Gearing Up for January 2023: Institutional Strategies for Implementing the NIH Data Management & Sharing Policy

Duke University:
Lindsey Spangler, Associate Dean, Research Integrity, Duke University
Tim McGeary, Associate University Librarian for Digital Strategies & Technology, Duke University

University of Delaware:
Jeffrey Friedland, Associate Vice President for Research Administration
John Huffman, Director of Cyberinfrastructure
Annie Johnson, Associate University Librarian for Publishing, Preservation, Research & Digital Access
Some initial thoughts from ARL and COGR

Today’s Moderators:
Cynthia Hudson Vitale, Director, Scholars & Scholarship
David Kennedy, Director, Costing & Financial Compliance, COGR
Introduction to COGR

- Council on Governmental Relations, est. 1948
  - 200+ member institutions, staff of seven
- COGR is a national expert on the crucial issues around research compliance, ethics, technology transfer, and financial sustainability. We work closely with all federal agencies to ensure the nation’s scientists and PIs can deliver the cutting-edge research necessary to make the world and nation a better place.
- Active Board and Committees, ~45 individuals from member institutions
- Regular collaboration with higher ed associations and other partners!
Background & Challenges


- As presented by NIH on Sept. 13 (@ FDP), the DMS policy applies to/requires:
  1. Applies to all research funded or conducted in whole or in part by NIH, that results in the generation of scientific data.
  2. Submission of a DMS Plan outlining how scientific data and any accompanying metadata will be managed and shared, taking into account any potential restrictions or limitations.
  3. Compliance with the awardee’s plan as approved by the NIH Institute and Center (IC).

- But the anticipated challenges are real:
  1. Varying expectations from varying ICs – will there be harmonization?
  2. Enforcement – what if the submitted plan is deemed inadequate?
  3. New administrative and cost burden – how to budget / how to pay?
COGR’s Approach

- **Education and Other Resources, including:**
  2. NIH Policy Summary, via the Matrix
  3. Readiness Guide (Living Document) – Culture, Roles & Responsibilities, etc.

- **Administrative and Cost Burden**
  1. Hypothesis – The burden is real and significant
  2. Cost Impact Survey – 40+ institutions, data analysis starts next week

- **Advocacy to NIH, including:**
  1. Harmonization
  2. How to Pay
  3. Enforcement
  4. Other Outreach, as we learn more
Association of Research Libraries
Convenes, Shapes, Informs and Influences for Systemic Change

Vision
Mission
Relationships
Resources
ARL’s Approach

➢ Advocacy
• Machine-readable DMP’s
• Persistent identifiers
• Well-documented and/or curated research data sharing
• Harmonization of policies

➢ Research and Resources
• Realities of Academic Data Sharing https://www.arl.org/realities-of-academic-data-sharing-rads-initiative/
  • Retrospective study of expenses to make research data publicly accessible in 5 disciplines across 6 institutions = ~3000 researchers (covering NIH, NSF, and DOE)
  • Campus mapping of institutional infrastructure and expenses for research data services at 6 academic institutions
• Implementing Effective Data Practices https://www.arl.org/implementing-effective-data-practices/
• Resources: https://www.arl.org/research-data-management-and-sharing-policies/ (more forthcoming)
COGR–ARL Collaborative Approach

- Cross-organization representation
  - COGR NIH DMS Working Group
  - ARL RADS Advisory Committee

- Resources and Webinars
  - Additional materials continue to be developed
  - Today’s Webinar and future sessions (as needed)
  - Other DMS activities and tasks (forthcoming)
Jeffrey Friedland, Associate Vice President
For Research Administration

John Huffman, Director of Research
Cyberinfrastructure

Annie Johnson, Associate University Librarian
for Publishing, Preservation, Research
& Digital Access
Supporting the New NIH DMS Policy
@ The University of Delaware

Jeff Friedland, John Huffman, and Annie Johnson
Jeffrey Friedland
Associate Vice President, Research Administration

John Huffman
Director of Research Cyberinfrastructure

Annie Johnson
Associate University Librarian for Publishing, Preservation, Research and Digital Access
Presentation Overview

- Background
- Data Management Support Roles
- Logistics
- Anticipated Challenges
- Education and Outreach
- Data Cyberinfrastructure
- Data Security, Storage and Access
University of Delaware

- 18,600 undergraduate students / 4,300 graduate students / 1,300 faculty
- 16 universities to hold the “triple crown” – Land Grant, Sea Grant and Space Grant designations
- R1: Doctoral Universities
- 2021 Research Expenditure Data $184 Million
FY21 Sponsored Research Expenditures

$184.3 MILLION

- 12.6% Department of Commerce
- 9.1% Department of Defense
- 6% Department of Energy
- 21.8% Health and Human Services
- 14.1% National Science Foundation
- Other
- 8.7% Other Federal
- 7.5% State of Delaware
- 18.6% Federal Flow-Through
A Few Facts from FY21

- 2091 Proposals
- $1.71B Requested in Funding
- 680 New Awards Received
- $335M New Awards
- 1958 Active Awards
Data Management Support Roles

Research Office

- Serves as the connector to other units
- Reviews to be sure DMP are submitted at proposal stage
- Works with PI/unit to budget for allowable data sharing/management costs
Data Management Support Roles

Library, Museums and Press
- Data Services Librarian (Daniel Peart)
- Research Data Steering Group
  - Library, Research Office, and IT
- UDSpace
  - Not a data repository!
Data Management Support Roles

Central IT

- Primary focus on infrastructure
- Assesses data security needs
Logistics

Research Office

● Connecting the dots; took lead on organizing meetings with Library and IT to assemble a plan
● Shared relevant resources, like those from COGR
Anticipated Challenges

- Awareness
- Last-minute nature of many NIH proposals
  - Inability to help develop a plan
  - Costing
- Lack of robust data storage options at UD
- Lack of understanding of secure data handling
Education and Outreach

- Lots of meetings!
- NIH DMSP landing pages on Library and Research Office websites
- Articles in Research Office newsletter
- Customizing DMPTool
Education and Outreach

- Promote DataWorks! DMP Challenge
- Integrating information into existing data workshops
- Conducting internal training for liaison librarians and Research Office staff
- Joint memo planned from Library/IT/RO
Data Cyberinfrastructure

- Infrastructure costs are not covered
- Projects with less funding may have less access to existing infrastructure
- Data storage is not data management
Data: Security, Storage and Access

Level III: High Risk
Encrypt at rest and transit, restricted and monitored access, explicit requirements

Level II: Moderate Risk
Share only with those who “need to know”

Level I: Low Risk
May be shared publicly

Security needs will help drive infrastructure choices

Domain Specific Repositories

Cloud Based Services and Infrastructure

Institutional Resources
Data Storage and Retention Costs

- Costs go up based on security needs
- DM and storage needs persist beyond funding
- Potential hidden costs of Cloud SaaS, IaaS and other cloud services
- UD is currently enabling NIH STRIDES access for commercial cloud vendors
Things to Consider

- Start now
- Collaboration is key
- One size does not fit all
Thank you!

For more information, contact:

Jeffrey Friedland: jeffreyf@udel.edu
John Huffman: jnhuffma@udel.edu
Annie Johnson: akjohnso@udel.edu
Research Data Management at Duke

Gearing up for January 2023

Tim McGeary, Associate University Librarian for Digital Strategies & Technology
Lindsey Spangler, Associate Dean, Research Integrity

Hosted by COGR and the Association of Research Libraries

September 30, 2022
Short History at Duke
2015 - NSF rejects Duke research proposal due to insufficient DMP, specifically due to lack of plans to deposit data

2015 - University Libraries, Trinity College of Arts & Sciences, and Office of Information Technology (OIT) collaborate on white paper to Provost and charge for Faculty Working Group for Digital Research Data Services

2016 - Interdisciplinary faculty working group meets for 9 months and submits recommendation to Provost; Provost accepts and funds recommendations

2017 - Research Data Services positions (four) hired by Duke University Libraries for university-wide partnership and services

2018 - Research Data Repository (RDR) launched

2022 - 200th data set published in RDR
Our mission is to curate, publish, and archive Duke digital research data from any discipline. We provide long-term public access to support research transparency, reproducibility and to foster new discoveries.
“Archiving data this way ensures reproducibility of the underlying science and long-term stability of the work, despite the natural turnover in my research group’s composition. It thus brings our research practice in compliance with data management requirements of funding agencies, and helps us respond promptly to external requests for our data. In addition, the very process of depositing gives us the opportunity to validate figures one last time, as we copy-edit page proofs.”

Patrick Charbonneau, PhD
Professor of Chemistry & Physics
Including several datasets published prior to 2017, RDR now hosts 205 datasets.
CDVS provides multiple programs supporting data driven research

- Data Science
- GIS & Mapping
- Data Visualization
- Data Management
Duke Office of Scientific Integrity

Established ASIST Office (Advancing Scientific Integrity, Services, and Training) in 2017

• Created a Data Management Plan Guidance document to help researchers develop a DMP

• Held weekly office hours and offered one-on-one support to assist in DMP development

• Identified resources to support data management
Research Data Policy Initiative

Started in spring 2020

https://research.duke.edu/research-data-policy-initiative

Facilitating efficient and quality research, ensuring data integrity, and fostering a culture of data sharing

For the past two years the Duke School of Medicine and the University have been working together strategically to develop a unified research data policy.

The timing with the NIH policy was advantageous in that some of the key research support stakeholders for our institutional policy are also those who will be most likely helping to build and support a workflow for NIH grants, paving the way for ongoing coordination for the NIH policy.
Education and Outreach
Training and Consultations

The School of Nursing sought presentations and training on data sharing policy.

Center for Data and Visualization Sciences (CDVS) partners with Duke Graduate School and Duke Office of Scientific Integrity to provide data management and sharing topics as part of RCR series.
Data Sharing Education

Total number of workshop attendees 2468. 2022 workshop series ongoing.
Researchers Speaking to Researchers

Dr. Mark Palmeri: An honest assessment of openness

This post is part of the Duke Research Data Curation Team’s ‘Researcher Highlight’ series.

In the field of engineering, a key driving motivator is the urge to solve problems and provide tools to the community to address those problems. For Dr. Mark Palmeri, Professor in Biomedical Engineering at Duke University, open research practices support the ultimate goals of this work, and helps get the data into the hands of those solving problems: “It’s one thing to get a publication out there and see it get cited. It’s totally another thing to see people you have no direct professional connection to accessing the data and see it impacting something they’re doing...”
Researchers Speaking to Researchers

Benefits of Data Management Plans in Clinical Research
A Data Management Exemplar Series Interview With Ryan Shaw, Ph.D., R.N.

https://research.duke.edu/story/benefits-data-management-plans-clinical-research
Making data sharing visible
Connections and Collaborations
...with the Research Community

<table>
<thead>
<tr>
<th>Category</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy</td>
<td>• Make it <strong>required</strong></td>
</tr>
<tr>
<td>Incentives</td>
<td>• Make it <strong>rewarding</strong></td>
</tr>
<tr>
<td>Communities</td>
<td>• Make it <strong>normative</strong></td>
</tr>
<tr>
<td>User interface</td>
<td>• Make it <strong>easy</strong></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>• Make it <strong>possible</strong></td>
</tr>
</tbody>
</table>

Adapted from Mellor Talk at BMTS meeting, January 2020.
... with Duke Research Support

- Clear and frequent communication is a key strategy for coordinating our response to the NIH policy
- Research support stakeholders that have been working directly with us on determining the NIH DMSP support landscape include:
  - Office for Research and Innovation
  - Office of Scientific Integrity
  - Office of Research Initiatives
  - Research Administration Offices
  - Duke University Medical Center Library and Archives
  - Duke University Libraries
- Shared resources and guides (e.g. the AAU-APLU guide, the COGR Guidelines, NIH Guidelines) as well as literature on emerging practices have also been useful tools to build shared understanding of the landscape, institutional resources, and next steps for increasing public access to data
- More planning and potential projects will likely emerge throughout the fall and into the launch of the actual requirements.
... with Information Professionals

- The Libraries Research Data Working Group includes representation from professional school libraries at Duke collaborating on how to best support liaison librarians working with researchers during implementation of the policy.

- Our community of practice has been asked by the National Library of Medicine and NIH to help educate NIH program officers, researchers, librarians and research support staff on what it means to curate data for FAIRness.
myRESEARCHhome: a Duke-created portal (with support from CTSA*) that helps members of the research community manage their research projects and compliance requirements in one place.

*UL1TR002553
myRESEARCHhome Data Management Planning request links to a REDCap survey where researchers can request additional assistance related to:

- Developing a data management and/or sharing plan for a funder requirement, service center, or a project/research group
- Depositing data into a repository
- Sharing or curating data
- Other needs

Survey is submitted to a central e-mail account which is managed by ASIST and CDVS who then reach out for follow-up consultation.
myRESEARCHpath: a Duke-developed roadmap (with support from CTSA*) that navigates individuals through the research project lifecycle and includes processes, policies, and resources to allow individuals to find information needed to facilitate their research.

*UL1TR002553
... Duke Researchers to Resources

The myRESEARCHpath page for Data Management Plan development includes many resources:

- Online learning
- Example DMS plans
- Policies, procedures, and guidance
- Training and job aids
- Resources for consultation and assistance
Budget Considerations

• NIH's website includes a list of allowable and unallowable costs—many of these functions may not exist within one unit at your institution—at Duke, we have experts/services in needed areas, but they exist across multiple offices and may even only be available within a specific school.

• Duke is looking at the most efficient, compliant, and cost-effective way to provide the necessary support to investigators with discussions ongoing in the following areas (non-exhaustive list):
  • Existing infrastructure that will require additional IT, data management, and research administration resources
  • A central service center to streamline the process and ease of use
  • The "Administrative Cap" and how to appropriately manage competing requirements with limited resources
  • Personnel that have expertise in proper costing, data management, and research administration are few and far between
Budget Considerations

COGR is working on two items that will help inform decisions related to costing for DMSP:

**Roles & Responsibilities Matrix**
- List of specific activities required across the data lifecycle
- Timing of requirements
- Role(s) for activities

**Administrative and Cost Burden Survey**
- Burden and resourcing needs for each stage of the data lifecycle
- Ability to answer for all levels of the institution from Grad Students and PIs to central offices, Deans, and VPRs
Compliance Monitoring

• We anticipate our pre-award office could review Plans to ensure the required elements are included before submission, but this would not include the adequacy of the content.

• NIH stated that they will monitor compliance with Plans over the course of the funding period via RPPR and that noncompliance with Plans may result in special award conditions or termination of the award (see FAQ D.1.)

• Duke will likely include some level of Plan monitoring as part of our general compliance review.

• Ongoing discussion as to how we review plan-by-plan compliance, but we hope through our "Intent-to-Submit" initiative, we can pick projects up early and guide them to appropriate resources before proposal submission.
DMP Boilerplate Language

The data will be deposited into the Duke Research Data Repository (RDR) an openly accessible preservation archive maintained by the Duke University Libraries. The RDR will assign appropriate metadata (Dublin Core) for discoverability and provide a Digital Object Identifier (DOI) for persistent access and unique identification of the data. Reuse conditions and expectations will be communicated to end users through the assignment of a standardized Creative Commons license or waiver.

The data will be preserved in the RDR for the long-term according to RDR policies and procedures. When the data are transferred to the RDR, data curators will review deposits to help ensure they are complete and in a structure and format that supports long-term preservation, access, and reuse. The RDR provides for automated backup of all data, which provides an added layer of protection and security for the data.
Curation in Practice

Data curation includes activities that may be taken to ensure data are fit for purpose and available for discovery and reuse.

✓ Open files
✓ Check documentation for completeness
✓ Assess file formats
✓ Perform high-level disclosure review
✓ Identify other data and code enhancements
Partnerships to support data sharing
Thank you!

Tim McGeary, Associate University Librarian for Digital Strategies & Technology
Lindsey Spangler, Associate Dean, Research Integrity
Hosted by COGR and the Association of Research Libraries
September 30, 2022