

2006 Information Management Committee Meeting Overview

The annual LTER Information Managers Meeting was held September 18-20, 2006 in Estes Park, Colorado and was attended by representatives from 26 sites, 15 representatives from ILTER sites, Co-Directors and Staff from the LTER Network Office, and IT partners ([link to participants](#)). The agenda included updates on current in- and outreach projects (items 1-5 below), working group reports and discussions (items 6-11), updates on participation in LTER NIS activities (items 12-13) and election of new IMExec members (item 14), as well as a visit with Henry Gholz. The first evening, representatives from the ILTER sites presented slides of their research sites and explained what challenges they have encountered building and maintaining information management systems within their organizations. Special emphasis was given in the meeting to create opportunities for exchanging tools, information and lessons, and building partnerships.

1. Databits. The Fall 2006 issue of the LTER Information Managers newsletter DataBits is available online at <http://lternet.edu/databits>. The newsletter is designed to engage the LTER IM community with a rotating editorship and authorship. [Brian Riordan](#) BNZ is editor and [John Campbell](#) (HBR) is co-editor and will become editor for the spring 07 issue.

2. IM Mentoring Website and IM Committee Website. A content management system for these websites is already in place. Working group leaders and/or members will be granted the authority by LNO webmasters to manage the content on these websites through a web interface. (contact [Marshall White](#), LNO)

3. IM Meeting Survey. Following the 2005 Annual LTER IM Meeting, IMExec created a survey to assess the success of our annual meetings. Questions focused on logistical support and the venue, participation from LNO staff, time allotted for working groups and communication with fellow IMs, and presentations of network-wide projects, goals, and broader impacts. Results from the 2005 survey helped IMExec shape activities for the 2006 meeting, including a poster and demonstration session, which proved to be very successful in 2006 and facilitated one on one communication between IMs. (contact [Jonathan Walsh](#), BES)

4. Poster and Demonstration Session. The following posters and demos were discussed.

Demonstrations (presenting author: title)

- Dave Balsiger (NTL): Cyberdashboard for sensor network management
- Matt Jones (NCEAS): SEEK, The Science Environment for Ecological Knowledge (SEEK)
- Wade Sheldon (GCE): Software tools for automated synthesis of LTER, USGS and NOAA climate and hydrologic data
- Jonathan Walsh (BES): Using 'Google Earth' to display LTER data
- John Porter (VCR), Chi-Wen Hsiao, Meei-Ru Jeng, Chau-Chin Lin (Taiwan): EML/statistical transformations

2006 LTER IMC Meeting Report, 3/3/2010

- Theresa Valentine (AND): Demo of the WatershedDB interactive website.
- Theresa Valentine (AND): Demo of a tool to create EML from ArcCatalog metadata.
- Suzanne Remillard, Don Henshaw and Theresa Valentine (AND): Applications of FLOW: enabling downloads of small watershed data, a zoomable-interactive Andrews research timemap, and GLITCH: General Linear Integrator for Time Changes
- Judy Cushing and Lee Zeman (Evergreen): The Canopy Databank Project
- Mark Servilla and Inigo San Gil (LNO): Trends

Posters (presenting author: title)

- Matt Jones (NCEAS): SEEK
- John Porter (VCR)
- Wade Sheldon (GES): "Software tools for automated synthesis of LTER, USGS and NOAA climate and hydrologic data"
- CCE LTER: Information Management (2004-2006), CCE
- Palmer LTER: Designing a Queriable Community Data System, PAL
- Research in Infrastructure Studies: Social and Organizational Perspectives on Ecological Data Management (PAL, CCE)
- Theresa Valentine (AND) --- WatershedDB
- Margaret O'Brien (SBC) --- publications list to EML using literature.xsd
- Suzanne Remillard, Don Henshaw and Theresa Valentine (AND): A metadata-driven information management system at the Andrews Forest LTER
- Ken Ramsey (JRN): The Jornada Basin LTER Information Management System: An XML Driven Website and Content Management System
- Judy Cushing and Lee Zeman (Evergreen): The Canopy Database Project
- Meei-Ru Jeng: A Metadata-Based Framework for Multi-Language Ecological Information Management in East-Asia Pacific International Long-Term Ecological Research
- Inigo San Gil (LNO): Climate Cooling Revisited
- Nicole Kaplan (SGS): LTER IM Articulation Work: Developing Community Web Recommendations
- Eda C. Melendez-Colom (LUQ): Incorporating Information Management into the Luquillo LTER Schoolyard Program

5. All-Scientists Meeting 2006. IMs across the community were encouraged to submit proposals for working groups in collaboration with PIs. These efforts proved successful and cultivated many interactions between PIs and IMs (see list below). In addition, several post-ASM proposals were coauthored by IMs and PIs.

- **Ancillary Day, Sept 20, 2006**
 - John Porter (VCR) and Nicole Kaplan (SGS): Design of Query Interface to LTER Data Catalog: Controlled Vocabularies Working Group
 - Christine Laney and Ken Ramsey (JRN): Information management issues and the LTER Trends project: A drawing board for making cross-site comparisons feasible
 - [Barbara Benson](#) (NTL), [James Brunt](#) (LNO), [John Porter](#) (VCR), and [John Vande Castle](#) (LNO): Planning Cyberinfrastructure for LTER Network Synthetic Science
- **ASM, Sept 21-23, 2006**
 - [Theresa Valentine](#) (AND) and [Jonathan Walsh](#) (BES): High performance geographical information systems for synthesis across ecological research sites
 - [John Porter](#) (VCR), [Peter Arzberger](#), [Barbara Benson](#) (NTL) [Paul Hanson](#) (NTL) [Chau Chin Lin](#) - INT, [Timothy Kratz](#) (NTL): Exploring the Effect of Scale-dependent Processes on Ecological Systems using Networked Sensors
 - [Kenneth Chiu](#) and [Barbara Benson](#) (NTL): Automatic Scaling of Sensor Networks at NTL
 - [Judy Cushing](#), [Nicole Kaplan](#) (SGS), [Judith Kruger](#), [Ken Ramsey](#) (JRN), and [Kristin Vanderbilt](#) (SEV): Cross-Site Synthesis of ANPP Grasslands Research - Challenges in Data Integration

- [Julia Jones](#) (AND), [Judy Cushing](#), [William Michener](#) (LNO), and [Kristin Vanderbilt](#) (SEV): Ecosystem Informatics Education and Research
- [Hen-biau King](#) (INT), [Barbara Benson](#) (NTL), [Chau Chin Lin](#) (INT), [Michael Mirtl](#) (INT), and Kristin Vanderbilt (SEV): International Information exchange, sharing, and cooperation activities
- [Susan Stafford](#) and [Peter Arzberger](#): Strategic and tactical ways to advance Mentoring Opportunities for the IMs in alignment with the strategic future directions of the LTER Network?

6. Web Site Design. Don Henshaw will be presenting web site design recommendations to the Exec Board for review and approval. This working document was vetted through IMC, NISAC and all LTER PIs. This working group developed a set of recommendations for individual site web pages, especially for new sites and for sites redesigning their existing pages. The primary goals are to improve access to site information and to emphasize membership in the LTER Network. (contact: [Nicole Kaplan](#), SGS).

7. ILTER Activities and Partnerships. IMs and IM representatives from ILTER sites were asked to share and discuss aspects of successful partnerships. Working groups reported a few important components that influenced the success of international partnerships and projects. Having face-to-face communication, at least to initiate collaboration, was deemed extremely important, as was the personal commitment of individuals involved. Successful collaborations occur between individuals, not fixed institutions or governments. Having a strong commitment from all parties and a clear set of goals and performance expectations for the collaboration was important. Availability of funding was also identified as a catalyst for successful collaboration. (contact: [Kristin Vanderbilt](#), SEV)

8. Controlled Vocabularies. This working group explored the use of hierarchical controlled vocabularies to facilitate browsing and searching of LTER datasets. They discussed how keywords are used in data discovery, different ways to develop keywords, how to test and evaluate existing resources for keywords, and how to relate keywords to the unit dictionary and attribute ontologies that are being developed within the community. The working group discussed next steps in terms of the evaluation of existing resources, such as GCMD, NBII, GEMET and WORDNET and tools for exploiting them to aid in LTER data discovery. Four suggestions included to enable auditing on Metacat to track requests so that we can see how researchers are attempting to use the existing system, help prepare the LTER IM group to deal with the complexities of thesauri and ontologies, continue to examine what others have done and to relate that work to our own efforts, and work with the SEEK Knowledge Representation group to represent attributes in an ontology template. (contact: [John Porter](#), VCR).

9. Trends. The information management issues as related to LTER Trends project and making cross-site comparisons feasible were discussed. There were challenges in finding long-term data sets, in creating ways to document the details and steps of data derivations, tracking data usage, and providing proper citations for data usage. Issues discussed by breakout groups included how Trends can be used by ILTER sites, and whether it might be useful to think about an international Trends effort. ILTER

representatives discussed the pros and cons of participating in TRENDS or developing an ILTER TRENDS to make comparisons on a global level. Participating IMs provided comments on the process of documenting and submitting data and metadata, and their satisfaction regarding the Trends project. The time commitments from IMs and PIs were discussed. It would be interesting to survey the sites on their feelings about the Trends project and for input regarding the process. It was suggested to create a sub-committee of the Trends Editorial Committee to address issues of derived data ownership; look towards NCEAS as an example. (contact: [Ken Ramsey](#), JRN)

10. Cyberinfrastructure Planning. The CI Strategic Plan developed as part of the LTER Network planning grant, was distributed for comments from the IM community by Barbara Benson. Benson and James Brunt asked the IMs to assess budget estimates for development of site level cyberinfrastructure at a level of minimum standard functionality and a list of priorities for site resources as well as to discuss what is missing in the CI plan that reflects functionality at the site level. John Vande Castle followed these discussions with an assessment of current status of functionality as reported in the latest CI survey and discussed redrafting the assessment survey to ensure that minimum standard functionality was addressed.

For new data collection, the advantages and disadvantages of using common database schemas, developing standards, providing training resources and other ways of reducing arbitrary heterogeneity from data acquisition through output were discussed. However, we would need to differentiate between local needs and necessary investments to conduct site specific science versus standardized collection. Another strategy included encouraging partnerships between sites and larger entities to develop successful research proposals. Larger institutions, even beyond LTER may provide the framework in which individual projects can prosper. (contact: [Barbara Benson](#), NTL)

11. LTER Cyberinfrastructure Assessment. From the spring 2006 CI survey, it was determined that everyone is managing information differently and we don't know why. More information is needed on collaborations, Roles/Functions/ Technology, Users of Data, Sensors, and Future Trends. It was suggested to include questions about what site capability you are most proud of, and what management methods are used for various types of data (e.g. Site, GIS, and proprietary/restricted/sensitive data). The next survey should also be linked closely to the CI strategic plan and IM review criteria (contact: [John Vande Castle](#), LNO)

12. EML Implementation & Harvesting. About half the sites are creating level4+ EML, and these sites seem comfortable with the process and are interested in moving ahead. For various reasons, about half the sites are still making discovery-level EML or less. Inigo's work with individual sites has been very valuable, but these are short-term, site-specific solutions. Even a site currently making high level EML will eventually update its IM system, so the long-term goal should be common implementations and shared database schemas (as is possible) -- recognizing the effort and expense involved. Successful utilization of EML in synthetic research is related to EML Quality Control as QC standards have not been created, and currently, parsers do not yield reports or even

consistent results. There is also a lot of interest in tools which make use of EML, or use-case activities which will engage scientists and create support (e.g., writing R scripts for Kepler). Some tools exist at a grass-roots level (and at various levels of development), which could be shared or encouraged (with resources). Again, communication is key to learning about activities or establishing common interests at different sites. Finally there are still shortfalls of EML/Metacat as a general data management tool. Sites generally have not standardized on EML for their internal applications and information systems – result is duplication of effort. Appears to be not yet appropriate for real-time data systems, or for tracking access of data (not metadata). (contact: [Inigo San Gil](#), LNO).

A list of EML projects for the next year was planned in working groups that focused on implementation of EML in these contexts: Creation, Quality Control, Use, and Management. As sites progress toward richer EML and tools are developed, we need to address additional recommendations for level 4&5 EML and reduce “workarounds”. Many sites would like to see more recommendations for handling units, in a unit dictionary either as a separate document or included in EML Best Practices. EML 2.0.2 (or 2.1?) needs bugs fixed and some enhancements in a timely manner. The EML Best Practices document will need to be improved and updated as EML implementation and utilization becomes more prevalent across sites and applications. (contact: [Margaret O’Brien](#), SBC)

13. ClimDB and HydroDB, and other “all site” databases. Coverage of precipitation, discharge, and air temperature data is strong across sites. We encourage sites to contribute relative humidity, soil temperature, wind speed & direction, and global radiation in datasets. The committee discussed briefly how the scientists would like to expand ClimDB. (contact: [Don Henshaw](#) AND). Other “all site” databases are maintained by LNO. Recent changes to the management of these databases include reporting participation of the all-site bibliography on the web site and planning to build web services into the personnel database, which will be followed by a request for comments period from the community. (contact: [James Brunt](#) LNO).

14. IMExec Committee. Regular members for the coming year include: Corinna Gries CAP (co-chair, IMC), Nicole Kaplan SGS (co-chair, IMC), Eda Melendez-Colom LUQ, Emery Boose HFR, Jonathan Walsh BES, Todd Ackerman NWT, Margaret O’Brien SBC, and John Campbell HBR. Ex officio members include: Barbara Benson NTL, James Brunt LNO, and Don Henshaw AND.