

MAGLAB NMR-MRI FACILITY USER DATA MANAGEMENT PLAN

Contact: Robert W. Schurko (rschurko@fsu.edu)

Funder: National Science Foundation Division of Chemistry and Division of Materials Research

Last modified: 2023-01-11

ABSTRACT:

Ensuring that publicly-funded research data is preserved and freely available safeguards efficient use of government resources and facilitates efficient delivery of scientific discoveries to maximize impact. The National Science Foundation (NSF) supports FAIR (Findability, Accessibility, Interoperability, and Reuse) data guiding principles [1], and considers data management planning as integral to any NSF-funded research. Therefore, products of research generated at MagLab User Facilities should be made available to the scientific community and general public. Specifically, this policy requires that all research products generated at the MagLab NMR-MRI User Facility be digitally accessible upon publication, or within 3 years. This data management plan (DMP) details resources available to facility users, and outlines procedures for managing data and the products of research in alignment with FAIR principles.

[1] <http://www.go-fair.org/fair-principles>

PRODUCTS OF THE RESEARCH

Raw data are single one-, two-, or three-dimensional (1D, 2D, 3D) NMR/MRI spectra or collections of NMR/MRI spectra. All data files contain the raw, digitized FID (and most often, an optimally processed spectrum or image), as well as numerous metadata including acquisition parameters, processing parameters, title, temperature, spinning rate, sample preparation, other protocols, etc. The data products, metadata, and file sizes vary depending upon the samples, experiment type, and users. Both internal and external users are responsible for providing metadata describing the sample names, mode(s) of preparation, experimental conditions, sample concentrations, etc. NMR-MRI Instrument Managers assist the users with the collection, storage, and organization of all data and metadata originating from work conducted at the MagLab.

DATA FORMAT

The vast majority of NMR/MRI data are acquired on Bruker spectrometers, and the data are stored as Bruker TopSpin files, which are accessible not only with Bruker software platforms, but numerous other commercial and home-built (e.g., within MatLab) software packages. If required, Topspin can also be used to convert the data to the open standard JCAMP-DX format [2], which is used for uploading to online repositories and can be accessed with the open source software OpenVnmrJ [3]. Software is available for all internal and external users for examining, analyzing, and creating presentations of this data (e.g., TopSpin, ssNake, and MatLab). Software packages provide ample "title" space for experimental metadata outside of the acquisition and processing parameters (sample name, temperature, preparation details, concentrations, chemical shift references, etc.).

[2] <http://www.jcamp-dx.org/>

[3] <https://openvnmrj.org/>

DATA SHARING AND ACCESS

RESPONSIBILITIES OF THE PRINCIPAL INVESTIGATOR

The Principal Investigator (PI) is the steward of the research data, will select the vehicle(s) for research products (publications or presentations), and will have ultimate authority in their initial use. The PI must notify the NHMFL upon successful dissemination involving data acquired at MagLab.



Research activities detailed in NMR-MRI user proposals and approved for magnet time are expected to result in presentations, publications, or other vehicles for dissemination of data and results. Details of experimental work and metadata (e.g., description of samples, sample preparations, experimental conditions, chemical shift references, etc.) should be included with published data. Journal articles should include digital object identifiers (DOIs) to indicate where relevant data and metadata can be accessed. Users are welcomed and encouraged to work in collaboration with NMR-MRI facility personnel to verify data or results before use in forums such as publications, meeting presentations, grant or patent applications.

It is the responsibility of the PI to ensure protection of privacy, confidentiality, intellectual property, national security, or other rights or requirements, and is encouraged to disclose these to NMR-MRI facility staff listed as collaborators to the extent necessary to facilitate compliance.

For research involving animals or human subjects, the PI's host institutional animal care and use committee (IACUC) or institutional review board (IRB) approval must be secured before a project is granted magnet time. Additionally, the PI must comply with all public access requirements that are laid out by other funding agencies sponsoring the research, in addition to NMR-MRI facility data management policies (see below).

DATA SHARING PRACTICES

Prior to publication, project data and metadata will be shared with registered MagLab users listed as project collaborators. Requests from other interested parties will be considered on a case-by-case basis, and subject to approval of the PI. The PI initial use authority does not control sharing data with NMR-MRI facility staff to gauge instrument performance, perform instrument calibration, or meet reporting requirements for the facility.

The MagLab is exploring the Open Science Framework (OSF) [4] to serve as a project management and data sharing platform between NMR-MRI facility personnel and external users. Users and staff are encouraged to use the OSF for data transfer, access, and storage, but it is not required. The PI can send/receive and share materials and data through their virtual delivery mechanisms of choice in consultation with NMR-MRI facility staff.

[4] <https://osf.io/>

DATA ACCESS POLICY

This policy applies only to data and metadata collected at the NMR-MRI User Facility under the NMR-MRI Users' Program. To balance the need to make data openly available to the community with user expectations that they will be able to publish results of their scientific efforts without fear of preemption, data are to be made available to the public as outlined, below:

1. All data and metadata associated with a user project will become publicly available when an associated manuscript is published, or within 3 years of the date the project was last assigned magnet time. Extension of the embargo period can be requested by the PI on a case-by-case basis. Exceptions will generally be granted if a publication or patent application is under review. Repository, entry, DOI, and other relevant accession information must be reported to the Director of the NMR-MRI facility at the time of publication, or conclusion of the data embargo period.
2. As exceptions to the requirements above, some data/metadata are not required to be made publicly available. These are data that will not form the basis of publishable research findings nor are associated with a user project. These include data from experiments known to be faulty in some regard (e.g., mishaps or flawed experimental design, data from preliminary experiments that are not intended to be delivered to NMR-MRI facility users, standards/calibration runs for which results are not needed to interpret legitimate project data, and data generated to verify successful operation of the instrument or demonstrate capability). Users should consult NMR-MRI Facility staff regarding the type of data collected and its suitability for public consumption.



DATA REPOSITORIES

FAIR guidelines [1] stipulate that data and associated metadata should be submitted to a discipline-specific, community-recognized, public repository. The project PI should select an appropriate repository, or choose from a recommended repository listed in the table, below:

Data-type, Field, or Funding Agency	Repository	Link to homepage
Generalist	Open Science Framework	http://osf.io/
Generalist	Zenodo	https://zenodo.org/
Protein structures (PDB)	Protein Data Bank	https://www.rcsb.org/
Generalist	Mendeley Data	https://data.mendeley.com/

If no suitable, community-recognized resource is available, data and associated metadata should be submitted to a recognized generalist repository. The journal, *Scientific Data* (Springer), recommends several generalist repositories [5]. Among them, the MagLab recommends the Open Science Framework, a free and open platform for research project management and a reliable data repository. The OSF supports the ability to embargo data and metadata in accordance with the policies outlined above. While embargoed, all submitted materials or datasets are given their own unique, persistent URLs. DOIs can be generated when projects or selected components are made public. These may be cited and accessed by the public, and are indexed in Google Scholar. The OSF is a flexible alternative to some field-specific repositories to efficiently, and wholly disseminate all data and metadata related to complex, large-scale projects spanning multiple disciplines.

[5] <https://www.nature.com/sdata/policies/repositories#general>

POLICIES FOR RE-USE, RE-DISTRIBUTION, AND PRODUCTION OF DERIVATIVES

Authors of any publications or presentations that utilize NMR-MRI facility data, results, software, or other resources are encouraged to cite relevant literature, include relevant DOIs, or otherwise acknowledge the researchers who generated the samples, data, results, software, or other materials.

In addition, all published manuscripts, datasets, and presentations must acknowledge the MagLab NMR-MRI Facility, and facility support (including NSF grant number) as outlined below:

“A portion of this work was performed at the NMR-MRI User Facility at the National High Magnetic Field Laboratory at Florida State University, which is funded by the National Science Foundation Cooperative Agreement (DMR-2128556) and by the State of Florida.”

For data collected from 2018-2022, the grant number is DMR-1644779. For data collected from 2023-2027, the grant number is DMR-2128556. Please include both grant numbers if your data were collected during both time periods.

ARCHIVAL OF DATA

The NMR-MRI facility provides data storage, backup, and archiving services for all user facility data. All raw data collection is performed on NMR-MRI Facility hardware and instruments and raw data are archived on MagLab storage infrastructure (the Z drive). Data archived on the Z drive are subjected to regular backups to magnetic tape to ensure data durability. Archival and backup servers are managed by the MagLab Computer Support Group. All data that has ever been archived has been retained and this will continue indefinitely. NMR-MRI facility users receive copies of raw data upon request as outlined in the Data Sharing Practices section, above. Users that collect data on-site can receive copies on physical media. Remote users are provided data through various virtual delivery mechanisms chosen by the PI, private repository submission (to be made public at the appropriate time), or through physical media delivered via parcel post.



DATA MANAGEMENT MAP

