

LTER Information Managers Report to the Coordinating Committee

Prepared by IMEXEC

15 April 2002

1. NIS (Network Information System). Since its formation in 1988, the LTER IM (Information Managers) Committee has worked to facilitate intersite research through the application of new techniques in computer science and the emerging field of ecoinformatics. Over the last 10 years, improved Internet connectivity at LTER sites and the advent of the World Wide Web have revolutionized the way in which LTER data and metadata are made available to other scientists and the general public. Though most LTER data sets were available on individual site web pages by the late 1990s, it was still difficult to analyze and synthesize data from different sites because most site servers return data in different formats. An integrated Network Information System was proposed by the LTER Information Managers in 1996 as a solution that would provide centralized access to important datasets from across the LTER Network. Since 1996 the NIS has gradually evolved from the planning stage, to a series of prototypes developed at various LTER sites, to near-production-level implementation of some of its modules. At present, administrative modules include personnel, site information (SiteDB), publications, and data table of contents (DTC); while scientific modules include climate (ClimDB) and net primary productivity. The current plan is to add scientific modules to the NIS from each of the LTER CC scientific theme meetings. Note the important distinction in terms of maintenance between modules that provide one-time "snapshots" of scientific data and modules that are regularly updated.

While the successes of modules like ClimDB (see below) and DTC show progress toward NIS goals, the LTER IM community continues to grapple with the challenge of how to balance priorities between (1) providing and supporting data products using current technologies, and (2) researching and developing new technologies that offer more advanced solutions for dynamically integrating data. The IM Committee recognizes a commitment to both approaches. On the one hand, there has been steady progress toward making parts of the NIS available as a production service, though progress has been limited by lack of long-term funding. On the other hand, the IM Committee continues to work closely with partnering institutions such as NCEAS (National Center for Ecological Analysis and Synthesis) and SDSC (San Diego Supercomputer Center) to introduce new technologies to the LTER Network. Two long-term goals that will affect the NIS are (1) the application of ontologies for guiding semantic integration of data and (2) the adoption of web services technologies that simplify the publishing of data and applications on the network. Both of these figure prominently in recent funding proposals by LTER or partnering institutions.

2. EML (Ecological Metadata Language). In 1999, the LTER IM Committee established a subcommittee to seek a standardized, machine-readable format for metadata across the network. This committee has collaborated with the KNB project (NCEAS, NET, SDSC) and the Networking our Research Legacy project (ASU) to develop the EML 2.0 specification, which is now nearing completion. The first of several planned workshops to migrate LTER sites to compliance with the new standard was held at ASU in January 2002. All but two site IMs were able to participate. Each site joined at least one of three hands-on groups that focused on building EML connectors to existing relational database catalogs, translating existing text metadata into EML, or entering new EML metadata using KNB's Morpho tool. All sites reported having a clearer plan for adopting EML after the workshop and two follow-up workshops are planned for summer 2002, after which the entire LTER Network is expected to be able to provide EML formatted metadata. Compliance with EML is a critical first step toward future implementation of advanced informatics being developed by a community of partners that includes the Network Office, several individual LTER sites, NCEAS, and SDSC.

3. ClimDB. ClimDB was the first NIS module to contain scientific data. The original prototype was developed at NTL with strong support from the LTER Climate Committee. Over the past year ClimDB was moved to AND, where it was considerably improved and a companion system for hydrological data (HydroDB) was developed with support from the U.S. Forest Service. ClimDB is a centralized database that makes meteorological data dynamically available over the Internet in standard downloadable and graphical formats. Recent enhancements include greater control over the harvesting process by individual sites, better quality assurance checking, and the ability to handle a much larger number of meteorological variables. At present 18 LTER sites are participating in ClimDB. Over the

coming year, a collaborative project between LTER and SDSC will prototype a web services implementation of ClimDB that will utilize EML and may serve as a model for future NIS modules. For more information on ClimDB, please see: <http://www.fsl.orst.edu/climdb>.

4. IM Mentoring Web Page. New LTER Information Managers face a steep learning curve. To aid newcomers, an IM Mentoring web page is being created on the LTER Intranet, and relevant resources are being identified or newly developed for this purpose. Organizing categories for the web page will include: (1) *Site IM Responsibilities*, including metadata standards, software options, and the Information Manager's role in proposal renewal and midterm review, (2) *Intersite IM Responsibilities*, including how to participate in administrative databases (SiteDB, DTOC, SiteBIB) and research databases (ClimDB, NIS modules for CC Science themes, etc.), (3) *LTER Data Policies* regarding data access and acknowledgment, (4) *Communications in the IM Community*, including resources available through NET, mail lists, discussion forums, and publications such as DataBits and Network News, (5) *Training Opportunities*, (6) *Important References and LTER IM Milestones*, and (7) *Funding Opportunities*.

5. SCI2002 Meeting. In an effort to promote cross-domain exchange as well as to identify a venue for publications, the next annual LTER Information Managers meeting will precede the SCI2002 6th World Multi-Conference on Systematics, Cybernetics and Informatics, scheduled for July 14-18, 2002, in Orlando, Florida. Two LTER sessions entitled *The Ecoinformatics Challenge: Meeting Ecological Information Needs for the Site, Network, and Community*, co-chaired by John Porter, Karen Baker and Susan Stafford, will result in twelve reviewed manuscripts in the conference proceedings. For more information on SCI2002, please see <http://www.iis.org/sci2002>.