NIH DATA SHARING

OCTOBER 2022

USC University of Southern California

OVERVIEW OF NIH DATA SHARING POLICY

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. 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."

Core principle: Data shared must be of sufficient quality to validate and replicate research findings.

OVERVIEW OF NIH DATA SHARING POLICY

Effective Date: **January 25, 2023**

This includes:

- Competing grant applications that are submitted to NIH for January 25, 2023, and subsequent receipt dates;
- Proposals for contracts that are submitted to NIH on or after January 25, 2023;
- NIH Intramural Research Projects conducted on or after January 25, 2023; and
- Other funding agreements (e.g., Other Transactions) that are executed on or after January 25, 2023, unless otherwise stipulated by NIH.

OVERVIEW OF NIH DATA SHARING POLICY

01 NIH requires all applicants planning to generate scientific data to prepare a Data Management and Sharing (DMS) plan that describes how the scientific data will be managed and shared.

02 Researchers are expected to budget for any costs associated with satisfying data sharing obligations.

03 Researchers are expected to call out any ethical, legal, or technical considerations that might limit the nature and extent of data sharing.

04 Although uniformity is desirable, there remains some flexibility for NIH Institutes, Centers, and Offices (ICO's) to add additional measures.

05 Data should be made available as soon as possible and no later than the time of an associated publication. Data must also be released in a timely way even if no publications arise from the research.

OVERVIEW OF NIH DATA SHARING POLICY

SPECIAL CONSIDERATIONS FOR DATA RELATED TO HUMAN SUBJECTS

As it relates to data obtained from human subjects, the Data Sharing policy encourages:

Addressing data management and sharing in the informed consent process.

Communicating any limitations on subsequent use of data to those individuals or entities preserving and sharing the scientific data.

Researchers to consider whether access to scientific data derived from humans should be controlled, even if de-identified and lacking explicit limitations on subsequent use.

DATA SHARING –

PLAN PREPARATION AND SUBMISSION REQUIREMENTS

USC University of Southern California

DATA SHARING PLAN: ELEMENTS

Should be no longer than two pages and describe how the scientific data will be findable and identifiable. Other key elements:

01 Data Type

Briefly describe the scientific data to be managed, preserved, and shared, including a summary of the types and amounts of data to be shared, and a description of which scientific data will be preserved and shared.

02 Related Tools, Software, and/or Code

Any indication of whether specialized tools are needed to access or manipulate shared scientific data to support replication or reuse, and names of the needed tools/software.

03 Standards

An indication of what standards will be applied to the scientific data and associated metadata (i.e., data formats, data dictionaries, data identifiers, definitions, unique identifiers, and other data documentation).

04 Data Preservation, Access, and Associated Timelines

Plans and timelines for data preservation and access, including the name of the repository; how scientific data will be findable and identifiable; when the scientific data will be made available to other users and for how long.

05 Access, Distribution, or Reuse Considerations

Plans should maximize the appropriate sharing of scientific data consistent with privacy, security, informed consent, and proprietary issues.

06 Oversight of Data Management and Sharing

Indicate how compliance with the Plan will be monitored and managed, frequency of oversight, and by whom (e.g., titles, roles).

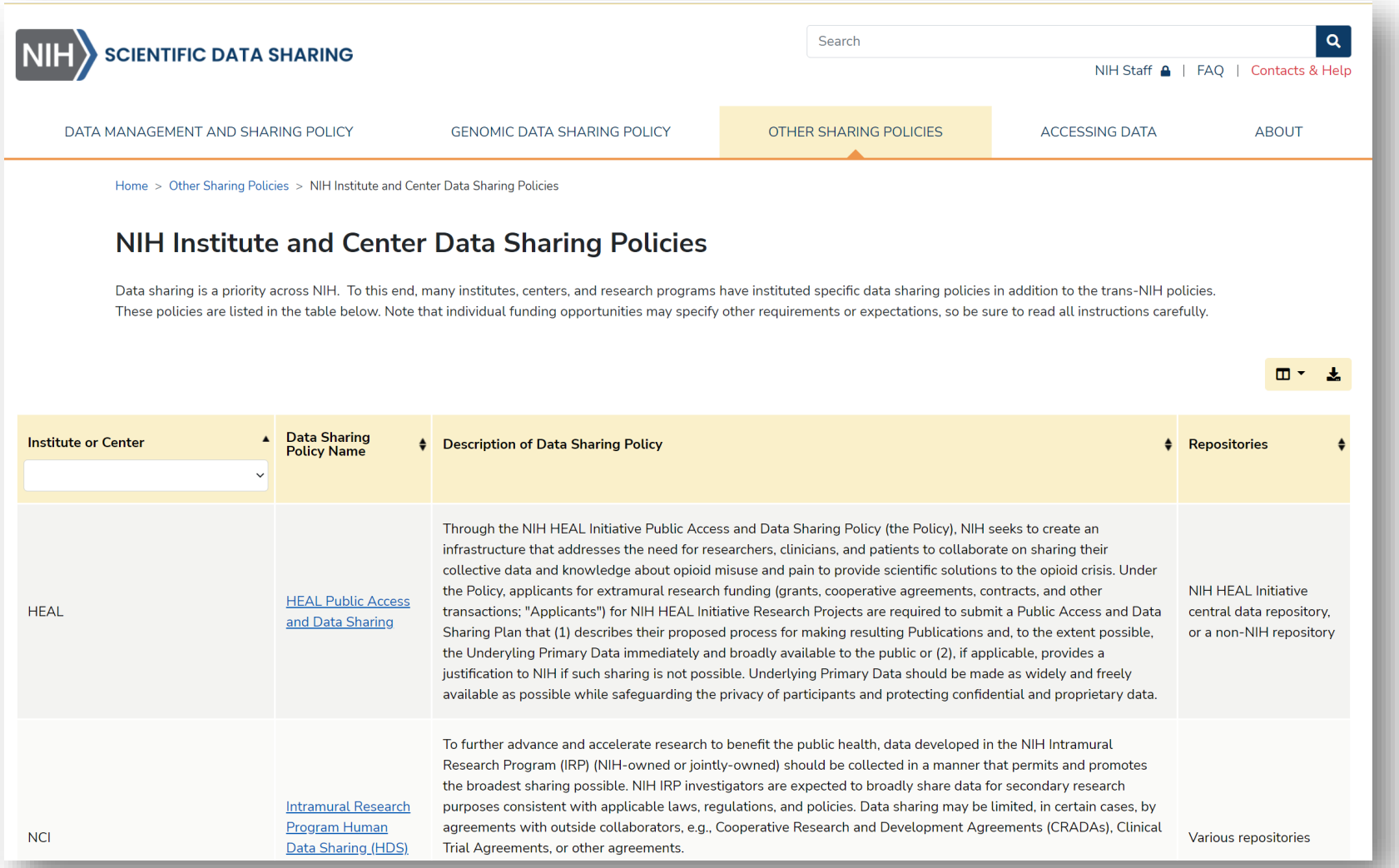
Please see Appendix One: Additional Guidance on Data Sharing Plans for additional detail.

DATA SHARING PLAN ELEMENTS: ADDITIONAL CONSIDERATIONS

07

ADDITIONAL CONSIDERATIONS

- Note that funding opportunities or ICs may have specific expectations (for example: scientific data to share, relevant standards, repository selection). View a list of NIH Institute or Center data sharing policies.



NIH SCIENTIFIC DATA SHARING

Search

NIH Staff | FAQ | Contacts & Help

DATA MANAGEMENT AND SHARING POLICY | GENOMIC DATA SHARING POLICY | OTHER SHARING POLICIES | ACCESSING DATA | ABOUT

Home > Other Sharing Policies > NIH Institute and Center Data Sharing Policies

NIH Institute and Center Data Sharing Policies

Data sharing is a priority across NIH. To this end, many institutes, centers, and research programs have instituted specific data sharing policies in addition to the trans-NIH policies. These policies are listed in the table below. Note that individual funding opportunities may specify other requirements or expectations, so be sure to read all instructions carefully.

Institute or Center	Data Sharing Policy Name	Description of Data Sharing Policy	Repositories
HEAL	HEAL Public Access and Data Sharing	Through the NIH HEAL Initiative Public Access and Data Sharing Policy (the Policy), NIH seeks to create an infrastructure that addresses the need for researchers, clinicians, and patients to collaborate on sharing their collective data and knowledge about opioid misuse and pain to provide scientific solutions to the opioid crisis. Under the Policy, applicants for extramural research funding (grants, cooperative agreements, contracts, and other transactions; "Applicants") for NIH HEAL Initiative Research Projects are required to submit a Public Access and Data Sharing Plan that (1) describes their proposed process for making resulting Publications and, to the extent possible, the Underlying Primary Data immediately and broadly available to the public or (2), if applicable, provides a justification to NIH if such sharing is not possible. Underlying Primary Data should be made as widely and freely available as possible while safeguarding the privacy of participants and protecting confidential and proprietary data.	NIH HEAL Initiative central data repository, or a non-NIH repository
NCI	Intramural Research Program Human Data Sharing (HDS)	To further advance and accelerate research to benefit the public health, data developed in the NIH Intramural Research Program (IRP) (NIH-owned or jointly-owned) should be collected in a manner that permits and promotes the broadest sharing possible. NIH IRP investigators are expected to broadly share data for secondary research purposes consistent with applicable laws, regulations, and policies. Data sharing may be limited, in certain cases, by agreements with outside collaborators, e.g., Cooperative Research and Development Agreements (CRADAs), Clinical Trial Agreements, or other agreements.	Various repositories

USC University of Southern California

DATA SHARING PLAN: DISCUSSION OF FACTORS THAT MAY LIMIT SHARING

Certain factors may limit data sharing and must be addressed if applicable. These include:

Informed consent will not permit or will limit the scope or extent of sharing and future research use

Existing consent (e.g., for previously collected biospecimens) prohibits sharing or limits the scope or extent of sharing and future research use

Privacy or safety of research participants would be compromised or place them at greater risk of re-identification or suffering harm

Explicit federal, state, local, or tribal law, regulation, or policy prohibits disclosure

Datasets cannot practically be digitized with reasonable efforts

DATA MANAGEMENT AND SHARING PLANS: HOW TO SUBMIT

The DMS plan should be submitted as follows:

EXTRAMURAL (GRANTS)
as part of the Budget Justification section of the application

INTRAMURAL
determined by the Intramural Research Program

EXTRAMURAL (CONTRACTS)
as part of the technical evaluation

OTHER FUNDING AGREEMENTS
prior to the release of funds

Although the plans are submitted before research begins, if any changes occur during the award or support period that affects how data is managed or shared, investigators should update the plan to reflect the changes. It may be helpful to discuss potential changes with the Program Officer. In addition, the funding institute or center will need to approve the updated plan.

USC University of Southern California

ADDITIONAL SUBMISSION DETAILS

The NIH will be making the following changes to forms to accommodate the new DMS plan requirements:

Data sharing plans and genomic data sharing plans will no longer be submitted to the “Resource Sharing Plan(s)” field.

A new “Other Plan(s)” field will be added to the PHS 398 Research Plan Form and PHS 398 Career Development Award Supplemental Form to allow for upload of the DMS plan.

An optional DMS Plan format page will be provided to assist applicants with the preparation of this attachment. Use of this format page is recommended, but DMS Plans generated using other approaches will be accepted.

The requested direct costs to support the activities proposed in the DMS Plan must be indicated as “Data Management and Sharing Costs” as follows:

- **R&R Budget Form: single line item in Section F. Other Direct Costs**
- **PHS 398 Modular Budget Form: as text embedded within the Additional Narrative Justification**

A brief summary of the DMS Plan and a description of the requested Data Management and Sharing Costs must be included within the budget justification attachment, as follows:

- R&R Budget Form: embedded within the Section L. Budget Justification attachment
- PHS 398 Modular Budget Form: embedded within the Additional Narrative Justification

These changes will be implemented with application form packages identified with a Competition ID of "FORMS-H" and associated application guide instructions. (Coming Fall 2022)

SELECTION OF A DATA REPOSITORY

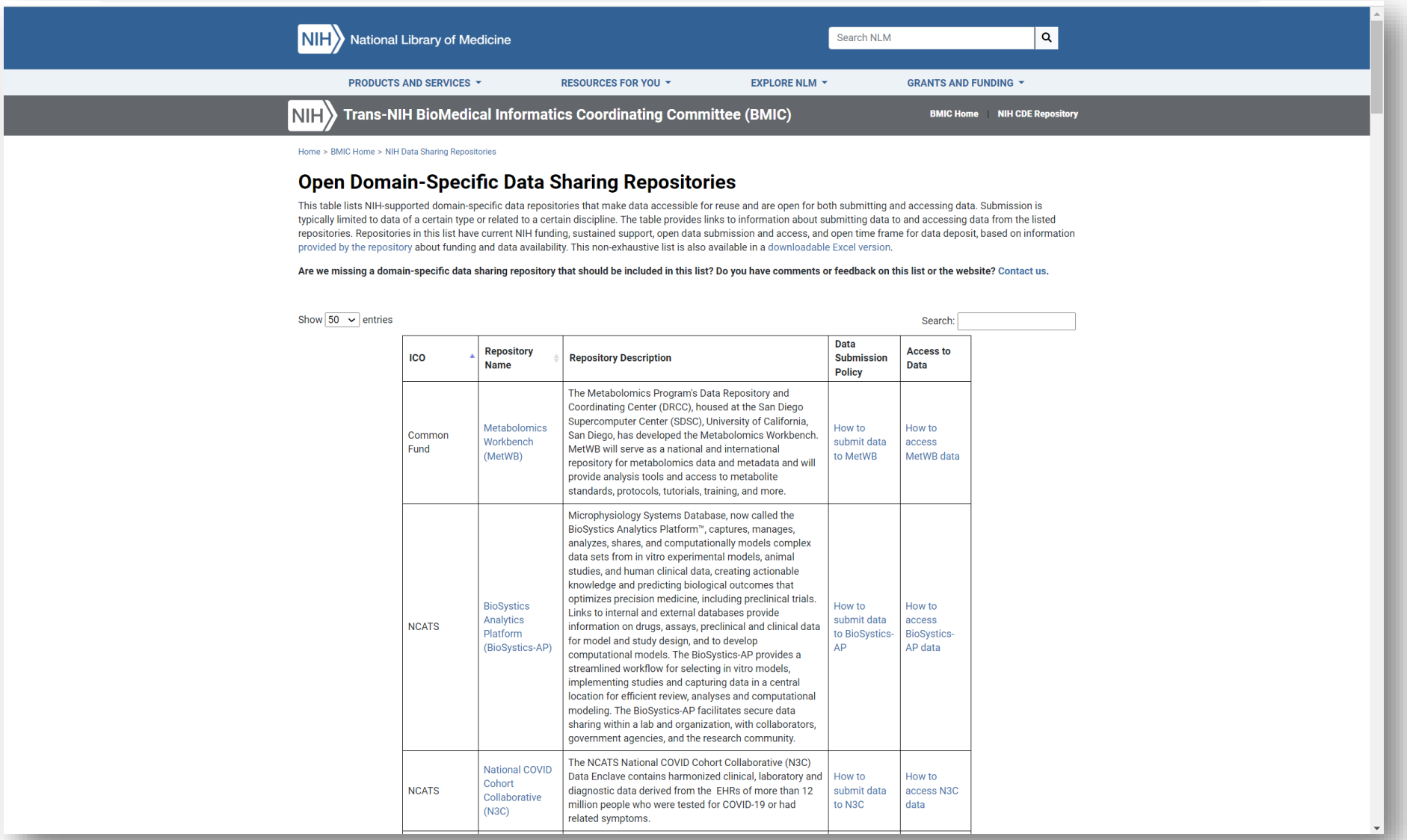
REPOSITORIES: CHARACTERISTICS AND SELECTION

For some programs and types of data, the NIH or the associated FOA may specify particular repositories.

For programs where no data repository is specified, researchers are encouraged to select a repository that is appropriate for the project, taking into consideration:

- Primary consideration should be given to data repositories that are discipline or data-type specific. NIH makes a list of such data repositories available (see https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html).
- If no such repositories exist, researchers should consider:
  - Small datasets (up to 2GB) may be included as supplemental material to accompany articles submitted to PubMed Central
  - Data repositories, including generalist repositories, that make data available to the larger research community
  - Cloud-based solutions for large datasets

OPEN NIH-SUPPORTED DATA SHARING REPOSITORIES



NIH National Library of Medicine

Search NLM

PRODUCTS AND SERVICES | RESOURCES FOR YOU | EXPLORE NLM | GRANTS AND FUNDING

NIH Trans-NIH BioMedical Informatics Coordinating Committee (BMIC)

BMIC Home | NIH CDE Repository

Home > BMIC Home > NIH Data Sharing Repositories

Open Domain-Specific Data Sharing Repositories

This table lists NIH-supported domain-specific data repositories that make data accessible for reuse and are open for both submitting and accessing data. Submission is typically limited to data of a certain type or related to a certain discipline. The table provides links to information about submitting data to and accessing data from the listed repositories. Repositories in this list have current NIH funding, sustained support, open data submission and access, and open time frame for data deposit, based on information provided by the repository about funding and data availability. This non-exhaustive list is also available in a downloadable Excel version.

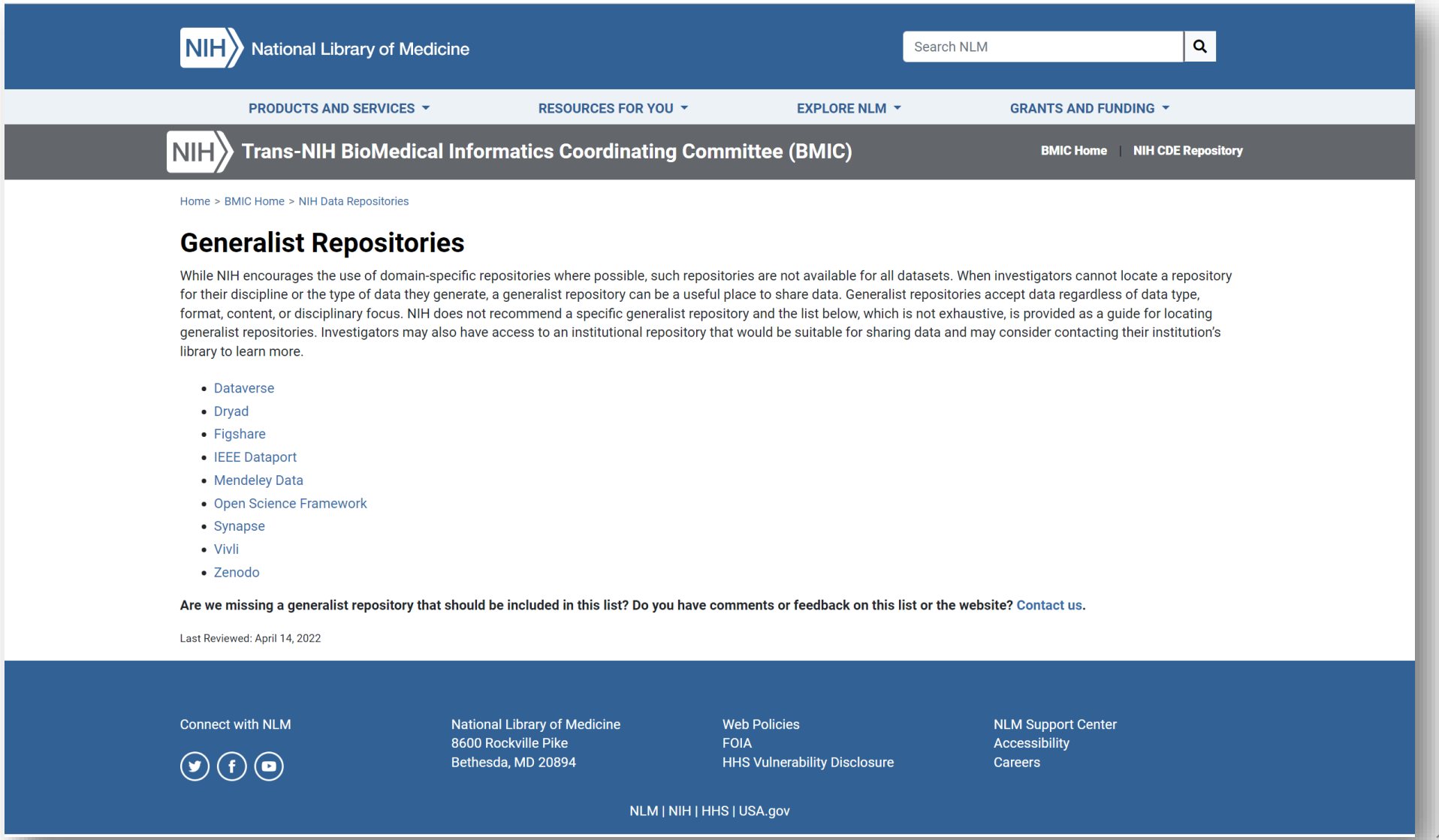
Are we missing a domain-specific data sharing repository that should be included in this list? Do you have comments or feedback on this list or the website? Contact us.

Show 50 entries Search:

ICO	Repository Name	Repository Description	Data Submission Policy	Access to Data
Common Fund	Metabolomics Workbench (MetWB)	The Metabolomics Program's Data Repository and Coordinating Center (DRCC), housed at the San Diego Supercomputer Center (SDSC), University of California, San Diego, has developed the Metabolomics Workbench. MetWB will serve as a national and international repository for metabolomics data and metadata and will provide analysis tools and access to metabolite standards, protocols, tutorials, training, and more.	How to submit data to MetWB	How to access MetWB data
NCATS	BioSystics Analytics Platform (BioSystics-AP)	Microphysiology Systems Database, now called the BioSystics Analytics Platform™, captures, manages, analyzes, shares, and computationally models complex data sets from in vitro experimental models, animal studies, and human clinical data, creating actionable knowledge and predicting biological outcomes that optimizes precision medicine, including preclinical trials. Links to internal and external databases provide information on drugs, assays, preclinical and clinical data for model and study design, and to develop computational models. The BioSystics-AP provides a streamlined workflow for selecting in vitro models, implementing studies and capturing data in a central location for efficient review, analyses and computational modeling. The BioSystics-AP facilitates secure data sharing within a lab and organization, with collaborators, government agencies, and the research community.	How to submit data to BioSystics-AP	How to access BioSystics-AP data
NCATS	National COVID Cohort Collaborative (N3C)	The NCATS National COVID Cohort Collaborative (N3C) Data Enclave contains harmonized clinical, laboratory and diagnostic data derived from the EHRs of more than 12 million people who were tested for COVID-19 or had related symptoms.	How to submit data to N3C	How to access N3C data

https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html

AVAILABLE THIRD-PARTY GENERALIST REPOSITORIES



NIH National Library of Medicine

Search NLM

PRODUCTS AND SERVICES | RESOURCES FOR YOU | EXPLORE NLM | GRANTS AND FUNDING

NIH Trans-NIH BioMedical Informatics Coordinating Committee (BMIC)

BMIC Home | NIH CDE Repository

Home > BMIC Home > NIH Data Repositories

Generalist Repositories

While NIH encourages the use of domain-specific repositories 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 can be a useful place to share data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus. NIH does not recommend a specific generalist repository and the list below, which is not exhaustive, is provided as a guide for locating generalist repositories. Investigators may also have access to an institutional repository that would be suitable for sharing data and may consider contacting their institution's library to learn more.

- Dataverse
- Dryad
- Figshare
- IEEE Dataport
- Mendeley Data
- Open Science Framework
- Synapse
- Vivli
- Zenodo

Are we missing a generalist repository that should be included in this list? Do you have comments or feedback on this list or the website? Contact us.

Last Reviewed: April 14, 2022

Connect with NLM

National Library of Medicine
8600 Rockville Pike
Bethesda, MD 20894

Web Policies
FOIA
HHS Vulnerability Disclosure

NLM Support Center
Accessibility
Careers

NLM | NIH | HHS | USA.gov

https://www.nlm.nih.gov/NIHbmic/generalist_repositories.html

REPOSITORIES: CHARACTERISTICS AND SELECTION – DESIRABLE CHARACTERISTICS FOR ALL DATA REPOSITORIES

UNIQUE PERSISTENT IDENTIFIERS

Assigns datasets a citable, unique persistent identifier (PID), such as a digital object identifier (DOI) or accession number.

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.

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).

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.

LONG-TERM SUSTAINABILITY

FREE AND EASY ACCESS

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.

PROVENANCE

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

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.

BROAD AND MEASURED REUSE

Makes datasets and their metadata available with broadest possible terms of reuse.

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.

RETENTION POLICY

Provides documentation on policies for data retention within the repository.

ADDITIONAL REPOSITORY CONSIDERATIONS FOR HUMAN DATA

https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository

When working with human participant data, including de-identified human data, here are some additional characteristics to look for:

FIDELITY TO CONSENT

Uses documented procedures to restrict dataset access and use to those that are consistent with participant consent and changes in consent.

RESTRICTED USE COMPLIANT

Uses documented procedures to communicate and enforce data use restrictions, such as preventing reidentification or redistribution to unauthorized users.

PRIVACY

Implements and provides documentation of measures (for example, 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 REVIEW

Makes use of an established and transparent process for reviewing data access requests.

BUDGETING FOR THE COST OF DATA REPOSITORIES

ALLOWABLE COSTS FOR DATA MANAGEMENT AND SHARING

All allowable costs must be submitted in the budget for the award and must be incurred during the performance of the award, even for scientific data and metadata preserved and shared beyond the award period.

Note that all allowable costs submitted in budget requests must be ***incurred during the performance period***, even for scientific data and metadata preserved and shared beyond the award period. For instance, if a DMS 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 before the end of the period of performance.

Reasonable and allowable costs include:

- Curating data and developing supporting documentation
- Preserving and sharing data through established repositories, such as data deposit fees necessary for making data available and accessible
- Formatting data according to accepted community standards, or for transmission to and storage at a selected repository for long-term preservation and access
- De-identifying data
- Preparing metadata to foster discoverability, interpretation, and reuse
- Local data management considerations, such as unique and specialized information infrastructure necessary to provide local management and preservation (for example, before deposit into an established repository)
- If the Data Management & Sharing (DMS) plan proposes deposition to multiple repositories, costs associated with each proposed repository may be included

Budget requests must not include infrastructure costs that are included in institutional overhead (e.g., F&A costs) or costs associated with the routine conduct of research.

PLAN ASSESSMENT: COMPLIANCE AND ENFORCEMENT

PLAN ASSESSMENT

The NIH ICO will assess the Plan, through the following processes:

- ☑ Extramural Awards: Plans will undergo programmatic assessment by NIH as determined by the proposed NIH ICO.
- ☑ Contracts: Plans will be included as part of the technical evaluation performed by NIH staff.
- ☑ Intramural Research Projects: Plans will be assessed in a manner determined to be appropriate by the Intramural Research Program.

Other funding agreements: Plans will be assessed in the context of other funding agreement mechanisms (e.g., Other Transactions).

PLAN COMPLIANCE AND ENFORCEMENT

- ☑ During the funding period, compliance with the Plan will be determined by the NIH ICO. Compliance with the Plan, including any Plan updates, may be reviewed during regular grant reporting intervals.
- ☑ After the end of the funding period, non-compliance with the NIH ICO-approved Plan may be taken into account by NIH for future funding decisions.


THANK YOU
& FIGHT ON!
USC University of Southern California

APPENDICES

APPENDIX ONE: ADDITIONAL GUIDANCE ON DATA SHARING PLANS

ADDITIONAL GUIDANCE: DEVELOPMENT OF DATA SHARING PLANS

01 DATA TYPE

Briefly describe the scientific data to be managed and shared:

- Summarize the types (for example, 256-channel EEG data and fMRI images) and amount (for example, from 50 research participants) of scientific data to be generated and/or used in the research. Descriptions may include the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing.
- Describe which scientific data from the project will be preserved and shared. NIH does not anticipate that researchers will preserve and share all scientific data generated in a study. Researchers should decide which scientific data to preserve and share based on ethical, legal, and technical factors. The plan should provide the reasoning for these decisions.
- A brief listing of 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.

02 RELATED TOOLS, SOFTWARE AND/OR CODE

Indicate whether specialized tools are needed to access or manipulate shared scientific data to support replication or reuse, and name(s) of the needed tool(s) and software. If applicable, specify how needed tools can be accessed.

03 STANDARDS

Describe what standards, if any, will be applied to the scientific data and associated metadata (i.e., data formats, data dictionaries, data identifiers, definitions, unique identifiers, and other data documentation).

ADDITIONAL GUIDANCE: DEVELOPMENT OF DATA SHARING PLANS

04 DATA PRESERVATION, ACCESS, AND ASSOCIATED TIMELINES

Give plans and timelines for data preservation and access, including:

- The name of the repository(ies) where scientific data and metadata arising from the project will be archived.
- How the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.
- When the scientific data will be made available to other users and for how long. Identify any differences in timelines for different subsets of scientific data to be shared.
    - Note that NIH encourages scientific data to be shared as soon as possible, and no later than the time of an associated publication or end of the performance period, whichever comes first. NIH also encourages researchers to make scientific data available for as long as they anticipate it being useful for the larger research community, institutions, and/or the broader public.

05 ACCESS, DISTRIBUTION, OR REUSE CONSIDERATIONS

Describe any applicable factors affecting subsequent access, distribution, or reuse of scientific data related to:

- Informed consent
- Privacy and confidentiality protections consistent with applicable federal, Tribal, state, and local laws, regulations, and policies
- Whether access to scientific data derived from humans will be controlled
- Any restrictions imposed by federal, Tribal, or state laws, regulations, or policies, or existing or anticipated agreements
- Any other considerations that may limit the extent of data sharing. Any potential limitations on subsequent data use should be communicated to the individuals or entities (for example, data repository managers) that will preserve and share the scientific data.

06 OVERSIGHT OF DATA MANAGEMENT AND SHARING

Indicate how compliance with the DMS plan will be monitored and managed.

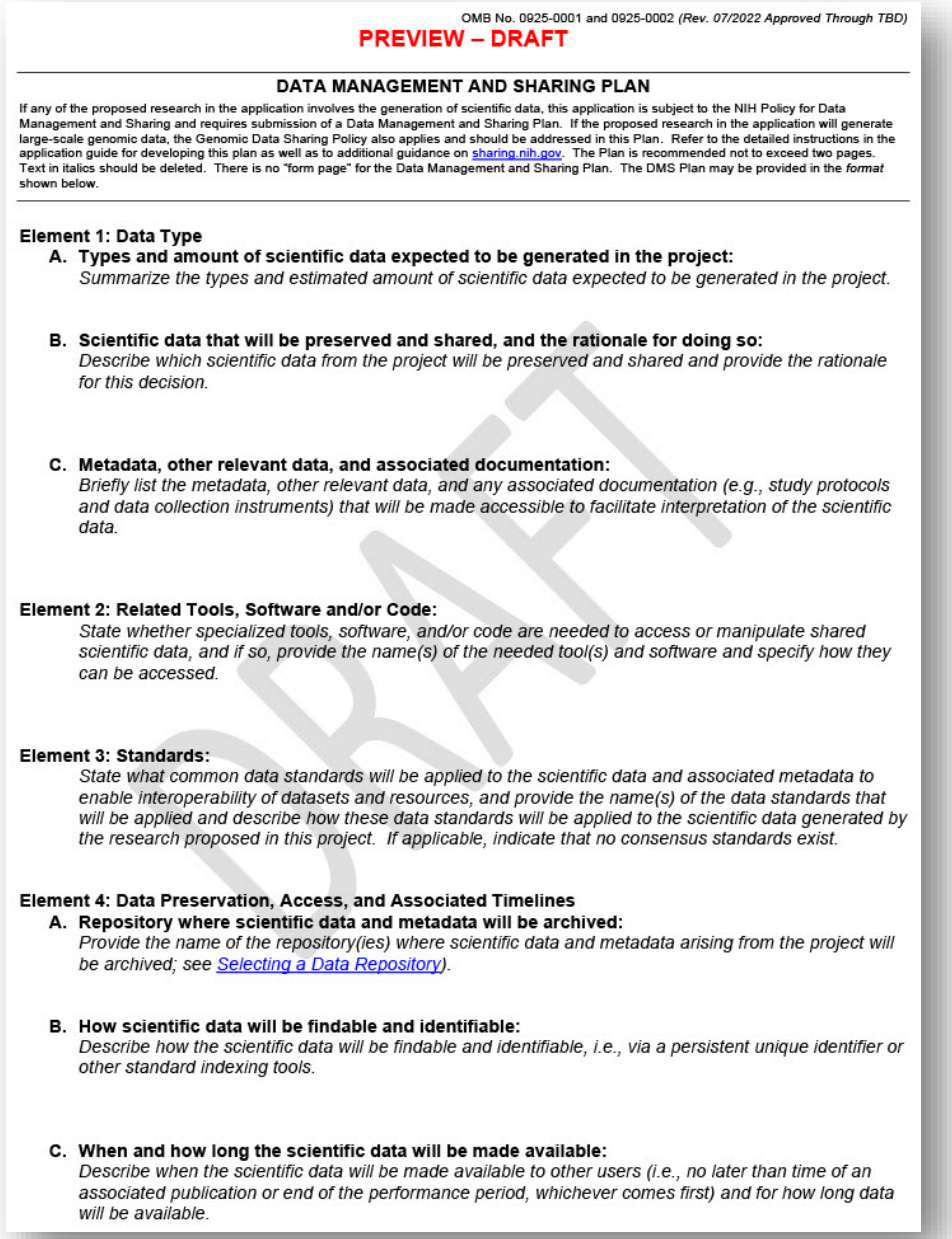
ADDITIONAL GUIDANCE: SPECIAL CONSIDERATIONS FOR RESEARCH INVOLVING HUMAN SUBJECTS

Award recipients must comply with any applicable laws, regulations, statutes, guidance, or institutional policies related to research with human participants and that protect participants' privacy. The DMS Policy encourages respect for participants by encouraging researchers and award recipients to:

- Address data management and sharing plans during the informed consent process to ensure prospective participants understand how their data will be managed and shared;
- Outline steps they will take for protecting the privacy, rights, and confidentiality of prospective participants (i.e., through de-identification, Certificates of Confidentiality, and other protective measures);
- Assess limitations on subsequent use of data and communicate these limitations to the individuals or entities (e.g., repositories) preserving and sharing the data; and
- Consider whether access to shared scientific data derived from humans should be controlled, even if de-identified and lacking explicit limitations on subsequent use. Sharing via controlled access may be specified by certain funding opportunity announcements (FOAs) or the funding NIH Institutes or Centers.

ADDITIONAL GUIDANCE: TEMPLATE DATA MANAGEMENT PLANS

https://grants.nih.gov/grants/forms/all-forms-and-formats.htm?category=format-pages&id=2099



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.

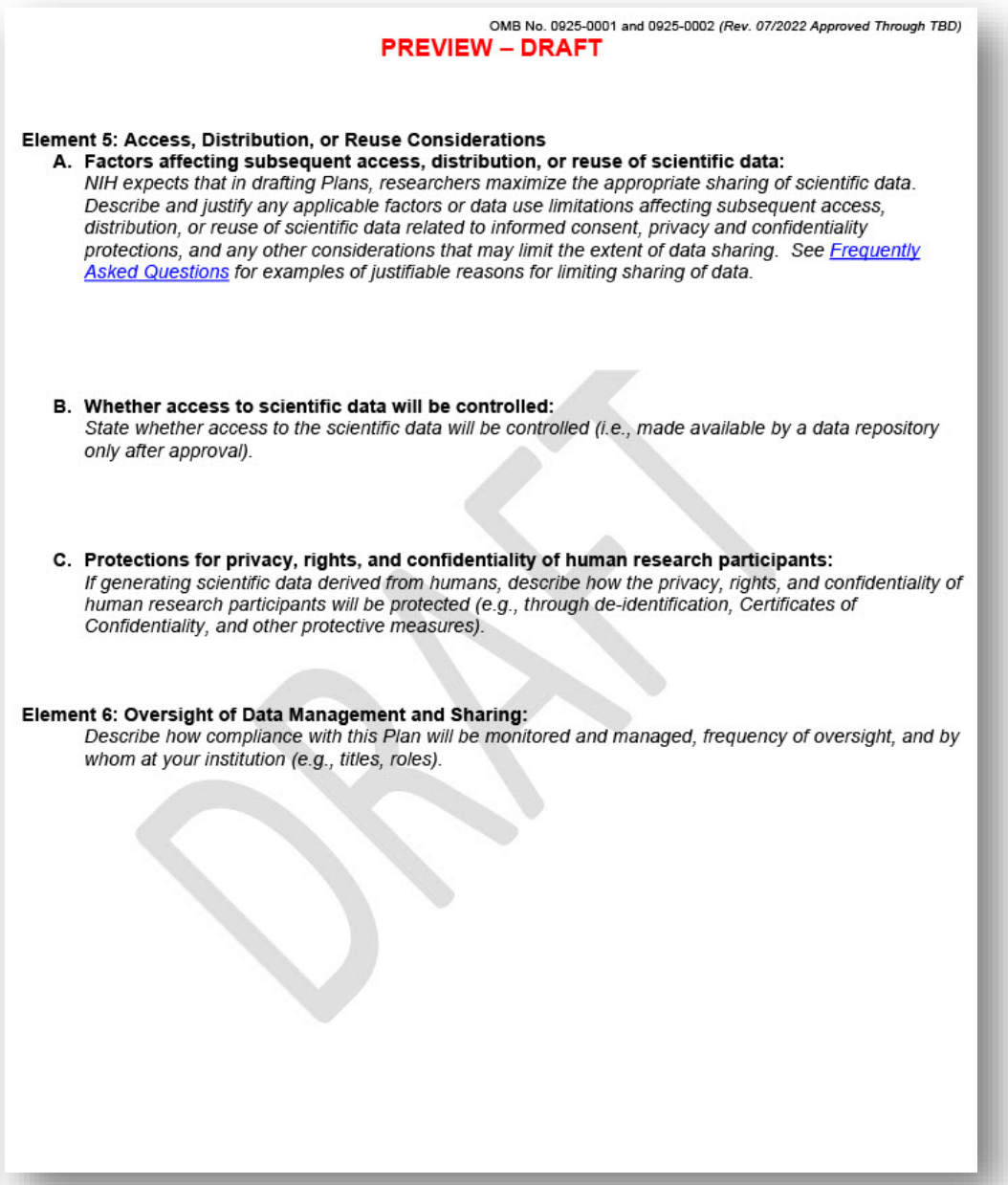
Element 3: Standards:
State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific 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.

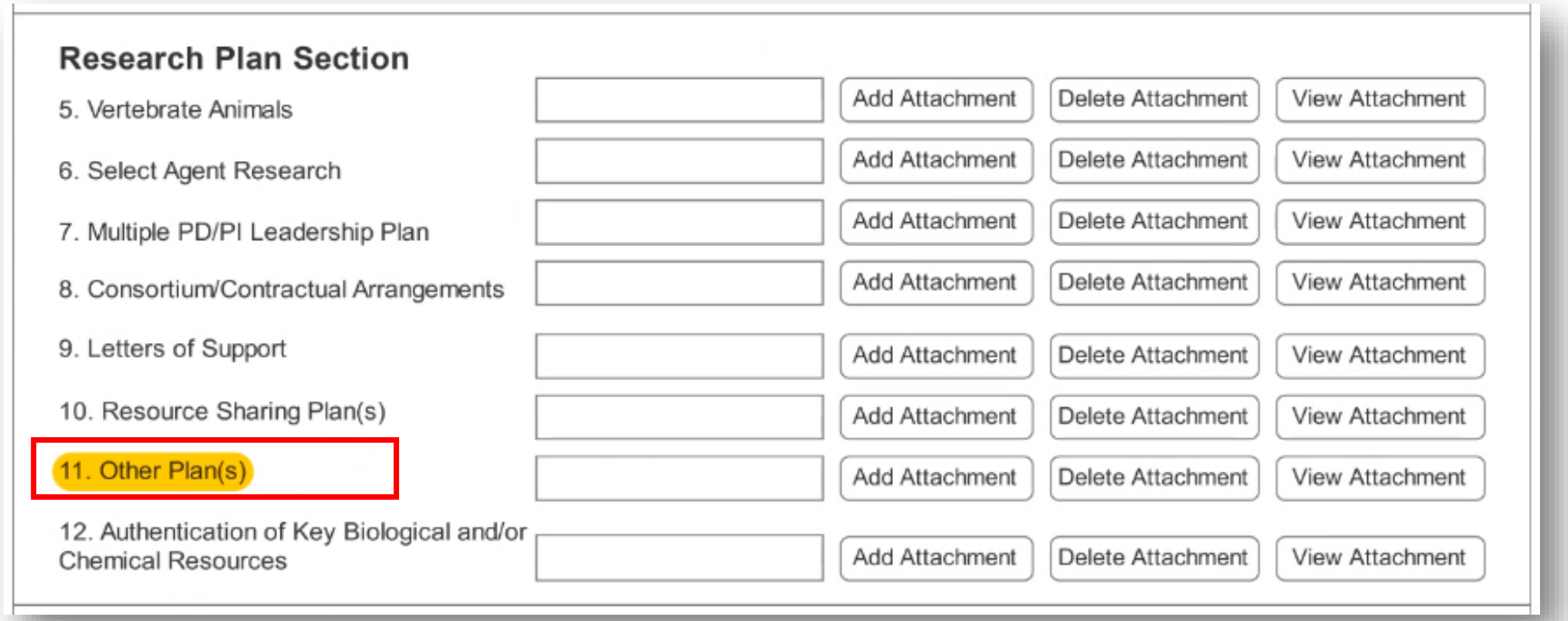
B. Whether access to scientific data will be controlled:
State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

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).

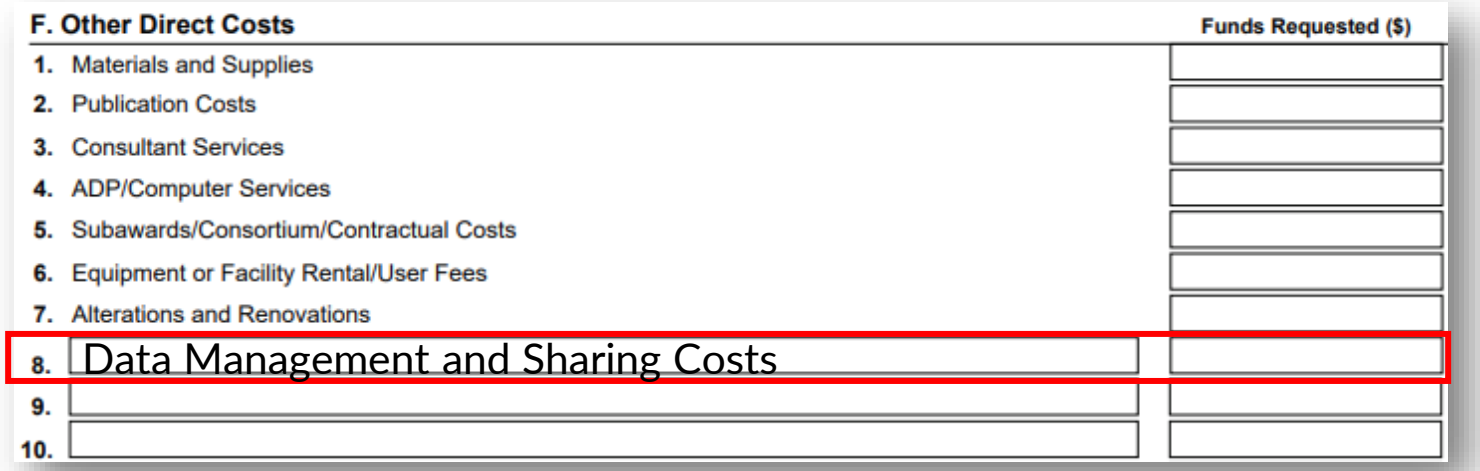
ADDITIONAL GUIDANCE: DATA MANAGEMENT PLAN SUBMISSION

- A new **"Other Plan(s)"** field will be added to the PHS 398 form to collect a single PDF attachment.
- Data Sharing Plans and Genomic Data Sharing Plans will no longer be submitted to the "Resource Sharing Plan(s)" field.



ADDITIONAL GUIDANCE: SUBMITTING DMS BUDGETS

- Direct costs to support the activities proposed in the DMS Plan must be indicated as **"Data Management and Sharing Costs"**
  - **R&R Budget Form**: line item in section F. Other Direct Costs



  - **PHS 398 Modular Budget Form**: within Additional Narrative Justification

APPENDIX TWO: BENCHMARKING OF AVAILABLE GENERALIST REPOSITORIES

USC University of Southern California

GENERALIST REPOSITORIES BENCHMARKING

DESCRIPTION	INDIVIDUAL COST	UNIQUE PERSISTENT IDENTIFIERS	LONG-TERM SUSTAINABILITY	METADATA	CURATION AND QUALITY ASSURANCE	FREE AND EASY ACCESS	BROAD AND MEASURED REUSE	CLEAR USE GUIDANCE	SECURITY AND INTEGRITY	CONFIDENTIALITY	COMMON FORMAT	PROVENANCE	RETENTION POLICY
Dryad	Base data publishing charge per data submission is $120 USD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Figshare	Can create free accounts and publish up to 20GB file sets at a time	✓	✓	✓	Partially met	✓	✓	✓	✓	Partially met	Partially met	✓	✓
Mendeley Data	Free up to 10 GB (if dataset is larger than 10 GB, a different repository is required)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Open Science Framework	Free for researchers to use (single file limit of 5 GB); Storage fees of $500 and above depending on space required	✓	Partially met	Partially met	Partially met	✓	Partially met	✓	✓	✓	✓	✓	✓
Vivli NOTE: Vivli is primarily for clinical trials, and as such, requires anonymizing the data before uploading.	Per study option for individual researchers--Free. There is a charge to 3rd party vendors for anonymizing the data, but it can be $5,000-$10,000 depending on the size of the data and the type of study,	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

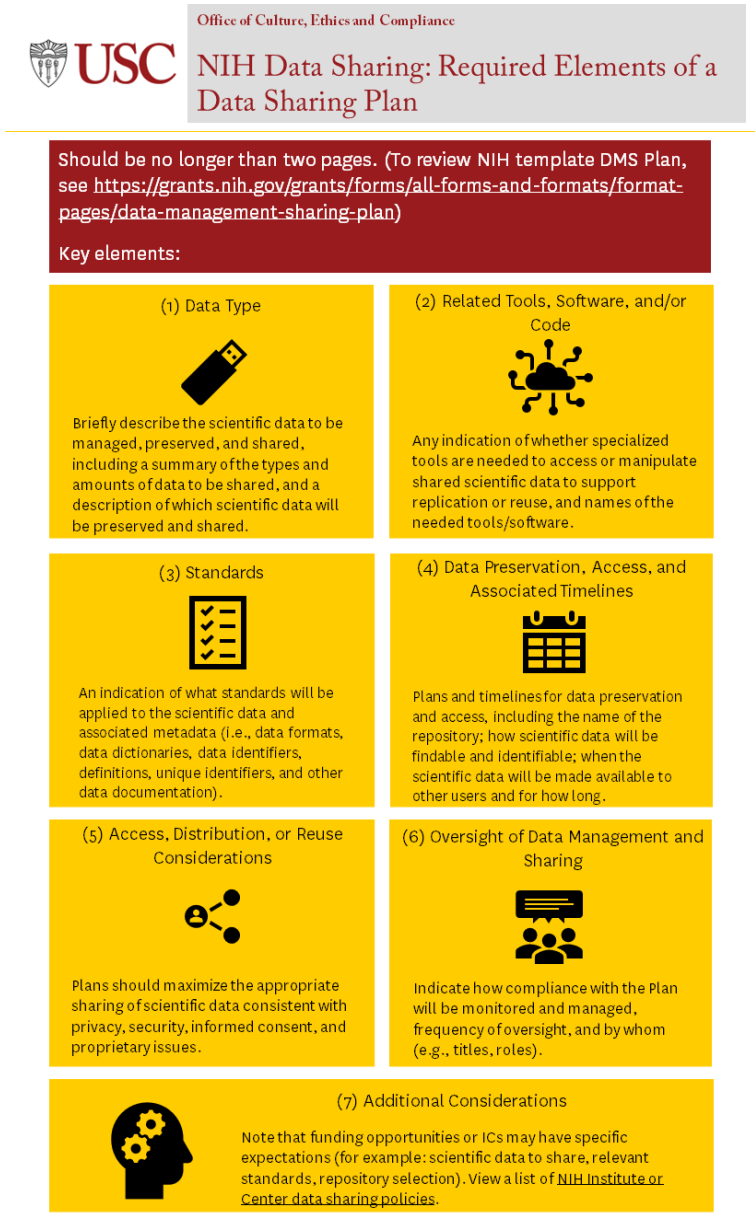
APPENDIX THREE:
USC RESOURCES

DATA MANAGEMENT POLICY: ONE-SHEETS

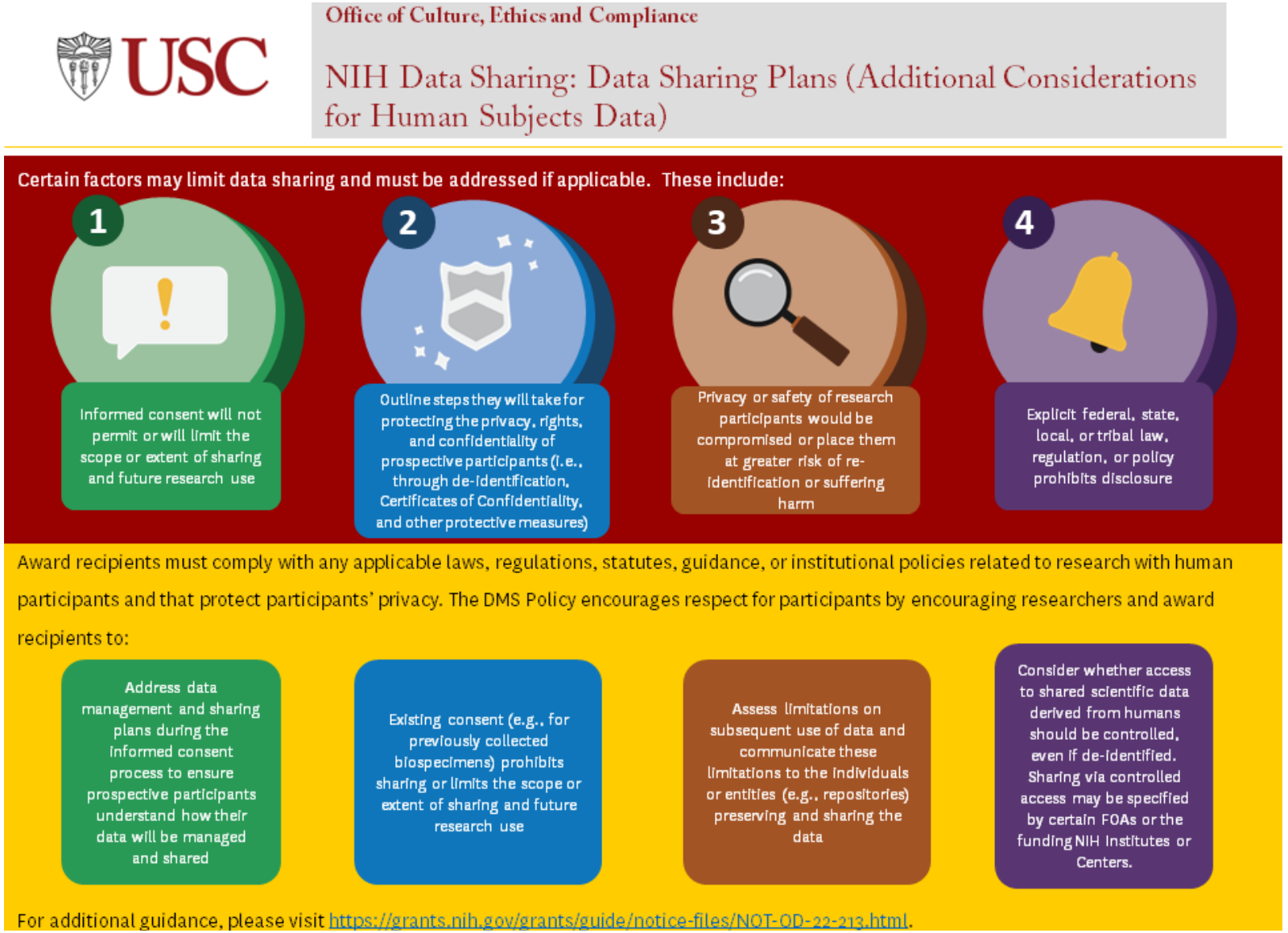
USC provides a library of one-sheets. Please visit the Department of Contracts and Grants website to download any of these one-sheets: https://dcg.usc.edu/nih-data-management/

- Required Elements of a Data Sharing Plan
- Data Sharing Plans (Additional Considerations for Human Subjects Data)
- Selection of a Data Repository
- Selection of a Data Repository (Additional Considerations for Human Subjects Data)
- How to Submit a DMS Plan
- Available Generalist Repository Overview
- Allowable Costs for Data Management and Sharing

ONE-SHEET: REQUIRED ELEMENTS OF A DATA SHARING PLAN



ONE-SHEET: DATA SHARING PLANS (ADDITIONAL CONSIDERATIONS FOR HUMAN SUBJECTS DATA)



ONE-SHEET: SELECTION OF A DATA REPOSITORY

USC University of Southern California

USC

Office of Culture, Ethics and Compliance

NIH Data Sharing: Selection of a Data Repository: Desirable Characteristics

Unique Persistent Identifiers

Assigns datasets a citable, unique persistent identifier (PID), such as a digital object identifier (DOI) or accession number.

Long-Term Sustainability

Has a plan for long-term management of data.

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.

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.

Broad and Measured Reuse

Makes datasets and their metadata available with broadest possible terms of reuse.

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

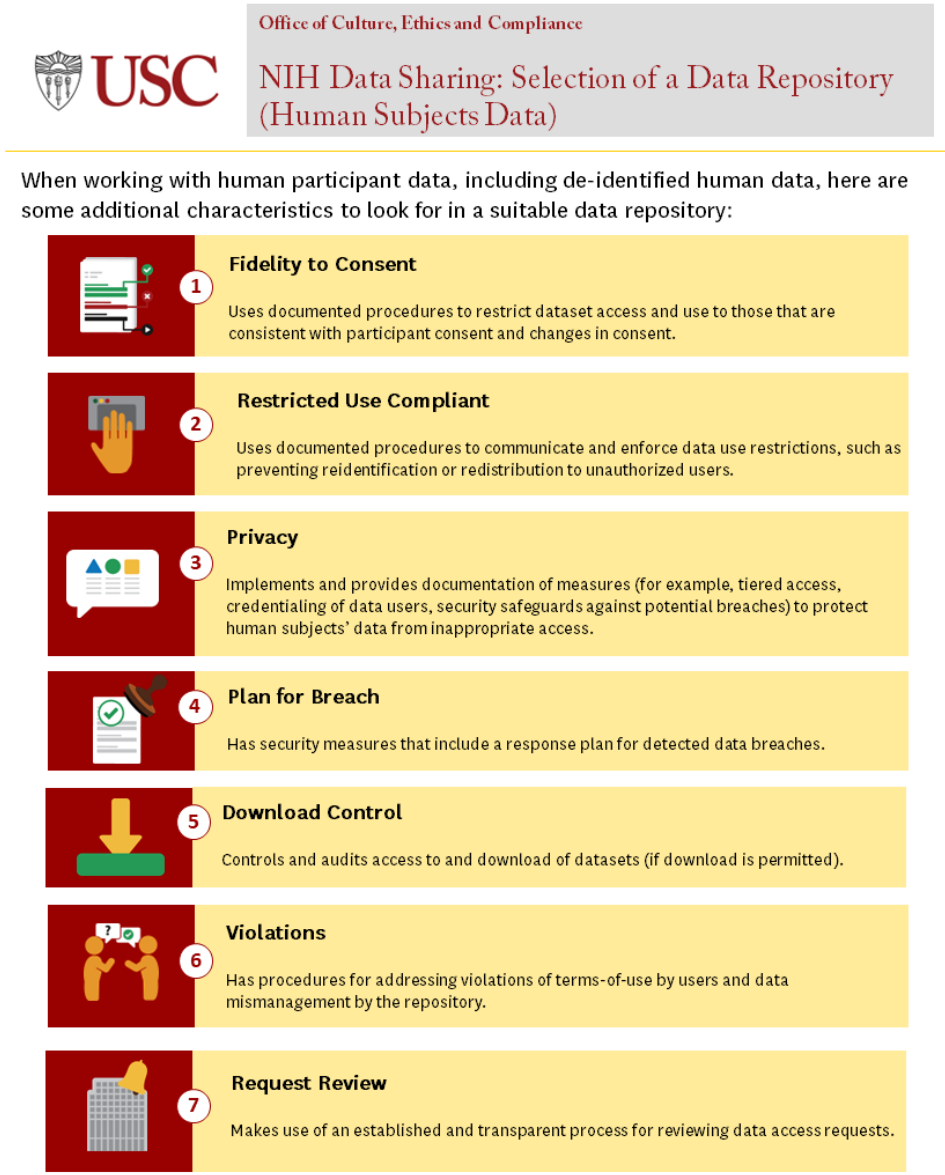
Has mechanisms in place to record the origin, chain of custody, and any modifications to submitted datasets and metadata.

Retention Policy

Provides documentation on policies for data retention within the repository.

Need more help? Call OCEC: 213-740-8258 Send an email: compliance@usc.edu Visit the website

ONE-SHEET: SELECTION OF A DATA REPOSITORY (ADDITIONAL CONSIDERATIONS FOR HUMAN SUBJECTS DATA)



USC

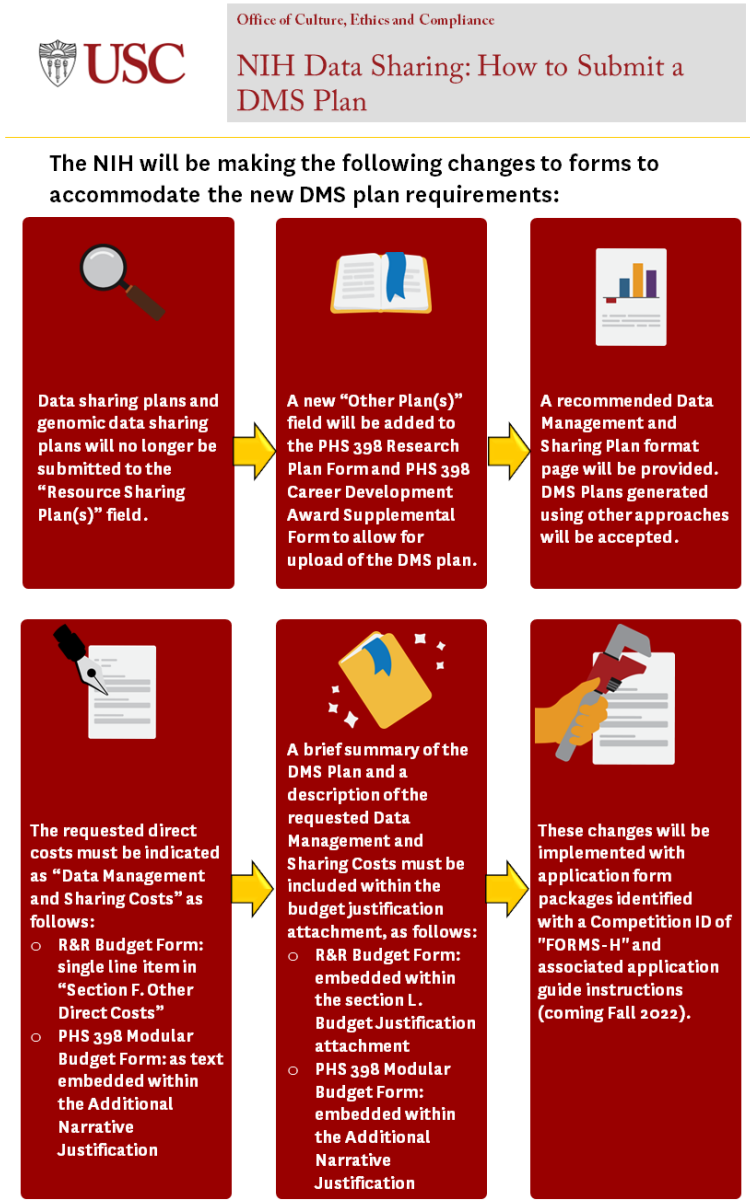
Office of Culture, Ethics and Compliance

NIH Data Sharing: Selection of a Data Repository (Human Subjects Data)

When working with human participant data, including de-identified human data, here are some additional characteristics to look for in a suitable data repository:

1. **Fidelity to Consent**
   Uses documented procedures to restrict dataset access and use to those that are consistent with participant consent and changes in consent.
2. **Restricted Use Compliant**
   Uses documented procedures to communicate and enforce data use restrictions, such as preventing reidentification or redistribution to unauthorized users.
3. **Privacy**
   Implements and provides documentation of measures (for example, tiered access, credentialing of data users, security safeguards against potential breaches) to protect human subjects' data from inappropriate access.
4. **Plan for Breach**
   Has security measures that include a response plan for detected data breaches.
5. **Download Control**
   Controls and audits access to and download of datasets (if download is permitted).
6. **Violations**
   Has procedures for addressing violations of terms-of-use by users and data mismanagement by the repository.
7. **Request Review**
   Makes use of an established and transparent process for reviewing data access requests.

ONE-SHEET: HOW TO SUBMIT A DMS PLAN



ONE-SHEET: AVAILABLE GENERALIST REPOSITORY OVERVIEW

USC

Office of Culture, Ethics and Compliance

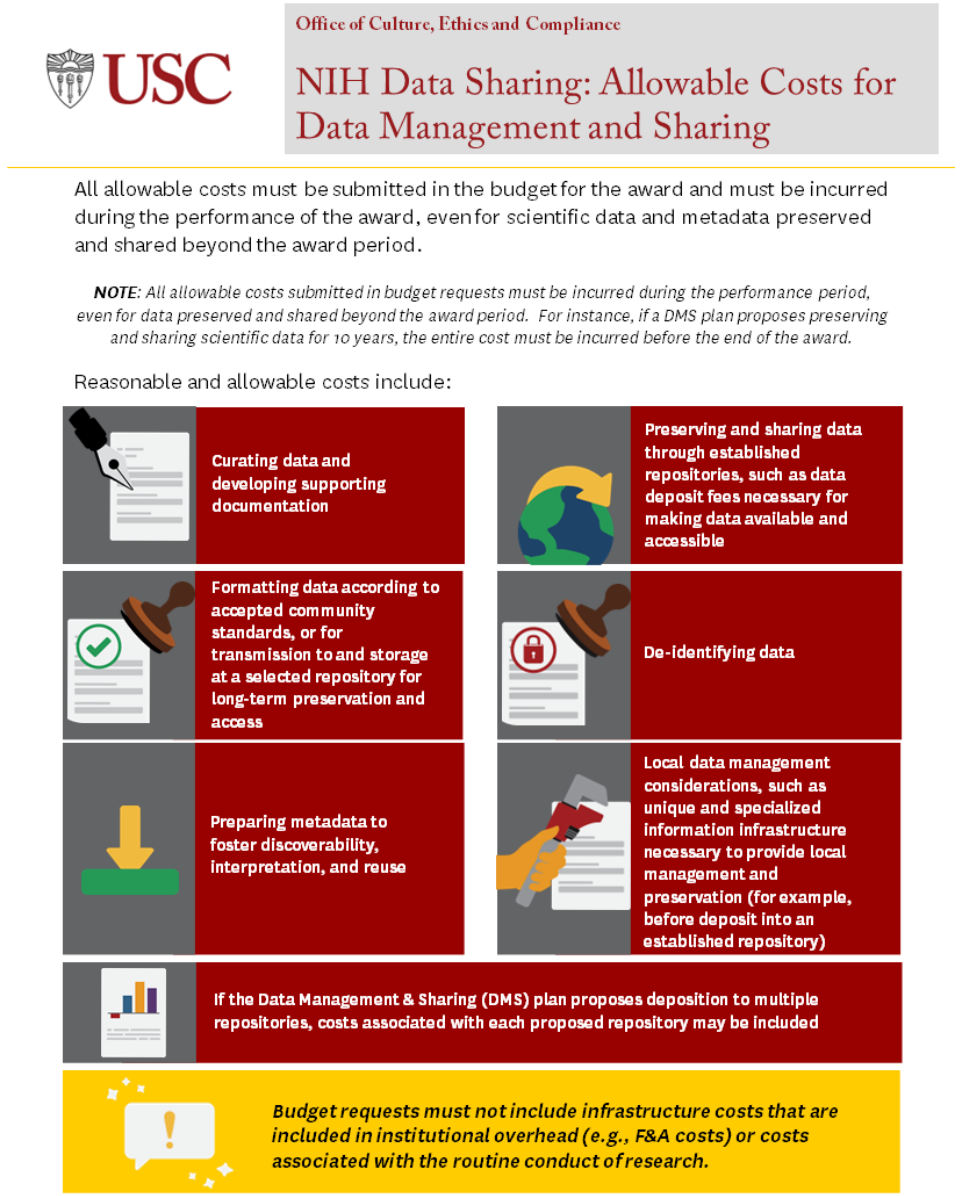
NIH Data Sharing: Available Generalist Repository Benchmarking

	Cost	Unique Persistent Identifiers	Long-term Sustainability	Metadata	Curation and Quality Assurance	Free and Easy Access	Broad and Measured Reuse	Clear Use Guidance	Security and Integrity	Confidentiality	Common Format	Provenance	Retention Policy
Dryad	Base data publishing charge per data submission is $120 USD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Figshare	Can create free accounts and publish up to 20GB file sets at a time	✓	✓	✓	Partially met	✓	✓	✓	✓	Partially met	Partially met	✓	✓
Mendeley Data	Free up to 10 GB (if dataset is larger than 10 GB, a different repository is required)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Open Science Framework	Free for researchers to use (single file limit of 5 GB); storage fees of $500 and above depending on space required	✓	Partially met	Partially met	Partially met	✓	Partially met	✓	✓	✓	✓	✓	✓
Vivli NOTE: Vivli is primarily for clinical trials, and as such, requires anonymizing the data before uploading.	Per study option for individual researchers--free. There is a charge from 3rd party vendors to anonymize data, which can be $5,000-$10,000 depending on the amount and type.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

For a full list, visit https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/generalist-repositories

ONE-SHEET: ALLOWABLE COSTS FOR DATA MANAGEMENT AND SHARING

USC University of Southern California



USC

Office of Culture, Ethics and Compliance

NIH Data Sharing: Allowable Costs for Data Management and Sharing

All allowable costs must be submitted in the budget for the award and must be incurred during the performance of the award, even for scientific data and metadata preserved and shared beyond the award period.

***NOTE**: All allowable costs submitted in budget requests must be incurred during the performance period, even for data preserved and shared beyond the award period. For instance, if a DMS plan proposes preserving and sharing scientific data for 10 years, the entire cost must be incurred before the end of the award.*

Reasonable and allowable costs include:

- Curating data and developing supporting documentation
- Preserving and sharing data through established repositories, such as data deposit fees necessary for making data available and accessible
- Formatting data according to accepted community standards, or for transmission to and storage at a selected repository for long-term preservation and access
- De-identifying data
- Preparing metadata to foster discoverability, interpretation, and reuse
- Local data management considerations, such as unique and specialized information infrastructure necessary to provide local management and preservation (for example, before deposit into an established repository)
- If the Data Management & Sharing (DMS) plan proposes deposition to multiple repositories, costs associated with each proposed repository may be included

Budget requests must not include infrastructure costs that are included in institutional overhead (e.g., F&A costs) or costs associated with the routine conduct of research.

APPENDIX FOUR:
NIH RESOURCES

USC University of Southern California

DATA MANAGEMENT POLICY: RESOURCES

Please review the following resources to review additional information from the NIH on scientific data sharing:

- NIH website on Scientific Data Sharing: https://sharing.nih.gov/
- NIH Guide notices related to Data Management and Sharing Plan requirements:
  - NOT-OD-21-013 – Final NIH Policy for Data Management and Sharing
  - 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-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-22-131 – 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-22-189 – Implementation Details for the NIH Data Management and Sharing Policy
  - NOT-OD-22-195 – New NIH "FORMS-H" Grant Application Forms and Instructions Coming for Due Dates on or after January 25, 2023
- Federation of American Societies for Experimental Biology (FASEB) Data Management Plan Challenge ($500 prize!!): https://www.faseb.org/resources/data-science-and-informatics/dataworks-dmp-challenge

DATA MANAGEMENT POLICY: ACTIVITIES NOT SUBJECT TO THE POLICY

Research projects <u>not</u> generating scientific data or non-research projects, including but not limited to:

- Training (Ts)
- Fellowships (Fs)
- Certain non-research Career Awards (e.g., KM1)
- Construction (C06)
- Conference Grants (R13)
- Resources (Gs)
- Research-Related Infrastructure Programs (e.g., S06)