

Technische
Universität
Berlin



Guidelines on the implementation of the Research Data Policy of the Technische Universität Berlin

23 October 2019

Service Center Research Data Management



Content

Preface	3
Point of Contact	3
Guidelines on the Handling of Research Data	4
I. Planning phase: before the research project	4
II. Execution phase: during the research project	5
III. Completion phase: after the research project	6

Preface

The Technische Universität Berlin considers research data as a valuable resource and an essential basis for scientific knowledge. On 23 October 2019, the Academic Senate of TU Berlin adopted the “Research Data Policy of the Technische Universität Berlin” pursuant to the “[Statute on the Safeguarding of Good Academic Practice at TU Berlin](#)”¹ and the “[TU Berlin’s Open Access Policy](#)”² and in accordance with the “[TU Berlin Transfer Strategy](#)”³, the “[Recommendations for Action for Knowledge and Technology Transfer in the Context of Open Science](#)”⁴ and the “[Code of Conduct for Research Involving Commercial Enterprises](#)”⁵. In the same meeting, the Academic Senate approved the “Guidelines on implementing the Research Data Policy of TU Berlin” drafted by the Service Center Research Data Management. TU Berlin supports its researchers with suitable offers for research data management according to financial possibilities. The guidelines⁶ supplement the principles described in the Research Data Policy of TU Berlin and provide practical advice for their implementation.

Point of Contact

The [Service Center Research Data Management \(SZF\)](#)⁷ is the central point of contact for all issues related to research data management at TU Berlin. Within SZF, the University Library, the Center for Campusmanagement (ZECM, formerly tubIT), and the Department V Research work together and bundle their competences to support the University’s researchers in handling research data. SZF is managed and coordinated by the University Library. SZF operates the research data infrastructure at TU Berlin, which is integrated into the University’s IT infrastructure, and develops further services as needed. The central technical infrastructure services include DepositOnce, the repository for research data and publications of TU Berlin, and TUB-DMP, a web tool for creating data management plans. Advisory services and a help desk complement the technical services. The research data management and the research data infrastructure of TU Berlin are aligned with the FAIR principles⁸, international guidelines for the optimal management of research data so that data is findable, accessible, interoperable, and reusable.

The SZF webpages serve as a central platform providing comprehensive information on the handling of research data and the services for research data management as well as advisory offers and contact information for the respective contact persons. Corresponding to the task sharing of SZF, the teams of [Department V Research](#)⁹ (see below) advise you on the requirements of funding organizations and project proposals. The [SFZ team](#)¹⁰ of the University Library advises you on the handling of research data and research data management. [ZECM](#)¹¹ advises you on the acquisition of infrastructure and the use of its services.

¹ https://www.tu-berlin.de/menue/ueber_die_tu_berlin/gesetze_richt_leitlinien/grundsaeetze_zur_sicherung_guter_wissenschaftlicher_praxis_an_der_tu_berlin/parameter/en/.

² https://www.tu-berlin.de/menue/ueber_die_tu_berlin/gesetze_richt_leitlinien/open_access_policy_der_tu_berlin/parameter/en/.

³ Transferstrategie TU Berlin. https://www.forschung.tu-berlin.de/fileadmin/f22/Transferstrategie_TU_Berlin.pdf (in German).

⁴ ZfgE/VD Handlungsempfehlungen Wissens- und Technologietransfer im Kontext von Open Science – Stand Dezember 2018. https://www.forschung.tu-berlin.de/fileadmin/f22/Einrichtungsdaten/V_D_allgemein/Handlungsempfehlungen_WTT_OpenScience_1811_KB.pdf (in German).

⁵ https://www.forschung.tu-berlin.de/fileadmin/f22/Einrichtungsdaten/V_D_allgemein/Code_of_Conduct-english.pdf.

⁶ These guidelines are based on the guidelines for research data policies of the Humboldt-Universität zu Berlin (2014): https://www.cms.hu-berlin.de/de/dl/dataman/hu-rdm-guidelines/at_download/file and the Friedrich Schiller University Jena (2016): <https://www.researchdata.uni-jena.de/fdmmedia/dokumente/fdm+politics/guidelines+and+recommendations+on+research+data+management.pdf>.

⁷ https://www.szf.tu-berlin.de/menue/ueber_das_szf/parameter/en/.

⁸ 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>.

⁹ https://www.forschung.tu-berlin.de/menue/abteilung_v/parameter/en/.

¹⁰ Help desk: team@szf.tu-berlin.de.

¹¹ <https://www.campusmanagement.tu-berlin.de/zecm/parameter/en/>.

Guidelines on the Handling of Research Data

Below are guidelines on the handling of research data for the different phases of a research project.

I. Planning phase: before the research project

In case of a **third-party funded project**, inform yourself in advance of any applicable guidelines regarding the reuse of research data originating in the project. Costs arising for the long-term archiving of data that extend beyond the University's basic infrastructure can and should be part of the funds applied for. Use the Research Department's advisory offers to learn about the different requirements of funding organizations and funding possibilities and to determine your individual needs.

- Advising on proposals for third-party funding and on DFG, BMBF, and joint projects: [Research Promotion Section](#)¹²
- Advising on EU projects: [EU Office](#)¹³
- Advising on doctorates: [TU-DOC – Office for Doctoral and Postdoctoral Services](#)¹⁴
- Contract negotiations in research and development projects, research cooperation, services, licensing contracts, consultancy contracts, etc.: [Research Contracts, Patents and Corporate Investments](#)¹⁵
- Advising on handling personal data in a project: [Data Protection Officer of TU Berlin](#)¹⁶

In order to receive adequate support for your new research project, it is recommended to announce the project during the proposal phase to the Research Department. This project announcement is achieved by using the [electronic project notification](#) (ePA). If you have any questions concerning the announcement, the [Research Promotion Section](#) will assist you.

Furthermore, for any project in which data is collected or forms the foundation of your research it is strongly recommended that you make an early examination of the requirements and possibilities of efficient and sustainable **research data management**. Contact the [SZF team](#) to find out about the handling of research data and about services for research data management and to develop a suitable strategy for your project.

The creation of a **Data Management Plan (DMP)** is recommended and increasingly required by funding agencies. A data management plan is a structured guideline outlining how to handle research data both during the project and after its completion. It documents the process how the research data is generated and how it is stored appropriately so that in later years it can be interpreted and verified and remains available, authentic, citable, and reusable.

During the proposal phase, a strategy for the sustainable archiving and availability of the research data should be determined and included in the data management plan. Clearly defined legal parameters and suitable security measures (such as contracts and licenses) for later use are also to be defined in the data management plan. To optimize research data management and as a basis for institutional assistance, a DMP should be created before the start of the research project and updated over the course of the project (keyword: *living document*).

For individual as well as for joint projects, you can use the web tool [TUB-DMP](#)¹⁷ that will assist you with the creation, versioning, and long-term storage of your data management plan. It contains templates in

¹² https://www.forschung.tu-berlin.de/servicebereich/menue/servicebereich_forschung/parameter/en/.

¹³ https://www.forschung.tu-berlin.de/eu_buero/menue/ueber_uns/parameter/en/.

¹⁴ <https://www.tudoc.tu-berlin.de/menue/nachwuchsbuero/parameter/en/>.

¹⁵ https://www.forschung.tu-berlin.de/forschungsvertraege_lizenzen_und_patente/menue/forschungsvertraege_lizenzen_und_patente/.

¹⁶ <https://www.tu-berlin.de/asv/menue/datenschutz/>.

¹⁷ https://www.szf.tu-berlin.de/menue/dienste_tools/datenmanagementplan_tub_dmp/parameter/en/.

the form of checklists with relevant questions, which you can answer in a step-by-step workflow as they pertain to your project. The tool also includes a template for creating a data management plan that is compliant with Horizon 2020. The [SZF team](#) will assist you with questions about your data management plans.

In all research projects and also in the publication of the research results **legal frameworks** must be observed. Certain research data, for example in the social or life sciences as well as in medicine, are subject to strict requirements, such as data privacy or previous review by an ethics committee. Copyright and the protection of third-party interests must also be ensured. For this reason, fundamental legal questions are to be settled in advance when planning a research project. The section “[Research Contracts, Patents, and Corporate Investments](#)” in Department V can assist you with this. To learn about data protection requirements, contact the [Data Protection Officer of TU Berlin](#) early on.

II. Implementation phase: during the research project

State of the art processes are to be used for the **storage and processing** of research data as well as for collaboration based on this data. This particularly includes compliance with data security as regards the availability, integrity, and authenticity of data. This requires, for instance, the use of data backup and protection, secure data exchange platforms, and versioning tools.

In the implementation phase of a project, datasets may evolve several stages (e.g. by selection, aggregation, integration). It is a good practice to label, document and keep the different versions at least for the duration of the project. Especially in case of text-based data, the use of versioning tools, as commonly used in software development (e.g. git, SVN), helps with the management of different **versions**.

ZECM provides the following services for research work at the University; further information can be found on the [ZECM website](#)¹⁸:

- Use of network file systems (incl. data backup)
- Archive storage services on tape drives
- Provision of virtual root servers (server hosting)
- Accomodation of real servers (server housing)
- Block storage services for servers (virtual hard disks over a dedicated storage network)
- Data exchange services
- Versioning services

These services are either included as basic equipment and are free of charge or offered at cost price. The same applies to the [collaborative working tools](#)¹⁹ provided by ZECM.

Different scientific disciplines and their research domains apply different methods in handling research data. This makes comprehensive recommendations for the concrete procedures hard to define. Therefore it is generally recommended to become acquainted in advance with the established **data formats, software, and standards** that are used in your scientific community for the documentation and annotation of research data (e.g. ontologies, controlled vocabulary, or metadata schemes). Using open, non-proprietary file formats supports the access to and long-term availability of research data.

Describing research data with **metadata** is fundamental for the reusability of research data. Metadata is data about data and describes the context in which the data was created. As a rule of thumb, metadata should answer the classic six questions: Who? What? Why? How? When? Where? Metadata are a prerequisite for enabling potential subsequent users to find data and assess its suitability for the intended use. Ideally, the description is structured and machine-readable. For this purpose, metadata

¹⁸ <https://www.campusmanagement.tu-berlin.de/menue/dienste/parameter/en/>.

¹⁹ https://www.szf.tu-berlin.de/menue/dienste_tools/kollaboratives_arbeiten/parameter/en/.

standards²⁰ and standardized terminologies exist in most disciplines. If these do not exist, generic standards, such as [Dublin Core](#)²¹, should be used to describe the data. They are developed and promoted by worldwide initiatives and help to make research results better findable and interoperable.

In collaborative projects or projects with large amounts of data, the use of **dedicated work environments** and portals for data management is advisable. Operating these infrastructures usually requires additional resources, but they provide the advantage of a uniform and central management for research data. Finding and sharing research data is thereby facilitated, but should be governed within the project consortium by a project-specific data policy.

III. Final phase: after completion of the research project

According to good academic practice, by the end of the project research data is to be stored and, if possible, made accessible, if there are no contradictory contractual, ethical, or legal regulations. Many funding organizations now place particular importance on **accessibility**, in order to enable the verification of the research results and the reuse of the research data. In keeping with its Open Access Policy, TU Berlin supports open access to research data. When publishing research data, TU Berlin recommends to follow the principle, “Accessible if possible, restricted if necessary”.

The following basic principles regarding the **publication of research data** should be observed:

- Individual websites (e.g. of projects, working groups, academic chairs, employees) are generally not a suitable location for the publication of research data. The long-term availability of such websites is often not ensured and the unique identification (keyword: persistent identifier) is usually not possible.
- When selecting data to be published, the DFG recommends (2015): “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).”²² In particular, data which forms the basis of a scientific article should be made accessible, if there are no contradicting data protection, legal or research ethics regulations.
- As is the case for scientific articles, research data should be assigned a unique persistent identifier (PID) upon publication. In this way, research data can be found and cited independently of a publication. Well known examples are DOI (digital object identifier) or URN (uniform resource name).
- To regulate the rights of use and utilization of research data, data should always be published with an appropriate license. The choice of a license should at least allow open access for scientific purposes. Any special requirements of the funding organization or of repositories are to be observed. Established free licenses in software are the [GNU General Public License \(GPL\)](#)²³, the [MIT license](#)²⁴ or the [Apache license](#)²⁵. [Creative Commons licenses](#)²⁶ are standard for texts, images, music and videos.

As a member of TU Berlin, you and your cooperation partners can use the interdisciplinary repository of TU Berlin, DepositOnce, to publish your research results (research data and publications). Research results, meaning consolidated data and all information needed to reproduce these results (such as notes, time histories/recordings, calculations, etc.) are stored in DepositOnce. Pursuant to the Statute on the Safeguarding of Good Academic Practice of TU Berlin, research data is stored for at least 10 years.

- All data in DepositOnce is provided with metadata (standard format Extended Dublin Core).

²⁰ See, for example <http://www.dcc.ac.uk/resources/metadata-standards>.

²¹ <https://dublincore.org/>.

²² DFG (2015), “Guidelines on the Handling of Research Data”, https://www.dfg.de/download/pdf/foerderung/antwortstellung/forschungsdaten/guidelines_research_data.pdf.

²³ <https://www.gnu.org/home.en.html>.

²⁴ <https://opensource.org/licenses/MIT>.

²⁵ <https://www.apache.org/licenses>.

²⁶ <https://creativecommons.org>.

- All datasets automatically receive a persistent Internet address (DOI).
- Various free licenses can be assigned to the datasets.
- Via the DOI, related research data and publications in DepositOnce can be linked to each other and then refer to each other.
- In accordance with the rules of good scientific practice, published research data can no longer be modified in DepositOnce. This maintains the data's citability and verifiability. DepositOnce utilizes a versioning in which new versions can be published while previous versions remain available. Every new version receives a new DOI; previous and current versions are automatically linked to one another and refer to each other.
- DepositOnce is committed to Open Access. The metadata are publicly accessible on the Internet and are broadly distributed and made searchable via standard interfaces (Google Scholar, etc.). An embargo can be placed on the research data itself.

In recent years, scientific communities worldwide have built discipline-specific research data infrastructures in which TU Berlin researchers are also involved. Meanwhile there is a large number of **discipline-specific repositories**. These may offer advantages compared to DepositOnce, such as discipline-specific metadata schemes and specific search options. If you already use a repository in your scientific community, you should continue to do so. The same applies if it seems reasonable to publish the research data in a discipline-specific repository and such a repository exists for your discipline. It is also possible to publish the data in special data journals. In some disciplines, it is common to publish data as a supplement to the respective article. However, this form of data publication has the disadvantage that the data can only be found via the article and does not form an independent, citable publication object.

When choosing a discipline-specific repository, the following criteria should be observed: long-term availability (at least 10 years), allocation of a persistent identifier (e.g. DOI, URN), licenses and usage rights of the data, reputation and visibility, costs. The portal [re3data.org](https://www.re3data.org/)²⁷ offers a helpful overview with comprehensive search and filter functions when you are searching for a suitable discipline-specific repository for your research data. The [SZF-Team](#) will assist you with questions regarding DepositOnce and repositories.

Research data is to be **published as soon as possible**. If applicable reasons exist, an embargo can be placed on data in DepositOnce. In this case only the metadata are published; the data itself is stored in the repository and is only visible after expiry of the embargo. Interested persons can request the data via email during the embargo. The embargo is determined by the responsible researchers whereby the requirements and guidelines of research funding agencies and repositories must be observed. Embargo periods should not exceed a maximum of 5 years after the project end. An embargo must be justified, for example in a file in the repository which also includes the expiration date of the embargo.

²⁷ <https://www.re3data.org/>.