



Leibniz-Institut für  
Astrophysik Potsdam

# Anforderungen an ein Tool zur Erstellung von Forschungsdatenmanagementplänen

Jochen Klar

# Projekt FDMP-Werkzeug

*Ziel des Projekts ist es, ein Werkzeug zur Verfügung zu stellen, das die strukturierte Planung, Umsetzung und Verwaltung des Forschungsdatenmanagements unterstützt und zusätzlich die textuelle Ausgabe eines Forschungsdatenmanagementplans ermöglicht.*

## Kollaboration

- Jochen Klar (AIP)
- Claudia Engelhardt (SUB)
- Jens Ludwig (SBB)
- Harry Enke (AIP)
- Heike Neuroth (FHP)

# Forschungsdatenmanagementpläne

- Optimierung des Forschungsdatenmanagements im Vorfeld
- Leitfaden für Umgang mit Forschungsdaten über den Projektverlauf
- Planung des Verbleibs der Forschungsdaten nach Projektende
- Erhöhung von Qualität und Effizienz von wissenschaftlicher Arbeit
- Mehr als nur Anforderung der Förderorganisation
- Im Idealfall Gewinn für die Forschenden

# Bestehende Tools

- DMPTool der California Digital Library
- DMPOnline des Digital Curation Centre
- DMP Checkliste an der Universität Bielefeld
  - Talk von Johanna Vompras
- TUB-DMP an der Technischen Universität Berlin
  - Talk von Fabian Fürste

# DMPTool

- [dmptool.org](http://dmptool.org)
- betrieben vom University of California Curation Center der California Digital Library (CDL)
- DMPTool existiert seit 2011, Version 2 seit Mai 2014
- ausgerichtet auf die verschiedenen Funding-Agencies in den USA (insb. NSF und NIH) und ihre Anforderungen an Datenmanagementpläne
- Moderner Webauftritt, Blog, Video, Outreach, Webinars
- MIT Lizenz, <https://github.com/CDLUC3/dmptool>





## SELECT DMP TEMPLATE

Select one of the funder DMP Templates listed to proceed to the next step. The type of template chosen can affect what information you will need to provide in the following steps.

Search

A - F

G - L

M - S

T - Z

All

- Alfred P. Sloan Foundation
- BCO-DMO NSF OCE: Biological and Chemical Oceanography
- Department of Energy: Office of Science
- DMP Template from DCC
- GoMRI Research Consortia DMP Template 2015
- Gordon and Betty Moore Foundation
- Institute of Education Sciences (US Dept of Education)
- Institute of Museum and Library Services
- Joint Fire Science Program
- National Institutes of Health
- National Science Foundation
  - NSF-AGS: Atmospheric and Geospace Sciences
  - NSF-AST: Astronomical Sciences
  - NSF-BIO: Biological Sciences (2013- )
  - NSF-CHE: Chemistry Division
  - NSF-CISE: Computer and Information Science and Engineering
  - NSF-DMR: Materials Research
  - NSF-EAR: Earth Sciences
  - NSF-EHR: Education and Human Resources
  - NSF-ENG: Engineering
  - NSF-GEN: Generic
  - NSF-PHY: Physics
  - NSF-SBE: Social, Behavioral, Economic Sciences
- NEH-ODH: Office of Digital Humanities
- U.S. Geological Survey DMP Guidance

<< Back

Next >>

## Department of Energy: Office of Science

Click on a section below to edit it at any time.

✔ = Complete

\* = Mandatory

### Template Outline

-  **Data types and sources** ✔
-  Content and format ✔
-  Sharing and preservation ✔
-  Protection ✔
-  Rationale ✔
-  Software & Codes ✔

### Instructions

### Links

The sections in the template outline are based on Suggested Elements of a DMP (see Links tab) provided by DOE, but DMPs are not required to follow this template. For the data types and sources suggested element, a brief, high-level description of the data to be generated or used through the course of the proposed research and which of these are considered digital research data necessary to validate the research findings may be included.

### Guidance

box size: [small](#) | [medium](#) | [full](#)

\*Requirement #1: DMPs should describe whether and how data generated in the course of the proposed research will be shared and preserved. If the plan is not to share and/or preserve certain data, then the plan must explain the basis of the decision (for example, cost/benefit

**B I U S**  $x_2$   $x^2$  |      |    

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Cancel Changes

Save Response

Save and Next

\*Requirement #1: DMPs should describe whether and how data generated in the course of the proposed research will be shared and preserved. If the plan is not to share and/or preserve certain data, then the plan must explain the basis of the decision (for example, cost/benefit considerations, other parameters of feasibility, scientific appropriateness, or limitations discussed in the Protection section). At a minimum, DMPs must describe how data sharing and preservation will enable validation of results, or how results could be validated if data are not shared or preserved.

Detailed instructions can be found on the [Office of Science Statement of Digital Management](#) page.

Check with your specific program to see if there are **specific program requirements**.

The term **digital data** encompasses a wide variety of information stored in digital form including: experimental, observational, and simulation data; codes, software and algorithms; text; numeric information; images; video; audio; and associated metadata. It also encompasses information in a variety of different forms including raw, processed, and analyzed data, published and archived data.

This statement focuses on **digital research data**, which are research data that can be stored digitally and accessed electronically. **Research data** are defined in regulation (2 CFR 200.315 (e)), continuing the definition from 2 CFR 215 (OMB Circular A-110) as follows:

"Research data is defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This 'recorded' material excludes physical objects (e.g., laboratory samples). Research data also do not include:

(A) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and

(B) Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study."

### Digital Research Data

The term **digital data** encompasses a wide variety of information stored in digital form including: experimental, observational, and simulation data; codes, software and algorithms; text; numeric information; images; video; audio; and associated metadata. It also encompasses information in a variety of different forms including raw, processed, and analyzed data, published and archived data.

#### See full definition

### Advanced Scientific Computing Research

ASCR considers software to be a data artifact that is covered by the Office of Science Statement on Digital Data Management. Research funded by ASCR may produce several different types of software (e.g., those that encode a mathematical algorithm, implement a specific tool or service, become part of a base Operating System). Other funded research may result in the creation or consumption of data that could be used to analyze application or system behavior. Software and data created by funded research must be released with sufficient descriptions to facilitate the validation of research results.

\*Requirements - The DMP must briefly discuss the applicable issues listed below:

- \* The Open Source License to be used, if applicable;
- \* If executable versions of the software will also be released, and if so what format will be used;
- \* How software can be found and accessed and the length of time the software will be publicly available;
- \* How any proprietary 3rd party software or libraries are used in the creation of this software;

### Office of Fusion Energy Sciences

The Office of Fusion Energy Sciences does not have any additional requirements, but they do provide additional guidance of software and codes.

Data management plans for codes should address the elements outlined in the SC Statement. With respect to codes, FES specifically:

- Encourages that codes developed with FES funding and used to analyze experimental data and/or to perform simulations be made available to the broader community via Open Source licensing.
- Recognizes the right of the code developers to establish a reasonable exclusive use period for the developed software; in addition, recognizes that sharing may be limited to stable versions of the code(s) and not to the currently-under-development "research" versions of the code(s).
- Recognizes that there may be code sharing limitations driven by proprietary interests, export control issues, intellectual property rights, and issues affecting US competitiveness.
- Acknowledges that code sharing does not create obligations for the code developers to provide software support, unless there is a collaboration agreement between the parties.
- Understands that code sharing should respect existing user license agreements established by the code developers



Export as PDF

Export as RTF

Print

**DMP Name:** A cool project

**DMP Template:** Department of Energy: Office of Science

**Owner:** Jochen Klar

**Institution:** Non partner institution

**Last Modified Date:** March 18, 2015

**Co-owner[s]:**

**Solicitation Number:**

### 1. Data types and sources

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

### 2. Content and format

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

### 3. Sharing and preservation

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam

# DMPTool

- Nutzerkonten direkt bei DMPTool oder per Single Sign-On der Institutionen
- Kollaboratives Arbeiten, Publizieren von DMP
- Templates werden durch Editoren gepflegt und mit Ressourcen angereichert
- Nur geringe Anreicherung der vorgegebenen Templates
- Aufgrund der unterschiedlichen Förderlandschaft nur schlecht auf Deutschland übertragbar

# DMPOnline

- [dmponline.dcc.ac.uk](http://dmponline.dcc.ac.uk)
- betrieben vom Digital Curation Centre (DCC)
- v.4 seit Dezember 2013
- Struktur und Features wie bei DMPOnline
- zugeschnitten auf das UK mit seinen Research Councils, aber auch Horizon 2020
- Angereichert durch
  - Vorgaben des Förderers
  - Hinweise der Institution
  - Hinweise des DCC
- AGPL3 Lizenz, Code auf github



## Data sharing

**B** *I* [List Icon] [List Icon] [Link Icon] [Table Icon]

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Save

Answered 2 minutes ago by Jochen Klar

## EC Guidance

Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).

In case the dataset cannot be shared, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

## Edinburgh guidance on Data Repository

For example a subject-specific data archive or the University of Edinburgh [DataShare repository](#)

If you do not plan to use an established repository, the data management plan should demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant.

## Edinburgh guidance on Method For Data Sharing

## Edinburgh guidance on Restrictions on Sharing

## DCC guidance on Method For Data Sharing

## DCC guidance on Restrictions on Sharing

## DCC guidance on Data Repository

# Aspekte

- DMPTool und DMPOnline sind in erster Linie Hilfsmittel zur Antragstellung
- Zentral betriebene Webangebote
  - Weitergabe von potentiell sensiblen Informationen
  - Eingeschränkte Anpassbarkeit durch einzelne Institutionen
- Generisch im Hinblick auf den disziplinären Hintergrund
- Tools im deutschsprachigen Raum sind an eine Institution gebunden
  - zugeschnitten auf die entsprechenden Infrastrukturen
  - nicht zugänglich für Nichtmitglieder
- Alle Tools sind nicht ohne Weiteres wiederverwendbar

# Erweiterte Anforderungen

- Ansprechen aller im Forschungsdatenmanagement involvierten Akteure
- Abdeckung des Forschungsdatenmanagement über die gesamte Projektlaufzeit
- Strukturiertes „Interview“
  - Entscheidungsbaum
  - Überspringen redundanter Fragen
  - Auswahlfelder, Checkboxen
- Anpassbarkeit der Inhalte
  - institutionelles Umfeld
  - disziplinspezifische Aspekte
- Einfache Installation in einem anderen Kontext

# Projektziele

- Konzeption
  - Inhaltliche Aktualisierung der WissGrid Checkliste
  - Zielgruppenausrichtung: Arbeitsgruppe, Projekt, Institution
  - Mehrsprachigkeit
  - Logische Strukturierung
  - Verknüpfung der Aufgaben, verschiedene Perspektiven
- Software
  - Realisierung einer Webanwendung für DMP
  - Schnittstellen (LDAP, DFN\_AAI, REST)
  - Offener Quellcode zu jedem Zeitpunkt der Entwicklung
  - Reibungslose Installation (z.B. durch IT-Abteilung)
  - zusätzlich zentrale Instanz

# Projektziele

- Nutzertests durch die Zielgruppen
- Kollaboration
  - Abteilung Optische Sonnenphysik am AIP
  - eResearch Alliance Göttingen
  - GESIS / Internationale Dateninfrastrukturen
  - ZPID / Projekt DataWiz
- Community
  - Abschlussworkshop
  - Schulungssession
- Laufzeit 18 Monate





Leibniz-Institut für  
Astrophysik Potsdam

Vielen Dank für die Aufmerksamkeit

Kontakt: [jklar@aip.de](mailto:jklar@aip.de), [@jochenklar](https://www.instagram.com/jochenklar)

[www.aip.de](http://www.aip.de)

Three thin, parallel yellow lines curve across the bottom of the slide, starting from the left and ending on the right, with a slight upward slope towards the right.