

Full, Open, and Timely Access to Data

A Policy Statement of the American Meteorological Society

(Adopted by AMS Council on 15 April 2019)

Introduction and Background

Full, open, and timely access to environmental data benefits science and society and is critical to the American Meteorological Society (AMS) community, including academic, government, nonprofit, and commercial interests. All of these groups benefit from the exchange of data related to the Earth system. For example, modern society relies on accurate weather forecasts for protection of life and property and as a basis for routine decisions. Indeed, prompt and accurate forecasts are not possible without the exchange of data among the 191 World Meteorological Organization (WMO) Member States and Member Territories under the World Weather Watch Program, per policies stated in WMO Resolution 40 and Resolution 25 [1]. Beyond weather forecasting, the research, education, and business communities increasingly need access to global environmental data. The global and complex nature of environmental problems requires that research be conducted collaboratively by distributed teams of investigators, which often necessitates sharing data, knowledge, and other resources. In addition, technical and policy ecosystems related to environmental data are changing rapidly, requiring the AMS stakeholder groups to reassess their approaches to data collection and sharing.

The purpose of this statement is to reaffirm AMS's commitment to a policy of full, open, and timely access to data that are critical to the advancement of atmospheric and related sciences, the provision of products and services for the benefit of society, and the promotion of commerce and private-sector activities. Adopting such policies could accelerate scientific discoveries, broaden and enhance participation in scientific enterprise, promote entrepreneurship, and benefit society.

AMS encourages its stakeholder communities to provide full, open, and timely access to environmental data and derived data products, as well as all associated information necessary to fully understand and properly use the data (metadata). In this context, *full* means that all data and metadata should be available, *open* means that it should be available to anyone who requests it, and *timely* means that it should be available as soon as possible, particularly in the case of data critical to human health and safety. These data are at the foundation of efforts to ensure public safety and national security, as well as efficient management and use of weather- and climate-sensitive sectors and systems such as water resources, transportation, and agriculture. Environmental data are used to protect critical infrastructure and support scientific publications, and they are essential for routine and high-impact weather forecasting and warning and climate monitoring. "Data," in this statement, refers to entities or outputs used as evidence of phenomena for the purposes of research or scholarship. The spectrum of data is diverse and includes in situ and remotely sensed observations, environmental predictions generated by numerical models, and data products derived from integrations of observational and model-generated sources.

For all data types, the provision of metadata and the use of widely accepted data and metadata format standards are critical. Data and metadata formats that have open and machine-readable specifications allow people to access and use data via a variety of services and software packages. Open formats encourage the sharing of data and their associated analysis tools. Maintaining the relationships between data, the software used to analyze them, and any publications or documents that result from the use of data and software is critical to establishing the transparency and trustworthiness of scientific work. These relationships are typically maintained through documentation of provenance, that is, the steps that were taken to transform, analyze, and visualize data.

In addition, AMS expects all scholarly papers published in its journals to contain sufficiently detailed references to public sources of information (literature and data) and methodology such that independent research can test the paper's scientific conclusions. This expectation is codified in the AMS "Data Archiving and Citation" guidelines within the AMS Publications' Author Guide [2], which provides recommendations for how data and metadata used to produce scientific conclusions presented in journal articles should be properly archived, cited, and made readily available to the scientific community.

Data discoverability, whatever the mechanism, is necessary to facilitate full, open, and timely access to data. Beyond increasing data discoverability through the scientific literature, there have been significant steps made by major search engines to use modern web standards for marking up datasets for discovery. The use of persistent web-accessible identifiers and community accepted metadata schema to identify and describe the dataset are recommended to take advantage of web-based data discovery and access.

The AMS concept of full, open, and timely access is to promote, to the greatest extent possible, data availability. Many governmental scientific agencies within the United States and internationally already have data policies in place [3], and have taken steps to ensure that data upon which scientific studies are based are preserved and made available at little or no cost to users. These policies emphasize the importance of data management planning and establish common expectations for data archiving and sharing. Such policies should continue to evolve as new data collection instruments are developed and as new uses and users of data emerge.

However, AMS recognizes that there are often significant costs associated with the collection, creation, storage, value-adding, and distribution of data. Therefore, it is reasonable for commercial data providers to charge for access to their data at rates commensurate with their value and costs.

Further, AMS recognizes that government agencies, academia, and private-sector entities face additional challenges in meeting the goals of open data. Some of these challenges include policies that inhibit the widest possible use of data, such as resource constraints, intellectual property rights, ethical and legal constraints related to human privacy or personal data ownership, and inadequate infrastructure to properly maintain and administer datasets. As one example, the European Union's General Data Protection Regulation (GDPR) regulates the use of data that may be directly or indirectly related to individuals from the EU [4].

To advance the goals of this statement and address these challenges, AMS proposes the following core principles and recommendations. These points are aligned with the growing international movement to

enable scientific data to be FAIR: Findable, Accessible, Interoperable, and Reusable [5]. These principles and recommendations are intended to inform national and international data policies and influence standard practices in academia, government, and the private sector in the pursuit of open access to data.

Principles

- **Full, open, and timely access to environmental data.** Public and private entities engaged in activities that ensure public safety, protect critical infrastructure, and support commerce should enable full, open, and timely access to environmental data.
- **Full, open, and timely access to scientific research data.** Access to data generated by scientists is fundamental to advancing basic and applied science and the independent validation of research findings.
- **Reasonable costs associated with data access and use.** Collection, creation, storage, value-added processing and distribution of data incur costs that must be recovered in some form. To promote full, open, and timely access to environmental data, cost recovery should be transparent and commensurate with the value and use of the data.
- **Well-documented data in formats that are standard, optimize understanding, and support interoperability and reuse.** The creation and maintenance of robust metadata and data in accepted standard formats that support interoperability are fundamental to data exchange and effective analyses. Data and data products should be archived, preserved, stewarded, and accompanied by these metadata that will ensure they will be fully usable and understandable for generations to come serving the need of science and society.
- **Pre-planned, value-added joint data collection capabilities.** In times of crisis or emergency, the government and supporting private-sector organizations should be able to activate existing pre-planned value-added joint data collection capabilities that enhance the ability to address the crisis, ensuring fair compensation and protection of intellectual property rights.

Recommendations

AMS recommends these guidelines to all institutions and individuals involved with environmental and related data:

- **Full, open, and timely access to environmental data**
 - Promote the importance and value of full, open, and timely access to environmental data, and why they should be integrated into practices of all data providers and users.
 - Advocate for more open data policies wherever barriers to such exist today.
- **Full, open, and timely access to scientific research data**
 - Enable data to be discoverable and accessed through well-understood, standardized, and modern technical approaches.
 - Use persistent identifiers to enable web-based discovery, access, and reference of data.
- **Reasonable costs associated with data access and use**
 - Encourage funding agencies, scientific societies, and other organizations to fully recognize the costs of data and address the need for adequate research funding to support those activities.

- Promote policies and practices that encourage research innovation by reducing barriers of cost, licensing, and other impediments to the open access concepts.
- **Well-documented data in formats that are standard, optimize understanding, and support interoperability and reuse**
 - Share expertise related to data preservation and stewardship at national and international meetings.
 - Promote the use of community and international conventions and standards for data and metadata for providing, archiving, referencing, and sharing to increase data use, interoperability, and discovery.
 - Encourage data sharing from relevant data collection and product development systems, including data from traditionally hard-to-obtain observing networks (e.g., hydrologic and air-quality observations from federal and state government networks).
 - Develop and promote ongoing community discussion, coordination, and shared responsibility between, data collectors, providers, archivists, and users that lead to data value assessment, data quality assurance and control, and protocols for reasoned decisions concerning what data to preserve, and for how long, within the resources available.
- **Pre-planned, value-added joint data collection capabilities**
 - Encourage processes that enable sharing of all privately and publicly held data during times of crisis or emergency, ensuring fair compensation to entities that own the data and/or operate proprietary data collection systems.
 - Encourage public–private partnerships to ensure data availability to support public safety, distribute costs, protect critical infrastructure, and advance science and related services.

References

[1] See: http://www.wmo.int/pages/prog/hwrrp/documents/wmo_827_enCG-XII-Res40.pdf and www.wmo.int/pages/prog/hwrrp/documents/Resolution_25.pdf.

[2] AMS Publications. 2017. Data Archiving and Citation <https://www.ametsoc.org/ams/index.cfm/publications/ethical-guidelines-and-ams-policies/data-archiving-and-citation/>

[3] See, for example:

- a. U.S. Executive Office of the President, Office of Management and Budget, "Open Data Policy- Managing Information as an Asset," <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf>
- b. Group on Earth Observations (GEO), "Data Sharing Principles," https://www.earthobservations.org/geoss_dsp.shtml
- c. NOAA, "Environmental Data Management (EDM) Framework," <https://nosc.noaa.gov/EDMC/framework.php>

[4] European Union, General Data Protection Regulation (GDPR)
https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

[5] Wilkinson, M. D., and Coauthors, 2016: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>.

[This statement is considered in force until April 2024 unless superseded by a new statement issued by the AMS Council before this date.]