

https://arcticdata.io

## Data Management Planning Workshop

Kathryn Meyer

& Matthew B. Jones





Data Data

NSF Award #1546024

POLAR 2018 Open Science Conference 19 June 2018



## **NSF Arctic Data Center**



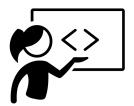




Data archive

Data discovery portal

Support services



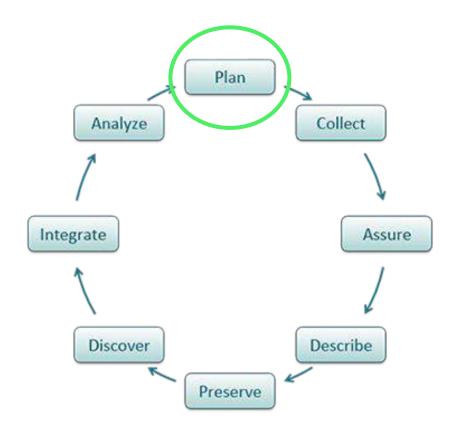
Training & Outreach



Data rescue



## **The Data Life Cycle**



3







Efficiency



Funder Requirement

Engagement



Share data



## What's in a Data Management Plan?

- Study design
- Data (including format)
- Metadata
- Policies for access, sharing & reuse
- Long-term storage & data management
- Budget



## Writing a Data Management Plan (DMP)



- 1. Engage your team
- 2. Plan from the start
- 3. Follow good advice
  - Arctic Data Center
  - Institutional Libraries
  - DataONE
- 4. Use good tools
  - DMPTool
  - DMPOnline
- 5. Review and Revise



## 10 Simple Rules for Writing a Good DMP

- 1. Determine the research sponsor requirements
- 2. Identify the data to be collected
- 3. Define how the data will be organized
- 4. Explain how the data will be documented
- 5. Describe how quality data will be assured
- 6. Present a sound storage & preservation strategy
- 7. Define the project's data policies
- 8. Describe how the data will be disseminated
- 9. Assign roles & responsibilities
- 10. Prepare a realistic budget



## **Find Research Funder Requirements**

- NSF
  - NSF 14-1, Grantee Standards, Section j (<a href="https://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/g">https://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/g</a> pg 2.jsp#IIC2j)
- NSF GEO
  - Directorate for Geosciences--Data Policies
     (<u>https://www.nsf.gov/geo/geo-data-policies/</u>)
- NSF Polar Programs
  - NSF 16-055 Dear Colleague Letter (<a href="https://www.nsf.gov/pubs/2016/nsf16055/nsf16055.jsp">https://www.nsf.gov/pubs/2016/nsf16055/nsf16055.jsp</a>)
- Check your funder for specific requirements



## **NSF Division of Polar Programs**

- Submit to Arctic Data Center within 2 years
  - AON program submit within 6 months
- Document your data well enough for reuse
- Exceptions for sensitive data
  - o i.e., social sciences, endangered species

9



## **NSF DMP Requirements: 5 Sections**

#### Products of research

 Types of data, samples, physical collections, software, curriculum materials, other materials produced during project

#### Data formats and standards

 Standards to be used for data and metadata format and content (for initial data collection, as well as subsequent storage and processing)

## Policies for access and sharing

 Provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements

### Policies and provisions for re-use

• Including re-distribution and the production of derivatives

## Archiving of data

Plans for archiving data, samples, research products and for preservation of access



## **DMP Tools & Resources Online**

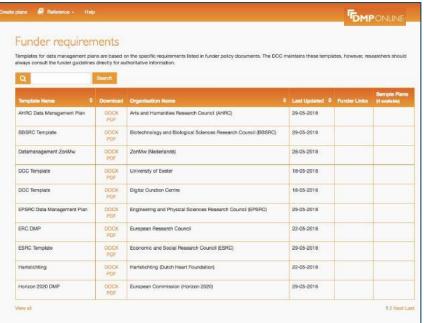


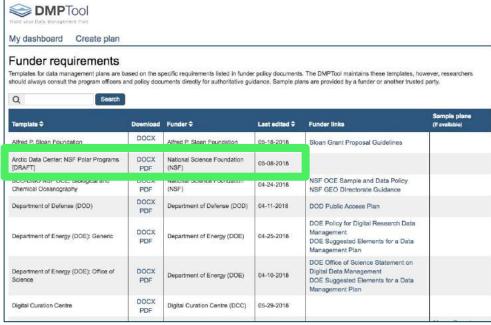


https://dmponline.dcc.ac.uk/



## **Similar DMP Resources**







## **NSF Arctic Data Center DMP Template**

#### Data and metadata formats



What format(s) will data and metadata be collected, processed, and stored in?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

Examp	le answer.
-------	------------

	_ data will be collected in	(Examples are handwritten lab notebooks, Microsof
Excel files,	CSV files, R scripts, etc.	Make sure to specify the collection format for each type of data
detailed in	your description of data.)	

All data will be transferred into the following formats for processing and storage: \_\_\_\_\_. (Examples are CSV files, NetCDF files, etc.)

Metadata will be collected in \_\_\_\_\_\_ . (Examples are handwritten lab notebooks, Microsoft Word files, etc.)

All metadata will be transformed from text into EML files by the Arctic Data Center online submission tool when submitting to the Arctic Data Center.

#### Guidance:

Arctic Data Center Data Format Policy: The Arctic Data Center primarily supports the upload of open-source, ubiquitous, and easy-to-read data formats. Examples of such formats are Comma Separated Values (CSV) files, text (TXT) files, PNG, JPEG or TIFF image files, and NetCDF files among many others. If you plan to submit to the Arctic Data Center, include your planned methods to create open-source, ubiquitous, and easy-to-read data. If you plan to work with any proprietary data formats such as Excel workbooks or MATLAB files, please include a plan to transform all data stored in these formats into an open-source format before submission to the Arctic Data Center. If you anticipate any data will not be able to be transformed into an open-source format, please provide your reasoning.



## **NSF Arctic Data Center DMP Template**

#### Data and metadata formats

What format(s) will data and metadata be collected, processed, and stored in?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.



= ;	xample answer.
	data will be collected in (Examples are handwritten lab notebooks, Microsoft Excel files, CSV files, R scripts, etc. Make sure to specify the collection format for each type of data detailed in your description of data.)
	All data will be transferred into the following formats for processing and storage: (Examples are CSV files, NetCDF files, etc.)
	Metadata will be collected in (Examples are handwritten lab notebooks, Microsoft Word files, etc.)
	All metadata will be transformed from text into EML files by the Arctic Data Center online submission tool when submitting to the Arctic Data Center.

#### Guidance:

Arctic Data Center Data Format Policy: The Arctic Data Center primarily supports the upload of open-source, ubiquitous, and easy-to-read data formats. Examples of such formats are Comma Separated Values (CSV) files, text (TXT) files, PNG, JPEG or TIFF image files, and NetCDF files among many others. If you plan to submit to the Arctic Data Center, include your planned methods to create open-source, ubiquitous, and easy-to-read data. If you plan to work with any proprietary data formats such as Excel workbooks or MATLAB files, please include a plan to transform all data stored in these formats into an open-source format before submission to the Arctic Data Center. If you anticipate any data will not be able to be transformed into an open-source format, please provide your reasoning.



## **NSF Arctic Data Center DMP Template**

#### Data and metadata formats

Dat	a and metadata formats
Wha	t format(s) will data and metadata be collected, processed, and stored in?
Note	e: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.
E.	xample answer.
	data will be collected in (Examples are handwritten lab notebooks, Microsoft Excel files, CSV files, R scripts, etc. Make sure to specify the collection format for each type of data detailed in your description of data.)
	All data will be transferred into the following formats for processing and storage: (Examples are CSV files, NetCDF files, etc.)
	Metadata will be collected in (Examples are handwritten lab notebooks, Microsoft Word files, etc.)
	All metadata will be transformed from text into EML files by the Arctic Data Center online submission tool
G	uidance:
7	Arctic Data Center Data Format Policy: The Arctic Data Center primarily supports the upload of open- source, ubiquitous, and easy-to-read data formats. Examples of such formats are Comma Separated Values (CSV) files, text (TXT) files, PNG, JPEG or TIFF image files, and NetCDF files among many



Values (CSV) files, text (TXT) files, PNG, JPEG or TIFF image files, and NetCDF files among many others. If you plan to submit to the Arctic Data Center, include your planned methods to create open-source, ubiquitous, and easy-to-read data. If you plan to work with any proprietary data formats such as Excel workbooks or MATLAB files, please include a plan to transform all data stored in these formats into an open-source format before submission to the Arctic Data Center. If you anticipate any data will not be able to be transformed into an open-source format, please provide your reasoning.



## **Hands-On: Create a DMP**

Login or create a DMPTool account (<a href="https://dmptool.org/">https://dmptool.org/</a>)

Draft your own Data Management Plan







#### **DMPTool** by the Numbers



29,887 Users



26,353 Plans



234
Participating institutions More

#### Top 5 templates

Digital Curation Centre

NSF-SBE: Social, Behavioral, Economic Sciences

NIH-GDS: Genomic Data Sharing

NIH-GEN: Generic

NEH-ODH: Office of Digital Humanities



#### Sign in options

Option 1: If your institution is affiliated with DMPTool.

Your institution

- or -

Option 2: If your institution is not affiliated with DMPTool.

**Email address** 

- or -

Option 3: If not affiliated and you need an account.

Create account with email address

About Terms of use & Privacy Privacy statement Accessibility ☑ GitHub 🗹 Contact us







#### Sign in options

Option 1: If your institution is affiliated with DMPTool.

Your institution

- or -

Option 2: If your institution is not affiliated with DMPTool.

Email address

- or -

Option 3: If not affiliated and you need an account.

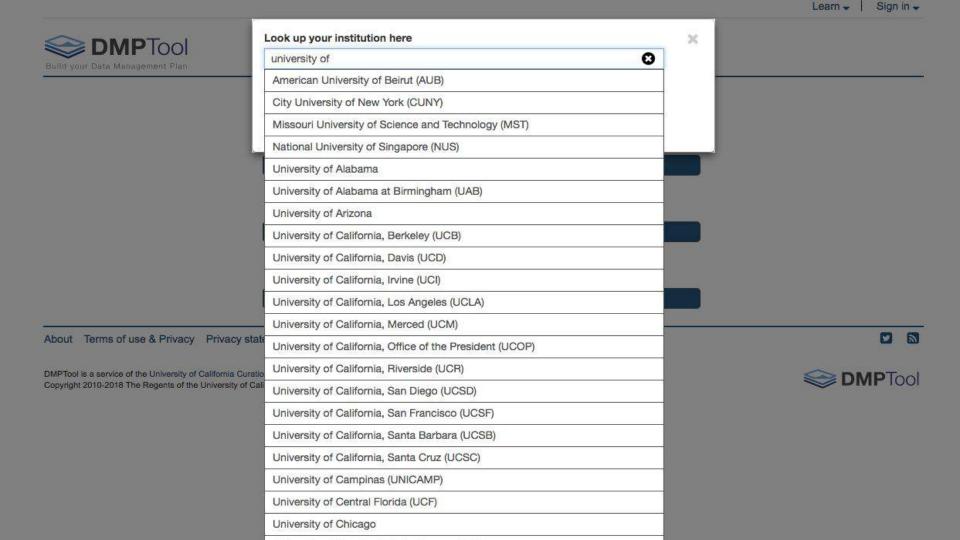
Create account with email address

About Terms of use & Privacy Privacy statement Accessibility ☑ GitHub ☑ Contact us











#### Sign in options

Option 1: If your institution is affiliated with DMPTool.

Option 2: If your institution is not affiliated with DMPTool.

Email address

- or 
Option 3: If not affiliated and you need an account.

Create account with email address

About Terms of use & Privacy Privacy statement Accessibility ☑ GitHub ☑ Contact us







Data Management Plan	* First name	* Last name	
	* Email		
	* Password		
	☐ Show password ☐ * I accept the terms and con-	ditions	
	Create account		

DMPTool is a service of the University of California Curation Center of the California Digital Library Copyright 2010-2018 The Regents of the University of California





Notice: Welcome! You have signed up successfully.

#### My dashboard

Create plan

#### Welcome

You are now ready to create your first data management plan. Click the 'Create plan' button to begin.

Copyright 2010-2018 The Regents of the University of California

There are no records associated

About Terms of use & Privacy Privacy statement Accessibility & GitHub & Contact us

DMPTool is a service of the University of California Curation Center of the California Digital Library











Funder Requirements

Public Plans

Participating institutions

FAQ

For researchers

Quick start guide

Data management general guidance

For Administrators 
Promote the DMPTool



Kathryn Meyer -

Learn -

Notice: Welcome! You have signed up successfully.

#### My dashboard

#### Welcome

You are now ready to create your first data management plan. Click the 'Create plan' button to begin.

There are no records associated

About Terms of use & Privacy Privacy statement Accessibility ☑ GitHub ☑ Contact us

DMPTool is a service of the University of California Curation Center of the California Digital Library Copyright 2010-2018 The Regents of the University of California

Create plan







#### My dashboard

#### Welcome

You are now ready to create your first data management plan. Click the 'Create plan' button to begin.



There are no records associated

About Terms of use & Privacy Privacy statement Accessibility & GitHub & Contact us





DMPTool is a service of the University of California Curation Center of the California Digital Library Copyright 2010-2018 The Regents of the University of California





#### Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

# What research project are you planning? If applying for funding, state the project title exactly as in the proposal. Select the primary research organization Begin typing to see a filtered list - or - My research organisation is not on the list or no research organisation is associated with this plan

#### Select the primary funding organization

Begin typing to see a filtered list - or - No funder associated with this plan



About Terms of use & Privacy Privacy statement Accessibility GitHub Contact us



DMPTool is a service of the University of California Curation Center of the California Digital Library Copyright 2010-2018 The Regents of the University of California



About Terms of use & Privacy Privacy statement Accessibility 2

DMPTool is a consider of the University of California Curation Contact of the California Digital Library

#### Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

#### What research project are you planning? Arctic Data Center DMP Demo Mock project for testing, practice, or educational purposes Select the primary research organization Begin typing to see a filtered list My research organisation is not on the list or no research organisation - or is associated with this plan Select the primary funding organization National Science Foundation (NSF) 0 No funder associated with this plan Which template would you like to use? We found multiple templates corresponding to your funder. Arctic Data Center: NSF Polar Programs [DRAFT] Create plan Cancel

GitHub 🖸

Contact us



* Project title				
	Center DMP Demo			
mock proje	ct for testing, practi	ce, or educatio	nal purpose	es
Funder				
National Scie	ence Foundation (N	SF)		
Grant number	r			
Project abstra	act			
B /	≣ -  ≣ -   &	<b>#</b> +		
D Z	v			
D 2	v			
<b>D</b> 1				
<b>D</b> 2				
<b>D</b> 1	,_			
5 7	v	1		
Principal In				
Principal Ir	nvestigator			
Principal Ir Name Kathryn Mey	nvestigator			
Principal Ir	nvestigator			
Principal Ir Name Kathryn Mey ORCID ID	nvestigator			
Principal Ir Name Kathryn Mey ORCID ID	n <b>vestigator</b> er			
Principal Ir Name Kathryn Mey ORCID ID	nvestigator er as.ucsb.edu			

#### Plan guidance configuration

To help you write your plan, DMPTool can show you guidance from a variety of organizations.

Select up to 6 organizations to see their guidance.

DMPTool

Find guidance from additional organizations below

See the full list

Submit



	etails		rlar	ove	rview	Write plan	Share	Downloa
* Proje	ct tit	lo.						
Arctic			ter [	OMP	Den			
□ mod	k pro	ject f	or te	esting	g, practi	ducati	onal purpose	es
Funde	3(							
Natio	nal S	cience	e Fo	ounda	ation (N	SF)		
Grant	numk	oer						
Projec	t ahe	tract						
В			_	1-	+ E	· =-		
D	1	:=	•	1=	• 0	ш,		
Princ Name	ipal	Inve	esti	igat	or			
			esti	igat	or			
Name Kathr	yn Me		esti	igat	or			
Name	yn Me		esti	igat	or			
Name Kathr ORCID	yn Me		esti	igat	or			
Name Kathr ORCID	yn Me	eyer						
Name Kathr ORCID Email meye	yn Me i <b>D</b> r@nc	eyer ceas.u	ıcsb	o.edu				
Name Kathr ORCID	yn Me i <b>D</b> r@nc	eyer ceas.u	ıcsb	o.edu				

#### Plan guidance configuration

To help you write your plan, DMPTool can show you guidance from a variety of organizations.

Select up to 6 organizations to see their guidance.

DMPTool

Find guidance from additional organizations below

See the full list

Submit

#### Arctic Data Center DMP Demo

Project details Plan overview

Write plan

Share

Download

#### Arctic Data Center: NSF Polar Programs [DRAFT]

This plan is based on the "Arctic Data Center: NSF Polar Programs [DRAFT]" template provided by National Science Foundation (NSF).

#### Instructions

Write plan

Types of data produced

- What types of data, samples, collections, software, materials, etc. will be produced during your project?
- What will be the approximate number and size of data files that will be produced during your project?
- What type of metadata (information others might need to use your data) will be collected during your project?
- Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

Data and metadata formats

- What format(s) will data and metadata be collected, processed, and stored in?
  - Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

Roles and responsibilities

- 1. What parties and individuals will be involved with data management in this project?
- 2. What will be the roles and responsibilities of each party and or individual with respect to management of the data
- 3. Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

Policies for access and sharing

 Will any of the data and/or related materials produced need provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements? If so describe them and detail any requested exceptions from the archiving requirements set for Arctic Sciences research.

## Arctic Data Center DMP Demo

Project details Plan overview

Write plan

Share

Download

#### Arctic Data Center: NSF Pola Programs [DRAFT]

This plan is based on the "Arctic Data Center: NSF A Rrograms [DRAFT]" template provided by National Science Foundation (NSF).

#### Instructions

Write plan

#### Types of data produced

- What types of data, samples, collections, software, materials, etc. will be produced during your project?
- What will be the approximate number and size of data files that will be produced during your project?
- · What type of metadata (information others might need to use your data) will be collected during your project?
- Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

#### Data and metadata formats

• What format(s) will data and metadata be collected, processed, and stored in?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

#### Roles and responsibilities

- 1. What parties and individuals will be involved with data management in this project?
  - 2. What will be the roles and responsibilities of each party and or individual with respect to management of the data
  - 3. Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?

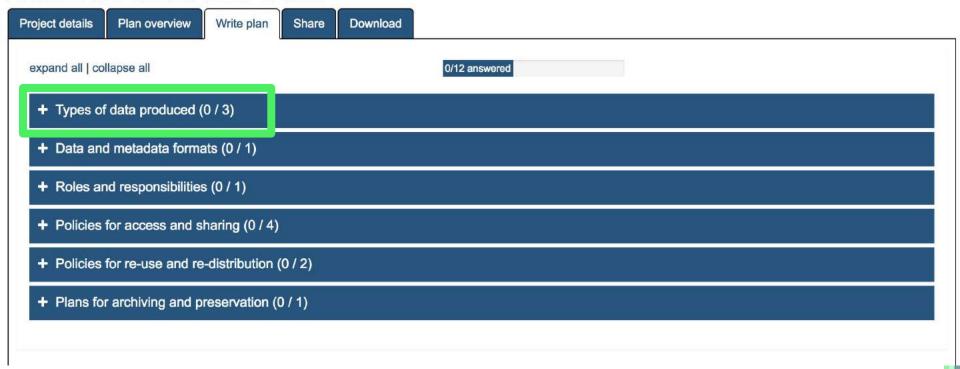
Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

#### Policies for access and sharing

 Will any of the data and/or related materials produced need provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements? If so describe them and detail any requested exceptions from the archiving requirements set for Arctic Sciences research.



#### Arctic Data Center DMP Demo



#### - Types of data produced (0 / 3)

What types of data, samples, collections, software, materials, etc. will be produced during your project?

B I 

Your answer text here



#### Save

#### NSF example answer

The researchers will collect and record \_\_\_\_\_\_. (Enter data types here. Examples are conductivity, temperature, and depth (CTD) data, gas flux data, aerial photos, modeled atmospheric data, etc.)

These data will include the variables \_\_\_\_\_\_. (Enter data variables here. Examples are water temperature, water salinity, photosynthetically active radiation, methane flux, soil albedo, etc.)

Additional data products that will be made available include \_\_\_\_\_\_. (Enter additional products here.

Examples are atmospheric model codes, educational materials, etc.)

#### - Data and metadata formats (0 / 1)

What format(s) will data and metadata be collected, processed, and stored in?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

#### Save

NSF example answer

\_\_\_\_\_ data will be collected in \_\_\_\_\_ . (Examples are handwritten lab notebooks, Microsoft Excel files, CSV files, R scripts, etc. Make sure to specify the collection format for each type of data

detailed in your description of data.)

All data will be transferred into the following formats for processing and storage: \_\_\_\_\_\_ . (Examples are CSV files, NetCDF files, etc.)

All metadata will be transformed from text into EML files by the Arctic Data Center online submission tool when submitting to the Arctic Data Center.

Guidance Comments

#### NSF DMPTool

#### Guidance

Arctic Data Center Data Format
Policy: The Arctic Data Center primarily

supports the upload of open-source, ubiquitous, and easy-to-read data formats. Examples of such formats are Comma Separated Values (CSV) files,

text (TXT) files, PNG, JPEG or TIFF image files, and NetCDF files among many others. If you plan to submit to the

Arctic Data Center, include your planned methods to create open-source, ubiquitous, and easy-to-read data. If you

plan to work with any proprietary data formats such as Excel workbooks or MATLAB files, please include a plan to

transform all data stored in these formats into an open-source format before submission to the Arctic Data Center. If

you anticipate any data will not be able to be transformed into an open-source format, please provide your reasoning.

#### Roles and responsibilities (0 / 1)

- 1. What parties and individuals will be involved with data management in this project?
- 2. What will be the roles and responsibilities of each party and or individual with respect to management of the data
- 3. Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?

Note: if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.

#### Save

#### NSF example answer

The following organizations and individuals will be involved with data management in this project:

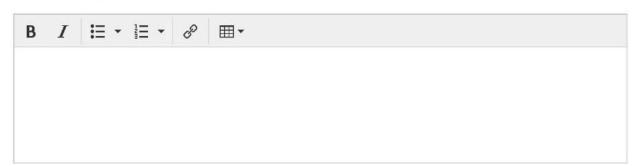
\_\_\_\_ will be responsible for \_\_\_\_\_. (Examples are collecting data, maintaining data storage and backup systems, interfacing with data repository personnel, etc. Make sure to specify the responsibilities for each organization/individual detailed above.)

The NSF Arctic Data Center will provide data archival, preservation, access and metadata authoring

Guidance Comments NSF **DMPTool** Guidance Arctic Data Center Identification Policy: The Arctic Data Center utilizes ORCiDs (https://orcid.org/) to identify individuals associated with each dataset. An ORCID will be required for the primary contact of each dataset. ORCiDs are not required for all associated parties but are encouraged so that proper identification and attribution can be given. Please plan on creating (when necessary) and recording ORCiDs for each individual involved with your project before submitting to the Arctic Data Center.

#### - Policies for access and sharing (0 / 4)

Will any of the data and/or related materials produced need provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements? If so describe them and detail any requested exceptions from the archiving requirements set for Arctic Sciences research.



#### Save

#### NSF example answer

\_\_\_\_\_ data are expected to need provisions for \_\_\_\_\_. (Examples are appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Make sure to specify all the types of data that are expected to need provisions.)

\_\_\_\_\_ data are expected to need provisions due to \_\_\_\_\_. (Examples are ethical restrictions, release of indigenous knowledge, etc. Make sure to specify explanations for all expected provisions detailed above.)

Because of these expected provisions, it is expected that \_\_\_\_\_ data will need to be exempted from the archiving requirements set for Arctic Sciences research.

Guidance Comments

## NSF DMPTool

#### Guidance

- NSF Office of Polar Programs Guidelines
- Arctic Data Center Guidelines on who must submit

#### Policies for re-use and re-distribution (0 / 2)

of data detailed in your description of data.)

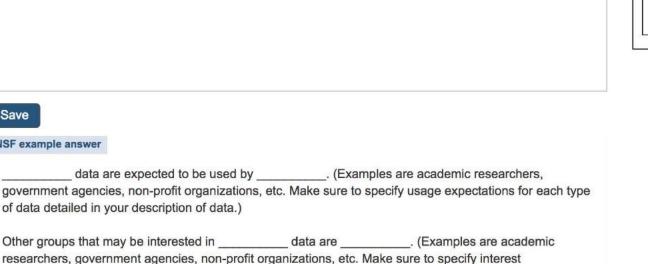
NSF example answer

How do you anticipate the data for this project will be used? Consider the following:

- 1. Which bodies/groups are likely to be interested in the data?
- 2. What and who are the intended or foreseeable uses/users of the data?

expectations for each type of data detailed in your description of data.)





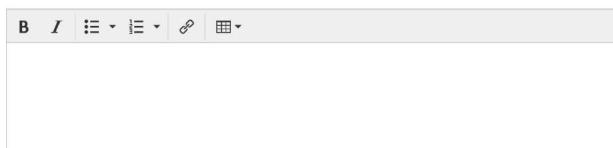


#### Plans for archiving and preservation (0 / 1)

What is the long-term strategy for maintaining, curating, and archiving the data?

**Note:** The Office of Polar Programs policy requires that metadata files, full data sets, and derived data products be deposited in a long-lived and publicly accessible archive.

**Note:** if you plan to submit data to the Arctic Data Center please refer to the guidance in the panel on the right.





NSF example answer

The data manager will follow the NSF Arctic Data Center guidelines to provide accurate and complete documentation for data preservation. The NSF Arctic Data Center will ensure that the data are curated in a relevant long-term archive and ensure data will be available after project funding has ended.





#### Guidance

Arctic Data Center Data Preservation
Policy: The Arctic Data Center ensures

the long-term preservation of the data entrusted to the repository. The guiding principles for the preservation plan follow:

- 1. Preserve the bits
- 2. Open science, open standards
- 3. Replicate data and metadata
- 4. Strong versioning
- 5. Frequent auditing
- 6. A wind down plan

#### Arctic Data Center DMP Demo

oject details	Plan overview	Write plan	Share Download	
Set plan	visibility			
Public or orga visibility by de	Charles the control of the control o	s intended for finis	hed plans. You must answer at least 50% of the questions to enable these options. Note: test plans	are set to private
Private:	visible to me, speci	fied collaborators	and administrators at my organization	
Organiza	ational: anyone at n	ny organization ca	ı view	
<ul><li>Public: a</li></ul>	nyone can view			
Email addre	60 GS		Permissions	*
meyer@ncea	as.ucsb.edu		Owner	
nvite co	llaborators			
* Email				
* Permission	ons			
○ Co-own	er: can edit project o	details, change vis	bility, and add collaborators	
Ceditor: c	an comment and m	ake changes		
<ul><li>Read on</li></ul>	ly: can view and co	mment, but not m	ike changes	
Read on		omment, but not m	ake changes	

#### Arctic Data Center DMP Demo





## **Summary**

- Good data management plans will save you time and effort overall
- Data management plans are not static revise as you do your research project
- Take advantage of DMP resources to create your plan
- The Arctic Data Center is available to assist with your DMP development



## **Arctic Data Center Support Team**

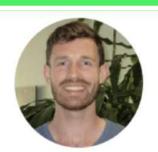
## support@arcticdata.io



Clark



Goldstein



Mullen



Chong



Meyer



Steves



Maier



Ochs



Train



Nguyen



Sun



Reevesman



Chen