

Trends In Data and Open Access

Moderator: Susana Macanawai

DATA CITATION RATES:

GIS DATA IN THE MARINE SCIENCES AND PUBLISHER CITATION REQUIREMENTS

Kristen B. LaBonte

Life & Environmental Sciences Librarian

University of California, Santa Barbara

Santa Barbara, CA 93106-9010

klabonte@library.ucsb.edu

Abstract:

While researchers are beginning to cite datasets, the trend to cite has not been widely adopted. In papers that use Geographic Information Systems (GIS), the citation of data can be easily identified when maps display background data that is not original to the author, such as bathymetry. This study assesses marine and aquatic GIS data citation from 2013 and illustrates scholarly trends in data citation. Author requirements by publishers are investigated to see if there is a relationship established between these requirements and data citations. English language academic papers indexed within *Aquatic Sciences and Fisheries Abstracts (ASFA)* are examined. This is a follow-up of a study conducted using *Web of Science*, published by the author in the 2013 IAMS LIC *Proceedings*.

Keywords: Data citation, GIS, geographic information systems, geospatial data, spatial data, marine, aquatic, publishers & publishing, authors & publishers, data literacy.

Research papers that utilize Geographic Information Systems (GIS) to create maps present useful opportunities for illuminating data citation trends. GIS maps are created with software where layers of data combine with each other. The end result is a map within an article that can be examined to see whether and how attribution is given for the use of these data layers. In maps such as these, original data collected in the field are combined with data from other sources. These other data layers are investigated here to determine whether and how data use is attributed. In the articles reviewed in this study, these layers typically include political boundaries, topography and bathymetry. In previous research, I found that GIS sources are frequently credited within the text of an article, but are not listed within the works cited or references sections (LaBonte, 2013). The present study examines a new set of articles, builds upon the 2013 research, and links citation rates to publisher data citation requirements.

The early work of Seiber and Trumbo (1995) illuminated the issue of a lack of data citation in the social sciences. They recommended requiring that editors develop a journal policy to determine whether sources of data are or are not from the author in order to “foster accurate citation of shared data.” In 2010, Weber et al. evaluated 307 journal titles in the field of Environmental Studies and found that only 6% gave data citation instructions.

The design of the present study was for a presentation to be delivered at the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMS LIC) Conference in September 2014 in Noumea, New Caledonia. I constructed a query in *Aquatic Sciences and Fisheries Abstracts*

(ASFA) using peer-reviewed English language articles from 2013 with the keyword terms ((marine OR aquatic) and (GIS OR “geographic information system*”). The Materials & Methods and References sections of each article in the set of search results were examined to identify data attribution. In addition, I used the “Find” function in Adobe Acrobat and Google Chrome to look for the specific keywords relating to data that might be cited. These words were: *GIS*, *ArcMap*, “*spatial data*”, *data*, *figure*, *layer*, and *map*. Of the 192 articles in the search results, 164 were ultimately evaluated. Thirteen articles did not have GIS maps, three had only original data in the maps, three were in non-English languages, and nine could not be supplied through Interlibrary Loan. In one article, GIS stood for the Greenland Ice Sheet, so it was not included in the study.

Seventy% of articles examined had some form of data attribution in the text, while 20% of the articles cited the data within the references or works cited sections (Table 1). The 164 articles were published in 83 different journals; of those, eight journals (10%) had a data citation requirement in the instructions for authors: seven journals (8%) within the reference or works cited sections and one journal (1%) within the text of the article (Table 2). The eight journals with a data citation requirement had thirteen articles included in this study and eight of them attributed where the data came from.

Articles examined	164
Data attributed within text	114
Data cited in references or works cited	33

Table 1: Articles with data attribution or citation.

Unique journals represented	83
Journals with a data citation or attribution requirement	8
- Within references or works cited	7
- Within text	1

Table 2: Journals with data attribution or citation requirements in the instructions for authors.

Some publishers require that authors use a certain manuscript preparation style, and that style may or may not have requirements or examples of data citation. Sixteen journals (19%) required the use of specific style manuals, and an additional six journals (7%) accepted any style, as long as the whole paper conformed to it (Table 3). Five of the eight style manuals that were referred to specified how to cite data (Table 4). There were 19 articles examined from journals with style manual requirements that have data citation instructions. Of these, 14 (74%) had citations within the text, and 2 (10%) also had citations in the list of references.

Unique journals represented	83
Style referred to in author instructions	24
- Specific manual referenced	16
- Any style referenced	6

Table 3: Journals that referenced style in the instructions for authors.

Style Manual	Data name for citation instructions
American Chemical Society	Data Sets
American Psychological Association	Data Sets
Chicago	Scientific Databases
Council of Science Editors	Database
Ecological Society of America	None
Entomological Society of America	None
Harvard	Datasets
ISO 690/1987	None

Table 4. Style manuals linked from author guidelines and data citation instructions.

Little research has been conducted to examine citation rates of data at this time. Mooney and Newton released a study in 2012 that showed similar results to the present study. While their study examined data citation in the broadest sense, reviewing 2010 social science, science and humanities articles, they came to the conclusion that “citation of digital research data is a rarefied activity.” If data citations are listed within the Reference or Works Cited sections, they are easy to discover and can be included within citation software, such as Thomson Reuters’ *Data Citation Index*. As stated by The Future of Research Communications and e-Scholarship in the final Joint Declaration of Data Citation Principles, “data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications”.

As more funding agencies require data archiving and the growth of data repositories continues, there will be a greater likelihood of publishers requiring data citation in a specified format. Librarians have a role in educating researchers and students on data literacy topics. Open Access Week is an appropriate venue, and librarians can also create web pages that link to sites like Data Pub, which gives directions for data citation and links to standards. It is hoped that editors and peer-reviewers will become acquainted with the publisher’s data citation requirements and take responsibility to vet data citations through the peer review process.

References

- Data Pub. (2014). *Data Citation*. California Digital Library. Retrieved from <http://datapub.cdlib.org/datacitation/>.
- LaBonte, K. (2013). GIS data citation rates: Are data being properly credited in lists of references? IAMSLIC 2013. Paper presented at IAMSLIC 2013, Dania Beach, FL, USA. <https://darchive.mblwhoilib.org/handle/1912/6657>.
- Mooney, H. & Newton, M. (2012). The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship and Scholarly Communication*, 1(1), eP1035. <http://dx.doi.org/10.7710/2162-3309.1035>.
- Sieber, Joan E. & Bruce E. Trumbo (1995). (Not) Giving Credit Where Credit is Due: Citation of Data Sets. *Science and Engineering Ethics*, 1, 11-20. <http://link.springer.com/article/10.1007%2F02628694#page-1>.

The Future of Research Communications and e-Scholarship (2014). *Joint Declaration of Data Citation Principles – Final*. Retrieved from <https://www.force11.org/datacitation>.

Weber, N., Piwowar, H., & Vision, T. (2010). Evaluating data citation and sharing policies in the environmental sciences. *ASIST 2010*. Paper presented at ASIST 2010, Pittsburgh, PA, USA. https://www.asis.org/asist2010/proceedings/proceedings/ASIST_AM10/submissions/445_Final_Submission.pdf.