National Institutes of Health
Office of Science Policy

Overview of the NIH Data Management and Sharing Policy

Presentation to NSC3
June 21, 2023

Ellen Wann, PhD
Data Management and Sharing Policy Team Lead
Scientific Data Sharing Policy Division
Office of Science Policy
National Institutes of Health

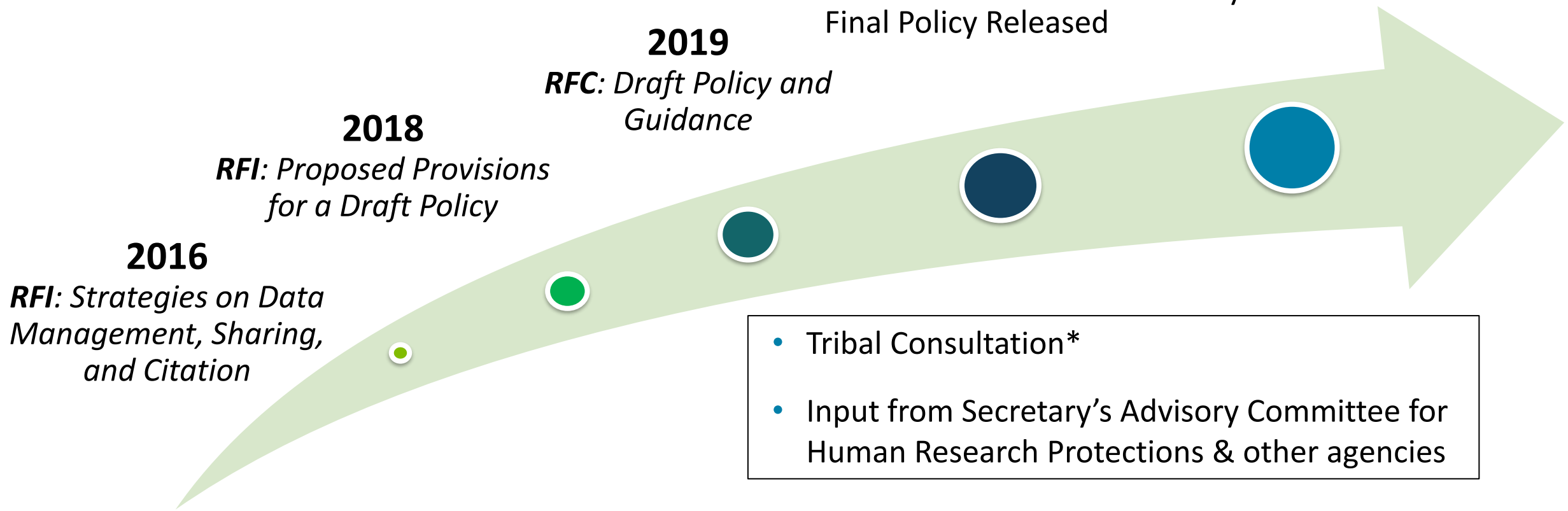
Cindy Danielson, PhD
Associate Director, Systems Integration,
Office of Research Reporting and Analysis
Office of Extramural Research
National Institutes of Health

Why does NIH Want Data to be Shared?

- **Advance rigorous and reproducible research**
  - Enable validation of research results
  - Make high-value datasets accessible
  - Accelerate future research directions
  - Increase opportunities for citation and collaboration

- **Promote public trust in research**
  - Foster transparency and accountability
  - Demonstrate stewardship over taxpayer funds
  - Maximize research participants' contributions
  - Support appropriate protections of research participants' data

Iterative Policy Development through Consistent Community Engagement



- Tribal Consultation*
- Input from Secretary's Advisory Committee for Human Research Protections & other agencies

*See "NIH Tribal Consultation Report: NIH Draft Policy for Data Management and Sharing"

NIH Policy for Data Management and Sharing

- **Submission of Data Management & Sharing Plan for all NIH-funded research** *(how/where/when)*
- **Compliance with the ICO-approved Plan** *(may affect future funding)*
- **Effective January 25, 2023** *(replaces 2003 Data Sharing Policy)*

Details [of the Policy] Matter!

- **Scope:** All NIH-supported research generating *scientific data*
  - **What's in: "**Recorded factual material… of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications"—relates to the proposed research questions and findings can include unpublished null results
    - May include qualitative data or data produced using fundamental basic science techniques
  - **What's out:** lab notebooks, preliminary analyses, case report forms, physical objects

- **Timelines:**
  - **When to share data?** no later than publication or end of award (for data underlying findings not published in peer-reviewed journals)
  - **How long to share data?** consider other relevant requirements and expectations (e.g., journal policies, repository policies)

Additional Expectations for Plans

- **SHARING SHOULD BE …**
  - **The default practice**
    - Data sharing should be maximized (with justifiable limitations)
    - All data should be managed; **not all must be shared**

  - **Responsibly implemented**
    - Plans should outline protection of privacy, rights, and confidentiality
    - Abide by existing laws, regulations, and policies
  - **Prospectively planned for at all stages of the research process**

Potential Limitations on Sharing

- **Data Management and Sharing Plans should maximize appropriate sharing:**
  - **Justifiable ethical, legal, and technical factors for limiting sharing of data include:**
    - Informed consent will not permit or limits scope of sharing or use
    - Privacy or safety of research participants would be compromised and available protections insufficient
    - Explicit federal, state, local, or Tribal law, regulation, or policy prohibits disclosure
    - Restrictions imposed by existing or anticipated agreements with other parties
    - Datasets cannot practically be digitized with reasonable efforts
  - **Reasons not generally justifiable to limit sharing include:**
    - Data are considered too small
    - Researchers anticipate data will not be widely used
    - Data are not thought to have a suitable repository
  - **Additional considerations:**
    - NIH respects Tribal sovereignty and supports responsible management/sharing of AI/AN participant data
    - SBIR/STTR Program Policy Directive permits withholding data for 20 years, as stipulated in agreements and consistent with program goals

Supplemental Information: Repository Selection

- Encourages use of established repositories
  - Generally improves FAIRness of data
- Helps investigators identify appropriate data repositories
  - E.g., use of persistent unique identifiers, attached metadata, facilitates quality assurance
- NIH ICs may designate specific data repository(ies)

See Selecting a Data Repository for details

Supplemental Information: Repository Selection
Specialized Data Repositories

- Prioritizes data-type and discipline-specific data repositories
- Refers to NIH-supported data repository list outlining:
  - Repository description (e.g., data-types accepted, research community served, tools available),
  - Supportive NIH IC(s),
  - Whether and when new data are accepted, and
  - How to submit data
- Examples include:

dbGaP	BioData Catalyst
GenBank	ImmPort
NIMH Data Archive	BioLINCC

Supplemental Information:
Allowable Costs

- **Reasonable costs allowed in budget requests (must be incurred during the performance period)**
  - Curating data/developing supporting documentation
  - Preserving/sharing data through repositories
  - Local data management considerations

- **<u>NOT</u> considered data sharing costs**
  - Infrastructure costs typically included in indirect costs
  - Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)

- Over time NIH **hopes to learn more about what constitutes reasonable costs** for various data management and sharing activities

See Budgeting for Data Management & Sharing for details

Plan Submission and Review: A Guide

Extramural Grant Awards*

Plan Submission

With application

Brief Plan description in Budget Justification

Full Plan as separate attachment

Plan Assessment

Peer reviewers comment on (not score) budget

NIH program staff assess Plans

Plans can be revised

Plan Compliance

Incorporated into Terms and Conditions

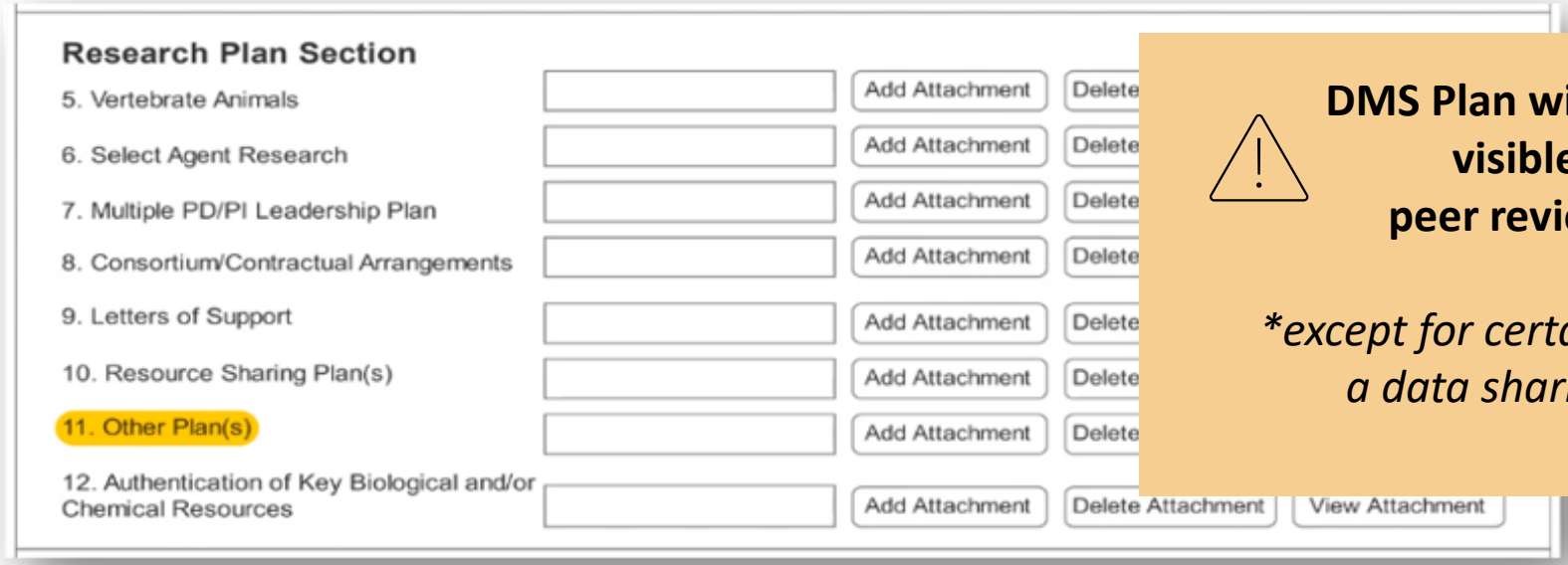
Monitored at regular reporting intervals – mechanisms and tools to support oversight under development

Compliance may factor into future funding decisions

**Analogous requirements for contracts, Other Transaction Awards, NIH Intramural Research Program*

Changes to DMS Plan Submission

- Updates made to FORMS-H version of grant application forms – updated instructions available now
- New "Other Plan(s)" field added to PHS 398 forms to collect a single DMS Plan PDF attachment
  - Separate "Data Sharing Plan" and "Genomic Data Sharing Plan" attachments no longer collected as Resource Sharing Plan(s)
  - Plans for sharing genomic data included as part of the DMS Plan in Other Plan(s)

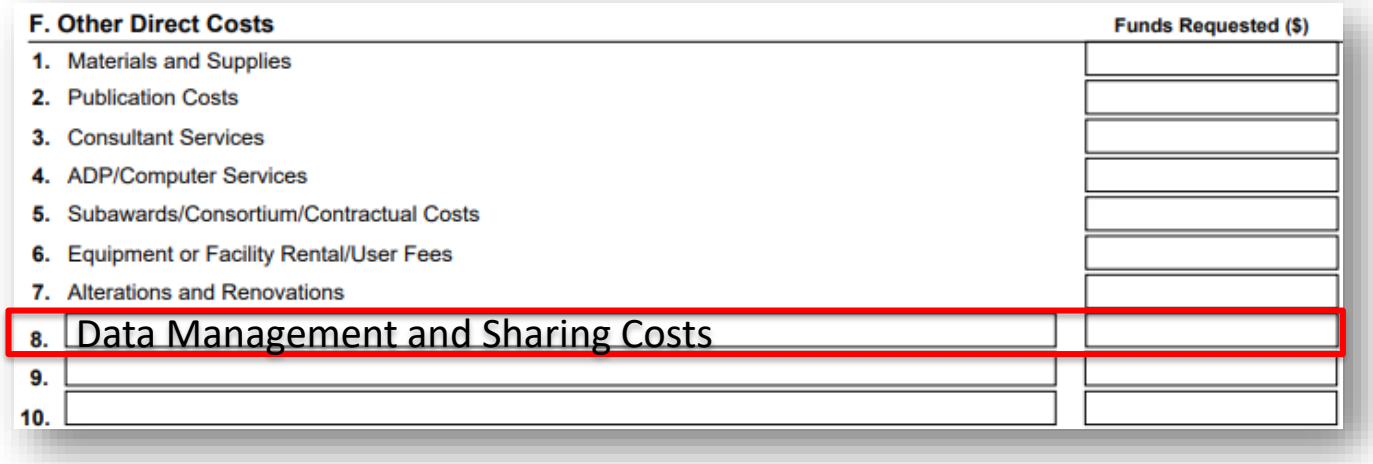


DMS Plan will NOT be visible to peer reviewers.

**except for certain FOAs with a data sharing focus*

Submitting DMS Budgets

- Direct costs to support activities proposed in DMS Plan must be indicated as “Data Management and Sharing Costs”
  - R&R Budget Form: line item in section F. Other Direct Costs



F. Other Direct Costs	Funds Requested ($)
1. Materials and Supplies	
2. Publication Costs	
3. Consultant Services	
4. ADP/Computer Services	
5. Subawards/Consortium/Contractual Costs	
6. Equipment or Facility Rental/User Fees	
7. Alterations and Renovations	
8. Data Management and Sharing Costs	
9.	
10.	

  - PHS 398 Modular Budget Form: within Additional Narrative Justification

2. Budget Justifications

Personnel Justification		Add Attachment	Delete Attachment	View Attachment
Consortium Justification		Add Attachment	Delete Attachment	View Attachment
Additional Narrative Justification		Add Attachment	Delete Attachment	View Attachment

Justifying DMS Budgets

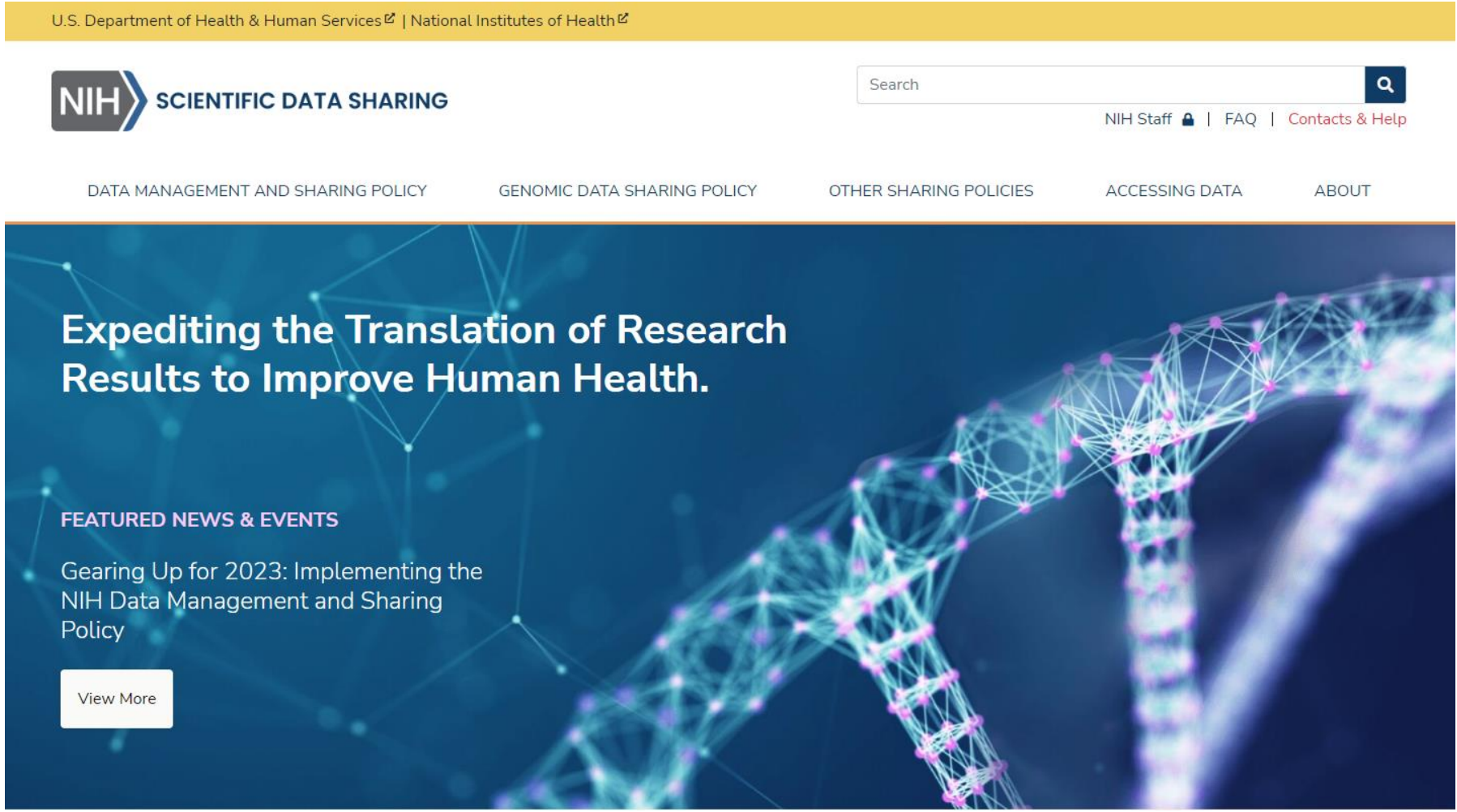
- Brief summary of DMS Plan and description of DMS costs (recommended ≤ ½ page) must be included within the budget justification attachment; peer reviewers may comment on requested DMS costs based on this information
  - R&R Budget Form: section L. Budget Justification

  - PHS 398 Modular Budget Form: Additional Narrative Justification

See Instructions in FORMS-H Application Guide

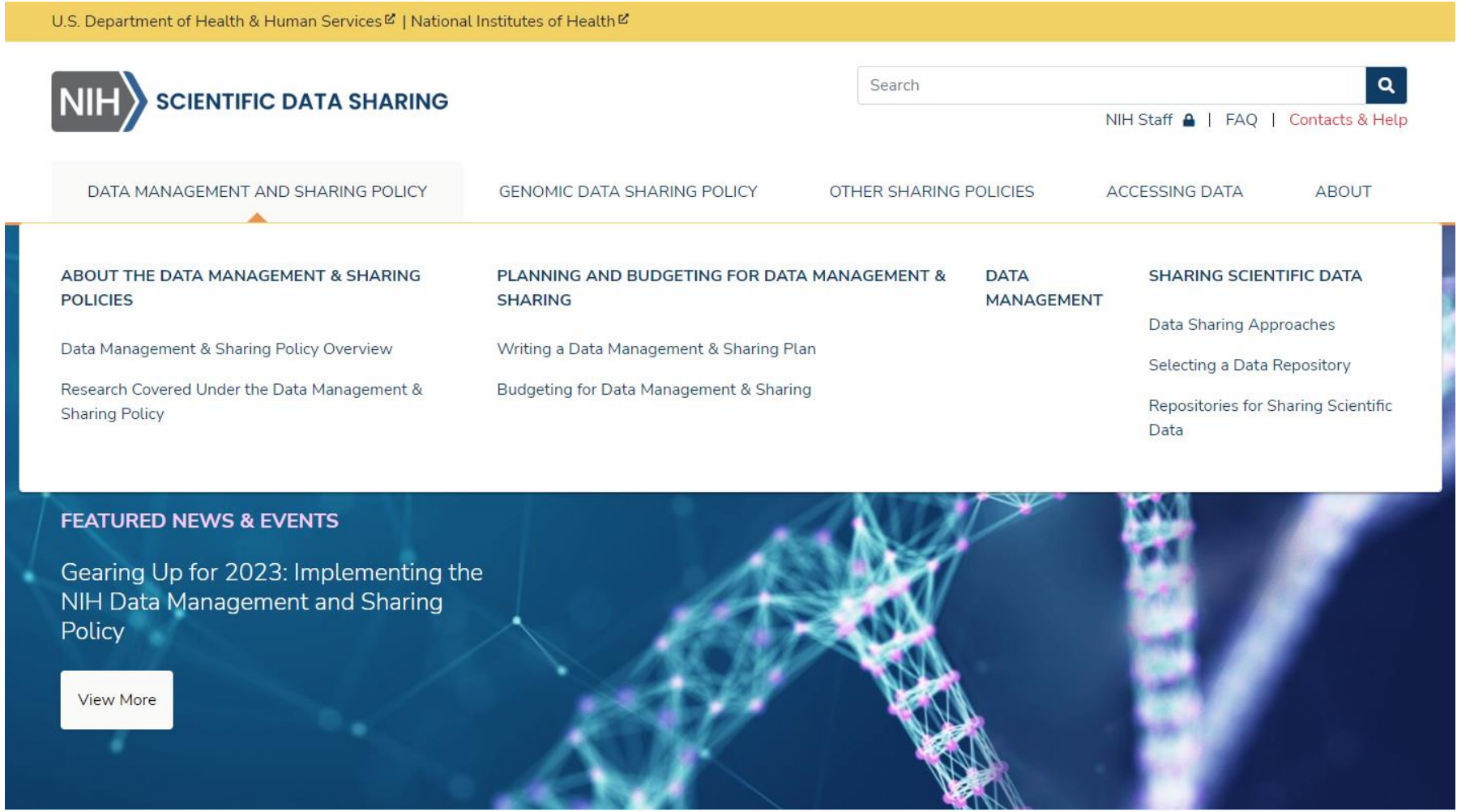
sharing.nih.gov

- Provides a central source of guidance related to multiple NIH data sharing policies
- Covers Data Management and Sharing, Genomic Data Sharing, Model Organisms, and Research Tools policies
- Content will be updated



Resources for…

- Provides a central source of guidance related to multiple NIH data sharing policies
- Covers Data Management and Sharing, Genomic Data Sharing, Model Organisms, and Research Tools policies
- Content will be updated

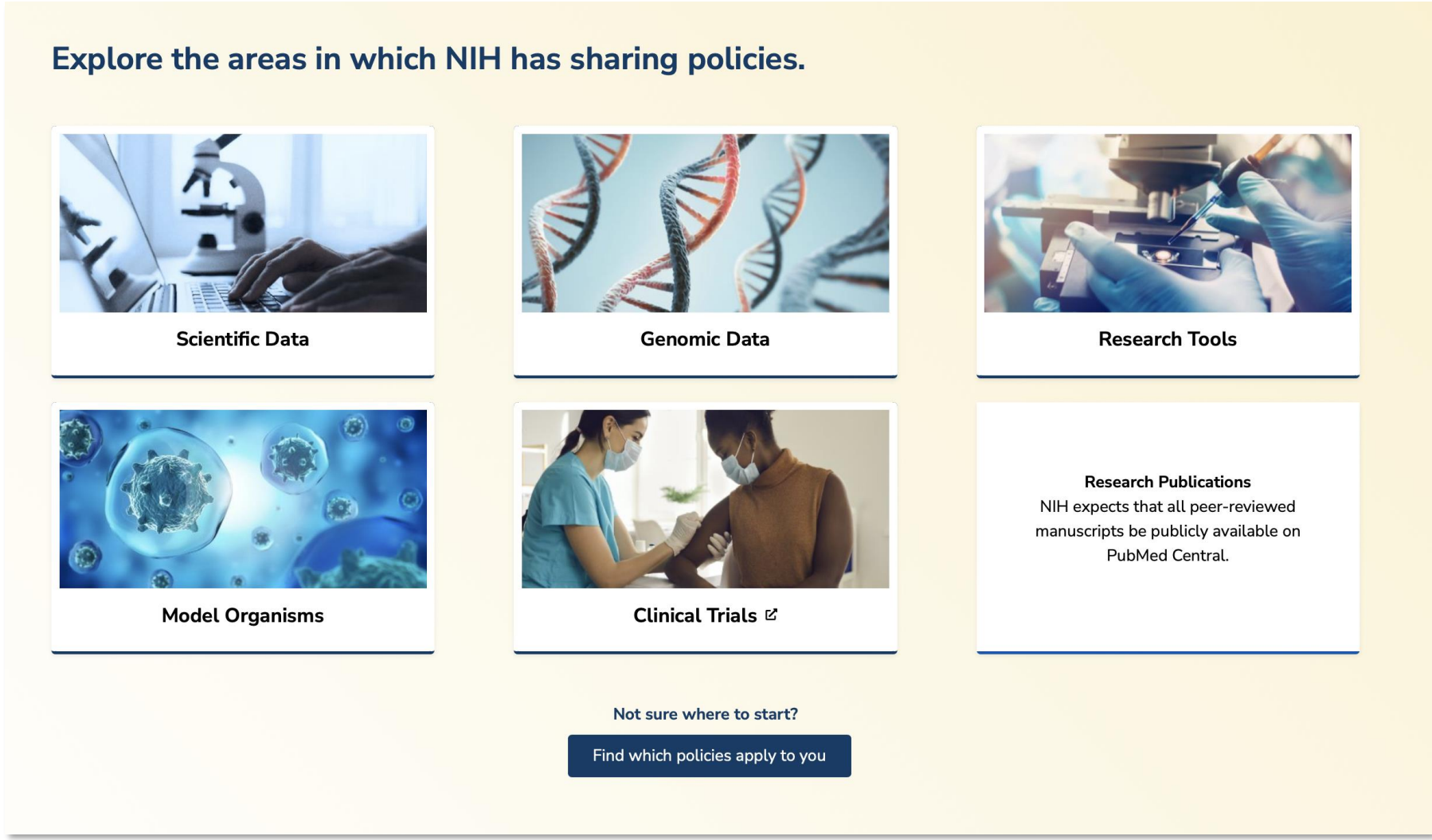


Resources for: Planning and Budgeting

- Expectations for writing a Data Management & Sharing Plan
- Budgeting for data management & sharing in your application for funding
- Includes instructions and format for submitting Plans

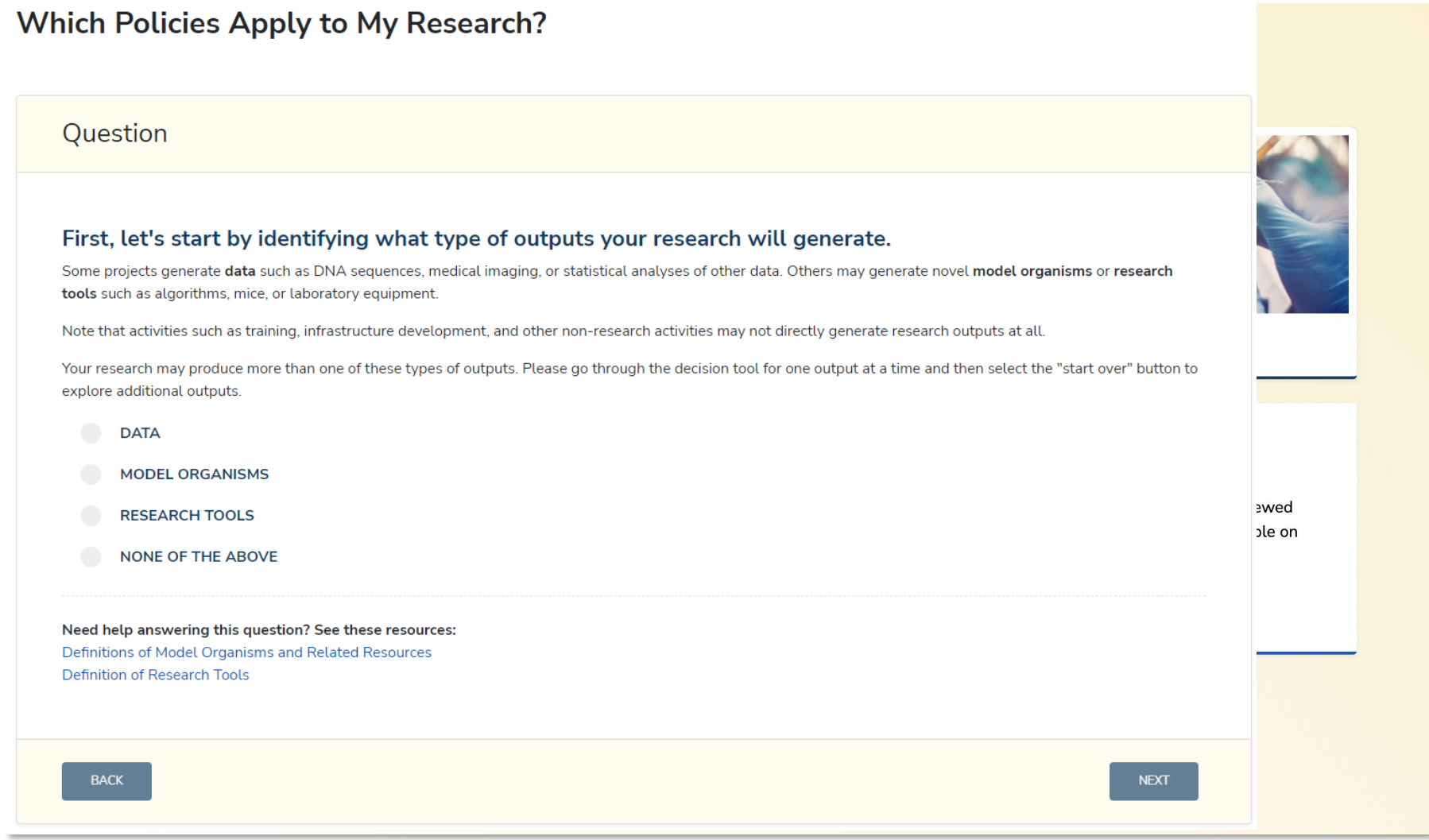
Resources for: Understanding Sharing Policies

- Explore NIH sharing policies
- Tool to find which policies apply to you
- FAQs to help understand each policy



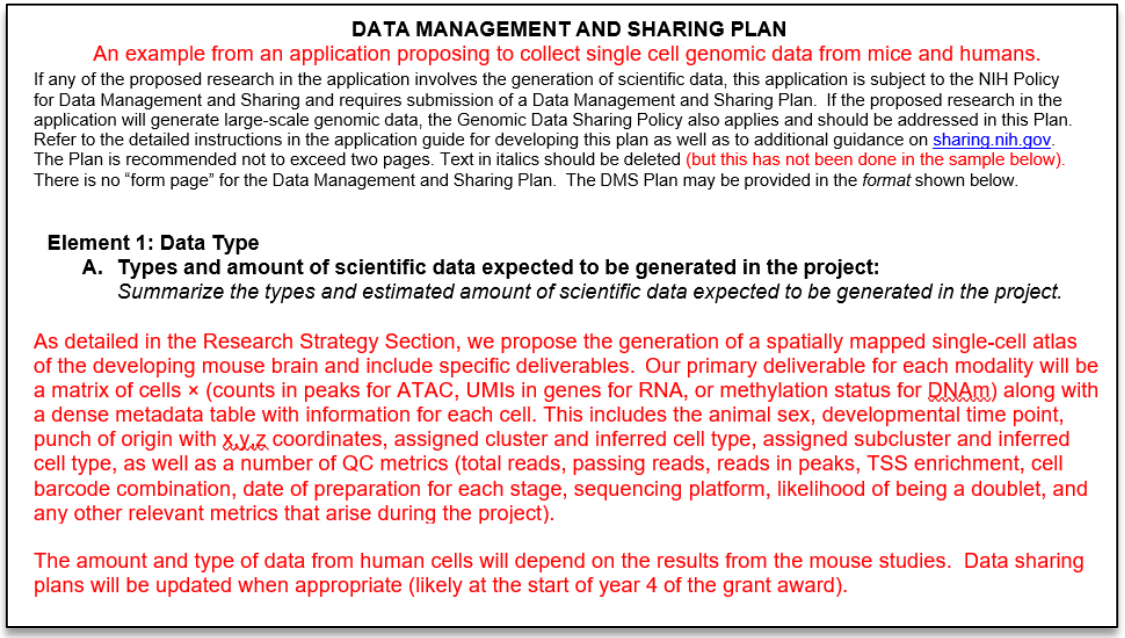
Resources for: Which Policies Apply to You?

- Explore NIH sharing policies
- Tool to find which policies apply to you
- FAQs to help understand each policy



Sample NIH DMS Plans Available

- 10+ sample NIH DMS Plans available for educational purposes, including:
  - Human clinical and/or MRI data (NIMH)
  - Human genomic data (NIMH, NHGRI, NIDDK)
  - Human & non-human genomic data (NIMH)
  - Secondary data analysis (NIMH, NIDDK)
  - Human clinical and genomics data (NICHD)
  - Human survey data (NICHD)
  - Model organism (Zebrafish) data (NICHD)
  - Technology development (NHGRI)
  - Clinical data (NIDDK)
  - Non-human basic research (NIDDK)



DATA MANAGEMENT AND SHARING PLAN

An example from an application proposing to collect single cell genomic data from mice and humans.

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted (but this has not been done in the sample below). There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the *format* shown below.

Element 1: Data Type

A. Types and amount of scientific data expected to be generated in the project:

Summarize the types and estimated amount of scientific data expected to be generated in the project.

As detailed in the Research Strategy Section, we propose the generation of a spatially mapped single-cell atlas of the developing mouse brain and include specific deliverables. Our primary deliverable for each modality will be a matrix of cells × (counts in peaks for ATAC, UMIs in genes for RNA, or methylation status for DNAm) along with a dense metadata table with information for each cell. This includes the animal sex, developmental time point, punch of origin with x,y,z coordinates, assigned cluster and inferred cell type, assigned subcluster and inferred cell type, as well as a number of QC metrics (total reads, passing reads, reads in peaks, TSS enrichment, cell barcode combination, date of preparation for each stage, sequencing platform, likelihood of being a doublet, and any other relevant metrics that arise during the project).

The amount and type of data from human cells will depend on the results from the mouse studies. Data sharing plans will be updated when appropriate (likely at the start of year 4 of the grant award).

See Writing a Data Management & Sharing Plan for details

Thank You!

Policy and Supplemental Information:

- NOT-OD-21-013 – Final NIH Policy for Data Management and Sharing
- NOT-OD-21-014 – Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan
- NOT-OD-21-015 – Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing
- NOT-OD-21-016 – Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research

Resources:

- NIH Data Sharing Website – sharing.nih.gov
- NIH Office of Science Policy DMS Policy Website – history and background on the NIH DMS Policy
- Frequently Asked Questions – sharing.nih.gov/faqs
- NIH Data Management and Sharing Policy Webinar Series – Implementation of the NIH DMS Policy
- News & Events – Latest news and upcoming events

Contact:

- Questions – sharing@nih.gov
- Follow us on Twitter – @NIH_OSP
- osp.od.nih.gov/blog/