Guidelines on the Handling of Research Data within the Leibniz Association

Adopted by the General Assembly of the Leibniz Association
on 29 November 2018.

I. Preamble

Research data forms the basis of scientific work. The increasing digitalisation of scientific processes calls for new approaches in the area of research data handling. Simply publishing the conclusions resulting from the analysis of collected research data is no longer sufficient. Instead, access to research data is becoming an increasingly important resource for researchers making use of the new opportunities offered by digital tools. Ensuring data access and interpretability creates a range of challenges on a number of different levels (for research funding bodies, research institutions and researchers and research support staff) and often requires a specific approach for each discipline. As a network of German research institutions in a range of different disciplines, the Leibniz Association emphasises the importance of a responsible and transparent approach to handling research data as part of a sustainable and quality-oriented research process. The proposals and measures set out below aim to promote the development of structures and processes within the Leibniz Association that meet this requirement.

II. Research data

Research data includes all data that is produced during research and that can be processed digitally. It takes a range of forms – depending in particular on the domain in question. It may include, for instance, measurement, (ongoing) survey and observational data, process-produced data, texts, data from polls, graphic visuals, software and simulations. Among other things, this research data includes raw data, aggregated data, metadata and descriptions of data structures.

Research data management is an active task and supports the activities linked to scientific discovery: the planning, generation, documentation, processing, archiving and, where relevant, publishing, of the data. Research data management is part of good scientific practice and ensures the quality, reproducibility, availability and reuse of research data and results. This kind of systematic management of research data is therefore a basis for scientific discourse and for maximising the data's reuse potential.

III. Conditions

The Leibniz Association sees responsible, sustainable handling of research data as a cornerstone of its research strategy. Taking account of differences in the scientific disciplines and their domain-specific standards is a prerequisite for ensuring the traceability of research results and their reuse in new research projects. In this context, the Leibniz Association has a special duty to take a case-by-case approach, since the diversity of its institutions reflects both the shared features and the differences between numerous scientific disciplines. Leibniz Association institutions have for many years been involved, through a number of professional scientific bodies (e.g. the German Council for Scientific Information Infrastructures and the German Data Forum), in drawing up discipline-specific regulations and recommendations for handling research data.

This approach is in line with developments in the political arena. The importance of research data has been strengthened among other things through recommendations issued by the German Council for Scientific Information Infrastructures (RfII), a body set up by the federal states and federal government of Germany. Even before this, the relevance of quality-assured, efficient, interoperable research data management had been emphasised at national and international level, in particular through recommendations from the High Level Expert Group on the European Open Science Cloud (EOSC) and through publications by the Research Data Alliance (RDA) and the FORCE11 group on FAIR (Findable, Accessible, Interoperable, Reusable) data principles. The German Research Foundation (DFG) also published guidelines on handling research data in September 2015, which reinforce these aims and call on professional associations to reconsider how they handle research data and to develop appropriate discipline-specific regulations on using and providing access to research data. Research data management is also subject to legal requirements, such as copyright, data privacy, liability and public services laws. Here, the Leibniz Association shares its experiences in dealing with sensitive data in discussions with lawmakers.

IV. Aims

The Leibniz Association identifies the following overarching and interconnected aims for handling research data, and acts in the interests of good scientific practice through the long-term storage and adequate documentation of all collected research data. It supports sustainable research data management as a framework for securing and documenting research data, in particular through:

- supporting appropriate organisational and discipline-specific standards and structures for sustainable handling of research data,
- providing technical infrastructure for Leibniz institutions to carry out research data management,
- recognising the collection and publication of research data as a scientific activity in its own right,

- promoting maximum accessibility and reuse of research data, in line with the principle "as open as possible, as closed as necessary" and
- training and qualifying staff in research data management.

To achieve the stated aims, research communities will need to move towards a greater perception and emphasis of activities associated with research data management. Among other things, this requires a citation culture for research data, with standardised metadata and persistent storage addresses.

The Leibniz Association identifies the following aims for handling research data:

Interoperability through the use of open standards and structured metadata

Research data management requires sustainable processes and systems designed for interoperability through the use of open standards. Particular attention should be paid to FAIR data principles, W3C Data on the Web Best Practices, discipline-specific community standards and initiatives, and the FORCE11 Joint Declaration of Data Citation Principles. Metadata about research data contains structured information, while vocabularies and ontologies represent a shared understanding of the structure and semantics of the data. They are all essential for the efficient sharing and effective reuse of research data and help make research data findable and interpretable as (interim) research results.

Traceability through adequate documentation

The Leibniz institutions support adequate documentation of research projects. The minimum requirement is to document the generation, processing, archiving and, where relevant, publication of research data using a research data management plan (or data management plan).

Infrastructure development requires long-term resource planning

The Leibniz Association pursues a strategy of establishing low-threshold access to research data in terms of costs and organisational and technical hurdles. To achieve this aim, suitable infrastructure services for research data management and the long-term storage of research data need to be set up, operated, expanded and, where relevant, connected to one another. Appropriate resources must be taken into account at the application and planning stage of research projects – both by applicants and by funding bodies.

Securing long-term availability and long-term archiving

When it comes to storing and archiving research data, the Leibniz Association recommends using a suitable subject-specific, institutional or generic repository, wherever legally possible. In accordance with its Guidelines on Safeguarding Good Scientific Practice, the Leibniz Association recommends storing research data for a period of at least ten years. The institutions of the Leibniz Association strive for a balance between long-term usability and justifiable expense.

A culture of responsibility as a basic principle throughout academia

The Leibniz Association supports a culture of responsibility throughout academia, in which compliance with data privacy regulations and ethical principles is a matter of course. This is particularly relevant in the context of research data from projects involving human participants (human subject research) and in the case of safety-related research. The Leibniz institutions support the reusability of research data, provided there are no legal or ethical constraints, in line with the principle "as open as possible, as closed as necessary".

V. Measures and recommendations

The tasks associated with sustainable research data management are diverse, and range from assuring data quality and interpretability to overcoming technological, structural and financial challenges associated with the accessibility and reuse of research data. For instance, there has been only a slow increase in open access options for research data in recent years. The reasons often given are a lack of (research) policy requirements, the pressure of competition in academia and a lack of incentive systems. To meet these challenges and achieve the stated aims, the following sections present possible measures and recommendations for the different levels: Leibniz Association, Leibniz institutions, researchers and research support staff. The measures and recommendations should not be seen as an exhaustive list, but rather as suggestions that can be adapted and expanded for specific disciplines.

Leibniz Association

- The Leibniz Association coordinates information sharing and collaboration with research data projects at national and international level. The specialist Sections play a particularly important role in this.
- The Leibniz Association supports organisational collaboration and, where applicable, technical networks between research data infrastructure at the different Leibniz institutions.
- The Leibniz Association recognises the collection and publication of research data as a scientific activity in its own right because of the genuine value it creates.
- The Leibniz Association welcomes the establishment of a voluntary citation system for research data within Leibniz institutions so as to highlight the status of research data (publications). In addition, the Leibniz Association recommends including this as a scientific outcome in appropriate surveys in annual reports.
- The Leibniz Association welcomes initiatives that aim to establish low-threshold access to research data in terms of costs and organisational and technical hurdles. In this area, it promotes sustainable resource planning at the research project application and planning stage.
- The Leibniz Association draws up proposals for ways of updating the standards for evaluation procedures to include the appraisal of research data publishing and research data management.

- The Leibniz Association recommends that research data handling skills be included in university curricula and supports this (for instance in existing partnerships between Leibniz institutions and universities).
- The Leibniz Association calls for research data to be made available, taking legal and ethical requirements into account – especially data generated using public funds. This should also be considered in future political decisions.
- The Leibniz Association supports a wide range of initiatives on the handling of research data. For instance, the Research Data Working Group and the LeibnizData network act as forums for exchanging ideas and as points of contact for research data issues within the Leibniz Association.
- The Leibniz Association will ensure that it regularly reviews the implementation of these guidelines and develops them further, taking account of suggestions from the Leibniz institutions and the research system. To this end, the Leibniz Association is drawing up a comprehensive position on open science, using these guidelines and the Open Access Policy as pillars.

Leibniz institutions
- We recommend that Leibniz institutions adopt their own research data guidelines based on these guidelines and the standards in their field. The guidelines should include appropriate measures for handling research data, with reference to current standards (FAIR data principles, data structure and metadata vocabularies, data management plans, community standards).
- Regardless of any currently foreseeable potential for reuse, it is recommended that research projects consider during the planning phase whether and how the research data generated can be made available to optimise reuse. To safeguard scientific results, the Leibniz institutions are advised to offer their researchers suitable workflows and advisory services on research data management in general, and on drawing up data management plans during the research project planning phase. Such services should be mentioned in the institution's research data guidelines.
- The scientific potential of research data should be highlighted within the Leibniz institutions. This can include e.g. subject-specific solutions for all areas of research data management, which should be communicated, shared and made available for discussion as transparently and openly as possible within the professional community inside and (where possible and sensible) outside the institute.
- Suitable organisational, technical and financial measures should be introduced to reduce the additional administrative burden on researchers associated with systematic research data management. For this reason, we recommend that the Leibniz institutions develop long-term expertise in research data management within their supporting infrastructure (libraries, archives, IT departments) and that they communicate this to scientists through contact points or advice centres. Leibniz institutions are also encouraged to get involved in establishing suitable, networked infrastructure for research data management.
- We recommend that Leibniz institutions use domain-specific standards and examples of research data documentation. When using metadata, vocabularies and ontologies for the

description of research data, the Leibniz Association institutions should be guided by the standards available in the specialist communities (e.g. the Metadata Standards Directory of the Research Data Alliance Metadata Working Group, W3C RDF vocabularies) and the use of generic approaches, for instance in the field of persistent identifiers for the permanent citability of research data (e.g. the Digital Object Identifier (DOI) standard).

- We recommend that Leibniz institutions create suitable processes and structures to help researchers respect legal and ethical requirements when dealing with research data (for instance in relation to compliance with data privacy legislation).
- We recommend that Leibniz institutions provide particular support for their young researchers and for research support staff with responsibilities in this field with regard to sustainable research data management, e.g. by providing suitable training.

Researchers

- Researchers should make use of the increasing possibilities and offers available for managing research data and take the initiative to ensure better research data management in their own projects. Examples of this are training and further training courses on research data management offered at national and international level. We recommend that researchers make use of these opportunities, where relevant assuming a multiplier role for this important topic within their working groups. The option of taking up such training and further training programmes should, where necessary, be offered to research support staff as well.
- We recommend that researchers use repositories and infrastructure for their data that meet today's standards for data publications (FAIR data principles), support data descriptions using metadata and vocabularies and assign persistent identifiers (e.g. DOI, ePIC handle, ARK, URN). They should reference the data they use by citing persistent identifiers.
- When planning their research phases or submitting applications for research projects, researchers should check whether they need to include resources for research data management, for instance the use of long-term archiving and data publishing services offered by repositories (which may charge fees).