Statement of Commitment
from Earth and Space Science Publishers and Data Facilities

Coalition on Publishing Data in the Earth and Space Sciences

Earth and space science data are special resources, critical for advancing science and addressing societal challenges – from assessing and responding to natural hazards and climate change, to use of energy and natural resources, to managing our oceans, air, and land. The need for and value of open data have been encoded in major Earth and space science society position statements, foundation initiatives, and more recently in statements and directives from governments and funding agencies in the United States, United Kingdom, European Union, Australia, and elsewhere. This statement of commitment signals important progress and a continuing commitment by publishers and data facilities to enable open data in the Earth and space sciences.

Scholarly publication is a key high-value entry point in making data available, open, discoverable, and usable. Most publishers have statements related to the inclusion or release of data as part of publication, recognizing that inclusion of the full data enhances the value and is part of the integrity of the research. Unfortunately, the vast majority of data submitted along with publications are in formats and forms of storage that makes discovery and reuse difficult or impossible.

Repositories, facilities, and consortia dedicated to the collection, curation, storage, and distribution of scientific data have become increasingly central to the scientific enterprise. The leading Earth and space science repositories not only provide persistent homes for these data, but also ensure quality and enhance their value, access, and reuse. In addition to data, these facilities attend to the associated models and tools. Unfortunately, only a small fraction of the data, tools, and models associated with scientific publications makes it to these data facilities.

Connecting scholarly publication more firmly with data facilities thus has many advantages for science in the 21st century and is essential in meeting the aspirations of open, available, and useful data envisioned in the position statements and funder guidelines. To strengthen these connections, with the aim of advancing the mutual interests of authors, publishers, data facilities, and end-users of the data, a recent Earth and space science data and publishing conference, supported by the National Science Foundation, was held at AGU Headquarters on 2-3 October 2014. It brought together major publishers, data facilities, and consortia in the Earth and space sciences, as well as governmental, association, and foundation funders. Further informational meetings were held with Earth and space science societies, publishers, facilities, and librarians that were not present at the October meeting. Collectively the publishers, data facilities, and consortia focused on open data for Earth and space science formed a working group: Coalition on Publishing Data in the Earth and Space Sciences. As one outcome, this group collectively endorsed the following commitments to make meaningful progress

toward the goals above. We encourage other publishers and data facilities and consortia to join in support.

Signatory data facilities, publishers, and societies, in order to meet the need for expanding access to data and to help authors, make the following commitments:

- We reaffirm and will ensure adherence to our existing repository, journal, and publisher policies and society position statements regarding data sharing and archiving of data, tools, and models.

- We encourage journals, publishers, and societies that do not have such statements to develop them to meet the aspirations of open access to research data and to support the integrity and value of published research. Examples of policies and position statements from signatory journals and societies are listed here.

- Earth and space science data should, to the greatest extent possible, be stored in appropriate domain repositories that are widely recognized and used by the community, follow leading practices, and can provide additional data services. We will work with researchers, funding agencies, libraries, institutions, and other stakeholders to direct data to appropriate repositories, respecting repository policies.

- Where it is not feasible or practical to store data on community-approved repositories, journals should encourage and support archiving of data using community-established leading practices, which may include supplementary material published with an article. These should strive to follow existing NISO guidelines.

Over the coming year, the signatory Earth and space science publishers, journals, and data facilities will work together to accomplish the following:

- Provide a usable online community directory of appropriate Earth and space science community repositories for data, tools, and models that meet leading standards on curation, quality, and access that can be used by authors and journals as a guide and reference for data deposition.

- Promulgate metadata information and domain standards, including in the online directory, to help simplify and standardize data deposition and re-use.

- Promote education of researchers in data management and organize and develop training and educational tools and resources, including as part of the online directory.

- Develop a working committee to update and curate this directory of repositories.

- Promote referencing of data sets using the Joint Declaration of Data Citation Principles, in which citations of data sets should be included within reference lists.

- Include in research papers concise statements indicating where data reside and clarifying availability.

- Promote and implement links to data sets in publications and corresponding links to journals in data facilities via persistent identifiers. Data sets should ideally be referenced using registered DOI's.

- Promote use of other relevant community permanent identifiers for samples (IGSN), researchers (ORCID), and funders and grants (FundRef).

- Develop workflows within the repositories that support the peer review process (for example, embargo periods with secure access) and within the editorial management systems that will ease transfer of data to repositories.

A major challenge today is that much more Earth and space science data are being collected than can be reasonably stored, curated, or accessed. This includes physical samples, information about them, and digital data (sometimes streaming at rates of terabytes per minute). Researchers and publishers are looking for guidance on what constitutes archival data across diverse fields and disciplines. The major data repositories provide leading practices that should help guide the types of samples, data, metadata, and data processing descriptions that should be maintained, including information about derivations, processing, and uncertainty.

To enable improved coordination and availability of open data, we encourage funders to support these commitments, ensure a robust infrastructure of data repositories, and enable broad outreach with researchers. As a general rule, data management plans promulgated by funders should indicate that release into leading repositories, where available, of those data necessary to support published results is expected at publication. The ultimate measure of success is in the replicability of science, generation of new discoveries, and in progress on the grand challenges facing society that depend on the integration of open data, tools, and models from multiple sources.

Signatories

American Astronomical Society
American Geophysical Union
American Meteorological Society
Biological and Chemical Oceanography Data Management Office, Woods Hole Oceanographic Institution (BCO-DMO)
Center for Open Science
CLIVAR and Carbon Hydrographic Data Office (CCHDO)
Community Inventory of EarthCube Resources for Geosciences Interoperability (CINERGI)
Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)
Continental Scientific Drilling Coordination Office (CSDCO)
COOPEUS
Copernicus Publications
Council of Data Facilities
Dryad
Elsevier
European Geosciences Union
Geochemical Society
Geological Data Center of Scripps Insitution of Oceanography
Geological Society of America
Geological Society of London
ICSU World Data System
Incorporated Research Institutions for Seismology (IRIS)
Integrated Earth Data Applications (IEDA)
John Wiley and Sons
Magnetics Information Consortium (MagIC)
Mineralogical Society of America
Neotoma Paleoecology Database
National Snow and Ice Data Center
Nature Publishing Group
OpenTopography
Paleobiology Database
Paleontological Society
Proceedings of the National Academy of Sciences
Rolling Deck to Repository (R2R)
Science
Springer
UNAVCO