

Data Management Plan Template: History and the Humanities

Abstract

This model was developed for researchers in history and in the larger field of humanities. It was designed to take into account the fact that research projects in these disciplines still primarily use analog research data during the active phases of a project.

Two versions of the model are proposed: Phase 1 is for the documentation of DMP sections joined with a funding or grant application. The headings documented in Phase 1 are primarily aimed at producing a DMP to support research data management (RDM) budgeting for the research project. Phase 2 may be considered once funding has been secured. The entire DMP is an evolving management document since the content of certain headings will only become clearer once the project is well underway.

Administrative Details

Template Author(s): Cynthia Lisée, Benoît Kelly, Université du Québec à Montréal

Published: April 9, 2021

DOI: [10.5281/zenodo.4684738](https://doi.org/10.5281/zenodo.4684738)

Contact: Portage Network - portage@engagedri.ca, portagenetwork.ca

License: [Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)



Version:

Version	Date	Changes
1.0	2021-04-09	Formatted for inaugural publication.

Phase 1: Data Management Plan for Grant Application

“Phase 1” is for the documentation of DMP sections joined with a funding application. The headings documented in Phase 1 are primarily aimed at producing a DMP to support research data management (RDM) budgeting for the research project.

Data Collection

Describe each set of research materials using the table provided. Repeat as many times as necessary for each new set.

Data Source	(e.g. the Archives of Ontario) If the data will be produced as part of the project, indicate this.
Data Type	(e.g. images, recordings, manuscripts, word processing files)
Data Granularity	(e.g. individual item; dataset, collection, corpus)
Data Creation Methodology	(if the data are produced as part of the project) (e.g., surveys and qualitative interviews or focus groups)
Data Producer	Explain 1) who created the research data if it is not collected data, or 2) who created an additional analytical layer to existing research data. Example: In the second case, one could use a finding aid prepared by the archive or a catalog raisonné of another researcher.
Is it sensitive data?	Archival records are generally reviewed by an archivist for privacy concerns before being made available to researchers. In cases where the information will be collected directly by the principal researcher, you should avoid disclosing any information that could identify a living person such as ethnic origin, personal beliefs, personal orientation, health status, etc. without permission. For further guidance, see the Human Research Data Risk Matrix .
Analog or digital format of research data/material during the project	(e.g. print, magnetic tape, artefact, .txt, .csv, .jpeg, .nvp, etc.) Find more information on file formats: UBC Library or UK Data Service .

<p>Does the research data require long-term preservation?</p>	<p>Research material that has heritage value or value to one or more research communities or to the public interest should provide for specific actions to ensure its long-term access. If so, explain here how long-term value is characterized.</p> <p>(The <i>Preservation</i> section provides an opportunity to reflect on all of the elements to be considered for this dataset, including in particular the preservation format).</p>
<p>Will the research data be shared?</p>	<p>If not, please justify why no form of sharing is possible or desirable. Sharing research materials promotes knowledge development, collaborations and reduces duplication of research efforts.</p> <p>(The <i>Sharing and Reuse</i> section provides an opportunity to consider all of the considerations for this dataset, particularly the dissemination format).</p>

Explain how the research data will be organized to facilitate understanding of its organization.

A classification system is a useful tool, especially if you work with original manuscripts or non-digital objects (for example, manuscripts in binders). Provide the description of your classification system in this plan or provide reference to the documents containing it.

Describe how digital files will be named and how their version(s) will be controlled to facilitate understanding of this organization.

Naming conventions should be developed. Provide the description of your naming and versioning procedure in this plan or provide reference to documents containing it. Read more about file naming and version control at UK Data Service.

Documentation and Metadata

What documentation is required to correctly read and interpret the research data?

Elements to consider in contextualizing research data: methodologies, definitions of variables or analysis categories, specific classification systems, assumptions, code tree, analyses performed, terminological evolution of a concept, staff members who worked on the data and tasks performed, etc.

To ensure a verifiable historical interpretation and the lowest possible bias in the possible reuse of the data, try to identify elements of implicit knowledge or that involve direct communication with the principal investigator.

Storage and Backup

Describe the storage conditions for your research data taking into account the following aspects:

<i>Master file and backup copies</i>	Follow the 3-2-1 backup rule: keep 3 copies of files (master file + 2 copies), stored on 2 types of media (e.g. institutional server + external drive), and 1 copy kept in an off-site location. Each storage medium has advantages and disadvantages. If needed, consult a <u>resource person</u> or contact the DMP Coordinator at support@portagenetwork.ca . Find more information on storage and backup practices at <u>UK Data Service</u> .
<i>Anticipated storage space</i>	(e.g. 2 tablets; 15 files of ~70 MB =~ 1 GB multiplied in 3 copies)
<i>Anticipated storage duration</i>	(e.g. for the duration of the project; 5 years after the end of the project; long term = well beyond the end of the project)
<i>Is the access to this research data restricted?</i>	If applicable, indicate what measures are being taken to manage this access (e.g. password protection, file encryption).
<i>Who can access the data?</i>	Describe functional roles. To make sure your research data is transmitted in a secure manner or through servers governed by Canadian or provincial legislation, either contact your <u>institution's library</u> or the DMP Coordinator at support@portagenetwork.ca .

What is the total cost for storage space in the active and semi-active phase of the project?

Consider the total volume of storage space expected for each media containing these files. If applicable, include hardware costs in the funding request. Include this information in the Responsibilities and Resources section as well.

Preservation

Describe the research data that requires long-term preservation by considering the following aspects:

Preservation reason	(e.g. heritage value; value for one or multiple research communities; public interest; policy requirement)
Preservation format	See recommendations of the Library of Congress . Note that converting from one file format to another for preservation purposes may result in loss of information. This type of operation must be mentioned in the <i>Documentation and Metadata</i> section.

Where will the research data be stored at the end of the research project?

For long-term preservation, you may wish to consider CoreTrustSeal certified repositories found in this [directory](#). However, as many repositories may be in the process of being certified and are not yet listed in this directory, reviewing the retention policy of a repository of interest will increase your options. For repositories without certification, you can evaluate their quality by comparing their policies to the standards of a certification. Read more on trusted data repositories at [The University of Edinburgh](#) and [OpenAIRE](#).

To increase the visibility of research data, opt for disciplinary repositories: search by discipline in [re3data.org](#).

For solutions governed by Canadian legislation, it is possible to browse by country of origin in [re3data.org](#). In addition, Canada's digital research infrastructure offers the Federated Research Data Repository ([FRDR](#)). Finally, it is quite possible that your institution has its own research data repository.

To make sure the selected repository meets the requirements of your DMP, feel free to seek advice from a [resource person](#) or contact the DMP Coordinator at support@portagenetwork.ca.

What are the costs related to the choice of deposit location and data preparation?

Preparing research data for eventual preservation involves different tasks that may have costs that are budgeted for preferably in the funding application. This could require a cost model or simply the basic sections of the [UK Data Service Costing Tool](#). Include this cost estimate in the *Responsibilities and Resources* section.

If necessary, contact a [resource person](#) or the DMP Coordinator at support@portagenetwork.ca.

Sharing and Reuse

Describe each research dataset that will be shared with other researchers or a broader audience while taking into account the following considerations:

<i>Is it sensitive data?</i>	If so, explain if this limits access. Refer to the <i>Ethics and Legal Compliance</i> section if necessary.
<i>Is the research data subject to intellectual property?</i>	If so, explain if this limits access. Refer to the <i>Ethics and Legal Compliance</i> section if necessary.
<i>Sharing requirement</i>	It depends on whether an institutional policy or the granting agency requires some form of sharing of research materials.
<i>Target audience</i>	(e.g. history researchers, researchers from various disciplines, general public)

Describe the proposed dissemination strategy to communicate the existence of the research data to its target audiences.

To optimize the dissemination of research material, follow the FAIR principles as much as possible. The [Australian Research Data Commons](#) offers an easy-to-use tool for assessing compliance with these principles. The [Digital Curation Centre](#) provides a detailed guide to data citation (both digital and physical).

To make the material retrievable by other tools and to cite it in scholarly publications, publish a data article in an open access journal such as the [Research Data Journal for the Humanities and Social Sciences](#).

If necessary, contact a [resource person](#) or the DMP Coordinator at support@portagenetwork.ca.

Example: Reporting in a recognized data repository, attribution of a perennial identifier such as DOI (see the [FREYA project guide](#)), reporting in mailing lists and social networks.

Responsibilities and Resources

For all research data management activities, consider who is responsible (individual or organization), based on what timeframe, whether staff training is required, and whether there are costs associated with these tasks.

If the DMP is at the funding application stage, focus on cost-incurring activities in order to budget as accurately as possible research data management in the funding application.

Activities that should be documented: drafting and updating the DMP; drafting document management procedures (naming rules, backup copies, etc.); monitoring document management procedure implementation; conducting the quality assurance process; designing and updating the succession plan; drafting the documentation; assessing the retention period; assessing data management costs; managing sensitive data; managing licences and intellectual property; choosing the final data repository location; and preparing research data for the final repository.

What is an overall cost estimate for the management of research materials?

Taking into account all the aspects in the previous sections, estimate the overall cost of implementing the data management plan. Consider both the management activities required during the active phase of the project and the preservation phase. Some of these costs may be covered by funding agencies.

Ethics and Legal Compliance

For each research dataset reported as containing sensitive data, identify the security issues that need to be considered to protect the privacy and confidentiality within your team.

If applicable, retrieve written consent from an institution's ethics approval process. If the research involves human participants, verify that the content of the various sections of this DMP is consistent with the consent form signed by the participants.

Describes anticipated data sharing issues within the team, their causes, and possible measures to mitigate them. Read more about data security at [UK Data Service](#).

If research data sharing is desired and possible, what difficulties do you anticipate in dealing with the secondary use of sensitive data?

If applicable, pay close attention to the appropriateness of the content in the *Sharing and Reuse* section with the consent forms signed by participants. Anonymizing data can reassure participants while supporting the development of a data sharing culture. Learn more about anonymization: [UBC Library](#), [UK Data Service](#), or the [Portage Network](#).

Also make sure that metadata does not disclose sensitive data.

Explain how access requests to research data containing sensitive data will be managed. What are the access conditions? Who will monitor these requests? What explanations should be provided to applicants?

Are there legal and intellectual property issues that will limit the opening of data?

Describe the allocation of copyrights between members of the research team and external copyright holders (include libraries, archives and museums). Verify that the licenses to use the research materials identified in the *Sharing and Reuse* section are consistent with the description in this section.

If commercial activities are involved in your research project, there are legal aspects to consider regarding the protection of private information. Compliance with [privacy laws](#) and [intellectual property legislation](#) may impose data access restrictions. This issue can be discussed with a [resource person](#).

Phase 2: Data Management Plan for Project Development

“Phase 2” may be considered once funding has been secured. The entire DMP is an evolving management document since the content of certain headings will only become clearer once the project is well underway. The questions that appear in gray have already been asked in “Phase 1”. You can copy your answers from this phase and update them, if necessary, when the answers may have changed.

Data Collection

Describe each set of research materials using the table provided. Repeat as many times as necessary for each new set.

<i>Data Source</i>	(e.g. the Archives of Ontario) If the data will be produced as part of the project, indicate this.
<i>Data Type</i>	(e.g. images, recordings, manuscripts, word processing files)
<i>Data Granularity</i>	(e.g. individual item; dataset, collection, corpus)
<i>Data Creation Methodology</i>	(if the data are produced as part of the project) (e.g., surveys and qualitative interviews or focus groups)
<i>Data Producer</i>	Explain 1) who created the research data if it is not collected data, or 2) who created an additional analytical layer to existing research data. Example: In the second case, one could use a finding aid prepared by the archive or a catalog raisonné of another researcher.
<i>Is it sensitive data?</i>	Archival records are generally reviewed by an archivist for privacy reasons before being made available to researchers. In cases where the information will be collected directly by the principal researcher, you should avoid disclosing any information that could identify a living person such as ethnic origin, personal beliefs, personal orientation, health status, etc. without permission. For further guidance, see the Human Research Data Risk Matrix .
<i>Analog or digital format of research data/material during the project</i>	(e.g. print, magnetic tape, artefact, .txt, .csv, .jpeg, .nvp, etc.) Find more information on file formats: UBC Library or UK Data Service .

<p><i>Does the research data require long-term preservation?</i></p>	<p>Research material that has heritage value or value to one or more research communities or to the public interest should provide for specific actions to ensure its long-term access. If so, explain here how long-term value is characterized.</p> <p>(The <i>Preservation</i> section provides an opportunity to reflect on all of the elements to be considered for this dataset, including in particular the preservation format).</p>
<p><i>Will the research data be shared?</i></p>	<p>If not, please justify why no form of sharing is possible or desirable. Sharing research materials promotes knowledge development, collaborations and reduces duplication of research efforts.</p> <p>(The <i>Sharing and Reuse</i> section provides an opportunity to consider all of the considerations for this dataset, particularly the dissemination format).</p>
<p><i>Will the dataset require updates?</i></p>	<p>If so, make sure to properly and timely document this process in the <i>Documentation and Metadata</i> section.</p>

Explain how the research data will be organized to facilitate understanding of its organization.

A classification system is a useful tool, especially if you work with original manuscripts or non-digital objects (for example, manuscripts in binders). Provide the description of your classification system in this plan or provide reference to the documents containing it.

Describe how digital files will be named and how their version(s) will be controlled to facilitate understanding of this organization.

Naming conventions should be developed. Provide the description of your naming and versioning procedure in this plan or provide reference to documents containing it. Read more about file naming and version control at UK Data Service.

Describe the quality assurance process in place to ensure data quality and completeness during data operations (observation, recording, processing, analysis).

Example : field notebook, information log, committee implementation, tools such as OpenRefine or QAMyData; consistency with standards.

Documentation and Metadata

What documentation is required to correctly read and interpret the research data?

Elements to consider in contextualizing research data: methodologies, definitions of variables or analysis categories, specific classification systems, assumptions, code tree, analyses performed, terminological evolution of a concept, staff members who worked on the data and tasks performed, etc.

To ensure a verifiable historical interpretation and the lowest possible bias in the possible reuse of the data, try to identify elements of implicit knowledge or that involve direct communication with the principal investigator.

What documentation strategy will enable you to regularly document the research data throughout the project?

You may want to systematically include a documentation section in project progress reports or link quality assurance activities to the documentation. It is good practice to ensure that data management is included in the tasks of designated individuals.

If applicable, indicate the metadata schema and tools used to document research data.

A metadata schema is very useful to systematize the description of research material while making it readable by computers, thus contributing to a better dissemination of this material (e.g.: see Réseau Info-Musée [link in French] or Cataloging Cultural Objects). However, their use may result in a loss of information on provenance or contextualization, so ensure that this documentation is present elsewhere.

Resource for exploring and identifying metadata schemas that may be helpful: RDA Metadata Directory.

Resources for exploring controlled vocabularies: CIDOC/ICOM Conceptual Reference Model, Linked Open Vocabularies. If needed, contact a DMP resource person at your institution.

Storage and Backup

Describe the storage conditions for your research data taking into account the following aspects:

<i>Master file and backup copies</i>	<p>Follow the 3-2-1 backup rule: keep 3 copies of files (master file + 2 copies), stored on 2 types of media (e.g. institutional server + external drive), and 1 copy kept in an off-site location.</p> <p>Each storage medium has advantages and disadvantages. If needed, consult a resource person or contact the DMP Coordinator at support@portagenetwork.ca. Find more information on storage and backup practices at UK Data Service.</p>
<i>Anticipated storage space</i>	(e.g. 2 tablets; 15 files of ~70 MB =~ 1 GB multiplied in 3 copies)
<i>Anticipated storage duration</i>	(e.g. for the duration of the project; 5 years after the end of the project; long term = well beyond the end of the project)
<i>Is the access to this research data restricted?</i>	If applicable, indicate what measures are being taken to manage this access (e.g. password protection, file encryption).
<i>Who can access the data?</i>	<p>Describe functional roles.</p> <p>To make sure your research data is transmitted in a secure manner or through servers governed by Canadian or provincial legislation, either contact your institution's library or the DMP Coordinator at support@portagenetwork.ca.</p>

What is the total cost for storage space in the active and semi-active phase of the project?

Consider the total volume of storage space expected for each media containing these files. If applicable, include hardware costs in the funding request. Include this information in the Responsibilities and Resources section as well.

Preservation

Describe the research data that requires long-term preservation by considering the following aspects:

Preservation reason	(e.g. heritage value; value for one or multiple research communities; public interest; policy requirement)
Preservation format	See recommendations of the Library of Congress . Note that converting from one file format to another for preservation purposes may result in loss of information. This type of operation must be mentioned in the <i>Documentation and Metadata</i> section.

Where will the research data be stored at the end of the research project?

For long-term preservation, you may wish to consider CoreTrustSeal certified repositories found in this [directory](#). However, as many repositories may be in the process of being certified and are not yet listed in this directory, reviewing the retention policy of a repository of interest will increase your options. For repositories without certification, you can evaluate their quality by comparing their policies to the standards of a certification. Read more on trusted data repositories at [The University of Edinburgh](#) and [OpenAIRE](#).

To increase the visibility of research data, opt for disciplinary repositories: search by discipline in [re3data.org](#).

For solutions governed by Canadian legislation, it is possible to browse by country of origin in [re3data.org](#). In addition, Canada's digital research infrastructure offers the Federated Research Data Repository ([FRDR](#)). Finally, it is quite possible that your institution has its own research data repository.

To make sure the selected repository meets the requirements of your DMP, feel free to seek advice from a [resource person](#) or contact the DMP Coordinator at support@portagenetwork.ca.

What are the costs related to the choice of deposit location and data preparation?

Preparing research data for eventual preservation involves different tasks that may have costs that are budgeted for preferably in the funding application. This could require a cost model or simply the basic sections of the [UK Data Service Costing Tool](#). Include this cost estimate in the *Responsibilities and Resources* section.

If necessary, contact a [resource person](#) or the DMP Coordinator at support@portagenetwork.ca.

Sharing and Reuse

Describe each research dataset that will be shared with other researchers or a broader audience while taking into account the following considerations.

<i>Is it sensitive data?</i>	If so, explain if this limits access. Refer to the <i>Ethics and Legal Compliance</i> section if necessary.
<i>Is the research data subject to intellectual property?</i>	If so, explain if this limits access. Refer to the <i>Ethics and Legal Compliance</i> section if necessary.
<i>Sharing requirement</i>	It depends on whether an institutional policy or the granting agency requires some form of sharing of research materials.
<i>Target audience</i>	(e.g. history researchers, researchers from various disciplines, general public)
<i>Data processing level</i>	Describe in what format the data is shared, i.e. raw, processed, analyzed, or final, or whether only metadata can be shared. These processing level options are not mutually exclusive. <ul style="list-style-type: none"> • <i>Raw</i>: data obtained directly from the field or an interview. • <i>Processed</i>: operations performed to make data ready for analysis or to de-identify individuals. • <i>Analyzed</i>: data resulting from a qualitative or quantitative analysis following a methodology and a conceptual framework. • <i>Final</i>: research data prepared for its preservation. • <i>Metadata</i>: information describing research data.
<i>User licence</i>	The holder of the research data intellectual property should grant a licence that clarifies how the research data may be used. The most commonly used licences are <u>Creative Commons licences</u> and <u>Open Data Commons licences</u> . Please note that once a licence is granted, even if it is subsequently changed, the use of data obtained under the former licence cannot be prevented.
<i>Required software</i>	If applicable, indicate the name and version of the software required to access research data.

Describe the proposed dissemination strategy to communicate the existence of the research data to its target audiences.

To optimize the dissemination of research material, follow the FAIR principles as much as possible. The [Australian Research Data Commons](#) offers an easy-to-use tool for assessing compliance with these principles. The [Digital Curation Centre](#) provides a detailed guide to data citation (both digital and physical).

To make the material retrievable by other tools and to cite it in scholarly publications, publish a data article in an open access journal such as the [Research Data Journal for the Humanities and Social Sciences](#).

If necessary, contact a [resource person](#) or the DMP Coordinator at support@portagenetwork.ca.

Example: Reporting in a recognized data repository, attribution of a perennial identifier such as DOI (see the [FREYA project guide](#)), reporting in mailing lists and social networks.

Responsibilities and Resources

For all research data management activities, consider who is responsible (individual or organization), based on what timeframe, whether staff training is required, and whether there are costs associated with these tasks.

Activities that should be documented: drafting and updating the DMP; drafting document management procedures (naming rules, backup copies, etc.); monitoring document management procedure implementation; conducting the quality assurance process; designing and updating the succession plan; drafting the documentation; assessing the retention period; assessing data management costs; managing sensitive data; managing licences and intellectual property; choosing the final data repository location; and preparing research data for the final repository.

Describe your succession plan to deal with significant disruptions.

Describe the process to be followed, the actions to be taken, and the avenues to be considered to ensure ongoing data management if significant changes occur. Here are possible events to consider: a principal investigator is replaced, a person designated in the assignment table changes, a student who has finished their project related to research data in this DMP leaves.

What is an overall cost estimate for the management of research materials?

Taking into account all the aspects in the previous sections, estimate the overall cost of implementing the data management plan. Consider both the management activities required during the active phase of the project and the preservation phase. Some of these costs may be covered by funding agencies.

Ethics and Legal Compliance

For each research dataset reported as containing sensitive data (see Research Data Collection section), explain how this data will be safely managed to protect the privacy and confidentiality within your team.

If applicable, retrieve written consent from an institution's ethics approval process. If the research involves human participants, verify that the content of the various sections of this DMP is consistent with the consent form signed by the participants.

Describes anticipated data sharing issues within the team, their causes, and possible measures to mitigate them. Read more about data security at [UK Data Service](#).

If research data sharing is desired and possible, what strategies will you implement to address the secondary use of sensitive data?

If applicable, pay close attention to the appropriateness of the content in the Sharing and Reuse section with the consent forms signed by participants. Anonymizing data can reassure participants while supporting the development of a data sharing culture. Learn more about anonymization: [UBC Library](#), [UK Data Service](#), or the [Portage Network](#).

Also make sure that metadata does not disclose sensitive data.

Explain how access requests to research data containing sensitive data will be managed. What are the access conditions? Who will monitor these requests? What explanations should be provided to applicants?

Are there legal and intellectual property issues that will limit the opening of data?

Describe the allocation of copyrights between members of the research team and external copyright holders (include libraries, archives and museums). Verify that the licenses to use the research materials identified in the *Sharing and Reuse* section are consistent with the description in this section.

If commercial activities are involved in your research project, there are legal aspects to consider regarding the protection of private information. Compliance with [privacy laws](#) and [intellectual property legislation](#) may impose data access restrictions. This issue can be discussed with a [resource person](#).