Earth and Space Science Data Should Be Widely Accessible in Multiple Formats and Long-term Preservation of Data is an Integral Responsibility of Scientists and Sponsoring Institutions

*Earth and space sciences data bases are a world heritage that should be made available to the scientific community and public as soon as possible (in some cases in real-time), should be organized and preserved in useable format, and should be conserved long-term for future use. The responsibility for achieving this falls upon individual scientists and their sponsoring institutions, and should be considered an integral part of conducting scientific research.*

Earth and space science data collection, analysis, and archiving are essential to our understanding of the natural environment and how it changes with time. AGU policy is grounded in the principle of full and open sharing of such data and associated metadata for research and education. Adherence to this policy will foster scientific advances, yield economic benefits, improve decision-making, enhance public safety and wellbeing, contribute to national and global security, and lead to a more informed public.

AGU extends this policy to data and derived products acquired with both public and private funding. To maximize scientific and educational returns, Earth and space science data should enter the public domain as soon as possible. The marginal cost of data dissemination is a legitimate charge to users. In limited circumstances, access to certain data might be restricted to protect security, confidentiality, or commercial value, but such restrictions should remain infrequent and temporary and be carefully justified on a case-by-case basis.

For some issues, such as responding to natural hazards, access to real-time data is critical. Further, assimilation of near-real-time data in models is becoming increasingly important for monitoring and predicting changes in the Earth’s environment and climate. Newly generated Earth and space science data and associated metadata should therefore be submitted to the appropriate governmental or intergovernmental data center promptly.

Documenting trends and long-term changes is essential for understanding many natural phenomena. Because the state of natural systems is never repeated, data losses, or missed data collection opportunities can never be corrected. Consequently, the value of Earth and space science data grows with time, placing a premium on long-term data curation. Because datasets are often later used for purposes other than those for which they were collected, accurate, complete, and, when possible, standardized metadata are as important...
as the data themselves. These should include the dataset version number to facilitate future reprocessing of both observational and model or simulation data.

The cost of collecting, processing, validating, and submitting data to a recognized archive should be an integral part of research and operational programs. Such archives should be adequately supported with long-term funding. Organizations and individuals charged with coping with the explosive growth of Earth and space digital data sets should develop and offer tools to permit fast discovery and efficient extraction of online data, manually and automatically, thereby increasing their user base. The scientific community should recognize the professional value of such activities by endorsing the concept of publication of data, to be credited and cited like the products of any other scientific activity, and encouraging peer-review of such publications.

Earth and space science data are a world heritage. Properly documented and archived, they will help future scientists understand the system that comprises the Earth's interior, atmosphere, oceans, biosphere, and space environment. Taking proper care of such data is our responsibility and our obligation to future generations.