

DATA SHARING PLANS

Some reference documents follow that might answer some of your questions about data sharing plans.

Included are:

a “Frequently Asked Questions” document created by Dorn and the Research Coordinators, with help from the rest of the network. It includes some examples of how some SG Programs are dealing with the new Data Sharing Plan requirement.

a Powerpoint about the new data sharing plan prepared by NOAA.

two other “Frequently Asked Questions” documents from two other granting agencies who have several years of experience dealing with DSP questions from grantees. Their answers won’t always be the same as NOAA’s, but they are still informative.

Other questions? contact Dorn Carlson at the National Office, dorn.carlson@noaa.gov, 301-734-1080.

Data Management Plan FAQs

This document is for Qs and As about NOAA's new data management plan requirements. It was copied from the collaborative FAQ document developed and shared on the BASECAMP Sea Grant Research Coordinators site on 5/1/2013.

For the starting point, Dorn used FAQs developed by NOAA from the powerpoint shown at the Research Coordinators meeting at Sea Grant week (available at https://geoide.noaa.gov/wiki/index.php?title=FAQ_for_Data_Sharing_for_NOAA_Grants_PD), and some actual questions I received. It's been supplemented with Qs and As excerpted from email dialogs on the subject in November and December.

This FAQ document is divided up into the following sections:

- GENERAL QUESTIONS
- HELP MANAGING DATA
- SEA GRANT RFP LANGUAGE (including some examples)
- SEA GRANT OMNIBUS PROPOSAL LANGUAGE
- MORE INFO

GENERAL QUESTIONS

Question: What is this new requirement?

Answer: The NOAA Directive on Data Management.

https://www.nosc.noaa.gov/EDMC/nao_212-15.php

It applies to data whose generation was funded by NOAA, including grants. The Directive says, Environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy

(such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements.

Question: What is meant by “environmental data”?

Answer(NOAA): 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.

Question: Are socio-economic studies in coastal communities included in the requirement?

Two examples might illustrate:

- 1) survey work to determine willingness to pay for improvements to beach access**
- 2) fisheries work that looks at the economic impact of a new vs. old harvest technique in different coastal zones (may include collection of harvest location basic oceanographic data)**

Answer(Dorn): As I understand the guidance, both examples might be covered by the requirement. The NOAA directive specifically says that social science data should be included if it is (from the Q&A above) “correlative” with other environmental data.

Question: What is meant by “sharing”?

Answer(NOAA): 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.

Question: What is considered “timely” data sharing?

Answer(NOAA): This will depend on the program awarding the grant or cooperative agreement, and the nature of the research project conducted. It is typically no later than two (2) years after the data are collected or created.

Question: What is meant by “independently understandable”?

Answer(NOAA): The data must be accompanied with documentation, metadata and, if needed, tools to read the data that allow a user to interpret the data properly. If there are concerns with understandability, they can be reported to NOAA, who will do an independent check.

Question: Who will determine if my data are visible, accessible and independently understandable?

Answer(NOAA): The person generating the data will have first responsibility for determining this. Common data quality standards in your scientific discipline may help you decide if the data are understandable. Ultimately, others who use your data will know whether they are visible, accessible and understandable to them. If there are concerns with data access or understandability, they can be reported to NOAA, who will do an independent check.

Question: What are examples of law, regulation, policies or security requirements that may limit my ability to share data?

Answer(NOAA): Policies applicable to protection of personally identifiable information, critical infrastructure information or proprietary trade information as well as regulations related to export control may impact your ability to share data, among other items.

Question: How must data be shared?

Answer(NOAA): This depends on the nature of the project and the data, and will be proposed by the investigator himself. Data sharing can be accomplished through:

- ✓ Data Archive: place where data are acquired, manipulated, documented, and distributed. NOAA facilities that archive data and make the data openly available should be considered.
- ✓ Data Enclave: controlled, secure environment in which eligible researchers can perform analyses using data resources
- ✓ Publishing: articles in scientific publications
- ✓ Researcher’s Efforts: investigator responds directly to data requests (e.g., posting data on a Web site)

Question: I’m a busy investigator. I don’t have time to process requests for my data. What should I do?

Answer(NOAA): In addition to publishing small datasets, there are several alternatives to responding to each separate request to share data (e.g., putting data in an archive or restricted access facility, and setting up a web site for data access). Archives and data enclaves provide technical assistance for users with questions or problems and may spare busy investigators time.

Question: Can I get additional funding to share my data?

Answer(NOAA): Unless otherwise noted in the federal funding announcement, funding to address data sharing must be requested as part of the proposal to collect/create data. The data sharing plans and related funds requested should consider the anticipated benefit of the data, the likely number of interested users of the data and the priorities of the program as outlined in the solicitation.

Question: I am the PI of a large Sea Grant program funded by an omnibus grant which in turn manages a number of individual research projects. Must every individual project have its own data sharing plan, or can I develop a program-wide data sharing plan?

Answer(NOAA): As the omnibus grant recipient, you have a responsibility to see that data sharing plans are followed for all research projects under your program. This may be done with a single Program-wide data sharing plan, individual plans for individual projects, or something in between, as long as all the relevant data generated is covered under some data sharing plan.

HELP MANAGING DATA

Question: What data management services or resources are available?

Answer(NOAA): There is information available at the NOAA Environmental Data Management Committee website reachable from www.nosc.noaa.gov. In general considering data sharing requirements prior to finalizing the methods for collecting/creating/storing the data will save time and effort later on. Unless otherwise noted in the federal funding announcement there is no specific data sharing plan template required.

Answer(from SG discussions): Several data centers are available for consideration. Many SG Programs are already involved in research projects funded by EPA, NSF or elsewhere that already have specified centers to which the data generated must be submitted.

Some of these data centers are below. Their websites often have general data management guidance, as well as instructions on using their services:

✓ NOAA-supported Data Centers. This document (https://www.nosc.noaa.gov/EDMC/documents/NOAA_Procedure_document_final_12-16-1.pdf) provides guidance for data managers who want to submit their data to a NOAA Data Center.

✓ In particular, NOAA NODC is what some universities consider “the place” to store all environmental data generated by NOAA grants. Website: <http://www.nodc.noaa.gov/General/NODC-Submit/submit-guide.html#polguide>

✓ The NSF-supported BCODMO (biological oceanography-chemical oceanography data management office). Website: <http://bcodmo.org/>

✓ The NSF-supported CUAHSI (Consortium of Universities for the Advancement of Hydrologic Science, Inc.). This a 501(c)3 research organization of than 130 U.S. universities (including some SG universities) who have an ongoing effort to provide a repository for hydrologic data produced by members (or others, I believe). I’m told they hope to become “the Google of hydrological data”. Website: <http://www.cuahsi.org/>

Question: The NOAA [or other federal] Program I apply to already requires an extensive Data Management Plan. Do I still need to do a Data Sharing Plan?

Answer(Dorn): Not necessarily—the one you did for the other Program could suffice, if you plan to follow that plan when managing your data.

Question: I am not collecting any data/information. What should I do?

Answer(NOAA): A statement indicating you are not collecting any data/information will be appropriate for your data sharing plan.

SEA GRANT RFP LANGUAGE

Question: what are SG Programs required to say in their RFP about data managment?

Answer(Dorn): Nothing. There is no requirement for any language in your RFPs. There is a requirement for those of us in NOAA to include language about data management in our Federal Funding Opportunity announcements, but my interpretation of the rules is that this does not extend to you.

What (if anything) you say in your RFP depends on how you plan to manage the data, and how you plan to communicate with your PIs about data management planning, as part of your competition and selection process. The only requirement is that, however you decide to handle

the data management, your omnibus proposal must include the plan (or plans).

If you are interested, this is what the NOAA procedural directive on grant data sharing (available here) says that NOAA must include in its FFOs. You can use this for information, but again, grantees are not required to include this in their RFPs:

The following language will be incorporated for [NOAA] competition announcements:

Environmental data and information collected and/or created under NOAA grants/ cooperative agreements must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner (typically no later than two (2) years after the data are collected or created), except where limited by law, regulation, policy or security requirements.

1. Unless otherwise noted in the federal funding announcement, a Data/Information Sharing Plan of no more than two pages shall be required as part of the Project Narrative. A typical plan should include descriptions of the types of environmental data and information created during the course of the project; the tentative date by which data will be shared; the standards to be used for data/metadata format and content; policies addressing data stewardship and preservation; procedures for providing access, sharing, and security; and prior experience in publishing such data. The Data/Information Sharing Plan will be reviewed as part of the NOAA Standard Evaluation Criteria, Item 1 -- Importance and/or Relevance and Applicability of Proposed Project to the Mission Goals....

Question: can you provide a sample statement to go in our SG Program's RFP?

Answer(Dorn): Here are some samples of what SG Programs are considering for their RFPs. All of these approaches are acceptable:

LOUISIANA SG: "Our tentative strategy is to hold off on the data management plans until we determine which proposals have been funded in our 2014-2016 omnibus. Once we decide which

among the funded proposals need a data management plan, we will request same to be returned with the proposal addendum."

MIT SG: Somewhere in the RFP process, we get the PI to sign off on a statement like this: Process for Providing Data and Information: Upon completion of the [State] Sea Grant funded project, the PI must provide [State] Sea Grant with electronic files containing all data and information collected during the project, as well as metadata that provide sufficiently detailed and easily understood information for another person to make use of the data and information. Files can be provided in the most appropriate format depending on the type of data (e.g., Excel spreadsheets for field observations or results of controlled experiments, source code for newly developed or enhanced models, shape files for spatial information, responses from surveys, etc.). [State] Sea Grant will hold these data on a secure server, and only after two years have passed will we provide the data to third parties who provide us with a written request for those specific data and information.

FLORIDA SEA GRANT: Here is the language from the draft Florida 2013 RFP.

9. NOAA Data Sharing Requirement (effective for all new NOAA funded research projects) Data and information collected and/or created under NOAA grants and cooperative agreements must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner (typically no later than two years after the data are collected or created), except where limited by law, regulation, policy or by security requirements.

The new requirement has two basic parts: (1) environmental data generated by a grant project must be made available after a reasonable period of exclusive use, and (2) the grant application must describe the plan to make the data available.

To comply with this new requirement, in the Full Proposal stage, the Principal Investigator must comply with the process described below, and in that Full Proposal, explain how the data and metadata will be provided. Funds may be budgeted in the project proposal for this task. For the

Statement of Interest, the Principal Investigator must simply agree in writing that they intend to follow this new NOAA requirement.

Process for Providing Data and Information: Upon completion of the Florida Sea Grant funded project, the PI must provide Florida Sea Grant with electronic files containing all data and information collected during the project, as well as metadata that provide sufficiently detailed and easily understood information for another person to make use of the data and information. Files can be provided in the most appropriate format depending on the type of data (e.g., Excel spreadsheets for field observations or results of controlled experiments, source code for newly developed or enhanced models, shape files for spatial information, responses from surveys, etc.). Florida Sea Grant will hold these data on a secure server, and only after two years have passed will we provide the data to third parties who provide us with a written request for those specific data and information.

TEXAS SEA GRANT: We are adding to our RFP:

Effective for all new NOAA funded research projects, data and information collected and/or created under NOAA grants and cooperative agreements must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner, except where limited by law, regulation, policy or by security requirements.

The new requirement has two parts: (1) environmental data generated by a research project must be made available after a reasonable period of exclusive use, and (2) the grant proposal must describe the plan to make the data available.

To comply with this new requirement, proposers must provide a “Data Management Plan”, explaining how the data and metadata will be provided. Funds may be budgeted in the proposal for this task.

Data Management Plans are not required at the pre-proposal stage. More details about Data Management Plans and data sharing will be sent to proposers along with letters of invitation to submit full proposals.

NEW YORK SEA GRANT: We are putting this in our RFP:

H. Data Management/Sharing Plan (NEW!)

New NOAA regulations require that data and information collected and/or created under NOAA (including Sea Grant) grants must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner (typically no later than two years after the data are collected or created), except where limited by law, regulation, policy or by security requirements. The new requirement has two basic parts: (1) environmental and socioeconomic data generated by a grant project must be made available after a reasonable period of exclusive use, and (2) the grant proposal must describe the plan to make the data available.

To comply with this new requirement, in NYSG’s full proposal stage, the principal investigator must include a data management plan that describes how the project’s data and metadata will be made available to others. Deposition of data in standard data archives (e.g., by discipline) or in available university archives is encouraged. This requirement for data archiving is in addition to the expected publication of research results in peer-reviewed journals. The proposed plan will be reviewed for compliance with NOAA requirements. If funds are needed for this task, they may be included in the project pre-proposal and full proposal budgets.

And on the pre-proposal form, we had this:

D. Data Management Plan: Indicate your recognition that an acceptable data management/sharing plan will be required as part of a full proposal, that compliance with the plan will be required if the project is funded, and that the preproposal’s budget estimate provides for this. _____ yes.

CALIFORNIA SEA GRANT: This will be included in our RFP:

NOAA Data Sharing Requirement (effective for all new NOAA funded research projects)

Because funds for our Core research program are provided by NOAA, all new Sea Grant awards will have to conform to NOAA's Directive on Data Management, available at

<https://www.nosc.noaa.gov/EDMC/PD.DMP.php> . This directive says:

Environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements.

Accordingly, data and information collected and/or created under Sea Grant grants and cooperative agreements must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner (typically no later than two years after the data are collected or created).

No action by proposers regarding this policy is required for submission of preliminary proposals. However, proposers should understand that in the Full Proposal stage, the Principal Investigator will be required to address this requirement explicitly and explain how data and metadata will be stored and provided, upon request.

SEA GRANT OMNIBUS PROPOSAL LANGUAGE

Question: Where in our omnibus application do we put our data sharing plan?

Answer(Dorn): The 2013 omnibus application instructions call for the data management plan to be included in the "Methodology" section of the 90-2 Project Summary Form of each individual project. If the Plan is lengthy and is already included in the Project Narrative, then what you put in the 90-2 (which is after all a project summary) could just be a brief summary description of the plan and directions to where the full detail can be found.

The omnibus instructions *also* call for an overall data management plan (which could be just a

summary and point of contact for all the individual project D.M.P.s in the omnibus). This overall D.M.P. could be stand-alone, but if it is, then the 90-2 for the main Management project should contain a summary of the plan and description of where the full Plan can be found. Alternatively, the main Management project could contain the full D.M.P., if it's short and simple enough for that to make sense.

Question: Can our Omnibus state rather simply: "XX Program will make an effort to ensure PIs meet the NOAA Guidance requirements for environmental data sharing as stated in?"

Answer(Dorn): I don't think it would be enough to say "XX Program will make an effort to ensure PIs meet the NOAA guidance...", for two reasons. (1) XX program is going to have to commit to following the NOAA directive, not just make an effort, and (2) XX program will need to specify how it intends to manage that data. For example, a statement in the omnibus like "All environmental data will be managed by one of the following methods: ...[list]... Each individual project description includes information on which data management option will be used."

A minimum requirement of a data management plan in an omnibus (or other) proposal would be that a person reading the proposal would know where to go to get the data. This might be in a table of all the individual projects and where or who to contact for the data, or better perhaps just a statement that one contacts the XX Program Research Coordinator with any data request. Then the RC maintains a list of his or her own, and sends any requesters where they need to go. This would be more work for the Program RC, but it would have the advantage of allowing RCs to know whether people were actually requesting data (perhaps useful later on, if NOAA tries to make the process more onerous), and to know if any PIs are improperly hoarding their data.

Question: I don't foresee us having the ability to include a technical review of that data management plan as part of our project selection process. This may be at odds with a statement in the PPT "If this is a new application for funding then your data sharing plan will be evaluated by experts in your field". When does this evaluation by experts in your field take place? During the technical review of PI proposals, or on the Omnibus proposal?

If it is to be part of the technical reviews conducted by the states what are the guidelines for the reviewers and panelists? By what metrics do we judge the plans, and is it part of our funding decision? What if our best proposals have lousy data management plans, is that sufficient cause to direct the funding to ‘less strong’ projects? (answer = no) If not, then why include it in the review?

Answer(Dorn): if you are going to accept responsibility for managing data at the Program level, then you need to include a data management plan in your omnibus application. The purpose of the information you get from the PI’s, then, is to make sure whatever you say in the omnibus is doable and consistent with what you agree to let each individual PI do. (And you don’t have to treat the data from all projects in the same way).

The relative weight you give to your own analysis of competing data management plans is up to you. If you want to say that you’ll accept any plan that meets minimum requirements (eg, it’s consistent with what you intend to write in your omnibus), that’s fine with me (Dorn). If you follow that pattern, then the real “data management plan technical review” NOAA describes will take place when your omnibus is reviewed at the national office.

Question: Our program office has neither the staff time nor resources to build a data warehousing effort that would be capable of handling the variety of environmental data (as def. by NOAA) in any meaningful way. Is it sufficient to point PIs to the NOAA guidance, and refer PIs to the archiving application system in place at NOAA Archives as one possible solution to their archiving requirement? Perhaps if they apply to the archiving process and they are accepted that would be sufficient. What if they are rejected? Is that an indication that long-term archival is less important? Most researchers keep their data for many many years, it’s really the public availability portion, and the need for maintenance of websites and such, that I see as problematic.

Answer(Dorn): Certainly the NOAA archives are an option you can point them to. The easiest data management plans to implement, I would think, are ones that don’t involve significant

changes from the way the data had been managed previously (except hoarding is no longer an option). So a PI who is used to sending his data to an archiving center might want to plan to continue doing that, and a PI who plans to publish his data in the peer-reviewed literature, or who offers to share a disk of it on request, could keep doing that. I think that data management planning will have to evolve at all levels over the next few years, in what individual PIs are expected to do, what SG Programs are expected to do to oversee data of PIs, what NSGO, and NOAA as a whole must do. But for now, I would suggest make as few drastic changes as possible.

MORE INFO

Question: I have another question. Who can I ask?

Answer(NOAA): Please email me your question at Ingrid.Guch@noaa.gov . I will either answer it or forward to someone who can.

Answer#2(Dorn): Or ask me, especially if the question is Sea Grant specific. dorn.carlson@noaa.gov. Or maybe write it into this document on Basecamp (would that work?).

Here are links to some of the original NOAA documents setting out the data management policy:

The NOAA Directive on Data Management. https://www.nosc.noaa.gov/EDMC/nao_212-15.php

NOAA “procedural directives”, including the one on grants data sharing requirements. <https://www.nosc.noaa.gov/EDMC/PD.all.php>

Here are a couple of more FAQs from agencies that have been doing data management plans for a while. Their rules are different from ours of course (NIH is much more concerned with protecting personal medical information for example), but they might touch on an issue you’re interested in that isn’t covered in our own FAQ. If you see something there and can’t figure out

how it would translate to NOAA requirements, ask Dorn or just write it into this document, and we'll try to figure it out.

NSF Data Management FAQs. <http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp>

NIH Data Sharing Policy.

http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

NOAA Data Sharing Policy for Grants and Cooperative Agreements

Ingrid Guch
NOAA Environmental Data
Management Committee



Purpose



- Provide information about the new data sharing policy to grantees, NOAA Federal Program Officers, NOAA Competition Managers and other interested parties

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.*

New FFO term



Provides Data Sharing Requirements and describes Data Information/Sharing Plan contents
Note: FFO Page limits for project narratives should be increased by 2 to accommodate the plan

New Language for NOAA Competition Announcements

Environmental data and information, collected and/or created under NOAA grants/cooperative agreements must be made visible, accessible, and independently understandable to general users, free of charge or at minimal cost, in a timely manner (typically no later than two (2) years after the data are collected or created), except where limited by law, regulation, policy or by security requirements.

1. Unless otherwise noted in this federal funding announcement, a Data/Information Sharing Plan of no more than two pages shall be required as part of the Project Narrative. A typical plan may include the types of environmental data and information to be created during the course of the project; the tentative date by which data will be shared; the standards to be used for data/metadata format and content; policies addressing data stewardship and preservation; procedures for providing access, data, and security; and prior experience in publishing such data. The Data/Information Sharing Plan will be reviewed as part of the NOAA Standard Evaluation Criteria, Item 1 -- *Importance and/or Relevance and Applicability of Proposed Project to the Mission Goals.*
2. The Data/Information Sharing Plan (and any subsequent revisions or updates) will be made publicly available at time of award and, thereafter, will be posted with the published data.
3. Failing to share environmental data and information in accordance with the submitted Data/Information Sharing Plan may lead to disallowed costs and be considered by NOAA when making future award decisions.

3

New Programmatic SAC



New NOAA Programmatic Special Award Condition: *Provides Data Sharing Requirements*

Environmental data and information, collected and/or created under this grant/cooperative agreement will be made visible, accessible and independently understandable to users, free of charge or at minimal cost, in a timely manner (typically no later than two (2) years after the data are collected or created), except where limited by law, regulation, policy or by security requirements.

- The Data/Information Sharing Plan (and any subsequent revisions or updates) will be made publicly available at time of award and, thereafter, will be posted with the published data.
- Environmental data and information produced under this award and which are made public must be accompanied by the following statement: *These environmental data and related items of information have not been formally disseminated by NOAA and do not represent and should not be construed to represent any agency determination, view, or policy.*
- NOAA may at its own discretion, use information from the Data/Information Sharing Plan to produce a formal metadata record and include that metadata in a catalogue to indicate the pending availability of new data.
- Failing to share environmental data and information in accordance with the submitted Data/Information Sharing Plan may lead to disallowed costs and be considered by NOAA when making future award decisions.

4

A Very Simple Data/Information Sharing Plan

Note: This is not an actual plan.

This project will generate environmental information. Publicly available GOES satellite data will be processed to generate maps of aerosols. Data products, along with information about the correct citation to use and the data/product limitations will be shared by December 31, 2014 along with documentation to read/use/understand the data and a point of contact for questions. This dataset will provide total column aerosol (aerosol optical depth units) for the Continental US at a 5KM resolution each day between Nov 30 2013 and Nov 29 2014. We will use [NetCDF](#) Climate and Forecast data/metadata conventions. We will publish the methodology for creating the products in the open literature. We will leave the data on the university website for at least three years. During that time period we will ask NOAA archives if they have interest/resources for archiving the data and work with them as needed. We have not had experience sharing data in the past, but we will work to set up a website similar to the format in <http://subunit.zzzuniversity.edu>. Our website will be password protected and available only to co-PIs until we determine it is beneficial to share the data more widely or our primary findings are accepted for publication or December 31, 2014, whichever comes first. If resources allow we will also include current daily aerosol maps as gif images on the website for real-time users.

6

Sharing-related items to minimize/prevent

- Grantees and research institutions being scooped
- Intellectual property misunderstandings
- Users confused or wasting time with data not fully quality controlled
- Users believing data to be more accurate than it is and making conclusions or decisions

7

Information Quality Act



- Environmental data and information produced under this award and which are made public must be accompanied by the following statement:
These environmental data and related items of information have not been formally disseminated by NOAA and do not represent and should not be construed to represent any agency determination, view, or policy.
 - In order to remove this disclaimer NOAA must verify that the data meet NOAA Information Quality Act guidelines and approve dissemination to the public.

7

What if data is to be shared after award closes?



- The NOAA FPO will verify the data sharing plan is available and being followed prior to closing the award file and take enforcement actions if necessary (NOAA could seek a post-award disallowance of costs if data is not shared per the data sharing plan)
- PIs who will seek future NOAA funds will be motivated to follow their data sharing plan even after the award has closed because they will need to indicate how and when they have made their data accessible and usable by the community in the past. This will be reviewed as part of the NOAA Standard Evaluation Criterion “Importance and/or relevance and applicability of proposed project to the mission goals.”

8

What if I don't have a logical place to make my data and/or data sharing plan widely available?



- Propose something reasonable.
 - If this is a new application for funding then your data sharing plan will be evaluated by experts in your field and they should understand the logistics that make sharing difficult.
 - If this is not a new application, consult with your NOAA FPO on ways to comply.

9

Question: What is meant by "environmental data"?

Answer: 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.

Question: What is meant by "sharing"?

Answer: 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.

Question: What is considered "timely" data sharing?

Answer: This will depend on the program awarding the grant or cooperative agreement, and the nature of the research project conducted. It is typically no later than two (2) years after the data are collected or created.

Question: What is meant by "independently understandable"?

Answer: The data must be accompanied with documentation, metadata and, if needed, tools to read the data that allow a user to interpret the data properly. If there are concerns with understandability, they can be reported to NOAA, who will do an independent check.

10

Question: Who will determine if my data are visible, accessible and independently understandable?

Answer: The person generating the data will have first responsibility for determining this. Common data quality standards in your scientific discipline may help you decide if the data are understandable. Ultimately, others who use your data will know whether they are visible, accessible and understandable to them. If there are concerns with data access or understandability, they can be reported to NOAA, who will do an independent check.

Question: What are examples of law, regulation, policies or security requirements that may limit my ability to share data?

Answer: Policies applicable to protection of personally identifiable information, critical infrastructure information or proprietary trade information as well as regulations related to export control may impact your ability to share data, among other items.

Question: How must data be shared?

Answer: This depends on the nature of the project and the data, and will be proposed by the investigator himself. Data sharing can be accomplished through:
Data Archive: place where data are acquired, manipulated, documented, and distributed. NOAA facilities that archive data and make the data openly available should be considered.
Data Enclave: controlled, secure environment in which eligible researchers can perform analyses using data resources
Publishing: articles in scientific publications
Researcher's Efforts: investigator responds directly to data requests (e.g., posting data on a Web site)

11

Question: I'm a busy investigator. I don't have time to process requests for my data. What should I do?

Answer: In addition to publishing small datasets, there are several alternatives to responding to each separate request to share data (e.g., putting data in an archive or restricted access facility, and setting up a web site for data access). Archives and data enclaves provide technical assistance for users with questions or problems and may spare busy investigators time.

Question: Can I get additional funding to share my data?

Answer: Unless otherwise noted in the federal funding announcement, funding to address data sharing must be requested as part of the proposal to collect/create data. The data sharing plans and related funds requested should consider the anticipated benefit of the data, the likely number of interested users of the data and the priorities of the program as outlined in the solicitation.

Question: I am the PI of a large [Cooperative Institute|Sea Grant/similar] program funded by an omnibus grant which in turn manages a number of individual research projects. Must every individual project have its own data sharing plan, or can I develop a program-wide data sharing plan?

Answer: As the omnibus grant recipient, you have a responsibility to see that data sharing plans are followed for all research projects under your program. This may be done with a single Program-wide data sharing plan, individual plans for individual projects, or something in between, as long as all the relevant data generated is covered under some data sharing plan.

12

Question: What web resources are available to help me do this and obtain more information?

Answer: There is information available at the NOAA Environmental Data Management Committee website reachable from www.nosc.noaa.gov. In general considering data sharing requirements prior to finalizing the methods for collecting/creating/storing the data will save time and effort later on. Unless otherwise noted in the federal funding announcement there is no specific data sharing plan template required.

Question: The NOAA Program I apply to already requires an extensive Data Management Plan. Do I still need to do a Data Sharing Plan?

Answer: Not necessarily, refer to the specific NOAA Program federal funding announcement to determine if a Data Sharing Plan is needed.

Question: I am not collecting any data/information. What should I do?

Answer: A statement indicating you are not collecting any data/information will be appropriate for your data sharing plan.

Question: I have another question. Who can I ask?

Answer: Please email me your question at Ingrid.Guch@noaa.gov. I will either answer it or forward to someone who can.

Thank-you for your support of the
new NOAA data sharing policy.

FREQUENTLY ASKED QUESTIONS ON DATA MANAGEMENT PLANS

from an NSF website, <http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp>

(NSF's rules are of course different from NOAA's, but you might want to look through this to get a general sense of the answers being provided by NSF, who has been requiring these plans for a couple of years now.)

1. What constitutes "data" covered by a Data Management Plan?

What constitutes such data will be determined by the community of interest through the process of peer review and program management. This may include, but is not limited to: data, publications, samples, physical collections, software and models.

2. Is a plan for Data Management required if my project is not expected to generate data or samples?

Yes. It is acceptable to state in the Data Management Plan that the project is not anticipated to generate data or samples that require management and/or sharing. PIs should note that the statement will be subject to peer review.

3. Am I required to deposit my data in a public database?

What constitutes reasonable data management and access will be determined by the community of interest through the process of peer review and program management. In many cases, these standards already exist, but are likely to evolve as new technologies and resources become available.

4. There is no public database for my type of data. What can I do to provide data access?

Contact the cognizant NSF Program Officer for assistance in this situation.

5. Should the budget and its justification specifically address the costs of implementing the Data Management Plan?

Yes. As long as the costs are allowable in accordance with the applicable cost principles, and necessary to implement the Data Management Plan, such costs may be included (typically on Line G2) of the proposal budget, and justified in the budget justification.

6. My institution's policy is that the data and all supporting materials from all research are owned and must remain with the institution if I leave. How does this policy affect what I can say about data management and access?

Data maintenance and archiving by an institution is one avenue by which data preservation and access can be achieved. However, the data access plan must address the institutional strategy for providing access to relevant data and supporting materials.

7. Does data management and access include supporting documentation and metadata, such as validation protocols, field notebooks, etc.?

All researchers are expected to be able to explain and defend their results. Doing so usually entails maintaining complete records of how data were collected. The manner in which one maintains such records and makes them available to others will vary from project to project. What constitutes reasonable procedures will be determined by the community of interest through the process of peer review and program management. These standards are likely to evolve as new technologies and resources become available.

8. How long should data be archived and made accessible?

What constitute reasonable procedures will be determined by the community of interest through the process of peer review and program management.

9. Does this policy mean that I must make my data available immediately, even before publication?

Not necessarily. The expectation is that all data will be made available after a reasonable length of time. However, what constitutes a reasonable length of time will be determined by the community of interest through the process of peer review and program management.

10. What are NSF's expectations regarding the release of data that include sensitive information (e.g., information about individuals or locations of endangered species)?

Such data must be maintained and released in accordance with appropriate standards for protecting privacy rights and maintaining the confidentiality of respondents. Within legal constraints, what constitutes reasonable data access will be determined by the community of interest through the process of peer review and program management.

11. My data include information of potential commercial value. Am I required to make that information available?

Not necessarily. It is NSF's strong expectation that investigators will 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. However, it is also necessary to protect intellectual property rights and potential commercial value. The Data Management Plan should describe the proposed approach, which will then be subject to peer review and program management. (For example, research use of sensitive data is often allowed through reasonable binding agreements that contain confidentiality provisions.)

12. Does NSF have particular requirements for archiving and accessibility of samples, physical collections and so forth?

No. If appropriate, your Data Management Plan should describe the types of samples, and/or collections, etc., that you will use, as well as personal, institutional or other repositories for archiving and providing access to others. What constitutes reasonable archiving and accessibility will be determined by the community of interest through the process of peer review and program management.

13. Does NSF have particular requirements for what types of samples, physical collections, and so forth should be saved?

No. What constitutes reasonable requirements will be determined by the community of interest through the process of peer review and program management. These standards are likely to evolve as new technologies and resources become available.

14. If data or samples are requested before I have completed all analyses on them, must I share them?

No. The expectation is that all data will be made available after a reasonable length of time. One standard of timeliness is to make the data or samples accessible immediately after publication. However, what constitutes a reasonable length of time will be determined by the community of interest through the process of peer review and program management

15. How does this policy relate to the issue of open access publishing?

Open-access publishing (making all published articles freely available) is a separate issue that is not addressed in the implementation of the data management plan requirement.

16. If I participate in a collaborative international research project, do I need to be concerned with data management policies established by institutions outside the United States?

Yes. There may be cases where data management plans are affected by formal data protocols established by large international research consortia or set forth in formal science and technology agreements signed by the United States Government and foreign counterparts. Be sure to discuss this issue with your sponsored projects office (or equivalent) and your international research partner when first planning your collaboration.

17. My proposal is interdisciplinary and there are multiple sets of guidance to follow on NSF's website (<http://www.nsf.gov/bfa/dias/policy/dmp.jsp>), which one do I follow?

All proposals are submitted to a lead program, with the option to specify other programs that the PI would like to consider the project. If the guidance appropriate to the lead program applies, it should be followed. Otherwise, provide a clear explanation of what you would do if the project were funded. Bear in mind that the merit review is conducted by colleagues from the communities of relevance and that your plan should be convincing to them. It should also be commensurate with the level of effort requested and required, and appropriate to the long-term value and benefit to your colleagues of any generated data products.

FREQUENTLY ASKED QUESTIONS ON DATA MANAGEMENT PLANS

from an NIH website, http://grants.nih.gov/grants/policy/data_sharing/data_sharing_faqs.htm

(NIH's rules are of course different from NOAA's. They need to worry a lot more about patient confidentiality, for example. But you might want to look through this to get a general sense of the answers being provided by NIH, who has been requiring these plans for several years now.)

A. March 5, 2003

1. **Why should I share my final research data?**

Data sharing achieves many important goals for the scientific community, such as

- reinforcing open scientific inquiry
- encouraging diversity of analysis and opinion,
- promoting new research, testing of new or alternative hypotheses and methods of analysis
- supporting studies on data collection methods and measurement
- facilitating education of new researchers
- enabling the exploration of topics not envisioned by the initial investigators
- permitting the creation of new datasets by combining data from multiple sources.

2. **Who benefits from data sharing?**

Everyone benefits, including investigators, funding agencies, the scientific community, and, most importantly, the public. Data sharing provides more effective use of NIH resources by avoiding unnecessary duplication of data collection. It also conserves research funds to support more investigators. The initial investigator benefits, because as the data are used and published more broadly, the initial investigator's reputation grows.

3. **Is data sharing 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 Justice. Moreover, many scientific journals require that authors make available the data included in their publications. In the biological sciences, protein and DNA sequences are made available to researchers through data archives, such as GenBank. Since 1996, NIH has required data sharing in several areas, such as DNA sequences, mapping information, and crystallographic coordinates.

4. **What do you mean by final research data?**

By "final research data", we mean recorded factual material commonly accepted in the scientific community as necessary to validate research findings. Final research data do not include laboratory notebooks, partial datasets, preliminary analyses, drafts of scientific papers, plans for

future research, peer review reports, communications with colleagues, or physical objects, such as gels or laboratory specimens.

5. Does "final research data" include data that were not originally produced under an NIH grant or contract?

Sometimes. For example, where NIH support is sought to transform or link datasets (as opposed to producing new data), the investigator should include a data-sharing plan in the application.

6. What do you mean by unique data?

By "unique data" we mean data that cannot be readily replicated. Examples of studies producing unique data include: large surveys that are too expensive to replicate; studies of unique populations, such as centenarians; studies conducted at unique times, such as a natural disaster; studies of rare phenomena, such as rare metabolic diseases.

7. What kinds of data are candidates for sharing?

Potentially all kinds of data are candidates for sharing, but unique data are especially important. Some biologic sciences already have data-sharing plans in place, such as genetic mapping. But other basic science data are also amenable to sharing. Data from human subjects (e.g., surveys, clinical studies) also can be shared if the identity and privacy of research participants can be protected.

8. Can you give me some examples of data that have been shared?

Examples of shared epidemiologic data include the Framingham Heart Study, the Honolulu Heart Program, the Atherosclerosis Risk in Communities, Epidemiology of Chronic Disease in the Oldest Old, and the Iowa 65+ Rural Health Study. Examples of shared data from clinical trials include the Asymptomatic Cardiac Ischemia Pilot, the Intermittent Positive Pressure Breathing Study, and the Safety and Efficacy Trial of Zidovudine for Asymptomatic HIV Infected Individuals. Examples of shared datasets from the basic sciences include a growing number of genome sequences and maps, as well as protein and nucleotide databases (see ENTREZ <http://www.ncbi.nlm.nih.gov/Database/index.html> and other resources for molecular biology at the National Center for Biotechnology Information at <http://www.ncbi.nlm.nih.gov>)

9. Data from my studies are generated from a very small number of rats, and I publish the final data. Am I expected to provide these data to other investigators as well?

Publishing these final data constitutes an acceptable mechanism for sharing data.

10. How soon after data collection am I obliged to share the final data?

Recognizing that the value of data often depends on their timeliness, data sharing should occur in a timely fashion. NIH expects the timely release and sharing of data to be no later than the acceptance for publication of the main findings from the final dataset. This time point will be influenced by the nature of the data collected. Data from small studies can be analyzed and submitted for publication relatively quickly. If data from large epidemiologic or longitudinal studies are collected over several discrete time periods or waves, data should be released in waves as data become available or main findings from waves of the data are published. NIH recognizes that the investigators who collected the data have a legitimate interest in benefiting from their investment of time and effort. NIH continues to expect that the initial investigators may benefit from the first and continuing use, but not from prolonged exclusive use. While NIH also understands that an institution's desire to exercise its intellectual property rights may justify a need to delay disclosure of research findings, a delay of 30 to 60 days is generally viewed as a reasonable period for such activity.

11. Does data sharing pertain only to published data?

No. Data-sharing plans should encompass all data from funded research that can be shared without compromising individual subjects' rights and privacy, regardless of whether the data have been used in a publication. Furthermore, data sharing prior to the publication of major results is encouraged in many instances, for example, when data are collected to provide a resource for the scientific community (as in the case of many large surveys).

12. Due to circumstances beyond my control (an earthquake!), I was unable to recontact a substantial portion of the sample in my longitudinal study. I was planning to put my data in an archive, but the resulting high rate of attrition makes the data minimally useful. Should I still archive the final dataset?

Investigators need to find a balance between the value of the final data and the costs associated with archiving. If the data are of limited usefulness, then it is probably not worth the expense and effort of putting them in an archive. However, if the investigator has published results based on this dataset, then the dataset should be shared.

13. I am preparing an SBIR application. Am I required to submit a data-sharing plan?

Yes. The specific nature of the data you will collect will determine whether or not you may share the final dataset. If the final data are not amenable to sharing, for example, if they are proprietary, then you need to explain this in your application. Under the Small Business Act, SBIR grantees may withhold their data for 4 years after the end of the award. The Small Business Act provides authority for NIH to protect from disclosure and nongovernmental use all SBIR data developed from work performed under an SBIR funding agreement for a period of 4 years after the closeout of either a Phase I or Phase II grant unless NIH obtains permission from the awardee to disclose these data. The data rights protection period lapses only upon expiration of the protection period applicable to the SBIR award, or by agreement between the small business concern and NIH.

14. I don't want to share my data, which were generated under an NIH grant. Can I be forced to do so?

When the PI and the authorized institutional official sign the face page of an NIH application, they are assuring compliance with policies and regulations governing research awards. NIH expects grantees to follow these rules and to conduct the work described in the application. Thus, if an application describes a data sharing plan, NIH expects that plan to be enacted. In some instances, for example, NIH may make data sharing a term and condition of award.

Under specific circumstances, your data also may be accessible through the Freedom of Information Act (FOIA). If your competitive grant was awarded after April 17, 2000 and if your data were cited in a Federal regulation or administrative order, then your data may also be accessible through FOIA. (See

http://grants.nih.gov/grants/policy/a110/a110_guidance_dec1999.htm).

15. Will the data-sharing plan affect the priority score of my application?

No. Reviewers will not factor the proposed data-sharing plan into the determination of scientific merit or priority score. Program staff is responsible for overseeing the data-sharing policy and for assessing the appropriateness and adequacy of the proposed data-sharing plan. Program concerns must be resolved prior to making any award.

16. My research, which seeks support from both the public and private sectors, will involve proprietary data. How do I deal with the data-sharing issue in my application?

NIH recognizes that there may be circumstances where a cofunder has requested restrictions on data sharing as a condition of funding. These restrictions should be identified in the application

and a proposal made about how data from the cofunded project will be shared. Should you believe that you are unable to share any of the data, your justification will be considered by NIH program staff.

17. I'm a busy investigator. I don't have time to process requests for my data. What should I do?

In addition to publishing small datasets, there are several alternatives to responding to each separate request to share data (e.g., putting data in an archive or restricted access facility, and setting up a web site for data access). Archives and data enclaves provide technical assistance for users with questions or problems and may spare busy investigators time.

18. Can I share data with colleagues under my own auspices?

Yes. Your data-sharing plans should indicate the criteria for deciding who can receive your data and whether or not you will place any conditions on their use. Data should be made as widely and freely available as possible while safeguarding the confidentiality of the data and privacy of participants. You should not place limits on the questions or methods others might pursue nor should you require co-authorship as a condition for receiving the data.

19. Should the data source be cited or acknowledged in papers that rely on shared data?

It is appropriate to acknowledge the source of data upon which a manuscript is based. Many investigators include this information in the methods and/or reference sections of their manuscripts. Journals generally include an acknowledgement section, in which the authors can recognize people who helped them gain access to the data. However, you should check the policies of the journal to which you plan to submit.

20. Should I consider contributing my research data to a data archive?

Maybe. Archives are organizations that collect and distribute data. They understand what is needed to prepare data for wider distribution and documentation for users. They provide stable, reliable, and cost-effective means for distributing data. They also provide protections for the dataset and technical assistance for requestors.

21. Where can I find guidance on preparing data for sharing and archiving?

Guidance is available from a variety of sources. For example, the Inter-University Consortium for Political and Social Research at the University of Michigan has prepared an excellent set of guidelines for preparing data for archiving. While these guidelines were written with social science data in mind, they are broadly applicable. See

<http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf> 

For molecular biology information, the National Center for Biotechnology Information (NCBI), a division of the National Library of Medicine (NLM) at the National Institutes of Health, is ready to assist researchers who have genome-specific and molecular data to submit. For more information about submitting and accessing NCBI data, see the NCBI Website at <http://www.ncbi.nlm.nih.gov/Genbank/index.html>.

22. How do I pay for preparing data for sharing and archiving?

NIH recognizes that it takes time and money to prepare data for sharing. You can request funds for data archiving and sharing as part of your grant application for collecting the data. If you have already collected the data, you may want to ask your NIH Project Officer about a competitive or administrative supplement. NIH recommends that you consider procedures and costs for data sharing during the application process rather than after the data have been collected.

23. Should I address data sharing in my NIH application?

Yes. By the October 1, 2003 application receipt date, NIH requests that all extramural applicants seeking \$500,000 or more in direct costs in any one year provide a data-sharing plan in their applications.

24. What do I need to include in my application and where do I put the information about data sharing?

Scientists submitting grant, cooperative, or contract applications should include a data-sharing plan, or provide justification for the absence of such a plan, in a brief paragraph to be placed immediately after the Research Plan Section (i.e., immediately after PHS 398 Section I. Letters of Support in the Research Plan Section of their application) so it does not count toward the application page limit. Additional information on data sharing might be included in other sections of the application, as appropriate. For example, if you are producing a large dataset that will become an important resource for the scientific community, you probably want to mention this in the significance section. If you are requesting funds to prepare, document, and archive the data, you would want to include relevant information in the budget and budget justification sections. In the Human Subjects section of the application, you should discuss the potential risks to research participants posed by data sharing and steps you will take to address those risks.

25. The informed consent form for my recently completed study states explicitly that only my research team will see the data provided and that we will not share the data. Am I now expected to share it?

No, but if you plan to collect additional data from those subjects under a grant with a data-sharing plan, you should revise the consent procedure to be consistent with the data-sharing plan. In preparing and submitting a data-sharing plan during the application process, investigators should avoid developing or relying on consent processes that promise research participants not to share data with other researchers. Such promises should not be made routinely or without adequate justification described in the data-sharing plan.

26. How can I protect the privacy of my subjects?

It is the responsibility of the investigators, their IRB, and their institution to protect the rights of participants and the confidentiality of their data. Data should be redacted to strip all individual identifiers, and effective strategies should be adopted to minimize risk of disclosing a participant's identity. Options to protect privacy include: withholding part of the data, statistically altering the data in ways that will not compromise secondary analyses, requiring researchers who seek data to commit to protect privacy and confidentiality, and providing data access in a controlled site, sometimes referred to as a data enclave. Some investigators use hybrid methods, releasing a redacted dataset for general use but providing access to more sensitive data through a user contract or data enclave. In most instances, sharing data is possible without compromising participant confidentiality and privacy.

27. Can institutions and investigators subject to the Federal Health Insurance Privacy and Portability Act (HIPAA) Privacy Rule share data in accord with the NIH Data Sharing policy?

Yes. NIH recognizes that data sharing may be complicated or limited, in some cases, by institutional policies or local IRB rules, as well as by local, state and Federal laws and regulations like the Privacy Rule. To protect the rights and privacy of people who participate in NIH-sponsored research, data intended for broader use should be free of identifiers that would permit linkages to individual research participants, and exclude variables that could lead to deductive disclosure of the identity of individual subjects. When data sharing is limited, applicants should explain such limitations in their data sharing plans.

28. I collect data on sensitive and, sometimes, illegal behaviors. Are these data too sensitive to be shared?

Not necessarily. The collection of sensitive data does not preclude sharing. For example, the National Center for Chronic Disease Prevention and Health Promotion at CDC operates the Youth Risk Behavior Surveillance System (YRBSS), available at <http://www.cdc.gov/nccdphp/dash/yrbs/>, which provides data on six health risk behaviors among youth: unintentional injuries and violence, tobacco use, alcohol and other drug use, sexual behaviors, dietary behaviors, and physical activity. Similarly, data from the National Survey of Family Growth, which includes statistical data on family life, marriage and divorce, contraception, sexual experience, pregnancy, and infertility, can be obtained from the National Center for Health Statistics.

Sensitive data can be shared so long as appropriate privacy safeguards are in place. Investigators must determine if and how the rights and privacy of the subjects can be protected. And investigators collecting data on sensitive and illegal behaviors should obtain a Certificate of Confidentiality (<http://grants.nih.gov/grants/policy/coc/>) to protect against the involuntary release of data that could identify research participants.

29. Can data from a clinical trial be shared?

It depends. Participants' privacy must be protected in accord with all applicable laws and regulations. Clinical trial datasets are frequently rich in items that could potentially identify individual subjects. For example, many early phase trials use small samples, which make it difficult to protect the privacy of the participants. Researchers who are planning clinical trials and intend to share the resulting data should think carefully about the study design, the informed consent documents, and the structure of the resulting data prior to the initiation of the study.

There are many precedents for sharing of clinical trial data. For example, data from a number of clinical trials supported by the National Heart, Lung, and Blood Institute (NHLBI) are available for research use (See <http://www.nhlbi.nih.gov/resources/deca/directry.htm>). The National Institute of Allergy and Infectious Diseases (NIAID) also lists their clinical trials datasets that they have made available through the National Technical Information Service (NTIS) for public use (See <http://www.niaid.nih.gov/research/aidsdata.htm>).

30. Is data on DNA and protein sequences archived?

Yes. For example, GenBank (<http://www.ncbi.nih.gov/Genbank/>) and Entrez (<http://www.ncbi.nlm.nih.gov/Entrez/>) archive gene sequencing data. The sharing of materials, data, and software in a timely manner has been an essential element in the rapid progress that has been made in the genetic analysis of mammalian genomes.

31. I did not request support for sharing data in my application, which was funded. Can I charge requestors for the costs associated with sharing the data?

Yes, as long as such costs are reasonable and not excessive and reflect actual costs associated with complying with the request. These expenses for preparing and shipping the data might include costs of personnel, computing time, supplies, and other directly related expenses. NIH requirements for accountability for various types of income under NIH grants are specified elsewhere, see http://grants.nih.gov/grants/policy/nihgps_2003/NIHGPS_Part8.htm#_Toc54600138

32. I am working on a select pathogen and cannot share the data for reasons of national security. Is this an acceptable reason for not sharing?

Yes.

33. If I am required to submit a revised data-sharing plan, what do I need to do?

As is the case with PIs who submit any additional or revised application material, your revised data-sharing plan must be signed by your institutional official and by you.

34. I want to request a dataset from a recent publication. How do I do this?

You should check the publication to see if reference is made to an archive, an enclave, or a Website where the data might be available. If no such information is provided, you may wish to send a letter to the PI to see if the data are available for sharing, and where you might be able to get the data and associated documentation.

B. February 16, 2004

1. I am a PI on a P30 center grant with a budget in excess of \$500,000 (direct costs) in each year. Some of the research projects that collect survey data benefit by the infrastructure support provided by the P30 but these research projects are not funded by NIH. Am I still expected to share data from these research grants?

If any NIH support (i.e., partial support) is provided for resource development, even if those research resources were developed primarily with non-NIH funds, then those research resources must be shared in line with NIH policy as if NIH funded the entire project. It should be emphasized that although a data sharing plan is only required of grants awarding direct costs of \$500,000 or more in any one year, data sharing itself (without a specific plan submission) continues to be a requirement of all NIH-funded grants. If the P30 maintains core resources that actually house and are the final repository of the data, e.g., a high throughput array analysis core, then any project using the center's resources would be subject to the center's data sharing plan.