



Enabling FAIR and Open Data in Earth and Space Sciences Publications

Shelley Stall (1), Kerstin Lehnert (2), Lesley Wyborn (3), Erin Robinson (4), Helen Glaves (5), Mark Parsons (6), Brooks Hanson (1), Joel Cutcher-Gershenfeld (7), Brian Nosek (8), and Lynn Yarmey (6)

(1) American Geophysical Union, Washington, D.C., United States (sstall@agu.org, bhanson@agu.org), (2) Lamont-Doherty Earth Observatory, Columbia University, Palisades, N.Y., United States (lehnert@ldeo.columbia.edu), (3) National Computational Research Infrastructure, Canberra, Australia (lesley.wyborn@anu.edu.au), (4) Earth Science Information Partners, Boulder, CO, United States (erinrobinson@esipfed.org), (5) British Geological Survey, Nottingham, United Kingdom (hmg@bgs.ac.uk), (6) Rensselaer Polytechnic Institute, Troy, N.Y., United States (parsom3@rpi.edu, yarmel@rpi.edu), (7) Heller School for Social Policy and Management, Brandeis University, Waltham, Mass., United States (joelcg@brandeis.edu), (8) Center for Open Science, University of Virginia, Charlottesville, United States (nosek@cos.io)

Our research ecosystem is diverse and dependent on many interacting stakeholders that influence and support the process of science. These include funders, institutions, libraries, publishers, researchers, data managers, repositories, archives and communities. Process improvement in this ecosystem commonly requires the support of most, or all, of these stakeholders.

In October of 2014 a Coalition on Publishing Data in the Earth and Space Sciences (COPDESS) was formed to connect the Earth and space science publishers and data facilities to help translate the aspirations of open, available, and useful data from policy into practice. Initially funded by the National Science Foundation, then the Alfred P. Sloan Foundation, COPDESS provides an organizational framework for Earth and space science publishers and data facilities to jointly implement and promote common policies and procedures for the publication and citation of data across Earth Science journals. The launch of the partnership was announced on 15 January 2015 and included a joint Statement of Commitment signed by key publishers and repositories.

More recently, the value of FAIR (Findable, Accessible, Interoperable and Reusable) and Open Data has encouraged funders to sponsor discussions with tangible agreements, specifically with publishers, that include the steps needed to move the ecosystem towards results. Work by many of these stakeholders over the past years have developed pilot efforts that are ready to be scaled with broader engagement.

Building on the work of COPDESS, and with funding from the Laura and John Arnold Foundation, a partnership of the American Geophysical Union, Earth Science Information Partners (ESIP), Research Data Alliance (RDA), DataCite, Center for Open Science (COS), National Computational Infrastructure (NCI), Australian National Data Service (ANDS), AuScope and key publishers including Science, Nature, Elsevier, PLOS, Wiley, and the Proceedings of the National Academy of Science (PNAS) have agreed to work together along with leading repositories to develop integrated processes, leveraging these pilots, to make FAIR and open data the default for Earth and space science publications.

Along with the initial COPDESS effort, this project builds on the work of ESIP, RDA, the scientific journals, and domain repositories to ensure that well-documented data, preserved in a repository with community agreed-upon metadata, and supporting persistent identifiers becomes part of the expected research products submitted in support of each publication. No longer will data be locked up in hard-to-discover supplements. Data that supports research publications will be available in public repositories and capable of being accessed via persistent identifiers. (Protected and sensitive data will still have appropriate metadata that is open and accessible, but the data will have proper access controls.) This is a significant policy shift in how the research products (in particular data) supporting the research results are identified and referenced in publications. This effort takes decisive steps to enabling FAIR and open data, and supporting research integrity and reproducible science.