Guidelines for Research Data Management
at the University of Bayreuth
last updated on 26.06.2023, FDM@UBT

Preliminary remarks

The University of Bayreuth (UBT) recognises the fundamental importance of research data and its documentation in order to ensure quality-oriented and compatible research and scientific integrity. The Guidelines for Research Data Management (RDM) at UBT are based on UBT's statutes on ensuring standards of good scientific practice and dealing with scientific misconduct[1] and were issued on 26.06.2023 by the Vice President for Digitalization, Innovation & Sustainability (VPN) and the Vice President for Research & Junior Scholars (VPF).

The present Guidelines supplement the **"Research Data Management Policy of the University of Bayreuth"[2]** and provide practical advice on its implementation. All web links mentioned were last accessed on 31.05.2023.

General notes

All data that is generated, collected, observed, simulated or derived in the research process is considered to be research data. Typical examples of research data are laboratory values, measurement data, audiovisual information, survey results, objects from collections, texts, methodological test procedures or simulations, source codes or protocols[3]. During the lifetime of research projects, research data can occur in different forms (format variants of primary data, processed data, published data) and be provided with different access rights, e.g. as open data, access-restricted data, or non-public data. For the subsequent use of research data, it is necessary to document the context of creation and the methods or tools used in the form of metadata. When publishing research data, the use of suitable, preferably subject-specific, data centres or repositories contributes to greater visibility and the possibility of subsequent use of the research results. In line with UBT's Open Access Strategy[4], UBT supports the principles of "Open Data"[5].

Contact

UBT supports researchers with suitable services to help them fulfil the RDM policy of the university. Under the leadership of VPN, VPF, CIO and the working group FDM@UBT, various service offerings have been designed. These will be further developed as needed and integrated into the existing service infrastructure. In doing so, UBT orients itself along national and international developments in RDM and cooperates with other higher education institutions. The FDM@UBT working group is made up of staff from three central institutes, namely IT Service Centre (ITS), University Library (UB) and Research Support Office.

Central technical infrastructure services are RDMO@UBT[6], a web tool for creating data management plans (DMPs), RDSpace@UBT[7], the institutional research data repository of UBT, and Collections@UBT[8], the university's collection management system. Information on RDM tools and services[9] at UBT as well as the

[1] https://www.amtliche-bekanntmachungen.uni-bayreuth.de/de/amtliche-bekanntmachungen/2022/2022-050.pdf
[2] https://www.forschungsfoerderung.uni-bayreuth.de/pool/dokumente/20161108_UBT-Leitlinien-Forschungsdaten-Management.pdf
[3] https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf
[4] https://www.ub.uni-bayreuth.de/de/download/openaccess-strategie.pdf
[5] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en
[6] https://rdmo.uni-bayreuth.de
[7] https://rdspace.uni-bayreuth.de
[8] https://collections.uni-bayreuth.de
[9] https://www.fdm.uni-bayreuth.de/de/Services/index.php

contact details of FDM@UBT[10] and consulting options are available via the group's web pages[11]. For enquiries in the RDM context, researchers can also contact the central email address fdm@uni-bayreuth. The group's consulting services are supplemented by training courses, e.g. at UBT's graduate schools or at individual departments. In accordance with the division of tasks within the FDM@UBT working group, the **Research Support Office** advise on the RDM requirements of funding organisations and on presenting RDM in project proposals, while the **University Library** provides advice on metadata and research data publications. The **IT Service Centre** advises on technical data management, the use of suitable tools, the procurement of infrastructure and the use of central technical services.

Recommendations for handling research data

Research data management (RDM) and the research data infrastructure at UBT are based on the FAIR principles[12] , which are internationally recognised principles for handling research data so that they are **F**indable, **A**ccessible, **I**nteroperable,and **R**eusable.

UBT recommends that the following aspects regarding the handling of research data be observed in the different phases of a research project:

1.   Planning phase: Before the start of the project

When planning a research project in which data are collected or data form the basis of the research, it is advisable to deal with aspects of RDM as early as possible. For this purpose, researchers can use the advisory services offered by the FDM@UBT group to develop a suitable strategy for the project. Aspects that should be considered in the planning phase are:

- Data management requirements in project applications, expectations of funding bodies
- Creation of data management plan
- Possibilities for publication or subsequent use of the research data
- Securing and long-term availability/archiving of research data

The **preparation and submission of a data management plan (DMP)** is recommended and required by some funding organisations. A DMP serves as a structured guideline for ensuring sustainable handling of the project's data during the project and beyond. All essential aspects of the life cycle of the data, from collection to long-term preservation, should be documented in such a way that the resulting data is comprehensible and interpretable, and that data reusability and long-term availability as well as citability is guaranteed.

The DMP is thus a central document of **quality assurance across all phases of the project** and should be adapted and updated according to the relevant requirements in the course of the project. **Data protection aspects** resulting from the requirements of the General Data Protection Regulation (GDPR/DSGVO) and ethical aspects should also be documented in the DMP. In the case of externally funded research, the requirements of the funding organisations on data management should be taken into account. The following questions can provide helpful **for the creation of a DMP:**

- What is the aim of the research project?
- Who is responsible for handling research data in the project?
- What research data is collected? Can data be re-used, and if so, how and when?
- Which (subject-specific) standards are used (e.g. data formats, metadata, ontologies)?
- How is the storage, securing, archiving and, if necessary, access to research data organised and documented?
- What agreements are there on the use and copyright aspects of the research data?

---

[10] https://www.fdm.uni-bayreuth.de/de/Ihre-Ansprechpartner/index.html
[11] https://www.fdm.uni-bayreuth.de/de/index.html
[12] Wilkinson, M. D. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci.Data 3*:160018 (2016). https://doi.org/10.1038/sdata.2016.18

- Are quality controls of the research data provided for and if so, how?
- Are there legal, temporal, ethical or other restrictions on the accessibility and re-use of research data?
- What resources are required for RDM in the project?

To support researchers at UBT in creating, versioning and saving the DMP, the web tool **RDMO@UBT**[13] is available. It contains templates with questions that can be answered in a step-by-step workflow. RDMO@UBT can be used for DMPs of both individual and collaborative projects. Please refer to FDM@UBT for all questions related to the DMP.

Research funding providers generally expect not only the **long-term preservation** of research data, but also data **publication,** insofar as no legal, ethical or other reasons stand in the way. A suitable strategy should therefore already be considered in the application phase and documented in the DMP. If the funding agency offers the possibility of doing so[14], it is advisable to apply for **costs** that go beyond the university's basic resources (for example, costs for processing the research data for subsequent use, for transferring the data to a public repository or repository membership fees). When planning and applying for the necessary equipment and IT services, the ITS should already be involved in the preparation phase of applications.

In all research projects, especially in the case of a planned publication of the findings, the applicable legal framework conditions, e.g. copyright concerns and legitimate interests of third parties, must be observed. It is therefore advisable to clarify **legal issues** in advance when planning a research project and to obtain any necessary permits. For the collection of sensitive research data, for example in the social sciences, life and health sciences, or with regard to dual use research, the **vote of the Research Ethics Committee**[15] of UBT needs to be obtained in order to protect the rights and safety of subjects, following the Declaration of Helsinki[16]. For personal data, the provisions of the GDPR also apply. If you have any questions about data protection issues, you can contact the university's data protection officer[17].

The use of research data generated in a project is primarily due to the researcher(s) who conducted the research. In case of research projects involving several cooperation partners, **agreements on the rights of use** of the research data and findings must be documented as early as the application phase. If the researcher collecting the data leaves UBT, the further handling of her/his research data must be clarified and documented.

To support the accessibility, interoperability and **long-term availability of research data**, the use of open, non-proprietary file formats is generally recommended.

2. Implementation phase: During the research project

The documentation of research findings includes all information relevant to their achievement, in accordance with professional practice. This includes used or emerging research data, materials, applied methods as well as self-programmed software (source code). The documentation includes all results - even those that do not support the research hypothesis - and must be protected against manipulation.

For the **processing and storage of research data,** common procedures in the research discipline should be used. This includes, in particular, the observance of data security with regard to availability, integrity (unalteredness) and authenticity. This can also include the use of data security and archiving, the use of secure data exchange platforms and documentation of the versioning process, e.g. by using versioning tools. It is advisable to continuously test the data security strategy during the project and to adapt it if necessary.

The DMP created in the planning phase serves as a reference in the active project phase and must be adapted and updated if necessary to changing framework conditions and requirements (in the sense of a "*living document*").

---

[13] https://rdmo.uni-bayreuth.de
[14] This option is offered by the DFG, for example: Information on funds that can be applied for
[15] https://uni-bayreuth.de/ethikkommission
[16] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/
[17] https://www.uni-bayreuth.de/beauftragte

In the implementation phase of a project, data sets can be subject to a multi-stage development (e.g. through selection, aggregation, integration). To ensure traceability, it is advisable to mark and document the **different versions** and to keep them at least for the active duration of the project. Especially for text-based data, the use of versioning tools, which are common in software development (e.g. Git[18], SVN[19]), facilitates the management of the different versions.

For research data to be reusable, it must be **described with metadata**. Metadata describe the context in which the research data was generated, who collected the data or under which conditions they can be re-used. As a rule of thumb, metadata *should answer the classic six questions (*Who? What? Why? How? When? Where?). (Subject-specific) metadata make it possible to find the research data, to assess their usability for one's own research question and to interpret the research data correctly. Ideally, the description should be structured and machine-readable[20]. If established metadata standards are available in the respective specialist community, it is advisable to use them to promote findability, visibility and reusability. Alternatively, generally applicable standards such as Dublin Core[21], MARC[22] or MODS[23] can be used.

In **collaborative research projects** or with large amounts of data, the use of dedicated working environments and portals for data management is a good option. Operating these tools usually requires additional resources but offers the advantage of uniform and centralised data management. This makes it easier to find and exchange data, but it should be regulated with the help of a coordinated data management scheme.

For the documentation of research processes, **UBT members** can use the interdisciplinary electronic laboratory notebook **SciNote@UBT[24]**. In SciNote@UBT, research activities can be managed and documented, and experimental data can be linked or stored in one place. In accordance with UBT's statutes on safeguarding standards of good scientific practice and dealing with scientific misconduct[25], research data is stored for a period of at least 10 years. The following applies to the use of SciNote@UBT:

- The documentation of research activities in SciNote@UBT is based on several levels (team → projects → experiments → tasks → protocols and associated results). All levels are provided with specific metadata, in part automatically.

- A rights management with user roles at team and project level can be used to regulate administration, joint editing, data exchange and, if necessary, inspection.

- In accordance with the rules of good scientific practice, data that is stored as final results of a research project can no longer be changed in SciNote@UBT.

- In addition, an integrated audit trail and versioning contributes to traceability and compliance with legal requirements.

If you have any questions about SciNote@UBT, the team of FDM@UBT[26] will be happy to help.

In the research data environment, the **ITS** provides the following **services,** among others, some of which are offered free of charge, or can be made available at marginal costs[27]:

- Use of network file systems (incl. data backup), e.g. for storage and data exchange

- Archive storage services on tape drives

- Block storage services for servers (virtual hard disks via a dedicated storage network)

- Provision of virtual root servers (server hosting)

---

[18] https://git-scm.com
[19] https://subversion.apache.org
[20] https://www.dcc.ac.uk/guidance/standards/metadata/list
[21] http://dublincore.org
[22] https://www.loc.gov/marc/
[23] https://www.loc.gov/standards/mods
[24] https://scinote.uni-bayreuth.de
[25] https://www.amtliche-bekanntmachungen.uni-bayreuth.de/de/amtliche-bekanntmachungen/2022/2022-050.pdf
[26] https://www.fdm.uni-bayreuth.de/de/Ihre-Ansprechpartner/index.html
[27] Marginal costs are the additional costs incurred when supply is expanded by a marginal unit.

- ▪ Accommodation of real servers (server housing)
- ▪ Data exchange services (DFN Cloud)

You can find more detailed service descriptions on the ITS website[28].

The **University Library** provides the following **services** in the research data environment:

- ▪ Advisory service for the long-term availability of research data

In addition, the UB is committed to the long-term availability of research data in cooperation with the Bavarian Library Network, as part of their project "Digital long-term availability for science and culture in Bavaria": In the spirit of long-term availability and good scientific practice, the aim is to preserve research data even beyond a period of 10 years. This is not only about the physical preservation of data, but also about taking into account other factors that ensure the long-term interpretability and readability of the data. The long-term availability of research data can be made possible by the Bavaria-wide cooperative infrastructure currently being set up at UBT.

3. *Closing phase: After completion of the research project*

Researchers are responsible for deciding whether, how and where to make their findings **publicly available,** taking into account the conventions of the respective field. Depending on the legal or contractual conditions, the researcher must choose the time, the scope and a suitable licence for the publication of the research data.

In the spirit of good scientific practice, research data must be saved and made accessible in the long term by the end of the project at the latest. It is recommended that research data be published according to the principle "***as open as possible, as closed as necessary***[29]". In particular, research data and materials on which a publication is based should be made available in recognised (subject-specific) data centres or repositories in accordance with the FAIR principles[30], provided that no data protection, research ethics, copyright or patent protection regulations prevent this.

- ▪ For the **provision of research data to be published,** the DFG advises in their Guidelines on the Handling of Research Data (2015)[31]: *"Data should be made accessible at a stage of processing that allows it to be usefully reused by third parties (raw data or structured data)".*

- ▪ **Research data** should be published as soon as possible. If there are good reasons, data can be published with an embargo period. Possible requirements of funding organisations or repositories must be taken into account. In the case of primary publication of research data via websites, long-term availability and identification via a unique and persistent identifier (PID) is often only possible to a limited extent. Research data should therefore be recorded in suitable data centres/repositories.

- ▪ Directories such as the Registry of Research Data Repositories (re3data[32]), the DFG Portal for Research Infrastructures (RIsources[33]), or the Directory for International Open Access Repositories (OpenDOAR[34]) can be helpful in the **search for suitable data** centres/repositories.

- ▪ As with scientific articles, research data should be assigned a **unique and persistent identifier (PID)** upon publication. The assignment of a PID (e.g. DOI, URN) is usually done by the repository or via the data journal in which the data are published. The use of PIDs ensures the long-term findability and citability of the data set. This makes the origin of the data traceable and the data source / data provider is credited.

- ▪ In order to regulate the **subsequent use of research data**, data should be published under a suitable licence. The prerequisite for granting a licence is the ownership of the rights to the data. Possible requirements of funding organisations or repositories must be taken into account. Creative

---

[28] https://www.its.uni-bayreuth.de/de/index.html
[29] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en
[30] https://www.nature.com/articles/sdata201618 or https://www.go-fair.org/fair-principles/
[31] http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf
[32] https://www.re3data.org/
[33] https://risources.dfg.de/
[34] http://v2.sherpa.ac.uk/opendoar/

Commons (CC) licences are widely used as open licences for research data[35] (suitable from version 4.0). Well-known free licences from the software environment are the GNU licences[36], the MIT licence[37] or the Apache licence[38].

- If **exploitation rights to research data** are transferred to third parties in connection with a publication, care should be taken to ensure that access to the data remains guaranteed.

- The **archiving of research data is** based on the recommendations of the DFG[39] for retention for at least 10 years after the end of the project. Shorter retention periods for parts of the data or comprehensible reasons for not retaining certain data must be explained in the DMP.

The interdisciplinary repository **RDSpace@UBT[40]** is available to **members of UBT** for publishing their research data. RDSpace@UBT is where research findings are stored, i.e. consolidated data and all information (such as scripts, spectra, calculations) necessary to reproduce the result. In accordance with UBT's statutes for safeguarding the standards of good scientific practice and for dealing with scientific misconduct[41], research data is stored for a period of at least 10 years. The following applies to the use of RDSpace@UBT:

- All data in RDSpace@UBT are provided with metadata (standard format Extended Dublin Core).

- All data sets automatically receive a persistent internet address (DOI).

- Various free licences can be assigned to the data sets.

- Using the DOI, related research data and publications can be linked to each other in RDSpace@UBT and then refer to each other.

- In accordance with the rules of good scientific practice, published research data in RDSpace@UBT can no longer be changed; this preserves citability and traceability.

- RDSpace@UBT has a versioning feature where new versions are published while the previous ones remain available. Each new version receives a new DOI. Previous and current versions are automatically linked and refer to each other.

- RDSpace@UBT is committed to Open Access. The metadata are freely accessible on the internet and are further disseminated and made searchable via standard interfaces (Google Scholar, etc.). The research data itself can be embargoed.

Some subject communities have built **subject-specific research data infrastructures** in recent years. These can offer advantages over institutional repositories such as RDSpace@UBT, such as subject-specific metadata schemes and subject-specific search options. If there is a repository that is recognised in the subject community, it is advisable to use it.

The following criteria should be considered when **selecting a trusted repository**[42]:

- Long-term availability (at least 10 years)

- Allocation of persistent identifiers

- Allocation of meta-data

- Information on data access and licensing

- Acceptance and visibility of the repository in the subject community

- Cost-free (if not: is there a clear cost regulation?)

---

[35] https://de.creativecommons.org
[36] https://www.gnu.org/licenses/
[37] https://opensource.org/licenses/MIT
[38] https://www.apache.org/licenses
[39] http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf
[40] https://rdspace.uni-bayreuth.de
[41] https://www.amtliche-bekanntmachungen.uni-bayreuth.de/de/amtliche-bekanntmachungen/2022/2022-050.pdf
[42] https://www.scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/

In addition to subject-specific repositories, there are also interdisciplinary repositories. The portal **re3data.org**[43] offers a good overview with extensive search and filter functions.

Research data can also be published in subject-specific or interdisciplinary data journals. In some disciplines, the publication of data as a supplement to the scientific article is also established. With this form of data publication, the data can only be found via the article and are therefore not independent, citable publication objects. Journals are also increasingly recommending the storage of research data in a repository. Here it is advisable to consider possible costs at an early stage.

If you have any questions about RDSpace@UBT, repositories or data journals, the team of the FDM@UBT working group[44] will be glad to help.

---

[43] https://www.re3data.org
[44] https://www.fdm.uni-bayreuth.de/de/Ihre-Ansprechpartner/index.html