BOF: ecosystem informatics, bio-informatics (and health/epidemiology)

Moderator: Judy Cushing, Evergreen State College

10 participants:
   6 in eco, incl Community GIS
   2 in genomics
   1 epidemiology
   1 both eco/benomics

Given the strong overlap\(^1\) between research issues in our two areas\(^2\), we looked at

1. rating common issues (low, med, high risk, short, medium, long term)

2. context, or driving forces of each of these “specialties”.
   - Industry factors driving CS/ECO/MB: W3C; economic costs of invasive species, bioterrorism, health/ecosystem/disease outbreaks; open source movement; need to identify customer base for industrial partners.
   - Bio/Eco driven to CS by NBII, LTER, advances in instrumentation (Genome sequencing, sensors), having a critical mass of CS with expertise in BIO/ECO.
   - CS/IT research driven to BIO/ECO by research areas that need data or applications (data mining, dbms, computational science, spatial and time series); having a critical mass of BIO/ECO with expertise in CS.

3. Discussion of how or if to grow the partnerships that now exist among research disciplines (e.g., CS-molecular biology; CS- ecology, ecology-statistics, medical, eco (environmental health) to
   - business (a long discussion ensued on how to articulate the business value (customers) of ecology applications, e.g., partner with us on solving temporal-spatial data structures problems – it would pay off with financial applications); consider cost of public health problems like asthma and overlap between environmental health and public health.
   - agencies e.g., NPS, NFS on data management for ecology
   - Sociological and educational issues such as metadata provision and best practices compliance
   - Great divide issues between science and policy. Expertise should flow not only from science to policy but vice versa!

\(^1\) Overlap of the two areas:
   - Terminology management with ontologies, glossaries
   - Data provenance, metadata, and annotation
   - Adaptive, flexible database schemas & schema management
   - Support for spatio temporal data & other domain specific data types
   - Uncertainty management & data cleaning
   - Management of mathematical and statistical models (with data) and modeling support

\(^2\) Differences of the two areas: for ecology: sensor networks and stronger emphasis on computational. For MB: problem of propagating uncertainty in derived data products for MB. Areas might rate six major research topics differently.
how we can be sure the case is made for domain-rich projects across domains, even where there is considerable overlap in the CS research area:

• devil is in the details for these complex domains, problems will not be solved unless tested with complex data and applications
• interesting problems might arise as sub-problems in metadata management, or spatial data types, for ecology that one might not see (at least not immediately) in bioinformatics.
• The disciplines need the infrastructure, and “we” need experience not just in building the systems but in their use, and extension.