



National Institutes of Health  
*Office of Science Policy*

# The NIH Data Management and Sharing Policy: Overview and Implementation Update

Presentation to the 2023 NCI DCCPS New Grantee Workshop  
October 30-31, 2023

## **Taunton Paine, MA**

Director, Scientific Data Sharing Policy Division  
Office of Science Policy  
National Institutes of Health  
[taunton.paine@nih.gov](mailto:taunton.paine@nih.gov)

## **Charlisse Caga-anan, JD**

Program Director, Genomic Program Administrator  
Division of Cancer Control and Population Sciences  
National Cancer Institute  
[charlisse.caga-anan@nih.gov](mailto:charlisse.caga-anan@nih.gov)

# 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

# Major NIH-wide Data Sharing Policies

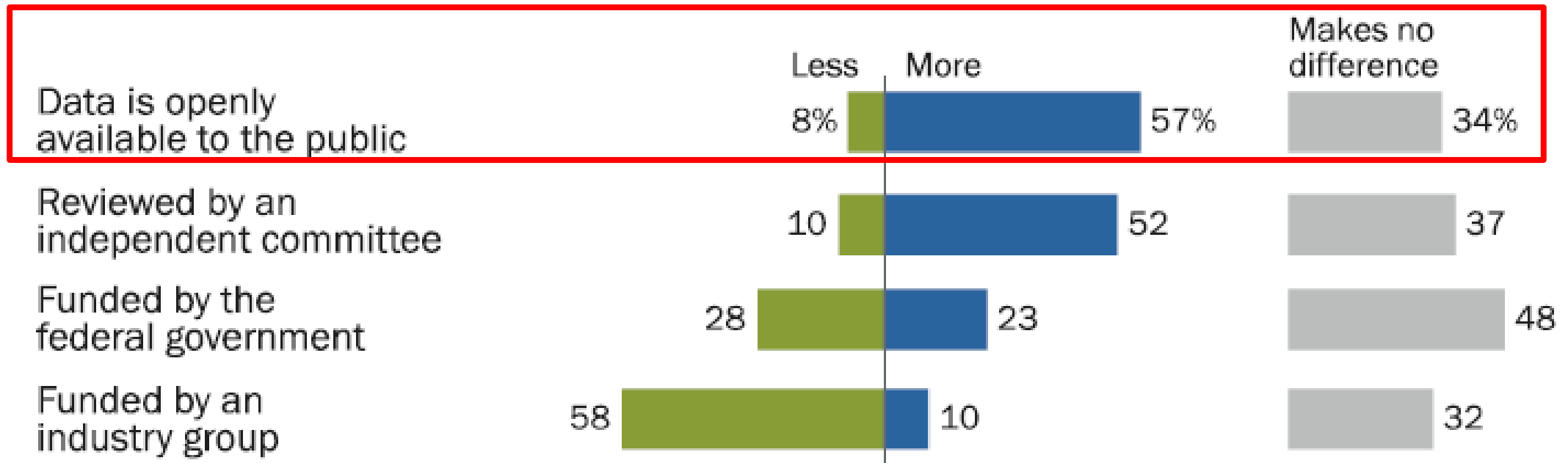
Policy	Expectations	Year
<b>NIH Data Sharing Policy</b>	Expects investigators seeking more than \$500K in direct support in any given year to submit a data sharing plan with their application or to indicate why data sharing is not possible.	2003
<b>Genomic Data Sharing Policy</b>	Expects sharing of large-scale human and non-human genomic data from NIH-funded studies through a publicly available data repository. All studies with human genomic data should be registered in <a href="#">dbGaP</a> , and the data should be submitted to an <a href="#">NIH-designated data repository</a> . Non-human data may be submitted to any widely used data repository.	2014
<b>Dissemination of NIH-Funded Clinical Trial Information</b>	Expects all investigators conducting NIH-funded clinical trials to register trials at ClinicalTrials.gov, and submit results information. Complementary to Part 11 regulations.	2016

# Data Accessibility: Still Work to Do

Author	Finding	Year
<a href="#">Tedersoo et al.</a>	<ul style="list-style-type: none"><li>• Evaluated data availability in <b>875 papers</b> across nine disciplines published 2000-2019</li><li>• Data requests successful <b>39.4%</b> on average; ranged 27.9–56.1% per field, <b>19.4% of requests declined</b> after repeated follow-up</li></ul>	2021
<a href="#">Errington et al.</a>	<ul style="list-style-type: none"><li>• Attempted to repeat <b>193 experiments</b> from <b>53 high-impact cancer biology papers</b>; <b>able obtain data for 32% of experiments</b></li></ul>	2021
<a href="#">Gabelica et al.</a>	<ul style="list-style-type: none"><li>• Requested data from <b>1,792 papers</b> published January 2019 with data availability statements; <b>6.8% of authors provided the requested data</b></li></ul>	2022
<a href="#">Narang et al.</a>	<ul style="list-style-type: none"><li>• Evaluated data availability for <b>213 NIH-funded pediatric clinical trial publications</b></li><li>• Individual-level participant <b>data available for 3.3%</b> of publications</li></ul>	2023
<a href="#">Hussey</a>	<ul style="list-style-type: none"><li>• Requested data from <b>52 papers</b> employing Implicit Relational Assessment Procedure over previous 5 years; <b>26.9% of authors provided the requested data</b></li></ul>	2023 (preprint)

# A Matter of Trust

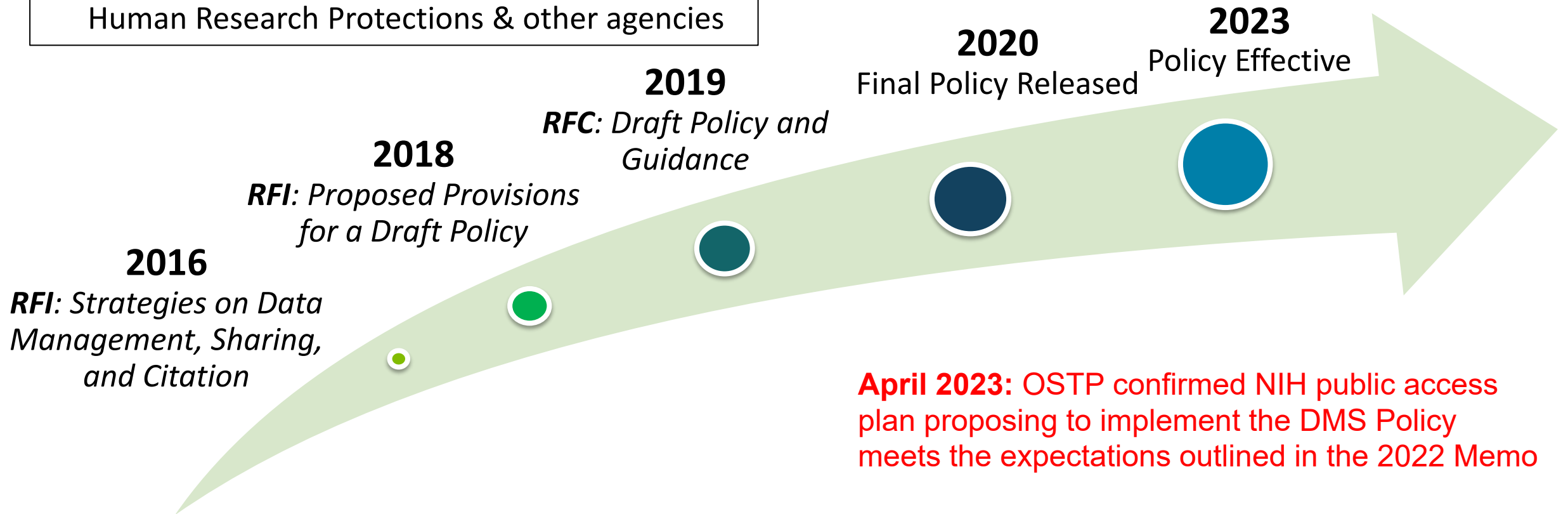
*% of U.S. adults who say when they hear each of the following, they trust scientific research findings ...*



[https://www.pewresearch.org/science/wp-content/uploads/sites/16/2019/08/PS\\_08.02.19\\_trust.in\\_scientists\\_FULLREPORT.pdf](https://www.pewresearch.org/science/wp-content/uploads/sites/16/2019/08/PS_08.02.19_trust.in_scientists_FULLREPORT.pdf)

# 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** (*replaced 2003 Data Sharing Policy*)

# Activities Subject to the DMS Policy

- **Applies to all research generating scientific data**, including but not limited to:
  - Research Projects
  - Some Career Development Awards (Ks)
  - Small Business SBIR/STTR
  - Research Centers
- **Does not apply to research projects not generating scientific data or non-research projects**, including but not limited to:
  - Training (Ts)
  - Fellowships (Fs)
  - Construction (C06)
  - Conference Grants (R13)
  - Resources (Gs)
  - Research-Related Infrastructure Programs (e.g., S10)

See [Research Covered Under the Data Management & 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

# Should all data be shared? And when?

- Am I expected to share all data generated during my research?
  - No. Not all data generated during NIH-supported research will constitute scientific data.
- When should scientific data be shared?
  - The DMS Policy expects scientific data to be shared by the earlier of:
    - > **Scientific data underlying peer-reviewed journal articles:** no later than the date on which the article is first made available in print or electronic format
    - > **Scientific data underlying findings not disseminated through peer-reviewed journal articles:** by the end of the performance period (or end of no-cost extension)
  - Scientific data may underlie unpublished key findings, developments, and conclusions; or findings documented within preprints, conference proceedings, or book chapters. Scientific data underlying null and negative findings are important to share even though these key findings are not always published
- **How long should shared data be available?**
  - Consider other relevant requirements and expectations (e.g., journal policies, repository policies) as a minimum timeframe for making data available—many repositories are indefinite

# Supplemental Information: Repository Selection

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



See [Selecting a Data Repository](#) for details

# 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

# Elements of the Optional DMS Plan Format Page

## Element 1: Data Type

- A. Types/amount of scientific data to be generated
- B. Scientific data to be preserved and shared, and the rationale for doing so
- C. Metadata, other relevant data, and documentation

## Element 2: Related Tools, Software and/or Code

## Element 3: Data Standards

## Element 4: Data Preservation, Access, and Associated Timelines

- A. Repository where scientific data/ metadata archived
- B. How scientific data will be findable and identifiable
- C. When and how long scientific data will be available

## Element 5: Access, Distribution, or Reuse Considerations

- A. Factors affecting subsequent access, distribution, reuse
- B. Whether access to scientific data will be controlled
- C. Protections for privacy, rights, and confidentiality of human research participants

## Element 6: Oversight of Data Management and Sharing

See [Writing a Data Management & Sharing Plan](#) for details and [Format Page](#)

# Sample NIH DMS Plans Available

- Many 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, NHGRI)
  - 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](https://www.nih.gov/sharing). 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 DNA<sub>m</sub>) 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



## 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
- **NOT 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



**PRACTICAL GUIDANCE  
WORKING WITH NCI/DCCPS PROGRAM STAFF**





# Advice for PIs preparing DMS Plans



- DMS Plan is not limited to two pages
  - Plan should balance length with informative detail
- DMS Plan needs to address other policies/requirements (e.g., Genomic Data Sharing Policy, requirements listed in the Notice of Funding Opportunity)
  - Other policies/requirements may not apply to subset of data
  - Other policies/requirements may be more specific regarding some plan elements
- Limitations to sharing should include a detailed explanation
  - Limitations often inform the choice of repository or other access considerations for some of the data
- Data sharing descriptions should no longer be included in the “Resource Sharing Plan” section
  - Submit the DMS Plan in a new “Other Plans” section

# Including the DMS Plan in the grant application

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

- “Other Plan(s)” field has been added for the DMS Plan
- “Resource Sharing Plan(s)” field will still be available for other sharing not related to data (e.g., model organisms)

---

# DMS Plan Assessment



---

## Study Section:

- Peer reviewers will not see the DMS Plan
- Peer reviewers will see the DMS costs in the budget justification and consider if it is reasonable

## NIH Staff:

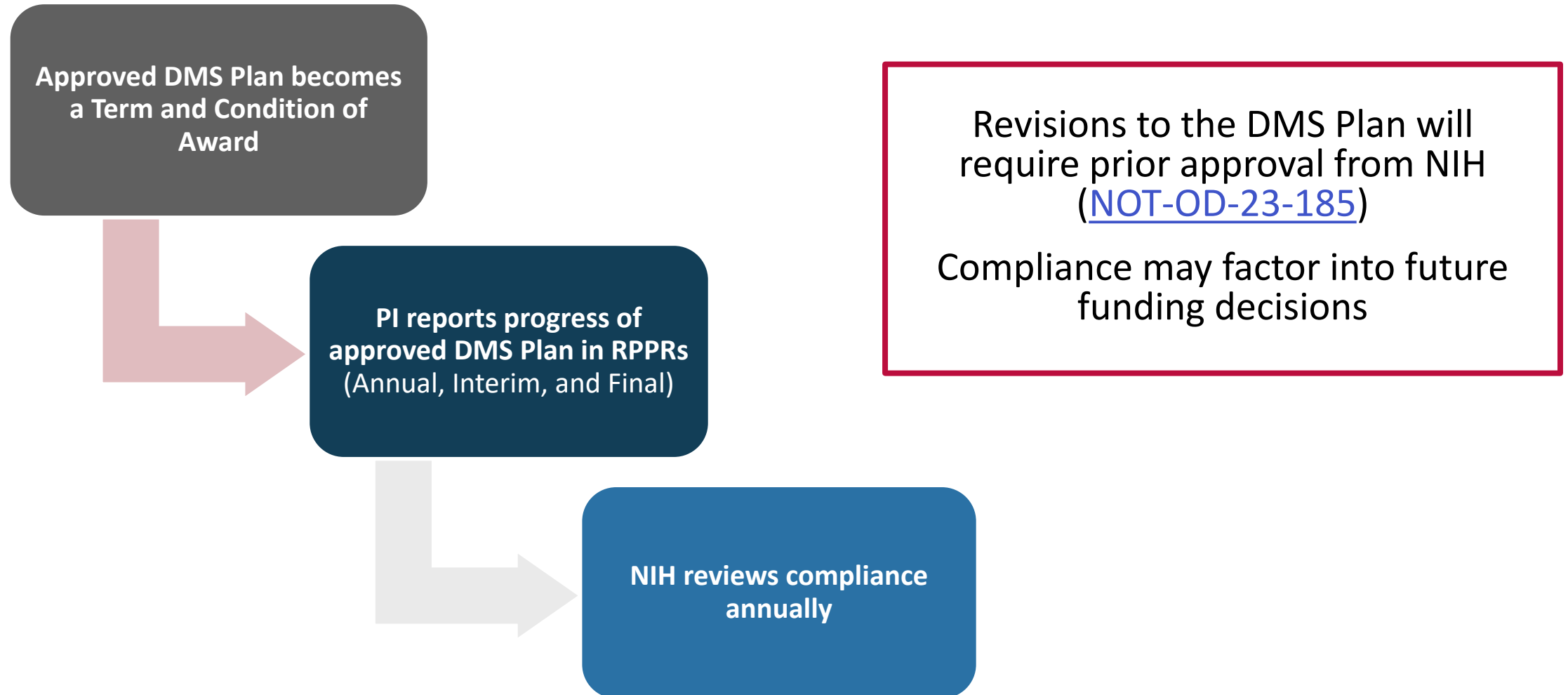
- Program officers review the plan to ensure the elements of the DMS Plan have been adequately addressed
- Applications will only be funded when the DMS Plan is considered complete and acceptable

# Before Award: Resolving DMS Plan Issues

- If additional details are needed, PIs should communicate with NIH staff to resolve issues with the DMS Plan
  - PI may potentially need to provide a revised DMS Plan
  - This should occur through standard Just-In-Time (JIT) process

**NOTE: PIs can contact NCI program staff for guidance and to review the DMS Plan before submitting a grant application**

# After Award: DMS Plan Compliance



# NCI DCCPS Resources Summary

- DCCPS Webpages:
  - Epidemiology and Genomics Research Program:  
<https://epi.grants.cancer.gov/funding/resource-data-sharing-plan.html>
  - Healthcare Delivery Research Program:  
<https://healthcaredelivery.cancer.gov/funding/data-sharing.html>
- NCI/NIH Program Staff

# 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

The screenshot shows the homepage of sharing.nih.gov. At the top, there is a yellow navigation bar with the text "U.S. Department of Health & Human Services | National Institutes of Health". Below this is the NIH logo and the text "SCIENTIFIC DATA SHARING". To the right of the logo is a search bar with the word "Search" and a magnifying glass icon. Further right are links for "NIH Staff", "FAQ", and "Contacts & Help". Below the navigation bar is a horizontal menu with five items: "DATA MANAGEMENT AND SHARING POLICY", "GENOMIC DATA SHARING POLICY", "OTHER SHARING POLICIES", "ACCESSING DATA", and "ABOUT". The main content area features a large blue background with a network diagram of nodes and lines. The headline reads "Expediting the Translation of Research Results to Improve Human Health." Below the headline is a section titled "FEATURED NEWS & EVENTS" with a sub-headline "Gearing Up for 2023: Implementing the NIH Data Management and Sharing Policy" and a "View More" button.

# 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](http://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](http://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](mailto:sharing@nih.gov)
- Follow us on Twitter – [@NIH\\_OSP](https://twitter.com/NIH_OSP)
- [osp.od.nih.gov/blog/](http://osp.od.nih.gov/blog/)

