

Sharing research data

Openness and knowledge sharing is a prerequisite for all research. It is an important research policy objective that the results of publicly funded research should be as open as possible. In this context, by research results we mean data that is created through research activities.

[The Research Council of Norway's policy for open access to research data](#) is intended to make research data available to relevant users, on equal terms, at the lowest possible cost. The policy guidelines apply to all data in projects funded by us – with some exceptions. The international FAIR principles have been developed as a set of guidelines to facilitate further use of research data. [FAIR](#) is an acronym for the words findable, accessible, interoperable and reusable. In other words, research data must be of a quality that makes it findable, accessible, interoperable and reusable. The management of research data in projects that receive funding from the Research Council must follow the FAIR principles, as much as possible.

As open as possible, as closed as necessary

Better access to research data enhances the quality of research, both because results can be validated and verified in a better way, and because datasets can be used in new ways and in combination with other datasets. Open access to research data also contributes to fewer duplications and unnecessary duplication of effort and will facilitate more interdisciplinary research.

Some datasets cannot simply be made openly available. The Research Council of Norway's policy therefore operates with clear exceptions. Data sets shall not be made openly available if doing so may threaten individual or national security, conflict with applicable data protection regulations or other legal provisions.

How data should be archived

As the main principle it is up to the R&D performing enterprise to decide which archiving solution to use. However, where appropriate, the Research Council of Norway may order the storage of data and/or metadata in specific national or international archives.

When we require storage in a specific archive, we will always state this in the call for proposals, and we will write this into the contract for the projects.

Requirements for data management plan in projects that manage data

A data management plan is a tool and framework for researchers who contribute to thoughtful, structured and documented data management, throughout the research process. A good data management plan makes the research data easier to retrieve and understand for others, creates awareness of data security, costs and quality, makes the research reproducible and increases the potential for reuse. The data management plan should be a living document that is regularly updated throughout the life of the project. Read more about data management plans on opendscience.no.

The requirement for a data management plan for projects receiving funding from the Research Council of Norway was introduced in 2018. The requirement applies to all projects that have received funding after 01.01.2018 and that manage data in their project. Projects must submit the first version of the plan when revising the grant application. An updated version is delivered together with the project's final report. The Research Council of Norway does not assess the content of submitted plans. It is the responsibility of the Project Owner to approve that the plan is in line with the institution's requirements and guidelines before it is submitted. Data management plans should, as far as possible, be public and published openly so that academic communities can better follow their peers' practices.

Based on the data management plans received, the Research Council will accept any costs for managing data as part of the operating costs of the projects we support. In addition, the Research Council of Norway will attach importance to providing funding for good infrastructures for data storage and data management, for example through the [National Financing Initiative for Research Infrastructure](#).

What a data management plan should include

This guide is a tool for projects that manage data and are required to submit a data management plan when revising the grant application and final reporting to the Research Council. It is based on Science Europe's '[Practical Guide to the International Alignment of Research Data Management](#)'.

The Research Council recommends using a service for data management plans that allows the project to generate a machine-actionable data management plan, for example according to the [RDA Common Standard](#). Until further notice, the project must upload the data management plan in the format of pdf, .doc(x) or similar. We are working on developing our systems to facilitate machine-actionable data management plans. We also recommend assigning your data management plan a persistent identifier, such as a DOI. Several services for data management plans offer this.

Administrative information

The data management plan should include information about the Project Owner's institution, project manager, project number, project title, funder and version.

Collection and/or use of existing data

The data management plan should include a description of how new data is collected or produced and/or how existing data will be (re)used. This will help make it easier for others to understand the technical context in which the data arises or is (re)used, ensures verifiability, and enables further reuse.

These points should be described:

- What methods or software are used if new data is collected or produced?
- How should data provenance be documented?
- Are there any limitations related to file format, licensed software or similar for (re)use of existing data?
- Are there any reasons why existing data sources are not reused?

File types and format

The data management plan should describe the types of data, data format and volume collected, produced and/or reused. This will help identify potential issues that may arise when storing, sharing and long-term preserving the data.

These points should be described:

- What types of data will the project collect, produce and/or reuse? Examples of type of data can be numerical, text, image, audio, video, etc.
- What format data is stored in when collecting and analyzing, such as .pdf, .xls(x), .doc(x), .txt, .rdf.
- On what basis are certain data formats chosen? For example, this may depend on expertise in specific formats, preference for open file formats is, standard formats accepted at data repositories, extensive use of specific formats in research communities or format given by equipment or software used.
- What is the (estimated) volume of data to be stored during collection and analysis, archived and possibly long-term preserved? This can be stated as storage space (bytes), number of objects, files, rows and columns.

Documentation, metadata, and data quality

Good organization of files, detailed documentation and focus on data quality are good craftsmanship in research projects that manage data. Anyone should be able to understand your project, data collection, analysis, and files based on the documentation you create. You may want or be required to share your data or your research colleagues are looking to verify, replicate or reuse your data. The data management plan should therefore include a description of (meta)data, how data quality is safeguarded during the project and say something about the documentation that will accompany the data.

These points should be described:

- What metadata will be used to help others identify and discover the data?
- What metadata standards will be used? Examples could be DDI, TEI, MARC, CMDI.
- How will data be organized throughout the project? This can be version control, file structure, conventions for file naming, etc.
- Is there other documentation necessary to facilitate reuse? It can be description of methodology, information about analysis and protocols, definitions of variables, electronic lab books, code books, readme.txt files, etc.
- How is reliability and quality of the data controlled and documented? This can include processes such as calibration of measuring instruments, repeated measurements and samples, standardized data capture, validation of data entry, peer review of data or use of controlled vocabulary.

Storage and data security during the project

When managing data, it is important to think about data security and how to store your data during the project, before collecting and analyzing. Check if your organization has guidelines on this. The data management plan should include a description of how (meta)data will be stored and backed up and how data security will be safeguarded during the project. This helps identify possible risks related to the data security and protection of your data, such as data loss or accidental or unwanted access.

These points should be described:

- Where will (meta)data be stored and backed up throughout the project, and how often will this be performed? Storing data on laptops, typical external hard drives, USB-sticks or similar is not recommended due to less protection and greater risk of data being lost.
- How should data be recovered in the event of an accident?
- Who will have access to the data during the project and how is access controlled? This is particularly important where the project is a collaboration with several research communities/institutions.
- If applicable, how should data security and risk management be handled in relation to sensitive data, such as personal data and data that underlies trade secrets?
- What institutional data protection policies apply?

Rights and legal requirements and codes of conduct

Sharing of research data can be limited by a number of legal and ethical factors related to, for example, privacy, protection of sensitive information and

commercialization. Clarification of rights to, and sharing restrictions for, research data is an important premise for being able to share data. When managing data, the data management plan should include a description of how research data should be handled in accordance with legal legislation and research ethics guidelines. This contributes to a conscious approach to sharing restrictions that apply to your data so that you do not inadvertently use means, such as a too restrictive license, that restrict the sharing of data more than necessary. In addition, clarification of rights to research data, especially in collaborative projects and between the individual researcher and institution, is orderly and ensures an appropriate division of responsibility.

These points should be described:

- Which legal entities have rights to and/or rights to determine the use of the research data?
- Will the data be openly accessible or with access restrictions, if so, what access restrictions? One example is that access to data is only granted via an authentication service.
- Will there be any purpose restrictions, such as that the data can only be used for non-commercial purposes, and if so, why?
- Which dedications to public domain or licenses should be applied to the research data?
- Where the project involves several partners and/or several legal or natural persons with rights to research data; How should rights to control data access be managed in the project?
- Where the research data falls under copyright or database protection under [the Copyright Act](#); What rights apply and how will this be managed in the project?
- When using data from a third party; What access and purpose restrictions, if any, apply to this data?
- What ethical issues can affect how data is stored and transferred, who has data access to view or use the data, and how long it should be kept?
- Which institutional, national and/or international guidelines for research ethics apply to the project? Examples may be approval from regional committees for medical and health research ethics (REK) or the Norwegian Food Safety Authority.

These points should be described if you are handling personal data:

- How are [GDPR](#) and the [Personal Data Act](#) complied with when handling/ processing personal data?
 - Is informed consent for long-term preservation and possibly sharing of personal data used?
 - Is anonymization, pseudonymization or encryption of personal data being considered for long-term preservation and/or sharing?
- Should a managed procedure be used for authorized access to personal data?

Data sharing and reuse

Data sharing is an important aspect of research integrity so that others can verify and reuse your data. Data sharing can take place at any time in your research project, but should be shared at the latest, for example, in the event of scientific publications where the data form the basis for scientific conclusions. It can be overwhelming to consider all aspects of data sharing, especially if you're new to it. Remember that data sharing does not mean that you should necessarily share openly, but on the principles of "as open as possible, as closed as necessary" and [FAIR](#) (findable, accessible, interoperable, re-usable).

The data management plan should include a description of when and how data will be made available. This will help you to make concrete choices about how to streamline your research where, for example, you collaborate with others and how you can make your data visible, it may be for colleagues or society as a whole. It pays off to assign a persistent identifier to your data so that it can be more easily retrieved and referred to. In addition, the data management plan should include a description of the methods or software necessary to access and (re)use the data.

Data sharing and reuse must be seen in the context of the other topics in the guide, but is highlighted here as a separate topic because it is an important aspect of the life cycle of your research data. Therefore, you will recognize some of the points that should be described from other parts of the guide.

These points should be described:

- How should the data be findable and how should it be shared? Examples may be that they are made available in a certified data repository, are indexed in a catalogue, that you use a secure data service, direct handling of data requests, etc.
- When should the data be shared? If, for example, an exclusive right of control granted by legal law is made that affects the time of sharing and, why and for how long? Examples may be that you wait until a scientific publication is available or that you want to protect intellectual property rights, such as patent law, until you have applied for a patent.
- Who can reuse the data? If access restriction is required, for example, that only certain groups/communities have access or a data sharing agreement will be used, it should be explained how and why and what measures are taken to minimize restrictions.
- How can the data be reused in a different context? For example, is there potential for commercial exploitation?
- Do potential users need specific tools, such as software, to access and (re)use the data? The sustainability of the software for future access to the data should be considered.
- Will a persistent identifier (DOI) for the datasets be used? Persistent identifiers should be applied to metadata and datasets so that they can be findable and referenced in a reliable and efficient manner. Using DOI also ensures that citations and reuse can be tracked. A certified data repository will often provide this to (meta)data deposited there.

Long-term preservation

Long-term preservation of your data is an important part of the scientific process. Your data may have value beyond the research project in which it occurs. Although they cannot be shared, they may have historical value for future researchers or, for example, be observational data that cannot be recreated. The national strategy on access to and sharing of research data refers to the researchers as responsible for making decisions about which data it is appropriate to preserve. The data management plan should include a description of the criteria used to determine which data should be retained for the long term and where and for how long it should be retained. This will help identify data in your project that you believe is valuable to preserve for posterity and establish good practice for long-term preservation of data in your research community.

These points should be described:

- What data must be preserved or deleted based on agreements, legal legislation and/or guidelines?
- What data should be preserved for a long time and what criteria are used to select these?
- What is potential future research purposes and/or users of the data?
- Where will the data be long-term preserved (for example, which data repository? If a specific data repository is not proposed, the plan should show that the data can be appropriately curated after the life of the project. It is recommended to refer to the policies and procedures of data repositories, including metadata standards and costs involved.

Data management – responsibilities and resources

Establishing clear responsibilities and establishing a good overview of cost and resource needs related to data management is important to avoid uncertainty about who is doing what, unpredictable costs and inefficient use of personnel. The data management plan should include a description of who will be responsible for data management in the project as well as the need for resources, both financial and time, that is dedicated to data management. This contributes to the accountability of different participants in the project, which is especially important where there are several partners, and supports the coordination of various activities related to data management. In addition, this will contribute to efficient planning and highlight the need for resources related to data management in the project, both internally in your organization and when applying for external project funding. [This checklist from the UK Data Service](#) can be useful when estimating costs in your project.

These points should be described:

- Which roles are assigned which responsibility for data management activities in the project? Examples of activities are data capture, metadata production, data quality, storage and backup, long-term preservation and data sharing. Responsible individuals should be disclosed, if possible.
- For collaborative projects; How is responsibility for data management between partners coordinated?
- Who is responsible for implementing the data management plan and for ensuring that the plan is reviewed and regularly updated? In our guidelines, it is the responsible institution that must approve the plan.
- How are the necessary resources budgeted and covered in the project to prepare data for sharing and long-term preservation (curating)? These can be costs related to storage, hardware, staff time, costs associated with preparing data for disposal and costs related to preservation at a data repository.

Tools and service providers to create a good data management plan

There are several providers and tools that generate data management plans for research projects. The solutions make it possible to update the data management plan during the project period. Here are examples of tools and services available to generate data management plans:

Data Stewardship Wizard (DSW), ELIXIR Norway

- [Data Stewardship Wizard \(DSW\). ELIXIR Norway](#)
- [Create a data management plan \(DMP\) | Sikt](#)
- [Digital Curation Centre](#)
- [easyDMP](#)
- [Argos \(openaire.eu\)](#)