Towards Linked Data for Oceanographic Science: The R2R Eventlogger Project, Controlled Vocabularies, and Ontologies at The MBLWHOI Library

Elizabeth Coburn  
MBLWHOI Library/Woods Hole Oceanographic Institution

Andrew R. Maffei  
Woods Hole Oceanographic Institution

Cynthia L. Chandler  
Woods Hole Oceanographic Institution

Follow this and additional works at https://escholarship.umassmed.edu/escience_symposium

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.

Repository Citation

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License. 
This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Towards Linked Data for Oceanographic Science: The R2R Eventlogger Project, Controlled Vocabularies, and Ontologies at The MBLWHOI Library

Elizabeth Coburn1, Andrew R. Maffei2, Cynthia L. Chandler2, and Lisa Raymond1
1. MBLWHOI Library/Woods Hole Oceanographic Institution; 2. Woods Hole Oceanographic Institution

Objectives
Research vessels coordinated by the United States University-National Oceanographic Laboratory System (US-UNOLS) collect data which are considered important oceanographic science research products. The NSF-funded Rolling Deck to Repository (R2R) project aims to improve access to these data and diminish barriers to their use. One aspect of the R2R project has been to develop a shipboard scientific event logging system. Eventlogger, which incorporates best practice guidelines, controlled vocabularies, a cruise metadata schema, and a scientific event log. Eventlogger facilitates the eventual ingestion of datasets into oceanographic data repositories for subsequent integration and synthesis by investigators. The careful use of controlled vocabularies and ontologies is an important feature of this system, as the use of internationally-informed, consensus-driven controlled vocabularies will make datasets more interoperable, discoverable and reusable as well as facilitate linking. The R2R Eventlogger project is led by Woods Hole Oceanographic Institution (WHOI), and the management of the controlled vocabularies is led by the Data Librarian in the Marine Biological Laboratory/Woods Hole Oceanographic Institution (MBLWHOI) Library.

Methods
The first target vocabulary is one for oceanographic instruments. Management of this vocabulary has consisted thus far of reconciling project vocabulary terms with the more widely used community vocabularies served by the NEERC Vocabulary Server v2.0 (NVS2.0): terms included in the SeaDataNet Device Catalogue (L22) and the SeaDataNet Device Category vocabularies (L05). Rather than adopt existing community terms, it is more often the case that local terms are mapped by the Data Managers in the NSF-funded Biological and Chemical Oceanographic Data Management Office (BICO-DMO) to community terms, which preserves any important information and meaning investigators impart through the process of assigning these local terms, and has less impact on researchers. New terms, those that cannot be mapped to the existing community vocabularies (often custom, or modified instruments) are submitted for review to the SeaVOX governance process for addition to the community vocabularies. These vocabularies and their mappings are an important part of the aforementioned Eventlogger system. Before a research cruise, investigators configure the instruments they intend to use for their science activities. The instruments available for selection are provided by the MBLWHOI Data Librarian, who curates UNOLS ship-specific lists of standard shipboard instruments using terms from the R2R Eventlogger Project Vocabulary. Non-standard shipboard instruments a researcher or investigator wishes to use can also be added, and these instrument terms will eventually be inducted into the R2R Eventlogger Project Vocabulary.

Results
The vocabulary for oceanographic instruments is currently being vetted locally by a domain expert, and the Eventlogger application is currently being tested across the UNOLS fleet. A large submission of suggested instrument terms to the SeaDataNet community listserv will follow. New tools for facilitating the management, mapping, and use of these controlled vocabularies are being developed, and new projects with eager partners are envisioned. Ideas for future controlled vocabularies for the ocean science community include: Cruise IDS, Persons, and Ships.

Conclusions
The promotion and use of controlled vocabularies and ontologies will pave the way for linked data in oceanographic science. By mapping local terms to authoritative and community-accepted terms, links are created whereby related datasets can be better discovered, and utilized. Librarians have an established history of working with controlled vocabularies and metadata. Libraries (have and will continue to) serve as centers for information discovery as well as a natural home for the management of standards.

References

Contact
Elizabeth Coburn ecoburn@whoi.edu