Checklist for Data Management Plan

Content

Introduction to SND:s Checklist for Data Managing Plan

Data management

The concept "data management" implies how to handle, organize, structure and store research data throughout the research process. A good data management strategy takes into account technical, organizational, structural, legislative and sustainability aspects. It thus helps researchers to keep the data collected and/or used within their project tidy, useable and safe, while at the same time ensuring their longevity.

In order to simplify the management of the data generated by data collection and to avoid time-consuming work afterwards, it is important with a clear structure of how data is going to be managed during the research project.

Detailed and structured documentation of the data might involve some additional work, but that extra time will pay off if one would have to go back to verify analysis and results. It is also crucial for long-term storage and availability of the material.

Data Management Plan

In order to simplify the work on data management, a data management plan (DMP) can be created early in the research process. A DMP is a formal document that provides a framework for how to handle the data material during and after the research project.

The content of a data management plan is designed in accordance with the specific research project. In support of designing such a plan, there are different tools like this checklist, which is presented in detail in this document, and several international online resources that have also drawn up checklists and tools to aid the researcher the creation of a data management plan (links listed last in the document).

Why creating a data management plan

Creating a data management plan (DMP) is an efficient way for the researcher to gain control over how data is managed during the research process. With structured data material, which is also well documented, it becomes easier for others to understand the material.

By thinking early about various aspects of data management, the researcher can ensure that the material is well-managed during the research period. This, in turn, facilitates the preparation of the material for archiving, and enables further research after the research project has ended.

Usually, it is easier to document a research material if it is done in close proximity to the steps in the research process that create or change the material. It is also easier to know how research data should be documented if consulting the institution that will ultimately preserve the material. Therefore, it is a good idea to contact a data repository early in a project and discuss documentation and data management.

Internationally, many research financiers require a DMP as part of the application and decision-making process. The argument for making data available are several, not least, data produced by public funds should be used to the greatest extent possible and available to the public. Unless there are legal, ethical or commercial barriers, data should also be openly available so that research results can be verified.

Checklist for Data Management Plan

Version and date	[At some points during the research project, it might be necessary to send the DMP to funding organizations and/or other project partners outside the research group. It is advisable to add a date and, where appropriate, a version number to the DMP, as well as write down to whom the DMP was sent. Make sure to save a copy of each DMP that is sent.]

Administrative data

Some of the information below are perhaps already documented somewhere else (in the project plan, application for research funding etc.) and do not necessary need to be specified in several places. Thus, project management is easier by keeping project related information in one place. Also, if data will be deposit to a data repository it is important to remember that all relevant information about the research project and the data are deposited.

Project name	[State the name on existing or planned project. If applying for funding or an application has been sent, make sure to use the same name on both document.]	
		CHECKLIST
Project Description	[Short description of the project. For example, the nature of the project, the research questions that are addressed and the purpose for which are data being collected or generated.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
PI / Researcher (person, institution or organization)	[Person, institution or organization that is responsible for the material and the intellectual content of the project. Name, telephone number,	Relevant for the project? ☐ Yes

	email contact details as well as organization. State researcher ID if possible, e.g. ORCID (http://orcid.org).]	☐ No ☐ Don't know
Participating researcher and/or organizations	[If there are other researchers and/or organizations involved in the project, state their name, telephone number and email contact details. If possible, state their roles in the project.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Project Data Contact	[State the person(s) that can answer questions related to the research project during and after the project. Name, telephone number and email contact details.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Owner of the material	[The organization that owns the data. If there are several organizations involved in the research project, name the organization with the main responsibility, how the ownership is regulated as well as who is responsible for what.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Producer	[Name the organization who has the administrative responsibility for the project. This can be an institution at a university.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Roles	[Write down who is responsible for what within the research project. This includes naming the individuals responsible for data management in the research project as well as writing and updating the data management plan. If there are external partners, write down what they are going to do and their responsibilities.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Funder	[State research funder if relevant. Later, also state the reference number of funding that has been granted.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Related Policies	[State whether there are any relevant policies/guidelines at the university/department, preferably with information about where those documents can be found, as well as the version	Relevant for the project? ☐ Yes ☐ No

	used. Does your funder or university have a Research Data Management policy? If information later in the data management plan relates to these guidelines, a reference/link can be made to this paragraph.]	☐ Don't know

Ethics and Legal Compliance		
Ethical review	[Is there a need for an ethical review?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Privacy officer	[If the data of the research project is going to have data that includes confidential information or information that requires informed consent, is there a requirement to notify a privacy officer?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Informed consent	[Is informed consent required?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Protection of the identity of participants	[How will the identity of the participants be protected? Describe the technical and procedural protections for information and how permissions and restrictions will be enforced.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Confidential information	[Is there any confidential information within the material that requires special treatment and/or limits the access to it during/after the project? Describe the technical and procedural protections for information and how permissions and restrictions will be enforced.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Intellectual property rights/Copyright	[Are there any copyright and/or intellectual property right issues to consider? Will permission be needed for the collection of the material that is	Relevant for the project? ☐ Yes

	going to be used? Name persons or entities who will hold these rights and note any constraints. Will these rights be transferred to another organization for data archiving and distribution?]	☐ No ☐ Don't know
Agreements with other organizations	[Are written agreements with other organizations needed?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Archiving	[Public records are documents that are either created at or has arrived to a university and shall in most cases be preserved. As long as there is not a legal support for the destruction of a document the basic principle is that it is forbidden to destroy them. Are there any routines at the university regarding public records that has to be followed? Who in the research group is responsible for making sure that the public records from the project are archived?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Restrictions	[Will there be any access restrictions? Describe the access procedures for dissemination. Will access be open or granted only to specific user groups?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Embargo	[Are there any embargo periods for the material?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Data Collection		
Existing data	[Are there any relevant existing data material that can be reused in the project? If so, describe the scope, quantity and format of the material. Will any of that material be reused and if so how will it be integrated? If new data are collected even though there are data that can be reused, explain why it is important to collect new data.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know

Type of data	[Describe the type/types of data to be used in the project, and whether they should be generated/collected or if data is going to be reused. Also describe the scope, quantity and format of the material.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Data collection	[If new data are going to be collected, how will data be captured (questionnaire, interview, observation, measurements, recordings etc.), where and under which time period? Who will be responsible for data collection? What resources will be required during the collection event in terms of personnel, instruments and software?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Documentation and Metadata		
Documentation	[A lot of information is produced during a research project describing methods used, decisions made as well as procedures during the data collection. What documentation will be created during the different phases of collection and analysis (e.g. log books, variable lists, protocols)? How will the documentation be structured? Consider what information will be needed to interpret and read the data in the future.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Metadata	[Some of the documentation that are created during a project is structured information with purpose to explain, describe and localize data. This type of documentation can be described as metadata. Describe the metadata that will be provided with the collected/generated/reused data.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Metadata standard	[What metadata standard(s) will you use and why? If you don't know what standard to use ask a data repository or other appropriate organization.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Terminologies, ontologies etc.	[Terminologies and ontologies (or similar) can be useful for categorizing and describing the material. Will established terminologies or ontologies be used in the project? If not, will new ones be created and perhaps mapped towards any established?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know

Data management during the project		
Folder structure	[How will you structure and name your folders?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Organizing your data	[How will you organize data that are collected/reused? Will the data be organized in simple files or more complex databases? Describe procedures for ensuring data quality during the project.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Data protection or security policy at university	[Relate to the guidelines and policies regarding data protection or data security that exist at the university and define what it implies for the project. If there are different security levels defined, which one applies for the data of the project and what security measures must be taken to protect the data?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
File naming	[How will you structure and name your files in a systematic and consistent way? Set naming conventions that all project members should follow during the project. Good file names provide useful clues to the content, status and version of a file, uniquely identify a file and help in classifying and sorting files.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
File format	[In what formats will the data be generated, maintained and in the end archived/made available? Is it possible to choose appropriate formats from the beginning of the project? Suitable formats should be commonly used, be well documented, open technical specification and not being proprietary.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Versioning	[What conventions for versioning datafiles should be followed during the project? Who will be responsible for securing that a "masterfile" will be maintained, documented and versioned according to the project guidelines? How to separate different versions of a datafile?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know

Storage and backup	[What physical resources and facilities will be used for the preservation and storage of the research data? Describe how and where data is stored as well as the backup procedures for the data. How will the data be recovered in the event of an incident?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Budget		
Staff	[Estimate what resources that will be needed to document the data material during the project? This include costs for staff to manage, process and document the material during the project as well as prepare the material for long term preservation and possibly dissemination.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Hardware and software	[Budget for possible costs to obtaining necessary hardware and software (for example systems for backup, security and software for documentation).]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Storage	[Budget possible costs which may occur to make storage possible.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Data sharing		
Making data available	[Describe how data will be shared including when, where and whether access to the data will be open or granted only to specific user groups. Will the data be accessible via a repository, domain specific database, an institutional repository at the university or will the research group have a dedicated website that the group will create and maintain?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Contact a data repository	[If applicable: contact a data repository. Contact a data repository, preferably at an early stage, to find out what can be done with the material to make a deposition easier at a later stage. Also, make clear	Relevant for the project? ☐ Yes ☐ No ☐ Don't know

	who's responsible for the contact with the repository.]	
Limitations (legal/ethical restrictions)	[Will all data or only parts of it be published/made available? Are there any legal/ethical restrictions that prevents the publication of all the material? Will these restrictions mean that action must be taken before the material can be made available?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Limitations (hardware & software)	[Will a specific software or tool be required to use the data material? Is it possible to deposit the software and/or tool together with the data material and what extra documentation is required if that is the case? If it is not possible to deposit the software/tool, what is required to be able to use the data material?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Delays	[Is there a risk of delayed publication/making data available (all or parts of)? Explain why and what might be needed to do to avoid this?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Citation	[Are there any specific requests of how data should be cited when reused?]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Persistent identifier (PID)	[If possible, state whether or not the data material will receive a persistent identifier (PID)? If data is made available via a data repository a PID, for example a Digital Object Identifier (DOI), will in most cases be given the data material.]	Relevant for the project? ☐ Yes ☐ No ☐ Don't know
Other issues		

Examples of online-resources: checklists and tools for Data Management Plans

http://datalib.edina.ac.uk/mantra/datamanagementplans/

http://www.dcc.ac.uk/resources/data-management-plans

http://www.icpsr.umich.edu/files/deposit/dataprep.pdf

http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/elements.html

http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html