NIH's New Data Management and Sharing Policy—How Academic Medical Centers Can Prepare for January 2023

Samuel, Sara; Paine, Taunton
https://dx.doi.org/10.7302/3731
https://hdl.handle.net/2027.42/171055

Downloaded from Deep Blue, University of Michigan's institutional repository
NIH's New Data Management and Sharing Policy—How Academic Medical Centers Can Prepare for January 2023

Taunton Paine & Sara Samuel
I have no relevant personal/professional/financial relationship(s) with respect to this educational activity

Taunton Paine, Director, Scientific Data Sharing Policy Division, Office of Science Policy, National Institutes of Health

I have no relevant personal/professional/financial relationship(s) with respect to this educational activity

Sara Samuel, Informationist, University of Michigan
Presenter 1: Taunton Paine

- Director, Scientific Data Sharing Policy Division, Office of Science Policy, NIH
- Worked in science policy for 10 years
- ORCID: 0000-0001-9037-4556
Presenter 2: Sara Samuel

- Informationist, Taubman Health Sciences Library, University of Michigan
- Worked in academic libraries for 16 years
- ORCID: 0000-0001-6808-2514
Special Thanks

- Cindy Danielson, PhD, Associate Director, Systems Integration, Office of Research Reporting and Analysis, NIH
- Amy Hafez, PhD, AAAS Science and Technology Policy Fellow, NIH Office of Science Policy
- Ellen Wann, PhD, Health Science Policy Analyst, NIH Office of Science Policy
- Diane Wilson, MPP, JD, MA, Assistant Director, Office of Regulatory Affairs, University of Michigan Medical School
Learning Objectives

1. Understand the specific guidelines, NIH's general expectations, and to whom they apply

2. Discuss essential elements of and best practices for Data Management and Sharing Plans, including examples

3. Identify the resources and stakeholders within an institution who will need to be engaged for researchers to develop adequate plans, along with the challenges they will face and the opportunities that the policy's requirements present
Overview of the New NIH Data Management and Sharing Policy

Taunton Paine
Director, Scientific Data Sharing Policy Division
Office of Science Policy, National Institutes of Health
NIH POLICY FOR DATA MANAGEMENT & SHARING:
Data Stewardship Goals

- **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
NIH POLICY FOR DATA MANAGEMENT & SHARING: An Iterative Policy Development Process

• Sought public comment repeatedly

• Tribal Consultation*
  *Details provided in “NIH Tribal Consultation Report: NIH Draft Policy for Data Management and Sharing”

• Intersection with other government agencies & Secretary’s Advisory Committee for Human Research Protections

2016: Solicited Community Input
RFI: Strategies on Data Management, Sharing, and Citation

2018: Solicited More Community Input
RFI: Proposed Provisions for a Draft Policy

2019: Solicited MORE Community Input
RFC: Draft Policy and Guidance

2020: Policy Release Date

2023: Policy Effective Date
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)

• Supplemental info available to assist

• Aims to foster data stewardship
NIH POLICY FOR DATA MANAGEMENT & SHARING: Policy Details

• **Scope:** All NIH-supported research generating *scientific data*
  – *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
  – *Does not include* lab notebooks, preliminary analyses, peer reviews, physical objects

• **Timelines:**
  – For *when to share data*, no later than publication or end of award (for unpublished data)
  – For *how long to share data*, consider relevant requirements and expectations (e.g., repository policies, retention requirements, journal policies) for minimum time frames
**SHARING SHOULD BE …**

- **The default practice**
  - Maximize appropriate data sharing; plans may justify exceptions (i.e., ethical, legal, technical factors)
  - All scientific data should be managed; not all scientific data must be shared

- **Responsibly implemented**
  - Plans should outline protection of privacy, rights, and confidentiality
  - Existing laws, regulations, and policies continue to apply

- **Prospectively planned for**
  - During informed consent, including communicating how data will be used and shared
  - Data submission, including whether access to data, even if de-identified, should be controlled
Supplemental Information to the Policy: 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
  - Refers to list of NIH-supported Data Repositories

- NIH ICs may designate specific data repository(ies)
Supplemental Information to the Policy:
Allowable Costs

- **Reasonable costs allowed in budget requests**
  - 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)
NIH POLICY FOR DATA MANAGEMENT & SHARING: Plan Submission and Review

Extramural Grant Awards*

Plan Submission
With application for funding in Budget Justification section

Plan Assessment
Peer reviewers only comment on (not score) budget
NIH program staff assess Plans
Plans can be updated

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 and intramural research
What’s Next?

- **Engage in outreach** to inform development of additional resources and supplemental information (including tribal-specific considerations)

- **Develop FAQs and other resources** to aid policy implementation

- **Clarify interactions** with other NIH data sharing policies (e.g., NIH Genomic Data Sharing Policy)

- **Develop resources** to inform data management and sharing costs (informed by efforts such as the [2020 NASEM report on forecasting costs](https://doi.org/10.17226/25693) & [April 2021 NASEM workshop on the culture of data management & sharing](https://www.nasm.org/))

- **Develop approaches** for incentivizing good data sharing practices
What’s Next?

– **Approaches and workflows** – Determine the appropriate roles, responsibilities, and processes by which ICs will assess Plans and monitor compliance

– **System changes** – Enhance award management systems and develop tools to support the submission, assessment, and compliance monitoring of Plans

– **Public posting of Plans** including how they will link to repositories, employment of persistent identifiers such as DOI, and FAIR principles

– **Planning communications and guidance** to ensure investigators, institutions, and NIH staff are prepared for the Policy
Elements and Common Approaches to Data Management and Sharing Plans

Taunton Paine
Director, Scientific Data Sharing Policy Division
Office of Science Policy, National Institutes of Health

Sara Samuel
Informationist, University of Michigan
NIH POLICY FOR DATA MANAGEMENT & SHARING:
Data Management & Sharing Plans

• Propose approach to data management & sharing
  – Attempt to maximize appropriate sharing
  – Be consistent with FAIR (findable, accessible, interoperable, reusable) principles
  – Be updated throughout the award

• Outline responsible data sharing
  – Plans should outline protection of privacy, rights, and confidentiality
  – Existing laws, regulations, and policies continue to apply

• Reflect applicable NIH Institute, Center, & Office data sharing expectations

• Address elements in 2 pages or less
NIH recommended elements of a Plan:

1. Data type
   - Identifying estimated type and amount of data to be generated (i.e., modality, level of aggregation, and degree of data processing)
   - Which data to be preserved and shared
   - Accompanying metadata, other relevant data, and associated documentation to be made available

2. Related tools, software, code
   - Tools and software needed to access and manipulate data

3. Standards
   - Standards to be applied to scientific data and metadata
NIH recommended elements of a Plan (cont.):

4. Data preservation, access, and associated timelines
   - Proposed repository to be used consistent with Supplemental Information on Repository Selection
   - How data will be findable and accessible (e.g., persistent unique identifier)
   - When data will be made available and for how long

5. Access, Distribution, and Reuse Considerations
   - Description of factors potentially affecting data access, distribution, or reuse related to informed consent or privacy and confidentiality protections
   - Whether access to human data will be controlled

6. Oversight of data management
   - Plan compliance will be monitored/managed and by whom
Examples

1. Data Type
2. Data Preservation, Access, and Associated Timelines
The research team will conduct approximately 20 semi-structured interviews ... Each interview will produce 2-3 pages of hand-typed notes in Word (.docx) format (converted from hand-written notes taken during the interview, which are promptly destroyed) and a one page analytic memo in Word (.docx) format, both convertible to .pdf for long-term storage.

In addition, the research team will conduct analysis of the interview data in the CAQDAS program ATLAS.ti, which will produce a .qdx file, which is the REFI-QDA Standard for interoperability between CAQDAS programs.

The research team may come into possession of additional documents ... These documents will be digitized as .pdf files for long-term storage.

Nicholas Bell, University of Pennsylvania and Georgetown University, “Why Do So Few Workers Take Trade Adjustment Assistance”
Example 2: Data Preservation, Access, and Associated Timelines

**Preservation**

Data will be publicly shared and preserved in the repository Zenodo (https://zenodo.org/).

Zenodo is freely available to anyone to use, and the data will be securely stored in the CERN Data Center. Once deposited in Zenodo and published, the data set will be assigned a DOI and will be findable via a web search, the Zenodo repository search feature, or by the assigned DOI.

**Access**

The data will be published concurrently with the first associated publication, but no later than the end of the award period. The data will be available for at least 20 years; Zenodo’s policies (https://about.zenodo.org/policies/) currently state “Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.”
Recommended practices for plans

• Start early
• Know your data
• Understand applicable requirements & guidelines
• Aim to share data in a repository
• Connect with data librarian
• Write data management and sharing costs into budget
  • Managers, cleaning up the data, sharing
• Harmonize other documents with the Plan
How Academic Medical Centers Can Prepare for January 2023

Sara Samuel
Informationist, University of Michigan
Institutional readiness

- Identify stakeholders
- Identify local resources
- Recognize and understand challenges
- Identify opportunities
- Actions to take now
Stakeholders

- Researchers funded by NIH
  - Need to learn new policy and be able to write a Plan
- Research support staff & units (research/grants administrators, IT, library)
  - Need to learn new policy to help researchers
  - Identify important things to look for in a Plan
  - Provide language to accurately describe storage and sharing options
- Institutional Research Leadership
- Graduate students
  - Policy can inform how they learn to do science
- Research participants
- Taxpayers
Local resources

• Librarians and Informationists
  • Many libraries now have data librarians
    • Identify repositories, provide info about recommended data practices, consult on Plans

• Information Technology
  • Identifying technical infrastructure available for storing/securing data

• Data Office
  • Some campuses may have a data office which can help with crafting data use agreements

• Institutional Data Repository

• Champions – committed and vocal to help with faculty buy-in
Other resources

• Online Tutorials or MOOCs - Learn about data sharing and data management
• Repositories – deposit guides
• DMPTool - [https://dmptool.org/](https://dmptool.org/)
  • Offers institutional affiliations
  • Member institutions can provide tailored guidance
Challenges

• **Educating stakeholders**
• Gathering resources for outreach and implementing
• Implementing effective outreach
  • Compliance vs “good science”
• Leading and coordinating
• Changing culture
  • Think about data at the beginning of a project, rather than at the end
• Addressing ambiguity: who owns researchers’ data?
Opportunities

- Recognizing data as valuable research output
  - Incentives, promotion & tenure
  - More opportunities for researchers to collaborate
- Building collaborations with different offices and groups on campus
- Preparing institution and researchers to fulfill growing data requirements from other funding sources or journals
- Creating or updating an institution-wide data policy
Actions to take now

1. Start conversations
Start conversations

• Who to talk to:
  • Deans & Directors
  • Research administrators
  • Grants administrators
  • Information Technology
  • Library
  • Data Office or honest broker

• What to say:
  • The NIH recently announced a new data policy. Every grant application after January 2023 will need to include a data management and sharing plan. I would like to talk more with you about this. When are you available to meet?
Actions to take now

1. Start conversations!
2. Review institutional policies
Review institutional policies

• Check your campus faculty/researcher handbook/guide
  • Research Data Policy
  • Institutional Data Policy
  • Data Ownership
  • Copyright Policy

• Examples:
  • Harvard University - https://researchdatamanagement.harvard.edu/policies
  • University of Wisconsin-Madison - https://data.wisc.edu/institutional-data-policy/
Actions to take now

1. Start conversations!
2. Review institutional policies
3. Conduct a Services Gap Analysis
Conduct a Services Gap Analysis

What services do you have? What will faculty & staff need?

- Planning for data
- Identifying data repository options
- Documenting data
- Protecting sensitive information
- Understanding funder and publisher data requirements
- Curation, documentation, and management of shared data

How do you get the services to the right people at the right time?

How will compliance be monitored at the institution?

DOI: [https://doi.org/10.31219/osf.io/tjybn](https://doi.org/10.31219/osf.io/tjybn)
Actions to take now

1. Start conversations!
2. Review institutional policies
3. Conduct a Services Gap Analysis
4. Create pathways to information
Create pathways to information

- Frequently Asked Questions
- Resource list or website
  - Point to both local and outside resources (don’t try to reinvent the wheel!)
- Workflows
  - Which service can help with each part of data management & sharing?
### Draft: Where to find information

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Related tools, software, code</th>
<th>Standards</th>
<th>Data preservation, access, and associated timelines</th>
<th>Oversight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of research and expected data [grant proposal?]</td>
<td>Check laboratory computers for program information used to analyze data</td>
<td>Check grant for mention of specific standards</td>
<td>PI decision to share in repository or not</td>
<td>Usually the PI</td>
</tr>
<tr>
<td>Consider ethical and legal restrictions for what you can share</td>
<td>Will you be writing scripts or code to help process the data?</td>
<td>Does your field of study have templates or standards for data collection?</td>
<td>Repository policies can provide helpful info regarding access and timelines</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IT departments can provide description of data storage and backups</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI decision regarding when to share, usually associated with study conclusion</td>
<td></td>
</tr>
</tbody>
</table>
Actions to take now

1. Start conversations!
2. Review institutional policies
3. Conduct a Services Gap Analysis
4. Create pathways to information
5. Make an outreach plan & implement it
Make an outreach plan & implement it

- Identify champions
- Decide on outreach approach – compliance vs “good science”
- Decide on various methods & timelines for each
  - Videos, flyers, digital signs, workshops, meetings, emails, newsletters
- Write an “elevator speech” and/or presentation to build awareness
Actions to take now

1. Start conversations!
2. Review institutional policies
3. Conduct a Services Gap Analysis
4. Create pathways to information
5. Make an outreach plan & implement it

Do you have any resources or additional suggestions? Please share in the chat!
Additional Resources

- re3data (Registry of Research Data Repositories) - [https://www.re3data.org/](https://www.re3data.org/)
- Briney KA, Coates H, Goben A (2020) Foundational Practices of Research Data Management. Research Ideas and Outcomes 6: e56508. [https://doi.org/10.3897/rio.6.e56508](https://doi.org/10.3897/rio.6.e56508)
Thank You - Questions?
Conference Content Copyright

All content included in this session is the property of the presenter(s) and is protected by United States and international copyright laws. Certain materials are used by permission of their respective owners. The course content may not be reproduced, transmitted, or shared in any way without the prior written permission of the presenter(s).

Access to this presentation should not be construed as a license or right under any copyright, patent, trademark or other proprietary interest of PRIM&R or third parties.