

## DEFF: DEFF DMP template (in English)

### Data collection - 1

What data types are collected or generated to answer the project's main objectives?

*Guidance:*

Here you can give a short description of the data you will create, or third-party data you plan to reuse. Consider the types of data your research will generate, for example qualitative survey data, computational models, software, databases, statistics, measurements, text, images, or audiovisual data.

Which file formats are the data in?

*Guidance:*

Make sure that you also consider the longevity of the file formats that you choose. Using open standards so your data can be read by a variety of software makes it easier to preserve and share with others. Questions related to preservation and sharing will be asked later on in the DMP.

How much data will be generated?

- 0-5 GB
- 50 GB - 500 GB
- 500 GB - 2 TB
- 5 GB - 50 GB
- > 2 TB

If you are using non digital data (e.g. physical objects), how are you storing those data?

### Data collection - 2

Which data collection standards will you use?

*Guidance:*

Outline how the data will be collected/created and which community data standards (if any) will be used.

Which methodology will you use to collect the data?

How will you structure and name your folders and files?

*Guidance:*

Consider how the data will be organized during the project, mentioning for example naming conventions, version control and folder structures.

How will you handle versioning?

What quality assurance processes will you adopt?

*Guidance:*

Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeat samples or measurements, standardized data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.

## Documentation and Metadata

How will you capture / create the metadata?

*Guidance:*

Metadata (i.e. description of the data, or, data about data) should be created to describe the data, and help discovery or retrieval of data. Consider how you will capture this information and where it will be recorded e.g. in a database with links to each item, in a 'readme' text file, in file headers etc

Read more here:

[DCC - Metadata](#)

Can any of this information be created automatically?

What metadata standards will you use?

*Guidance:*

Researchers are strongly encouraged to use community standards to describe and structure data, where these are in place.

Read more here:

[DCC - Metadata Standards](#)

[DCC - Metadata](#)

[DDI Alliance](#)

What metadata, documentation or other supporting material should accompany the data for it to be interpreted correctly?

*Guidance:*

Describe the types of documentation that will accompany the data to provide secondary users with any necessary details to make data sets understandable and reusable, increase visibility of the data and promote transparency in research. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data. As a minimum there must be some basic information that will help find data, including who created the data, a title, a date and any access restrictions.

What information or software needs to be retained to enable the data to be read and interpreted in the future?

## **Ethics and legal compliance**

Have you gained consent for data preservation?

*Guidance:*

Investigators carrying out research involving human participants must ensure that consent is obtained to allow data to be shared and reused.

Have you gained consent for data sharing?

How will sensitive data be handled to ensure it is stored and transferred securely?

*Guidance:*

Managing ethical concerns may include: anonymization of data; referral to departmental or institutional ethics committees; the Danish Data Protection Agency; and formal consent agreements. Ethical issues may affect how you store data, who can see/use it and how long it is kept. Read more here:

[Danish Data Protection Agency](#)

[Danish Data Protection Agency - The Act on Processing of Personal Data/](#)

How will you protect the identity of participants?

How are copyright and IPR of newly generated data regulated?

*Guidance:*

State who will own the copyright and IPR of any new data that you will generate. For multi-partner projects, IPR ownership may be worth covering in a consortium agreement.

If the data is suitable for reuse. How will they be licensed?

*Guidance:*

Digital Curation Centre (DCC) has made a guide to help decide how to apply a licence to research data, and which licence would be most suitable. The guide concentrates on the UK context, though some aspects apply internationally.

Read more here:

[How to License Research Data](#)

If you are using third-party data, how do the permissions you have been granted affect licensing?

*Guidance:*

If purchasing or reusing existing data sources, consider how the permissions granted to you affect licensing decisions.

Will data sharing be postponed / restricted e.g. to seek patents?

*Guidance:*

Outline any restrictions needed on data sharing e.g. to protect proprietary or patentable data.

## **Storage, backup and security**

Where will the data be stored?

Do you have access to enough storage or will you need to include charges for additional services?

How will the data be backed up?

*Guidance:*

Back-up of your unique data is more critical than copies of secondary data. The more important the data and the more often it is used, the more regularly it needs to be backed up. Fully managed file services with automated back-up are very robust and save you the time and effort of implementing your own system. Such services could be used in combination with portable storage or cloud computing to meet particular needs.

Who will be responsible for backup and recovery?

How will the data be recovered in the event of an incident?

What are the risks to data security and how will these be managed?

How is security for sensitive data guaranteed?

*Guidance:*

All data should be managed securely, but special attention should be given to sensitive data (e.g. personal data not already in the public domain, confidential information or trade secrets). Here you should outline any appropriate security measures and note any formal standards that you will comply with e.g. [ISO 27001](#).

Read more here:

[ISO Standards](#)

How will you control access to keep the data secure?

How will you ensure that collaborators can access your data securely?

If creating or collecting data in the field, how will you ensure its safe transfer into your main secured systems?

## **Selection, preservation and sharing**

What data should be retained?

*Guidance:*

Decide which data to keep and for how long. This could be based on any obligations to retain certain data, the potential reuse value, what is economically viable to keep, and any additional effort required to prepare the data for data sharing and preservation. Remember to consider any additional effort required to prepare the data for sharing and preservation, such as changing file formats.

What data must be retained for contractual, legal, or regulatory purposes?

What data must be destroyed for contractual, legal, or regulatory purposes?

How will you decide what other data to keep?

What are the foreseeable research uses for the data?

Where e.g. in which repository or archive will the data be held?

*Guidance:*

Consider how datasets that have long-term value will be preserved and curated beyond the lifetime of the grant. Also outline the plans for preparing and documenting data for sharing and archiving. If you do not propose to use an established repository, the DMP should demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant.

How long will the data be retained and preserved?

Have you estimated in time and effort to prepare the data for sharing / preservation?

What is the potential value of long term preservation?

Can you share your data?

- Yes
- No (please clarify where it will be stored and preserved)

Which data will be shared and how? Specify where the data and associated metadata, documentation and code are deposited.

*Guidance:*

Consider where, how, and to whom data should be made available. The methods used to share data will be dependent on a number of factors such as the type, size, complexity and sensitivity of data. Consider how people might acknowledge the reuse of your data (e.g. via citations) so you gain impact.

With whom will you share the data, and under what conditions?

Will you share data via a repository, handle requests directly or use another mechanism?

How will the identity of the person accessing the data be ascertained?

Will you pursue getting a persistent identifier for your data?

*Guidance:*

Everything that exists digitally is at risk of getting lost, for example when websites reorganise or the digital object is moved to another location on the web. Therefore, it is necessary to use identifiers which can help uniquely identify the object.

Read more here:

[Digital Object Identifier](#)

[Danpid - The Danish service for persistent identifiers to objects](#)

When will you make the data available?

How will potential users find out about your data (e.g. by adding keywords)?

Are there restrictions on data sharing?

- Yes (If yes, please continue)
- No

What action will you take to overcome or minimize restrictions?

*Guidance:*

Outline any expected difficulties in sharing data, along with causes and possible measures to overcome these. Restrictions may be due to confidentiality, lack of consent agreements or IPR, for example.

For how long do you need exclusive use of the data?

Will a data sharing agreement (or equivalent) be required?

*Guidance:*

Consider whether a non-disclosure agreement would give sufficient protection for confidential data.

## **Responsibilities and Resources**

Who is responsible for implementing the DMP, and ensure it is reviewed and revised?

*Guidance:*

Outline the roles and responsibilities for all activities e.g. data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Consider who will be responsible for ensuring relevant policies will be respected. Individuals should be named where possible.

Who will be responsible for each data management activity?

How will responsibilities be split across partner sites in collaborative research projects?

Will data ownership and responsibilities for Research Data Management be part of any consortium agreement or contract agreed between partners?

Is additional specialist expertise on Research Data Management (or training for existing staff) required?

*Guidance:*

Carefully consider any resources needed to deliver the plan, e.g. software, hardware, technical expertise, etc. These costs can usually be written into grant applications but need to be clearly outlined and justified.

Do you require hardware or software which is additional or exceptional to existing institutional provision?

Will charges be applied by data repositories or archive?

- No
- Yes (If charges are applied, how you are going to cover for the costs?)