A structured approach to research data management - use RDMO to create your own data management plan (DMP)

Nina Dworschak
Forschungsdatenreferentin der Goethe-Universität Frankfurt

Gerald Jagusch
Head of Research Data Services
Universitäts- und Landesbibliothek Darmstadt
Outline

1. Introduction to Research Data Management & Data Management Plans
2. Research Data Management Organizer (RDMO): How to start?
3. Hands-on-experience I: Create your own DMP
4. Introduction to Metadata & Archiving: What you need to know!
5. Hands-on-experience II: Finish your DMP
6. Alternatives for RDMO
7. Further information
Short introduction to Research Data Management (RDM) in connection with Data Management Plans (DMP)
What is research data?

Research data is ....

“all data that is generated, processed or used in the course of a scientific process."

"Data collected, observed, simulated, derived or generated during research."

“also includes the necessary software and documentation for traceability."
How do research sponsors define the term research data?

- **EU Horizon 2020:**

  “Research data is information (particularly facts or numbers) collected to be examined and considered, and to serve as a basis for reasoning, discussion or calculation.”

- [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm)
"...covers all areas of data management, in particular the planning of a data collection, data generation and processing, data integrity, documentation and retention, and data sharing."

Go Fair – Global Open FAIR Initiative
A data management plan refers to all stages of the data lifecycle.
What is a data management plan (DMP)?

A DMP...

• helps to administer research data.
• determines steps in research data handling during a research project.
• concerns generation, storage, support, publication etc. of research data.
• is a dynamic document, that has to be updated in the course of the project.
The Research Data Management Organiser (RDMO) enables institutions as well as researchers to plan and carry out their management of research data. RDMO can assemble all relevant planning information and data management tasks across the whole life cycle of the research data.
Example: Goethe-RDMO

RDMO
A tool to support the planning, implementation, and organisation of research data management.

Welcome to the RDMO of Goethe University
The Research Data Management Organizer (RDMO) is a tool for the structured planning and organization of research data management of projects. The Goethe University supports its scientists with the provided web application in the creation and maintenance of data management plans. The goal of RDMO is to plan, control and document the handling of data in your scientific project in a structured way. In addition, the collected information can be downloaded as a written document in the form of a report or data management plan. This is supposed to simplify the application process for research funding organizations such as the EU, DFG and BMBF through RDMO.

The RDMO project is funded by the DFG (Deutsche Forschungsgemeinschaft). Find out more about the project and its development goals at rdmorganiser.github.io.

Login
Username
Password
Login
Research data lifecycle: sub-steps

Creating data:

• design research
• plan data management (formats, storage etc.) → DMP
• plan consent for sharing
• locate existing data
• collect data (experiment, observe, measure, simulate)
• capture and create metadata
• take care about data privacy protection/legal situation
Processing data:

- enter data, digitise, transcribe, translate
- check, validate, clean data
- anonymise data where necessary
- describe data (metadata)
- manage/organise and store data
- care about versioning
Research data lifecycle: sub-steps

Analysing data:

• interpret data

• derive data

• care for plausibility and reproducibility of datasets

• produce research outputs/author publications

• prepare data for preservation
Research data lifecycle: sub-steps

Preserving data

- migrate data to best format
- migrate data to suitable medium
- back-up and store data
- add metadata (better while creation) and documentation
- clarify the financing (better during project planning)
- give the data to data preservation specialists, e.g. data repositories/archives
Giving access to data:

- distribute data
- share data/ e.g. via repositories
- control access
- establish copyright/licenses
- promote data
Research data lifecycle: sub-steps

Reusing data:

- follow-up research
- new research
- undertake research reviews
- standardize findings
- teach and learn
Create your own DMP

1. Enter tudmo.ulb.tu-darmstadt.de or rdmo.server.uni-frankfurt.de,
2. create a new DMP project for your research,
3. answer the questions in the parts „General Information“ and „Digital Data“.
Umfrage 1

Konnten Sie für Ihre Forschungsdaten passende Antworten auswählen?

A: Ja
B: Nein
C: Nur zum Teil
D: Ich war mir unsicher
Metadata: necessity of good documentation

As time goes by, information is forgotten.

Overview of versions and different formats is lost.

Documentation is the essential element of good data management and facilitates data exchange.

Despite initial additional effort, future work is made easier.

---

Metadata: Metadata vs. Primary Data

<table>
<thead>
<tr>
<th>sample</th>
<th>temperature</th>
<th>conductivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbC2</td>
<td>120</td>
<td>40</td>
</tr>
<tr>
<td>DDc1</td>
<td>56</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Problem:

Metadata explains data, but:

Is metadata itself self-explanatory?
Metadata

Possible use-cases for metadata

1. Internal use
   • users are familiar with data

2. External use
   • published data, users unfamiliar with data

3. Automated use
   • computers have very limited capacity for interpretation
Tips for writing quality metadata

- Do not use jargon
- Define technical terms and acronyms
- Clearly state data limitations
- Use Standards whenever possible, e.g.:
  - ISO 639  Code for the representation of the names of languages.
    e.g. eng – English; ger/deu – German; fre/fra - French
  - ISO 8601  Codes for the representation of date and time.
    e.g. 2014-06-19T13:15:30Z; 2014-06-19T13:16:30+01:00
- Always remember: a computer will read your metadata
- Do not use symbols that could be misinterpreted (! @ # % { } | / \ < > ~)
- When copying and pasting from other sources, use a proper text editor (e.g., vim, Notepad) to eliminate hidden characters
Archiving & Publishing: How to select a suitable repository

http://openarchaeologydata.metajnl.com/about
What are repositories?

Repositories are databases, in which you can **archive**, **document** and **publish** your research data.

Types of repositories:

- subject-specific
- generic
- institutional
- media-specific (texts, certain data formats, films, etc.)
Reasons for using a repository

A repository makes it possible to...

• archive data securely
• manage your data and metadata
• share data (e.g. with your team)
• publish data
• search for data yourself
  • keep an overview of the state of research
  • include data of other authors in your own work
Archiving & Publishing: How to select a suitable repository

Selection Criteria:

- visibility in your community
- long-term availability
- accessibility
- persistent identifiers (e.g. DOI, URN)
- licenses offered
- review process

Subject-specific or institutional repository?
Archiving & Publishing: re3data.org

How to find a suitable repository to...

... archive or publish your own data?

... look for interesting data of others?

Registry of research data Repositories = Re³Data

Search filter (selection):

- scientific discipline
- territories
- data types
- authorizations
- identifiers
- terms of use
- access to data
Who owns data and who may use it? Issues and Suggestions

Issues:

• **No default** legal positions on how data may be used in any given context
• **Different jurisdictions**
  • apply **different standards** of creativity, originality, labor
  • allow for **different reusage** when using copyright affected material (especially for academic purposes!)
• **Different aspects of a dataset** may be treated quite differently, e.g.
  • data values themselves
  • field names / structure and data model
  • visualizations and reports derived from the data

Practical Suggestions for action:

• Take advantage of **advice** at an **early stage**
• Making **contractual agreements** early on
• **Involve all** potential rights holders and **balance interests**
• Do not release data without making **clear terms of use**
Licensing

Use **standardized licenses** and attach it to the data when publishing

Only the actual right holders can **grant licenses**

4 conditions commonly found in open licenses:

- **Attribution (BY):** Licensor must be given due credit for the work when it is distributed, displayed, performed, or used to derive a new work.

- **Share alike / Copyleft (SA):** Any new works derived from the licensed one must be released under the same license, and only that license.

- **Non-commerciality (NC):** The licensee is prevented from exploiting the work commercially.

- **No derivatives (ND):** The work may be edited / changed but the edited version may not be distributed.
Finish your own DMP

1. Reenter [tudmo.ulb.tu-darmstadt.de](http://tudmo.ulb.tu-darmstadt.de) or [rdmo.server.uni-frankfurt.de](http://rdmo.server.uni-frankfurt.de),

2. answer the questions in parts „Metadata and documentation“ and „Storage, archiving and publication“ for your project.
Umfrage 2

**Frage 1**
- Sind die Fragen für Ihre Zwecke vollständig?
  A Ja
  B Nein

**Frage 2**
- Wer hat gegenwärtig Zugang zu Ihren Forschungsdaten?
  A Nur ich
  B Meine Forschergruppe und ich
  C Ich publiziere bereits alle Forschungsdaten
  D Das weiß ich nicht

**Frage 3**
- Für welche Art von Repositorium würden Sie sich entscheiden?
  A Institutionell
  B Allgemein (z.B. zenodo)
  C Fachspezifisch
  D Ich bin unentschlossen
RDMO and other webtools

Research Data Management Organiser

Zur Planung, Umsetzung und Verwaltung des Forschungsdatenmanagements Pilotphase
GFBio Data Management Plan Tool

💡 BASIC QUESTIONS ABOUT DATA MANAGEMENT IN ONE PLACE

🗂 DYNAMIC DATA MANAGEMENT PLAN CREATION

😊 PERSONAL SUPPORT

Start DMPT
Data Management Plan Tool

Get a free GFBio account to save your dynamic DMP.

Features and Services
DMP CLARIN-D – browser based approach for DMPs

https://www.clarin-d.net/de/aufbereiten/datenmanagementplan-entwickeln
### Alternatives for RDMO?

<table>
<thead>
<tr>
<th>Name</th>
<th>Betreiber</th>
<th>Sprache</th>
<th>Templates</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARGOS</td>
<td>OpenAIRE</td>
<td>Englisch</td>
<td>• EU Horizon 2020</td>
</tr>
<tr>
<td>DMPonline (DMP Roadmap)</td>
<td>Digital Curation Center</td>
<td>Englisch</td>
<td>• EU Horizon 2020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• wichtigste UK-Förderorganisationen</td>
</tr>
<tr>
<td>DMPTool (DMP Roadmap)</td>
<td>University of California Curation Centre</td>
<td>Englisch</td>
<td>• wichtigste US-Förderorganisationen</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• DCC Template</td>
</tr>
<tr>
<td>Data Stewardship Wizard</td>
<td>ELIXIR CZ/ ELIXIR NL</td>
<td>Englisch</td>
<td>• FAIR Open Science</td>
</tr>
</tbody>
</table>
You team of experts research data at JGU:

Members:
1. Dr. Anne Vieten (Department for Research and Technology Transfer; team coordination and first contact person)
2. Dr. Elisabeth Klein (University library, currently: digital teaching)
3. Karin Eckert (University library)
4. Dr. Jörg Steinkamp (Data centre)
5. Henrike Backhaus (mainzed; provisional)

Website: https://www.forschungsdaten.uni-mainz.de/
Contact: forschungsdaten@uni-mainz.de
RESEARCH DATA MANAGEMENT - JGU

Services:
Research data management (RDM)
• Introduction to RDM
• Requirements of funding programmes for RDM and Open Data
• Creating data management plans
• Creating project-specific RDM policies

Data publication and digital preservation
• JGU and external repositories (open data)
• Data publication (open data)
• Metadata standards
• Prerequisites for long-term preservation
• Persistent identification (DOI, URN, etc.)

Technical support on RDM
• Data archiving
• Data versioning – GitLab
• Data base hosting
TUdata – RDM service team at TU Darmstadt

- Central service team consisting of ULB & HRZ staff
- Contact for all questions on RDM:
  - ✓ Proposals
  - ✓ IT
  - ✓ Organisation
  - ✓ Training
  - ✓ Awareness & Outreach

- www.tu-darmstadt.de/tudata
- tudata@tu-darmstadt.de
A structured approach to RDM - use RDMO to create your own DMP

Nina Dworschak & Gerald Jagusch

https://www.ub.uni-frankfurt.de/forschungsdaten/rdmo.html
Further resources

Do you want to learn more?

Useful resources:

- [www.forschungsdaten.info](http://www.forschungsdaten.info) (in German)
- [https://www.dcc.ac.uk/information/researchers](https://www.dcc.ac.uk/information/researchers) (in English)

**DFG**
Video Tutorials:
https://rdmorganiser.github.io/dokumentation/

**Videos**

*Screencast: Wie beantworte ich einen Fragenkatalog? (Video-Download)*

13.10.2020 | A structured approach to RDM - use RDMO to create your own DMP | Nina Dworschak & Gerald Jagusch | 42
Thank you for attending!

Please evaluate this workshop: