

# Data Policy

Bill Cook  
Bridging Data Lifecycles  
April 4, 2012



# Publisher's perspective

- Evaluate data in support of a manuscript
- Publish article with associated data
  - As integral component of article
  - As “supplemental” information
  - As referenced information
- Retain (curate)



# Public Policy efforts

- OSTP RFI closed, initial report issued April 3
  - No findings, simply an update
- EGU General Assembly
  - GI1.3 Data Assimilation and Data Publishing in Geosciences
  - Convened by Pangea



# AGU in particular

- A broad range of communities contributing to journals as authors, editors and reviewers
- Variations in data availability
- Need a policy that covers all communities



# Excerpts from existing policy – Citing Data

..data cited in AGU publications must be permanently archived in a data center or centers that meet the following conditions:

- are open to scientists throughout the world.
- are committed to archiving data sets indefinitely.
- provide services at reasonable costs.

Data sets that are available only from the author, through miscellaneous public network services, or academic, government or commercial institutions not chartered specifically for archiving data, may not be cited in AGU publications.



# Excerpts from existing policy – Preserving/Archiving Data

AGU does not expect to archive data sets subject to this policy, except on a for-fee basis and for sets of a small size. It is not AGU's intention to serve as an archive for large data sets that should be housed in data centers.

AGU maintains a deposit service for supplementary material of different types in order to provide long-term access to small supporting data sets and graphics files that are published concurrently with, and are an electronic component of, some AGU journal articles.



# Working draft of policy

In accordance with the core values of the Union, the AGU has adopted the following Policy to enhance the review process and to promote the transparency of science:

An inherent principle of publication is that others should be able to evaluate and build upon the authors' published results. AGU strongly encourages authors to archive their data ([http://www.agu.org/pubs/authors/policies/data\\_policy.shtml](http://www.agu.org/pubs/authors/policies/data_policy.shtml)) in publicly endorsed community archives (e.g., [http://www.agu.org/pubs/authors/manuscript\\_tools/journals/auxiliary\\_material/data\\_centers.shtml](http://www.agu.org/pubs/authors/manuscript_tools/journals/auxiliary_material/data_centers.shtml)).

On request, supporting data products must be made available to editors and peer reviewers for the purposes of evaluating the manuscript. Peer reviewers may be asked to comment on the terms of access to materials, methods and/or relevant data products. Restrictions on the availability of materials or information must be disclosed to the editors at the time of submission.

AGU encourages authors to make materials, relevant data products and associated protocols promptly available to readers without undue qualifications in material transfer agreements. Data product availability and limitations must be specified in the paper.

AGU reserves the right to refuse publication in cases where authors do not provide adequate assurances that they can comply with this Policy.



# Data Identifiers

- CrossRef looking at data DOIs
  - Can identify data today
- Would be persistent, similar to article DOIs
- Would help to reach content, but not to consolidate or curate
- Need to include data repositories as prime holders of data and DOIs

