

Source: RDA DMP Common Standard for machine-actionable Data Management Plans  
(<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard> )

## **DATA MANAGEMENT PLAN REQUIREMENTS**

### **fields corresponding to the RDA DMP**

#### **FWF:**

### **1. Data Characteristics: Description of the data**

- What kinds of data/source code will be generated or reused (type, format, volume)? (dataset - type, dataset - distribution - byte size, distribution - format)
- How will the research data be generated and which methods will be used? (dataset - description)
- How will you structure the data and handle versioning? (host - support\_versioning)
- Who is the target audience? (dmp - description)

### **2. Documentation and Metadata**

#### **2.1 Metadata standards:**

- What metadata standards (if any) will be in use and why? (see [Digital Curation Centre](#)) (dataset - metadata, metadata\_standards\_id)

#### **2.2 Documentation of data:**

- What information is needed for the data to be findable, accessible, interoperable and re-usable ([FAIR](#)) in the future? (distribution - access url, available\_until, data access, host - url, license)
- Is the data machine-readable?
- How are you planning to document this information? (dmp\_id)

#### **2.3 Data quality control**

- What quality assurance processes will you adopt? (dataset - data\_quality\_assurance)
- How will the consistency and quality of data collection be controlled and documented? (distribution - access)

### **3. Data Availability and Storage**

#### **3.1 Data sharing strategy:**

- How and when will the data be shared and made accessible? (distribution -data access, distribution - available\_until)
- What repository will you be using? (host - description, geo location, certified with, title, url)
- What persistent identifier will be used? (dataset - identifier, type; metadata standard id - identifier)

#### **3.2 Data storage strategy:**

- What data are to be preserved for the long-term, and what data will not be stored? (dataset - description, type, preservation\_statement; sensitive\_data)
- How and where will the data be stored and backed up during the research? (host - backup frequency, backup type, storage type)

- How and where will the data be stored after the project ends? (host - description, storage type, title, url)
- For how long will the data be stored? (distribution - available\_until)
- Are there any costs that need to be covered for storage? (cost - currency-code, description, title, value)
- At what point during or after the project will the data be stored?
- Are there any technical barriers to making the research data fully or partially accessible? (dataset - technical\_resource, security\_and\_privacy)

#### 4. Legal and Ethical Aspects

##### 4.1 Legal aspects:

- Are there any legal barriers to making the research data fully or partially accessible? (dataset - distribution, technical\_resource)
- Who owns the data? (contributor - \_id, mbox, name, role)
- What licence for reuse are you planning to attach to the data? (distribution - license)
- Are there any restrictions on the re-use of the data? If so, why? (dataset - distribution)

##### 4.2 Ethical aspects:

- Are there any ethical barriers to making the research data fully or partially accessible? (ethical\_issues\_description, ethical\_issues\_exist, ethical\_issues\_report)
- If applicable, how are you planning to deal with sensitive data during and after the project? (dataset - sensitive\_data) Consider "[Ethics for researchers](#)" published by the European Commission or "[The European Code of Conduct for Research Integrity](#)".

## Academy of Finland:

### **1. General description of data:**

1.1 What kinds of data is your research based on? (project - description) What data will be collected, produced or reused? (dataset - description, type) What file formats will the data be in? (distribution - format) Give a rough estimate of the size of the data produced or collected (distribution - bytesize, download url)

- Also explain what kinds of existing data you will use, for example, the types of texts, images, photographs, measurements, statistics, physical samples or codes.

1.2 How will the consistency and quality of data be controlled? Explain how the data collection, analysis and processing methods used may affect the quality of data and how you will minimise the risks related to data accuracy. (dataset - data\_quality\_assurance, security\_and\_privacy)

### **2. Ethical and legal compliance:**

2.1 What ethical issues are related to your data management, for example, in handling sensitive data, protecting the identity of participants, or gaining consent for data sharing? (dmp - ethical\_issues\_description, ethical\_issues\_exist, ethical\_issues\_report, dataset - description, distribution, personal\_data security\_and\_privacy, sensitive\_data)

- Describe how you will maintain high ethical standards and comply with relevant legislation when managing your research data. Ethical issues must be considered throughout the research data life cycle. If you are handling personal or sensitive information, describe how you will ensure privacy protection and data anonymisation or pseudonymisation.

2.2 How will data ownership, copyright and IPR issues be managed? Are there any copyrights, licences or other restrictions that prevent you from using or sharing the data? (contributor - contributor\_id, mbox, name, role, contact - contact\_id, mbox, name, distribution - license, license\_ref, start\_date)

- Describe who will own the data and how the ownership issues have been agreed. Describe who can issue permissions to (re)use the data.

### **3. Documentation and metadata**

3.1 How will you document your data to make them findable, accessible, interoperable and reusable for you and others? What kinds of metadata standards, files or other documentation will you use to help others understand and use your data? (dmp\_id, metadata - description, language, metadata\_standards\_id)

## 4. Storage and backup during the research project

4.1 Where will your data be stored, and how will they be backed up? (host - description, backup\_frequency, backup\_type, storage\_type, title, url; contributor - contributor\_id, mbox, name, role, contact - contact\_id, mbox, name)

- Describe where you will store and back up your data during your research project. Methods for preserving and sharing your data after your research project has ended are explained in more detail in Section 5.
- Consider who will be responsible for backup and recovery. If there are several researchers involved, create a plan with your collaborators and ensure safe transfer between participants.

4.2 Who will be responsible for controlling access to your data, and how will secured access be controlled? (security\_and\_privacy, dataset - data\_access, contributor - contributor\_id, mbox, name, role, contact - contact\_id, mbox, name)

- Describe who has access to your data, what they are authorised to do with the data and how you will ensure the safe transfer of data to your collaborators.

## 5. Opening, publishing and archiving the data after the research project

5.1 What part of the data can be made openly available or published? Where and when will the data, or their metadata, be made available? (distribution - available\_until, access\_url, data\_access; host - description, storage\_type, title, url)

- Describe whether you will publish or otherwise make all your data or only parts of them openly available. If your data or parts of them cannot be opened, please explain why.

5.2 Where will data with long-term value be archived, and for how long? (preservation\_statement, host - description, storage\_type, title, url, distribution - available\_until)

- Briefly describe what data to archive and for how long – as well as what data to dispose of after the project. Describe the access policy to the archived data.

## 6. Data management responsibilities and resources

6.1 Who will be responsible for specific tasks of data management during the research project life cycle? (contributor - contributor\_id, mbox, name, role; contact - contact\_id, mbox, name; technical\_resource; cost - currency - code, description, title, value)

- Estimate also the resources (e.g. financial, time and effort) required for data management.

- Estimate the need to hire expert help to manage, preserve and share the data. Consider the additional computational facilities and resources that need to be accessed, and what the associated costs will amount to.