

It's Time to Share: Proposed NOAA Data Policy

Information Briefing for NOS Seminar Series
September 7, 2011

Ingrid Guch

Director, NESDIS Cooperative Research Program





NOAA Data Sharing

- Recent signing of NAO 212-15 **Environmental data will be visible, accessible and independently understandable to users**, except where limited by law, regulation, policy or requirements
- **Numerous calls for data sharing policy for grants in the past**, IE GAO Recommendation (2007): To ensure that the agencies maximize opportunities to make data available in a manner useful to other researchers, the Secretaries of Commerce and Energy, the NASA Administrator, the NOAA Administrator, and the NSF Director should consider using the grant review process, where their program offices are not currently doing so, to facilitate further data sharing by **(1) evaluating researchers' data-sharing plans as part of the grant review process** and **(2) using evidence of researchers' past data-sharing practices to make future award decisions**. The use of such criteria in the grant review process should be clearly conveyed to researchers before they submit research proposals and after award decisions have been made.



Interested Parties

- NOAA leadership
- Grantees
- NOAA Federal Program Officers
- NOAA Competition Managers
- NOAA Councils (Research, Observing Systems, CIO, Grants Management, Environmental Data Management)
- NOAA General Council
- SAB/DAARWG
- DoC Federal Acquisition Law Division
- OMB
- Cooperative Institutes, Sea Grant, GEO, other large programs/partnerships
- State and local managers
- Federal partners
- Disaster response/mitigation/prediction
- Many, many more

Data Sharing Policy for NOAA Grants – Schedule and Milestones



Today's Date ▼

Limited Comment Period

- Desired Outcome this Fiscal Year (per NOAA senior management)
 - ***Data collected/created under a NOAA Grant or Cooperative Agreement will be understandable and available for interested users in a timely manner at minimal cost***
- Current Status:
 - Procedural Directive Approved by EDMC
 - Briefed Research Council and NOAA Data Management Conference attendees
- Next Steps:
 - Notify NOSC, OCIO of intent to publish PD
 - Publish PD along with training materials for grantees
 - Education and outreach introducing PD to grantees and program managers
 - Annual evaluation of PD effectiveness, changes incorporated as needed



Policy Development Team

- Lead: Ingrid Guch/NESDIS
- Key Members: Chris Miller/OAR-CPO, Dorn Carlson/OAR-SeaGrant, Jim Sargent/NMFS, Beth Turner/NOS
- Councils Briefed: EDMC (owner), RC (including CI Subcommittee), NOSC, OCIO, GMAC, SAB DAARWG
- DoC/FALD Reviewers: Jeff Joyner and Elizabeth Becker



Context

- Environmental Data in the context of the Grants Data Sharing Policy at NOAA has one or more of the following attributes:
 - Potential broad utility
 - Significant NOAA funds were spent creating/collecting the data
 - Are a reference data set
 - Are associated with community buy-in

Additional attributes help ensure benefits of sharing data outweigh costs (financial as well as risk of PI being “scooped”). More information about attributes will be in training materials.



Context

- Sharing data refers to making data visible, accessible, and independently understandable to users in a timely manner at minimal cost, except where limited by law, regulation, policy or by security requirements. NOAA facilities that archive data and make the data openly available should be considered for the disposition of the data.



Policy in a nutshell

Unless otherwise specified by NOAA program in FFO or other grant documents...

- all environmental data created using NOAA grant funding must be shared no later than 90 days after the end date of the project (at the same time the final project report is due)
- a 2-page Data Sharing Plan is required to be part of the project narrative. Typically includes:
 - Types of environmental data created
 - Standards to be used for data format and content;
 - Policies addressing data stewardship and preservation;
 - Previous data sharing experience; and
 - Procedures for providing access, sharing, and security.



Publications

- Balancing grantee need to publish with community need for timely data sharing
 - This policy says “unless otherwise indicated in the FFO, data must be shared no later than 90 days after the project end date”
 - Pros: Clarifications/communications/enforcement actions, if needed, can coincide with review of grant final progress report also due at that time
 - Cons: Many NOAA programs specified a preference for sharing data “*no later than 2-years after the data were collected*”, similar to NSF programs.
 - Mitigation: NOAA programs who believe they can track/reward/encourage compliance with a later sharing date can specify alternate timeframe in FFO.



Expectations

- Balancing desire for archival-quality data with community need for timely data
 - Sharing data no later than 90 days after the project end date means that it may not be “archival quality”
 - Pros: Pre-publication data, with appropriate disclaimers or marker papers, can be used by others - beyond what the data producers could do themselves in a similar time period, and frequently for scientific purposes outside the original goals of the project. Information from others can be given to data producers to improve or better understand the data
 - Cons: Additional resources/time needed to share data in pre-publication state
 - Mitigation: Data Sharing Plan required with proposal submission will ensure that PIs anticipate sharing pre-publication quality data prior to collection and can plan project to minimize cost of doing this.



Outreach and Training Fliers, Website, Conferences

Coming Soon: New Grants Policy on Data Sharing at NOAA

All environmental data and information collected or created under NOAA grants or cooperative agreements awarded after [date] must be made visible, accessible, and independently understandable to users in a timely manner, except where limited by law, regulation, policy or by security requirements.

Frequently Asked Questions

What is meant by "environmental data"? Environmental data are recorded and derived observations and measurements of the physical, chemical, biological, geological, and geophysical properties and conditions of the oceans, atmosphere, space environment, sun, and solid earth, as well as correlative data, such as socio-economic data, related documentation, and metadata. Media, including voice recordings and photographs, may be included. Environmental Data in the context of the Grants Data Sharing Policy at NOAA also has one or more of the following attributes: potential broad utility, significant NOAA funds were spent creating/collecting the data, are a reference data set, or are associated with community buy-in. Refer to the solicitation and/or contact the funding program if you are unsure if the data you are collecting/creating meets the criteria.

Why share data? The greater the availability of the data, the more quickly and effectively user communities can develop innovative practical applications for public benefit. In many cases these applications will be in areas not originally anticipated by the principal investigator. In addition, NOT making data available that supports scientific findings may provide reason to doubt the validity of the findings and limit their usefulness. More information about data sharing concerns are in a 2007 GAO Report, <http://www.gao.gov/new.items/d071172.pdf>. "Climate Change Research: Agencies Have Data-Sharing Policies but Could Do More to Enhance the Availability of Data from Federally Funded Research." The GAO report specifically recommends NOAA consider evaluating data sharing plans as part of the grant review process.

Data sharing is widely accepted as a good practice. National scientific organizations have made a commitment to the sharing and archiving of data through their ethical codes (e.g., the American Sociological Association) or publication policies (e.g., the American Psychological Association). More than 15 years ago, the National Academy of Sciences described the benefits of sharing data. (See <http://books.nap.edu/catalog/2033.html>) For many years, the National Science Foundation (NSF) Economics Program has required data underlying an article arising from an NSF grant to be placed in a public archive. Similar expectations exist at the National Institute of Health, and the National Institute of Justice. Moreover, many scientific journals require that authors make available the data included in their publications.

Some examples of methods to share data:

- Formal Archive (example: www.ncdc.noaa.gov)
- Portal (example: www.geodata.gov)
- Project Website (example: <http://cimss.ssec.wisc.edu/aeri/>)
- Informally between researchers on a peer-to-peer basis

Consider the anticipated benefit of the data, the likely number of interested users of the data, and the priorities of the funding program when choosing a method to share data. In addition there are several reports that contain information about sharing environmental data, including:

Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age (National Academies, 2009)

Environmental Data Management at NOAA: Archiving, Stewarding, and Access (2007), Committee on Archiving and Accessing Environmental and Geospatial Data at NOAA, National Research Council

A template for a data sharing plan is not available; however relevant sections from a Data Management Plan are consolidated below and may be helpful.

General Description of Data Collection

1. Name and Purpose of Project:
2. What data will be generated in the project? (Give a short description, including amount, if known, and the content of the data).
3. Project Time Frame or Date(s) of Collection:

June 20-22, 2011

NOAA Data Management Conference



Contact Information

- Ingrid.guch@noaa.gov 301-763-8127x152
(Ingrid is collecting comments/questions)
- NOAA Environmental Data Management Committee: www.edmc.noaa.gov
- NOAA Observing System Council:
www.nosc.noaa.gov
- NOAA CIO Council
http://www.cio.noaa.gov/IT_Groups/noaa_cio_CIOCouncil.html (or google NOAA CIO Council)



References

- <http://www.nature.com/nature/journal/v461/n7261/full/461168a.html>
- http://www.sois.uwm.edu/MetricsPreCon/documentation/Piwowar_Chapman_Sharing.pdf
- http://bing.exp.sis.pitt.edu:8080/webdav/archives_072211/lis2000_archive/journal.pone.0007078.pdf
- <http://www.gao.gov/new.items/d071172.pdf>

Backup

Challenges



- Some NOAA grants (Cooperative Institutes, Sea Grant) may be for tens of millions of dollars, hundreds of projects for multiple years, while other NOAA grants are for a single year and projects less than \$50K – the data sharing plans will necessarily be very different
 - The implementation of this policy will have each Data Sharing Plan associated with a single NOAA award rather than project-based.
 - Pros: Federal Program Officers for Cooperative Institutes and Sea Grant will not need to oversee hundreds of data sharing plans. Encourages Cooperative Institutes and Sea Grant to provide their scientists with standard methods of data sharing rather than on a project by project basis.
 - Cons: Some project scientists may not know they are supposed to be following the larger award's data sharing plan. Some larger awards' data sharing plans may be so general they are not effective.
 - Mitigation: NOAA Federal Program Officers for Cooperative Institutes and Sea Grant will need to spend more time than single-award grantees /FPOs on the data sharing plan to make sure it is likely to be effective.



NSF

– Data Sharing Policy

- Investigators are expected to **share with other researchers**, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. See [Award & Administration Guide \(AAG\) Chapter VI.D.4.](#)

– Data Management Plan Requirements

- Beginning **January 18, 2011**, proposals submitted to NSF must include a supplementary document of no more than two pages labeled “Data Management Plan”. This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. See [Grant Proposal Guide \(GPG\) Chapter II.C.2.j](#) for full policy implementation.



NSF Grant Proposal Guide

Plans for data management and sharing of the products of research.

- Proposals must include a supplementary document of no more than two pages labeled “Data Management Plan”.
- This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see [AAG Chapter VI.D.4](#)), and may include:
 - the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project
 - the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);
 - policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
 - policies and provisions for re-use, re-distribution, and the production of derivatives; and plans for archiving data, samples, and other research products, and for preservation of access to them.



NASA

- Promotes the **full and open sharing of all data with the research and applications communities, private industry, academia, and the general public.**
- Adopted a data policy with 11 bullets including “All NASA Earth science missions, projects, and grants and cooperative agreements shall include data management plans to facilitate the implementation of these data principles.”



NSF Award and Administration Guide

Dissemination and Sharing of Research Results

- a. Investigators are expected to promptly prepare and submit for publication, with authorship that accurately reflects the contributions of those involved, all significant findings from work conducted under NSF grants. Grantees are expected to permit and encourage such publication by those actually performing that work, unless a grantee intends to publish or disseminate such findings itself.
- b. Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. Privileged or confidential information should be released only in a form that protects the privacy of individuals and subjects involved. General adjustments and, where essential, exceptions to this sharing expectation may be specified by the funding NSF Program or Division/Office for a particular field or discipline to safeguard the rights of individuals and subjects, the validity of results, or the integrity of collections or to accommodate the legitimate interest of investigators. A grantee or investigator also may request a particular adjustment or exception from the cognizant NSF Program Officer.
- c. Investigators and grantees are encouraged to share software and inventions created under the grant or otherwise make them or their products widely available and usable.
- d. NSF normally allows grantees to retain principal legal rights to intellectual property developed under NSF grants to provide incentives for development and dissemination of inventions, software and publications that can enhance their usefulness, accessibility and upkeep. Such incentives do not, however, reduce the responsibility that investigators and organizations have as members of the scientific and engineering community, to make results, data and collections available to other researchers.
- e. NSF program management will implement these policies for dissemination and sharing of research results, in ways appropriate to field and circumstances, through the proposal review process; through award negotiations and conditions; and through appropriate support and incentives for data cleanup, documentation, dissemination, storage and the like.