01001100 01010100 01000101 01010010

LTER DataBits

Information Management Newsletter of
The Long Term Ecological Research Network

01001100 01010100 01000101 01010010

Feature Articles

News Bits

Calendar

DataBits: An electronic newsletter for Information Managers. ----- Fall 1999
http://www.lternet.edu/documents/Newsletters/DataBits/99fall.html

Featured in this issue: an overview of the recently funded NSF/KDI Biodiversity partnership plans in addition to a variety of approaches to metadata, conferencing, visualization and training. DataBits continues as a semi-annual electronic publication of the Long Term Ecological Research Network. It is designed to provide a timely, online resource for research information managers and to incorporate rotating co-editorship. Availability is through web browsing as well as hardcopy output. LTER mail list IMplus will receive DataBits publication notification. Others may subscribe by sending email to majordomo@lternet.edu with two lines "subscribe databits" and "end" as the message body. To communicate suggestions, articles, and/or interest in co-editing, send email to databits-ed@lternet.edu.
----- Co-editors: Karen Baker (PAL) and Denise Steigerwald (MCM)

Feature Articles

A Knowledge Network for Biocomplexity: Building and Evaluating a Metadata-based Framework for Integrating Heterogeneous Scientific Data

- Matthew B. Jones, National Center for Ecological Analysis and Synthesis (NCEAS)
with text from KDI proposal by SAndelman, JBrunt, JHelly, MJones, OJReichman, MSchildhauer, RWaide, MWillig

Catalyzed by societal concerns, and facilitated by technology advances, scientists focused on complex ecological systems have generated an explosion of ecological and environmental data. When integrated with data from other disciplines (e.g., meteorology), these data have the potential to greatly enhance understanding of biocomplexity. However, broad-scale and synthetic research is stymied because these data are largely inaccessible due to their spatial dispersion, extreme structural and semantic heterogeneity, and complexity.

Knowledge Networking
In our recently successful proposal to the Knowledge and Distributed Intelligence (KDI) program at the National Science Foundation, our consortium (NCEAS, LTER, SDSC, and Texas Tech) has proposed to integrate the distributed and heterogeneous information sources required for the development and testing of theory in ecology and its sister fields into a standards-based, open architecture, knowledge network. This network will extend recent advances in metadata representation to provide conceptually sophisticated access to integrated data products drawn from distributed, autonomous data repositories. In addition, the knowledge network will include advanced tools for exploring complex data sets from which multiple formulations of hypotheses can be tested.

The existence of such a network will lead to broadened understanding of biocomplexity and ecological systems, and allow the application of that understanding to societal issues. In developing this network, we will create a new community of environmental scientists who will be able to focus on complex, multi-scale problems that, to date, have proven to be intractable. We will perform foundational research in computer science and informatics to create new tools for discovering, retrieving, interpreting, integrating, and analyzing data from these diverse sources. Our prototype network will be useful across a variety of disciplines and will provide a basis for the growth of multidisciplinary research groups focused on biocomplexity.

Technology

The infrastructure for this network has been divided into three principal components: data, information, and knowledge. The "data" component, which will be lead by the LTER Network Office, encompasses a physical infrastructure that we have proposed will link over 30 distributed, autonomous field sites to simplify the discovery and exchange of data. We are exploring the Storage Request Broker (SRB), developed at the San Diego Supercomputer Center, for achieving interoperability among different storage infrastructures at the sites.

The "information" component, which will be lead by NCEAS, is comprised of a set of tools for structuring and interpreting metadata, thereby allowing more sophisticated handling of the data resources found across the sites. This metadata framework will consist of a metadata exchange syntax (XML), a set of interoperable metadata catalogs at the various sites, and a centralized metadata catalog that harvests metadata from the sites to provide more efficient search and retrieval of metadata. Finally, we will be extending this metadata framework to include semantic information about the content of data sets that will allow us to partially automate the process of integration of heterogeneous data sets.

The "knowledge" component will consist of an advanced set of tools, to be developed at SDSC, for automating the process of modeling scientific hypotheses. This framework will build upon the facilities in the previous two components, and will provide scientists with a mechanism for expressing scientific hypotheses in a structured language that will ultimately lead to the evaluation of the hypothesis using data from the network.

Impact

The results of the proposed research will have broad implications for our ability to understand and manage sustainably the complex ecological systems and biological resources on which all humans depend. Information on biocomplexity is voluminous and complex, but currently is inaccessible to research scientists and policy makers. The intellectual advances in information science that we propose will, for the first time, provide an accessible infrastructure for identifying, integrating, managing, and, ultimately, synthesizing the nation's ecological and biodiversity information resources.

Project Status

Further information will be made available on the web as the project progresses at: http://www.nceas.ucsb.edu/kdi

Background reading

-Biron, P. V. and A Malhotra. 1999. XML schema part 2: datatypes. World Wide Web Consortium Working Draft 06 May, 1999. Available at: http://www.w3.org/1999/05/06-xmlschema-2.

-Brickley D., R.V. Guha, eds. 1999. Resource Description Framework (RDF) Schema Specification. W3C Proposed Recommendation 03 March 1999. Available online at http://www.w3.org/TR/PR-rdf-schema.

-Daniel, R. Jr., C. Lagoze, and S. D. Payette. 1998. A metadata architecture for digital libraries. Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries. IEEE Computer Society, Los Alamitos, California. pp. 276-288.

-Kent, R. L. 1999. RDF Schema in OML/CKML. Available at http://asimov.eecs.wsu.edu/CKRMI/RDF-Schema.html.

-Lassila, O. 1998. Web Metadata: A matter of semantics. IEEE Internet Computing 2:30-37.

-Lassila, O. and R. R. Swick. 1999. Resource Description Framework (RDF) Model and Syntax Specification: W3C Recommendation 22 February 1999. World Wide Web Consortium. Available at http://www.w3.org/TR/REC-rdf-syntax.

-Morgenstern, M. 1997. Metadata for Heterogeneous Databases. In: The Second IEEE Metadata Conference, Silver Springs MD. Available at http://www.llnl.gov/liv_comp/metadata/md97.html.

-Sheth, A., and J. Larson. 1990. Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases. ACM Computing Surveys 22: 183-236.

-Srivastava, J. 1999. Warehouse creation— a potential roadblock to data warehousing. IEEE Transactions on Knowledge and Data Engineering. Volume 11(1): 118-126.

Extensible Markup Language and Data Management Across the Internet

- Peter McCartney, Central Arizona - Phoenix LTER

Extensible Markup Language (XML) has often been described as the successor to HyperText Markup Language (HTML), the language of the World Wide Web. What is different about XML from HTML, however, far outdistances what is similar. In HTML, there are a finite number of tags. In XML, users of the language define their own tags and attributes. The list of tags can be limitless and custom-tailored to the specific application. In HTML, the information contained in the tags describes

only how the text is to be rendered on the screen - it conveys no information about the content, or meaning, of the text. In contrast, XML tags are used exclusively to describe the content, or structure of the data. An XML document cannot be rendered for display by a browser as can HTML - it requires the addition of a style sheet to provide display instructions for each element type in the document. XML also enhances the value of documents by embedding metadata within them. Metadata tags enable automated processing of web documents through the use of a common language that is abstracted from any hardware or software environment.

As with HTML, XML documents can be created and tagged entirely with an ASCII text editor or through the use of an editor designed specifically to handle XML documents such as Microsoft's XML Notepad. More importantly, XML documents can also be created as output from other programs such as a database application. Mainstream software support for easy creation of XML documents is still months away, but press releases indicate that products like MS Office will include native XML support in one form or another. For generating XML from existing data sources, customized solutions may be needed to accommodate the hardware and software environments at these sources. The reward, however, is that the result is a data format that is independent of those environments and has firm industry commitment for support. Tools for parsing XML are already readily available and expanded XML support is expected in the next crop of Internet browsers and development languages.

The role of XML in information management includes the following:

- **Long-term access to primary data.** Passing of information from site databases to NET including bibliographic data, personnel data, site description data and even monitoring data has been based on structured ASCII files and or web based document management tools have been implemented. XML can provide a universal language for defining content standards for data exchange files that takes advantage of extensive industry support and development. Recent enhancements to the XML standard such as "data islands" permit embedding XML formatted content inside conventional HTML documents, thus leveraging the value of existing web distribution systems. For example, basic citation information might be embedded in an online document to facilitate easy indexing in an online bibliographic catalog.

- **Identification and documentation of data.** Creation and publication of metadata to document data archives is essential to provide a means to locate, identify and make use of archived datasets. XML is well suited for defining and implementing metadata standards.

- **Mining our gray literature.** Much of the knowledge about our data products (and often the primary data themselves) is buried in text form inside countless reports, theses, and other manuscripts with limited distribution. Even when they have been made available in digital form, the information content of these documents is not readily apparent nor easily searched. The value of such documents would be enhanced through the use of XML to tag their content so that the information can be more efficiently accessed.

- **Data comparability and scalability.** Primary data are typically very contextual, reflecting the specific research questions posed by individual researchers for a given subject. As synthetic

research addresses broader issues at increasingly larger scales, there is a need for solutions to facilitate the integration of primary datasets that may differ in units, scales, or precision. Resource Definition Framework (RDF) is an XML based language being developed for encoding machine-readable metadata. This capability builds a foundation for future software that might enact upon such metadata to automate some the tedious steps required to make data from two or more sources compatible for analysis.

In summary, XML provides a non-proprietary way to transfer information across the internet and by virtue of extensive industry support, can be used to produce standardized output that can be used in quite different applications such as display, data exchange, and search engines.

IBM Data Explorer (DX) Open Source Project

-Tony Fountain, SDSC Liaison to the LTER Network

A powerful visualization tool is now available for FREE. IBM DX is a general-purpose visualization toolkit that includes utilities for manipulating, transforming, and animating various types of data and images. It supports a number of visualization operations including color mapping, volume rendering, contours and isosurfaces, as well as 3-dimensional plotting. It runs under a variety of operating systems, including AIX, Irix, Linux, Solaris, and Windows 95/NT. One of the most distinctive characteristics of DX is the Visual Program Editor (VPE) with its point-and-click interface. By using the VPE, analysts can explore and analyze data and create images and animations without looking at lines of code. The VPE features a collection of predefined modules, each of which corresponds to a processing task. These modules are linked via the point-and-click interface to create applications. The VPE also provides tools for creating web pages based on visualizations, including interactive Java-based applications. Additional information on IBM DX project, including tutorials, can be found at http://www.research.ibm.com/dx/.

On May 18, 1999 IBM announced that it was discontinuing its support of DX as a program product and shortly thereafter announced the release of the DX source code to the public. In order to capitalize on this opportunity, an organization was formed to coordinate and promote the continued development and distribution of DX. The open source version of IBM's Data Explorer is named 'OpenDX'. Additional information, including an image gallery and the downloadable software, is located at http://www.opendx.org.

Researchers at the San Diego Supercomputer Center (SDSC) have used the commercial version of DX for a number of projects and the release of OpenDX has generated new enthusiasm. Researchers are now free to explore DX's visualization capabilities without the financial burden of previous licensing agreements. The lack of licensing fees also means that application deployment and distribution is simplified. SDSC has created a web page as a resource for OpenDX information and examples: http://www.sdsc.edu/dx. One specific SDSC project that is exploring the use of OpenDX is the collaboration between SDSC and the LTER Network Office. Several prototype

visualizations of LTER data have been created to explore the utility and performance of remote visualization tools. A sample of these applications can be viewed at http://www.sdsc.edu/sdsc-lter.

SDSC has no immediate plans to offer training in OpenDX, however classes may be organized in the future as interest grows. In the meantime, visualization enthusiasts are encouraged to investigate the upcoming course on AVS 5. AVS 5 is similar to IBM DX in both functionality and programming model. More information about the NPACI-SDSC course on AVS 5 can be found at http://vis.sdsc.edu/events/workshops/avs5/index.html.

Network Conferencing Technologies: Opportunities for LTER

-John H. Porter, Virginia Coast Reserve LTER and James Brunt, LTER Network Office

Scientists within the LTER Network engage in a wide array of collaborative activities. Many of these activities are supported by the LTER Network Office in the form of meetings, electronic news groups and conference calls. However, a new range of network technologies which take advantage of high-speed Internet connections (such as those provided by the vBNS), are coming of age and can be used to extend the level of collaboration between LTER researchers working at different sites while reducing the need for face-to-face meetings.

Collaboration technologies allow the following types of interactions:

- Audio - The Internet acts like a toll-free telephone line. Typically there are slightly longer delays than experienced on normal telephone lines.
- Video - Live video can be transmitted across the Internet. This can vary from small, low resolution, images to full-screen, high-resolution video, depending upon the hardware used and the network bandwidth available.
- Application Sharing - The screen for an application running on one computer can be displayed on the screens of other computers. Optionally, control of the program can be passed among collaborators.
- Whiteboards - Whiteboards are a graphical surface that can be drawn on by collaborators using a variety of graphical editing tools (pens, straight lines, boxes, polygons, text) similar to those found in most painting programs (e.g., Microsoft Paint). In addition, graphics can be "pasted" in from other programs. The results are shared across virtual meeting participants.
- Chat - Messages typed at one computer appear in the chat window all meeting participants or selected collaborators. This can be used to augment audio communications (exchanging notes while another is speaking) or to replace audio communications when network resources will not support audio.
- File Transfers - Files can be rapidly and transparently sent to meeting participants by a simple command.

There are several ways of using collaboration technologies and the types of equipment, software

and network bandwidth requirements vary between them. These are:

- Desktop Conferencing - communications are between individual computer desktops, where each participant has their own collaboration workstation.
- Group Conferencing - communications between groups where individuals within each group share a collaboration workstation.
- Point-to-point - A one-on-one conversation between individuals or groups involving only two collaboration workstations.
- Multipoint - A conversation between three or more individuals or groups, each with their own collaboration workstation.

Point-to-Point Desktop Conferencing Point-to-point desktop conferencing requires the lowest level of hardware, software and network resources. A minimal configuration supporting audio, video, application sharing, whiteboard and chat capabilities would be a 486 or above PC workstation connected to a 28.8 modem, running Microsoft NetMeeting with a sound card, microphone and inexpensive (<$80) video camera. However, the University of Virginia recommends:

- a PC with a Pentium 200 (or faster) MHz MMX CPU
- 48 Mbytes memory
- Video Display with 2Mbytes video memory
- Sound card
- Video capture card
- Camera (with connector that matches video capture card)
- Speakers and microphone OR headset

Microsoft NetMeeting software has the advantage that it supports all of the major collaboration capabilities and is available free of charge. However there is a variety of commercial software (e.g., Netspeak Webphone, Intel Videophone) which support audio and video communications. However most of these products are restricted to point-to-point conferencing for at least their audio and video capabilities.

Group Conferencing Doing group conferencing requires a step up in terms of peripheral devices such as displays, cameras and microphones (those used for desktop conferencing are typically too small or crude to be used for group communication). This type of conferencing also requires much higher network bandwidth to support larger, faster video displays. UVA recommends the following:

- A multimedia-ready Windows PC Pentium II 300 MHz CPU (or better) 128 Kb RAM
- Wireless keyboard and mouse (radio preferred over infrared)
- VCON Armada Escort 25 Pro system ($700) for video (includes video card, camera, software, headset)
- Switched network connection

Multipoint Conferencing Multipoint conferencing requires two major differences from point-to-point conferencing. First, the conferencing software must support multiple simultaneous connections of

audio and video. Microsoft NetMeeting (the first choice for point-to-point desktop conferencing software) does not support multiple simultaneous audio and video connections, although it does support multiple application sharing, chat and whiteboard connections. With the aid of JAVA-based WWW tools, it can be used in conjunction with a multipoint control unit to participate in group sessions, although only one individual can be seen at a time. CUSEEME software supports multiple audio and video connections without the need for external (e.g., JAVA) programs. Second, you need to have a multipoint control unit or reflector that provides the linkages between multiple sites. Multipoint control units and reflectors vary widely in their cost and capabilities. Software for support of multiple groups can cost $8,000-$10,000 and supporting multiple groups (with high-resolution video connections) can require dedicated use of a UNIX workstation costing $12,000. However, there are some public-domain software solutions that may meet needs for multipoint desktop conferencing using existing UNIX servers.

Recent Activities In conjunction with Jim Beach (KSU/KDI) and John Porter (VCR), NET has been making some efforts at testing point-to-point and multipoint desktop conferencing solutions. Indy Burke (SGS), Bob Waide (NET) and Eda Melendez (LUQ) have also participated in test sessions. Microsoft NetMeeting has been used on almost a daily basis between KSU and VCR. Experience over a range of network traffic conditions supports the following conclusions: 1) a telephone always has less of a delay than does the teleconferencing software, 2) despite the delay (which ranges from a fraction of a second to several seconds) it is quite possible to communicate and work on a project, especially if you are patient, 3) network dropouts can occur, causing disruption of audio. It is therefore important to have a reliable network end-to-end. 4) video, although not essential to audio communication, does improve the quality of communications and 5) whiteboard and application-sharing technologies can be very useful for communicating information or demonstrating software. The following table gives a letter grade for the different applications that NetMeeting supports:

Audio	C- it works, but can be very choppy or unreliable. Always has delay.
Video	C+ performance similar to audio, but our expectations are less.
Application Sharing	A- very useful for demos or sharing graphical information
File Transfer	B+ nice and fast. Puts files in special directories.
Whiteboard	B works well, but less useful than application sharing
Chat	B workable, but not spectacular

University of California Addresses Technical Training

- Karen Baker, Palmer LTER

The University of California (UC) has proactively addressed the issue of providing technical training

to a wide group of faculty, staff and students in a single unified but cost-effective manner. It is difficult to supply a wide variety of technical training in computers and software packages due to individual seminar costs, to retraining requirements resulting from personnel turn-over as well as to constant new product shifts. Early in 1998 UC developed a cross-campus co-operative effort with CBT Systems, (http://www.cbtsys.com/catalog/catalog.htm), a major provider of interactive software for training in information systems.

UC has invested in web-based training with CBT in order to provide an alternative opportunity for staff education in computer courses ranging from introductions for beginners to technical courses for certification. CBT was chosen because it has a large number and wide range of types and levels of training modules as well as a standard friendly user interface. In addition the price, site license, and technical support are good. Although available at all UC campuses except for UCSF (ie UCSD: http://www-cbt.ucsd.edu, UCD: http://cbt.ucdavis.edu), it is handled separately by each location so both presentation and module availability varies by campus.

Currently, the training is available via web access from a PC. Availability on CD also exists while availability for MAC and UNIX systems is planned.Topics covered currently include Microsoft General, Office 97, Information Technology Core Concepts, Internet and Intranet Skills, Cisco, IBM, INFORMIX, Java, Lotus, Marimba, Netscape, Novell, Oracle, SAP, Sybase, C/C++, Internet Security, Internetworking, UNIX, and Centura with a range of 4 hour courses within each topic. The course modules for any one topic are structured for end-user, application designers, and/or system administrators. So for instance, the Microsoft modules available for the end-user, including old and new versions of software, include a section on Microsoft Applications with modules on Office 95, Office 97, Windows NT 4.0, Exchange 4.0, Exchange 5.0 and Outlook 97.

This is a multi-year pilot effort provided through co-operation between academic and administrative computing, telecommunications, human resources,staff education, medical centers and a variety of individual departments. If the approach proves useful, the campus programs may expand to include additional modules available through CBT.

News Bits

LTER Metadata Standards Committee

- Peter McCartney, Central Arizona - Phoenix LTER and Don Henshaw, H.J. Andrews LTER

LTER has long recognized the importance of metadata to data survival. The recent NSF KDI award to NCEAS, LTER-NET, and SDSC will usher in changes in the role of metadata from textual documentation to more structured information that can support future applications that automate much of the data management and access process. To prepare the LTER network for these new developments, a Metadata Standards

committee was formed at the 1999 IM meeting to review existing content standards for ecological metadata and to make recommendations for upgrading site metadata to meet the structure and machine-accessibility requirements of the KDI project, NIS expansion, and other cross-cutting data access developments. Members of the committee are P. McCartney - CAP (leader), J. Brunt - NET, J. Campbell- HBR, D. Henshaw - AND, J. Porter - VCR, D. Steigerwald - MCM, M. Jones - NCEAS.

In addition to reviewing content standards, the committee will review software options available to sites for preparing structured metadata. XML-based solutions will be examined closely and the committee will develop some demonstration applications of this technology to enhance awareness of this technology. It is anticipated that conversion of existing metadata to structured formats will represent a significant workload for some sites. The committee will draft a proposal for future technology supplements to mitigate these costs which will be forwarded by the LTER IM Committee to the CC Committee.

LTER Site Information System Elements Survey

- *Karen S. Baker, Palmer LTER*

The LTER information managers decided at their recent summer meeting to gather information about existing site information system designs in preparation for upcoming KDI activities. Within the short timeframe available at the meeting, a survey was designed and the majority of representatives completed the two page form. These forms have been transcribed and will be available online in November for correction and update. Under discussion are the organization of input forms and structure of the database.

Ecological Technology Trail: A Concept for the ASM2000

- *Chris Wasser, Shortgrass Steppe LTER*

As part of the All Scientists Meeting prior to the 2000 ESA meetings, the LTER Information Managers have developed a concept for hands-on demonstrations of cutting edge tools and technologies associated with the ecological sciences. The target audience for this "workshop" would be both scientists and informatics professionals, but the focus will be on tools and technologies that can be implemented today to support ecological science. We envision this session to include both an indoor computer-based demonstration area and an outdoor trail of relevant technologies for ecology.

The indoor section would be focused on 4 or 5 kiosks demonstrating software and hardware tools such as: computer visualizations, voice recognition software, video conferencing, GIS/GPS applications (possibly connected to outdoor sites), and examples of VBNS usages. Professionals from the LTER network and commercial vendors of these products would staff these kiosks. Complementing the indoor program would be a short walk with multiple stations for discussions and demonstrations of field-based tools. We would imagine these stations would include tools such as: GPS units, microloggers, wireless technology (perhaps connecting

visitors to demos inside), weather station equipment, and others. Professionals from the LTER network and commercial vendors of these products would staff these stations.

Data Management Collaboration: LTER sites BES and CAP

- Peter McCartney, Central Arizona - Phoenix LTER and Sam Walker, Baltimore Ecosystem Study LTER

In Summer 1999 two LTER sites, Baltimore Ecosystem Study (BES) and Central Arizona - Phoenix Collaboration (CAP) organized two workshops on data management. One held May 3 at BES was attended by visitors from other LTER sites (Brian Kloeppel, CWT; Peter McCartney, CAP; Russel Watkins, CAP) and included topics 'virtual sampling' and 'patch structure'. The second workshop (http://caplter.asu.edu/intranet/dm/workshop99.htm) was held June 24-25, 1999 at CAP with Sam Walker (BES) and Chris Daniel (BES) as guests. Both workshops provided opportunity for participation from general site scientists as well as more focused discussion among data management staff. While each workshop addressed local issues, the similarities between CAP and BES provided much fertile ground for input from visitors. Some of the topics covered included problems in urban remote sensing classification, sampling design and database issues related to both site's 200 point monitoring programs, and general data management procedures. The opportunity for intense and focussed collaboration between data management led to greatly-broadened minds and two tangible products. The first was a draft proposal for a collaborative research project on developing classification standards for urban remote sensing imagery (A follow-up meeting will take place in September 1999 at BES). The second was substantial input that greatly enhanced an ASU proposal to NSF Biological Databases and Informatics program. The general response from these workshops was that they provided a valuable format for cross-site communication and generating collaborative research. Based on the success of these meetings, McCartney and Walker recommended to the 1999 IM committee meeting that other sites might consider similar cross-site visitation as one solution for enhancing intersite communication.

Upcoming BES Data Management Exchanges

- Sam Walker, Baltimore Ecosystem Study LTER

BES plans for September and October 1999 include continued focus on data management and cross-site collaboration. Following earlier data management exchanges, a visit to BES is planned by CAP scientists Russell Watkins, Michael Ramsey, and Will Stefanov in late September to discuss remote sensing research opportunities between the two cities. Plans are to outline a workshop and develop some standardized classification procedures. The goal is to promote the effective comparison of urban landscapes through remote sensing technology. Cross-site collaborations also will take place with Ned Gardiner, Coweeta Data Manager, during an October 6th visit to BES. The annual BES site meeting is scheduled for October 14th and 15th. An agenda, directions and other information can be found at http://baltimore.umbc.edu/lter. The Data Management Team is scheduled to present a workshop on metadata development and creation on October 14th from 4-5 pm.

Palmer and McMurdo Education Outreach

- Karen S. Baker, Palmer LTER

The Palmer and McMurdo LTER sites are participants in the Teachers Experiencing Antarctica and Arctic (TEA) program supported by NSF. The TEA program, a joint collaboration between Office of Polar Programs and the Division of Elementary, Secondary and Informal Education makes it possible for teachers to travel to the field with polar research teams. Program goals include providing teachers with a research experience as a part of their continued professional development and taking polar experience back to their classrooms. The McMurdo LTER works with teacher Barb Schulz who deployed several years ago and returns to the field again this November. Teachers Besse Dawson and Mimi Wallace associated with the Palmer LTER have field participation January 1998 and January 2000, respectively. Such collaborations depend in many ways upon technology and are a unique opportunity to foster technology transfer. Thus, support from and coordination with information management can be an important aspect of education outreach efforts. To facilitate communications, each LTER site has a designated education contact (http://www.lternet.edu/oppts/education/contacts.html).

FAQ: What is Z39.50?

- Peter McCartney, Central Arizona - Phoenix LTER

Z39.50 is an information retrieval protocol that supports communication among different information systems. When we think of sharing information from multiple databases, two kinds of solutions immediately spring to mind - (1) compiling one central database that contains a copy of everything in the contributing databases or (2) using software that can simultaneously query multiple databases and return a single, combined result, without requiring the user to know the detailed query syntax for each of those databases. Bibliographic specialists developed the Z39.50 search engine to solve the latter approach to query multiple library holdings (http://lcweb.loc.gov/z3950/). The system consists of a Z server which provides a universal connection to each contributing database. A profile defines the mappings between the database fields and a set of standardized fields. In the case of library databases, this standard is based on the USMARC bibliographic format. A Z client application provides an interface through which users define a boolean query using the standardized fields. The query is submitted to multiple Z servers which parse the query and execute it using the local system's syntax. Results are then received by the search engine where they are compiled and returned to the user. Recent projects by the Federal Geographic Data Committee, University of Kansas Natural History Museum, and the San Diego Supercomputing Center have demonstrated the power in extending the use of Z39.50 to other categories of data such as metadata catalogs (http://www.fgdc.gov), biological collections (chipotle.nhm.ukans.edu/documentation/applications/CollectionAttributes.htm), and systematics databases (chipotle.nhm.ukans.edu/documentation/applications/SpeciesAnalyst/). It is likely that Z39.50, or a similar technology, will represent the ultimate solution for sharing data from the LTER sites.

Good Read: Federated Database Vocabulary

- Karen S. Baker, Palmer LTER

As we come to terms with networked data systems within the LTER, there is a need for a broadened stock of words to capture and describe the changing conceptual vistas of information management. Sheth and Larson (ACM Computing Surveys 22:183-236) help the community into the 1990's by presenting a pertinent vocabulary with which to address "Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases". A range of topics are discussed from semantic heterogeneity to types of federation and architecture schemas for cooperation among independent sites/systems.

Noted: Ecologist Outreach at Spokane ESA Meeting

- Karen S. Baker, Palmer LTER

The Spokane-Review Newspaper (7August99) carried the story about the ESA annual conference being held for the first time in Spokane this year. It was reported that last year an ESA member was queried "So you're here for that conference, those environmentalists?" and the ESA member had responded with "No! We're scientists!" This interchange alludes to a scientific objectivity expected of ecologists. In fact, the ESA literature defines "ecology is the study of the relationships between living organisms - including human - and their physical environment". It is interesting to note that after this exchange, specifically as an outreach for the Spokane meeting, the society invited nonscientists from the host city to participate as well as discounted education meeting costs for educators. This can be noted as a proactive response to a communications need.

Calendar

*99 Sep 16-19 OBFS Meeting; Mountain Lake Biological Station, VA
*99 Sep 30-Oct 03 CC Meeting; Hubbard Brook
*99 Oct KDI LTER/NCEAS/SDSC/KU begins
*99 Nov 03 FGDC Biological Working Group
*99 Nov 09-10 Biological Scale Process Modelling Workshop II; SD, CA
*99 Nov 19-21 LTER Education Workshop; Kellogg Station, MI
*00 Aug 02-04 All Sci Meeting; Snowbird, Utah
*00 Aug 01-04 LTER IM Meeting; Snowbird, Utah

LTER Information Manager Meeting 04-06Aug99

The recent LTER Data Manager Meeting in Spokane had a full agenda developed at the Spring DataTask Meeting. Meeting organization was by Chris Wasser and the meeting committee. Section leaders came with presentations "in hand". Two recommendations will be forwarded to the LTER CC meeting: 1)Future technology supplements be made available specifically for developing structured metadata; 2)The standing committee name of 'Data Manager Committee' (DM) be updated to 'Information Manager Committee' (IM) and the DataTask Executive Committee name be updated to 'IM Exec'.

LTER Education Workshop 19-21Nov99

The LTER Education Workshop is being organized by the LTER Education Chair Diane Ebert-May and Network Office Staff Patty Sprott. Applications to attend the upcoming meeting were due 15 Sep. Two-person LTER site teams will assemble at Kellogg Station, MI for this proposal writing workshop.