University of Basel

Research Data Management Network

# FAIR Data

In 2016, Mark D. Wilkinson et al. published the *FAIR Guiding Principles for Scientific Data Management and Stewardship* to optimize the publication and reusability of research data. FAIR is an acronym for **F**indable, **A**ccessible, **I**nteroperable and **R**eusable and consists of a set of 15 principles.

The Swiss National Science Foundation and other research funders require data to be published in a FAIR way. In order to comply with the (complex) FAIR principles, we offer the following recommendations:

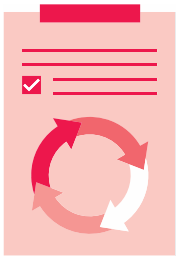
| **Findable** | **Accessible** |
|---|---|
| • Add metadata to your data, so that the data can be found.<br>• Add a persistent identifier (e.g. a DOI) to be able to reference data unambiguously. | • Publish the data online in a repository for research data.<br>• If data publication is not possible (e.g. due to legal restrictions), pubish metadata online and clearly state the restrictions for accessibility. |
| **Interoperable** | **Reusable** |
| • Use metadata standards and standardized controlled vocabularies to describe your data.<br>• Use file formats recommended for data archiving, as these are generaly widespread and readable by many softw are programs and can still be read in the long term. | • Provide sufficient documentation on your data, e.g. on the context of origin and all data processing.<br>• Add a license that is as open as possible to make it clear how the data may be reused. In the context of research data, CC licenses are widely used; CC0 or CC-BY are particulary recommended. |

**Check it out:** Use a self-assessment checklist to improve the FAIRness of your dataset.

University of Basel

Research Data Management Network

# Data Management Plan (DMP)



A DMP is a valuable tool designed to help you prepare for effective data management in the research process. It also serves as a checklist to ensure that important aspects of data handling are addressed during the research project and in the long term after its completion. Nearly all funding institutions now require a DMP, either already as part of a grant application or for approved projects.

As a dynamic, 'living' document, a DMP should be as concise as possible. Numerous templates and online tools are available to assist you in the development of your DMP.

**A DMP usually provides information on the following aspects:**

| **Data collection & documentation** | **Ethical, legal & security issues** |
|---|---|
| • Data type, format and volume<br>• Reuse of data<br>• Collection standards and methods<br>• File organization and versioning<br>• Documentation and metadata | • Data protection<br>• Data security<br>• Copyright issues<br>• Research ethics |
| **Data storage & preservation** | **Data exchange & reuse** |
| • Storage location & capacity<br>• Backup<br>• Responsibilities<br>• Planning and selection of data for archiving | • Access/publication<br>• Repositories |

**DMP templates by the University of Basel:**
On the website of the Research Data Management Network, you can find templates for writing DMPs in life sciences and humanities: https://researchdata.unibas.ch/en/planning/dmp

**DMP Tools**
- ARGOS
- Data Stewardship Wizard and DS Wizard of the SIB
- DMPTool
- DMPonline | DCC
- RDMO GitHub

University of Basel

Research Data Management Network

**By starting data management planning early in your project, you can…:**

- assign clear roles and responsibilities for data management within your team.
- estimate necessary infrastructure, services, legal advice, and budget requirements (sometimes reimbursable by your funder).
- identify open questions concerning your data that may need clarification during later project phases.
- increase research efficiency, saving time and resources in the long run.
- prevent loss of data by enhancing data security.
- avoid data duplication through effective data organization.
- establish continuity by ensuring that new project members can easily orient themselves through well-organized data and thorough documentation.
- boost the visibility and outreach of your research data through data publication.
- enhance the reproducibility and verifiability of your research.

University of Basel

Research Data Management Network

# Data Management Tools for Active Research

During active research, various tools and platforms can be used to organize your data. This page lists a non-exhaustive selection of domain-specific data management solutions that are made available by the University of Basel, or are freely available online.

## Natural sciences

### Data management with LabKey

- LabKey is a general purpose web-based data management platform for quantitative data. It is an open-source software developed by a commercial company specialized in clinical data management. Despite being rooted in clinical data, the platform is generic enough to suit most natural (and other quantitative) sciences. Please contact sciCORE to get started.

### Microscopy image repository with OMERO

- OMERO is a client-server software for managing, visualizing and analyzing microscopy images and associated metadata. The data repository not only supports the management of microscopy image data but also allows for their visualization, annotation, archiving, sharing and export of images. OMERO supports over 120 different microscopy file formats. At the Biozentrum, an OMERO server is jointly operated by the department's Imaging Core Facility (IMCF) and Biozentrum Research IT.

### Electronic Lab Notebook (ELN)

- An ELN is a software tool that typically replicates an interface similar to a page in a paper lab notebook. In an ELN you can enter protocols, observations, notes, and other experimental data. The Biozentrum Wiki can be used as an ELN by members of the Biozentrum and other life science groups of the University of Basel. There is a multitude of other ELN tools available; the ELN Finder can help you select a suitable tool. Beware of legal, ethical and/or financial constraints when choosing an ELN.

University of Basel

Research Data Management Network

# Humanities & social sciences

### Bibliographic data management with Zotero

- Zotero is a free, open-source literature management program for collecting, managing and citing various online and offline sources. It supports the editing of bibliographic information and literature lists in scientific publications. The University Library offers trainings.

### Archival photo management with Tropy

- Tropy is a free, open-source application to manage, describe and organize photographs of research materials, especially archival sources. An annotation tool allows to transcribe and contextualize research data. RISE provides support and trainings.

### Surveys with EvaSys

- EvaSys is suitable for the secure execution of large surveys and evaluations. It can be used for both paper-based and online surveys. The advantage lies in the various automation functions. The results of the survey can be further processed in other programs (e.g. SPSS) or summarized into any partial results. For more information and support, contact the EvaSys Survey Services of the University of Basel.

For further information about tools and platforms for research in the humanities & social sciences, contact the RISE team.

University of Basel

Research Data Management Network

# Data Organization

Organize your data with a clear folder structure and consistent file names. This allows you and others to easily locate, access and use your data, to avoid duplication, and to ensure that your data can be backed up.

## Use suitable file formats

Careful selection of file formats is crucial for ensuring that your files remain accessible, interoperable, properly archivable, and can still be used after many years. Aspects of a suitable file format are:

- No restrictive licenses
- Readable by many software products (at best by open source/code software)
- No encryption or digital rights management (DRM) – in case sharing is allowed
- Established in the research community
- Openly accessible documentation

## Use a file naming convention

Use consistent, meaningful file names. Include the following elements:

- Project/event or research team/department (abbreviation)
- Author (whole name or initials)
- Keyword(s) describing the content of the file
- Date (YYYY-MM-DD) and/or version

Not all characters are allowed for folder and file naming! Do not use spaces and/or special characters in the file name. Max. absolute path length: 255 characters.

## Use a folder structure

- Organize your data logically in a hierarchical folder structure.
- Separate active and completed work and delete any temporary files after use.
- Don't make the folder structure too nested, this can lead to long and complicated file paths.

University of Basel

Research Data Management Network

## Add documentation and metadata

Good documentation (e.g. in a readme file) facilitates the reuse of data – both for others and for yourself. An important element of documentation is metadata. Metadata are structured and machine-readable information about a (data) object. Metadata are therefore data about data. Metadata include, for example, information on title, creator, date, license, file format and persistent identifier. If there is no domain-specific or institutional standard for documentation, it is recommended to define an internal project standard.

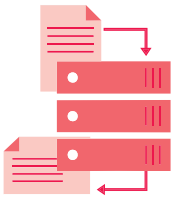
Check out the website of the Research Data Management Network for more information on...:

- Folder structures, file naming conventions and suitable file formats
- Documentation

University of Basel

Research Data Management Network

# Storing Active Research Data



Choosing the right storage for active data depends on your specific needs, budget, and requirements. Effective active data storage ensures fast, reliable, and efficient data access, providing high performance and responsiveness for both applications and users.

**Central data storage options:**
University of Basel employees can request a group share from IT-Services (ITS) for everyday research documents (e.g. office documents, small to mid-scale research data). For large projects requiring significant storage capacity (tens of TBs or more), sciCORE offers high-performance storage solutions. Both ITS and sciCORE central storage are regularly backed up.

**Advantages of central storage services:**
- Professional administration
- Service availability monitored and ensured by the infrastructure provider (ITS or sciCORE)
- Regular and professional maintenance, including infrastructure updates and expansions
- Established and verified disaster recovery concepts
- Comprehensive data backup policy

Check out the website of the Research Data Management Network for more information on storing research data. For specific information and technical support fill out the contact form.

University of Basel

Research Data Management Network

# Sharing Research Data with External Collaborators

During active research, data sharing with external collaborators is often necessary. Switch offers services for efficient data exchange, which support large file transfers and promote seamless collaboration. By leveraging these tools, researchers can maintain productivity and data integrity across institutional boundaries.

**Important advice on sharing personal data**

Be aware that sharing personal data requires specific solutions. In case you want to share personal data, contact IT-Services.

- Use encryption when sharing personal data; avoid unencrypted email, cloud services, or flash drives.
- Share personal data only with authorized collaborators.

## Switch Filesender

Students and university staff can use the secure web application Switch Filesender to exchange large files with anyone:

- Maximum file size per upload: 300 GB
  (multiple files can be uploaded simultaneously; the 300 GB limit applies to all files together)
- Maximum number of e-mail recipients: 50
- Maximum storage period for files: 60 days
- Supports data encryption with one click

## Switch Drive cloud storage

Switch Drive enables collaborative work for students and university staff. Users can share individual files and folders for continuous collaboration. Switch Drive features include:

- 100 GB online storage per user
- Protected access
- Synchronized files and folders across multiple devices
- Access via desktop client, browser or mobile application

University of Basel

Research Data Management Network

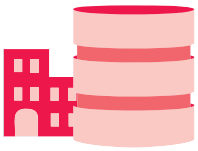
Switch Drive is usually bound to one person. If your project has a changing management, there is the possibility to get a project folder, where larger amounts of data (up to 1 TB) can be shared. It can be requested from IT-Services.

Servers are located in Switzerland, making Switch Drive a good alternative to many commercial cloud providers. However, Switch Drive is not a suitable storage location for unencrypted personal or other sensitive research data. Additionally, Switch Drive does not offer automatic backups, therefore users are responsible for backing up their own data.

University of Basel

Research Data Management Network

# Publishing Research Data



Alongside publishing your research findings, it is considered good practice – and often a requirement – to make the data that support your findings available to the research community. This promotes reproducibility of your findings and enables the reuse of your data in future research.

Before sharing your data, ensure you are allowed to! Beware of legal and ethical restrictions, especially for personal or other sensitive data. To publish your data, follow the FAIR principles and use a suitable repository that allows you to add proper documentation and metadata.

When selecting a repository, check the requirements of your research funding organization. Many require data publication in accordance with FAIR principles and the use of non-commercial repositories.

## Recommended repositories

**Subject-specific repositories:**
- DaSCH – Swiss National Data and Service Centre for the Humanities
- SWISSUbase, especially for social sciences and linguistics

**Interdisciplinary repositories:**
- Zenodo is a research data repository created, hosted and operated by CERN and OpenAIRE.
- Dryad is a research data repository, originally funded by the National Science Foundation U.S., governed by a nonprofit membership organization.

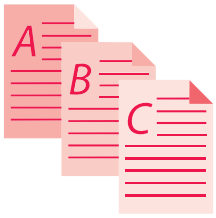
**Institutional repositories:**
The University of Basel does not provide an institutional repository and recommends using existing subject-specific or general repositories. If collaborating with other universities, you might have the option to publish data through their repositories.

**Repository registries:**
- The Swiss National Science Foundation (SNSF) provides a list of generalist, discipline-specific and institutional data repositories.
- The Open Access Directory provides a list of repositories for open data of various disciplines.
- Deposition Databases by Elixir, provides a list of repositories for biomolecular data.
- Find more repositories via re3data: a registry of research data repositories.

University of Basel

Research Data Management Network

# Data Preparation Guidelines for Publishing in a Repository




Thorough data preparation is essential for ensuring that your dataset is findable, accessible, interoperable, and reusable (FAIR). The following step-by-step guidelines will help you prepare your data for publication, ensuring compliance with the University of Basel's guiding principles.

## Data Review and Cleaning

Review

- [ ] Check your data for accuracy and completeness, and ensure all necessary components of the dataset are included.
- [ ] Use descriptive and consistent file naming conventions.

Cleaning

- [ ] Remove redundant records or variables.
- [ ] Correct inconsistencies and errors.
- [ ] Standardize variable naming and coding schemes

## Data Protection and Compliance

- [ ] When working with personal data, ensure compliance with relevant data protection laws and regulations.
- [ ] Follow the terms and conditions of the informed consent form that the research subjects have signed.
- [ ] Check that your data are properly anonymized or pseudonymized.
- [ ] Consult the University's data protection officer if in doubt.

University of Basel

Research Data Management Network

## Data Formatting

Open formats improve the interoperability and reusability of your data.

- [ ] Use open, non-proprietary, widely accepted formats (e.g., UTF-8, PDF, TIFF)
- [ ] Avoid closed formats unless necessary, and provide conversion instructions and/or soft ware details. You can also upload your dataset in two different file formats, if possible, to increase interoperability.

## Metadata Creation

Metadata are structured and machine-readable information about your data, and they should be as accurate and comprehensible as possible.

- [ ] Embed metadata into your data (e.g., title, authors, data source, data collection method, data type, and data format)
- [ ] Check the metadata fields that your repository provides and make sure that you have all relevant metadata/information ready.
- [ ] Follow standards such as Dublin Core, DataCite, or discipline-specific standards.
- [ ] If your dataset contains sensitive information and cannot be shared, consider creating a metadata-only record.

## Documentation

Create a detailed documentation that guides other researchers through your data, for example as a simple README file.

- [ ] Provide context (such as relation to other projects, publications, funders etc.), project objectives, and methodology of data collection.
- [ ] Document licenses applied to your own data and usage rights for third-party data.
- [ ] Describe data preparation and all further processing steps (data analysis, cleaning, interpretation).
- [ ] Describe data quality, and potential assumptions/limitations/biases.
- [ ] Consider adding a codebook that explains variables, codes, file names, standardized vocabulary, and file structure.
- [ ] Prepare a summary/abstract for the repository, which describes your dataset as accurately as possible so that other researchers can understand what the data is about, particularly its content, scope and specificity.

University of Basel

Research Data Management Network

## Copyright

- [ ] Be mindful of copyright laws – ensure that the data you want to publish are your own.
- [ ] If you want to publish third-party data that you have reused, verify that its license allows you to publish it.
- [ ] Document usage rights for third-party data in README file and add clear attribution where required.
- [ ] Consult legal support if uncertain.

## Licensing

Choose a license for your data. This determines how others are allowed to use your data.

- [ ] Where possible, data and products should be released under open licenses. Check with your selected repository which licenses are possible.
- [ ] For Creative Commons, it is recommended to use CC0 or CC-BY.
- [ ] Document license in the metadata and the README file.

## Data Archiving

- [ ] Back up data in a secure location (e.g., institutional storage) before publication.
- [ ] Maintain version control for updates.

University of Basel

Research Data Management Network

# Personal Data Management
## Laws and Regulations


What qualifies as personal data? What are you (not) allowed to do with it? Does your research project involve personal data? Do you know how to handle and store it?

Any researcher working with personal data is subject to (Swiss) data protection law and needs to be aware of the relevant laws and regulations.

### What is personal data?

Personal data is any information relating to an identified or identifiable natural person (data subject). There are different categories of personal data:

| **Ordinary personal data:** | **Sensitive personal data:** |
|---|---|
| Personal data is information referring to a definite or definable natural person. Examples include:<br>• Name<br>• Date of birth<br>• Email address<br>• Phone number<br>• Sex<br>• Information about particularly characteristic features, e.g. “only woman working in a certain team” | Personal data which are particularly sensitive in relation to fundamental rights and freedoms, which include but are not restricted to:<br>• Political, religious or philosophical beliefs<br>• Trade union-related views or activities<br>• Health<br>• Sexual life and/or sexual orientation<br>• Processing of genetic and/or biometric data<br>• Data of vulnerable persons, e.g. children, ethnic minorities etc. |

### Key laws

The Law on Information and Data Protection of the Canton of Basel-Stadt (IDG-BS), the Swiss Federal Act on Data Protection (FADP) and the General Data Protection Regulation of the European Union (GDPR) protect privacy and regulate the handling of personal data.

The Human Research Act (HRA) protects the dignity, privacy and health of human beings involved in research concerning human diseases and the structure and function of the human body.

University of Basel

Research Data Management Network

## Which law applies when you process personal data?

- Research at the University of Basel: In general, the Cantonal Law on Information and Data Protection (IDG-BS) and the Ordinance to the IDG-BS apply to public bodies of the Canton of Basel-Stadt and its municipalities (e.g. University of Basel).
- Private Research and Research at the ETHZ/EPFL: In general, the Swiss Federal Act on Data Protection (FADP) and the Ordinance to the FADP apply to private persons and federal bodies (e.g. ETHZ).
- Cross-cantonal research: When processing personal data, the data protection law of the canton where the processing occurs (e.g. collection, storage, revision, or disclosure) typically applies. This means that, in cases involving multiple locations, more than one cantonal data protection law may apply.
- Research with an international dimension:
  - Often the data protection law at the foreign research location is also applicable.
  - If you process personal data of EU individuals or monitor the online behavior of users based in the EU, GDPR is applicable.

In case of doubt, ask the **Data Protection Officer (DPO)** of the University of Basel: datenschutz@unibas.ch.

University of Basel

Research Data Management Network

# Personal Data Management
# Principles and Requirements for Processing Personal Data

In data protection law, "processing" means any operation which is performed on personal data, whether automated or not. This includes all actions with data, regardless of technology, medium (e.g. paper, digital, ...), or duration (temporary or permanent). Examples of data processing include collecting, storing, preserving, using, modifying, disclosing (e.g. granting access, forwarding or publishing), archiving, and destroying data.

Research involving personal data at a public institution, such as a university, is possible if certain conditions are met and the necessary precautions are taken. The following is a summary of the principles and requirements for processing personal data at the University of Basel:

## Main principles of data protection

- **Rule of Law:** The processing of personal data must be based on a legal basis. The University statutes (§ 1) assign the University of Basel the responsibility of conducting research, which includes processing personal data.
- **Purpose limitation:** Data processing must always be carried out for a specific, stated purpose.
- **Proportionality:** Data processing must be necessary for the intended purpose and proportionate to the impact on privacy.
- **Accuracy (integrity):** Those who process personal data must ensure its accuracy.
- **Perceptability of data processing:** Data subjects must be aware that their personal data are collected and processed.
- **Transparency:** Data subjects must be adequately informed about data processing, including which activities are being performed and their purpose. They also have the right to request information about their data at any time, without justifications.
- **Data security:** The data processing must comply with both technical and organizational security requirements.

University of Basel

Research Data Management Network

## Informed Consent

Informed consent from the data subject is required for specific processing related to an individual, which is typically the case in research.

**The declaration of consent must…:**

- specify the type and scope of the data and data processing
- include a reference to the intended use
- mention any possible disclosure of data to third parties
- list the data subject's rights (e.g. right of access, right to restriction of data processing, right to rectification, blocking or deletion of data)
- be voluntary; with no disadvantages for the data subject in case of refusing consent.

**More information, fact sheets and templates:**
can be found on the website of the University's Data Protection Officer.

University of Basel

Research Data Management Network

# Personal Data Management
## Data Masking

In research, personal data have to be anonymized or pseudonymized as soon as the processing purpose permits. Personal data may only be processed with the informed consent of the person concerned. This guide explains the basic terms.

### Definitions

**Anonymization** means that the reference to a person is irreversibly (= definitively) removed in such a way that it is no longer possible to identify the individual without disproportionate effort. Anonymized data are no longer regarded as personal data.

**Pseudonymization** refers to the removal of personal references, whereby a specific key (i.e., a table that translates pseudonym to person) is retained for the re-personalization of the information. If data is pseudonymized, the conditions under which a person may be identified and how the key is stored must be regulated (key management). Unlike anonymized data, pseudonymized data remain personal data.

**Encryption** is a method of protecting data from unauthorized access. It means that a plain text is converted (encoded) into a non-interpretable character string (ciphertext) by means of an encryption procedure. One or more keys (= codes) are used as crucial parameters for encryption. Secure encryption makes it practically impossible to retrieve the plain text without the appropriate key. However, secure encryption must be done correctly. When in doubt, contact your IT.

**Useful data anonymization tools:**
- ARX
- Argus
- scdMicro
- QualiAnon

University of Basel

Research Data Management Network

## Data Masking Example:

| | ID | Name | Date of birth | Disease | Treatment |
|---|---|---|---|---|---|
| **Collected data** | DS001 | Alan Smith | 01.06.2001 | Appendicitis | Operation |
| **Anonymized data** | DS001 | | | Appendicitis | Operation |
| **Pseudonymized data**<br>• Research Table | DS001 | | | Appendicitis | Operation |
| • Mapping table (Key) | DS001 | Alan Smith | 01.06.2001 | | |
| **Encrypted data** | ZTUVqwEAb Uf+M0tKeQX zQVjnI3VMM vJTdpgB9eIM bzs | xu05js8QCrn 90lj1uaflSzUy bf8uMEvrDT+ KSXvJY24 | lYtVtyocsRqT l6QANxqfDc jxxvRHSWEG jQY54ddSvzE | TOP2AMUh /Fgs2XhoCY 2Zv30qaKP wZ/4A7t1Fp LcRBng | P7Hz796XwSk PXeQvGv2PG cR8XUmG/lC0 kSkv6n/bnlc |

University of Basel

Research Data Management Network

# Data Archiving after Project Completion

When it comes to storing, archiving and publishing research data, misunderstandings often arise due to unclear definitions and inconsistent usage of these terms by different service providers. For example, many repositories refer to "archiving" or even "long-term storage" of data, yet only retain data for a duration of 5 to 20 years.

To ensure clear communication, the Research Data Management Network of the University of Basel has defined the terms as follows:

| Term | Description | Provider | Time |
|---|---|---|---|
| Active storage | Active data management in a filing system within a running research project | IT-Services, sciCORE, SWITCHdrive | 1-5 years |
| Mid-term storage/ deposition/deep storage | Finalized data(sets), to be reused for further (internal) projects | sciCORE: deep storage | 1-10 years |
| Publication | Curated and disseminated (selected) data(sets), understandable and reusable for others | Repositories | 5-20 years |
| Preservation, (long-term) archiving | Highly valuable data and digital objects to be kept forever (e.g. cultural heritage data, singular data such as climate data) | University Library (only for cultural heritage data) | 100+ years |

University of Basel

Research Data Management Network

## How long do I have to keep my data after the end of a project?

- The University of Basel requires its researchers to keep the data for at least 5 years. (Regulation relating to academic integrity at the University of Basel, Art. 4, 8).
- The SNSF recommends a retention period of at least 10 years. (FAQ Preservation and long-term storage of research data).
- All data relating to clinical trials must be retained for at least 20 years after completion or premature termination of the clinical trial (Ordinance on Clinical Trials, ClinO, Art. 45).

In certain cases, retaining data for the long-term is advisable. Long-term archiving involves addressing key challenges related to preserving digital materials, including format obsolescence, media degradation, and evolving technologies.

## Evaluating data for long-term archiving:

Long-term archiving is not suitable for all types of data. The following considerations can help guide the evaluation process.

### Reasons for long-term archiving:

- The data holds significant value and potential future importance for researchers and society.
- The data is unique and irreplaceable (e.g. weather observations or interviews with contemporary witnesses).
- The data has not yet been fully explored or analyzed scientifically.

Long-term archiving is not recommended for data from standard procedures or measurement results that can easily and inexpensively be reproduced as needed.

University of Basel

Research Data Management Network

# Finding and Reusing Research Data

Finding and reusing research data can be a multi-step process that involves identifying suitable datasets through specialized repositories, metasearch engines, or data journals, and critically evaluating their relevance, quality, and reuse conditions.

## Finding research data

Here are some steps to help you find research data:

- Clearly define the research question or topic for which you need data. This will help you identify the specific type of data you require.
- Determine the appropriate sources where you can find research data:
  - **Domain-specific, interdisciplinary or institutional data repositories:** Research data repositories are specialized online platforms where research data can be uploaded and shared together with the related metadata. Re3data (Registry of Research Data Repositories), which offers many filter options, can be used to find research data repositories.
  - **Metasearch engines specialized in research data:** Metasearch engines enable the simultaneous retrieval of several research data repositories. Be aware that none of the metasearch engines will query all available repositories. Therefore, not all relevant datasets may be found. Examples for metasearch engines are:
    - b2find.eudat.eu
    - datacite.org (international, non-profit)
    - Google Dataset search (especially for open government data)
    - Mendeley Data
    - OpenAire (especially data from EU projects)
  - **Data journals:** Check whether there are data journals relevant to your field. Data journals publish articles that describe datasets in detail, but without interpreting them. A list of data journals is provided by forschungsdaten.org.

University of Basel

Research Data Management Network

# Reusing research data

If you discover a dataset that might be relevant to your research question, you can use the documentation and metadata provided with the dataset to evaluate the following key considerations:

- Who compiled the data? (assess the reliability of the source)
- How and for what purpose were the data generated?
- Under what conditions were the data generated?
- How is the dataset structured?
- Are the data suitable for answering my research question?
- Do I have access to the data?
- Can and may I combine or enrich the data with other datasets?

Remember to comply with all applicable data usage guidelines, obtain any necessary permissions or licenses, and ensure your use of the data adheres to ethical and legal requirements.

## Citation

If you reuse a dataset for a research project, it is important to cite it properly. The citation should include the following information:

- Author
- Title of the dataset
- Year of publication
- Repository
- Version (if indicated)
- Persistent identifier

## Persistent identifiers

Persistent identifiers allow scientists and repositories to uniquely identify persons or cite datasets. Here are some examples of widely used persistent identifiers:

- **DOI: Digital Object Identifier**
  A unique alphanumeric string assigned by a registration agency (e.g. DataCite, Crossref) to identify content and provide a persistent link to its location on the Internet.
- **ARK ID: Archival Resource Key Identifier**
  ARK ID is particularly useful for hierarchically structured data objects.
- **ORCID: Open Researcher and Contributor Identifier**
  It provides an identifier for authors as they engage in research, scholarship, and innovation activities. It is highly recommended that you also create an ORCID for yourself.

University of Basel

Research Data Management Network

# Moving On: Managing Data at the End of a Project or Upon Leaving the University

If you have submitted your PhD thesis, completed a research project and/or are leaving the University of Basel, it is essential to address key questions regarding your datasets. We recommend to start planning a few months ahead to ensure your data is transferred and/or stored appropriately. This includes coordinating with your project leader/supervisor and IT services.

Important note: Research data generated within the framework of research projects at the University of Basel generally remain the property of the University of Basel.

## Checklist to guide you through the process

### Update the Data Management Plan (DMP)

- ☐ Ensure the DMP reflects the current status of your data, and future needs such as permissions and access controls for collaborators or supervisors, or long-term storage solutions.

### Clean up data on the storage infrastructure

- ☐ Copy all personal documents to a private storage.
- ☐ Remove redundant files (e.g. drafts and earlier versions of published papers).

### Review and enhance data documentation

- ☐ Update the README files.
- ☐ Compile any additional documentation required to understand and reuse the data.

University of Basel

Research Data Management Network

## Identify data that can or must be published

- ☐ Consider the legal obligations and the requirements of your funding institutions and publishers.
- ☐ Organize the publication of eligible datasets in an appropriate repository.
- ☐ Include comprehensive documentation and metadata for published data.

## Plan for data preservation

- ☐ Check the retention period specified by the university and funding agencies (e.g., the SNSF currently recommends 10 years).
- ☐ If applicable, arrange for data deletion after the required retention period.
- ☐ Assign a responsible person in charge of long-term data preservation.
- ☐ Choose a suitable storage solution for the data to be preserved.
- ☐ If you need future data access, discuss access arrangements with you project leader/ supervisor and IT.

## Organize data transfer

- ☐ If data must be transferred, plan and execute the transfer process in coordination with relevant stakeholders.

After having completed these steps, arrange a handover appointment with your project leader/supervisor. Ideally, both parties sign a confirmation that the data have been handed over appropriately. Such a document could also outline agreements regarding future use of the data.

More detailed information can be found in the document on Management of Digital Research Data at the End of a Project, developed by the RDM Network of the University of Basel.