2-2019

Creating Successful Data Management Plans for your Grant Proposal

Amy E. Gay
Binghamton University--SUNY, aegay@binghamton.edu

Elizabeth A. Brown
Binghamton University--SUNY, ebrown@binghamton.edu

Creating Successful Data Management Plans for Your Grant Proposals

Elizabeth Brown, Director of Assessment & Scholarly Communications

Amy Gay, Digital Scholarship Librarian

What we will talk about today

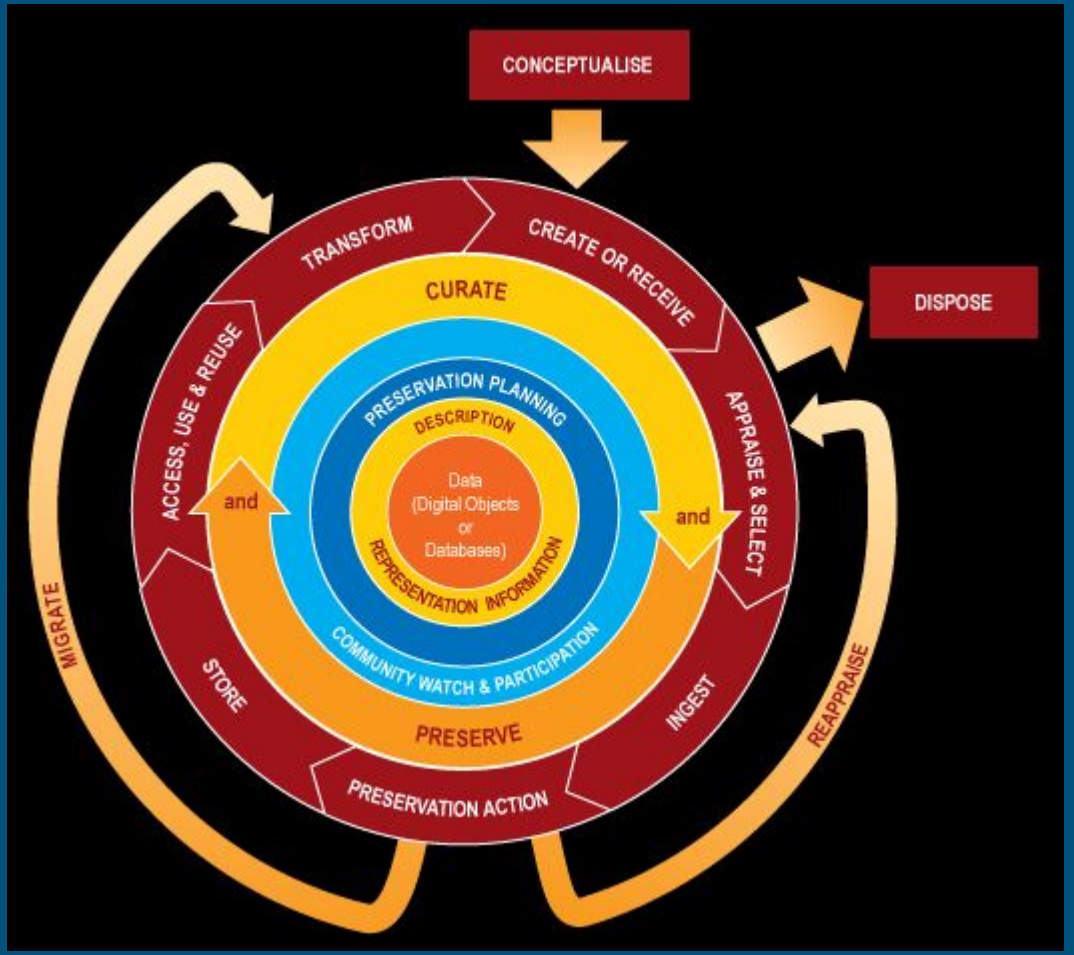
I. Brief Overview:
   A. Research data lifecycle & its components
   B. Data management plans - overview and examples
   C. Preservation needs & options
   D. Open data for reuse

II. DMPTool:
   A. Hands-on & walkthrough
   B. Best Practices & Additional Campus Resources

Research and Data Lifecycle

Concentric circles show each activity, with the outer circle showing general activities.

Source:
Digital Curation Center, UK



Some considerations for Research Data Management

1. What are your data storage needs during and after the project is completed?
2. Do you have security access and storage needs (such as HIPAA, FERPA) requirements?
3. How will ethical issues be managed?
4. How will you manage copyright and IPR (Intellectual Property Rights)?
5. How will data be shared?
6. What standards, documentation and metadata will accompany the data?
7. What resources will be needed to fulfill the data management plan?

Data Management Plans (DMPs): A Component of the Research Data Lifecycle

What is a Data Management Plan (DMP)?

- A formal plan that outlines how the data, including both physical and digital data, will be handled over the course of the research project and beyond
- A formal plan that is now required by many funding agencies
- Meant to encourage open and timely sharing of research data and ensure that data are stored in a secure and accessible manner
- Still good to have, even when not required

Capabilities of DMPs

What DMPs do:

- ➔ Help follow good data management practices
- ➔ Ensure data integrity throughout a project
- ➔ Help prevent against data loss
- ➔ Increase transparency & enable data reuse
- ➔ Another opportunity to promote your proposal

DMPs do not:

- ➔ Require that you share all data with every person who requests it
- ➔ Replace grant applications--they are just a component of proposals and awards
- ➔ Update themselves

Common Fields for DMPs [Formal Language]

Two main topics to address:

1. What data are generated by your research?
2. What is your plan for managing the data?

Other aspects to keep in mind:

- Description of data that will be produced and data that will be retained
  - Samples, geospatial coordinates, software code, databases, documentation, etc.
- Period of Data Retention
  - Address timeliness of access to data and for how long
- Data Formats and Dissemination
  - Describe data formats, media, and dissemination approaches to make data and metadata available to others (and privacy / other rights or requirements)
- Data Storage and Preservation of Access
  - Effective Physical and cyber resources for preservation and storage

Questions/Considerations when writing your DMP [Informal Language]

- Data types and formats - how much data will you create?
- Project storage, post-project data storage needs - how, where do you store data?
- Proprietary file formats - how much data has open file formats?
- Data sharing and requests for data - who, how, when others may ask for data
- Metadata schema - fields and structure of how metadata is collected
- Preservation - how long post-project will you store data? Are there sites you plan to use?

Preservation & Repositories - Institutional Options

The Open Repository @Binghamton (ORB) is an institutional repository of articles, data and other scholarly and creative works that captures (collects) the university's intellectual output.

- ★ Global visibility
- ★ Reliable digital storage and preservation for multiple file types and sizes
- ★ Works are directly discoverable through all major search engines (i.e. Google)
- ★ Not blocked by subscription walls

Rosetta is a system to manage, preserve and make accessible the massive amounts of information that is available in digital formats.

- ★ Scalable, expandable, and flexible
- ★ Searchable and discoverable through our Primo (FindIt!) library search interface
- ★ Private (dark) archive preservation storage capabilities
- ★ Reliable storage for multiple file types and sizes

Data Preservation and Access - Rosetta

Rosetta - Libraries digital preservation system

- ISO: 14721 Standard; based on NASA’s OAIS (Open Archival Information System) Model
- Metadata uses a modified Dublin Core schema
- Library of supported file formats
- Files are never deleted!

Data Preservation and Access - the ORB

ORB (Open Repository at Binghamton): orb.binghamton.edu

- Institutional Repository with journal articles, presentations, image galleries, journal hosting
- Uses a modified Dublin Core metadata schema
- Access can be controlled - files can be openly available or restricted to metadata only
- Harvestable by OAI-PMH (Open Archival Initiative-Protocol for Metadata Harvesting)

Where to share your data

Community Owned Digital Preservation Tool Registry (COPTR): Describes tools useful for long term digital preservation

Directory of Open Access Repositories (DOAR): Global directory of academic open access repositories

Open Science Framework (OSF): Great for collaborative projects, customized private and public sharing of data, accepts all types of data

SourceForge - code and programming

DataONE - earth and environmental data

ICPSR - social science datasets

NCBI (PubMedCentral) - health and medicine data

Figshare: Accepts all types of data, large amounts of data, 5GB size files

GitHub: Coding and software creation

Open Data - Where is it?

Data.gov - government-funded research

github/awesomepublicdatasets - all types of public datasets grouped by subject

Public data projects can give you insight into how to share your own data

Subject Data Repositories - where are they?

Re3data.org

Nature recommended lists

Open Access Directory

Note: Manuscript preprint servers are different from data repositories - most preprint servers DON’T accept data

Additional Resources

Ten Simple Rules for Creating a Good Data Management Plan

Data Management Plans for NEH Office of Digital Humanities

Proposals and Awards 2019

Data Management Plan Templates

My Dashboard Create plan Admin Features

Create a new plan

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

* What research project are you planning?

* Select the primary research organization

Binghamton University - or -

* Select the primary funding organization

Begin typing to see a filtered list - or -

Create plan Cancel

My Dashboard

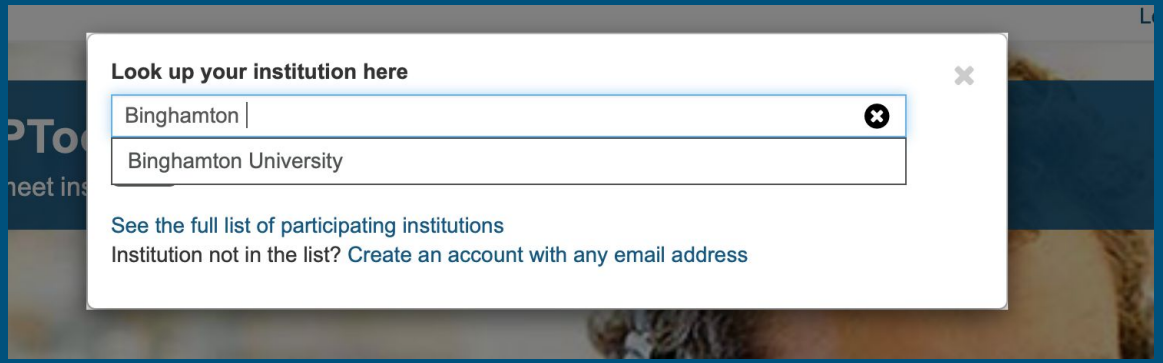
The table below lists the plans that you have created, and that have been shared with you by others. You can edit, share, download, make a cop

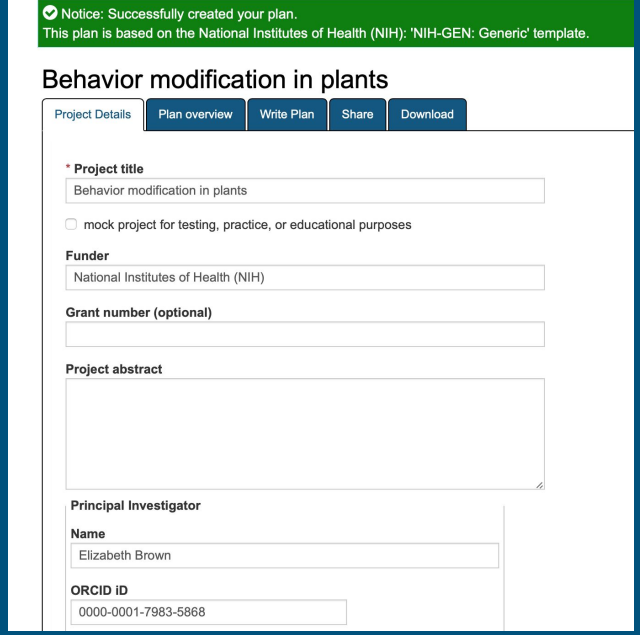
Project Title	Template	Edited	Role
sample	NSF-SBE: Social, Behavioral, Economic Sciences	10-23-2018	Owner
big star project	NSF-AST: Astronomical Sciences	10-23-2018	Owner
Sample NIH-GEN dmp	NIH-GEN: Generic	10-18-2018	Owner
test amp	NSF-GEN: Generic	10-20-2016	Owner

Why use DMPTool?

- Created to make completing a DMP easier
- Funder templates built into platform
- Examples of DMP language in system
- Save and keep track of your DMPs you create
- Make an account to keep track of DMPs you create and share
- Created by the California Digital Library (CDL)
- Administrator: Elizabeth Brown, Libraries

Create a DMP with DMPTool





My Dashboard

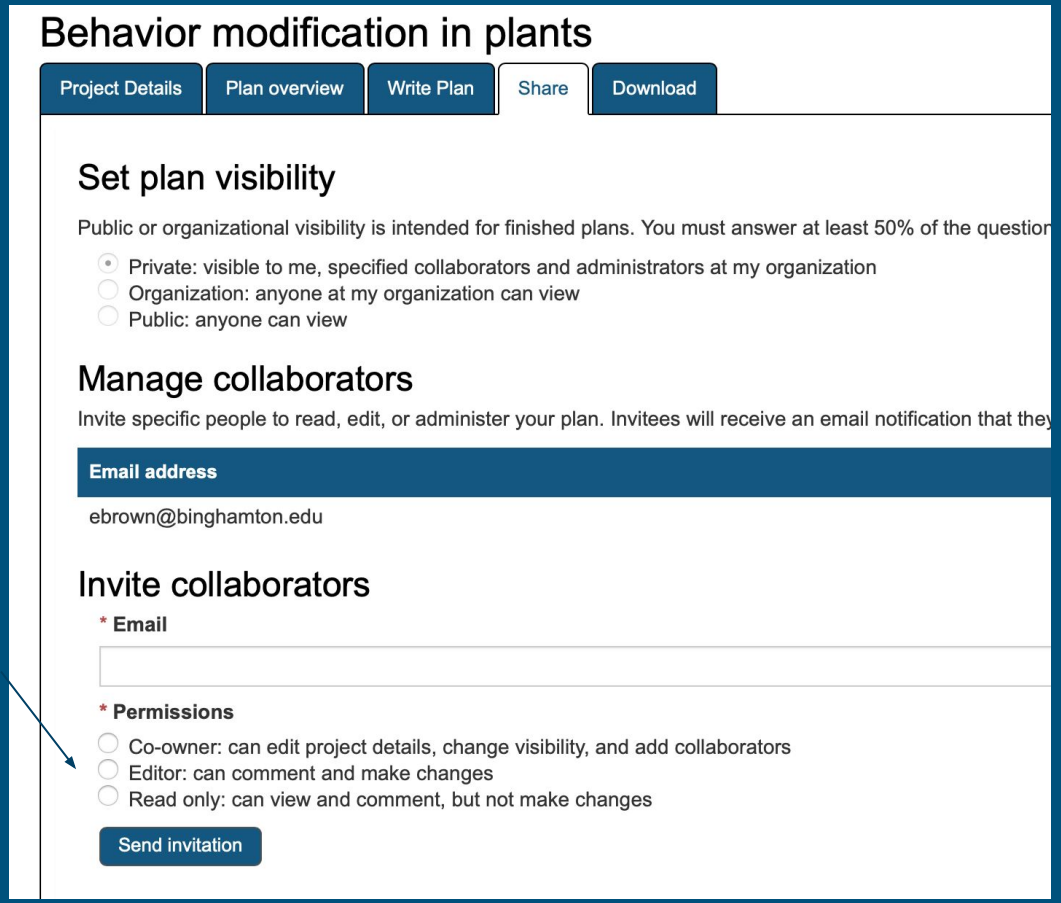
The table below lists the plans that you have created, and that have been shared with you by others. You can edit, share, download, make a cop

Project Title	Template	Edited	Role
sample	NSF-SBE: Social, Behavioral, Economic Sciences	10-23-2018	Owner
big star project	NSF-AST: Astronomical Sciences	10-23-2018	Owner
Sample NIH-GEN dmp	NIH-GEN: Generic	10-18-2018	Owner
test amp	NSF-GEN: Generic	10-20-2016	Owner

DMPTool allows you to share files

Set plan visibility

Invite collaborators to share plans



What if your funder is not in DMPTool?

- Are there links to successful program proposals? These may show DMP and data details.
- Is your funder based in the US or international? Some countries have data sharing guidelines which you can use to help write the DMP.
- Do your collaborators have DMPs they have used from this funder or program?
- DMPTool has a “no-funder”option

Reviewing DMPs - campus websites

Research Data Management Support

Created by Elizabeth (Beth) Brown (Libraries)

University Brief: NSF Data Management Plans

Created by Michael (Mike) Jacobson (Research Foundation)

Facilities and Resources Information:

Compiled by ITS Staff

When to revise a DMP

1. Your project does not receive funding, reapplying to the same proposal program
2. Staffing/students/collaborators change
3. Acquire new equipment
4. Applying to a new proposal program
5. Working with collaborators outside the US for the first time
6. Working with human subjects/human consent
7. Size and amount of project data changes substantially
8. Publication and other venues for sharing research change
9. Updated policies from funders or other regulatory requirements

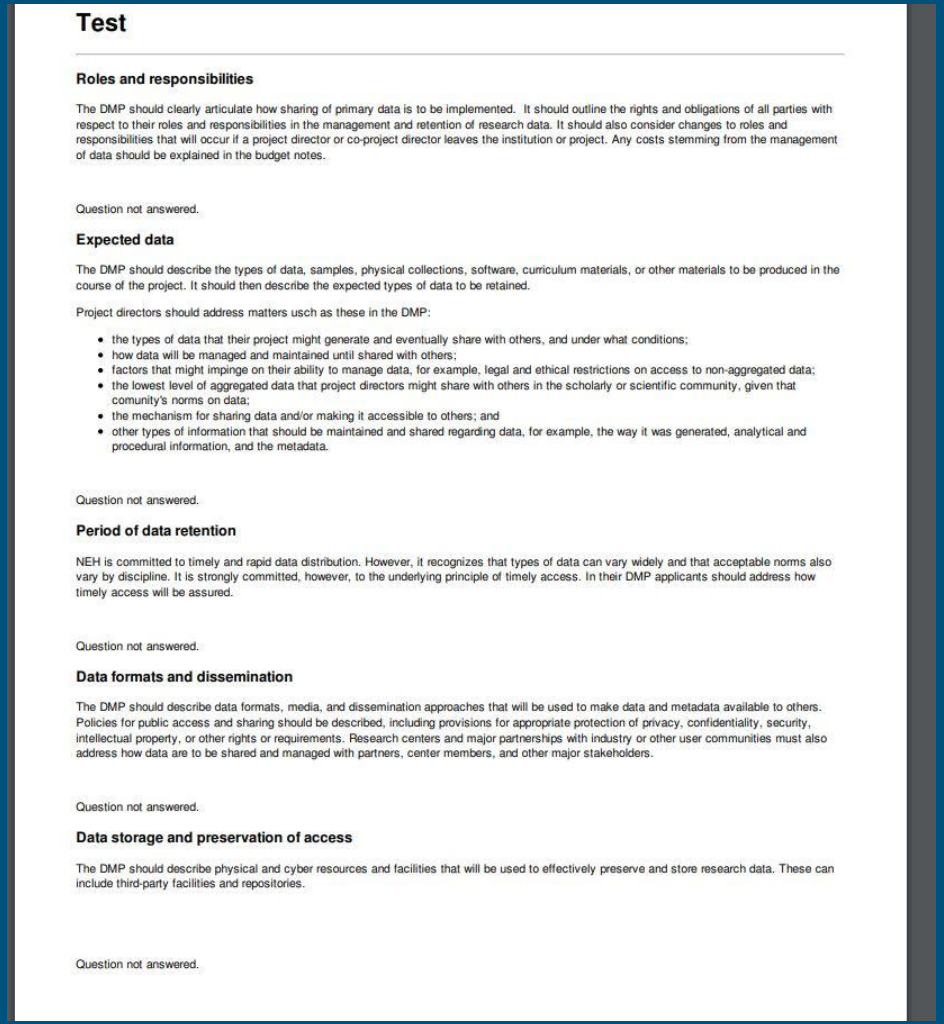
Thank you!

Elizabeth Brown
Director of Assessment and Scholarly Communications
ebrown@binghamton.edu
607-237-1937
e.brown27 (Instagram)

Amy Gay
Digital Scholarship Librarian
aegay@binghamton.edu
607-777-3495
@imabooknerd (Twitter)

DMP Example

Example from National Endowment for the Humanities (NEH) on DMPTool



Test

Roles and responsibilities

The DMP should clearly articulate how sharing of primary data is to be implemented. It should outline the rights and obligations of all parties with respect to their roles and responsibilities in the management and retention of research data. It should also consider changes to roles and responsibilities that will occur if a project director or co-project director leaves the institution or project. Any costs stemming from the management of data should be explained in the budget notes.

Question not answered.

Expected data

The DMP should describe the types of data, samples, physical collections, software, curriculum materials, or other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

Project directors should address matters usch as these in the DMP:

- the types of data that their project might generate and eventually share with others, and under what conditions;
- how data will be managed and maintained until shared with others;
- factors that might impinge on their ability to manage data, for example, legal and ethical restrictions on access to non-aggregated data;
- the lowest level of aggregated data that project directors might share with others in the scholarly or scientific community, given that comunity's norms on data;
- the mechanism for sharing data and/or making it accessible to others; and
- other types of information that should be maintained and shared regarding data, for example, the way it was generated, analytical and procedural information, and the metadata.

Question not answered.

Period of data retention

NEH is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by discipline. It is strongly committed, however, to the underlying principle of timely access. In their DMP applicants should address how timely access will be assured.

Question not answered.

Data formats and dissemination

The DMP should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

Question not answered.

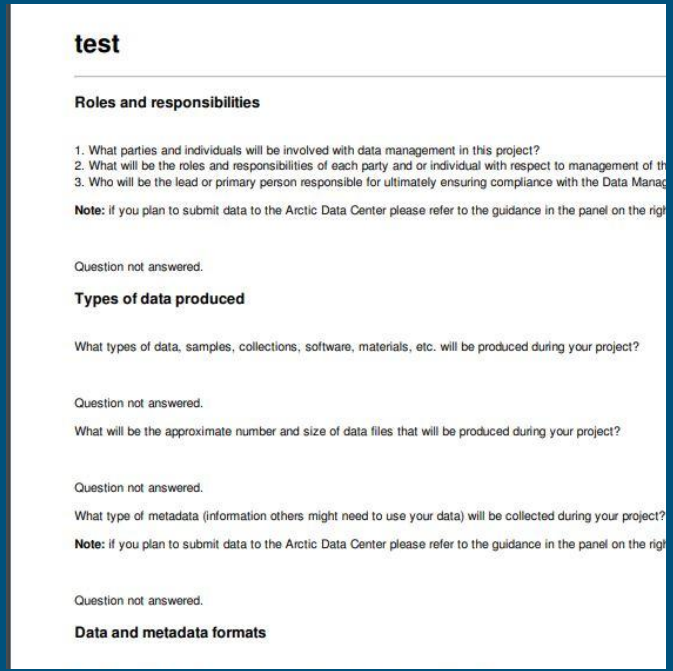
Data storage and preservation of access

The DMP should describe physical and cyber resources and facilities that will be used to effectively preserve and store research data. These can include third-party facilities and repositories.

Question not answered.

DMP Example

Example from National Science Foundation (NSF) on DMPTool



test

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 th
3. Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Manag

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

Question not answered.

Types of data produced

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

Question not answered.

What will be the approximate number and size of data files that will be produced during your project?

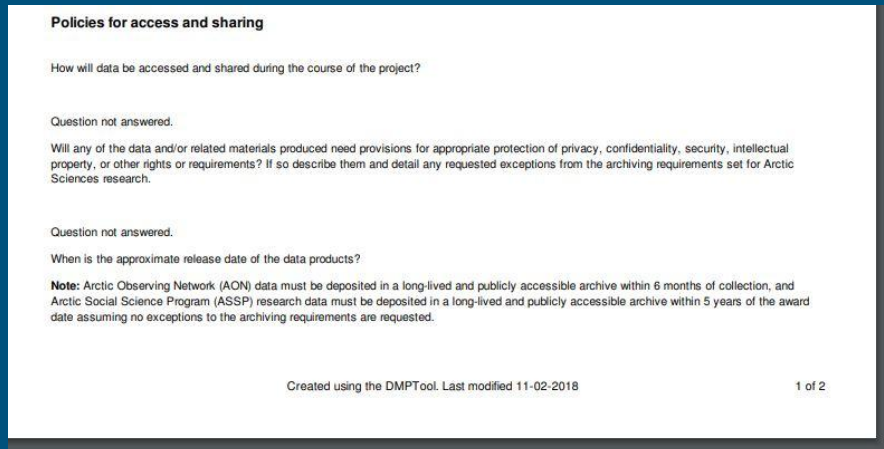
Question not answered.

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 righ

Question not answered.

Data and metadata formats



Policies for access and sharing

How will data be accessed and shared during the course of the project?

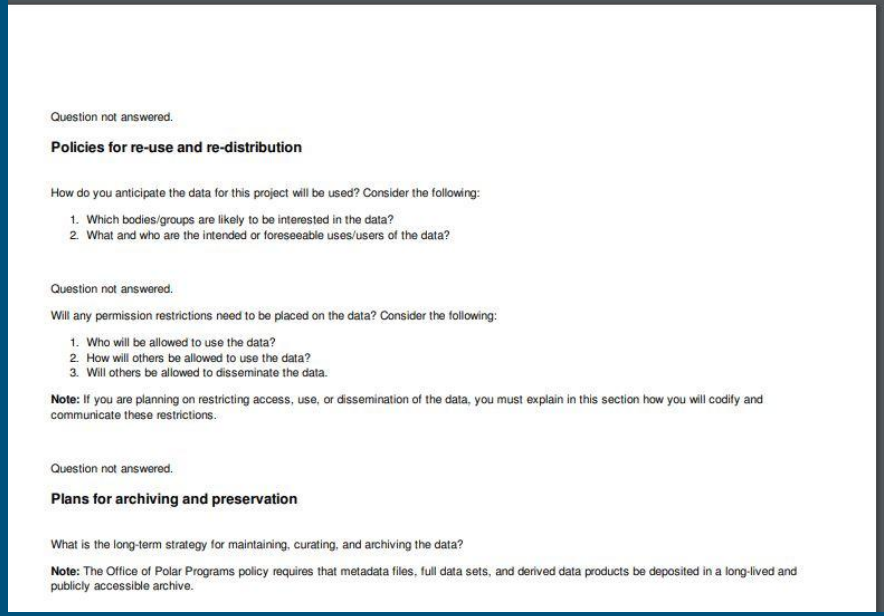
Question not answered.

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.

Question not answered.

When is the approximate release date of the data products?

Note: Arctic Observing Network (AON) data must be deposited in a long-lived and publicly accessible archive within 6 months of collection, and Arctic Social Science Program (ASSP) research data must be deposited in a long-lived and publicly accessible archive within 5 years of the award date assuming no exceptions to the archiving requirements are requested.



Question not answered.

Policies for re-use and re-distribution

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?

Question not answered.

Will any permission restrictions need to be placed on the data? Consider the following:

1. Who will be allowed to use the data?
2. How will others be allowed to use the data?
3. Will others be allowed to disseminate the data.

Note: If you are planning on restricting access, use, or dissemination of the data, you must explain in this section how you will codify and communicate these restrictions.

Question not answered.

Plans for archiving and preservation

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.