EGU22-9964, updated on 05 Jul 2022
https://doi.org/10.5194/egusphere-egu22-9964
EGU General Assembly 2022

Citing large numbers of diverse datasets

James Ayliffe[1], Martina Stockhause[2], Shelley Stall[3], Deb Agarwal[4], Justin Buck[1], Caroline Coward[5], and Chris Erdmann[3]

[1]National Oceanography Centre, British Oceanographic Data Centre, United Kingdom of Great Britain – England, Scotland, Wales (jamayl@noc.ac.uk)

[2]Deutsches Klimarechenzentrum, Hamburg, Germany (stockhause@dkrz.de)

[3]American Geophysical Union, Washington DC, United States of America (SStall@agu.org)

[4]Lawrence Berkeley National Laboratory, Berkeley, United States of America (DAAgarwal@lbl.gov)

[5]NASA, Jet Propulsion Laboratory, Pasadena, United States of America (caroline.m.coward@jpl.nasa.gov)

In Earth and Biological sciences, data are often preserved and publicly available in data repositories where the data are citable by DOIs and published under a Creative Commons CC-BY license. Researchers combine many datasets across disciplines, repositories, and regions to better understand processes, patterns, and drivers. Citing these many datasets is difficult as the large number does not fit into the references section of a paper but the licenses of the datasets require that credit is given to their creators.

The Data Citation Community of Practice (CoP) was formed to target such challenges in data citation and other scholarly work that will support indexing and measuring the impact. The CoP identified a container as a solution for large numbers of data citations that holds the citations and its internal format, which is referred to as a 'reliquary'. The existing dataset collection methods have been gathered and evaluated using concrete citation use cases. Requirements for the reliquary content have been identified and applied to the use cases. In this presentation, we will report on the current progress on an approach to building a reliquary.

Reliquaries are an important part of enabling cross-disciplinary analysis of large amounts of data stored in many repositories. The challenge with a reliquary will be to design a method that works across diverse repositories and domain citation practices and to enhance the indexing system to direct credit to the reliquary content and authors. The CoP is in the process of setting up a Research Data Alliance (RDA) Working Group on Complex Citations in the Earth, Space, and Environmental Sciences to broaden the discussion and to find further use cases for evaluation and interested early adopters.