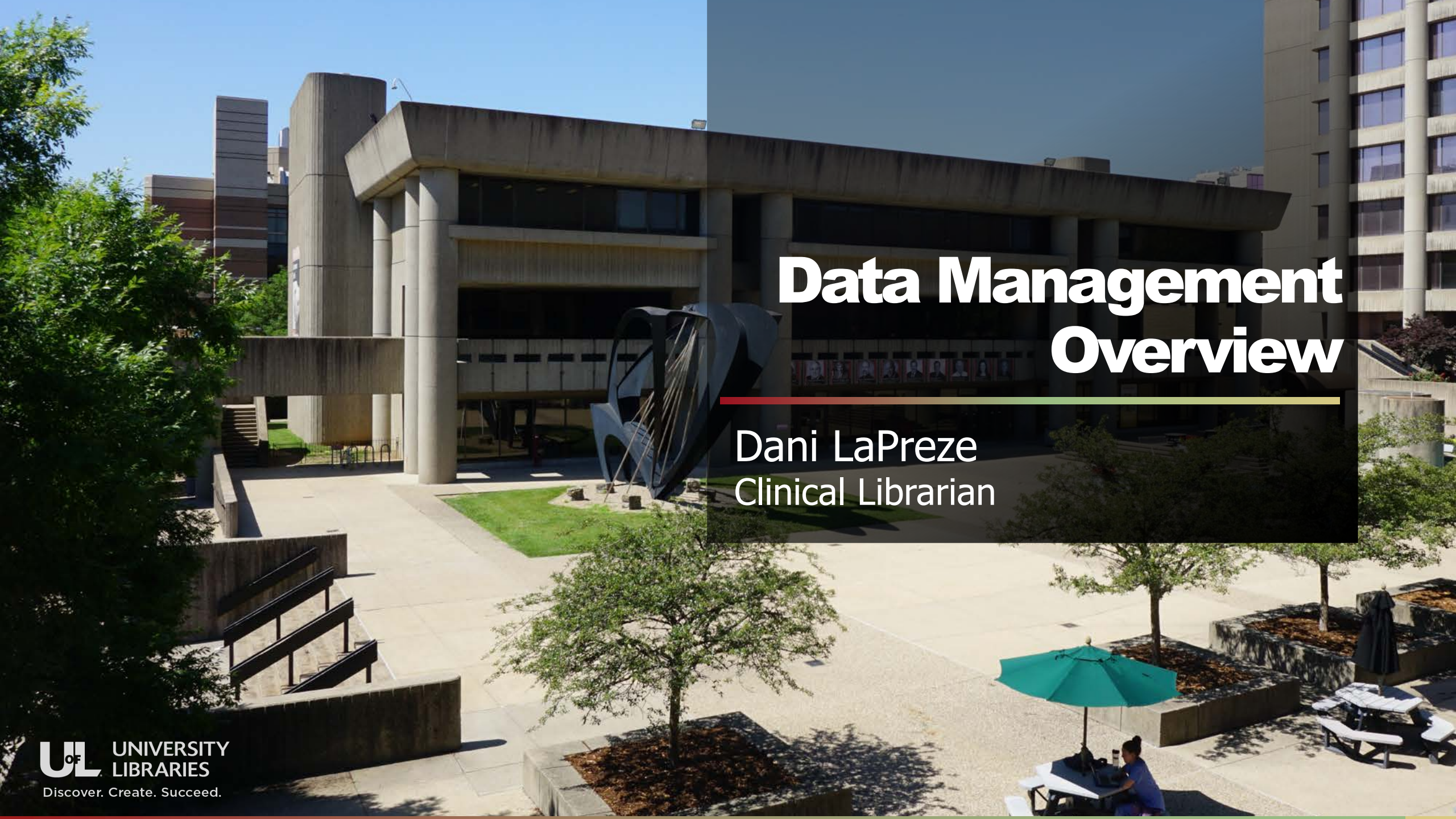
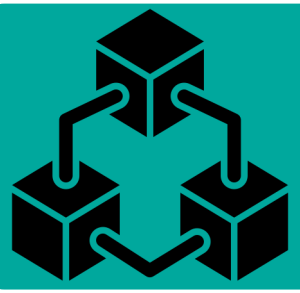


# Data Management Overview

Dani LaPreze
Clinical Librarian

UNIVERSITY LIBRARIES
Discover. Create. Succeed.

# Agenda



## Research Data Management & FAIR Data

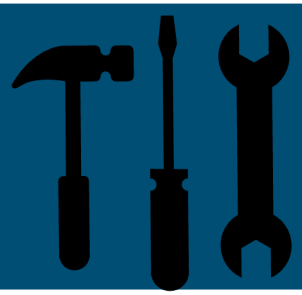
- Overview of RDM
- What is FAIR Data and why should you care?

## Policy Overview

- UofL
- NIH

## Data Management Plans

- NIH Requirements
- Allowable Costs
- Selecting a Repository



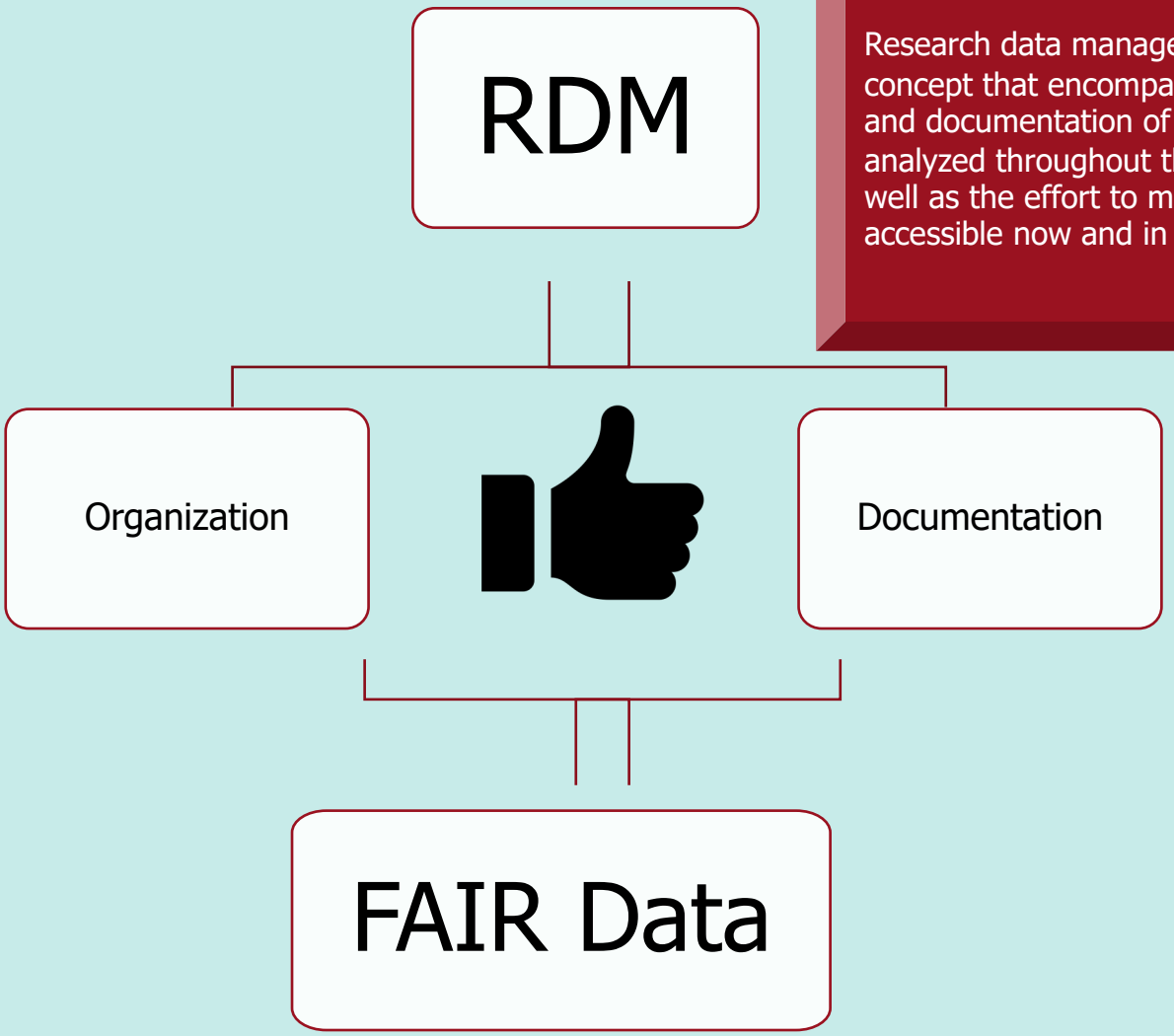
## Resources

- Templates
- Resources & Toolkits

# What RDM is

RDM is not simply putting data on a thumb drive, uploading it to a cloud server, or any other temporary means of preservation or storage. These are first steps, but true RDM requires more.

Research data management (RDM) is a broad concept that encompasses the organization and documentation of data collected and analyzed throughout the research process, as well as the effort to make that data easily accessible now and in the future.




# What RDM is NOT

# What is FAIR Data?

2016 -- **The FAIR Guiding Principles for scientific data management and stewardship** is published

> *"There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals."*

Since its publication, FAIR data principles have become best practices when sorting out what to do with data at the conclusion of a research project.

# What is FAIR Data?

## Findable

- (meta)data are assigned a globally unique and persistent identifier
- data are described with rich metadata
- metadata clearly and explicitly include the identifier of the data it describes
- (meta)data are registered or indexed in a searchable resource

## Accessible

- (meta)data are retrievable by their identifier using a standardized communications protocol
- the protocol is open, free, and universally implementable
- the protocol allows for an authentication and authorization procedure, where necessary
- metadata are accessible, even when the data are no longer available

## Interoperable

- (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- (meta)data use vocabularies that follow FAIR principles
- (meta)data include qualified references to other (meta)data

## Reusable

- meta(data) are richly described with a plurality of accurate and relevant attributes
- (meta)data are released with a clear and accessible data usage license
- (meta)data are associated with detailed provenance
- (meta)data meet domain-relevant community standards

The basic principle is that data should be easy to find, access and apply to future research, regardless of the circumstances. Promotion of open access and reusability are key with these principles.

For example, if you access a data set, but the program it was saved on no longer works, then the data is not interoperable. If the data set does not have correct licensing, it is not reusable. If there is a paywall, the data is not accessible.

FAIR principles are best practices, and though not universally mandatory, many funders and publishers require some form of open access regarding data.

# Management & Sharing of Research Data Policy

Highlights

- Applies to all research AND researchers at UofL
- Applies regardless of funding or source of support
- Principal Investigators of research projects, are responsible for the collection, management and retention of research data.
- Follow discipline-specific best practices for sharing & managing data
- Adhere to FAIR data principles
- Research data must be archived for a minimum of five years after the conclusion of a study (for UofL; may be longer per FOA or journal requirements)
- Ownership of data is governed by UofL's Policy on Intellectual Property

# Data Management & Sharing (DMS) Policy

Highlights

**Section I. Purpose**

- Establish expectations for DMSP's
- Encourage FAIR data practices
- Sets requirement for prospective planning

**Section II. Definitions**

- 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. 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.

- Data Management: The process of validating, organizing, protecting, maintaining, and processing scientific data to ensure the accessibility, reliability, and quality of the scientific data for its users.

- Metadata: Data that provide additional information intended to make scientific data interpretable and reusable (e.g., date, independent sample and variable construction and description, methodology, data provenance, data transformations, any intermediate or descriptive observational variables).

**Section III. Scope**

- Applies to all research funded or conducted in whole or in part by NIH *that generates scientific data*

**Section IV. Effective Date**

- January 25, 2023

# Data Management & Sharing (DMS) Policy

Continued

**Section V. Requirements**
- Submit a plan
- Comply with approved plan

**Section VI. DMSP's**
- Generating scientific data = needing a DMSP
- Revisions allowed (during RPPR)

**Section VII. Managing & Sharing**
- Researchers are expected to maximize the appropriate sharing of scientific data, acknowledging legal, ethical, technical factors
- If you are using human subjects should outline in their plan how this data will be protected (privacy, participant rights, confidentiality, etc.)
- Data repository selection
- Preservation & sharing timelines

**Section VIII. Compliance & Enforcement**
- Plan + Award = Term & Condition
- Compliance determined by the NIH ICO (institute, center, or office)

# Supplemental Information

## Related Announcements

NOT-OD-23-012 - Reminder: FORMS-H Grant Application Forms & Instructions Must be Used for Due Dates On or After January 25, 2023 - New Grant Application Instructions Now Available

NOT-CA-23-007 - Request for Information (RFI): Soliciting Input on the Use and Reuse of Cancer Metabolomics Data

NOT-OD-22-214 - Supplemental Information to the NIH Policy for Data Management and Sharing: Responsible Management and Sharing of American Indian/Alaska Native Participant Data.

NOT-OD-21-013 - Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data.

NOT-OD-22-189 - Implementation Details for the NIH Data Management and Sharing Policy

NOT-OD-22-104 – Notice of Extension of the Public Comment Period for NOT-OD-22-064 DRAFT Supplemental Information to the NIH Policy for Data Management and Sharing: Responsible Management and Sharing of American Indian/ Alaska Native Participant Data

NOT-OD-22-064 - Request for Public Comments on DRAFT Supplemental Information to the NIH Policy for Data Management and Sharing: Responsible Management and Sharing of American Indian/ Alaska Native Participant Data

NOT-OD-22-029 - Request for Information on Proposed Updates and Long-Term Considerations for the NIH Genomic Data Sharing Policy

NOT-HG-21-023 - Notice Announcing NHGRI Guidance for Third-Party Involvement in Extramural Research

NOT-HG-21-022 - Notice Announcing the National Human Genome Research Institute's Expectation for Sharing Quality Metadata and Phenotypic Data

NOT-OD-21-014 – Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan

NOT-OD-21-015 – Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing

NOT-OD-21-016 – Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research

NOT-OD-20-013 - Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance

NOT-MH-21-265 - Notice of Biospecimen Sharing Policy for the National Institute of Mental Health, Including Requirements for Induced Pluripotent Stem Cell Resource Development and Sharing

NOT-OD-22-131 –Request for Public Comments on DRAFT Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data

NOT-OD-22-195 – New NIH "FORMS-H" Grant Application Forms and Instructions Coming for Due Dates on or after January 25, 2023

NOT-OD-22-198 – Implementation Changes for Genomic Data Sharing Plans Included with Applications Due on or after January 25, 2023

# What should a DMSP include?

## Data Type

- Will you produce new data, or will you be acquiring data from another source?
- General summary of types & estimated amount of data to be generated
- Description of data to be preserved & shared
- Brief listing of metadata & associated documentation that will be made accessible

## Tools, Software, Code

- Are specialized tools or programs necessary to access or use data?
- How can specialized tools be accessed?
- Will specialized tools remain available long-term?

## Standards

- Always use filed-established standards
- Data formats, data dictionaries, data identifiers, & other data documentation
- If no establish standards, this can be mentioned in the plan & one (such as Dublin Core) can be adopted

## Preservation, Access, & Timelines

- Name of repository(ies) where data will be archived
- How will the data be findable and identifiable?
- When will data be made available to other users and for how long?

## Access, Distribution, & Reuse

- Describe factors relating to access, distribution, & reuse
- Informed consent, privacy, or confidentiality protections
- Will human-derived data access be controlled?
- Are there any restrictions imposed by laws, regulations, policies, or existing/anticipated agreements?

## Data Management & Sharing Oversight

- Who will ensure compliance with the DMSP?

Specifically laid out in the NIH policy are 6 elements that must be included in a DMSP. Other funders may have other requirements, but these 6 are pretty universal with regards to data management and preservation. But double check your researchers needs.
These 6 elements can be used as headers in the document, which makes for easier writing and reading.

# Allowable Costs

Reasonable, allowable costs may be included in NIH budget requests when associated with:

**1.Curating data and developing supporting documentation,** including formatting data according to accepted community standards; de-identifying data; preparing metadata to foster discoverability, interpretation, and reuse; and formatting data for transmission to and storage at a selected repository for long-term preservation and access.

**2.Local data management considerations,** such as unique and specialized information infrastructure necessary to provide local management and preservation (e.g., before deposit into an established repository).

**3.Preserving and sharing data through established repositories,** such as data deposit fees necessary for making data available and accessible. For example, if a Data Management and Sharing Plan proposes preserving and sharing scientific data for 10 years in an established repository with a deposition fee, the cost for the entire 10-year period must be paid prior to the end of the period of performance. If the Plan proposes deposition to multiple repositories, costs associated with each proposed repository may be included.

# Repositories

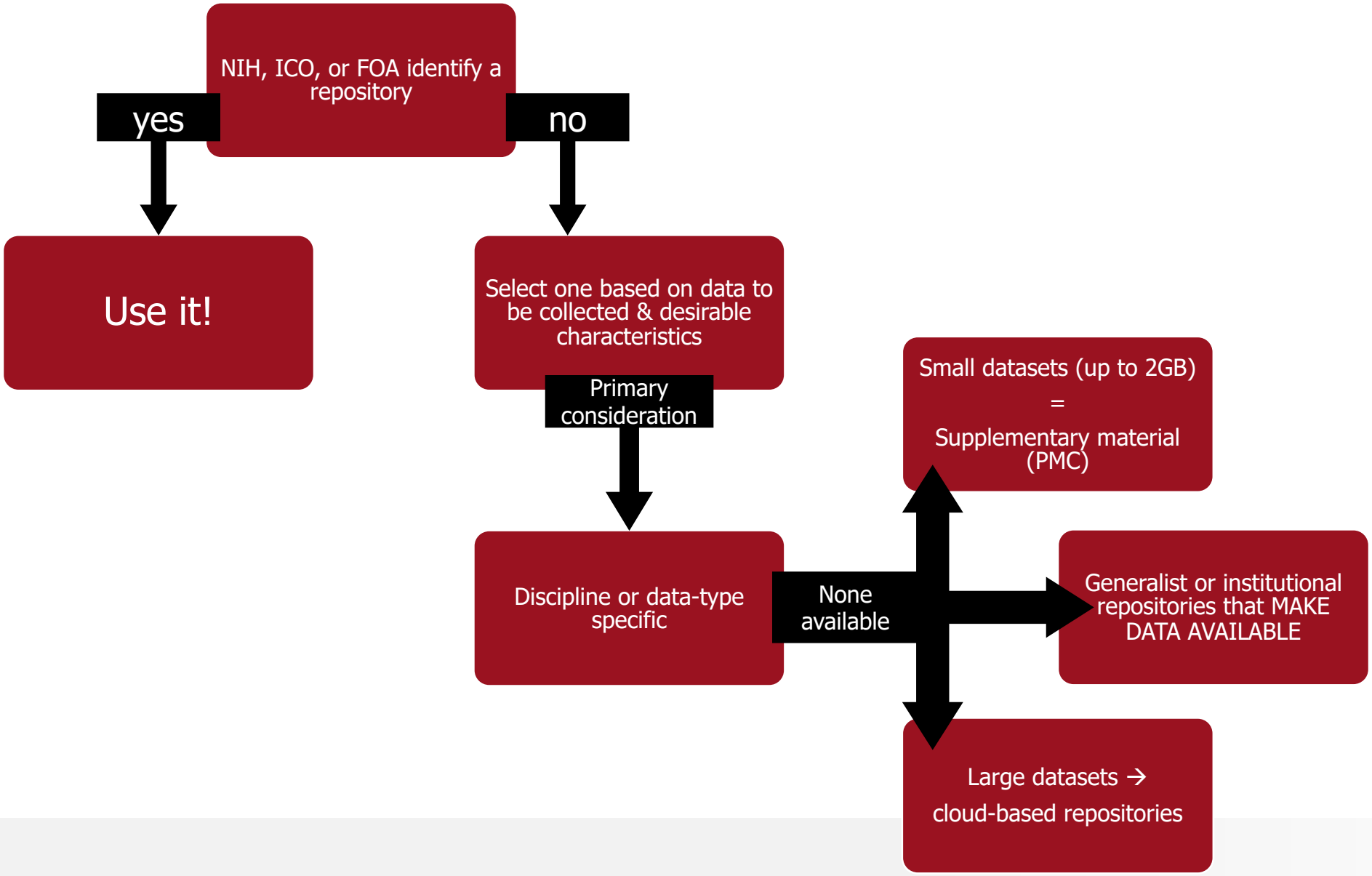
NIH promotes the use of established data repositories because deposit in a quality data repository generally improves the FAIRness of the data.

NIH will not necessarily provide data repositories – check Funding Opportunity Announcement (FOA)

If no repository is provided in the FOA, NIH has an "if not this, then that" guide.

Regardless of which repository is selected, NIH encourages researchers to select data repositories that exemplify desired characteristics and and ensure that data are managed and shared in ways that are consistent with the FAIR data principles.

# Selecting a Repository Flowchart



# Selecting a Repository (direct from the policy)

1.For some programs and types of data, NIH and/or Institute, Center, Office (ICO) policy(ies) and Funding Opportunity Announcements (FOAs) identify particular data repositories (or sets of repositories) to be used to preserve and share data. For data generated from research subject to such policies or funded under such FOAs, researchers should use the designated data repository(ies).
2.For data generated from research for which no data repository is specified by NIH or the NIH ICO (as described above), researchers are encouraged to select a data repository that is appropriate for the data generated from the research project and is in accordance with the desired characteristics, taking into consideration the following guidance:
1.Primary consideration should be given to data repositories that are discipline or data-type specific to support effective data discovery and reuse. NIH makes a list of such data repositories available
(see https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html).
2.If no appropriate discipline or data-type specific repository is available, researchers should consider a variety of other potentially suitable data sharing options:

1. Small datasets (up to 2 GB in size) may be included as supplementary material to accompany articles submitted to PubMed Central (see https://www.ncbi.nlm.nih.gov/pmc/about/guidelines/#suppm).
2. Data repositories, including generalist repositories (see https://www.nlm.nih.gov/NIHbmic/generalist_repositories.html) or institutional repositories, that make data available to the larger research community, institutions, or the broader public.
3. Large datasets may benefit from cloud-based data repositories for data access, preservation, and sharing.

# Desirable Characteristics for Data Repositories

**Unique Persistent Identifiers:** Assigns datasets a citable, unique persistent identifier (PID), such as a digital object identifier (DOI) or accession number, to support data discovery, reporting (e.g., of research progress), and research assessment (e.g., identifying the outputs of federally funded research). The unique PID points to a persistent landing page that remains accessible even if the dataset is de-accessioned or no longer available.

**Long-Term Sustainability:** Has a plan for long-term management of data, including maintaining integrity, authenticity, and availability of datasets; building on a stable technical infrastructure and funding plans; and having contingency plans to ensure data are available and maintained during and after unforeseen events.

**Metadata:** Ensures datasets are accompanied by metadata to enable discovery, reuse, and citation of datasets, using schema that are appropriate to, and ideally widely used across, the community(ies) the repository serves. Domain-specific repositories would generally have more detailed metadata than generalist repositories.

**Curation and Quality Assurance:** Provides, or has a mechanism for others to provide, expert curation and quality assurance to improve the accuracy and integrity of datasets and metadata

**Free and Easy Access:** Provides broad, equitable, and maximally open access to datasets and their metadata free of charge in a timely manner after submission, consistent with legal and ethical limits required to maintain privacy and confidentiality, Tribal sovereignty, and protection of other sensitive data.

**Broad and Measured Reuse:** Makes datasets and their metadata available with broadest possible terms of reuse; and provides the ability to measure attribution, citation, and reuse of data (i.e., through assignment of adequate metadata and unique PIDs).

# Desirable Characteristics for Data Repositories

**Clear Use Guidance:** Provides accompanying documentation describing terms of dataset access and use (e.g., particular licenses, need for approval by a data use committee).

**Security and Integrity:** Has documented measures in place to meet generally accepted criteria for preventing unauthorized access to, modification of, or release of data, with levels of security that are appropriate to the sensitivity of data.

**Confidentiality:** Has documented capabilities for ensuring that administrative, technical, and physical safeguards are employed to comply with applicable confidentiality, risk management, and continuous monitoring requirements for sensitive data.

**Common Format:** Allows datasets and metadata downloaded, accessed, or exported from the repository to be in widely used, preferably non-proprietary, formats consistent with those used in the community(ies) the repository serves.

**Provenance:** Has mechanisms in place to record the origin, chain of custody, and any modifications to submitted datasets and metadata.

**Retention Policy:** Provides documentation on policies for data retention within the repository.

# Additional Considerations When Storing Human Data

**Fidelity to Consent:** Employs documented procedures to restrict dataset access and use to those that are consistent with participant consent (such as for use only within the context of research on a specific disease or condition) and changes in consent.

**Restricted Use Compliant:** Employs documented procedures to communicate and enforce data use restrictions, such as preventing reidentification or redistribution to unauthorized users.

**Privacy:** Implements and provides documentation of appropriate approaches (e.g., tiered access, credentialing of data users, security safeguards against potential breaches) to protect human subjects' data from inappropriate access.

**Plan for Breach:** Has security measures that include a response plan for detected data breaches.

**Download Control:** Controls and audits access to and download of datasets (if download is permitted).

**Violations:** Has procedures for addressing violations of terms-of-use by users and data mismanagement by the repository.

**Request for Review Process:** Makes use of an established and transparent process for reviewing data access requests.

# NIH Template

OMB No. 0925-0001 and 0925-0002 *(Rev. 07/2022 Approved Through TBD)*

**PREVIEW – DRAFT**

## DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted. There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the format shown below.

**Element 1: Data Type**

**A. Types and amount of scientific data expected to be generated in the project:**
*Summarize the types and estimated amount of scientific data expected to be generated in the project.*

**B. Scientific data that will be preserved and shared, and the rationale for doing so:**
*Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.*

**C. Metadata, other relevant data, and associated documentation:**
*Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.*

**Element 2: Related Tools, Software and/or Code:**
*State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.*

**Element 3: Standards:**
*State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.*

**Element 4: Data Preservation, Access, and Associated Timelines**

**A. Repository where scientific data and metadata will be archived:**
*Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived; see Selecting a Data Repository).*

**B. How scientific data will be findable and identifiable:**
*Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.*

**C. When and how long the scientific data will be made available:**
*Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.*

OMB No. 0925-0001 and 0925-0002 *(Rev. 07/2022 Approved Through TBD)*

**PREVIEW – DRAFT**

**Element 5: Access, Distribution, or Reuse Considerations**

**A. Factors affecting subsequent access, distribution, or reuse of scientific data:**
*NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See Frequently Asked Questions for examples of justifiable reasons for limiting sharing of data.*

**B. Whether access to scientific data will be controlled:**
*State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).*

**C. Protections for privacy, rights, and confidentiality of human research participants:**
*If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).*

**Element 6: Oversight of Data Management and Sharing:**
*Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).*

**DATA MANAGEMENT AND SHARING PLAN**

An example from an application proposing to collect clinical and MRI data from human subjects.

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted (but this has not been done in the sample below). There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the *format* shown below.

**Element 1: Data Type**

**A. Types and amount of scientific data expected to be generated in the project:**

*Summarize the types and estimated amount of scientific data expected to be generated in the project.*

Demographic, clinical, and MRI, $^{1}$H fMRS and fMRI imaging data will be acquired from 110 affected youth and 110 matched healthy controls (described in detail in sections C.3 and C.4 of this application). All data will be de-identified prior to receipt by the repository, but the information needed to generate a global unique identifier for the NIMH Data Archive (NDA) will be collected for each subject.

**B. Scientific data that will be preserved and shared, and the rationale for doing so:**

*Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.*

Sufficient data from this project will be preserved to enable sharing via NDA data of sufficient quality to validate and replicate research findings described in the Aims. NIMH requires data measured from human subjects to be shared using the NDA.

**C. Metadata, other relevant data, and associated documentation:**

*Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.*

In addition to the subject level data described above, all $^{1}$H fMRS and fMRI task related paradigm designs and experiment definitions will be deposited in the NDA.

**Element 2: Related Tools, Software and/or Code:**

*State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.*

The clinical data will be analyzed with custom Python code written using the statsmodels, numpy, and pandas packages, all of which are freely available. $^{1}$H fMRS spectra will be analyzed with LCModel 6.3 software using LCMgui, which is freely available from http://s-provencher.com/lcm-test.shtml. fMRI images will be analyzed using the SPM8 toolbox (https://www.fil.ion.ucl.ac.uk/spm/software/spm8/) for MATLAB (http://www.mathworks.com/products/matlab/). While MATLAB is commercial software, most universities have site licenses available and the SPM8 toolbox is free. It is also possible that the toolbox might run in Octave, an open-source alternative to MATLAB (https://www.gnu.org/software/octave/index), but we have not tried it. All code will be shared on our lab website at https://github.com/labname the main readme.md file for the project will also include instructions and parameter choices for the GUI-based analyses.

**Element 3: Standards:**

*State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.*

Participant age, sex, ethnicity, height, weight, socioeconomic status, and other demographic data will be collected using the following instruments as defined in NDA:

1) Research Subject and Pedigree (ndar_subject01)
2) Demographics Short Form (demsf01)
3) Ethnic Group Questionnaire (ethgrp01)
4) Height and Weight (height_weight01)
5) Hollingshead Socioeconomic Rating Scale (ses01)
6) Pubertal Development Scale (pds01)
7) Edinburgh Handedness Inventory (edinburgh_hand01)
8) WASI-2 (wasi201).

In compliance with NOT-MH-20-067, the following data will be collected to facilitate aggregation of this data set with other data sets:

1) DSM Crosscutting for Youth (dsm5crossch01)
2) RCADS-25 (rcads2501)

The clinical assessments we plan to collect for this study include:

1) Kiddie-SADS-Present and Lifetime Version (ksads_pl01)
2) Children's Yale-Brown Obsessive Compulsive Scale (cybocs01)
3) Schedule for Obsessive-Compulsive and Other Behavioral Syndromes (Hanna. Schedule for Obsessive-Compulsive and Other Behavioral Syndromes, Ann Arbor: University of Michigan, 2010, new data dictionary will be defined in NDA)
4) Dimensional Obsessive Compulsive Scale (docs01)
5) Yale Global Tic Severity Scale (yale01)
6) Child Behavior Checklist (cbcl01)
7) Multidimensional Anxiety Scale for Child Parent and Self (masc_p01)
8) Conners 3 (conners3_ps01)
9) Adolescent Depression Rating Scale (doi:10.1186/1471-244X-7-2, new data dictionary will be defined in NDA)

$^{1}$H fMRS and fMRI data will be shared with the Image (image03), Imaging Work Flow (iwf01), and Imaging Collection (imagingcollection01) data dictionaries as defined in NDA.

**Element 4: Data Preservation, Access, and Associated Timelines**

**A. Repository where scientific data and metadata will be archived:**

*Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived; see Selecting a Data Repository).*

All data will be deposited to NDA starting 12 months after the award begins and will be deposited every six months thereafter following the usual NDA data submission dates.

**B. How scientific data will be findable and identifiable:**

*Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.*

Data will be findable for the research community through the NDA Collection that will be established when this application is funded. For all publications, an NDA study will be created. Each of those studies is assigned a digital object identifier (DOI). This data DOI will be referenced in the publication to allow the research community easy access to the exact data used in the publication.

**C. When and how long the scientific data will be made available:**

*Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.*

The research community will have access to data when the award ends. As required by NDA, studies will also be created that contain the data used for every publication. Those studies will be shared when the pre-print is available. NDA studies have digital object identifiers (DOI) to aid in findability. We will include that DOI in relevant publications. NDA will make decisions about how long to preserve the data, but that data archive has not deleted any deposited data up to now.

**Element 5: Access, Distribution, or Reuse Considerations**

**A. Factors affecting subsequent access, distribution, or reuse of scientific data:**

*NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See Frequently Asked Questions for examples of justifiable reasons for limiting sharing of data.*

All research participants will be consented for broad data sharing.

**B. Whether access to scientific data will be controlled:**

*State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).*

To request access of the data, researchers will use the standard processes at NDA, and the NDA Data Access Committee will decide which requests to grant. The standard NDA data access process allows access for one year and is renewable.

**C. Protections for privacy, rights, and confidentiality of human research participants:**

*If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).*

The NDA GUID tool allows researchers to aggregate data from the same research participant without different laboratories having to share personally identifiable information about that research participant. The NDA data dictionaries do not permit personally identifiable information to be shared. NDA maintains a Certificate of Confidentiality.

**Element 6: Oversight of Data Management and Sharing:**

*Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).*

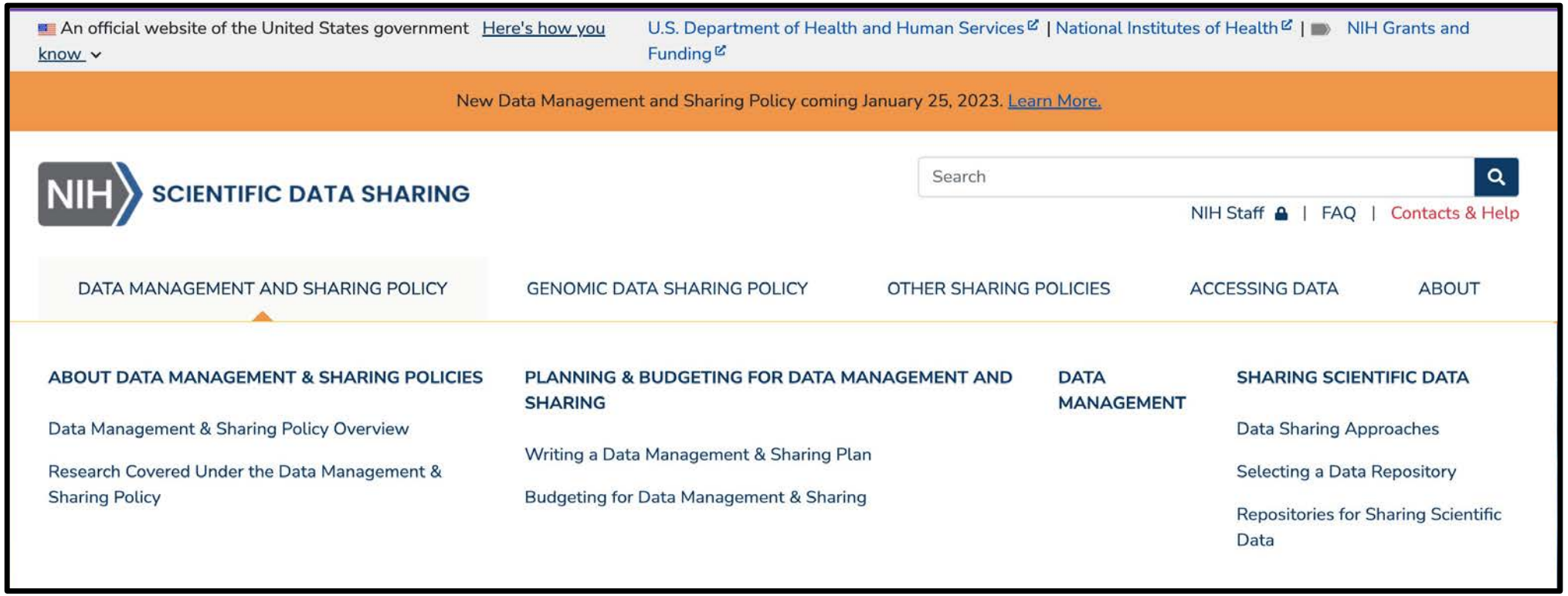
The Office of Sponsored Programs at University X that will be administering this award has created a data management and sharing plan compliance system as part of their process for submitting the annual NIH progress report. That Office is collecting information related to the number of research participants that are deposited each reporting year. The Office of Sponsored Programs will also look for the NDA data DOIs from NDA Studies and will include that information in the annual progress report.

**Validation Schedule (this section is required by NIMH)**

If funded, within 6 months of the Notice of Award date we will submit a Data Submission Agreement signed by the principal investigators and an institutional business official, as well as define and complete the Data Expected section of this project. Uploads of all initial demographic, clinical, and raw structural MRI, $^{1}$H fMRS and fMRI research data will be completed using the second submission cycle deadline following the Notice of Award date. Subsequent data uploads will be harmonized, validated, and submitted biannually on the standard January 15th and July 15th submission deadlines.

We also plan to use the NDA validation tool as a quality control measure in the laboratory. The data manager in charge of submitting data to NDA will help researchers in the group validate their data once every month.

# Resources




https://sharing.nih.gov/

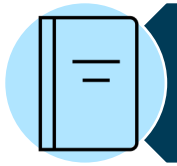
# Resources

DMPTool

Build your Data Management Plan

- Templates
- Funder requirements
- Public DMP's
- Intuitive use

Academic paper format



Use headers

Cover all necessary points

Be as descriptive as necessary

If a section is not answered, give reasoning

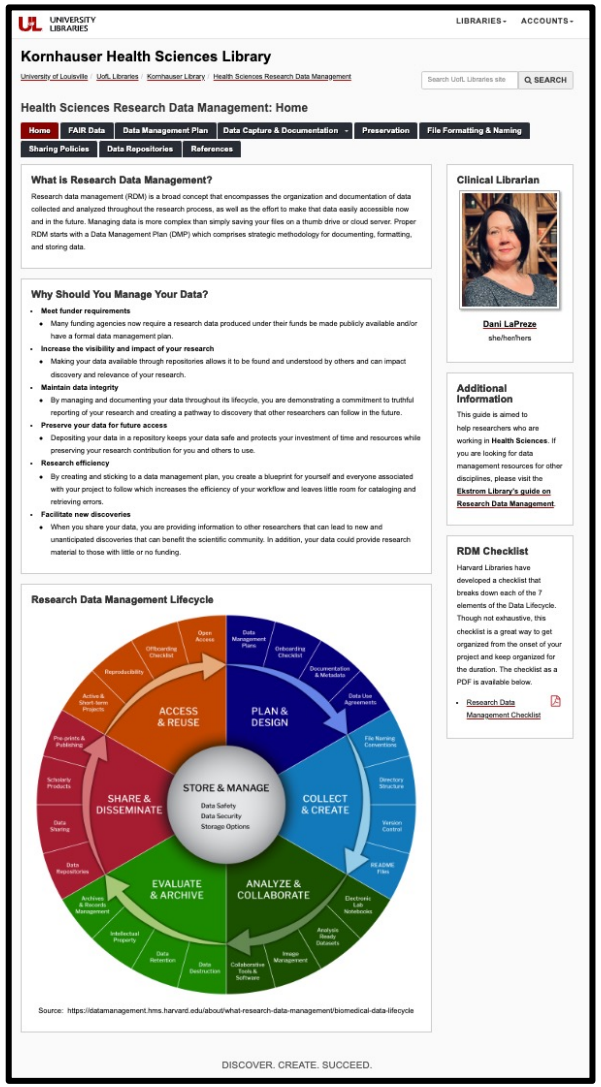
https://dmptool.org/

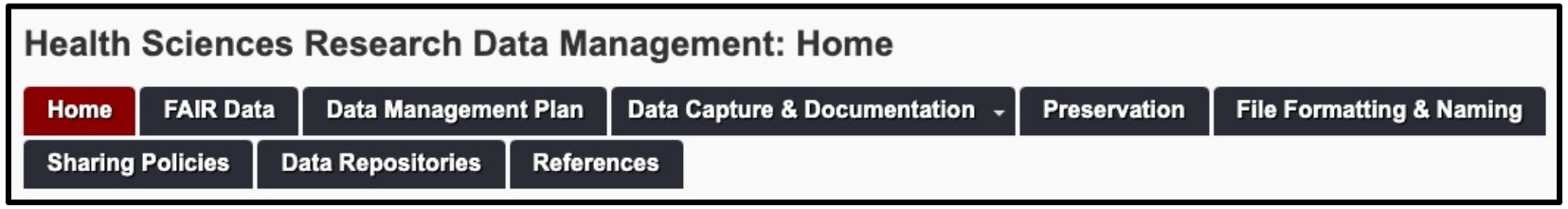
# Resources


Policy Readiness Checklist for Librarians
Data Management and Sharing Plan Checklist for Researchers
•DMSP Checklist
•*extended reference (coming soon!)*
Example DMS Plans
•Searchable Catalog of DMPs
•Annotated Example DMP in NIH 2023 Format
•Spreadsheet Version of DMP Catalog
•More Information
*Repository Finder (coming soon!)*
Glossaries
•Data Terms related to the NIH DMS Plan and Policy
•Grant Glossary for Librarians
•Grant Speak in Context

https://osf.io/uadxr/wiki/home/

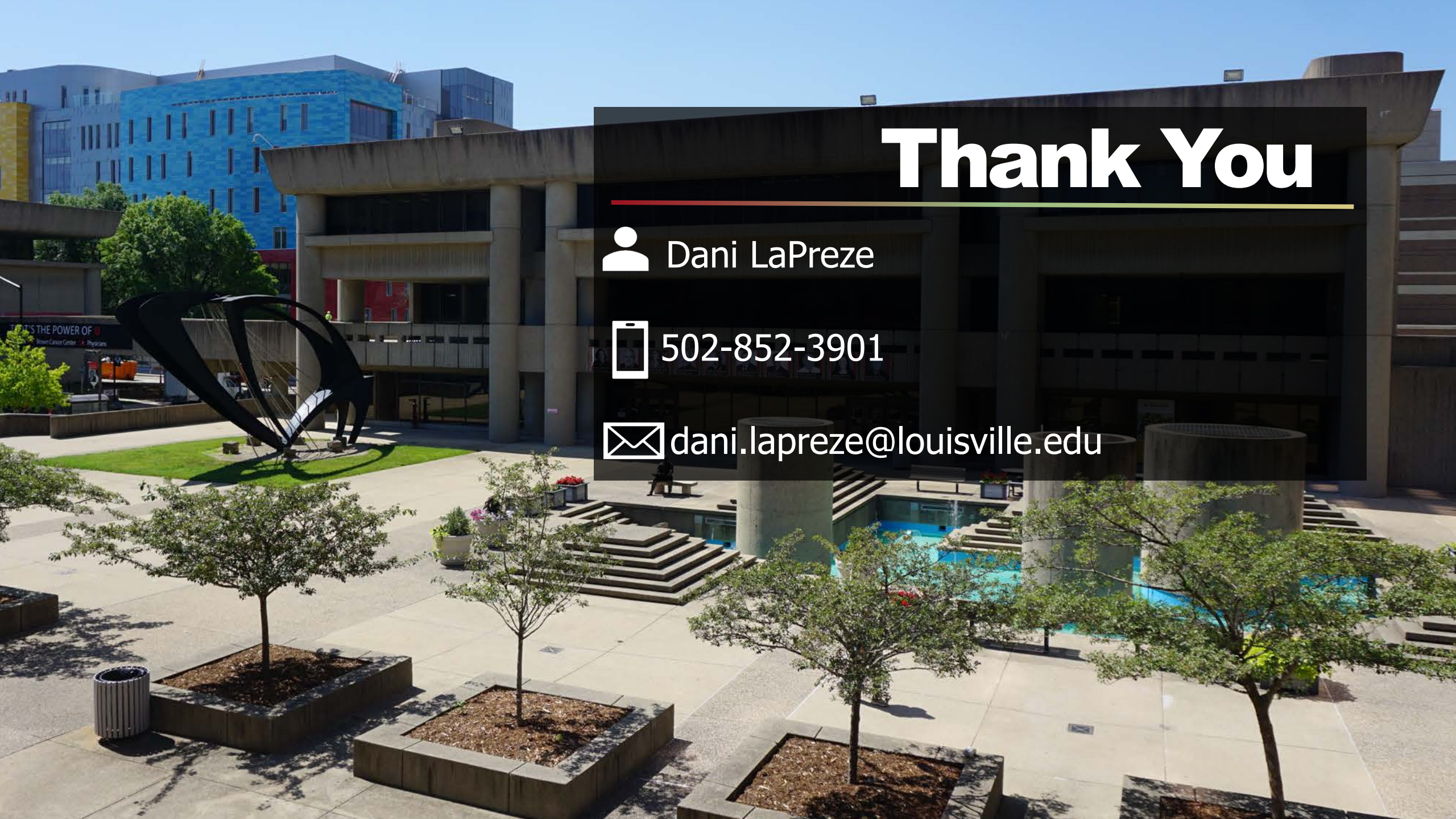
# Resources




- Quick access to some of the information covered earlier
- My contact information
- Good resources for students and novice to expert researchers

https://library.louisville.edu/kornhauser/datamanagement/home



# Thank You


Dani LaPreze

502-852-3901

dani.lapreze@louisville.edu