SPA Series: Data Management Sharing Plans

What does it mean for Northwestern researchers?
Questions

- What does the NIH DMSP say?
- What does this mean for Northwestern?
- What resources are available to Northwestern researchers?
What is a Data Management Plan?

A description of how you plan to describe, preserve and share your research data?

This is often required by funding agencies.
Funders that require DMPs
Final NIH Policy for Data Management and Sharing

Notice Number:
NOT-OD-21-013

Key Dates

Release Date: October 29, 2020
Effective Date: January 25, 2023

What the policy says

• Effective **January 25, 2023** for all grants
• Data management and sharing plans **submitted with grant applications**
• Evaluated by NIH staff (reviewers can comment)
• “Shared scientific data should be made accessible **as soon as possible**, and no later than the time of an associated publication, or the end of the award/support period, whichever comes first.”
• **Exceptions** for data that cannot be shared must be described in the data management plan
Goal: Make NIH-funded research data...

• **Findable**: searchable and has a unique ID
• **Accessible**: backed up, preferably in a repository
• **Interoperable**: Described in common standards for your field
• **Reusable**: Properly formatted, documented and licensed for reuse
Implementation Details
Elements of an NIH DMSP

1. Data Type
2. Related Tools, Software and/or Code
3. Standards
4. Data Preservation, Access, and Associated Timelines
5. Access, Distribution, or Reuse Considerations
6. Oversight of Data Management and Sharing

## Elements of an NIH DMSP

1. **What data, code etc. will you produce?**
2. **How will you assure others can use your data?**
3. **How will you preserve and provide access?**
4. **How will you make sure you follow your plan?**

Implementation Details

• DMSPs will be submitted in the "other plans" field
• **Additional details** in NIH Application Form Instructions by Fall 2022.
• Insufficient plans may be revised through the **Just in Time** process
• Plans can be updated during annual **Research Performance Progress Report** (RPPR) process

Budget Details

• **Direct costs** to support data management designated as “Data Management and Sharing Costs”
  • R&R Budget Form: single line item in section F. Other Direct Costs
  • PHS 398 Modular Budget Form: as text embedded within the Additional Narrative Justification

• Include summary of DMSP in **budget justification**

What does this mean for Northwestern University?
## NIH-funded research at Northwestern

<table>
<thead>
<tr>
<th>Unit</th>
<th>number_grants</th>
<th>number_PIs</th>
<th>total_cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feinberg School of Medicine</td>
<td>761</td>
<td>407</td>
<td>$514,333,307</td>
</tr>
<tr>
<td>Weinberg College of Arts and Sciences</td>
<td>113</td>
<td>77</td>
<td>$46,248,119</td>
</tr>
<tr>
<td>McCormick School of Engineering</td>
<td>47</td>
<td>31</td>
<td>$22,770,975</td>
</tr>
<tr>
<td>Northwestern School of Communication</td>
<td>25</td>
<td>15</td>
<td>$8,035,478</td>
</tr>
</tbody>
</table>

Data retrieved from NIH reporter by searching “Northwestern University” in the Organization field. Subawards and R&D contracts were removed from the dataset.
The Data Lifecycle

- Plan
- Store
- Analyze
- Visualize
- Archive
- Publish
NIH Policy

Plan

Store

Archive Publish

The Data Lifecycle

Analyze Visualize
Where we support

The Data Lifecycle

Plan
Store

Archive
Publish

Analyze
Visualize
Researcher Perspectives

• Never written a DMSP before
• Need help framing their work from the data perspective
• Biomedical data has privacy concerns
• Varying levels of support
• Compliance could affect future funding
Culture Shift
# Common Misconceptions

<table>
<thead>
<tr>
<th>Objection</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is just another hoop to jump through!</td>
<td>Planning is good for research.</td>
</tr>
<tr>
<td>This is an unfunded mandate!</td>
<td>Write data management costs into your budget.</td>
</tr>
<tr>
<td>What is I can’t comply with what I wrote because something changed about my project?</td>
<td>The DMSPs are living documents. You can update your DMSP to reflect changes in your work.</td>
</tr>
<tr>
<td>I can’t share my data because…</td>
<td>You don’t have to. There are many reasons why data can’t be shared. You just have to explain why.</td>
</tr>
</tbody>
</table>
## Resources

<table>
<thead>
<tr>
<th>Archive/Publish</th>
<th>Plan</th>
<th>Store</th>
<th>Analyze/Visualize</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Zenodo</td>
<td>• DMPTool</td>
<td>• Globus</td>
<td>• Quest Analytics Nodes</td>
</tr>
<tr>
<td>• Dryad</td>
<td></td>
<td>• OneDrive</td>
<td>• Quest High Performance Computing Cluster</td>
</tr>
<tr>
<td>• Microsoft Azure</td>
<td></td>
<td>• SharePoint</td>
<td></td>
</tr>
<tr>
<td>• Amazon Glacier</td>
<td></td>
<td>• Amazon S3</td>
<td></td>
</tr>
</tbody>
</table>
Writing a DMSP

- Templates for writing Data management plans for all funding agencies
- Institute-specific guidance where available
- NIH format page
Document Sharing and Data Storage

https://www.it.northwestern.edu/service-catalog/collab/doc-collab/storage-finder/index.html
Dryad Storage

• **Coming Soon:** Data Sharing Repository Service for Campus

• **About Dryad**
  – Launched 2009
  – Membership started in 2019
  – Non-profit
  – Community governance
  – Open-sourced
  – Based at the California Digital Library
Dryad features continued...

- Recommended by federal funding agencies
- Recommended by scientific journals
- Excellent user experience
- Enforces data sharing best practices
- Promotes FAIR data principles (Findable Accessible, Interoperable, Reusable)

Dryad technical details

- 300 gb per data submission
- Reviewed by Dryad data curators before published
- All datasets are given a DOI and will be indexed by the Thomson-Reuters Data Citation Index, Scopus, and Google Dataset Search.
Data Consultations

• University Libraries support
  – Kelsey Rydland & Sarah Thorngate

• Research computing services (RCS)
  – Tobin Magle

• Submit a support request!
  – bit.ly/rcsconsult
Galter Health Sciences Library & Learning Center's NIH Data Management and Sharing Plan (DMSP) Resources

• Feinberg's resources for consultations on NIH DMSPs:
  – Sara Gonzales, data librarian
  – Galter DataLab’s Data Clinic

• Data management and Bioinformatics trainings:
  – 10 classes ranging from basics of data management to Introduction to SQL for Clinical Researchers
  – Biosciences and Bioinformatics trainings

• LibGuides:
  – Biosciences and Bioinformatics
  – Creating an NIH Data Management and Sharing Plan
  – Data and Data Management

• Class: Getting to Know the NIH Data Management and Sharing Policy

• DigitalHub, Feinberg's institutional repository
  – https://digitalhub.northwestern.edu/
Thank you!

- RCS/Library assistance
- Feinberg assistance
  - [https://galter.northwestern.edu/contact/data/clinic](https://galter.northwestern.edu/contact/data/clinic)