Central Arizona Phoenix Long-Term Ecological Research

<ecological informatics at CAP LTER>

Peter McCartney, Corinna Gries, Matthew Luck
Center for Environmental Studies

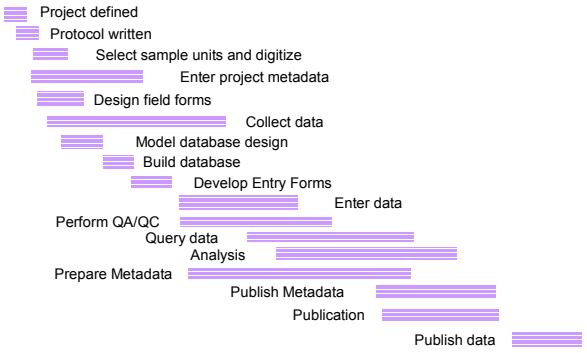
Arizona State University

<abstract>

If we envision environmental information from a network perspective, we see LTER sites as nodes interacting within a cross-cut fabric of networks including other departments and projects within the institution, regional partners, other LTER sites, and so on. A variety of tools and solutions have been used at CAP to create a system through which information are generated, managed and disseminated within the CAP LTER research cycle. More advanced methods are now being applied to better integrate the products of CAP research with those of other collaborators within and beyond ASU. These efforts and the collaborative projects through which they are being carried out illustrate some of the network connections that integrate only data, but also new partnerships for research.

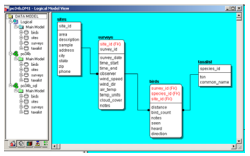
<flow of data through the research cycle>

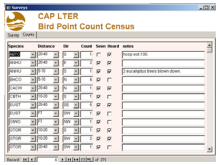
While every research project has its own unique characteristics, the creation of ecological data follows a typical trajectory within the project. The key to successful data management is recognizing the important steps in this process and their dependencies. As at most LTER sites, the idealized sequence depicted here for CAP LTER datasets should take approximately 2-3 years.



<quality control through data design>

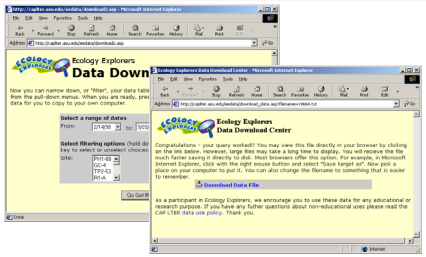
Databases are designed with entity/relationship modeling software using the written protocol and sample field collections forms as guides. Quality control is enforced at the database server through rigid enforcement relational constraints and domain rules.





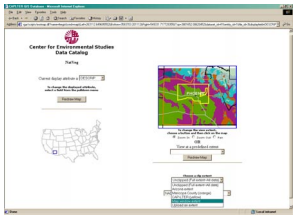
Consistent, re-usable design templates for data entry applications facilitates the use of common interface aids such as pull-down lists, editing masks, event scripts and reports for proofing sheets.

<data access>



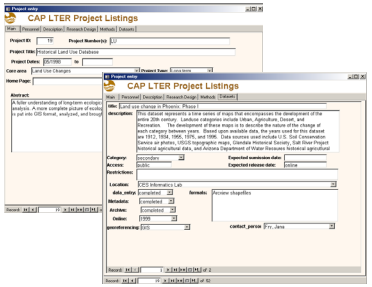
CAP LTER data are stored in relational databases. To make querying data easier for researchers and other users such as K-12 participants, stored views are created to hide complicated table joins. Web forms simplify the query selection process.

Advanced systems for publishing data should do more than provide access – they should assist users in navigating, visualizing, querying and analyzing data. The key to all of this is **metadata**. In the current CAP LTER data catalog, server applications read metadata to open and display selected data, allowing to the user to browse, subset, and re-project data prior to downloading.



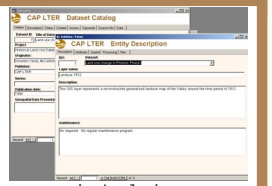
<project definition>

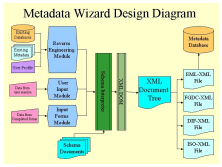
Data management begins with the definition of a new research activity. For each dataset expected to be created or acquired by that project, a record is generated in a tracking table where its progress may be monitored and reported via the web site.



<metadata>

Metadata generally falls into two categories –the semantic meaning of the data and the syntax of its organization. The latter can be mined from the metadata created by the data modeling software used, eliminating much typing. The meaning of the data, however, must be provided by the researchers.

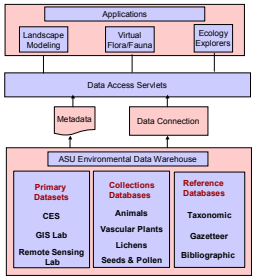


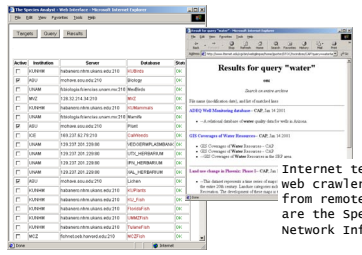


With funds from NSF, CES is developing a tool to combine automated generation of metadata from data structure with a "wizard" or "interview" type of interface for researchers to supply contextual information.

<Integration>

Under a separate grant from NSF, CES is building an infrastructure to integrate environmental data resources across ASU departments. Based on XML encoded metadata and re-useable code, the system will access data from multiple sources and support multiple kinds of applications.





Internet technologies like Z39.50 and http web crawlers enable ASU data to be accessed from remote search applications. Shown here are the Species Analyst and the LTER Network Information System