NIH Data Management and Sharing Getting Ready

June 9, 2022

COGR

An Association of Research Institutions

Panelists

Michelle Christy, Interim Director, Contracts & Grants Administration Committee

Cynthia Hudson Vitale, *Director, Scholars and Scholarship at the Association of Research Libraries (ARL)*

Yvette Seger, *Director of Science Policy at the Federation of American Societies for Experimental Biology (FASEB)*

Twila Reighley, *Associate Vice President for Research and Innovation at Michigan State University*

Current Status

Current status

- NIH policy dates to 2003; OSTP mandate from 2008
- some policies are in place; additional policies are coming; expect revisions along the way
- huge opportunities for the advancement of science with expanded data sharing

NIH-funded researchers will need to collect and manage results with the intention of sharing the data no later than the end of the award (often throughout the life of the award) – (excludes F, T, K, infrastructure – other IC changes)

Challenges

- assess the current state – educate PIs & change management
- multiple offices are likely involved – is there a team?
- costs will vary and could be significant
- monitoring & compliance requirements will evolve

COGR

Data Management & Sharing Working Group

People:

Melissa Korf (*Harvard*), Cynthia Hudson Vitale (*ARL*), Tom Burns (*JHU*), Stephanie Endy (*Brown*), Jennifer Lassner (*U-Iowa*), JR Haywood (*MSU*), Suzie Allard (*UTK*), Mike Legrand (*UC-Davis*), Joe Gindhart (*Wash-U St. Louis*), Alicia Reed (*KU*), Gina Cregg (*KU*), Alessia Daniele (*Cornell*), Jeff Silber (*Cornell*), Walter Goldschmidts (*CSHL*), Lizbet Boroughs (*AAU*), Jim Luther (*FDP*), Toni Russo, Michelle Christy, David Kennedy (*COGR*)

Goals:

Education & Resources - assisting institutions in complying with the requirements

Advocacy - implementation issues, e.g., harmonization across NIH ICs, monitoring the costs that just be borne by the institutions, and other issues as they arise

Cost of Compliance - cost of compliance survey and report

COGR

What is your level of "readiness" for the NIH DMS implementation?

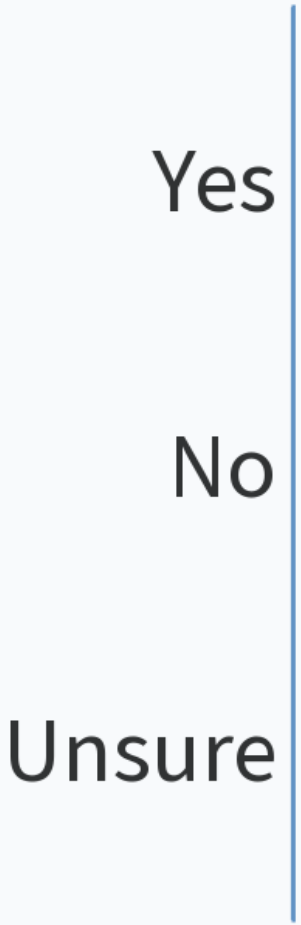
We're just starting this work at my institution

We're underway and making progress

We've got this - we'll be ready

I don't know, but I hope someone else is on this.

Are you involved in your institution's data management & sharing efforts?



Readiness Guide

Chapters:

1. Briefing Sheet – Released!

2. Policy Matrix – End of June

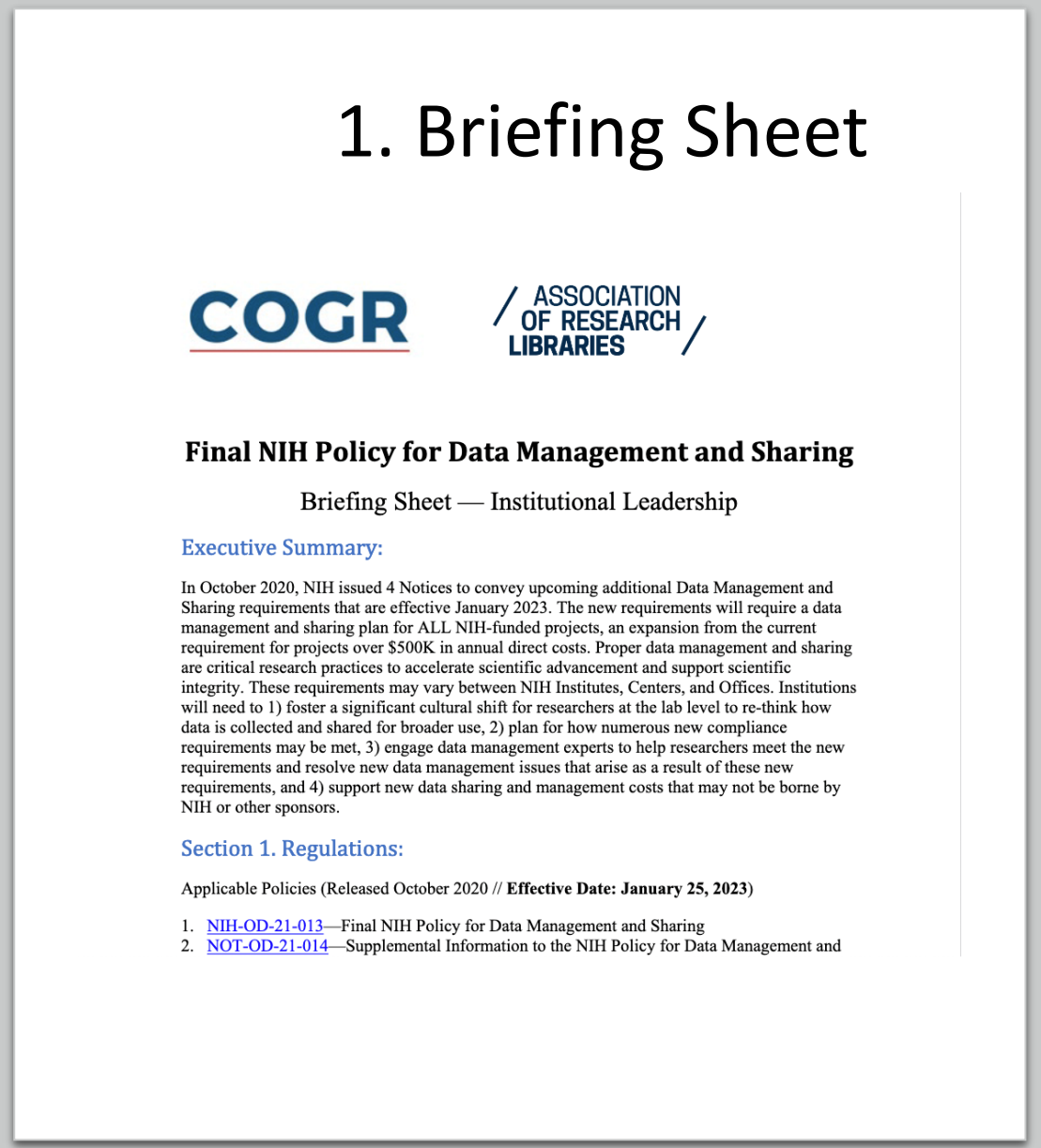
3. Roles & Responsibilities – on deck

4. Culture Change – summer release

5. Costing Issues – summer release

Other topics: data management & storage, DMS plans, human subject research, research security & data sharing, monitoring & compliance

1. Briefing Sheet



COGR / ASSOCIATION OF RESEARCH LIBRARIES

Final NIH Policy for Data Management and Sharing

Briefing Sheet — Institutional Leadership

Executive Summary:

In October 2020, NIH issued 4 Notices to convey upcoming additional Data Management and Sharing requirements that are effective January 2023. The new requirements will require a data management and sharing plan for ALL NIH-funded projects, an expansion from the current requirement for projects over $500K in annual direct costs. Proper data management and sharing are critical research practices to accelerate scientific advancement and support scientific integrity. These requirements may vary between NIH Institutes, Centers, and Offices. Institutions will need to 1) foster a significant cultural shift for researchers at the lab level to re-think how data is collected and shared for broader use, 2) plan for how numerous new compliance requirements may be met, 3) engage data management experts to help researchers meet the new requirements and resolve new data management issues that arise as a result of these new requirements, and 4) support new data sharing and management costs that may not be borne by NIH or other sponsors.

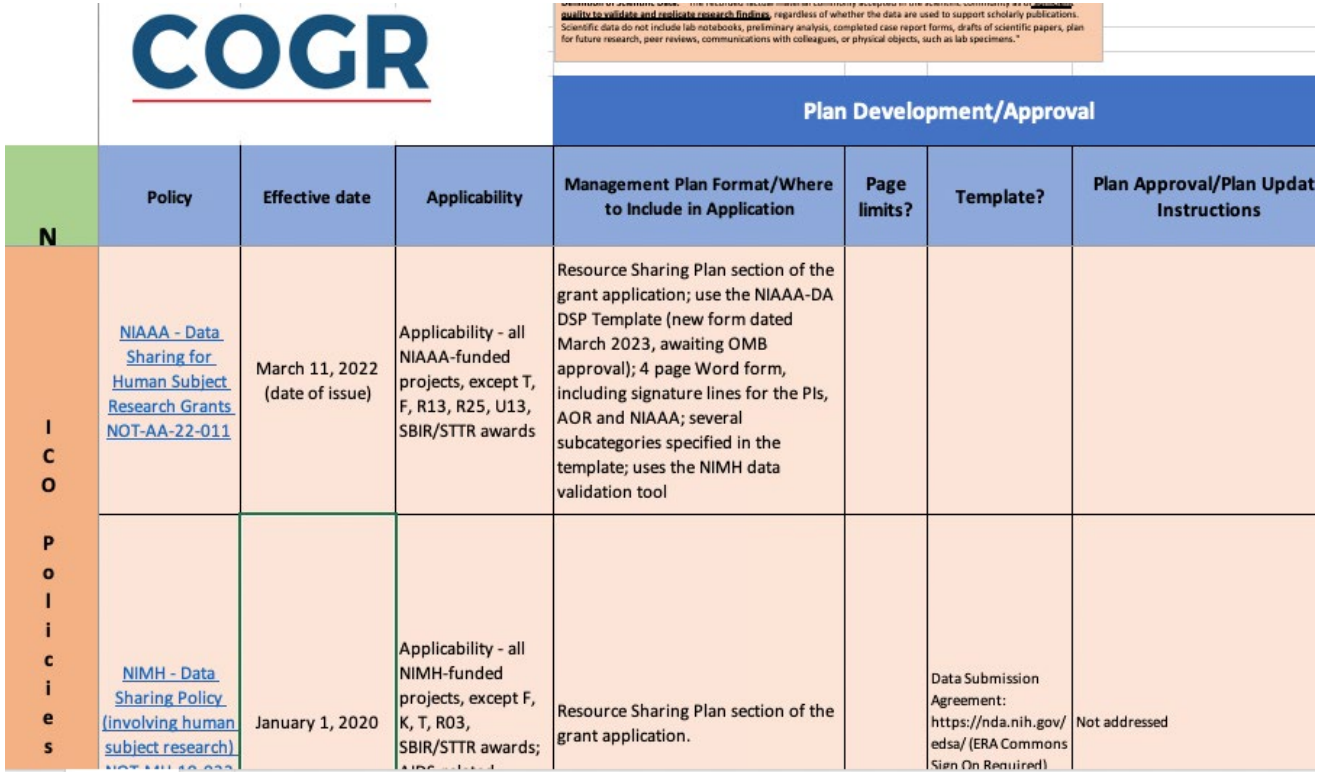
Section 1. Regulations:

Applicable Policies (Released October 2020 // **Effective Date: January 25, 2023**)

1. NIH-OD-21-013—Final NIH Policy for Data Management and Sharing
2. NOT-OD-21-014—Supplemental Information to the NIH Policy for Data Management and

COGR

2. NIH Policy Matrix



COGR

Scientific data do not include lab notebooks, preliminary analysis, completed case report forms, drafts of scientific papers, plan for future research, peer reviews, communications with colleagues, or physical objects, such as lab specimens."

				Plan Development/Approval			
N	Policy	Effective date	Applicability	Management Plan Format/Where to Include in Application	Page limits?	Template?	Plan Approval/Plan Updat Instructions
ICO Policies	NIAAA - Data Sharing for Human Subject Research Grants NOT-AA-22-011	March 11, 2022 (date of issue)	Applicability - all NIAAA-funded projects, except T, F, R13, R25, U13, SBIR/STTR awards	Resource Sharing Plan section of the grant application; use the NIAAA-DA DSP Template (new form dated March 2023, awaiting OMB approval); 4 page Word form, including signature lines for the PIs, AOR and NIAAA; several subcategories specified in the template; uses the NIMH data validation tool			
	NIMH - Data Sharing Policy (involving human subject research)	January 1, 2020	Applicability - all NIMH-funded projects, except F, K, T, R03, SBIR/STTR awards;	Resource Sharing Plan section of the grant application.		Data Submission Agreement: https://nda.nih.gov/edsa/ (ERA Commons Sign On Required)	Not addressed

- Includes ”Final policy” from 2021, 4 supplements, and FAQs
- NIH Institutes and Centers are issuing their own implementations and specifics (NIMH 2019, NIAAA 2022)
- Reminder – policies effective no later than January 2023
- Estimated release - end of June.

COGR

3. Roles and Responsibilities

#	Activity				Role									
	Lifecycle Public Data Access Activities	Lifecycle / Timing	Reference	NIH / Peer Review / Program Staff	VPR	PI	Library	Postdoc / Grad Student	Dept. Grant Support Staff	Pre-award (Central Office)	Central Oversight & Monitoring	Proc. / Other	IT	Costing
1	**DMP Development**		NIH Notice and IC specific			A	C	R						
a	Review IC specific data sharing expectations (e.g., scientific data to share, relevant standards, repository selection, timelines) that apply and should be reflected in a Plan	At Proposal												
b	Ensure all Elements from Notice are addressed (Data Type, Related Tools, Software and/or Code, Standards, Data Preservation, Access, and Associated Timelines, Access, Distribution, or Reuse Considerations, Oversight of Data Management and Sharing)	At Proposal												
c	Adjust DMS plan as needed	JIT												
d	Peer reviewers may comment on the proposed budget for data management and sharing (comments do not impact the overall score)	Peer Review		Peer Reviewers										
e	NIH Program Staff review of Plans, review of updates, and compliance monitoring.	Peer Review		NIH Program Staff										
f	DSM Plan included in Award T&C's	Award Start-Up												
g	DSM adjusted as needed based on science (updated during the course of the award/support period to reflect any changes in the management and sharing of scientific data)	Life												

Tentative Plan: COGR to publish a Blank Excel Sheet that Institutions can customize "Roles", "Responsibilities", and associated level of detail using the Responsibility assignment matrix: Responsible, Accountable, Consulted, and Informed

3. Roles & Responsibilities – Planned Process

Through Working Group Volunteers and Engagement with Institutions

1) **Evaluate Life-Cycle Construct – Current Plan**
   a) DMP Development
   b) Develop Budget requests (Curating data/developing supporting documentation, Preserving/sharing data through repositories, & Local data management considerations)
   c) Data Curation & Metadata Curation FAIR, Data dictionary, etc.
   d) DMP Monitoring & Compliance through closeout
   e) Data Storage (during life of project)
   f) Data Storage (post-closeout for publication)

2) **Determine Level of detail of Responsibilities**

3) **Continue to Refine "Considerations" Document**

4) **Issue Draft to Membership *(Target for Version 1 = July 15th)***

5) **Revise and Update Accordingly**

Volunteers Welcome

COGR

4. Cost of Compliance Survey

- Similar approach as Research Security
  - New hires, Effort, IT, Training, etc.
- User-friendly survey & tool (Alchemer)
- All COGR members are encouraged to participate (Summer kick-off) – contact dkennedy@cogr.edu
- At the core is “How to Pay?”

New DMS Website

Updated as new NIH policies become available

Readiness Guide chapters will be released here as they become available

Feedback always welcome

Academic Libraries: partners for the implementation of the NIH Data Management and Sharing Policy

Cynthia Hudson Vitale

Director, Scholars and Scholarship

ASSOCIATION OF RESEARCH LIBRARIES

The Association

- Founded December 1932
- 126 members from libraries and archives representing research intensive institutions
- 20 expert staff
- One mission

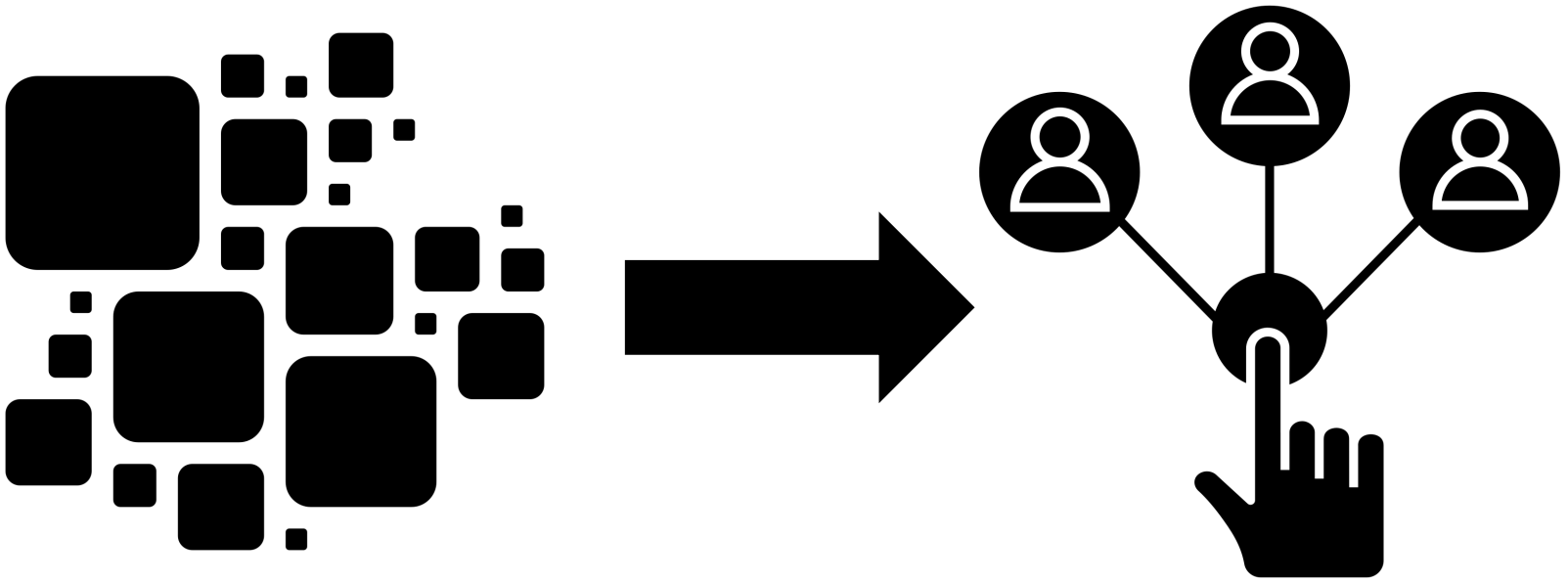
ARL Convenes, Shapes, Informs and Influences for Systemic Change

Vision
Mission
Relationships
Resources

Research Data Lifecycle

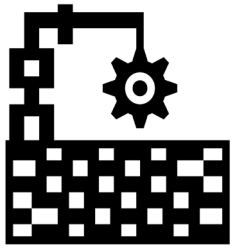
Lifecycle viz developed by Harvard Longwood [LMA Research Data Management Working Group]

Campus Coordination

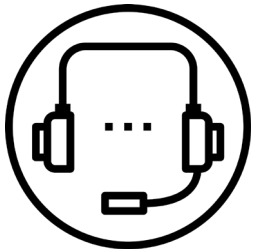


ASSOCIATION OF RESEARCH LIBRARIES

Tactics for supporting the NIH Data Management and Sharing Policy

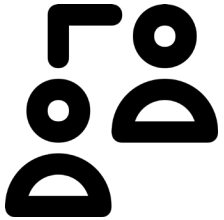


Infrastructure



Services & Processes

Training & Education

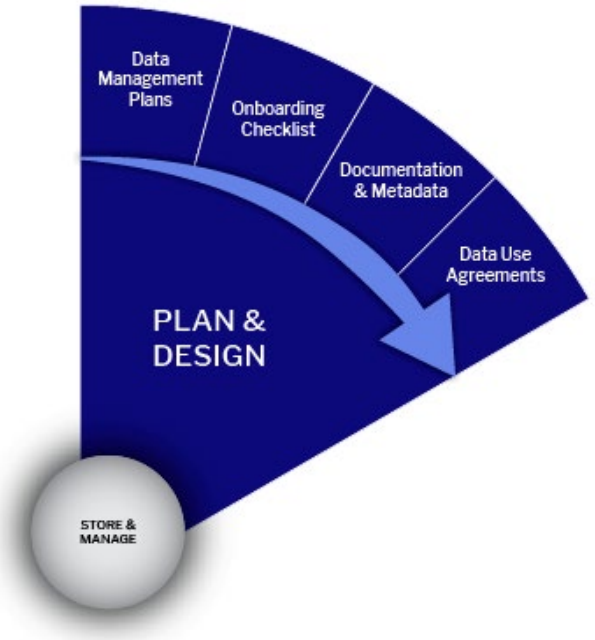


Coordination & Communication

Governance & Compliance

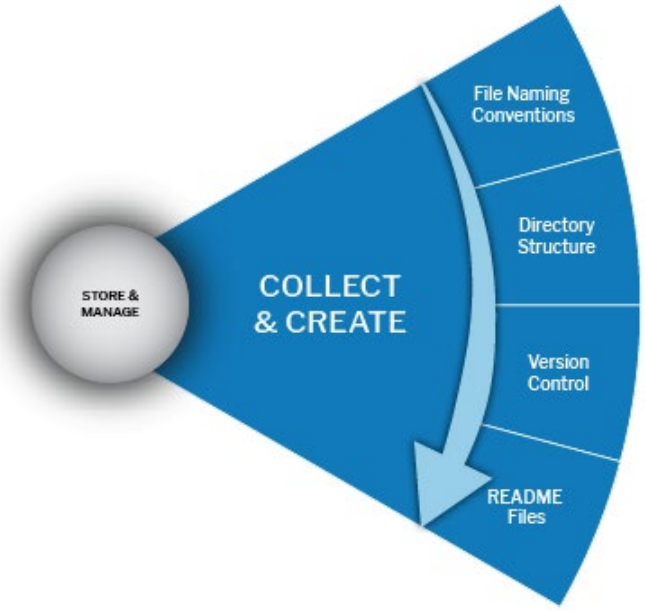
Joint data collection between AAHSL, AAMC, and ARL

Plan & Design



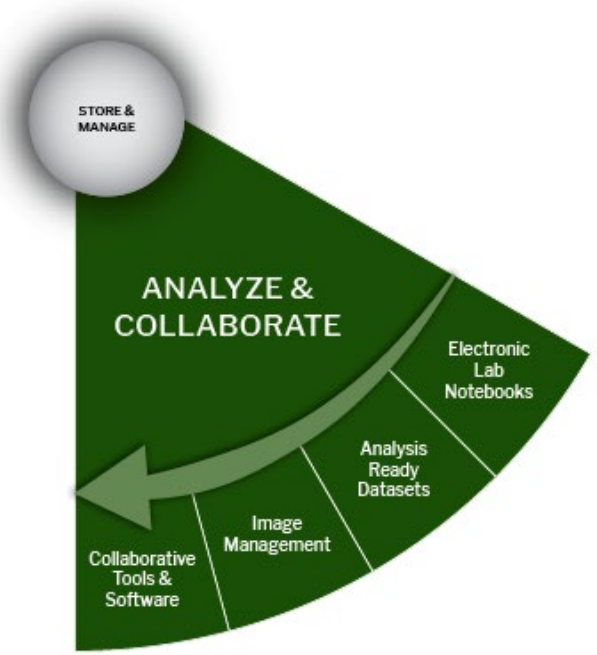
- Research project management and developing research skills
- Ethics, copyright and compliance
- Data management planning

Collect & Create



- Active data management
- Literature review searching
- Citation management
- Text and data mining

Analyze & Collaborate

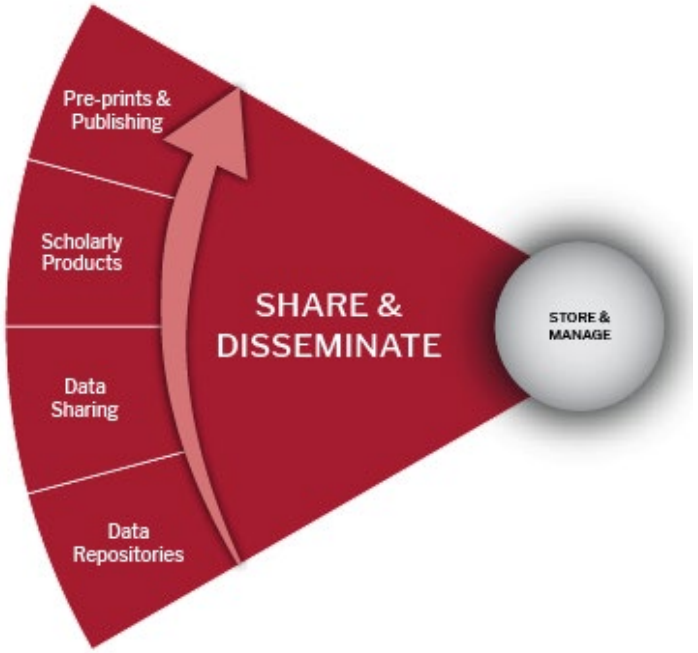


- Data transfer
- Collaboration tools
- Data analysis and visualization

Evaluate & Archive

- Long-term data retention
- Licensing/IP
- Data destruction

Share & Disseminate



- Public access
- Data curation
- Data sharing through repositories

Access & Reuse

- Journal and data metrics
- Preservation

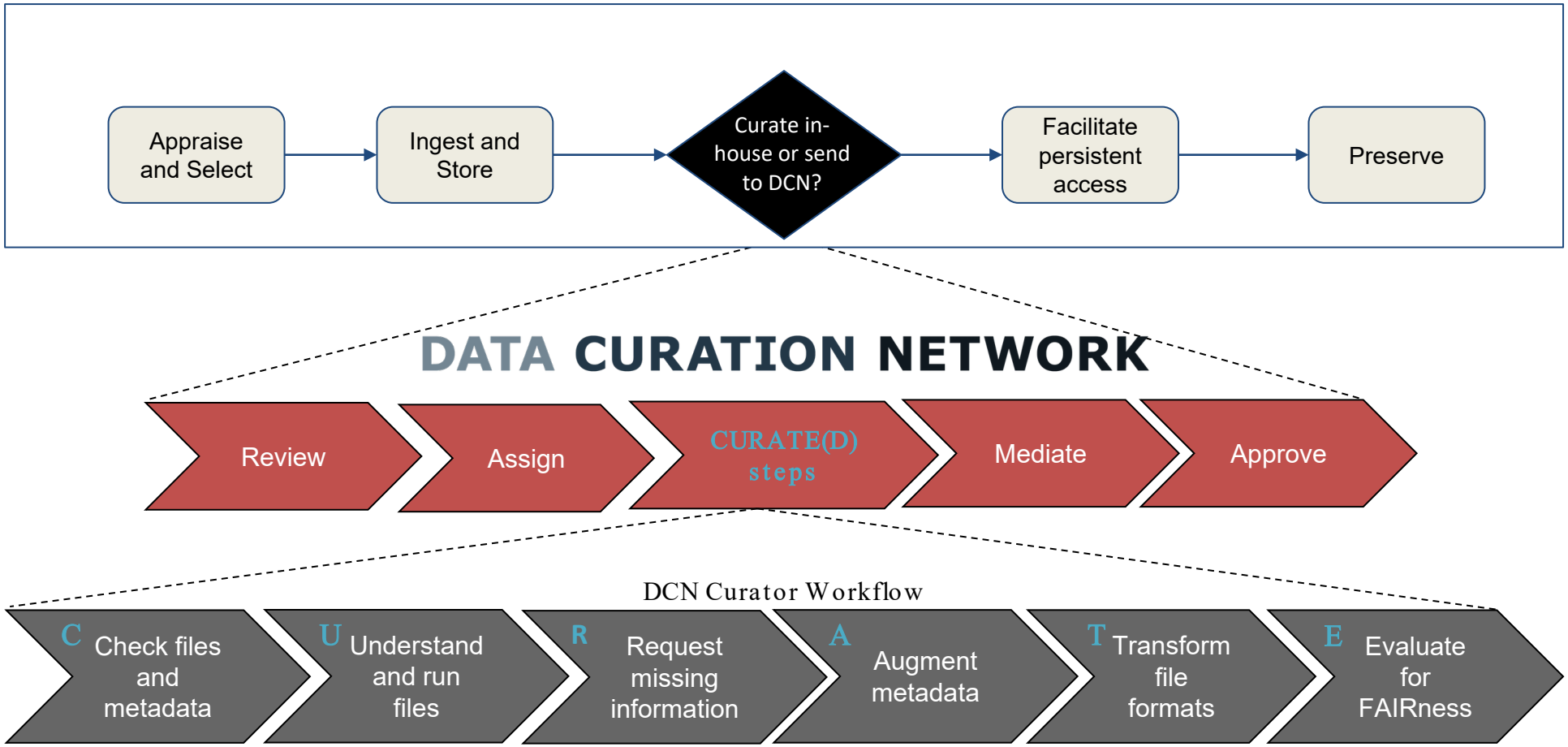
Cross-institution Coordination

DATA
CURATION
NETWORK

Trusted, community-led
network of curators
advancing open research
by making data

Ethical. Reusable. Better.

datacurationnetwork.org


Appraise and Select
Ingest and Store
Curate in-house or send to DCN?
Facilitate persistent access
Preserve
DATA CURATION NETWORK
Review
Assign
CURATE(D) steps
Mediate
Approve
DCN Curator Workflow
C Check files and metadata
U Understand and run files
R Request missing information
A Augment metadata
T Transform file formats
E Evaluate for FAIRness
*CURATE(D) = Document curation process throughout

ASSOCIATION OF RESEARCH LIBRARIES

Realities of Academic Data Sharing (RADS): Research Phases

1. Assess public access to research data repository use
2. Conduct a retrospective study of public-access research data practices of faculty (5 disciplines) on academic campuses
3. Collect financial information on expenses related to public access to research data

Within 5 specific disciplines:

- **environmental science, materials science, psychology, biomedical sciences, and physics**

Realities of Academic Data Sharing research has been generously funded by NSF EAGER grant #2135874:
Completing the Lifecycle: Developing Evidence Based Models of Research Data Sharing

ASSOCIATION OF RESEARCH LIBRARIES

Expected Outputs

- Data and information about where funded researchers are sharing their research data – along with a workflow for other institutions to do the same
- Models for institutional support for public access to research data
- Disciplinary case studies and decision-making factors influencing public access to research data
- Data, information, and case studies on costs for public access to research data and the possible differentiators to those expenses

Realities of Academic Data Sharing research has been generously funded by NSF EAGER grant #2135874:
Completing the Lifecycle: Developing Evidence Based Models of Research Data Sharing

Thank You!

cvitale@arl.org

www.arl.org

ASSOCIATION OF RESEARCH LIBRARIES

FASEB DATAWorks!
Advancing Data Sharing and Reuse

FASEB DataWorks!

Building a Culture of Data Sharing and Reuse

Yvette R. Seger, PhD

FASEB - *28 societies representing over* ***115,000 scientists***

FASEB DataWorks!

A new initiative that brings the biological and biomedical research communities together to advance human health through data sharing and reuse.

The new workspace consists of 4 program areas, initially supported by a $1.5M investment by FASEB

DATAWorks! FASEB Salon

DATAWorks! FASEB Prize

DATAWorks! FASEB Community

DATAWorks! FASEB Help Desk

DATAWorks! FASEB
Advancing Data Sharing and Reuse

Bringing the Community Together

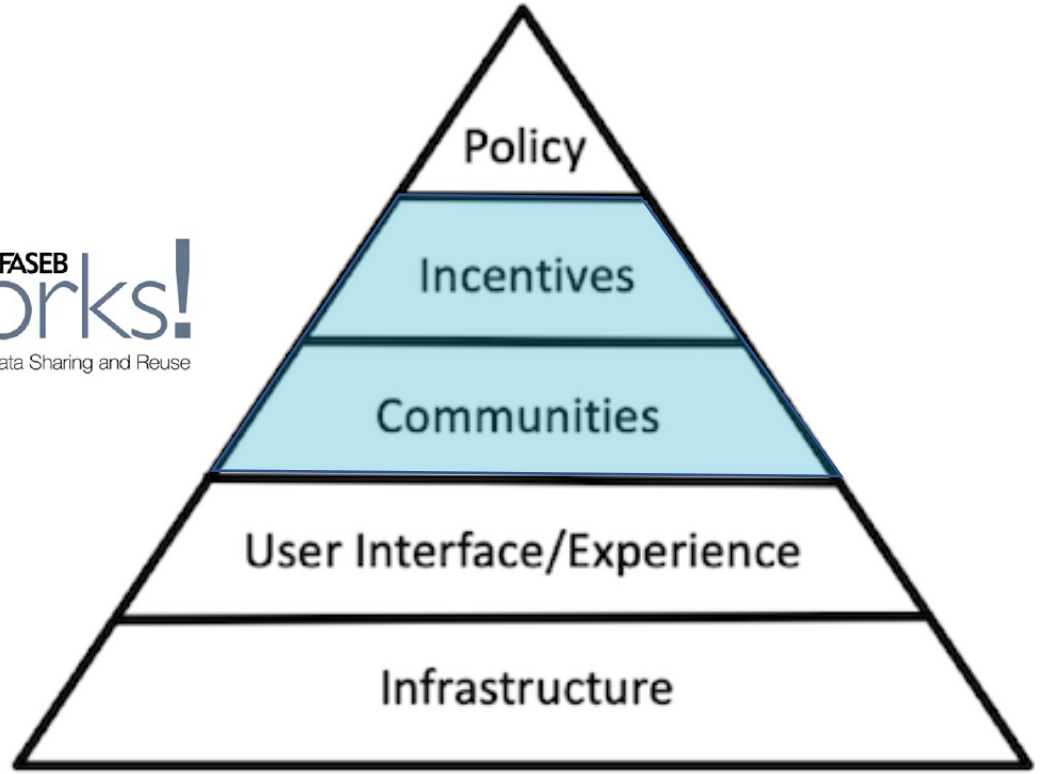
We know data management and sharing is possible.
The expertise is out there. The benefits are clear.

DataWorks! is a convener.

We bring the biological and biomedical community together to advance data management and sharing.

FASEB DATAWorks!
Advancing Data Sharing and Reuse

Building a Culture of Data Sharing and Reuse



Make it required

Make it rewarding

Make it normative

Make it easy

Make it possible

Listening to the Scientific Research Community

Opportunities
Acknowledgement
Financial Support
Infrastructure
Ease of Use
Training
Collaboration

Barriers
Cost
Data Format
Staff Resources
Training
IT/Software
Storage

Symposium and listening tour:

- Individual Researcher
- University Administrator
- Scientific Societies
- Funding Organizations

FASEB DATAWorks!
Advancing Data Sharing and Reuse

Listening to the Scientific Research Community

Opportunities
Acknowledgement
Financial Support
Infrastructure
Ease of Use
Training
Collaboration

DATAWorks! Prize

Annual recognition prize for scientific discoveries made possible through data sharing and reuse

DATAWorks! Salon

Monthly conversation spaces for researchers to learn and engage together

DATAWorks! Advancing Data Sharing and Reuse

DATAWorks! FASEB Prize

Highlighting the Power of Data Sharing and Reuse in the Biological & Biomedical Sciences

$500,000 Prize Purse

Up to 12 monetary prizes recognizing team achievement in data sharing or reuse practices

www.herox.com/dataworks

Register by June 28, 2022

Submissions Open May 11 – July 19, 2022

DataWorks! Prize is a partnership between FASEB and NIH

NIH National Institutes of Health
Office of Data Science Strategy

DATAWorks! FASEB
Advancing Data Sharing and Reuse

DataWorks! Prize

Goal: recognize and reward leaders in data sharing and reuse and create opportunities for broader research community to learn from their achievements

Submissions currently open – 5/11-7/19

www.herox.com/dataworks

Register by June 28, 2022

FASEB DATAWorks!
Advancing Data Sharing and Reuse

DATAWorks! FASEB

Salon

Since Initiative Launch – September 2021

Introduction to DataWorks!

What is data sharing?

Creating scientific opportunity through data reuse

9 Salons

Establishing a data sharing culture within a research team

"Challenging" Data Types

What's a DMP?

What are the FAIR and CARE principles?

How are funding organizations support data sharing and reuse

How to develop a Data Management and Sharing Plan

DATAWorks! FASEB
Advancing Data Sharing and Reuse

DATAWorks! FASEB

Community

DataWorks! Community will enable biological and biomedical researchers and teams to hone skills and mentor peers in data management and sharing.

Three-Month Cohort Program:

Month 1: Core course work in data management, community standards, disciplinary practices, and data science training
Month 2: Application of data curation skills
Month 3: Development of discipline-specific data sharing capstone project

Anticipated in Fall 2022

DATAWorks! FASEB

Help Desk

DataWorks! Help Desk will provide guidance for the biological and biomedical research community to navigate and adopt data sharing and reuse policies and practices.

Phase 1 Anticipated in Early 2023

DATAWorks! FASEB
Advancing Data Sharing and Reuse

Our Vision

Compliance

Culture Change

FASEB DATAWorks!
Advancing Data Sharing and Reuse

DataWorks! Partners

DataWorks! Advisory Committee

- **Parker Antin, PhD** (Chair) – University of Arizona
- **Maryann Martone, PhD** (Vice Chair) – University of California San Diego
- **Tim Clark, PhD** – University of Virginia
- **Kristi Holmes, PhD** – Feinberg School of Medicine, Northwestern University
- **Naim Matasci, PhD** – Ellison Institute for Transformative Medicine, University of Southern California
- **Ross Poldrack, PhD** – Stanford Data Science, Stanford University
- **Jason Williams** – DNA Learning Center, Cold Spring Harbor Laboratory

Ex Officio

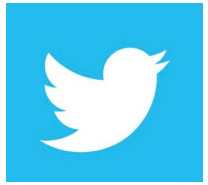
- **Patricia L. Morris, MS, PhD** – FASEB President
- **Cherié L. Butts, PhD** – FASEB Treasurer-Elect

FASEB DATAWorks!
Advancing Data Sharing and Reuse

Keep Up with DataWorks!

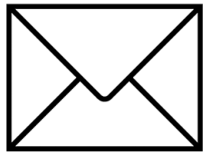
High-level Program Updates: www.faseb.org/dataworks

GitHub Repository: https://github.com/FASEB-DataWorks



@FASEBorg

#FASEBDataWorks



dataworks@faseb.org

FASEB DATAWorks!
Advancing Data Sharing and Reuse

FASEB DATAWorks!
Advancing Data Sharing and Reuse

Questions?

Contact us: dataworks@faseb.org

NIH Data Management and Sharing

What we have learned and may anticipate for the research administrator

Twila Fisher Reighley
Assoc. VP for Research
Sponsored Programs Administration
Michigan State University

June 9, 2022

Michigan State University: funding and practices

2021 NSF HERD Survey

- MSU $710M Total R&D Expenditures

Policies and practices related to research data:

- MSU Technology: MSU-institutional-data-policy-MSUT
- Faculty handbook:
  - In context of rights and responsibilities, consistent with standards and conformity with regulations, etc.: Faculty rights and responsibilities
  - HRPP “maintain security and storage:” Human research protection and data
  - In context of Misconduct: Misconduct procedures definitions
- Best practices including designating PI responsible for maintenance/retention of research data: RIO and research data

Approach to exploring experiences to date

Prep for policy implementation (and for COGR session):	Sent an email to colleagues at 22 universities asking about their experience with similar requirements.	Email response from six colleagues	Some had specific examples (more in next slide).
		Follow-up discussion with about 50% of institutions	Generally, not much administrative involvement with the requirements.
	Contacted MSU pre- and post-award staff:	The team dealing with close-outs had some examples.	
		From pre-award, an additional example at the JIT phase.	
	Discussed with two MSU faculty with large NIH awards:	Reminders, challenges, and advice (separate slides)	

What we have learned so far:

NIH:

- NIMH: It is already required; faculty/data designees are contacted and sometimes administrators are contacted too.
- NIAAA: asked for updated info (on required template) as part of JIT
- NCI and NIH Brain: experienced scientific data-related reminders, questions, and comments
- NIDA: AOR certification needed with 2-day turnaround related to requirements related to genomic data
- NIH (when institute not identified):
  - Projects ≥ $500K direct costs in a project year require data sharing
  - See requests at RPPR or final reporting

NASA:

- Data Use Agreement completed separately from the Notice of Grant Award and not referenced in the NOGA.
  - Postaward office was not aware of DUA but required to follow-up when tasks were not completed.
    - *Through discovery technology transfer office knew about the DUA*.

What we have learned from faculty feedback:

Expect significant impact

- Data management and sharing can require extensive first-time and ongoing submission hours

Plan for a data expert/data scientist

- Budget for personnel time or service center support

Budgeting as direct costs is allowable (subject to an institution's direct vs. indirect allocability)

- Sometimes peer reviewers are not supportive

Various NIH caps have not been increased, which is challenging

- Modular DC $250K, approval to submit if DC>$500K.

Faculty may be expected to input one project's data to different topical NIH databases.

Learned from faculty feedback *(continued)*:

Expect to reconcile data and add metadata and footnotes for context:

- More coding was necessary for imaging data and videotaped interactions.
- There was not a structure for nested data
  - Data tracked by individuals, twins, family, community, etc.
- Even adult ages were expected to be provided in months
  - Programming may help to convert the data effectively.
- NIH labeling implied different timeframes for follow-up visits.
- Challenging to get NIH client support in reconciling data
  - NIH sent frequent reminders to investigator and designated data contact for data delivery.
- Still reconciling items after project end date.
- NIH expected more on cooperative agreements and large grants.

What we have learned so far *(continued)*:

Expect the change will impact pre- and post-award offices:

- From a comment NIH Michelle Bulls made at FDP, we do anticipate that grants management may get more involved in compliance and that comment is supported specifically by:
  - NIH Policy NOT-OD-21-013 Section VIII. Compliance and Enforcement:
    - *Extramural Awards: The Plan will become a Term and Condition of the Notice of Award. Failure to comply with the Terms and Conditions may result in an enforcement action, including additional special terms and conditions or termination of the award, and may affect future funding decisions.*
- At some institutions, the efforts may be led by the library, but at others, it may need pre- or post-award leadership.

Preliminary Approach at MSU

Cross campus workgroup – MSU's representation:

- VPR areas: regulatory support, human research protections, preaward and postaward, and cyber-enabled research
- Academic areas: research dean, biomedical research informatics core director, and a faculty member
- Other units: IT research cyberinfrastructure, libraries, university counsel

Charge:

- Take steps to ensure MSU is prepared to comply with NIH requirements.

Preliminary MSU work through sub-workgroups:

- Costing
- Human subjects, privacy
- Data management, archiving, and sharing

Planning for a faculty survey:

- Colleague shared a sample initiated through his institution's library.

Resources

Also learning from info others have shared:

- Reviewing NIH info: NIH Data Sharing
- Reviewing COGR info: COGR NIH Data Mgmt. and Sharing Policy Resource
- Reviewing other universities' sites on data; for instance:
  - Cornell University
  - Duke University
  - Stanford University
  - University of Arizona
  - University of Michigan
  - University of Pennsylvania

Questions?

Thanks!

Twila Reighley
reighley@msu.edu