Data Availability

Data availability speeds knowledge discovery, increases the efficiency of the scientific community, and reduces the time that researchers spend looking for and synthesizing data. Enhanced data accessibility and integration are critical components of the Sea Grant program (Kelly et al. 2007) and are now being emphasized by the recently announced NSF Cyberinfrastructure initiative (NSF 2008). This new emphasis recognizes the importance of providing high-quality, interoperable, and easily usable data to the scientific research community. The BCO-DMO database is an example of a community database, which is typically hosted at a university, that has been developed and operated by a university, and is designed to support the scientific research community. The BCO-DMO database is a collaborative effort among the National Science Foundation (NSF), the University of Maine, the Woods Hole Oceanographic Institution (WHOI), and WHOI’s Institute of Oceanography (IO). The database is currently supported by the NSF’s East Coast Regional Node (ECRN) for the Global Earth Observation System of Systems (GEOSS), which is part of the NSF’s Cyberinfrastructure program.

Interoperability and Visualization

One important design specification for the new BCO-DMO data system is that components should be chosen to increase interoperability with other data systems regardless of boundaries associated with disparate research disciplines, funding agencies, or nationalities. The current data management system, using the JGOFS/GLOBEC distributed, object-oriented, relational, data management system, and access to nautical and geographic data will be provided via any standard Web browser through a GIS application (Open Source, OGC implementing MapServer). Storing metadata in a relational database permits data-set descriptions to be generated in compliance with a variety of metadata content standards and provides mechanisms for exchange of data with other repositories.

Conclusions

Robust metadata records are the key to accurate use of data. Metadata records capture the information required to answer the who, what, where, why, when, and how questions that are asked about a data set. It is important to know who collected, analyzed and contributed the data and where, when and how those data were acquired. Ideally, term dictionaries and controlled vocabularies are used to populate the metadata database fields. Access to metadata and supporting documentation aids data discovery (through development of powerful search engines) and helps to ensure efficient and accurate use of data. Data re-use is especially dependent on metadata availability.

References
