**Subject:** Update on NIH DMS Plan Activity

Dear colleagues,

Penn is joining the [Dryad Data Platform](https://datadryad.org/stash) as an institutional member to provide easy access for our faculty to a generalist data repository in anticipation of needs based on the [2023 National Institutes of Health (NIH) Data Management and Sharing Policy](https://sharing.nih.gov/data-management-and-sharing-policy). The Office of the Vice Provost for Research (OVPR), working with Penn Libraries and the Office of Research Services (ORS), has evaluated available generalist data repositories and selected to move forward with Dryad. We ask that you disseminate this information to your faculty, and we will have a brief discussion about this topic at next week’s Provost’s Council for Research meeting.

**Background -** As you are aware, NIH has issued a Data Management and Sharing (DMS) policy, **effective January 25, 2023**, to promote the sharing of scientific data. Under the DMS policy, NIH expects that investigators and institutions will:

* Plan and budget for the managing and sharing of data
* Submit a DMS plan for review when applying for funding
* Comply with the approved DMS plan

**What Scientific Data Need to Be Shared -** Scientific data are the recorded factual material commonly accepted in the scientific community as sufficient quality to validate and replicate research findings, regardless of whether the data 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. There are justifiable reasons for limiting data sharing, which should be described in the DMS plan.

Data plans for research subject to the NIH Genomic Data Sharing policy will now be included as part of the DMS plans. Genomic data sharing considerations will be addressed in the DMS plan, such as where and when genomic data will be shared.

At a minimum, scientific data supporting a publication must be shared by the time of publication (when the publication first appears, either online or in print). Other scientific data must be shared by the end of the research project or protocol. The Office of Intermural Research (OIR) encourages sharing high-quality scientific data not included in a publication, including “negative results.”

**Selecting a Data Repository**

Text

Description automatically generated

While NIH encourages using [domain-specific repositories](https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html) where possible, such repositories are not available for all datasets. When investigators cannot locate a repository for their discipline or the type of data they generate, a [generalist repository](https://www.nlm.nih.gov/NIHbmic/generalist_repositories.html) can be a useful place to share data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus.

OVPR evaluated three generalist data repositories, Dryad Data Platform, FigShare, and Dataverse. All three generalist data repository options provide the minimum functionality Penn expects. With that said, after considering each generalist data repository, OVPR has selected to join the Dryad Data Platform as an institutional (large, research-intensive) member.

Dryad Data Platform is a curated resource that makes research data discoverable, freely reusable, and citable and provides a general-purpose home for a wide diversity of data types. There are no data publishing deposits for faculty once Penn is an institutional member. Dryad is open source, and researchers’ software code can be publicly available on GitHub. The platform is based on an underlying Ruby-on-Rails data publication platform called Stash, encompassing three main functional components: Store, Harvest, and Share. Contents are free to download and re-use under a Creative Commons Zero (CC0) license. Contents are preserved for the long term to guarantee access to contents indefinitely.

There is a limit of 300GB per data publication uploaded through the Dryad web interface. Dryad can accept larger submissions, but the submitter needs to contact Dryad for assistance. They recommend that individual files not exceed 10GB to ensure they are easily accessed and downloadable by users.

Dryad is a non-profit, shared community working together to promote data publishing, curation, and preservation, providing significant flexibility. Dryad does not accept submissions that contain personally identifiable human subject information. Human subjects’ data must be appropriately anonymized and prepared under applicable legal and ethical guidelines.

Best,

Mike

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| A picture containing text, clipart  Description automatically generated | **Michael Borda, Ph.D. (He/Him/His)**  Associate Vice Provost for Research  Executive Director of Strategic Initiatives      **Phone** 215-898-0394  **Mobile** 302-983-6042  **Web** [http://research.upenn.edu](http://research.upenn.edu/)  **E-mail** [mjborda@upenn.edu](mailto:mjborda@upenn.edu)  1 College Hall, Suite 118  Philadelphia, PA 19104 |