ELSI Research and the New NIH Data Management and Sharing (DMS) Policy

May 23, 2023

Elena Ghanaim, Dave Kaufman,
Rene Sterling, Nicole Lockhart, Imani McGregor, Bri Foster

NIH NHGRI

1

The goal of today is to discuss the new Data Management and Sharing (or DMS) Policy, especially as it relates to projects funded by the NHGRI ELSI program.

NHGRI Staff

Elena Ghanaim

Dave Kaufman

Rene Sterling

Nicole Lockhart

NIH NHGRI

2

- Elena Ghanaim, M.A., NHGRI's Policy Advisor for Data Science and Sharing.
- Dave Kaufman, Ph.D., Rene Sterling, Ph.D., M.H.A., and Nicole Lockhart, Ph.D., Program Directors for the NHGRI ELSI Research Program

Outline

1. Basics of the DMS Policy
2. Preparing and Submitting a DMS Plan
3. Resources
4. Discussion

NIH NHGRI

Basics of the Data Management and Sharing Policy

NIH
NHGRI

The overall spirit of the Data Management and Sharing Policy is to promote the responsible sharing of scientific data. Responsible data management and sharing has many benefits, including accelerating the pace of biomedical research, enabling validation of research results, and providing accessibility, for example to more junior researchers, to high-value datasets.

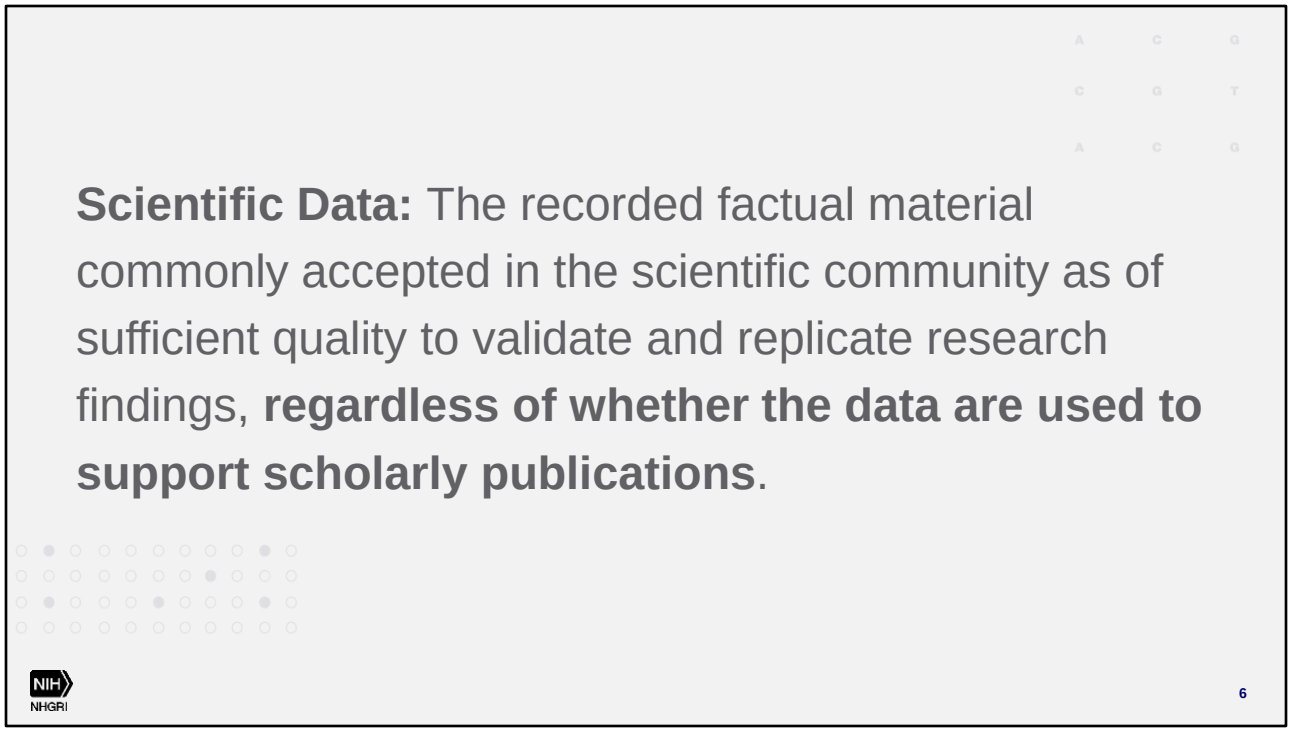
The DMS Policy is the latest policy aimed at achieving NIH's longstanding commitment to making the results of NIH-funded research available. It applies to the vast majority of NIH-funded activities, and more specifically, applies to research that generates scientific data**.**

If you're familiar with NIH's data sharing policies of the past, this Policy applies regardless of the scientific data type and regardless of the funding amount.

Therefore, nearly all applicants for NIH funding are now required to submit a Data Management and Sharing Plan outlining how scientific data and any accompanying metadata will be managed and shared, **taking into account any limitations or other restrictions.**
This Plan will be a component of the application, and **will be reviewed programmatically**.

If an application is funded, then compliance with the approved Plan becomes a term and condition of award.

NIH's website, https://sharing.nih.gov, contains a variety of helpful information and resources, including some sample DMS Plans. NHGRI recently contributed a sample Plan for survey and interview data to this website, which may be particularly helpful to this audience. https://sharing.nih.gov also covers the GDS Policy requirements. If are proposing to generate large-scale genomic data over the course of your study, then that Policy would also apply to your research.

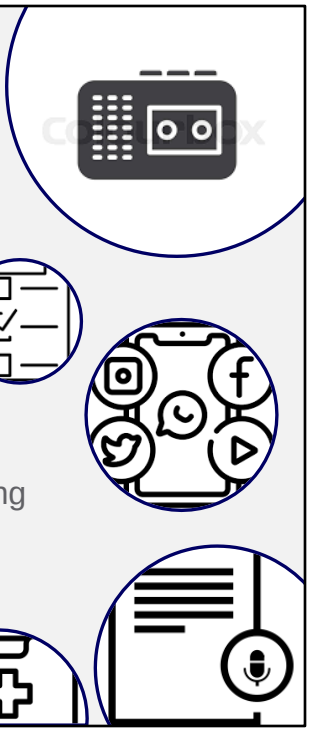


Scientific Data are the recorded factual material commonly accepted in the scientific community as of *sufficient quality to validate and replicate research findings*, regardless of whether the data are used to support scholarly publications.

Scientific data do not include things such as: laboratory notebooks, preliminary analyses, plans for future research, communications with colleagues, or physical objects, such as laboratory specimens.

Examples of ELSI scientific data

- Voice recordings
- Raw and coded transcripts from interviews/focus groups, stakeholder meetings, Delphi processes, deliberative events
- Data from social media, websites, public records
- Raw and recoded survey responses
- Outcomes data from interventions
- Ethnographic observations
- Data extracted from source articles in a meta-analysis, scoping review, lit review
- Documents and data arising from materials used in a legal, conceptual or normative analysis
- Data from participants' health care records



NIH NHGRI

Since today's webinar is specifically geared toward our ELSI researcher community, our PDs, based on review of prior awards, created a list of some of the data types that appear in ELSI work and would likely be considered "scientific data."

Two things that are important to emphasize: (1) this Policy only applies to new applications, submitted on or after the effective date of January 25, 2023. If you currently have an NIH grant, this does not change the expectations for your active award. (2) The Policy expects that, when planning for data sharing, researchers will maximize the sharing of data as appropriate.

What kinds of (ELSI) applications need to comply?

- Most NIH funding mechanisms
- Includes K01s, K99s, R00, R03, R21, R01, U-grants most other research grants
- Does NOT include:
  - T32 and fellowship (e.g., F99, F31) grants
- Administrative supplements are subject to the policies governing the parent award.

NIH NHGRI

9

This slide has been tailored to some of the common Activity Codes that NHGRI uses to support ELSI research. NIH has a comprehensive listing of the activity codes to which the DMS Policy does and does not apply (https://sharing.nih.gov/sites/default/files/flmngr/List-of-Activity-Codes-Applicable-to-DMS-Policy.pdf). Given that the Policy applies to any NIH-funded activities that generate scientific data, it applies to most (but not all) NIH funding mechanisms.

The Policy does apply to K01s, K99s, R awards, U-grants, etc.

It does not apply to the mechanisms that support training or education of individuals, such as the T32 or F99/F31 fellowship grants.

Administrative supplements will follow the requirements of the parent award. Therefore, if your original award was issued before the effective date of the DMS Policy, it will not apply to administrative supplements for the award. However, if you are seeking a supplement, including a competing revision, that changes a parent award's approved approach to data management and sharing for which the DMS Policy did apply, the DMS Plan of the parent award should be updated.

Factors to Consider: Model of Access

Sharing through **controlled access** if…

- Explicit limitations on sharing
- Data pose particular risks to participants or groups
- Data cannot be sufficiently de-identified
- Unanticipated risks are discovered after initial planning

Sharing through **unrestricted access** if….

- De-identified
- Institutional review finds unrestricted sharing poses very low risk
- Explicitly consented for an unrestricted model of access

NIH NHGRI

Considerations for protecting participant privacy (not mentioned on this slide) are: de-identifying data prior to sharing, establishing data sharing and use agreements, and understanding legal protections that are in place to ward against disclosure and misuse (such as certificates of confidentiality).

This slide goes into detail about one of the other factors to consider: the appropriate means of access for a given dataset.

The DMS Policy expects researchers to consider whether access to scientific data from participants should be controlled (which is when there are measures in place such as requiring data requesters to verify their identity and the appropriateness of the proposed research use to access protected data). This is even if data are de-identified to a regulatory standard.

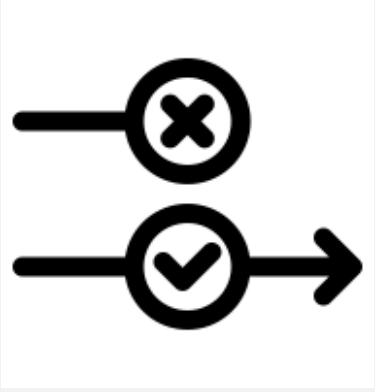
A controlled access model can be used in cases where there are explicit limitations on subsequent use, as imposed either by laws, regulations, or policies, or by informed consent or other agreements. Although having explicit limitations on use is not a prerequisite for this mode of sharing. It can be used in instances where data, even when de-identified to established standards, cannot sufficiently reduce the possibility of re-identification. If, due to previously unanticipated approaches or technologies, it is found that there are risks to participant privacy; updates to the DMS Plan can and should be communicated to the NIH immediately.

Participant data can in certain instances, be shared in an unrestricted manner. For instance, when the dataset can be de-identified and the institutional review finds sharing poses very low risk to privacy of individuals in the study. Also, in instances where participants were explicitly consent for unrestricted access.

NIH asks investigators to consider the various options for protecting participant privacy when crafting their DMS Plan.

Justifiable limitations on sharing

- Consent
- Privacy or safety would be compromised
- Federal, state, local, Tribal law or policy
- Existing or anticipated agreements
- Other technical, ethical, legal factors



NIH NHGRI

11

A common question we've received is whether *all* data must be shared under this new Policy.

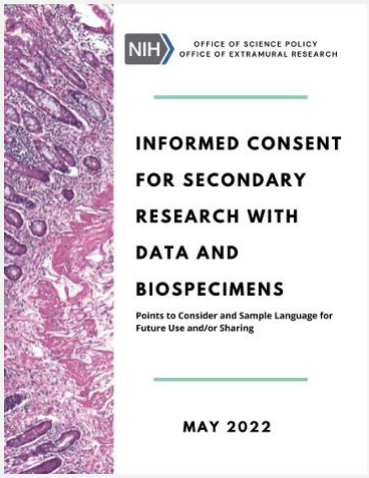
The Policy (and NHGRI) encourages and expects that data will be maximally shared. However, there is flexibility to allow for instances in which it is most appropriate to limit data sharing (either entirely, or by employing a controlled access model and enforcing use restrictions). There may be cases where no individual-level data, will be shared due to privacy or other concerns such as stigmatization or participant community concerns.

The Policy itself calls out several categories of justifiable limitations to sharing scientific data (listed on the slide).

When the proposal is for sharing to be limited in some manner,

or the investigators believes its most appropriate not to share the data, this should be described and justified in the DMS Plan. Investigators should consider whether de-identification, controlled access, Certificates of Confidentiality, or other protections alleviate certain concerns, but these must be balanced with the other considerations listed on this slide.

Informed Consent



- DMS Policy encourages researchers to plan for how data management and sharing will be addressed
- **NIH Resource provides sample language, which can be adjusted**
- Broad consent not required
  - Researcher and IRB to determine appropriate scope of sharing
- Building respectful partnerships with American Indian/Alaska Native communities

NIH NHGRI

12

The DMS Policy did not add any expectations for informed consent; but it does "strongly encourage researchers to plan for how data management and sharing will be addressed in the informed consent process, including communicating with prospective participants how their scientific data are expected to be used and shared."

NIH has published a new resource for informed consent for secondary research with data and biospecimens that provide some sample language. This may be a helpful starting point if you're going to be addressing this notion of data sharing in your consent forms for the first time or perhaps in a new way as a result of this Policy. This can and should be adjusted as needed.

Also, it’s important to note that Broad Consent, which is an approach for consent that permits current and future access

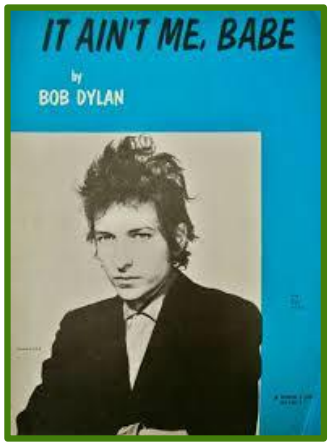
and use of samples or data for research without necessarily specifying what the focus of such studies might be, is not a requirement of the NIH data sharing policies; it is up to the researcher and their IRB to determine whether any limitations on sharing would be appropriate and to communicate these to participants via the informed consent process.

Additionally, we encourage researchers who may be engaging and including Tribal communities to understand the relevant considerations and best practices for developing partnerships with American Indian/Alaska Natives from the outset of a research project, including around the plans for data sharing. NIH has also created a resource on this to support researchers in this endeavor.

Preparing and submitting a Data Management & Sharing Plan and budget in a grant

NIH NHGRI

No data being collected or T32 or F-series grant?



Research grant collecting scientific data?

NIH National Institutes of Health

If you are not sure whether you need a data sharing plan for a given application, first check the Notice of Funding Opportunity that you'll be applying to.

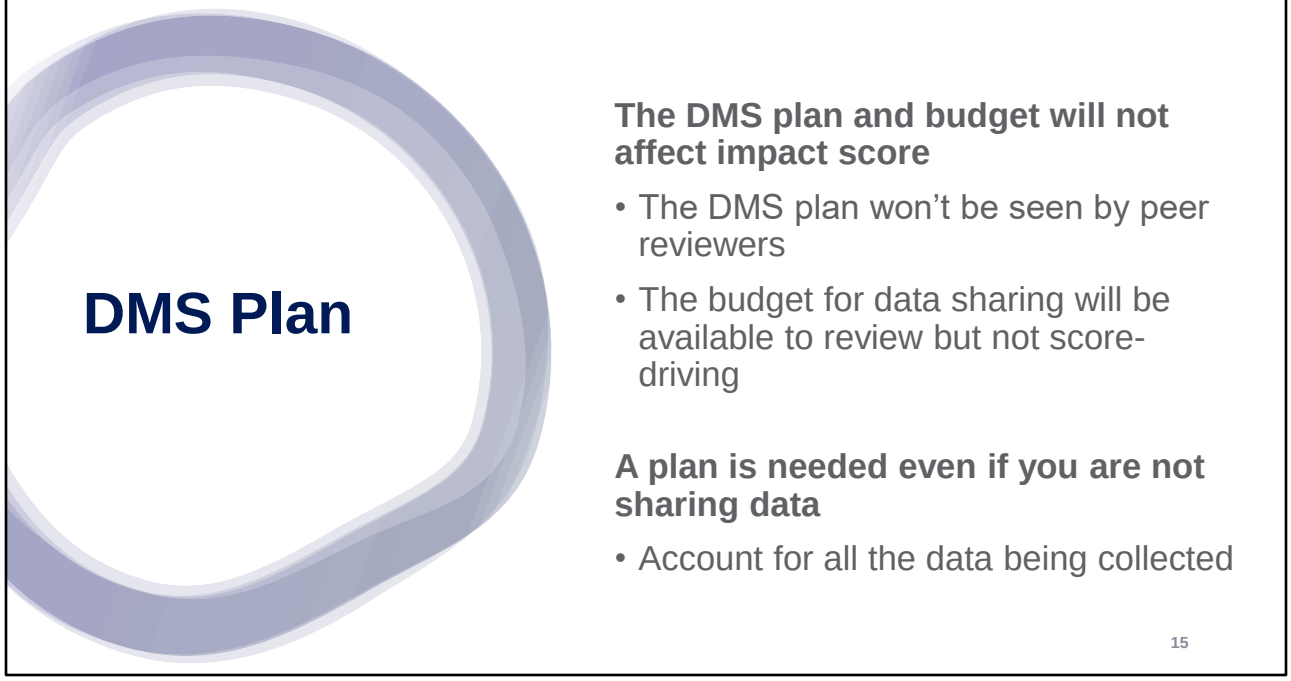
Generally speaking only 2 types of ELSI applications don't need a Plan:

- applications where no scientific data are being collected do not need a DMS Plan
- applications for Training grants (T32) and Fellowship (F-series grant) do not need a DMS Plan

In contrast, for most grants - Research grants that are collecting Scientific Data, as broadly defined by the DMS Policy, Its Your Thing, and you need a sharing plan

A link to the complete list of activity codes that fall under the

plan is here:
https://sharing.nih.gov/sites/default/files/flmngr/List-of-Activity-Codes-Applicable-to-DMS-Policy.pdf



Two important points of emphasis about Data Management and Sharing Plans:

The Data Management and Sharing plan and budget will not affect the impact score of your grant.

The plan itself is not seen by reviewers - it is not included in their packet of materials.

The budget for data sharing will be available to review but is not score-driving

Second, a plan is needed even if you are not sharing

data.

In fact it is probably even more important if you have data you are not sharing, as you have to account for your decisions to share or not share given data in the Plan

Before Starting the DMS Plan

- Check the funding opportunity
  - Repositories to use
  - Sharing requirements
  - Budget limits or requirements
- Consider preliminary conversations with IRB, communities affected, program officers as needed
- Decide on a repository

16

Before you start to write the plan here are 3 things to help prepare

1. Check the funding announcement you are applying to, and look to see if there are requirements about what repository to use, or any novel sharing or budget requirements
2. Consider speaking with your IRB, any communities that might be affected by the data sharing requirement, and your program officers as needed
3. Finally, decide on a repository to use for sharing.

Where to Share?

NIH encourages use of the established repository that is most appropriate for your data type and discipline

Repository specified by an NIH Institute or the Funding Notice
NHGRI – AnVIL

Discipline or data-type specific repository

Other potential options for small less complex data sets

NIH NHGRI

17

One question we are hearing frequently is *where applicants should share their data*

If the funding opportunity does not specify a repository, NIH encourages researchers to select the repository that is most appropriate for their data type and discipline

Several ideas for repositories exist, and we'll put links into the chat for

- NIH's resource on Repositories for Sharing Scientific Data [https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data],which includes a listing of generalist repositories that accepts all data types, [https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/generalist-repositories]
- Nature's Data Repository Guidance [https://www.nature.com/sdata/policies/repositories] and

- the global Registry of Research Data Repositories. [https://www.re3data.org/]

You can use more than one repository for a given study.

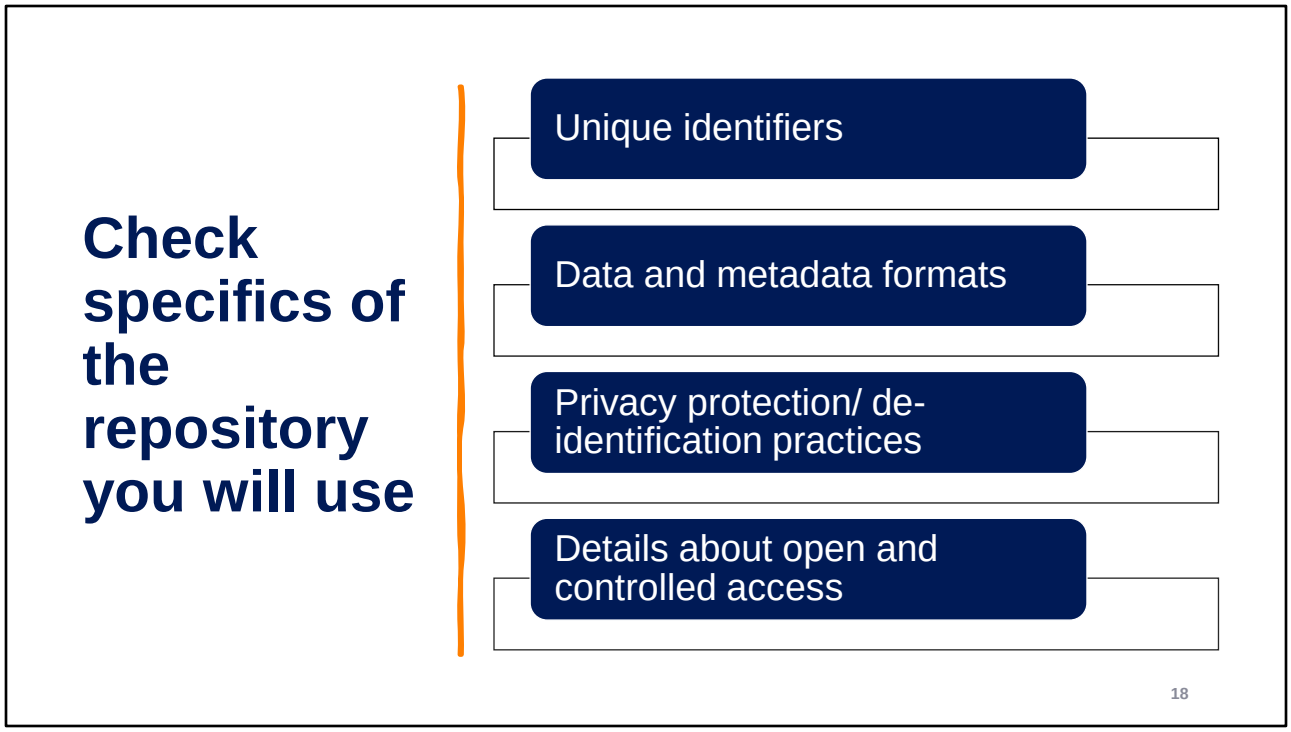
However, whenever possible we encourage using a single location for all data and documentation.

When selecting a repository, consider factors including the accessibility and stability of the repository and the privacy and security it offers, and curation services it might provide.

A full list of things to look for is linked https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository

NIH's preferences is that researchers first consider putting data in the repository supported by Institute.

1. AnVIL https://anvilproject.org/ is NHGRI's primary repository,- it supports sharing a variety of data types including data collected by ELSI investigators
2. You can propose an alternative to AnVIL, and justify your choice.
3. There are other options for small datasets, including attaching the data as an appendix to a publication. However such a choice may not maximize access to the data and you may be asked to revise your plan.



When you are deciding on a repository, look at the detailed information they provide about depositing protecting and sharing data.

Platforms vary across many important features, which may influence your choice.

Also, a repository's recommendations & requirements may provide information that is helpful when you are writing your plan

Writing the DMS Plan

- 2 pages
- The template NIH provides is optional but helpful
- Sample plans are available
- DMPTool
- No hyperlinks in plan
- Includes genomic data sharing if applicable
- Resource Sharing Plan is separate

NIH NHGRI

19

Form:

A few quick thoughts on writing the plan

Strive for 2 pages in length
NIH provides a blank template plan which is optional to use but very helpful as it lays out a structure and reminds you of details to include https://grants.nih.gov/sites/default/files/DMS-Plan-blank-format-page.docx

Several institutes including NHGRI have posted **sample plans online** https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/writing-a-data-management-and-sharing-plan#sample-plans

The Samples do not cover all possible data types and sharing issues.
Do not copy/paste the sample Plan.
Consider the specifics of your application – focusing on data you'll generate, and the circumstances they can be shared under.

Another online resource called
DMPtool https://dmptool.org/public_templates?page=1&search=nih outlines NIH requirements, and provides some template text that may be useful.

Some specific advice –
First do not include any hyperlinks In your data sharing and management plan. Doing so can cause your entire grant to be rejected before it is reviewed.
Second, most ELSI grants will not collect genomic data but if you are collecting such data, then the genomic data sharing plan gets embedded within the larger DMSP.
Finally a Resource Sharing Plan may be still required in addition (check your NOFO). These Resource Sharing Plans are now being used to describe dissemination of other grant products like publications, software, websites, or research tools.

Elements of the DMS Plan

1. Data types
2. Related tools, software and code
3. Standards
4. Data preservation, access, timelines
5. Access, distribution, reuse considerations
6. Oversight of data management and sharing

NIH NHGRI

20

There are six main elements of the data management and sharing plans. An overview of each section is below. The NIH template also provides detail.

- **Data types -** Briefly describe the scientific data to be managed and shared:
  - Summarize the types and amount of data you expect to generate
  - Which data will be preserved and shared, and rationale, based on ethical, legal, and technical factors
  - Metadata, other relevant data, and associated documentation to share. This includes data collection instruments, interview guides, survey data dictionaries; qualitative coding schemes, analysis methods, and any important contextual details

needed to use the data. Plans should identify proprietary measures that cannot be publicly shared

- **Related tools, software and code**
    - List specialized tools, software, or code needed to access or analyze the shared data – for example Nvivo might be needed to read transcript coding files.
    - How to access them (open source, commercial**)**
    - You don't have to provide access to the tool
- **Standards**
    - Common formats and standards you will use to share data to maximize use and interoperability
- **Data preservation, access, timelines:** this includes
    - Repositories where data will be stored and shared, for each data type
    - How data will be findable and associated with your study
    - When data will be made available, and for how long
- **Access, distribution, reuse considerations:** what factors will affect subsequent access, distribution, or reuse of your scientific data?
    - How are you maximizing data sharing, including Privacy and confidentiality provisions used
    - Any limitations on sharing, resharing and reuse of data
    - Justifications for those limitations
- **Oversight of data management and sharing**
    - Who on the grant is responsible for monitoring compliance.
    - How will the applicant organization monitor and manage compliance with the Plan.

How much detail do I need to provide in the plan?

- Check the NIH Sample Plans for examples
- Address each data type that is being treated differently
- Be clear which data will and will not be shared.
- Clear rationale for anything not shared or limited

NIH NHGRI

21

Throughout the plan, in each section, make sure you address data types that are being treated differently

- Check the NIH Sample Plans for examples
- Throughout the plan address each data type that is being treated differently
- Be clear which data will and will not be shared.
- Provide clear rationale for anything not shared or limited in sharing

Submitting the plan

A new "**Other Plan(s)**" field added to the PHS 398 form to collect a single PDF attachment

Research Plan Section				
5. Vertebrate Animals		Add Attachment	Delete Attachment	View Attachment
6. Select Agent Research		Add Attachment	Delete Attachment	View Attachment
7. Multiple PD/PI Leadership Plan		Add Attachment	Delete Attachment	View Attachment
8. Consortium/Contractual Arrangements		Add Attachment	Delete Attachment	View Attachment
9. Letters of Support		Add Attachment	Delete Attachment	View Attachment
10. Resource Sharing Plan(s)		Add Attachment	Delete Attachment	View Attachment
11. Other Plan(s)		Add Attachment	Delete Attachment	View Attachment
12. Authentication of Key Biological and/or Chemical Resources		Add Attachment	Delete Attachment	View Attachment

NIH NHGRI

22

You may be working with your administrative or Scientific Officer when attaching the Data Sharing Plan to your grant

A specific field "Other Plans" has added to the form, this is where the DMS Plan gets attached to your grant application

Budgeting for Data Sharing Activities

You may request funds for data management and sharing

Allowable

- Curating, de-identifying data
- Formatting to standards
- Transmission & storage
- Repository deposit fees
- Creating supporting docs, metadata
- Some local infrastructure

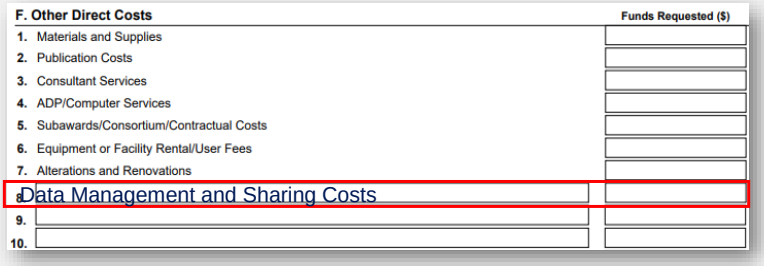
Not allowable

- Infrastructure covered by institutional overhead costs
- Routine conduct of research
- Double charged direct and indirect costs

NIH NHGRI

Submitting the budget

Show Amount on R&R Budget Form: line item in section F. Other Direct Costs



Justify costs PHS 398 Modular Budget Form: within Additional Narrative Justification

NIH NHGRI

24

For applications with detailed budgets you'll need to add a budget line " Data Management and Sharing Costs"
Show the amount requested for data management and sharing activities, under Funds Requested
If no cost will be requested, enter "0" in the "Funds Requested", do not leave it blank
Details should be outlined in the budget justification attachment

For modular grants simply elaborate in the narrative justification attachment
If no costs will be incurred, still include a line item and enter "0" for the requested dollar amount

Your DMS Plan must be approved by NIH prior to a grant award;

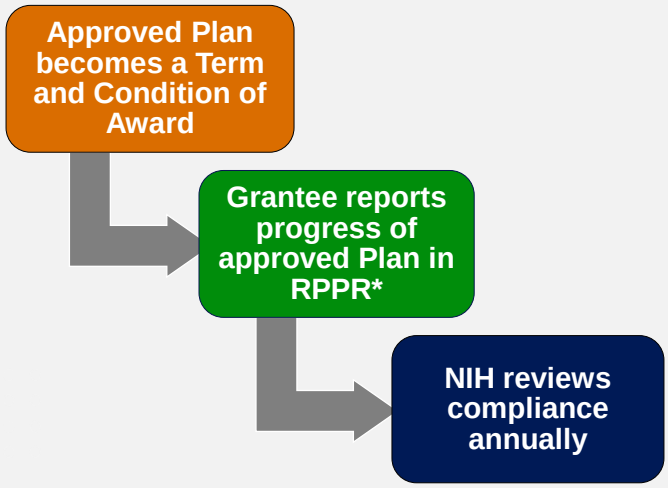
This is an internal review by program officers and grant officers

Applicants will be notified by program directors, and expected to communicate with NIH and revise the plan as needed .

We also realize that plans may need to be updated or revised over the course of a project for a variety of reasons for example, changes in community permissions, or addition of a data set from a supplement.

Those revisions need to be mutually agreed upon as well.

Monitoring Compliance



NIH NHGRI

***RPPR:** Research Performance Progress Report (RPPR) - Annual, Interim, and Final

The plan that we approve becomes a term and condition of your award.

As a grantee you will; report on your compliance with the plan as part of your annual progress reports or RPPRs.

NIH will review your compliance annually.

Failure to comply may result in an enforcement action and affect future funding decisions.

Communication

- Contact us well in advance of your application to discuss questions
- Make sure your AO knows how to upload the plan
- We will contact you after programmatic review if we are considering the grant for funding.

NIH NHGRI

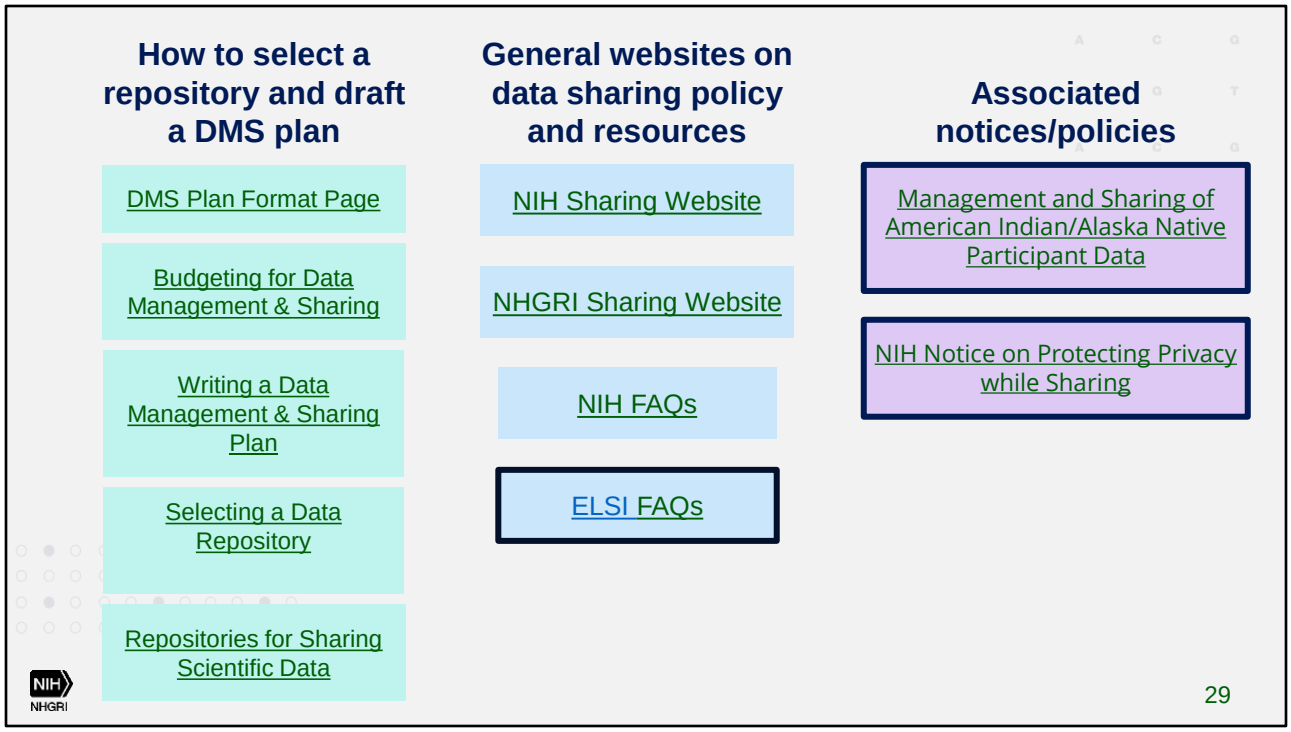
27

- Contact us in advance of your application to discuss questions
- Make sure your Administrative Officer at your institution knows how to upload the plan
- We will contact you if we are considering the grant for funding – but you are always welcome to reach out

Resources

NIH
NHGRI

28



There are many online resources:

How to select a repository and draft a DMS plan
DMS Plan Format Page
Budgeting for Data Management & Sharing
Writing a Data Management & Sharing Plan
Selecting a Data Repository
Repositories for Sharing Scientific Data

General websites on data sharing policy and resources

NIH Sharing Website
NHGRI Sharing Website
NIH FAQs
ELSI FAQs

Associated notices/policies

Management and Sharing of American Indian/Alaska Native Participant Data

NIH Notice on Protecting Privacy while Sharing

We also hope to be resource for you…please reach out and be patient as we learn the ropes together

ELSI Program Officers:

dave.kaufman@nih.gov Dave Kaufman
lockhani@mail.nih.gov Nicole Lockhart
rene.sterling@nih.gov Rene Sterling

Q & A

NIH NHGRI