Report to the Executive Board
April 30 2009

Collaborative Development of a Project Database for LTER Sites
Margaret O'Brien (SBC) and Corinna Gries (CAP)

BACKGROUND: Experiments or projects are central to all research endeavors, and LTER sites usually encompass many. Research projects are commonly documented at the site-level in order to maintain records for annual reporting, field site permissions, coordination of research activities in space and time, and for linking datasets and other products to individuals. In a centralized database at the network level, however, well-documented research projects with a common search interface provide a wealth of information on past and present research activities and opportunities for cross-site collaboration. The goal of this working group was to build such a database and provide user friendly access for management by individual sites and search across all sites. The database was established at the Network office and interfaces built using an architecture which allows individual LTER sites to access information without further programming effort. The network-adopted schema (EML) was used as the basis for the database storage and for information exchange via web service.

IMPLEMENTATION: The workshop was conducted in two parts: an initial 2-day workshop was held to outline specifications and needs of individual sites and the network, and a longer workshop (5 days) in which code would be created. It was not necessary that the participants of the two workshops be the same, although significant overlap was planned.

Workshop 1: November 17-19, 2008
Attended in person at the network office by: Margaret O'Brien (SBC), Corinna Gries (CAP), Jonathan Walsh (BES), Ken Ramsey (JRN), Suzanne Remillard (AND), Kristin Vanderbilt (SEV); and via Polycom VTC: Wade Sheldon (GCE), John Porter (VCR), Sven Bohm (KBS), Jason Downing (BNZ).

Workshop 2: April 20-24, 2009
Attended in person at the network office by Margaret O'Brien (SBC), Corinna Gries (CAP), Jonathan Walsh (BES), Wade Sheldon (GCE), Raul Aguilar (CAP); and via VTC by John Porter (VCR) and Sven Bohm (KBS), as available. Network support was provided by James Brunt (LNO) and Mark Servilla (LNO).

PRODUCTS: A database was designed and built using an adapted version of the EML schema's project module and the open-source native XML database, "eXist" (http://eXist-db.org). This approach takes advantage of our network standard, and uses a framework which is already equipped to deliver data as a web service. The group took advantage of the participant's different skills and interests to accomplish the tasks. The following examples represent some of the uses of the database:
1. General browsing, i.e., search by site or within the entire LTER by person, time, place, lat/long or keyword.
2. Maintain records of datasets, images, publications or other material associated with a project or experiment
3. Maintain material used in reporting to funding bodies
4. Record permits associated with a project, either as the grantor or the granted
5. Export EML-project "trees" for inclusion in datasets

The database has now been installed at the LNO, and has been implemented via its web services at several sites. The design maintains individual site identity, but the common look-and-feel promotes a network identity. Approximately 25 projects have been added by the sites participating in code creation and by one "early adopter" (MCR). The following links are active for demonstration purposes. All demonstrations include the "Search all LTER Projects" form. Since the content available for searching is still limited, we have suggested terms to be used in the initial free text searches: "water, soil, plant, arthropod, temperature, salinity".

GCE: http://gce-lter.marsci.uga.edu/public/research/project_search.htm
CAP: http://caplter.asu.edu/home/projects/WStest/projectSearch.jsp

SBC: http://sbc.lternet.edu/research/SBC_projectDB_demo.html
BES: http://beslter.org/frame7-page_13.html
MCR: http://mcr.lternet.edu/Project.html

Additionally, we have created documentation for the database specifications and use cases, plus a User Guide for Information Managers accessing the web services, and recommendations for best practices when creating database entries. Forms for editing and maintenance of project data were also initiated. The forms, guides and recommedations will be further developed as feedback is received from sites. All material is available on the LTER IM Committee website (http://intranet.lternet.edu/im/project/LTERProjectDatabase).

The flexibility of the infrastructure means that additional features are possible, such as displaying projects' sites on a map, linking to research projects from an individual's entry in a personnel database, exporting to other markup languages (e.g., GML) or outputting a printed document. All modern web-programming languages are capable of using XML as input, so individual sites are likely to find other uses for this database. Collaboration among ten sites and the LNO provided us with a wide variety of use cases, experiences and expertise. The mode of operation in which a motivated group dedicates a week to writing code for a common need was efficient (although grueling), and resulted in a faster output than the model in which information managers meet only occasionally from their home institutions or at an annual meeting. The development group plans to introduce the rest of the IMC to the database at the ASM later this year.