Getting ahead of the curve

Responding to emerging data management plan requirements

Rick Gilmore, Brianna Wham, & Rocco Zinoble

2021-09-30 07:43:25

Agenda

- Introductions
- What is a data management plan?
- When do I need one?
- How do I create one?
- Beyond checking the box…
- Discussion

Invited experts

- Rocco Zinoble, College of the Liberal Arts
- Briana Wham, Penn State Libraries

What is a data management plan?

A plan and a promise

…Thus, if an application describes a data-sharing plan, NIH expects that plan to be enacted. If progress has been made with the data-sharing plan, then the grantee should note this in the progress report. In the final progress report, if not sooner, the grantee should note what steps have been taken with respect to the data-sharing plan.

(NIH, n.d.)

In the case of noncompliance (depending on its severity and duration) NIH can take various actions to protect the Federal Government's interests. In some instances, for example, NIH may make data sharing an explicit term and condition of subsequent awards.

(NIH, n.d.)

Plan components

- What
- Where
- How
- Who

Examples

- Play & Learning Across a Year (PLAY) Project (R01HD094830). Resource Sharing and Data Sharing Plan.
- EAGER: Expanding public access to restricted research data (NSF 2032713). Data Management Plan.

When do I need one?

DMPs at NSF

- Required now
- 1-2 pp document
- (NSF, n.d.-a)
- (NSF, n.d.-b)

Current NIH policy

Starting with the October 1, 2003 receipt date, investigators submitting an NIH application seeking $500,000 or more in direct costs in any single year are expected to include a plan for data sharing or state why data sharing is not possible.

New NIH policy

- Effective January 25, 2023
- (NIH, 2021)

The final DMS Policy does not create a uniform requirement to share all scientific data. Unlike a requirement for submission of Plans, which can be implemented across various funding mechanisms and types of research with little variation, appropriate data sharing is likely to be varied and contextual. Through the requirement to submit a Plan, researchers are prospectively planning for data sharing, which we anticipate will increasingly lead researchers to integrate data sharing into the routine conduct of research.

Accordingly, we have included in the final DMS Policy an expectation that researchers will maximize appropriate data sharing when developing Plans. The final DMS Policy retains the Draft Policy's factors (i.e., ethical, legal, or technical) that may necessitate variations in the extent of scientific data preservation and sharing, and researchers should convey such factors in their Plans.

We believe this will provide the necessary flexibility for researchers to accommodate the substantial variety in research fields, projects, and data types that this expectation will encompass.

Resources

- NIH Data Sharing Policy and Implementation Guidance
- National Academy of Sciences and Engineering, "Changing the Culture of Data Management and Sharing: A Workshop," in March 2021

Elements

- 1-2 pp
- Data type(s)
- Related Tools, Software and/or Code
- Standards
- Data Preservation, Access, and Associated Timelines
- Access, Distribution, or Reuse Considerations
- Oversight of Data Management and Sharing

How do I create one?

DMPTool.org

- Data Management Toolkit

FAQs

What is/is not 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.

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

When to consider?

We encourage investigators to consider, while developing their Plans, how to address data management and sharing in the informed consent process, such that prospective participants will understand what is expected to happen with their data

"[s]hared 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."

Where to share?

The final DMS Policy strongly encourages the use of established repositories to the extent possible. This reflects NIH's preference that scientific data be shared and preserved through repositories, rather than kept only by the researcher or institution and provided on request, with the recognition that this is not always a practical or even a preferred approach.

How to pay for?

Personnel costs required to perform the types of data management and sharing activities described in the final Supplemental Information are allowable. Regarding the availability of data beyond the end of the project, which is crucial to achieving the goals of the DMS Policy, the final Supplemental Information clarifies that fees for long-term data preservation and sharing are allowable…

but funds for these activities must be spent during the performance period, even for scientific data and metadata preserved and shared beyond the award period. NIH funds cannot legally be spent after the award period.

Beyond checking the box…

Motivation

Sharing scientific data accelerates biomedical research discovery, in part, by enabling validation of research results, providing accessibility to high-value datasets, and promoting data reuse for future research studies.

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

Ethical & practical considerations

- Protecting participants AND accelerating discovery
- Data and materials sharing benefits science
- Most data can be shared in some form, with appropriate protections and restrictions
    - If you plan ahead for it
- (Brakewood & Poldrack, 2013); (Cychosz et al., 2020); Gilmore, Xu, & Adolph, 2021; (Gilmore, Cole, Verma, Aken, & Worthman, 2020); (Gilmore & Qian, 2021); (Gollwitzer et al., 2020); (Meyer, 2018)

Where to share

- Institutional repositories
  - PSU ScholarSphere
- Data repositories/archives/enclaves
  - Databrary
  - ICSPR
  - Open Science Framework (OSF)
  - Dataverse
  - NIH Data Archive

Plan your work, work your plan

- Embedding sharing and quality assurance (QA) into ongoing workflows
- Make sharing with collaborators ~ sharing with others
- (Soska et al., 2021)

Your thoughts?

Materials

Code

This talk was produced on 2021-09-30 in RStudio using R Markdown. The code and materials used to generate the slides may be found at https://github.com/psu-psychology/2021-10-01-odds-data-mgmt.

Information about the R Session that produced the code is as follows:

```
sessionInfo()
```

```
## R version 4.1.0 (2021-05-18)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Big Sur 11.6
##
## Matrix products: default
## LAPACK: /Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods
## [7] base
##
## loaded via a namespace (and not attached):
##  [1] digest_0.6.27     R6_2.5.0          jsonlite_1.7.2
##  [4] magrittr_2.0.1    evaluate_0.14     rlang_0.4.11
##  [7] stringi_1.7.3     jquerylib_0.1.4   bslib_0.2.5.1
## [10] rmarkdown_2.9     tools_4.1.0       stringr_1.4.0
```

References

Brakewood, B., & Poldrack, R. A. (2013). The ethics of secondary data analysis: Considering the application of belmont principles to the sharing of neuroimaging data. *NeuroImage*, *82*, 671–676. https://doi.org/10.1016/j.neuroimage.2013.02.040

Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., … Weisleder, A. (2020). Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01365-9

Gilmore, R. O., Cole, P. M., Verma, S., Aken, M. A. G., & Worthman, C. M. (2020). Advancing scientific integrity, transparency, and openness in child development research: Challenges and possible solutions. *Child Development Perspectives*, *14*(1), 9–14. https://doi.org/10.1111/cdep.12360

Gilmore, R. O., & Qian, Y. (2021). An open developmental science will be more rigorous, robust, and impactful. *Infant and Child Development*. https://doi.org/10.1002/icd.2254

Gollwitzer, M., Abele-Brehm, A., Fiebach, C., Ramthun, R., Scheel, A. M., Schönbrodt, F. D., & Steinberg, U. (2020, September). *Data management and data sharing in psychological science: Revision of the DGPs recommendations*. https://doi.org/10.31234/osf.io/24ncs

Meyer, M. N. (2018). Practical tips for ethical data sharing. *Advances in Methods and Practices in Psychological Science*, *1*(1), 131–144. https://doi.org/10.1177/2515245917747656

NIH. (n.d.). NIH data sharing policy and implementation guidance. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm. Retrieved from https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

NIH. (2021). NOT-OD-21-013: Final NIH policy for data management and sharing. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html. Retrieved from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html