LTER DATABITS, Spring 1994

Scientific Information Management Newsletter of Long-term Ecological Research

---

[Downlaod .pdf version](#)

Table of Contents

Welcome to the Spring issue of Databits. Appropriately networks and their uses dominate this issue of Databits as publication on the network becomes a real possibility. The paper version of Databits will continue to be produced, but Databits will be implemented as an electronic publication as well. Past issues of Databits are available on the LTERNET gopher and indexed versions are available on the Virginia Coast Reserve Information System. Thanks to everyone who took the time out of busy schedules to write something for Databits!

1994 Data Managers' Workshop Scheduled

The 1994 Data Managers' Meeting will be held in Seattle from September 22 through September 24. Members of the planning committee for the meeting (dmmoc@lternet.edu) are Karen Baker, Barbara Benson, Rick Ingersoll, David Jones, Rudolf Nottrott, and John Porter.

The first day of the meeting will be closed for LTER Data Managers to conduct LTER business. The next two days will involve outreach to other groups including participants from other ecological research networks and outside speakers on topics. The first outreach day will have the theme Spatial Data Management Issues. The second outreach day's theme will be Intersite Data Access (including data

exchange formats and linkage of databases).

Rudolf Nottrott is arranging for a meeting room which has Internet access with the expectation that many of the presentations will include demonstrations. There will be time in the agenda for LTER sites to spotlight developments in the theme areas. Let the organizing committee know if your site has accomplishments you want to present.

Site Flashes

AND -- Susan Stafford is now the Director for the NSF Biological Instrumentation and Resources Division for this calendar year. Susan will be making monthly visits home to Corvallis, but will play a limited role in Andrews activities. Susan has been greeted with many new challenges as the Division Director, but hopes to breathe new life into the BIR Database Activities Program. She is planning on attending the September Data Manager's meeting as an NSF representative.

The Andrews has initiated a complete revamping of its climatic and hydrological measurement program. Measurement collection at the two primary Meteorological Stations will become more standardized in terms of variables collected and methods used. A third Meteorological Station will be installed at a high elevation site this summer, and a fourth is planned for 1995. New secondary stations for the collection of precipitation and temperature will also be added, and existing precipitation and temperature networks will be modified (automated with Campbell CR10 recorders) or eliminated. All existing Andrews meteorological data is also being reformatted. New standardized formats will help pave the way for the development of an interactive data access system for all of the Andrews climatic and hydrological data.

New phone lines have been run to the new Andrews headquarters, and we are in the process of setting up our network connection between the Andrews and Oregon State University. The initial installation will use a pair of Netblasers to provide dial-up support for TCP, IPX, and APPLETALK communications at 14.4 baud. Later this year, we expect to upgrade the link to a dedicated 56kb connection with dial-up backup.-- *Don Henshaw & Mark Klopsch, Andrews Forest LTER, DHhenshaw@lternet.edu & MKlopsh@lternet.edu)*

NTL-- Our site has acquired a Sun SPARC 10 computer and we have installed Oracle database software on the Sun. The data management staff has been busy learning Unix and designing the Oracle database. Prior to getting Oracle, our databases had resided in Ingres on a centralized campus Vax cluster. The staff at the computing center performed some of the tasks which we are now taking on (backup, space management, etc.).

We are excited about making our database more directly accessible to the researchers here. We are planning to buy Oracle Data Browser which will provide an easy interface for end users to the Oracle databases.

We have recently gotten a Gopher Server running on our Sun (thanks to Rudolf and El Haddi for some

timely help). Mark MacKenzie is planning on putting the "tour" of images many data managers saw last summer in Madison into WWW format.--*Barbara Benson, North Temperate Lakes LTER, BBenson@lternet.edu*

NWT--We continue to make progress in terms of centralization and standardization of data quality control/quality assurance, archival, and backup protocols. We now have 63 data sets archived in a single directory on our workstation and each of these files is identical in structure, in ASCII format, and includes the metadata. Moreover, we are retroactively applying recently formalized rigorous protocols to pre-centralization data sets, particularly our climate data.

Although they have been in existence for only about one year, our data management "manuals" are now 192 and 250 pages in length. They include general information on data entry and processing procedures, as well as backup and archival policies. All entry and processing programs are included along with detailed documentation for each of these programs. These manuals are "living" documents and new hard copies are generated once every few months.

Our data management staff has also provided tutorials (Unix, vi, etc.) to LTER investigators and graduate students to assist them in making the transition from VMS to Unix. (The University of Colorado will, for the most part, phase out VMS as the primary campus operating system within the next several months.) In many cases, the initial panic is being replaced with excitement as these researchers discover some of the neat things that can be done within the Unix environment.

Our site was visited by representatives from most of the other sites immediately prior to the All Scientists' Meeting. This field trip was quite successful and the attendees had a "taste" of the Niwot Ridge experience and its fickle weather --- snow, wind, thunder, lightning, and even bright sunshine!

Our 2 major "new" research initiatives, the 100+ year snowfence experiment and the subnivean shelter, are now beyond the planning stages. The snowfence was erected and the subnivean shelter constructed during this past autumn. Instrumentation of the latter is underway. These projects will provide long-term data that will enable us to answer heretofore unanswered questions relevant to ecological processes in alpine tundra. -- *Rick Ingersoll, Niwot Ridge LTER, RIngersoll@lternet.edu*

PAL-- The third year of Palmer LTER field work is in progress off the Antarctic peninsula. The Spring cruise in August93 and the annual cruise in January94 have been completed. The second on-station field season with weekly sampling is in progress and will end March94. The work includes oceanographic, krill and bird surveys. Two unix workstations were used for the first time at sea and on station in addition to macintosh and ibm/pc's allowing more rapid and complete data processing.

There was continued focus on dataforms. Documentation completion while in the field was encouraged and largely successful during both cruises this season. It stimulated further consideration of definitions and core datasets. Metadata has been made available within a gopher structure. Also, a draft data policy was developed while considering data storage and coordination. With additional hardware available, a

second remote campus site appletalk network was linked to the central palmer lter appletalk network using tunneling across the campus internet.*-- Karen Baker, Palmer LTER, KBaker@lternet.edu.*

VCR--Our renewal proposal is finished at last! During the writing process we made extensive use of the network, from e-mail to gopher servers, to keep our widely-flung PIs involved in the process. At a final wrap-up session, five computers were arrayed around a room so that PIs could directly make changes to specific parts of the proposal. FRAMEMAKER was used and its book composition features made integration of different sections and page numbering a snap. The proposal text is currently available on our Gopher system and we are working on adding it to our WWW hypertext server. At UVA there is increasing interest in SGML and electronic texts in general. The library is even giving courses on creating electronic texts!

Connectivity Team Strikes Again!

The Division of Environmental Biology at the National Science Foundation recently invited the LTER connectivity team to host a workshop/demonstration on "Connectivity", or "How to Use the Global Electronic Communications Systems for Data Management" purposes. The workshop was intended to be a "realistic" look at connectivity the LTER network as a demonstration of how data and information are managed and communicated via networking. The event was aimed at building interagency understanding and use of the Internet as way to facilitate collaboration and foster networks of scientists among the federal agencies. The workshop was held on December 16 1993 at NSF's new networked building in Arlington VA.

The workshop was designed to demonstrate a user-friendly, detailed look at how one can electronically search, retrieve, manipulate, analyze and utilize science data from a variety of sources, globally, via the INTERNET. The connectivity team, yours truly (SEV), Rudolf Nottrott (NET), and John Porter (VCR) demonstrated the use of electronic mail, telnet, FTP, remote computing, Gopher, Khoros, and Mosaic on SunOS and DOS platforms. The workshop was attended by representatives from government organizations and public organizations, including the Army, the Cousteau Society, EPA, NBS, NOAA, NPS, USFS, SCS, the Smithsonian Institution and World Wildlife Fund, The agency representatives came from every level of technical sophistication.

As with other similar exercises, the connectivity team had complete confidence in the integrity of the network and showed up in Washington D.C. with little more than our laptops. Hopes were dashed as once inside the newly constructed building the group discovered entrails of wiring closets scattered about, periodic slowdowns and whispered talk of yet another network outage. Fortunately, the network rallied and remained alive long enough to transfer all necessary programs and data across to DEB's Sun workstation and PC. Like thieves in the night we moved the DEB Division Director's computers down to the conference room several floors below. Technical issues resolved, all was fine until we realized that all three of us had prepared to give the same presentation which made for some interesting scrambling and a late night. Finally, we broke it up so that I presented an overview of connectivity, John Porter gave an overview of network tools and Rudolf Nottrott succinctly described how the Network Office systems

facilitated communications within LTER. All our preparations paid off when the NSF network server crashed at the beginning of our presentation! We were quickly able to dial up a local Gopher server at the University of Maryland and connect to the LTERNET Gopher to continue the demonstration. Fortunately the network recovered in time for the Mosaic and KHOROS demonstrations! *-- James Brunt, Sevilleta LTER, jbrunt@lternet.edu*

Overt Activities of a Network Agent

Ed. Note: The following article arrived in an unmarked envelope during the night. After decoding with a Captain Planet decoder ring, it turned out to be the work of Harvey Chinn, the Network Special Agent:) Assuming a favorable outcome for the revised budget for 1994, I will continue working part-time under Caroline Bledsoe, LTER Research Coordinator. My efforts will focus on several topics that extend or broaden the work I have already done. Much of this will occur in cooperation with the Network Office staff and the data managers.

1) Investigate software systems to enhance productivity, collaboration, and synthesis by scientists, students, and managers.

Systems to be explored include Collage, Envision, GopherMOO, integrated tool suites, Khoros, Mosaic, and Mud/MOO. We expect to develop a series of test cases of ecological synthesis involving LTER scientists, students, and, yes, real data.

Another aspect will be on-going study of the data format and exchange issue with Rudolf Nottrott, which we expect to lead to a number of practical tests involving various groups of the data managers.

2) Phase 2 of the All-Site Bibliography includes working out the details of obtaining updates on a timely basis and consideration of Network standards for the format and content of citations in the site bibliographies.

In addition, I will continue to improve the filter programs as needed

3) Investigate methods for achieving distributed, on-line databases.

This will prepare for Phase 3 of the All-Site Bibliography; if satisfactory a method is available, we will try implementing it for the bibliography, again, with the assistance of one or more data managers.

4) Investigate and test methods for linking the various related on-line databases, namely the Bibliography, the Core Data Set Catalog, and the Personnel Directory. This would likely be for a hypertext system, such a Mosaic.

As always, I am pleased to receive assistance in the form of (pointers to) programs, techniques, or

advice or volunteers to assist in the various tests.

I look forward to a second exciting year working with LTER to improve the usefulness of its networked resources. *--Harvey Chinn, harvey@lternet.edu*

Merging onto the Electronic Superhighway

What is an electronic superhighway

"What is an electronic superhighway?" Despite extensive news coverage of pending legislative initiatives, that question is seldom asked, and never fully answered. Given the fundamental changes that building a national network could have, some confusion and ambiguity is inevitable. We find ourselves in a position similar to the director of transportation for New York City -- at the time where horseless carriages first started to show up! We might correctly anticipate a drop-off in watering troughs and street sweepers, but fuel injection, traffic jams, smog, and parking lots might slip by our planning process! Nonetheless, the technical elements which make the superhighway possible are well enough defined that we can make some positive statements about how the highway will work.

The key element that makes an "electronic superhighway" work is that sounds, pictures and mail all can be reduced to a digital form. Once in a digital form they can be transported by a common network. Thus a single network connection can feed information to your phone, your television set, your personal computer and your FAX machine. More importantly, the digital link can operate in both directions so that each home becomes an information provider, as well as an information consumer.

How does it really work? Start by converting sounds and images into a digital form. A simple way to do it is to break sounds or pictures into small pieces. For each piece, assign a number that describes how loud (or bright) that piece is. The next step is to "package" the series of numbers so that it can be transmitted over a network. The "envelope" that your data is packed into needs to have several pieces of information on it. First it needs to say where the data came from (like a return address), secondly it needs to say what type of data is stored in this particular packet and finally there needs to be an address to which the data is being sent. Once the information is packaged, its ready to be sent across the network. Notice that all the network needs to be able to deal with is the "envelope" -- there is no need for it to interpret the data itself. By analogy, the post office doesn't need to read or understand your mail in order to deliver it!

The processes used in the proposed network will be more complicated than described here and different than those used on the Internet as well. Some sort of compression of data will be needed to reduce the amount of data that needs to be transmitted. For some types of data, where they must be transmitted in real time, such as video and sounds, virtual circuits may be needed. Also a single mail message or terminal session may be broken up into multiple "packets," so the network also needs sequencing and error-correction data included in the "envelope." However, the basic rules of digitization and transmission given in the simplified example still hold.

The next step needed for an electronic superhighway is speed. As FAX machines prove, you can send pictures even across a regular phone line. However, a phone line is much too slow to transmit a movie. To do that, you need to be able to send approximately 24 frames per second. Even the fastest FAX machine takes almost a minute per frame -- too slow for even the most patient video addict!

The speed of transmission is described as "bandwidth" which is described as bits per second. A standard phone line has a bandwidth of 26,000 bits per second (26 KB/S). A very slow network connection is 56 KB/S, whereas a faster one (such as a "T1" phone line) is 1500 KB/S. A really fast long-distance line, such as the "T3" links that form the backbone of the Internet are 43,000 KB/S. A standard local area network (LAN) can handle 10,000 KB/S, while a fast one can achieve 100,000 KB/S. How fast will the electronic superhighway be? Like conventional highways, speed will vary depending on whether you are on the "interstate" or "secondary roads." The backbone of an electronic superhighway will be FAST -- starting at 1,000,000 KB/S. The "exit" leading to most homes, however, will be much slower, somewhere around the speed of the "T1" line. Universities and commercial businesses will have service at "T3" speeds.

Who will own the highway? Well, its a good bet that, since there are no proposals for extensive federal funding, all the roads will have electronic toll-booths! One big change is that, since there is no difference between the way different types of digital data are transmitted, there is no need for separate phone or video providers. Many of the challenges yet to be overcome on developing the superhighway are problems in the transition from specialized service providers to general service providers.

Test driving the electronic superhighway.

The plans for the electronic superhighway are still being drafted, but there are still ways to "test drive" the electronic superhighway by checking out some of the secondary roads and the "Route 66" of electronic highways, the Internet. Taking advantage of interactive information services provided by existing commercial services such as America On-line, Compuserve or Prodigy is one form of test driving. These services are analogous to a shopping plaza. You travel down a highway to reach them, but they are, in essence, centralized. All the data is stored on a single computer system (although for some services, the system itself is made up of computers linked by networks). Nonetheless, they provide a small sampling of the potential services that will be available.

Gopher, World Wide Web and Wide Area Information Servers are available on the Internet and give a glimpse of the future of interactive information services. Unlike the commercial services, these systems are highly distributed, where each item in a menu may be provided by a different computer, sometimes separated by several time zones. Gopher systems provide one of the closest analogies to the sorts digital information that will be available, but not necessarily the format or organization. Using Gopher-based information systems is like attending a flea market, the information you want is probably out there somewhere, but you have to rummage through lots of boxes to find it! On the electronic superhighway, the success of future commercial information services will depend on their ability to organize and

present information, as much as on the sorts of information they have.

Some other (dull) ways to check out the what the "superhighway" will do are to watch TV or make a long-distance phone call. Both these forms of data will be carried over the network. If you really want a glimpse of the future, make a long-distance phone call while you are watching TV, and then pretend that the face of the person you are talking to is on the screen. Imagine what sorts of displays you'll be able to generate for your callers when you aren't properly attired to take their call. How about grafting your face onto Arnold Schwartznegger's body with a digital drawing program!

How will the "real" electronic superhighway be different than the "test drives" experienced using current networks? One area where you can expect improvement is in user interfaces. You can expect to see more "point and click" and menu-based interfaces so that computer literacy is not a prerequisite for accessing information. Additionally, the types of information coming to you will be quite different. Currently, most home access to national networks is via modem. The relatively low speeds of modems restrict transfers to text and simple sounds. However, at faster transfer speeds pictures, complex sounds and even animation all become possible.

Despite the tremendous technical changes, the biggest single difference between today and tomorrow will be where information originates. Current information systems are geared towards a one-way flow, from a centralized information source to the user. However, the superhighway will have lanes running both directions! Although few of us are likely to set up major commercial information nodes, it is certain that we will all contribute to the information available over the network.

Avoiding Potholes!

I've talked about the advantages of the proposed electronic superhighway. Fast communication, fingertip access to information on national and international scales and integration of existing information sources. However, the promise of the electronic superhighway is balanced by some dangers as well. Progress towards an electronic superhighway will involve careful steering to avoid the "potholes."

The first "pothole" will be making the network universally available at reasonable prices. To do otherwise is to divide the nation into information "haves" and "have nots" and diminish the value of the network for everyone. For example, how useful would you find mail if only a few people had a mailbox, or a telephone if most of your friends had none? A restricted network could greatly act to widen the gap between rich and poor, especially when employment opportunities may depend on network access, either for locating employment opportunities or for using the network for "telecommuting." The gap would be especially devastating for schools, where lack of network access would be as debilitating, in a relative sense, as a lack of textbooks.

There are two basic approaches being proposed to avoid the "access" pothole: competition and regulation. Proposals to stimulate competition between service vendors and keep prices to a practical minimum are on the table. However, although competition between companies "laying the pavement"

may keep prices down, it is unlikely to provide universal access. This is because it is more profitable to build the high volume "main roads" between major metropolitan areas than the "back roads" to smaller towns and rural areas. In a strictly competitive environment, network access would be far from universal. Regulation will be needed to assure universal access by requiring network providers to go into markets and to regulate costs in markets that have only a single provider. The evolving network may take familiar forms, such as having single, regulated, local providers coupled to multiple, unregulated long-distance providers. However this "phone company" model is by no means the only one available. For one thing, many communities have existing telephone, cable-TV and electrical power networks. All of these can be adapted to serve as information carriers permitting competition at even a local level.

A second pothole is that networks may work TOO well! The electronic equivalent of junk mail could swamp electronic mail services. The Internet, the best existing precursor of the electronic superhighway is restricted to research and educational uses so advertising is not such a problem. However, a strictly enforced system of assigning priorities to messages needs to be used to keep a few ambitious mass marketers from swamping us!

"Laying the pavement" for the electronic superhighway will pose many challenges. Current communications laws dictate an ornate web of specialized providers. However, creation of an electronic superhighway demands that most of these distinctions be eliminated. Subtle distinctions in new legislation could give an overwhelming advantage to specific companies, guaranteeing that lobbying for such distinctions will be fierce. Obtaining a system that is both practical and equitable will be a formidable challenge for our politicians!

A condition for using the network for financial transactions or confidential messages is that it must be unquestionably secure. One approach to security is through encryption, where each transmission is scrambled at the source and descrambled by the right recipient. Interestingly, there is a major debate about how good scrambling algorithms should be allowed to be! Private companies want to use state-of-the-art "unbreakable" encryption schemes, however law enforcement agencies want less advanced schemes so that "wiretaps" would still be possible (given the appropriate court order!).

Policies controlling the ownership of information and how users will be charged for it will have a great effect on the form of the network. Imagine the situation where a company spends millions of dollars on computers to give timely stock market information, a single subscriber downloads that information and then resells it to everyone else. Alternatively, imagine a situation where a subscriber to a stock market service does extensive analyses on information obtained from the service, but is unable to distribute the results because the data provider prohibits redistribution of the data or products based on the data. Rules to control the redistribution of information need to be carefully structured. If too strict, they inhibit the exchange of information for which the network was designed. If too loose, they eliminate the incentive for information providers.

One of the major advantages of the proposed superhighway is that everyone can be both an information provider and an information user. However, a final "pothole" to be avoided is protecting you from

providing too much information! Detailed logs of your activities on the network could leave you with very little personal privacy. Two-way television monitoring was a critical element of control in Orwell's "1984," and the information coming from your house (as you pay bills, make purchases, order prescriptions, watch programs and call friends) could make it very easy for a "Big Brother" to know all about you! Orwell's "Big Brother" needed human observers to do the spying -- now computers armed with programs similar to the profiles used to spot tax frauds or locate potential customers, could allow even a small number of people could exert tremendous control over the lives of millions of people. Laws prohibiting detailed monitoring, coupled with eternal vigilance, are the only way to avoid this "pothole."
-- John Porter, Virginia Coast Reserve LTER, JPorter@lternet.edu

Publication Information