

NIH Data Management and Sharing (DMS) Policy

<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>



NIH DMS Policy: Goals

Goal: To make the results and outputs of NIH-funded research available to the public through effective and efficient data management and data sharing practices.

Why? Data sharing enables researchers to:

- Rigorously test the validity of research findings,
- Strengthen analyses through combined datasets,
- Reuse hard-to-generate data,

Good data management practices are foundational to effective data sharing.

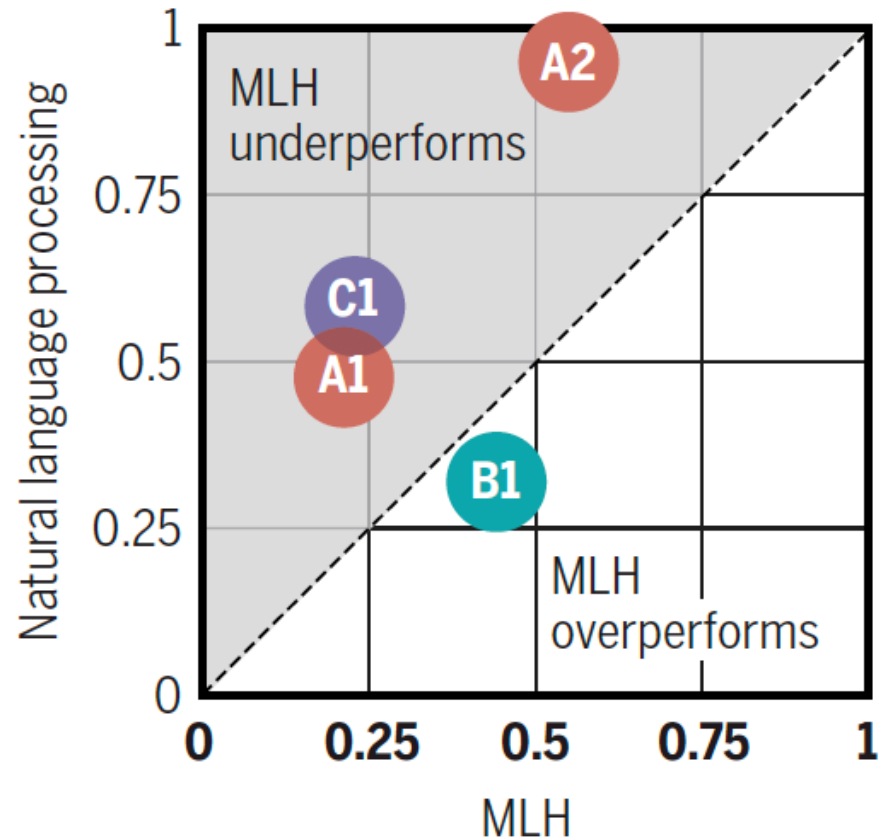
<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>



Example of Why This Matters: Reproducibility in Machine Learning for Health (MLH)

Evaluation metrics

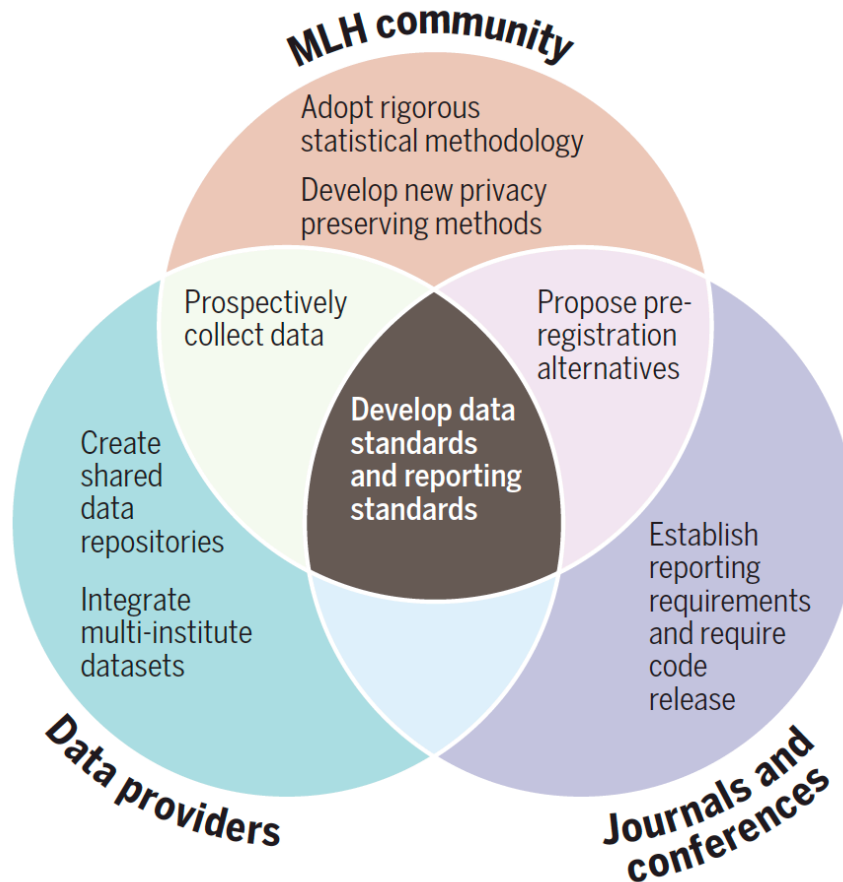
- A** **Technical reproducibility**
 - 1 Code available
 - 2 Public dataset
- B** **Statistical reproducibility**
 - 1 Variance reported
- C** **Conceptual reproducibility (replicability)**
 - 1 Multiple datasets



McDermott, MBA et al. *Sci. Trans. Med.* 13, eabb1655 (2021); <https://www.science.org/doi/10.1126/scitranslmed.abb1655>



Improving Reproducibility in MLH



McDermott, MBA et al. *Sci. Trans. Med.* 13, eabb1655 (2021); <https://www.science.org/doi/10.1126/scitranslmed.abb1655>



NIH DMS Policy: Scope

The DMS Policy takes effect 1/25/2023:

- **Applies to ALL new NIH research grants**, but not to training/fellowship, infrastructure & other non-research awards

Requires:

- 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.*
- Compliance with the plan.

<https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies>



NIH DMS Policy: Details

- Shared scientific data made accessible ASAP – no later than the time of an associated publication or end of performance period, whichever comes first.
- DMS plans will be reviewed by NIH program staff (not peers).
- Peer reviewers will only consider a proposal if data management budget is acceptable.
- *Costs for DMS can be included in the budget (discussed later in this presentation)*



NIH DMS Policy: Definition of Data

Definition of Scientific Data: The 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.

<https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies>





DMS Plan Components

Two-page or less description of:

- Which scientific data you will preserve or share
 - Data type, amount, rationale, metadata
- How you will accomplish that:
 - Tools, software, code
 - Data standards
 - Repository for scientific data and metadata
 - Persistent Unique Identifiers
 - Timelines
- Who will be responsible for managing compliance with the plan.



DMS Budget

- May include reasonable costs related to data management and sharing in budget proposals as **direct cost**.
 - **Best to avoid modular budgets if you can**
 - Costs must be incurred during performance period
- NIH has specified the following as allowable:
 - Curating data and developing supporting documentation
 - Local data management considerations
 - Preserving and sharing data through established repositories



DMS Budget

Budget requests may NOT include:

- Infrastructure costs included to institutional overhead
- Costs associated w/ routine conduct of research (e.g. collecting or accessing research data)
- Costs that are double-charged or inconsistently charged as both direct and indirects



Data Repositories

- FOA will either indicate repository
- OR
- If FOA doesn't specify, use discipline or data-type specific repository
- OR
- If FOA doesn't specify and there is no domain or data-type specific repository:
 - Small datasets (up to 2GB) may be included as supplementary material for articles submitted to PubMed Central
- OR
- Use a **generalist / institutional repository** that makes data available to larger research community / broader public



Compliance

Failure to comply could result in additional terms or conditions on the grant or even termination of a grant – and could impact an investigator's ability to obtain future NIH funding.



Why?

Open sharing of research data is essential not only to ensuring ongoing public support for the work we do (and that the public pays for), but also to ensuring that our work as researchers is of the **greatest benefit possible to the people it is intended to serve.**



Next Steps...

- Encourage researchers to register for an [ORCID](#)
- Familiarize yourself w/ the [DMS policy](#) and [DMS plan template](#)
- Review [allowable costs](#) for data management and sharing budgeting
- Encourage your investigators to adopt [LabArchives](#) (rollout forthcoming)
- Encourage researchers to utilize [GitLab](#) for code



More to Come...

- Institutional Data Repository
- Data deidentification
- Town Hall webinars
- Trainings
- Monthly updates to research community



University Resources

- [River Campus Libraries](#) and [Miner Library](#)
 - Writing a DMS plan
 - Organizing and storing data
 - Assigning metadata
 - Accessing data
- [Office of the Vice President for Research](#)
 - General information on DMS policy, communications from VPR, and institutional resources
- [CIRC](#)
 - Assistance with code repositories

