Data Management Plans

A data management plan (DMP) is a document that describes how data is handled during a research project and after the project is completed[https://en.wikipedia.org/wiki/Data_management_plan]. Data management is most effective when it is embedded into research early. Therefore, FoDaKo already offers advice before a research project starts and a good planning should be part of any data management.

Main objectives of a data management plan

Data management plans are as different as research data itself. Main objectives may be for example [partly from cessda, Research data management plan]
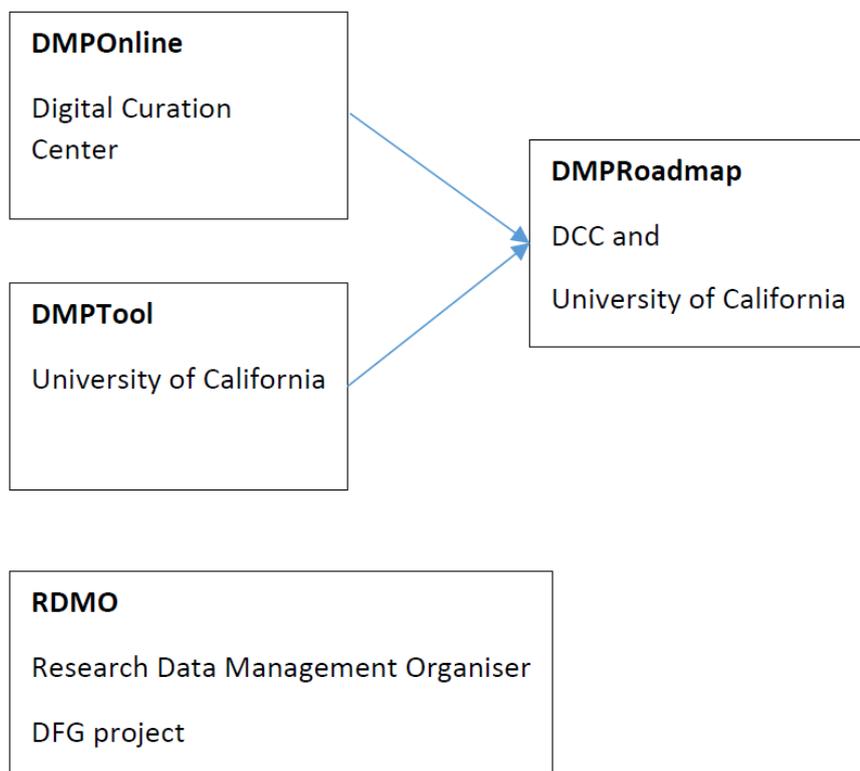
- a description of the expected data
- strategies to protect research participants and procedures to obtain informed consent
- assignment of intellectual property rights for use and conditions of re-use
- a specification of the corresponding metadata
- ordering and naming conventions
- workflows for generation, ingestion, sharing and publication of research data
- a planning of the long-term archiving, e.g. safety and security concepts for the bitstream preservation and a catalogue of jobs for the curation of data
- a specification of responsibilities

Another purpose of a data management plans that becomes more important is its use as an annex to a proposal for promotion. In the past few years, more and more institutions offering research funding expect a data management plan together with a proposal. Thereby data management plans not only facilitate handling of research data but also enhance the chance to get a promotion or is a mandatory deliverable at the beginning of the project.

Software tools

Of course, a data management plan can be written with a word processor. Nevertheless, some special software exists for that purpose too and can facilitate the creation and expansion of a data management plan. An advantage of all these tools are the questionnaires, which are a good starting point and can prevent authors of data management plans from forgetting something important.

General open-source tools are DMPonline, DMPTool and RDMO (Research Data Management Organiser). All three are hosted in the Internet, i.e only a Web browser is necessary to use them. DMPonline was developed by the British Digital Curation Center (DCC), DMPTool by the University of California Curation Center (UC3). Both tools are well established and have many users. After registration one can create data management plans, which are stored on a British or American server. DCC and UC3 have decided to cooperate and to develop a common tool, DMPRoadmap, which is still under way.

DMPOnline

Digital Curation Center

DMPTool

University of California

DMPRoadmap

DCC and

University of California

RDMO

Research Data Management Organiser

DFG project

RDMO is also still under way but may already be used. The software has been developed by a DFG project named RDMO too. The software is hosted by the AIP Potsdam but can also be installed by local providers. The developers of RDMO even want to extend their software to a real organiser as this is specified in the name. Announced features are

- metadata extraction, i.e. metadata can be extracted from the data management plan and can be filled-in into the data repository
- task definition with remainders sent by email

Already now, its verbose questionnaire may be a help.

Additionally, some scientific communities have been working on subject-specific tools. Examples are the DMPTY Wizard of CLARIN-D for the humanities and DataWiz for psychology.

RDMO in the FoDaKo project

The Universities of Düsseldorf, Siegen and Wuppertal offer secured RDMO instances for their researchers. Login to RDMO has been realized via Shibboleth:

	URL	User name	Password
Düsseldorf			
Siegen	https://rdmo.e-science-service.uni-siegen.de/		
Wuppertal	https://rdmo.zim.uni-wuppertal.de/	Your local ZIM account name	Your local ZIM password

Shibboleth Identity Provider
Bergische Universität
Wuppertal

RDMO ZIM BUW
RDMO ZIM BUW

› Haben Sie Ihr Passwort vergessen?

› Benötigen Sie Hilfe?

Benutzen Sie bitte hier Ihren ZIM-Account.
Benutzername

Passwort

☐ Löschung aller bereits freigegebenen Informationen für diesen Dienst

Login

After login Your projects list is empty, see screenshot below. You can create a data management plan by filling-in web forms; click to *Create new project* on the right.

My Projects

Name	Role	Created	

My Data

Name
Christian SchÃ¤ffer

My Groups

Options

Create new project

For the new project select questionnaire "RDMO" unless you really need another one, see screenshot below.

Create new project

Title

The title for this project.

Description

A description for this project (optional).

Catalog

RDMO

The catalog which will be used for this project.

Create project   Cancel

After creation of the new project click to *Answer questions* and fill-in the data management plan. You may interrupt your work after any *Save* or *Save and proceed* and continue later. Leaving out questions is definitely not a problem. You may also navigate with the *Overview* on the right.

Questionnaire for project *Medtest*

General / Topic

What is the main research question of the project?

Development and evaluation of an attachment piece for pin fixation with Kirschner wires with a surgery roboter. Fictitious project for the purpose of showing RDMO

Please give some keywords describing the research question.

Please enter the keywords line by line. You can add keywords using the green button and remove them using the red cross (×).

| orthopedics | × |
| mechatronics | × |

Add keyword

Back   Skip   Save   Save and proceed

If the first version of the data management plan shall serve as an annex for a research promotion proposol to DFG, you may focus on the following questions that should be answered since similar topics are mentioned in the DFG Guidelines on the Handling of Research Data. The questions are here in the same sequence as they appear in the RDMO questionnaire.

Does your institution have rules or guidelines for the handling of research data? If yes, please briefly outline them and refer to more detailed sources of information if necessary. Please also indicate, if the rules / guidelines are mandatory or optional.

This question can only be reached if you have added a project partner, see screenshot below.

Questionnaire for project *Medtest*

General / Project partners

Please fill in the form for each project partner. The different project partners will be referred to in following questions. You can add a new project partner using the green button. Once created, you can edit or delete project partners using the buttons in the top right corner.

Add project partner

Back   Skip   Save   Save and proceed

Answers to this question for the three universities Düsseldorf, Siegen and Wuppertal can be found on the FoDaKo "Grundsatzpapiere" webpage.

Are there requirements regarding the data management from other parties (e.g. the scholarly/scientific community)?

Which are these additional requirements regarding data management?

What kind of dataset is it?

Which individuals, groups or institutions could be interested in re-using this dataset? What are possible scenarios?

Which file formats are used?

Where is the dataset stored during the project?

Will this dataset be published or shared?

If no, please explain why not. Please differentiate between legal and contractual reasons and voluntary restrictions.

Questionnaire for project *Medtest*

Content classification / Data

The following questions collect information on the data that is produced or used in the project. They also help to estimate the value of the data in terms of potential re-use and long-term preservation.

Before data is newly created, it is advisable to check if there is existing data that could be re-used. This way, redundant collection or creation of research data is prevented. This saves efforts and costs. Furthermore, in the case of personal data, the German Federal Data Protection Act allows the collection of personal daa only when there are no other reasonable means to clarify the research question (re-use of existing data would be such a reasonable means). Also, there shall be collected no more information than necessary.

The information regarding the data collected, produced or used in the project is gathered along datasets. The definition of these datasets is an important conceptional decision that has to be made individually and carefully for each project.

Please fill in the form for each dataset. The different datasets will be referred to in following questions. You can add a new dataset using the green button. Once created, you can edit or delete datasets using the buttons in the top right corner.

Add dataset

Back   Skip   Save   Save and proceed

The data management plan, regardless if already complete or not, can be viewed using the link *View answers* (In the project overview on the right).

Medtest

Description

Development and evaluation of an attachment piece for pin fixation with Kirschner wires with a surgery roboter

Tasks

Task	Description	Time frame
Inform data security officer	At least one dataset contains personal data. Please contact the data security officer in charge.	

Views

View	Description	
Bielefeld	DMP template from the University of Bielefeld.	👁
CITEC DMP	DMP template from the University of Bielefeld for CITEC funded research projects.	👁
DFG (English)	Template in English language for data management plans that shall be added to a proposal for promotion to Deutsche Forschungsgemeinschaft (DFG), following http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_research_data.pdf, section 1	👁
DFG (German)	Template in German language for data management plans that shall be added to a proposal for promotion to Deutsche Forschungsgemeinschaft (DFG), following http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_research_data.pdf, section 1	👁
DMPonline template	Template from DMPonline, online https://dmponline.dcc.ac.uk	👁
DMPTool template	Template from DMPtool, based on "NSF-GEN: Generic", online: https://dmptool.org	👁
Horizon 2020 FAIR Data Management Plan template	Template for Horizon 2020, from "Guidelines on FAIR Data Management in Horizon 2020", online: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf	👁

Snapshots

Snapshot	Description	Created	
Version 1	Annex to proposal	April 10, 2018, 2:09 p.m.	👁 ✏ ↩

Members

User	E-Mail	Role	
Torsten Rathmann	rathmann@uni-wuppertal.de	Owner	
Christian SchÃ¤ffer	schaeffe@uni-wuppertal.de	Author	✏ ✖

Options

Update project information
Delete project

Answer questions

View answers

Create snapshot

Add member

Export project as XML

Back to projects overview

Versions of data management plans can be stored using the *Snapshot* function. Returning to a former version is possible. With the *Members* function a project's owner can change the role of other RDMO users. In this way, other users can be given rights and the data management team of a project be enlarged. Four different roles are possible:

Role	Description
Guest	is only allowed to read incl. export (download of the data management plan)
Author	is allowed to change the data management plan

Manager	same as author but with the additional right to store snapshots
Owner	has all rights incl. deletion of the whole project and can give rights to other users

Special output filters are in the *Views*. They adapt the output to the needs of funding institutions, e.g. DFG, or let the data management plan look like output from other data management plan tools, especially DMPonline and DMPTool. Views are only output functions; they do not change the answers in the database of RDMO. For example, the output of the View "DFG (English)" is shown in the screenshot below.

Data management plan

Data generation and ingestion

What research data will be generated or evaluated during the scientific research project?

Dataset "Patient records": Anonymised X-ray images and survey data from patients

Dataset "Animal experimentation": X-ray images from experiments with pigs before surgery, after surgery and after metal extraction

Dataset "Informed consents": Documents containing the letters of agreement signed by the patients

Dataset "Engineering data": Blueprints and specification sheets for the new attachment piece

Which file formats will be used?

Dataset "Patient records": X-ray images are in the format DICOM 3 (Digital Imaging and Communications in Medicine), survey data in PDF-A

Dataset "Animal experimentation": DICOM 3

Dataset "Informed consents": PDF-A

Dataset "Engineering data": Blueprints are in AutoCAD format DXF, specification sheets in PDF-A

Which measures of quality assurance will be taken for this dataset?

Dataset "Patient records": assuring the quality of X-ray images according to DIN 6868-157; check if the DICOM 3 and PDF-A files can be read; check if the DICOM 3 files really contain images from before and after the surgery and after metal extraction

Dataset "Animal experimentation": assuring the quality of X-ray images according to DIN 6868-157; check if the DICOM 3 files can be read; check if the DICOM 3 files really contain images from before and after the surgery and after metal extraction

Dataset "Informed consents": check if the PDF-A files can be read

Dataset "Engineering data": check if the PDF-A and DXF files can be read

Does your institution have rules or guidelines for the handling of research data? If yes, please briefly outline them and refer to more detailed sources of information if necessary. Please also indicate, if the rules / guidelines are mandatory or optional.

Project partner Universität Wuppertal: Grundsätze zum Umgang mit Forschungsdaten an der Bergischen Universität Wuppertal, Verkündungsblatt der Bergischen Universität Wuppertal, published by the rector of the University of Wuppertal on the 27.08.2015, https://bscw.uni-wuppertal.de/pub/bscw.cgi/d9959781/am15093.pdf; Handlungsempfehlungen in Ergänzung zu den Grundsätzen zum Umgang mit Forschungsdaten an der Bergischen Universität Wuppertal, Verkündungsblatt der Bergischen Universität Wuppertal, published by the rector of the University of Wuppertal on the 27.08.2015, https://bscw.uni-wuppertal.de/pub/bscw.cgi/d9959792/am15094.pdf

Project partner Universität Siegen: Forschungsdaten-Policy der Universität Siegen, passed by the rectorate of the University of Siegen on the 30.03.2017, http://www.fodako.nrw/fileadmin/redaktion/Forschungsdatenmanagement/Siegen_forschungsdaten-policy.pdf; Ordnung zur Sicherung guter wissenschaftlicher Praxis der Universität Siegen, passed on the 21.12.2015, http://www.fodako.nrw/fileadmin/redaktion/Forschungsdatenmanagement/Siegen_ordnung_zur_sicherung_guter_wissenschaftlicher_praxis.pdf

Project partner Universität Düsseldorf: Forschungsdaten-Richtlinie der Heinrich-Heine-Universität Düsseldorf, passed on the 26.11.2015, https://www.uni-duesseldorf.de/home/fileadmin/redaktion/ZUV/Justitiariat/Amtliche_Bekanntmachungen/2015/2015_12_16_AB_36.pdf

Project partner Robomeditech GmbH: None

Are there requirements regarding the data management from other parties (e.g. the scholarly/scientific community)?

No

Storing

Where is the dataset stored during the project?

Dataset "Patient records": The dataset is stored on a server of the Department of Orthopaedies, University of Düsseldorf.

Dataset "Animal experimentation": The dataset is stored on a server of the Department of Orthopaedies, University of Düsseldorf.

Dataset "Informed consents": The dataset is stored on a server of the Department of Orthopaedies, University of Düsseldorf.

Dataset "Engineering data": The dataset is stored on a server of the School of Mechanical Engineering of the University of Wuppertal.

Where will the data (including metadata, documentation and, if applicable, relevant code) be stored or archived after the end of the project?

Dataset "Patient records": ['Generic data center: Research data repository of the University of Düsseldorf']

Dataset "Animal experimentation": ['Generic data center: Research data repository of the University of Düsseldorf']

Dataset "Informed consents": ['Generic data center: Research data repository of the University of Düsseldorf']

Dataset "Engineering data":

Access

Which subset of these data could be relevant for other research contexts? Which individuals, groups or institutions could be interested in re-using this dataset? What are possible scenarios?

Dataset "Patient records": The dataset could be interesting for orthopaedic surgeons.

Dataset "Animal experimentation": The dataset could be intersting for orthopaedic and veterinary surgeons, especially specialists who deal with fractures of the extremities.

What data will be made available to other researchers for reuse and under which license?

Dataset "Patient records": license with ['Attribution (BY)']

Dataset "Animal experimentation": license with ['Attribution (BY)']

If data will not be shared, please explain why. Please differentiate between legal and contractual reasons and voluntary restrictions.

Dataset "Informed consents": The dataset contains personal data and has to be protected. Additionally, it does not contain research data but only the informed consents.

Dataset "Engineering data": The dataset contains data that shall be integrated in a patent and data which are protected by existing patents. Part of the data shall be published in the patent specification. All other ways of publication are excluded by contract with Robomeditech GmbH.

When will the data be published (if they are)?

Dataset "Patient records": Dec. 1, 2020

Dataset "Animal experimentation": Oct. 1, 2020

Third-party rights

Which datasets contain personal data? Will the data be anonymised or pseudonymised?

Dataset "Patient records", anonymisation/pseudonymisation: Yes, before / at the beginning of the data analysis

Dataset "Informed consents", anonymisation/pseudonymisation: No

Does the project use and/or produce data that is protected by intellectual or industrial property rights?

yes

Does copyright law apply to this dataset? Which kind of work is protected?

Dataset "Patient records": ['work of literature, scholarship or the arts (acc. to § 2 UrhG)']

Dataset "Animal experimentation": ['work of literature, scholarship or the arts (acc. to § 2 UrhG)']

Dataset "Informed consents": ['work of literature, scholarship or the arts (acc. to § 2 UrhG)']

Dataset "Engineering data": ['work of literature, scholarship or the arts (acc. to § 2 UrhG)']

Do other intellectual property rights apply to this dataset? Which kind of property rights?

Dataset "Patient records":

Dataset "Animal experimentation":

Dataset "Informed consents":

Dataset "Engineering data": ['patent', 'registered design']

Several file formats for export are offered.

Für Daten aus der Sprachwissenschaft, Computerlinguistik, etc. / DMPTY - ggf. für das DMW-Projekt (Sprachatlas, Universität Siegen) interessant? (Information von T. Trippel, Universität Konstanz)

Der **CLARIN-D** Datenmanagementplan wurde auf der Grundlage von DMPOnline und vor dem Hintergrund der Beratungspraxis entwickelt. Es stellte sich heraus, das gerade der Bereich der Nutzungsrechte bei der Nachnutzung vorhandener Daten komplex ist. Ein wesentliches Ziel von DMPTY besteht daher auch darin, die Hoster und die Datenersteller ins Gespräch miteinander zu bringen und durch den Datenmanagementplan nicht nur Geldgebende davon zu überzeugen, dass man weiß, was man tut, sondern bei Datenerstellenden ein Bewusstsein für Arbeitsabläufe und notwendige kontinuierliche Aktivitäten für das Datenmanagement zu schaffen, sowie im Dialog mit den Hostern bereits im laufenden Projekt die Datenübernahme vorzubereiten.

CLARIN hat dabei einen Schwerpunkt auf allen Geisteswissenschaften, die irgendwie mit Sprache arbeiten (Historiker und Literaturwissenschaftler, die Philologien, Sprachwissenschaften, etc.) Bei Umfragen, Erhebungen, Video- und Audiointerviews kommen zudem noch Fragen der Persönlichkeitsrechte hinzu. (Pseudo-)Anonmyisierung ist in dem Fall auch ein Thema, aber auch Zustimmung der Informanten zur (Nach-)Nutzung der Daten. Weitere Informationen gibt es unter https://www.clarin-d.net/de/schulungen-und-support/rechtliche-fragestellungen

Es gab Kontakte zu RDMO. Nach den CLARIN-D Erfahrung muss das Angebot gerade im Bereich der Geisteswissenschaften so niederschwellig wie irgend möglich sein, mit dem Ziel, den Informationsfluss zwischen den Datenhaltungsexpertinnen und den Datenbereitstellenden zu ermöglichen. Dann erübrigen sich häufig komplexe Fragen oder man entwickelt die individuell passenden Lösungen zusammen, wo es keine Standardprozeduren gibt.

Für Forschungsdaten aus dem Bereich der **Psychologie** wird **DataWiz** weiterentwickelt. Ein Assistenzsystem zur forschungsbegleitenden Unterstützung des Datenmanagements. Mehr Info hier: http://psydok.psycharchives.de/jspui/bitstream/20.500.11780/3727/1/GEBF_DataWiz.pdf