University of Massachusetts and New England
Area Librarian e-Science Symposium

2013 e-Science Symposium

Apr 3rd, 12:00 AM - 1:45 PM

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

L. Raymond
MBLWHOI Library/Woods Hole Oceanographic Institution

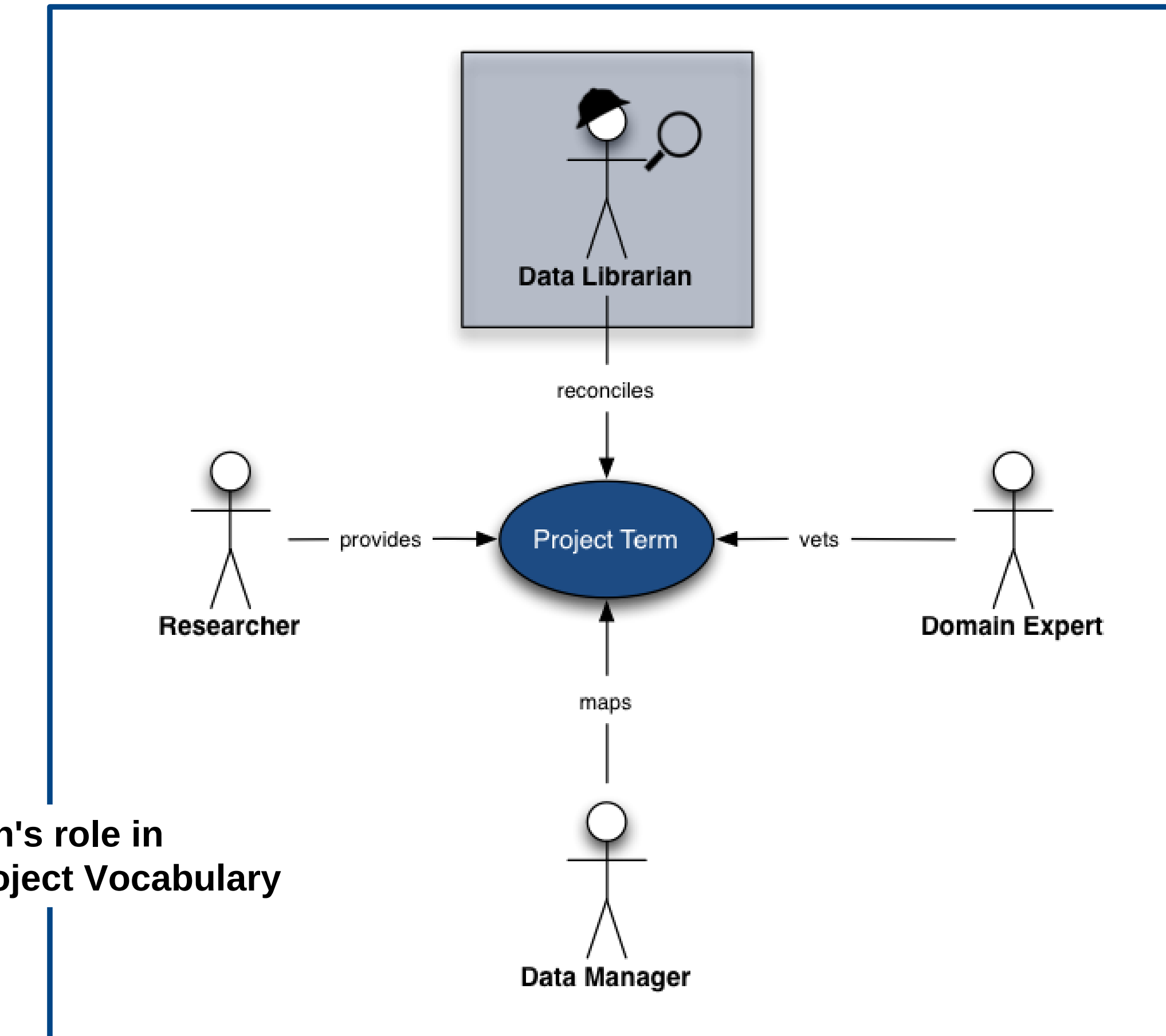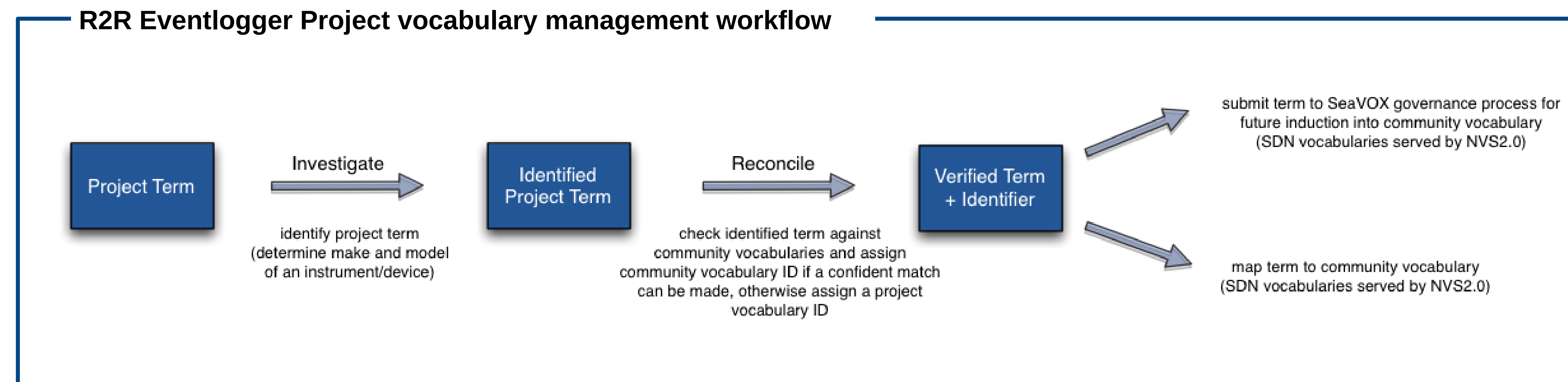Part of the Library and Information Science Commons, and the Oceanography and Atmospheric Sciences and Meteorology Commons

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

Elizabeth Coburn[1], Andrew R. Maffei[2], Cynthia L. Chandler[2], and Lisa Raymond[1]
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.

R2R Eventlogger Project vocabulary management workflow

The MBLWHOI Data Librarian's role in managing the R2R Eventlogger Project Vocabulary

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 NERC 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 (BCO-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 for instruments 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.
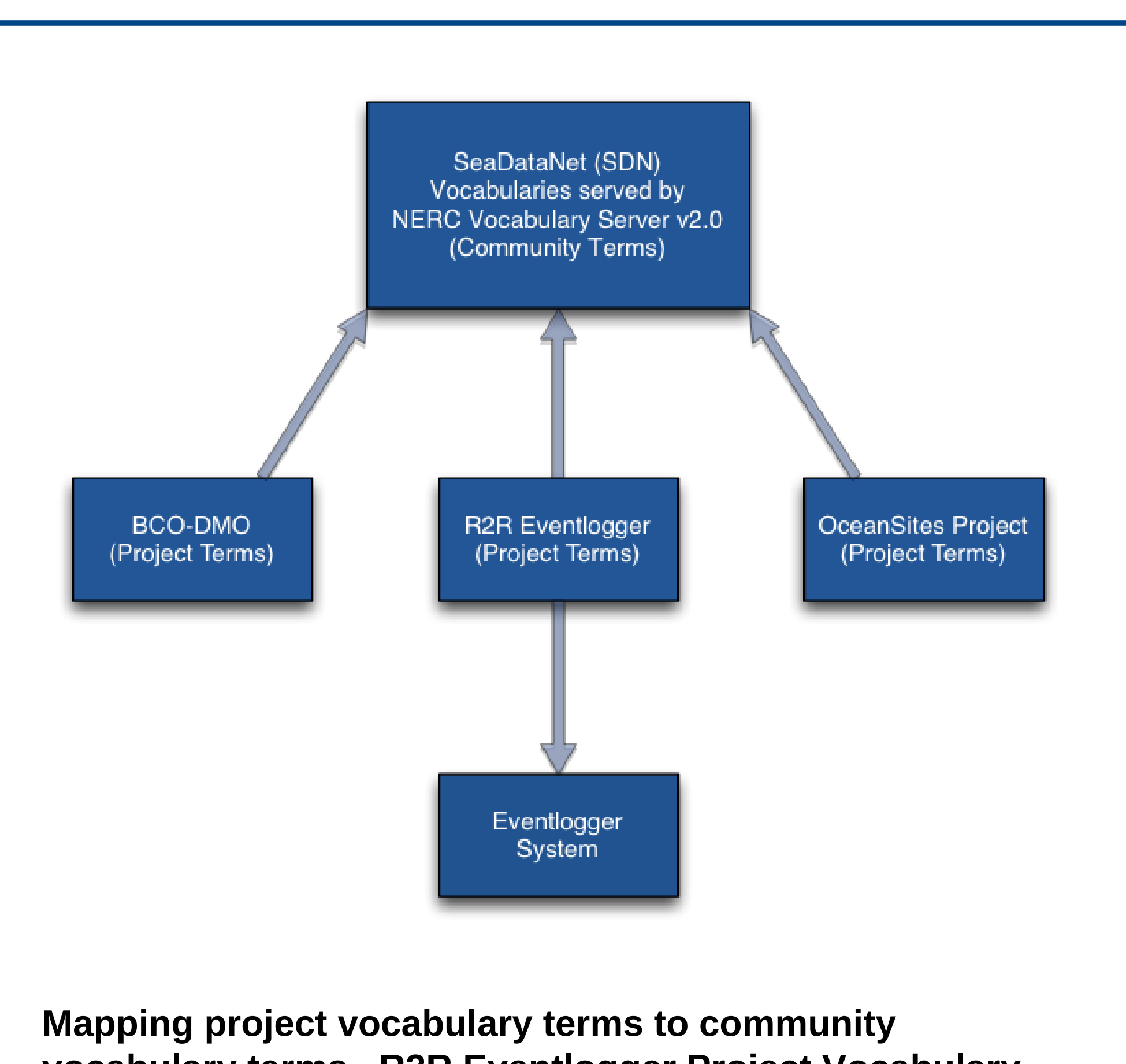
References

Maffei, Andrew R., Cynthia L. Chandler, Janet Fredericks, Nan Galbraith, Laura Stolp. Rolling Deck to Repository (R2R): A Controlled Vocabulary and Ontology Development Effort for Oceanographic Research Cruise Event Logging. EGU2011-12341. Poster presented at the 2011 EGU Meeting.

Contact
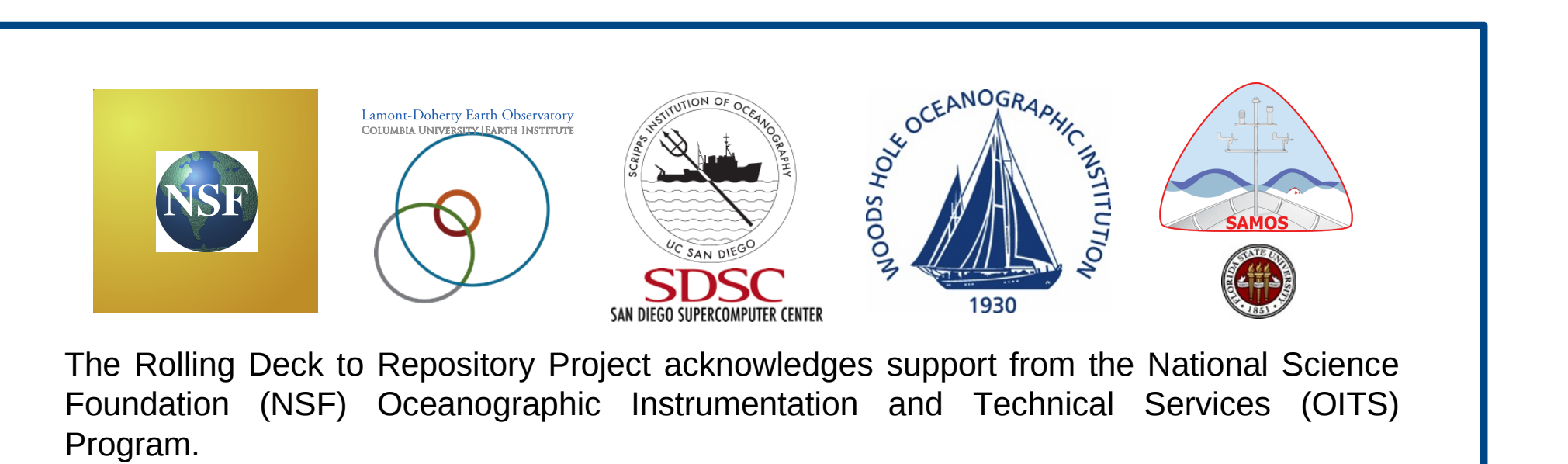Elizabeth Coburn ecoburn@whoi.edu

Mapping project vocabulary terms to community vocabulary terms. R2R Eventlogger Project Vocabulary terms are mapped up to SeaDataNet Device Category terms (L05) and SeaDataNet Device Catalogue terms (L22), both are served by the NERC Vocabulary Server v2.0 (NVS2.0). Eventlogger Project terms are used in the R2R Eventlogger System.

Currently, the Data Librarian uses Google Drive to manage R2R Eventlogger Project vocabulary terms (left). A new Drupal site, Project Vocabulary Mapper (right), is in the prototype stage, and will potentially meet the long-term vocabulary management needs of Data Managers, the Data Librarian and others, as well as foster collaboration among stakeholder institutions.