

NIH Data Management and Sharing Policy 2023: Overview and Resources

Wasila Dahdul, Data Curation Librarian
Digital Scholarship Services

UCI Libraries



Final NIH Policy for Data Management and Sharing

Notice Number:

NOT-OD-21-013

Key Dates

Release Date:

October 29, 2020

Effective Date:

January 25, 2023

Previous NIH Sharing Policies

- **2003: Data Sharing Policy** requires investigators seeking \$500,000 or more in NIH funding to submit a data sharing plan (or rationale for not sharing). Superseded by the new policy in January 2023.
- **2008: Public Access Policy** requires NIH-funded scientists to submit final peer-reviewed journal manuscripts to PubMed Central (PMC) no later than 12 months after publication.
- **2014: Genomic Data Sharing Policy** requires investigators generating large-scale genomic data to submit a genomic data sharing plan.

What is the new NIH policy?

NIH's New Data Management & Sharing Policy (DMSP)

Effective January 25, 2023

- Requires researchers seeking NIH funding to prospectively **submit a plan outlining how scientific data from their research will be managed and shared.**
- Researchers should “maximize the appropriate sharing of scientific data.”
- Data should be shared as soon as possible, and no later than the time of an associated publication or end of performance period (whichever comes first).
- This plan represents the minimum requirements. NIH ICOs may expect more specificity in their plans - check funding announcements for information.

NIH's New Data Management & Sharing Policy (DMSP)

- Applies to all research funded in whole or in part by NIH that generates scientific data:
 - Extramural grants
 - Contracts
 - Intramural research projects
 - All other NIH funding mechanisms
- Exception: funding that does not generate data (e.g., training grants).

“Scientific data”

Defined as “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.”

Does 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.”

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

What is a data management plan
(DMP)?

DMP Basics

A document that addresses how you will **manage** and **secure** your data throughout the lifecycle of a research project

Can be both a **required document** for grants as well as a **guide to daily activities**

What's a good DMP?

A good DMP should have a clear, organized and effective system to manage data ***throughout*** the project.

Include plans for the data ***after*** the research is complete.

*The most important component of most federal data management plans is on **data sharing and data preservation**.*

Elements of the NIH DMSP

Elements of the plan

DMSP: A brief document (2 pages or less) submitted with the application for NIH funding for projects generating scientific data that describes:

- Data type(s) and metadata (data description)
- Related tools, software, and/or code
- Standards for the data/metadata
- Data preservation, access, and associated timelines
- Access, distribution, or reuse considerations
- Oversight of data management and sharing

Source: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-014.html>

1. Data Type(s)

- Summarize the **types** and **amounts** of data you are collecting.
- Which scientific data from the project will be **preserved** and **shared**? Note: it doesn't have to be everything! You are expected to maximize the appropriate sharing of data; the plan should supply the **rationale**.
- What **metadata** or **documentation** (study protocols, data collection instruments) will be made accessible to facilitate interpretation of the data?
- Collecting **sensitive data**? Campus-specific guidance or policies may apply.

2. Related Tools, Software and/or Code

- Are **specialized tools** needed to access or manipulate your scientific data to support replication or reuse? If so, which ones?
- How can these tools be **accessed**? Are they open source or do they require a license?
- Are these tools likely to remain available for as long as the scientific data remain available?

3. Standards

- 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)?
- Some fields have community-developed standards while others do not. Indicate if no standards have been established.
- Research Data Alliance: [Directory of Metadata Standards](#)

4. Data Preservation, Access & Timelines

- What research **data repositories** will you use for your data?
- How will your data be **findable** and **identifiable** (e.g., persistent unique identifier)?
- When will it be made **available** to others and for how long?

*Note: NIH encourages data to be made available **as soon as possible**, and no later than the time of an associated publication or end of the performance period, whichever comes first.*

Selecting a data repository

Primary consideration should be given to data repositories that are **discipline or data-type specific**

- Search for NIH Institute and Center required or recommended data repositories, <https://sharing.nih.gov/other-sharing-policies/nih-institute-and-center-data-sharing-policies>
- List of NIH-supported [Open Domain-Specific Data Sharing Repositories](#) and [Other NIH Data Resources](#)

Use a **generalist repository** if no appropriate discipline or data-type-specific repository is available

- UC's [Dryad Digital Repository](#) makes data openly available to the larger research community
- List of [NIH generalist repositories](#) (GREI), including **Dryad**
- Small datasets (up to 2 GB in size) may be included as supplementary material for articles submitted to PubMed Central

Source: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html>

Desirable characteristics of data repositories

- Unique Persistent Identifiers
- Confidentiality
- Long-Term Sustainability
- Common Format
- Metadata
- Provenance
- Curation and Quality Assurance
- Retention Policy
- Free and Easy Access
- Security and Integrity
- Broad and Measured Reuse
- Clear Use Guidance

Source: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html>

5. Access, Distribution, Reuse

Describe any factors affecting access, distribution or reuse including:

- Informed consent - will you get consent to share? *Note: NIH recommends addressing data management and sharing during the informed consent process.*
- How are you ensuring privacy and confidentiality (de-identification, etc)?
- Will access to the data be controlled?
- Is your data subject to any restrictions on access (HIPAA, Tribal or state laws, etc.)?

Not all data needs to be shared

- informed consent limitations
- existing agreements prohibit sharing
- privacy or safety of research participants need protection
- explicit law or regulation prohibiting sharing
- data cannot be digitized with reasonable effort

6. Oversight of Data Management and Sharing

- How will you monitor compliance with this plan?
- Who will monitor compliance and how often?
- A list of possible data management roles:

<https://tinyurl.com/DMProles>

How are these plans evaluated/enforced?

- NIH ICO staff will assess the plans, grant reviewers can comment on the proposed budget for data management and sharing.
- Plans will be part of the Terms and Conditions for extramural awards and non-compliance can result in termination of award and impact future funding decisions.
- Compliance will be monitored by the relevant NIH institute during progress reports.

Data management resources at UCI



- Create data management plans that meet funder requirements
- UC campuses partner with the California Digital Library to offer DMPTool
- DMP templates for 22 federal and private funders available, including NIH and NSF
- Free for UCI researchers
- <https://dmptool.org>



Create Data Management Plans that meet requirements and promote your research

**77,521 Users****343 Participating Institutions****75,318 Plans**[Sign in / Sign up](#)**Email address ***

For SSO, use institutional address.

[Continue](#)

- Sign in with UCI email
- Link to ORCID ID

[Problems signing in? Contact us.](#)**Latest News from DMPTool**














Supporting FAIR Data with Integrations

[View all news](#)

Funder Requirements

Templates for data management plans are based on the specific requirements listed in funder policy documents. The DMPTool maintains these templates, however, researchers should always consult the program officers and policy documents directly for authoritative guidance. Sample plans are provided by a funder or another trusted party.

Search

Template Name 	Download	Organization name 	Last Updated 	Funder Links	Create a new plan 	Sample Plans (if available)
NIH-GDS: Genomic Data Sharing	 	National Institutes of Health (nih.gov)	10-25-2021	NIH Genomic Data Sharing Policy [PDF] NIH Public Access Plan [PDF] NIH GDS Policies and Supplemental Information		
NIH-GEN DMSP (Forthcoming 2023)	 	National Institutes of Health (nih.gov)	03-14-2022	Final NIH Policy for Data Management and Sharing		Elements of an NIH Data Management and Sharing Plan Selecting a Repository for Data Resulting from NIH-Supported Research Allowable Costs for Data Management and Sharing
NIH-GEN: Generic (Current until 2023)	 	National Institutes of Health (nih.gov)	10-25-2021	NIH Data Sharing Policy and Implementation Guidance NIH Public Access Plan [PDF] NIH Data Sharing Policies by Program		NIH Examples of Data Sharing Plans

- NIH “Generic” template for the 2023 policy
- Specific ICO templates forthcoming

NIH Data Sharing Plan

[Project Details](#)

[Contributors](#)

[Plan overview](#)

[Write Plan](#)

[Share](#)

[Download](#)

[expand all](#) | [collapse all](#)

0/12

+ Data Type (0 / 3)

+ Related Tools, Software and/or Code (0 / 2)

+ Standards (0 / 1)





+ Data Preservation, Access, and Associated Timelines (0 / 3)

+ Access, Distribution, or Reuse Considerations (0 / 2)

+ Oversight of Data Management and Sharing (0 / 1)

Briefly describe the scientific data to be managed, preserved, and shared.

A general summary of the types and estimated amount of scientific data to be generated and/or used in the research. Describe data in general terms that address the type and amount/size of scientific data expected to be collected and used in the project (e.g., 256-channel EEG data and fMRI images from ~50 research participants). Descriptions may indicate the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing that has occurred (i.e., how raw or processed the data will be)

B *I*    

Save

NIH example answer

This project will produce _____ [Data type, e.g., imaging, sequencing, experimental measurements] data generated/obtained from _____ [e.g., instrument, method, survey, experiment, data repository]. Data will be collected from ____ [number] of research participants/specimens/experiments, generating ____ [number] datasets totaling approximately ____ [amount of data] in size. The following data files will be used or produced in the course of the project: _____ [list input data files, intermediate files, and final, post-processed files]. Raw data will be transformed by ____ [analysis, method] and the subsequent processed dataset used for statistical analysis. To protect research participant identities, _____ [e.g., individual, aggregated, summarized] data will be made available for sharing.

Comments & Guidance

Guidance

NIH

DMPTool

NIH Guidance

The final DMS Policy defines S factual material commonly ac community as of sufficient qu: research findings, regardless c support scholarly publication: laboratory notebooks, prelimi report forms, drafts of scientif research, peer reviews, comm physical objects, such as labor

Even those scientific data not are considered scientific data Policy's scope. We understand not necessarily mean that the however, indicating that scien independent of publication is underlying null or negative fin

NIH template also available for download in Word or PDF

- Contains the guidance and sample answers aligned with the new NIH Data Policy
- [View and download](#) for offline editing

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.*

NIH Guidance

The final DMS Policy has [specific definitions for what Scientific Data is, and what proposals are considered to be producing scientific data](#). Per the [Policy](#), "Even those scientific data not used to support a publication are considered scientific data and within the final DMS Policy's scope. We understand that a lack of publication does not necessarily mean that the findings are null or negative; however, indicating that scientific data are defined independent of publication is sufficient to cover data underlying null or negative findings."

NIH Genomic Data Sharing (GDS) Policy Considerations

Check if your research is subject to NIH GDS (Genomic Data Sharing) policy using [this criteria](#) and list those data and the levels of processing here.

Individual NIH Institutes and Centers (IC) may have additional expectations or requirements for genomic data sharing as well. Please check the [IC-specific genomic data sharing requirements](#).

Example Answers:

This project will produce _____ [Data type, e.g., imaging, sequencing, experimental measurements] data generated/obtained from _____ [Data modality, e.g., instrument, method, survey, experiment, data source]. Data will be collected from ____ [number] of research participants/specimens/experiments, generating ____ [number] datasets totaling approximately ____ [amount of data] in size. The following data files will be used or produced in the course of the project: _____ [list input data files,



- General-purpose data repository that makes data and code discoverable, freely reusable, and citable
- UC campuses partner with CDL to co-develop Dryad
- Publisher and researcher workflow integrations
- Curation and quality assurance workflow during submission
- Free for UCI researchers
- <https://datadryad.org/stash>



DRYAD

[Explore Data](#)[About ▼](#)[Help ▼](#)[Login](#)

Make the most of your research data

- Sign in with UCI email
- Link to ORCID ID

Dryad is a community-owned resource
[Learn more about our organizational memberships](#)

[Submit Now](#)

Data from: Genetic identification of source and likely vector of a widespread marine invader

Krueger-Hadfield, Stacy A., University of Alabama at Birmingham

Kollars, Nicole M., College of Charleston

Strand, Allan E., College of Charleston

Byers, James E., University of Georgia

Shainker, Sarah J., College of Charleston

Terada, Ryuta, Kagoshima University

Greig, Thomas W., National Ocean Service

Hammann, Marieke, College of Charleston

Murray, David C., College of Charleston

Weinberger, Florian, GEOMAR Helmholtz Centre for Ocean Research Kiel

Sotka, Erik E., College of Charleston

Publication date: April 24, 2018

Publisher: Dryad

<https://doi.org/10.5061/dryad.fn53k>

Citation

Krueger-Hadfield, Stacy A. et al. (2018), Data from: Genetic identification of source and likely vector of a widespread marine invader, Dryad, Dataset, <https://doi.org/10.5061/dryad.fn53k>

Abstract

The identification of native sources and vectors of introduced species informs its ecological and evolutionary history and may guide policies that seek to prevent future introductions. Population genetics represents a powerful set of tools to identify origins and vectors, but can mislead when the native range is poorly sampled or few molecular markers are used. Here, we traced the introduction of the Asian seaweed *Gracilaria vermiculophylla* (Rhodophyta) to the United States, identifying its source and likely vector. We used a combination of population genetics and

Data Files



Download dataset

> April 24, 2018

Related Works

Article

<https://doi.org/10.1002/ece3.3001>

Metrics



330 views



78 downloads



2 citations

Keywords

Gracilaria vermiculophylla

Oysters

Protocols.io

- A secure platform for developing and sharing reproducible methods
- With the UC-wide pilot, UCI researchers have access to Premium features at no cost for private collaboration (through May 2024)
- <https://www.protocols.io>

UNIVERSITY
OF
CALIFORNIA



protocols.io

UCI Libraries Guide to Research Data Management

UCI Libraries

University of California Irvine / Research Guides / I Am Sharing My Research / Research Data Management / New 2023 NIH Data Sharing Policy

Research Data Management: New 2023 NIH Data Sharing Policy

* UC Irvine access only

URL: <https://guides.lib.uci.edu/datamanagement>

[Home](#) / [Data Management Plans](#) / [Managing Data](#) / [Sharing Data & Preservation](#) / [Data Repositories](#) / [Working with Sensitive Data](#) / [Research Computing Resources at UCI](#)

[COVID-19 Data](#) / [Tools](#) / [New 2023 NIH Data Sharing Policy](#)

Overview of 2023 NIH Data Management and Sharing Policy

The National Institutes of Health (NIH) has issued a [Final NIH Policy for Data Management and Sharing](#) to promote the sharing of scientific data. Starting on **January 25, 2023**, NIH requires researchers to submit a plan outlining how scientific data from their research will be managed and shared. The policy also establishes the expectation for maximizing the appropriate sharing of scientific data generated from NIH-funded or conducted research, with justified limitations or exceptions.

The new policy requires submission of a 2-page (max) [Data Management and Sharing Plan \(DMSP\)](#) describing the following:

1. Data type
2. Related tools, software, and/or code
3. Standards
4. Preservation, access, and associated timelines
5. Access, distribution, and reuse considerations
6. Oversight of data management and sharing

NIH has provided supplemental resources to help the research community prepare for the new policy. These include:

- [NIH Scientific Data Sharing website](#)
- [NIH FAQ about the new policy](#)
- [Examples of DMS Plans provided by NIH](#)
- [Allowable Costs for Data Sharing](#)
- [Selecting a Repository](#)


Template for NIH DMS Plan

A [template with guidance and sample answers](#) aligned with the new NIH Data Policy has been developed by the DMPTool. The template can be downloaded from Google docs and edited in Word, or accessed and completed using the [DMPTool](#) (see [quick start guide](#)).

Need more help?

[Request a consultation](#) for NIH data policy and data management-related questions, or email Data Curation Librarian Wasila Dahdul (wdahdul@uci.edu) for assistance.

Resources

 **DRYAD**

FAQ for NIH 2023 DMS Policy

https://guides.lib.uci.edu/datamanagement/NIH_2023_data_sharing_policy

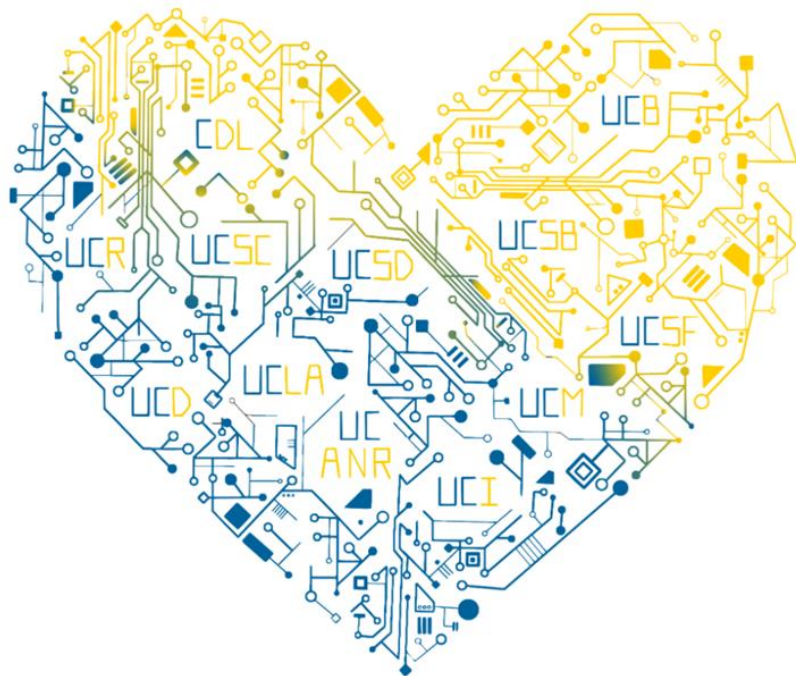
UC Love Data Week, Feb 13-17, 2023

HOW DO YOU YOUR DATA?

February 13 - 17, 2023

UC Love Data Week is a week-long offering of presentations and workshops focused on data access, management, security, sharing, and preservation. Whether you're working on qualitative or quantitative data, we've got events for you! All members of the University of California community are welcome to attend. Make sure to register with your UC-campus email.

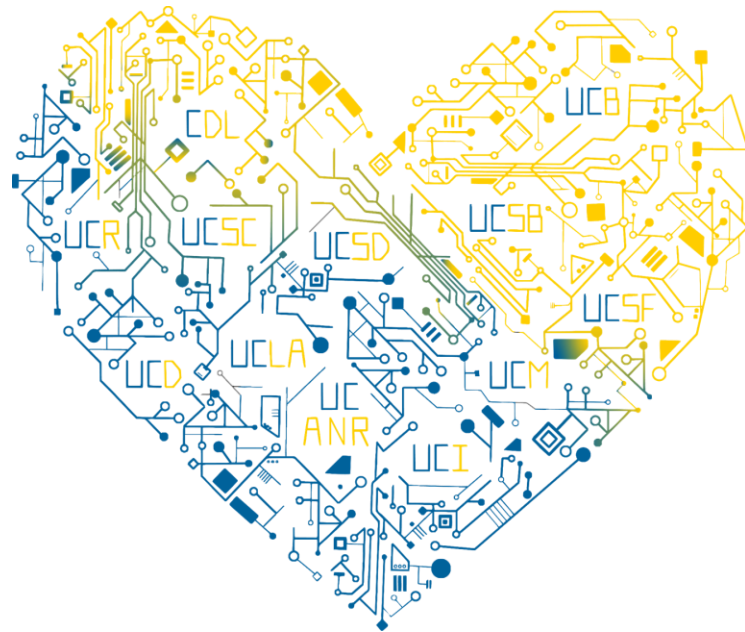
**WORKSHOP
INFORMATION**



For more information: <https://uc-love-data-week.github.io>

NIH workshops during UC Love Data Week

- Managing and Sharing Data for NIH Projects
 - Feb 13, 1:00-2:00 pm
 - [registration link](#)
- Writing a Data Management and Sharing Plan for NIH
 - Feb 13, 11:00 am - 12 pm
 - [registration link](#)



Need help?

Email Wasila Dahdul, Data Curation Librarian: wdahdul@uci.edu

Schedule an appointment: <https://spaces.lib.uci.edu/appointments/wasila>

The logo for UCI Libraries, featuring the text "UCI Libraries" in white on a blue rectangular background.

UCI Libraries



Data Sharing Plan Submissions and Budget Considerations
Maria Diaz Romero
January 10th, 2023

3 Important things to know for your proposal

1. **ALL** funded proposals that will produce scientific data are required to submit a DMS Plan, regardless of funding amount, effective 1/25/23.
2. A DMSP that complies with the new format must be submitted with your proposal.
3. A DMSP becomes a binding agreement.

Budget Considerations-Allowable Costs

Reasonable, allowable costs may be included in NIH budget requests for:

1. Curating data/Developing supporting documentation
 - Formatting data
 - De-identifying data
 - Preparing metadata
 - Formatting data for repositories
2. Preserving/sharing data through repositories
 - Data deposits fees necessary to make data available and accessible
3. Local data management considerations,
 - Activities necessary to manage /preserve data before deposited.

Budget Considerations (cont.)

- **Unallowable Costs:**
- Infrastructure costs that are included in institutional overhead (for instance, Facilities and Administrative costs)
- Costs associated with the routine conduct of research, including costs associated with collecting or gaining access to research data.

Allowable Cost Resource:

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

<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-015.html>

Final Budget Considerations (cont.)

- ALL Costs MUST be incurred during the performance period of the award.
- Assistance estimating costs for data preparation and archiving
 - UCI Libraries
 - NIH Budget Estimator Tool:
(<https://nda.nih.gov/niaaa/application.html/>)

A	B	C	D	E
	NDA Data Submission Cost Estimation Tool https://nda.nih.gov/ Need help? ndahelp@mail.nih.gov			
<i>Replace sample answers with information tailored to your research project.</i> PLEASE NOTE: This estimation is for the entire project, not per year.				

Budget and Budget Justification

- R&R/Detailed Budget - A line item in the budget form, Section F. Other Direct Costs.

7. Alterations and Renovations	
8. Data Management Sharing Plan	5,000.00

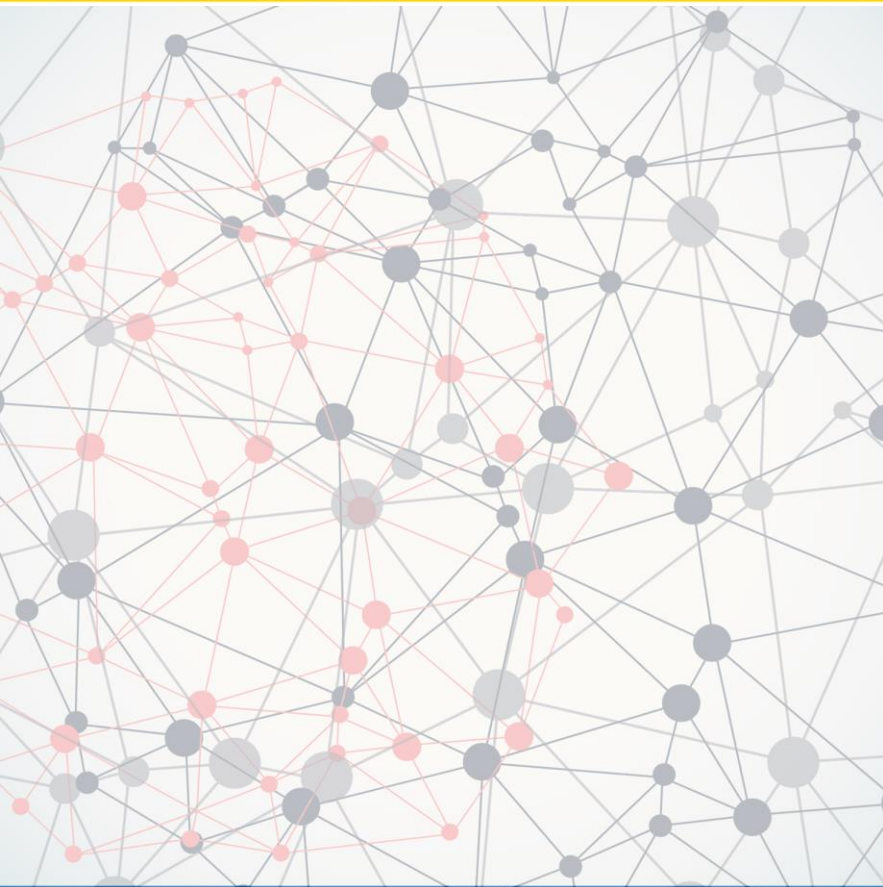
- Modular Budget – Additional Narrative Justification.

Personnel Justification		Add Attachment
Consortium Justification		Add Attachment
Additional Narrative Justification	Include DSMP Budget Justification	Add Attachment

- Include a brief summary of the DMS Plan and a description of the requested DMSP costs in the budget justification.



NIH 2023 Data Management and Sharing Policy – IRB Considerations
Anu Mathur, MS, CIP
January 10th, 2023



If the new **2023 NIH DMS Policy** is applicable to your research but are unsure whether IRB review is required, please visit the [UCI IRB / Human Research Protections \(HRP\) website](#) for extensive guidance with this determination.

- [**Do You Need IRB Review**](#) – quick links guidance document
- [**Current IRB Consent Templates**](#)
- If your research project qualifies for IRB review, please note the following Participant Consent considerations.

Considerations related to the UCI Consent Form:

- If study is collecting genomic data and is also subject to NIH GDS Policy add language from the associated GDS Consent Addendum
 - *Excellent tool to help determine which NIH Policies apply to your research: <https://sharing.nih.gov/other-sharing-policies/which-policies-apply-to-my-research>*
- **Risks Section:**
 - Consider **breach of confidentiality**: If identifiers retained, coded etc.
 - If **identifiers retained**, or there are circumstances in which **re-identification** may be possible, note associated risks.
 - If data shared with **unrestricted access**, consider associated risks.
- **Certificate of Confidentiality (CoC):**
 - A CoC is automatically issued for NIH funded Human Subject's Research.
 - Add CoC specific language from our current Consent template

Consent Form Considerations (cont.):


- **WHAT HAPPENS IF I WANT TO STOP TAKING PART IN THIS STUDY?:**
 - If identifiers will be retained for future research, add appropriate current template text.
- **HOW WILL INFORMATION ABOUT ME AND MY PARTICIPATION BE KEPT?**
 - **Subject Identifiable Data:** specify if subject identifiers will be linked to research data.
 - **Data Storage:** describe how data will be maintained in consideration of your DMS Plan.
 - **Data Retention:** If the research plan includes storing information / biospecimens indefinitely, include current required language.
- **Future Research Use:** specify that information / biospecimens will be shared with other researchers per current template.
- **Use of Biospecimens:** If collecting and sharing, add this information.
- **Genetics:** Required if the study involves genetic testing or access to genetic information

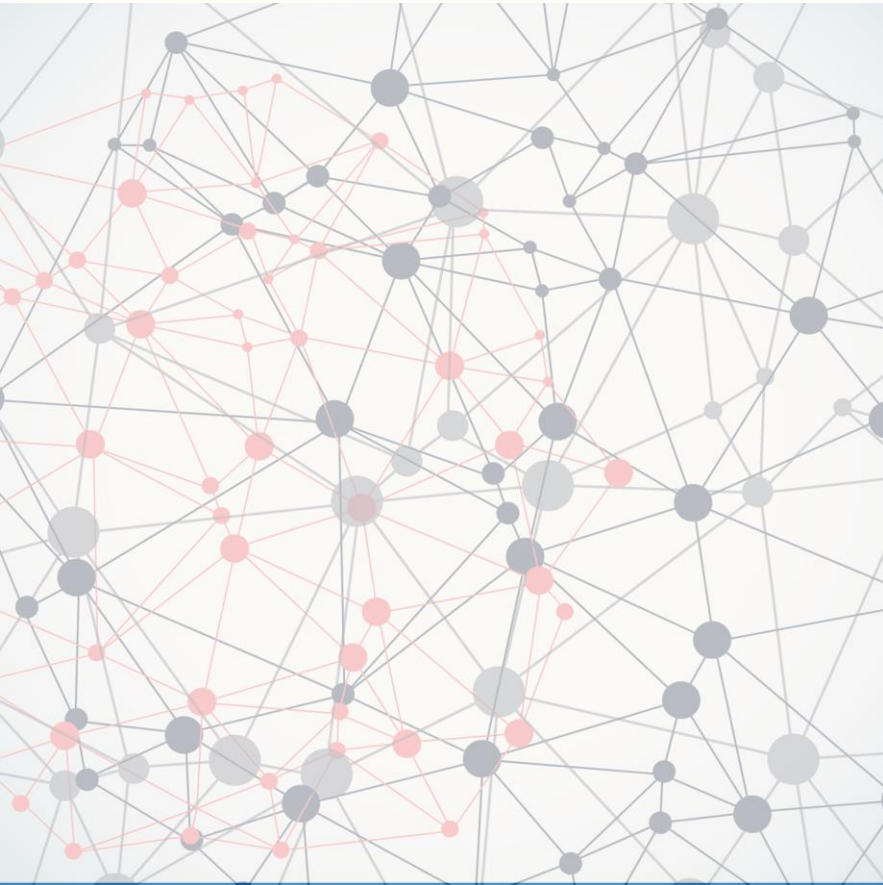
Kuali Research Protocols (KRP) is UCI's IRB, sIRB and hSCRO submission and management system.

Considerations related to the IRB submission via KRP:

- **Project Funding:**
 - Select “Grant / Contract”
 - Specify National Institutes of Health (NIH)
- **Research Procedures:**
 - Remember that all procedures used to collect research data should align with your 2023 DMS Plan.
- **Risk Assessment:**
 - Consider **breach of confidentiality**: If identifiers retained, coded etc.
 - If **identifiers retained**, or there are circumstances in which **re-identification** may be possible, note associated risks.
 - If data shared with **unrestricted access**, consider associated risks.

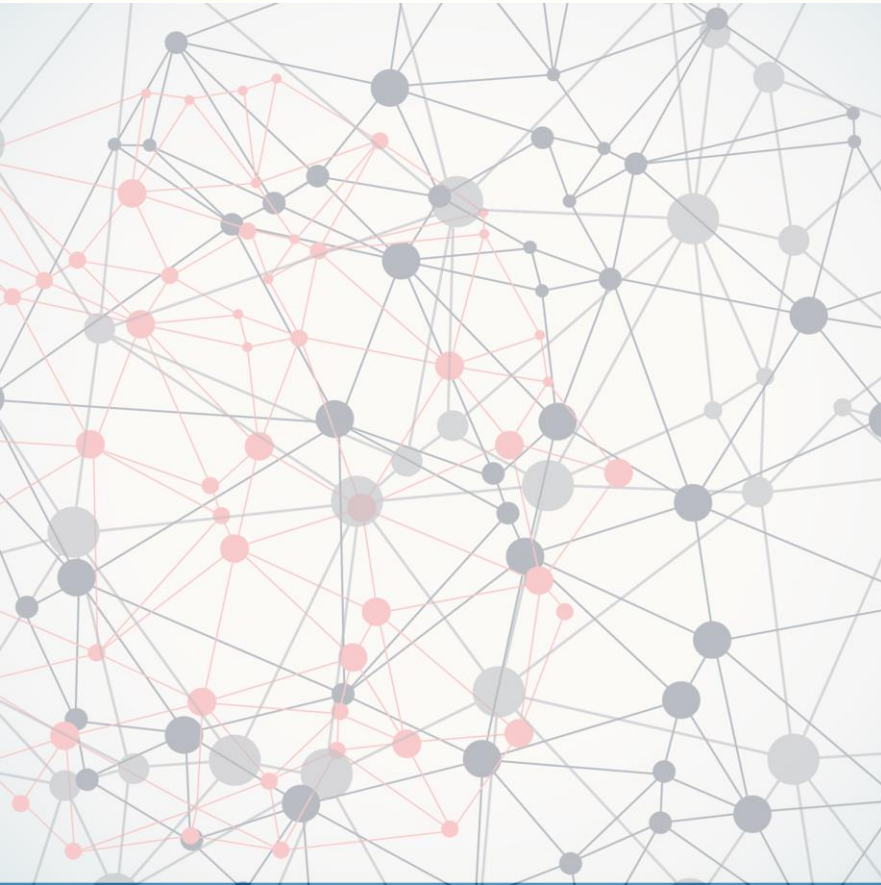
KRP submission (cont.):

- Certificate of Confidentiality (CoC):
 - Indicate situations when identifiable information may be disclosed
- Confidentiality of Data: 
 - Specify if identifiers retained, coded etc.
 - **Identifier retention:** Consider the appropriate timeframe (indefinite or specified) for data and biospecimen storage most appropriate per your DMS Plan.
 - Specify appropriate **steps to safeguard data** per your DMS Plan
 - Please review the [UCI OIT Information Security webpage](#) for further information.



PLEASE NOTE: UCI HRPP has Consent Language specifically addressing the 2023 NIH DMS Policy which should be inserted in the main Consent Form:

NEW!! [Consent Language: NIH Data Management and Sharing \(DSM\)](#)



2023 Data Management & Sharing Policy **Frequently Asked Questions (FAQs):**

<https://sharing.nih.gov/faqs#/data-management-and-sharing-policy.htm>

Questions?

Thank You!