Emerging Role of Librarians in Data Publication

Lisa Raymond
*Woods Hole Oceanographic Institution*

Cyndy Chandler
*Woods Hole Oceanographic Institution*

Roy Lowry
*British Oceanographic Data Centre*

See next page for additional authors

Follow this and additional works at: [http://escholarship.umassmed.edu/escience_symposium](http://escholarship.umassmed.edu/escience_symposium)

Part of the [Library and Information Science Commons](http://escholarship.umassmed.edu/library_and_information_science_commons)

Raymond, Lisa; Chandler, Cyndy; Lowry, Roy; Urban, Ed; Moncoiffe, Gwenaelle; Pissierssens, Peter; Norton, Cathy; and Miller, Holly, "Emerging Role of Librarians in Data Publication" (2012). University of Massachusetts and New England Area Librarian e-Science Symposium. 5.

[http://escholarship.umassmed.edu/escience_symposium/2012/posters/5](http://escholarship.umassmed.edu/escience_symposium/2012/posters/5)

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.
Presenter Information
Lisa Raymond, Cyndy Chandler, Roy Lowry, Ed Urban, Gwenaelle Moncoiffe, Peter Pissierssens, Cathy Norton, and Holly Miller

Keywords
Data publication, Data citation, Data management

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.
Emerging Role of Librarians in Data Publication

Lisa Raymond¹, Cyndy Chandler¹, Roy Lowry², Ed Urban³, Gwenælle Moncoiffé², Peter Pissierssens⁴, Cathy Norton⁵, Holly Miller⁵

1. Woods Hole Oceanographic Institution (WHOI); 2. British Oceanographic Data Centre (BODC); 3. Scientific Committee on Oceanic Research (SCOR); 5. IOC Project Office of IODE; 6. Marine Biological Laboratory (MBL)

Purpose

This poster demonstrates the procedures and tools developed to deposit datasets in an Institutional Repository (IR) and assign Digital Object Identifiers (DOIs).

Description

Current literature on the topic of data publication suggests that success is best achieved when there is a partnership between scientists, data managers, and librarians. The Marine Biological Laboratory/Woods Hole Oceanographic Institution (MBL/WHOI) has developed tools and processes to automate the ingest of metadata from BCO-DMO for deposit into the Institutional Repository (IR) of the Woods Hole Open Access Server (WHOAS). The system also incorporates functionality for BCO-DMO to request a Digital Object Identifier (DOI) from the Library. This partnership allows the Library to provide data access and preservation services for datasets generated by MBL/WHOI scientists.

Method

Basic Workflow
1. A dataset is submitted by a scientist to BCO-DMO.
2. BCO-DMO staff enter metadata for the dataset in their database.
3. BCO-DMO staff store the dataset itself in their repository.
4. At some later time, a scientist contacts BCO-DMO with a request to cite a dataset.
5. BCO-DMO staff trigger a submission to WHOAS via a special submit button in BCO-DMO’s web interface.
6. BCO-DMO’s database system pulls data and metadata from its database and assembles a METS package.
7. BCO-DMO’s database system submits the METS package to WHOAS via SWOR.
8. WHOAS ingest data automatically and return a handle URL to BCO-DMO’s system.
9. BCO-DMO’s system parses the handle URL and generates a corresponding DOI.
10. BCO-DMO staff sends the DOI to the scientist who made the citation request.
11. After a short period following the submission, WHOAS notifies the BCO-DMO manager via email.
12. The WHOAS manager creates a DOI for the submitted dataset record.

Conclusions

The assignment of persistent identifiers enables accurate data citation. The Library can assign a DOI to appropriate datasets deposited in WHOAS. A primary activity is working with authors to deposit datasets associated with published articles. The DOI would ideally be assigned before submission and be included in the published paper so readers can link directly to the dataset, but DOIs are also being assigned to datasets related to articles after publication. WHOAS metadata records link the article to the datasets and the datasets to the article.

Related Websites

PublishedOceanData http://publishedoceandata.net/
Woods Hole Open Access Server https://darchive.mblwholibrary.org/
British Oceanographic Data Management Office http://www.bco-dmo.org/

Acknowledgments

Funding provided by the Jewett Foundation

1 E-Science Symposium 2012

Diagram showing the evolution of a scientific dataset through its different stages from raw data collection (Fieldwork) to the generation of master files which must be preserved from accidental loss, self-described, exchangeable, re-usable, and citable (by deposition in an accredited data center), through to the preparation of analysis data subsets for the purpose of scientific investigation and communication, data subsets which must in turn be openly accessible and citable for the purpose of traceability.