expanded the LTERnet mail forwarding system to include scientists of the NSF Land-Margin Ecosystem Research (LMER) Program (approximately 50 members), members of the Long-Term Study Section of the Ecological Society of America (LTSS; 500 members, 140 of whom have e-mail addresses), and numerous other small research groups within the ecological community. To reflect this important outreach activity of the LTER Network, we have extended the LTERnet communications server to ESnet (Ecological Sciences Network). Thus, for example, a message to the members of the Long-Term Study Section of the ESA can now simply be addressed to LTSS@Esnet.edu (which is the same as LTSS@LTERnet.edu).

During 1992, John Magnuson (NTL) is on sabbatical to do research at the Network Office GIS/RS lab. His work, in collaboration with John Vande Castle, is on comparative landscape ecology of disparate ecological systems using satellite data of all LTER sites.

The system presently includes on-line versions of the LTER Core Data Set Catalog and the LTER Personnel Directory (as well as digital copies of selected publications). The integrated database capability allows information retrieval and maintenance of the databases over electronic networks, as well as generation of e-mail addresses from the personnel directory. The expansion of NSS functions makes these capabilities widely available within LTER and to the greater ecological community. The LTER Core Data Set Catalog and the Personnel Directory can be searched directly or by electronic mail for specific key words, investigators, core areas and other information. Direct access to these facilities via the international Internet is simple and is recommended for all queries. We have taken great care, however, to enable access to our databases via electronic mail.

The basic hardware configuration of the LTER NSS has not changed substantially since last year (A core of two Sun SPARC-2 computers with multi-gigabyte file storage serving integrated mail, database and GIS activities for the Network; connecting via TCP/IP Ethernet LAN to optical disk and optical disk jukebox mass storage systems, 9 track and 8mm tape systems, film recorders, color printers and plotters etc.)

.

4.1.13 NIN

The Baruch Institute Marine Laboratory, located in Georgetown, SC, is under construction and will be reoccupied in late February 1993. The new laboratory has been designed to withstand severe hurricanes and associated storm surges. The facilities offer expanded computational, data management, and archival facilities, in addition to increased space for resident and visiting scientists.

Scott Chapal continues much of his work on developing quality assurance/quality control techniques for ecological data sets. We look forward to his presentation on the topic at the Albuquerque symposium.

4.1.14 NTL

We have made progress on incorporating our meteorological data into our standard data flow. A protocol for error screening the electronically collected data from the NTL-LTER weather station has been developed including a system of data flags to indicate data quality. An Ingres database was created for the meteorological data and a substantial portion of these data have been screened and loaded into Ingres.

We are implementing a user interface for our GIS data. The user interface was designed to permit the user to (1) select an area of interest on a regional map and then see superimposed on the map the boundaries of each GIS coverage that includes the area of interest and (2) select one of these coverages and view associated documentation. The interface is being constructed using the modules and command macros of ARC/INFO.

4.1.15 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 geostatistical 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.

4.1.16 PAL

The first field season from September91 to March92 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 May92 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.

4.1.17 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.

4.1.18 VCR

In February 1992 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.

We have also been taking steps to make our site increasingly accessible over the Internet. This is needed because the VCR/LTER serves an increasingly widespread PI community, with investigators from over 7 institutions in 3 states. We have instituted a daily electronic-mail calendar that keeps everyone informed on activities by LTER researchers. Additionally, we have established the Virginia Coast Reserve Information System (VCRIS), a gopher-based information server which contains current calendars, weather data, the site bibliography, site administrative documents and provides gateways to other systems (such as LTERNET and NSF's STIS).

4.2 Workshop Agenda

A. Friday Evening

1. 6:00 -- Meet for informal dinner excursions

2. 7:30-9:00- INFORMAL meeting
    a) Site flashes
    b) Highlights of recent CC meetings
    c) Agenda highlights and revisions
3. Mixer

B. Saturday Morning

1. 08:00-08:15 Introduction
    a) Agenda adjustments
    b) Logistics for reimbursement
2. 08:15-09:00 Introductions of participants
3. Reports (5 min, max)
4. 09:00-10:15 Activity Reports (5 min max)
    a) Data Management Symposium (Michener)
    b) Climate Database (Briggs)
    c) Interaction with Chinese Ecological Research Network (Brunt)
    d) LTERNET (Nottrott)
        1) LTERnet access through the Sprintnet-->LTERnet connection
        2) LTERnet Bulletin Board
        3) Network Resource Databases
            a> Climate modeling database
            b> Variability database
            c> Remote-sensing database
            d> LTER personnel directory
            e> On-line Data Set Catalog
            f> LTER History/reference file

g> Maintenance and expansion of the LTERnet on-line dataset catalog
e) Electronic Data Collection Report (Ingersoll & Chapal)
f) Interactive Data Access Report (Benson)
g) DATABITS (Porter)
h) LTER DM site reviews (Briggs/Michener)
i) NSF Database Activities Grant proposal (someone from AND)
j) DIRT proposal (Briggs)
k) GPS Network proposal (Michener/Porter)

5. 10:15-10:30 Break

6. 10:30-12:00 Reports (30 min max)
a) All-site bibliography (Bledsoe/Andrews)
b) National Center for Ecological Synthesis and Analysis
c) The "Role of Data Management" at the 1993 All Scientists' Meeting

C. Saturday Afternoon

1. 13:00-13:10 Discussion Topic 1 - Intersite Synthesis Efforts

2. 13:10-13:30 Presentations and demonstrations
a) Data exchange
1) Central Plains system (Kirchner)
2) Cedar Creek system (Elhaddi)
b) Interactive Data Access
1) ANDREWS LTER IDA system

3. 13:30-16:00 Working Groups (preliminary list)
a) Data exchange facilitation
b) Interactive Data Access
c) MSI evaluation
d) Promoting use of LTER databases
e) Data Access Policies at LTER sites
f) Intersite research bibliography (Bledsoe/Andrews)
g) Network Resource Databases
h) Facilitating intersite information manager travel and communication

4. 16:00-16:30 Wrapup and Action Items

5. 16:30-17:00 Presentations and Demonstrations
a) Data input software
1) Clarion (Porter)
2) EasyEntry (Michener)
3) Spreadsheets (Benson)
b) LTERnet Database Query by e-mail and by direct access; examples LTER Dataset Catalog and Personnel Directory
c) Databases of Research Site Locations
d) GPS (Michener, Porter)

D. Sunday Morning

1. 08:00-09:45 Discussion Topic 2 -- Site Data Management

2. 09:45-10:00 Break

3. 10:00-11:30 Working Groups (preliminary list)
a) Solutions to Common Data Management Challenges
1) new sites
2) older sites
b) Additional funding opportunities for data management
1) National Center for Ecological Synthesis and Analysis
c) Quality Assurance/Quality Control

d) Management of Spatial Data
1) Special Meta-data issues for GIS and remote-sensing databases
2) GPS Data Management
e) SQL Database round-table (exchange of problems -- and solutions!)
f) UNIX and DOS round-table

4. 11:30-12:00 Wrapup and Action Items

E. Sunday Afternoon

1. 13:00-13:30 Discussion Topic 3 -- Outreach

2. 13:30-14:45 Working Groups
a) The "Role of Data Management" at the 1993 All Scientists' Meeting
b) Data management outside LTER network
1) What role can LTER DM activities play?
2) How can we facilitate communications?
c) Maintenance and expansion of the LTERnet on-line dataset catalog
d) Chinese DM training
e) Leadership in Biological Information Management
f) Relationships with professional societies
1) ESA Long-term studies section, LTERnet mail forwarding system

3. 15:00-15:20 Working Group Wrapup

4. 15:30-16:30 Workshop Wrapup
a) Agenda for next DM workshop
b) Writing Assignments

5. 16:00-17:00 DataTask meeting (all welcome!)

4.3 Software Demonstrations

4.3.1 Interactive Data Access Program Developed at CPR

The interactive data access program, ltermenu, that we are developing has been converted to run under the X windowing system, and will soon run on personal computers under MS-DOS. The program previously ran only under Sunview. Ltermenu enables one to examine the CPER data sets by accessing the data and data descriptions using the Internet. The program uses menus to select and display data or data descriptions. The data can be displayed in tabular form, or as graphs. Data files can be downloaded to the local machine and saved, and data items selected by the investigator can be extracted from the data file and saved.

As part of a supplement through the University of Michigan we are developing a library of functions which can be used to transfer data between computers attached to the Internet. The routines can be compiled to use either BSD Unix stream sockets, or to use the X/Open Transport Interface (XTI). XTI is a transport layer independent interface compatible with ATT’s Transport Layer Interface (TLI). These routines can be used to transfer entire files, to transfer records of data, and to pass messages between concurrently executing programs. The functions can be accessed from either Fortran or C, and provide a base on which distributed data management tools easily can be built.

Attribute-Value Syntax for Describing Data

We are in the process of converting our data descriptions to a new format (Figure 1). The new format provides considerably more flexibility in describing data files than the previous format. It is based on a C-style

description of data structures. The structures can be simple variables, such as FileName, or more complicated structures, such as Column_format (Figure 2). Structures can be nested with structures. The complex structures use {} to delimit the beginning and end of the structure. A structure consists of an optional type specifier, an attribute label or "tag", and a simple or complex data item.

By default, the data “values” are character strings. The special types INTEGER, REAL, TEXT, and LIST (perhaps more later) can also be used to identify the type of a data element. A TEXT item has one or more character strings enclosed by {} braces. INTEGER and REAL flag numeric types. A LIST is one of more character strings, and is handled much like TEXT. TEXT is differentiated from LIST to help in the process of reformatting the metadata for printing.

We have written a parser that will return the values of items that match an attribute or, for complexstructures, the specific value for an attribute. The parser assumes nothing about the names of attributes, nor about the complexity of the structures, so one can add new structures at any time. One can also include an item in one structure, such as “Units” within the Item structure, without including a Units record in all Item structures in the file.

The general syntax for a data description is

<TYPE_DECLARATION> Attribute (value | value {)

Things enclosed in brackets <> are optional and | denotes alternative selections. The default type for an attribute is STRING, which is a string of characters beginning with a non-whitespace character and ending with a newline.

SYNTAX RULES:

1. Complex attributes are structures and are identified by having { as their last character.

2. Names of attributes can contain no whitespace characters

3. The value for a simple attribute starts with the first non-whitespace character following the attribute and ends with a newline character. The interpretation of newline characters as the end of a value can be overridden by “escaping” them with a \. The definition of a complex attribute ends with a }.

4. The brackets {} enclose a value that is itself made up of 1 or more attribute-value pairs or, for fundamental types such as TEXT, embedded newline characters.

5. The $ character flags the start of an attribute name that is defined within the file. For example, the EXTERNAL_FUNCTION named DisplayData takes as its argument the name of the data file, which is declared elsewhere in this file.

6. All possible elements within a complex attribute do not need to be defined. For example, Missing_codes is defined only for the Item called NUMBER. The parser returns a flag indicating that an attribute is not defined if the attribute is omitted from an instance of a complex attribute.

The value of a complex attribute can be thought of as the tag for a structure. A generic parser for such a description can look up and return the values for specific attributes, such as

Column_format.[ID].Start_column

or

Column_format.[NUMBER].Missing_codes.BMDP

with. used to separate hierarchical levels within the data structures. The parser would look up the data by matching the elements of the string (ID, etc.) against the attributes and values. The specification for a value within a particular structure uses brackets ([]) to identify the value of the structure, as in:

Column_format.[Item]ID.Start_column

This additional specification could be used to allow the label ID to be associated with more than 1 attribute within Column_format, as long as the ID attributes were of different types.It would also allow one to formulate a request for all values of Variables using the wildcard character *, as in:

Column_format.*.Variable

Such a request could be used to generate a menu of data items that could be selected from

the data

file.

The fundamental types would be INTEGER, REAL, TEXT, LIST, EXTERNAL_FUNCTION and perhaps some other types we would agree upon as a group to facilitate data exchange.

EXTERNAL_FUNCTION is included as a fundamental type to facilitate using the data description as a class in an object oriented data access system. Such a system would be able to handle non- ASCII data in a logical fashion. For example, if the data were a bit map for an image, the Display- Data function could be a program to display the image on a workstation. If the data were stored as binary values, the DisplayData function could be a program to convert the data to ASCII

and then display it. If a site frequently exported data to spreadsheet programs then a function like CreateDIF could identify the appropriate program to call to convert the data to a DIF format.

This new syntax was developed specifically to facilitate the management of non-ASCII data, such as map images, and to facilitate the exchange of data with other sites. The interactive data access program we are developing, ltermenu, is being converted to use the new format for data descriptions. We have been cooperating with the VCR site to enable them to use the metadata tools that we develop at their site. The ltermenu program and data description software that we develop will be made publicly available.

FIGURE 2. An example of a data description file.

```
FileName  btpw71f
Data_set_id    btemp71
Data_set_type  ASCII
Title     1971 belowground temperature (pawnee) summary data file
```

```
Fortran_format      (a2,i2,a3,3i2,i3,i4,21f5.1)
INTEGER    Number_data_items  3

Column_format {
    Item datype {
        Variable datype
        Type  a
        Definition data type (‘59’)
        Units      <something ... here if not currently an n/a>
        INTEGER Start 1
        INTEGER End 2
    }

    Item date {
        Variable date
        Type  i
        Definition date of sampling
        INTEGER Start 3
        INTEGER End 8
     }

     Item rain {
        Variable rain
        Type  f
        Definition Amount of rainfall
         INTEGER Start 9
        INTEGER End 12
    }
}

Missing_value_code  M

Description {
    Principal_investigator {
        Name George Van Dyne
        TEXT Address {
            some department
            CSU ... etc
        }
        Phone     491-0000
    }

    Technician Ray Souther
     Date_of_beginning_study      71/01/01
     Date_of_ending_study    71/12/31
    Sampling_frequency       daily after 26/4/71

    TEXT General_purpose {
        Minimum, maximum and average soil temperatures in
        degrees Centigrade are recorded daily for belowground
        depths of 1, 2.25, 4, 8, 20, 40, and 72 inches.
    }

    LIST Keywords {
        temperature
        soil
    }

    Data_form          nrel-59
    Location_of_data_forms                microfilm
    TEXT Data_entry_instruction {
        A record will appear for each day of the year. Missing
        code is ‘-99.9’ for temperatures and ‘-99’ for time.
        This data is created by program ‘bgtemp’.
    }

    TEXT Permanent_locations {
        Central Plains Experimental Range Station,
        Nunn, CO Elevation 1652m Exclosure E1/2 Section 23 T10N R66W 6th P.M.
    }

    TEXT Related_data_sets {
        Pawnee Standard Weather measurements are recorded
        at same time and location as the belowground temperature measurements
     }

    TEXT Restrictions_on_use {
```

```
        See Bill Lauenroth
    }

    Programs {
        Program deptmps {
            Name deptmps
            TEXT Description {
                This program, written in fortran 66 for use on the CYBER, is
                not presently used. It is stored on magnetic tape n0001 at
                NREL. It was replaced by program 'bgtemp' in 1986.
            }
        }

        Program bgtemp {
            Name bgtemp
            TEXT Description {
                This program, written in fortran 77 for use on the VAX, uses
                field data as input. It produces a summary data file
                consisting of maximum, minimum, and average temperatures in
                degrees Centigrade. A julian date is added to each record.
                A record exists for each day of the year.
                The missing code is '-99.9' for temperatures and '-99' for
                time. In addition, a human readable file is created that
                includes daily records with temperatures in degrees Centigrade
                and monthly averages. Summaries of monthly averages are given
                at the end of each year and the end of the entire data set.
                Missing code for this file is 'M'. This program is located on
                the VAX and on magnetic tape n0001.
            }
        }
    }
    Programmer          jerry d. peltz (program deptmps); cinda a. liggon
}

EXTERNAL_FUNCTION   DisplayData    /bin/cat $FileName
EXTERNAL_FUNCTION   SelectData     ~LTER/bin/Extract $THIS_FILE
```

FIGURE 3. Another example of a data description file.

```
FileName  ~LTER/DATA/BG/Arth/BGAR78
Title     Belowground arthropods
Location{
    Site ESA
    Treatment G
}

Fortran_format (a3,3i2,a4,i3)
Column_format  {
    Item ID {
        Variable  ID
        Definition     initials of the data collector
        INTEGER Start_column     1
        INTEGER End    3
        Missing_value  ..
        }
```

Retrieving values from the data description

To get a list of all the values for Column_format.Item you search for:

```
Column_format.Item
```

To get the starting column for an item called ID you search for:

```
Column_format.[ID].Start
```

To get all of the information in the ID structure you search for:

```
Column_format.[ID].*
```

Data is returned in a structure

```
struct Value {
```

```
    int Type;
    int ValueType;
    char *String[100];
    int n;
    union Numeric {
        float Float;
        long Int;
    } Numeric;
};
```

4.3.2 CDRLIB: A library of Functions for Scientific Data Manipulations

Demonstration of data conversion utilities based on CDRLIB was presented by Abderrahman A. El Haddi.

CDRLIB is a library of functions used for manipulating scientific data and converting it between different software applications. CDRLIB was developed at Cedar Creek LTER.

Below, two utilities are used to demonstrate how CDRLIB can be used as a fast data integration tool.

Convert

This utility can be used to convert data among different formats. An intermediary format can be used as a common bridge among all other formats. This we call header format. The basic idea is to have an ASCII file with variable names, variable label, formats and any documentation. The scheme is flexible enough and supplies the variable names “var1,var2,..,varn” if they are omitted. The variable labels and formats are also optional. If the format is specified, the data is read accordingly and warnings are issued if any data point does not conform to the format. If no format is specified, the data format will be determined automatically and an optimal format is used.

Convert will convert data among various formats for different architectures. Examples of supported formats include: Statview, Cricket Graph, JMP, SAS, DIF, etc.

Buildsas

This utility is dedicated to building SAS files once the data is in header format. This is a useful utility for building large data libraries since it can be used from shell scripts or batch files. The following are some examples of how buildsas can be used to convert ASCII files to SAS files without explicitly going through a SAS data step.

Listing1: Example of an ASCII file: nx82e01

```
D 1 31 2 3 1982 08 11 12 128.3760 1.4214 0.5720 4.1427 6
D 1 32 8 80 1982 08 11 30 1359.0402 2.3909 0.7030 10.9234 12
D 1 33 9 0 1982 08 11 8 91.1410 1.4772 0.7104 4.3808 8
D 1 34 9 0 1982 08 11 14 150.5520 1.7398 0.6592 5.6960 10
D 1 35 3 6 1982 08 11 19 317.5710 2.2132 0.7516 9.1445 11
D 1 36 1 0 1982 08 11 16 171.9200 1.8389 0.6632 6.2896 10
D 1 37 4 10 1982 08 11 16 318.6420 1.6625 0.5996 5.2722 8
D 1 38 1 0 1982 08 11 15 206.0030 1.9013 0.7021 6.6947 7
D 1 39 1 0 1982 08 11 14 165.5400 1.7846 0.6762 5.9572 9
D 1 40 6 28 1982 08 11 21 594.2659 1.7493 0.5746 5.7503 10
D 1 41 5 16 1982 08 11 17 353.8029 1.7799 0.6282 5.9291 6
D 1 42 2 3 1982 08 11 18 206.3650 1.5221 0.5266 4.5819 9
```

We can convert nx82e01 to the header file nx82e01h by issuing the following command:

```
convert -f r -t h -v ‘field expt plot ntrt nitradd year month \
        day SR totbio shanwinr evenness diversty adjSR’ -o nx82e01h nx82e01
```

convert will use the default dictionary to extract variable labels and any notes and insert them with the new output file “nx82e01h”. This new file is shown in listing 2.

Listing 2: An example of a header file obtained using convert.

```
## file name :nx82e01h : archival media:
##
##
##
```

```
## Cedar Creek Long Term Ecological Research
## Ecology, Evolution and Behavior
## University of Minnesota, 318 Church St. S.E., MPLS, MN 55455
##
## Header format[Column(i) : variable abbreviation : variable description :format]
##
##
## Column01 : field : Old field name :char[%1s]
## Column02 : expt : Experiment number :int[%12d]
## Column03 : plot : Plot number :int[%5d]
## Column04 : ntrt : Nitrogen Treatment :int[%4d]
## Column05 : nitradd : Nitrogen fertilizer addition (g/m2/yr) :int[%4d]
## Column06 : year : Sampling year :int[%4d]
## Column07 : month : month of the year (1=Jan,...,12=Dec ) :int[%2d]
## Column08 : day : day of the month :int[%2d]
## Column09 : SR : Species richness :int[%3d]
## Column10 : totbio : Field avg. total biomass (g/m2) :float[%10.4f]
## Column11 : shanwinr : Shannon-Wienner Index :float[%7.4f]
## Column12 : evenness : Evenness :float[%7.4f]
## Column13 : diversty : Diversity :float[%8.4f]
## Column14 : adjSR : Adjusted Species Richness(ASR=SR(for prop>0.01)) :int[%3d]
##
## Additional notes:
##
## prop = species mass / total species biomass
##
## DATA:
##
D 1 31 2 3 1982 8 11 12 128.3760 1.4214 0.5720 4.1427 6
D 1 32 8 80 1982 8 11 30 1359.0402 2.3909 0.7030 10.9234 12
D 1 33 9 0 1982 8 11 8 91.1410 1.4772 0.7104 4.3808 8
D 1 34 9 0 1982 8 11 14 150.5520 1.7398 0.6592 5.6960 10
D 1 35 3 6 1982 8 11 19 317.5710 2.2132 0.7516 9.1445 11
D 1 36 1 0 1982 8 11 16 171.9200 1.8389 0.6632 6.2896 10
D 1 37 4 10 1982 8 11 16 318.6420 1.6625 0.5996 5.2722 8
D 1 38 1 0 1982 8 11 15 206.0030 1.9013 0.7021 6.6947 7
D 1 39 1 0 1982 8 11 14 165.5400 1.7846 0.6762 5.9572 9
D 1 40 6 28 1982 8 11 21 594.2659 1.7493 0.5746 5.7503 10
D 1 41 5 16 1982 8 11 17 353.8029 1.7799 0.6282 5.9291 6
D 1 42 2 3 1982 8 11 18 206.3650 1.5221 0.5266 4.5819 9
D 1 43 2 3 1982 8 11 23 373.5160 2.3324 0.7439 10.3024 13
D 1 44 9 0 1982 8 11 16 280.5980 1.7822 0.6428 5.9430 8
D 1 45 7 50 1982 8 11 23 759.1290 2.2171 0.7071 9.1807 11
```

Now that we have a header file we can convert it to any format we desire. The following is a script file that will build SAS files from header files. Of course the user does not have to know the contents of this script file, but instead type ‘‘bs infile outfile’’

```
#! /bin/csh
#
# name: bs
#
# Purpose:
#
# Builds SAS files from header files
# for multiple files use mbs instead to process multiple files
#
#
# first arg : header file name ( input )
# 2nd arg : SAS file name ( output )
#
#

if ( $#argv != 2 ) then
 echo “Usage: $argv[0] infile outfile”
 exit -1
endif

if ( $?SASLIB == 0 ) then
 setenv SASLIB ‘pwd‘
 echo “Using default SASLIB=$SASLIB”
endif

if ( $?SSD == 0 ) then
 setenv SSD ‘ssd01’
 echo “Using default $SSD as default extension for SAS files”
```

```
endif

if ( -f $SASLIB/$2.${SSD} || -f $SASLIB/$2.ssd ) then
 echo “File $SASLIB/$2.ssd[01] exists “
 echo “ “
 echo -n “Erase it ? (y/n) “
 set ans=$<
 echo $ans
 if ( “$ans” != “y” && “$ans” != “Y” ) then
 exit -1
 endif
endif

buildsas -i $1 -l $SASLIB -o $2 -p ${1}.sas -v -e -f

egrep -i “warn|err” ${1}.log

# Clean up my mess.

\rm junk@@* *.log *.sas
```

The script “B.S.” can now be used to build multiple files with a single command such as “mbs file1 file2 ... filen ‘“or “mbs *”. Note how grep is used to extract any SAS Errors or warnings.

Buildsas will write a SAS program and SAS is then invoked using a system call.

Listing 2 shows the contents of mbs (build multiple SAS files.)

```
#! /bin/csh
#
# name: mbs
#
# Purpose:
#
# Build SAS files for an entire set of files using the
# header file also as a name for the SAS file with the *.ssd01
# as an extension. Note that ssd01 can be set using setenv.
#
# bs expects the first argument to be the header file and
# the second argument to be the SAS file ( output ) without the (ssd )
# extension.
#

foreach file ( $argv )

 bs $file $file
 \rm junk*

end
```

Using the example above, we can now convert the nx82e01h header file to a SASFILE by issuing the command:

```
mbs nx8?*h*
```
or
```
bs nx82e01h nx82e01h
```

4.3.3 GQL (Graphical Query Language) by Andyne Computing

Barbara Benson demonstrated a front-end package which allows users to access a variety of SQL-based database management systems (DBMS) through a simple graphical user interface. The North Temperate Lakes site is considering purchasing this software to make the data which reside in Ingres databases readily accessible to researchers. The GQL product provides access from Macintosh, Windows, or Unix Motif platforms. GQL generates both ad hoc and predefined queries. The query results are displayed in spreadsheet format where the user can do local sorting and manipulation or create reports. The data can be seamlessly integrated with other applications.

4.3.4 Sprintnet->LTERnet Access

Rudolf Nottrott demonstrated access to the LTERnet system using a dial-up connection through a local (Honolulu) phone number. Using the Sprintnet<->LTERnet connection, Rudolf demonstrated queries (direct and by e-mail) of the Core Dataset Catalog and the LTER Personnel Directory. The communications software used was Procomm Ver. 2.4.2., a ”user supported” software package available for a contribution to the developers ($25-50, Datastorm Technologies). Call Datastorm, (314) 449-9401, or contact Rudolf Nottrott for a trial copy of Procomm 2.4.2.

Instructions on how to use the Sprintnet<->LTERnet access can be obtained by sending any message to Sprint@LTERnet.edu (Sprint@LTERnet on Bitnet). For an overview of other LTERnet functions send any message to Overview@LTERnet.edu.

4.3.5 Clarion Database Application Generator Demonstration

John Porter presented a brief talk on Clarion Professional Developer, a database application (DBA) generator. What is a DBA? Lets take it one word at a time. As the database part of the name suggests, it has to do with storing, manipulating and outputting data. All database programs have certain capabilities in common, with few exceptions they let you input, sort, query (search) and output data. The application part of DBA refers to a set of programs that make it so that a specific database can be manipulated with a minimum level of user training using pre-specified menus and report formats. Database applications are needed because the commands built into most database programs can be quite complex and cryptic to novices. After all, you don't want a salesman to have to spend hours or days to learn enough about database commands to search for the price of a part.

It used to be that database applications were very costly to develop. Expert database programmers were required to develop the thousands of lines of database code needed to make easy-to-use (but not to program) menus and report formats. Enter the DBA generator. A DBA generator provides a set of tools that simplify the writing of DBA's so that even non-programmers can create usable applications. Tools typically include a “screen painter,” which allows you to create menus and input screens by moving blocks of text or input fields around on the screen and a report generator, which allows you to design headings, columns and grouping and total fields for reports. What the generator does is convert the output from the screen painter and report generator into a program by actually writing code in the language used by the database program (most database programs have their own languages). If its necessary to have the application do something that cannot be specified by the screen painter, you can add your own code to the program generated by the application generator.

These days most professional database programs have application generators which vary from poor to very good. Clarion Professional Developer 2.1 seems to be a good one. John found it relatively easy to generate simple menu-driven applications. Its screen painter was easy to use, once John got the hang of the function and control keys. Fortunately, there's on-line help in case you forget! It has several nice features. It has one screen that gives you a list of data files that are available for use in the application and then a “tree style” diagram that lists each of the menus and subroutines in a diagrammatic form. This makes it easy to keep track of what menu item does what. Subroutines or sub- menus that have been referenced but not yet designed are clearly labeled “ToDo.” This is a nice feature because applications (at least in my experience) aren't written so much as evolved. The tree diagram makes it quick and easy to figure out what menu or report you need to edit, even in applications whose previous draft was written weeks or months earlier. To modify a data file menu or report, you just move to the name in the tree diagram using cursor keys and hit enter. The design screen then comes up automatically with the appropriate menu or database displayed. There are lots of things you can do to customize menus (such as select screen colors), but mostly that is unnecessary as the Clarion defaults are pretty good.

One of the things that distinguishes Clarion is that you can generate standard PC.EXE files of your application. This means that you can give copies of your applications investigators and students who don't own a copy of Clarion. Another advantage of using compiled code is that it runs fast! Unlike DBASE IV, which has a command line interface where individual commands can be entered, Clarion is all menu and program driven. Even the commands to create a database file are in the form of a program (usually one generated by a Clarion design screen). Clarion has several basic tools. A designer takes care of the basic design of data-

bases, menus and reports. A compiler takes the code generated by the designer and lets you run and debug it within Clarion. If bugs do appear (unlikely in automatically generated code) an editor jumps you to the point in the program causing the problem. Once the program has been tested, the translator converts it to an .EXE file (this isn't done earlier because the conversion takes a while and you don't want to redo it every time you make minor changes in the program). Other utilities are a “help” editor that facilitates adding on-line help to your applications, a “scanner” that displays data files in a spreadsheet form, a database sorting and reformatting utility and a file translation program that lets you read data into a Clarion database from several database or ASCII formats.

Help in Clarion is generally good and context sensitive. John’s only gripe is that sometime you “just can't get there from here” if you are trying to check out a program feature that is not really pertinent to the current screen display. There are nice tutorials. A freely distributable product information disk containing the tutorials of several Clarion products is available from Clarion.

The Clarion report program requires some special explanation. It can create reports that combine data from several different database files (including databases in non-Clarion formats). Unlike regular Clarion applications which are simple .EXE files, reports generated by the report program (but not the somewhat simpler report designer built into the Clarion application designer) require that you distribute a “runtime” program called RRUN along with your report specification.

4.4 List of Data Management Workshop Participants

Susan Stafford (AND)

Oregon State University, Department of Forest Science,

Forestry Sciences Lab 020, Corvallis OR 97331-7501

LTERnet address: SStafford (or SStaffor), E-Mail: stafford@fsl.orst.edu

Phone: (503) 737-6581, FAX: (503) 737-1393

Donald L. Henshaw (AND)

USDA Forest Service, Pacific NW Station,

3200 SW Jefferson Way, Corvallis OR 97331

LTERnet address: DHenshaw, E-Mail: henshaw@fsl.orst.edu

Phone: (503) 750-7335, FAX: (503) 750-7329

James Laundre (ARC)

Marine Biological Laboratory, Ecosystems Center, Woods Hole MA 02543

LTERnet address: JLaundre, E-Mail: jimL@lupine.mbl.edu

Phone: (508) 548-3705, FAX: (508) 457-1548

Mark Klingensmith (BNZ)

University of Alaska, Forest Soils Laboratory,

305 O'Neill Building, Fairbanks AK 99775-0740

LTERnet address: MKlingensmith (or MKlingen), E-Mail: mklingen@taiga.lter.alaska.edu, fnmlk@alaska (Bitnet)

Phone: (907) 474-7036, FAX: (907) 474-7439

Abderrahman A. El Haddi (CDR)

University of Minnesota, Department of Ecology, Evolution and Biology

100 Ecology Bldg, 1987 Upper Buford Circle, St. Paul MN 55108

LTERnet address: AElhaddi, E-Mail: elhaddi@cdr.lter.umn.edu, elhaddi@cs.umn.edu

Phone: (612) 625-7930, FAX: (612) 625-4490

Thomas B. Kirchner (CPR)

Colorado State University, Natural Resource Ecology Laboratory and Department of Range Science,

Fort Collins CO 80523

LTERnet address: TKirchner (or TKirchne), E-Mail: tom@chloris.nrel.colostate.edu

Phone: (303) 491-1986, FAX: (303) 491-1965

Gil N. Calabria (CWT)

University of Georgia, Institute of Ecology, Athens GA 30602

LTERnet address: GCalabria (or GCalabri), E-Mail: gil@sparc.ecology.uga.edu, gcalabri@uga (Bitnet)

Phone: (706) 542-5691, (706) 542-2968, FAX: (706) 542-6040

Cindy Veen (HBR)

USDA Forest Service, Forestry Sciences Laboratory

P.O. Box 640, Durham NH 03824

LTERnet address: CVeen, E-Mail: c_Veen@unhh.unh.edu, or S24L06A (USFS)

Phone: (603) 868-5692, FAX: (603) 868-1538

Emery R. Boose (HFR)

Harvard University, Harvard Forest, Petersham MA 01366

LTERnet address: EBoose, E-Mail: harvard.forest (Omnet)

Phone: (508) 724-3302, FAX: (508) 724-3595

Richard A. Lent (HFR)

Harvard University, Harvard Forest, PO Box 68, Petersham MA 01366

LTERnet address: RLent, E-Mail: harvard.forest (Omnet)

Phone: (508) 724-3302, FAX: (508) 724-3595

Barbara Nolen (JRN)

New Mexico State University, Department of Biology,

Box 30001, Dept. 3AF, Las Cruces NM 88003

LTERnet address: BNolen, E-Mail: bNolen@nmsu.edu

Phone: (505) 646-4465, FAX: (505) 646-5665

Lolita Krievs (KBS)

Michigan State University

W.K. Kellogg Biological Station

Hickory Corners MI 49060-9516

LTERnet address: LKrievs, E-Mail: krievs@msukbs (Bitnet)

Phone: (616) 671-2214, FAX: (616) 671-2351

John M. Briggs (KNZ)

Kansas State University, Division of Biology, Ackert Hall, Manhattan KS 66506-4901

LTERnet address: JBriggs, E-Mail: jmb@andro.konza.ksu.edu

Phone: (913) 532-6629, FAX: (913) 532-6653

Eda C. Melendez (LUQ)

Terrestrial Ecology Division, P.O. Box 363682, San Juan PR 00936

LTERnet address: EMelendez (or EMelende), E-Mail: e_melendez@upr1.upr.clu.edu

Phone: (809) 767-0350, FAX: (809) 758-0815

Rudolf W. Nottrott (NET)

University of Washington, College of Forest Resources, AR-10, Seattle WA 98195

LTERnet address: RNottrott (or RNott), E-Mail: rNott@lternet.edu

Phone: (206) 543-8492, FAX: (206) 685-0790

Caroline S. Bledsoe (NET)

University of California-Davis, Department of Land, Air and Water Resources / Hoagland Hall

Davis CA 95616

LTERnet address: CBledsoe, E-Mail: csBledsoe@ucdavis.edu

Phone: (916) 752-0388, FAX: (916) 752-1552

William K. Michener (NIN)

J.W. Jones Ecological Research Center, Route 2, Box 2325, Newton GA 31770

LTERnet address: WMichener (or WMichene), E-Mail: michener@inlet.geol.scarolina.edu, a299050@univscvm (Bitnet)

Phone: (912) 734-4706

Barbara J. Benson (NTL)

University of Wisconsin-Madison, Center for Limnology

680 N. Park Street, Madison WI 53706

LTERnet address: BBenson, E-Mail: bbenson@macc.wisc.edu

Phone: (608) 262-2573, (608) 262-3088, FAX: (608) 265-2340

Martha Andrews (NWT)

University of Colorado, INSTAAR, Campus Box 450, Boulder CO 80309-0450

LTERnet address: MAndrews

Rick Ingersoll (NWT)

University of Colorado, INSTAAR, Campus Box 450, Boulder CO 80309-0450

LTERnet address: RIngersoll (or RIngerso), E-Mail: ingersoll_r@cubldr.colorado.edu

Phone: (303) 492-4771, (303) 492-0566, FAX: (303) 492-6388

Karen Baker (PAL)

UCSD/SIO, A-018, La Jolla CA 92093-0218

LTERnet address: KBaker, E-Mail: karen@crseo.ucsb.edu

Phone: (619) 534-2350, FAX: (619) 534-2997

James W. Brunt (SEV)

University of New Mexico, Department of Biology, Castetter Hall,

Albuquerque NM 87131-1091

LTERnet address: JBrunt, E-Mail: jBrunt@sevilleta.unm.edu

Phone: (505) 277-9342, (505) 277-9372, FAX: (505) 277-0304

John H. Porter (VCR)

University of Virginia, Department of Environmental Science, Clark Hall

Charlottesville VA 22903

LTERnet address: JHporter, E-Mail: jhp7e@virginia.edu

Phone: (804) 924-7761, FAX: (804) 982-2137