

# Handling research data at the Humanities, SDU

## Contents

General principles pertaining to research data.....	2
Examples of why research data should be made publicly accessible.....	3
How to make research data publicly accessible.....	3
Where to publish research data? .....	4
Data storage options for active projects.....	4
Best practices for storage of data .....	4
Long-term preservation/archiving .....	5
Archiving as an alternative to deletion .....	5
How to preserve data after the project (e.g. PhD Project) has ended.....	6

The Department of Culture and Language and the Department of Design, Media and Educational Science support the *Sorbonne Declaration on Research Data Rights*<sup>1</sup>. We are thus committed to practice and promote data handling according to the FAIR principles whenever relevant and in consideration of the special circumstances pertaining to research within the Humanities listed in the Open Science Policy. This implies that researchers at our departments must handle their research data in accordance with the FAIR principles.

Supporting the Sorbonne declaration places a premium on the idea of sharing data, the preparation of data in an open and FAIR manner, the recognition of the value of data and of the work entailed in making data sets publicly accessible.

A key element in adhering to the FAIR principles is that data as a minimum are made publicly accessible as metadata (i.e. a description of the content of a given data set, methods used and contact information of the researcher) allowing for the data to be discovered by other researchers and can be made available if the researcher agrees to it.

- All data sets that underpin research outputs e.g. individual articles or books should be made publicly accessible.
- Unique data sets, costly data sets and data collected longitudinally should be made publicly accessible.
- Examples of data that are not meant to be published are temporary data, explorative data, small data sets and data sets which are easily to recreate, and data sets restricted by GDPR.
- All data underpinning publications should be retained for a minimum of 5 years after publication of the results, as stated in the Danish Code of Conduct for Research Integrity.<sup>2</sup>
- It is recommended that data sets which are not published are deleted when there is no longer use for them.

---

<sup>1</sup> The *Sorbonne Declaration on Research Data Rights*: <https://sorbonnedatadeclaration.ent.upmc.fr/>

<sup>2</sup> See <https://ufm.dk/publikationer/2014/the-danish-code-of-conduct-for-research-integrity>

- That researchers that hold rare datasets and datasets with a unique value are strongly encouraged to archive these in the Danish National Archives (Rigsarkivet) - if accepted by the Archives.<sup>3</sup>

Researchers can consult the SDU Research Data Management Support<sup>4</sup> for inquiries about which data to publish and the best solution for publication and storage.

## General principles pertaining to research data

Research data must be:

- Acknowledged as valuable output of research that should be made openly available and re-usable, where possible.
- Covered by a data management plan when commencing a new research project.
- Stored securely and appropriately.
- Made **F**indable, **A**ccessible, **I**nteroperable and **R**eusable (FAIR).
- Retained for a minimum of five years after publication of the research.
- Considered archived at Danish National Archives (Rigsarkivet) if accepted, instead of deleted or anonymized at the end of the project, according to current legislation.<sup>5</sup>
- Managed in line with ethical protocols, including confidentiality.
- Managed in compliance with legal requirements for privacy and data protection GDPR.

### The FAIR principle of data handling

- **Findable** means that others can discover your (meta)data. This implies that they should be published in a repository that assigns a persistent identifier (e.g. a DOI number) and that relevant metadata should be assigned to data preferably according to a community-specific metadata standard.
- **Accessible** means that your data can be made available to others directly by download or through contact with the author. A data license (e.g. Creative Commons) or a clear data accessibility statement needs to be attached, and data should be archived in long-term storage. *In case of personal data, then at least the metadata should be published open access along with the contact details and preferably the ORCID of the principal investigator.*
- **Interoperable** means that your data can be integrated with other data. The use of standardized metadata, standard terminology and broadly used open formats for (meta)data is encouraged.

---

<sup>3</sup> See guidelines: <https://www.rigsarkivet.dk/aflever-data/>

<sup>4</sup> See <https://sdunet.dk/da/research/research-data-management-support>

<sup>5</sup> Only a requirement for personal data.

- **Re-usable** means that your data can be used for new research. (Meta)data should be well- documented and released under a clear usage license, which is as open as possible (e.g. Creative Commons, MIT, GPL, etc.).

### **Examples of why research data should be made publicly accessible**

- Required by funders or publishers (e.g. Horizon Europe).
- An obligation towards collaborators.
- To expose the department's research and increase the impact.
- To enable new research and collaborations.
- For public interest.
- For secondary data analysis in other projects.
- For use in teaching and student projects.
- To heighten credibility and accountability of research at the department.
- To improve transparency and reproducibility.
- To prevent or detect research fraud, as well as biased and selective analyses and publication.
- To make replication of statistical and psychometric analyses possible

### **Data which should not be made publicly accessible**

The data types below do not have to be publicly accessible. They are however, subject to the creation of a Data Management Plan and the FAIR principles as per the Open Science Policy

- Administrative data.
- Data from third parties, data repositories and administrative registers with conditions limiting reuse, publication and dissemination (e.g. materials on loan from museums or archives) or copyrighted materials.
- Publicly available data including literature (news media, books and journals)
- Studies included in systematic reviews and meta-analyses. The exemption does not apply to documentation of searches, selection of studies for review and analyses in tables, figures and similar supplementary material routinely published online with reviews.

### **How to make research data publicly accessible**

- As a minimum data which underpin a publication should be made publicly accessible as metadata (a description of the content of a given data set, methods used and contact information of the researcher) allowing that data can be discovered by other researchers and can be made available if the researcher agrees to it.

- A more comprehensive way to publish data is that the researcher makes the data that underpins a publication publicly accessible in full, implying that data should be stored in a public free trusted data repository (e.g. Zenodo, Dataverse etc.) with documentation (e.g. code book) that allows graphs to be reproduced and results to be recalculated.
- All datasets should receive a license for reuse, e.g. Creative Commons<sup>6</sup>.
- All datasets that can be shared should receive a persistent identifier (e.g. a Digital Object Identifier - DOI).
- All data should include all necessary documentation and metadata.
- Use open and long-lived file formats such as .csv alongside R, SPSS, STATA, SAS or other files for statistical or data management software.
- Use Danish National Archives/Danish Data Archive (Rigsarkivet/Dansk Data Arkiv), only for preservation (if accepted).<sup>7</sup>

## Where to publish research data?

The university library offers guidance and links to a directory of academic open access

Repositories on how to publish research.

It is recommended that if a researcher cannot identify a repository of special relevance for his/her research field, he or she uses the Zenodo repository.<sup>8</sup> This repository is created on behalf of the EU, maintained by CERN and allows for direct upload and management of the site by the researcher. Furthermore, it is citable, meaning that each upload of metadata or full data sets receives a unique object identifier - a DOI.

For publishing data we recommend to use some of the repositories mentioned in this link:

<https://sdunet.dk/en/research/research-data-management-support/databases-and-repositories-for-research-data#datarepositories>

## Data storage options for active projects

See the General Data Protection Regulation (GDPR) guidelines. Research Data Management support ([rdm-support@bib.sdu.dk](mailto:rdm-support@bib.sdu.dk)) and system administrator Erik B. Madsen ([erikm@sdu.dk](mailto:erikm@sdu.dk)) can advise on options.

## Best practices for storage of data

---

<sup>6</sup> See concerning licenses: <https://creativecommons.org/licenses/?lang=da>, (in Danish) <https://creativecommons.dk/http://sciencecommons.org/projects/publishing>

<sup>7</sup> <https://www.rigsarkivet.dk/aflever-data/>

<sup>8</sup> <https://zenodo.org/>

- Use of predefined file structures.
- Use of file versioning systems.
- Sample labelling and tracking.
- Discipline-specific metadata standards.
- File naming, dating and versioning according to best available methods and practices. The university library can advise on this, contact Research Data Management Support.<sup>9</sup> See also: <https://guides.library.stanford.edu/data-best-practices>
- Use of lab notebooks, preferably electronic, for experimental data.

Department staff is encouraged to develop additional recommendations for best practices, based on experience with implementation of the open science policy.

### **Long-term preservation/archiving**

All data should be stored for a minimum of five years after publication of the research (required permissions should be obtained or extended, including from the Danish Data Protection Agency). Beyond this minimum requirement, several types of research data should be preserved for long term access and reuse, including (the list is not exhaustive):

- If it would be unethical to subject humans or animals to unnecessary repetition of experiments, trials, observations or other research activities.
- If it would be unethical or indefensible to waste research funds and human resources that could be put to better use (i.e. prevention and cure of disease) on unnecessary repetition of experiments, clinical trials and observational research. Data and materials that are impossible or hard to reproduce.
- Data and materials that are costly to produce, in terms of funding, time or human resources.
- Data and materials that can be reused in new projects, serve as benchmarks, as reference or are of public interest.

### **Archiving as an alternative to deletion**

Valuable data and materials should be preserved by archiving in the Danish National Archives (Rigsarkivet). Preserving your data and materials in this archive fulfils legal requirements of deletion when a data processing permission expires.

---

<sup>9</sup> Write to: [rdm-support@bib.sdu.dk](mailto:rdm-support@bib.sdu.dk)

Documentation of archived data is required, using guidelines from the Danish National Archives (Rigsarkivet) (in Danish).<sup>10</sup>

If the Danish National Archives (Rigsarkivet) declines to archive the data, the person responsible for the data is responsible for deleting or anonymising the data if the data can be linked to an individual.

### **How to preserve data after the project (e.g. PhD Project) has ended**

- Document the data, using guidelines from the Danish National Archives (Rigsarkivet).
- According to the permission from the Danish Data Protection Agency, personal and sensitive data should either be archived in the Danish National Archives (Rigsarkivet) or permanently deleted or anonymised before the permission expires.
- Data is encouraged to be offered to the Danish National Archives (Rigsarkivet) with the least restrictive conditions for access possible. Use Guidelines of the Danish National Archives (Rigsarkivet) for documenting, reporting and archiving research data are available in Danish: (See [link](#)).

---

<sup>10</sup> See, <https://www.rigsarkivet.dk/aflever-data/>