

# Policy on Research Data Management

Version & Date: 1.0 / 10 April 2024

Ownership:

Participating departments and committees: Good Scientific Practice Committee

Frequency of reviews: Bi-annually

Approved by: Board of Directors

Status: Active

Confidentiality level: Public

## INTRODUCTION

The Policy on Research Data Management (RDM) is composed of 3 levels of increasing granularity. While the actual policy (1) is a more static document, the guides (2) and toolboxes (3) are intended to help complying with the RDM Policy. They should be seen as living documents that will be improved and specified in iterative cycles, based on experience and feedback by the MPI-CBG community.

### (1) Policy

The actual Policy on Research Data Management regulates fundamental principles of proper research data management. It mandates and enforces compliance with standards, in line with good scientific practice and contractual obligations.

### (2) Guides

The Guides specify and supplement the RDM policy by addressing key messages and standards. They provide background and procedures for better understanding and support of implementing in day-to-day work.

### (3) Toolbox

The Toolbox is a collection of best practices, related documents, and references to further training, reading and support.

## STRUCTURE

For clarity, the structure and content of the different levels are presented in the overview below.

### Policy

1. Background & purpose
2. Basic principles
3. Definitions
4. Scope of application
5. Handling of research data
  - *Data Management Plan*
  - *Data Documentation*
  - *Data Storage*
  - *Storage Period*
  - *Deletion of data*
  - *Reusability*
  - *Intellectual property Rights*
  - *Use and transfer of data*
6. Handling of published data
7. Responsibilities, rights and obligations
8. Enforcement

### Guides

- Key message
  - Keywords, tags\*
  - Background\*
  - Procedure
- |    |  |
|----|--|
| 1: | Data Management Plan                                 |
| 2: | Documentation – ELN                                  |
| 3: | Data storage & organization                          |
| 4: | Analog data  |
| 5: | Open Access, FAIR principles, licensing, citeability |
| 6: | Research Data as basis for publications              |
| 7: | Research Integrity Culture                           |
| 8: | Enforcement  |

### Toolbox\*

- Best practices
- Related documents
- Training
- Further reading
- Support

\* Supplements provided in online formats

# Policy on Research Data Management

---

**Version & Date:** 1.0 / 10 April 2024

**Ownership:**

**Participating departments and committees:** Good Scientific Practice Committee

**Frequency of reviews:** Bi-annually

**Approved by:** Board of Directors

**Status:** Active

**Confidentiality level:** Public

## BACKGROUND & PURPOSE

The Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG) is a leading research institution focusing on the primary research question "How cells form tissues" with the aim of deciphering fundamental cellular functions in the development and evolution of tissues and organisms.

The MPI-CBG is committed to the highest research standards and responsible acting in science, it promotes Good Scientific Practice (GSP) and Open Science. The institute recognizes the long-term value of research data for scientific progress and its potential to benefit society as a whole.

The MPI-CBG appreciates the fundamental importance of research data management (RDM) for maintaining high quality research and ensuring scientific integrity, continuity and reproducibility. Research data is the basis, integral part and result of every research project. It must be correct, easy to find, reproducible and reusable. Research data management plays a central role here, as it does in the publication and defense of scientific processes and research results.

The purpose of this Research Data Management Policy<sup>1</sup> is to establish guidelines for the appropriate and effective management of research data to guide current and future researchers and to ensure the integrity of research and scientific progress at the MPI-CBG.

## BASIC PRINCIPLES

The basic principles for safeguarding Good Research Practice, including proper Research Data Management (RDM), are specified in the Code of Conduct "Guidelines for Safeguarding Good Research Practice", published by the DFG in 2019<sup>2</sup>, and the binding rules of conduct for Good Scientific Practice – how to handle scientific misconduct "Responsible acting in science" of the *Max Planck Gesellschaft* (MPG)<sup>3</sup>.

## DEFINITIONS

### Researchers:

- (i) All staff members of MPI-CBG who are active in research, including (emeriti) directors, research group leaders, scientific service leaders, scientists at the postdoctoral and predoctoral level, and research technicians. Active researchers are typically employees of the MPG holding a working contract, or fellows holding a stipend provided by the MPG. As a rule, the workplace is at the MPI-CBG or the Center of Systems Biology Dresden (CSBD).
- (ii) Scientific collaborators and guests who are temporarily researching at MPI-CBG, and using the institutional infrastructure. They typically hold an MPG guest contract while being employed by another institution, receive a fellowship provided by a third party, or are enrolled as students. Alternatively, the latter may have a contract as a student employee.

<sup>1</sup> This policy was developed on the basis of: (1) Grasse, M., López, A., & Winter, N. (2018). Musterleitlinie für Forschungsdatenmanagement (FDM) an Hochschulen und Forschungseinrichtungen. <https://doi.org/10.5281/zenodo.1149133>; (2) Hiemenz, B.; Kuberek, M. 2019. Strategischer Leitfaden zur Etablierung einer institutionellen Forschungsdaten- Policy. <http://dx.doi.org/10.14279/depositonce-8412> und (3) "Model Policy for Research Data Management (RDM) at Research Institutions/Institutes" (<https://doi.org/10.14324/000.learn.26>) in: LEARN (ed.) LEARN Toolkit of Best Practice for Research Data Management. (pp. 133-136).

<sup>2</sup> DOI 10.5281/zenodo.3923601

<sup>3</sup> <https://www.mpg.de/199493/regelnWissPraxis.pdf>

### Group leaders:

Researchers at MPI-CBG holding a group leader position and responsibility for mentoring young researchers (postdocs, PhD students), students, technicians or other scientifically working members or guests in their own group. Group leaders, include (emeriti) directors, research group leaders, and scientific service leaders.

### Research Data:

Research data are all (digital and analog) data which are subject, steps, or results of the scientific process, including metadata. Examples for research data are experimental design and planning, measurements, chromatograms, microscope images, blots, audiovisual information, movies, mathematical data, collections, simulations, source codes, protocols, compounds, plasmids, organisms, software, and many more. The range of data reflects the diversity and methodical spectra of the scientific disciplines and research processes. Research data can take different forms during the course of research projects (different variants of the primary data, processed data including negative and ambiguous results, shared data, published data) and be provided with different access rights, for example as open, restricted and non-public data.

### Research data management (RDM):

The management of research data includes the planning, acquisition, analysis, documentation, publication and storage of data. It ensures access, reuse, reproducibility and quality assurance of all research data on which scientific results are based.

### Data management plan (DMP):

The DMP is a formal document that specifies how the research data of a project is to be handled. A DMP is written before starting the project, or as soon into the project as possible, and updated during the project as need occurs.

### Phrases:

“must”, “required” or “shall” means an absolute requirement, whereas “must not” or “shall not” means an absolute prohibition. “recommended” or “should” means that there may be valid reasons to deviate from the preferred path, in which case full implications must be understood and carefully weighed.

## SCOPE OF APPLICATION

The guideline is directed at researchers at MPI-CBG. It shall be taken into account for all research work, including individual projects, collaborative projects, and projects with internal and/or external partners, and thus covers all research carried out at MPI-CBG, regardless of whether the necessary resources are provided by the MPI-CBG or by third parties.

Guidelines or stipulations of the Max Planck Society, or specific agreements with third-party funders in relation to data management, take precedence over this guideline.

## HANDLING OF RESEARCH DATA

Research data must comply with the FAIR principles – this means they must be Findable and Accessible and allow for Interoperability and Reuse when possible.

**DATA MANAGEMENT PLAN:** As a basis for complying with FAIR principles a data management plan (DMP) is crucial. MPI-CBG endorses establishment and maintenance of Data Management Plans as routine practice for research conducted at MPI-CBG, and as a vital part of onboarding procedure, grant applications, and cooperative research projects.<sup>4</sup>

**DATA DOCUMENTATION:** Research data need to be documented in a manner that is complete, fraud-resistant, legible, and prompt. This also applies to non-publishable data and inconclusive results and investigations. MPI-CBG mandates documentation of all research work in electronic repositories and storage, which is provided by MPI-CBG, and accessible to the group leader. Experimental work must be documented in the Electronic Lab Notebook (ELN). Theoretical work must also be comprehensively documented, with source code placed in the MPI-CBG Git/GitLab repositories and technical reports maintained, and version controlled, either within Git/GitLab or central storage.<sup>5</sup>

**DATA STORAGE:** The integrity of research data must be preserved. Research data must be stored in a correct, complete, unaltered and reliable manner, they must be protected from unauthorized access, destruction, theft, and manipulation. Both digital and analog research data shall be stored in a suitable repository or archiving system in a safe and accessible manner. The data should be completed with appropriate metadata. Primary research data including metadata are stored on institutional storage infrastructure, according to pre-defined location and structure, in concordance with the FAIR principles.<sup>6,7</sup>

---

<sup>4</sup> See Appendix: GUIDE 1 – Data Management Plan

<sup>5</sup> See Appendix: GUIDE 2 – Documentation – Laboratory notebooks

<sup>6</sup> See Appendix: GUIDE 3 – Data storage and organisation

<sup>7</sup> See Appendix: GUIDE 4 – Analog data

**STORAGE PERIOD:** Research data and documents must be stored and made accessible in accordance with applicable legal and contractual provisions, irrespective of whether research results are published, conclusive, or not. In general, the minimum retention period for research data is 10 years, starting with the date of publication. The minimum retention period for research data not intended for publication or further research is 10 years, starting from the leaving date of the responsible group leader.

**DELETION OF DATA** after the retention period has expired must be traceable and documented. The decision whether to retain or delete the data shall take into account the interests and contractual provisions of (third-party) funders, participating researchers and collaboration partners.

**REUSABILITY:** Research data shall be freely licensed and made openly available in accordance with intellectual property rights, provided that no third-party rights, statutory provisions or other intellectual property rights prohibit this. Research data intended for subsequent use shall be made available on a suitable storage (repository) in a citable form by tagging with persistent identifiers (DOI or handle), and complementing with relevant metadata. The origin of reused data is thus clearly traceable and the corresponding source is rewarded. Citation rules are to be observed.<sup>8</sup>

**INTELLECTUAL PROPERTY RIGHTS:** Copyright guarantees protection of intellectual creations and output. Rights of use and exploitation of research data are defined in contractual agreements between the researcher and the MPG. Rights of use and exploitation can also be defined by additional agreements (such as grant agreements or consortium agreements) and should also be addressed in the RDM plan. Whether research data are subject to copyright law depends on whether the requirements as to the extent of own intellectual creation are met, and need to be evaluated on a case-to-case basis.

**USE AND TRANSFER OF DATA:** Group leaders who transition to a new hosting entity for continuing their research do have the right to transfer data generated by their entire research group to the new location. A structured off-boarding procedure guided by the institute must precede the data transfer in any case. Group members such as postdoctoral or predoctoral scientists must conclude documented rights of use agreements with the responsible group leader. In any case, data remain securely stored at MPI-CBG for at least 10 years, implying that only copies may be taken.

## HANDLING OF PUBLISHED RESEARCH DATA

Published research data need to be readily available for re-use of data, reproduction, and defense of the research result.

Any data associated with a publication must be retained, organized, and described sufficiently in a specific storage location on the MPI-CBG fileservers with a copy of relevant data.<sup>9,10</sup> That implies not only processed but also primary data. The compilation of the data shall be completed before submission of the manuscript, and available in the final version upon acceptance. This copy will be made read only and archived, and referenced to the publication and its ID. Published data must be stored for at least 10 years starting with the date of publication.

## RESPONSIBILITIES, RIGHTS AND OBLIGATIONS

Responsibility for research data management during and after the project period lies with the Max Planck Institute of Molecular Cell Biology and Genetics, and its researchers who create, access, maintain, store, or share research data; they are thus accountable as data creators and data users. Ultimately, the group leaders are primarily responsible for proper research data management, while the institute is responsible for providing infrastructure, training and guidance.

### Responsibilities of the researchers

Researchers at MPI-CBG

- i. handle research data in line with the principles and requirements of this and related institute's policies, MPG regulations, and any other legal requirements.
- ii. collect, document, store and archive research data and the associated documentation such that it is findable and that access, re-use or legitimate deletion is possible.
- iii. create a Data Management Plan (DMP) for each research project, which is maintained over the course of the project.
- iv. plan, as far as possible, the further use of the data, especially after completion of the project, including usage and exploitation rights, assignment of appropriate licenses, and the storage, archiving and deletion of data.

<sup>8</sup> See Appendix: GUIDE 5 – Open Access, FAIR principles, licensing, citeability

<sup>9</sup> See Appendix: GUIDE 6 – Special requirements for research data as basis for a publication

<sup>10</sup> Further details will be regulated in the MPI-CBG publication policy

- v. meet all relevant organizational, regulatory, institutional and other contractual and legal provisions both in relation to research data and associated research documents, and stand up for them.
- vi. grant their group leader access to research data upon request and properly hand over research data that they have created at the MPI-CBG upon leaving the institute
- vii. adhere to supplementary agreed processes and responsibilities in joint research projects.

### Responsibilities of the group leaders

#### Group leaders at MPI-CBG

- i. ensure proper research data management in line with the principles and requirements of this policy, and any other legal or MPG requirements, by holding an annual group briefing to recall the key points of this policy, and by training and advising group members during daily work.
- ii. review the accuracy and completeness of the DMPs, the primary data and the data documentation and storage with each group member, at least on an annual basis.
- iii. compile and store all data for publication in line with this policy and additional institutional requirements.
- iv. participate in the RDM on-boarding procedure, a yearly internal RDM review and in the RDM off-boarding procedure.
- v. ensure proper access to and transfer of research data they generated at MPI-CBG upon leaving the institute. Rights to access and copy for other co-investigators to research data resulting from own work shall be addressed individually.

### Responsibility of the institute<sup>11</sup>

#### The Max Planck Institute of Molecular Cell Biology and Genetics

- i. provides funding, resources and support for research, services, the operation of organizational units, infrastructure and employee qualification.
- ii. promotes compliance with the recommendations for good scientific practice. To this end, the institute provides training, support and advice for responsible research and proper RDM, in accordance with institutional and other applicable policies.
- iii. develops mechanisms and provides services to securely save and store research data so that access to research data can be guaranteed during and after the completion of research projects.
- iv. provides access to the services and infrastructure described above, so that researchers can comply with the requirements of the institute, third-party funders and other legal entities, and can fulfill their responsibilities as described in these guidelines.
- v. regularly reviews the implementation of and compliance with applicable RDM policies, as well as necessary updates to the same.
- vi. stores (analog and digital) research data generated at MPI-CBG in line with legal requirements

## ENFORCEMENT

All MPI-CBG staff members are required to comply with this Policy in carrying out their duties for the MPI-CBG. Compliance with this policy is part of the regular performance appraisals. Labor law measures may be taken in the event of a violation.<sup>12</sup>

This Policy on Research Data Management was approved on **10 April 2024** by the MPI-CBG Board or Directors. It enters into force on **1 July 2024** and will be updated every 2 years by decision of the Board of Directors.

Dresden,

18.04.2024

  
\_\_\_\_\_  
Prof. Anne Grapin-Botton  
Managing Director

<sup>11</sup> See Appendix: GUIDE 7 – Research Integrity Culture

<sup>12</sup> See Appendix: GUIDE 8 – Enforcement

# Research Data Management Policy

## Guides

<b>Version &amp; Date:</b>	1.0 / 10 April 2024
<b>Responsible:</b>	Coordinator, Good Scientific Practice & Research Output
<b>Participating departments and committees:</b>	Good Scientific Practice (GSP) Committee
<b>Frequency of reviews:</b>	Bi-Annually
<b>Approved by</b>	Board of Directors
<b>Status:</b>	Active
<b>Confidentiality level:</b>	Public

The guides and toolboxes are appendices to the MPI-CBG Research Data Management Policy, which was adopted by the MPI-CBG Board of Directors on **10 April 2024**. While the actual policy (1) is a more static document, the guides (2) and toolboxes (3) are intended to help you comply with the RDM Policy. They should be seen as living documents that will be improved and defined in iterative cycles, based on experience and feedback by the MPI-CBG community.

However, the Guides also include binding key messages and procedures. They are listed below for clarity. Further content of the guides, such as background, literature, and the toolboxes, can be found in the English version.

### GUIDE 1: RESEARCH DATA MANAGEMENT PLAN

#### KEYMESSAGE

MPI-CBG requires establishment and maintenance of Data Management Plans as the routine practice for research conducted at MPI-CBG and as a vital part of onboarding procedure, grant applications, and cooperative research projects.

#### PROCEDURE

MPI-CBG provides a DMP template which is pre-filled according to the most current MPI-CBG requirements and infrastructure for RDM. DMPs are "living documents" and should be maintained as follows:

- (1) Scientific newcomers start writing their DMP within the first 3 months at MPI-CBG.
- (2) The DMP should be updated in context with TAC reports (PhD students) and the annual appraisal meeting (Postdocs).
- (3) The group leader is responsible for initiating the RDM draft, and for reviewing and discussing the DMP on an annual basis.
- (4) Collaborative projects (with internal and external partners) require a DMP that is mutually agreed upon and which will be updated during the project.
- (5) DMPs shall be stored on the file server, in the project folder it belongs to, and updated as needed.
- (6) It is required to finalize the DMP upon completion of a project. Before leaving the institute, the DMPs of all running projects must be updated.

### GUIDE 2: DOCUMENTATION – ELN

#### KEYMESSAGE

The MPI-CBG has selected *Labfolder* as its centralized Electronic Lab Notebook (ELN) and mandates its usage for the documentation of web lab scientific work and the compilation of research data and record keeping. Source code generated must be maintained within the MPI-CBG source code repository (Git/GitLab).

#### PROCEDURE



### Wet Lab Research

- (1) Use of the MPI-CBG ELN Labfolder is mandatory for all wet lab research at the MPI-CBG, irrespective the scientific discipline or the career level, as it provides a central repository of this information for the MPI-CBG. All benchwork, regardless of group, must be documented in Labfolder.
- (2) The group leader (GL) shall have the administrative rights for the ELNs of their group members. S/he is responsible for the overall structure (folders), granting access to the ELN to laboratory members, and ensuring termination of access upon departure of a group member.
- (3) Researchers are responsible for documenting their research activities, data and results as outlined in the RDM policy, thus ensuring accuracy and completeness.
- (4) The GL is responsible for reviewing the accuracy and completeness of the ELN according to discipline specific requirements.
- (5) The GL is responsible for reviewing the primary data associated with lab notebook entries for accuracy and completeness
- (6) If additional analog data, or data stored in a repository or other systems, belong to the complete documentation, the content must be regularly imported into the ELN if technically possible or entries made indicating its location, if not possible, using persistent identifiers (PID, e.g. DOI or handle) if they exist to maintain continuity.
- (7) The ELN is generally not designed to store large data sets, these should be stored instead on central MPI-CBG resources (the file server and project spaces, see Guide 3). The ELN shall indicate the location on the central MPI-CBG fileservers where the corresponding primary data is stored.
- (8) Access to the ELN, as well as training for ELN usage and proper documentation, is part of the onboarding process and shall be completed within the first three month after assuming work at the institute.

### Theoretical Work

Groups that primarily engage in theoretical work, such as those focused on computer science, mathematics, modeling, and related disciplines, should ensure comprehensive documentation of their research in accordance with the specific standards and requirements of their discipline. This can vary by each group but:

- (1) The use of the MPI-CBG ELN Labfolder is mandatory for any wet lab or benchwork carried out. It is highly recommended to utilize the ELN for the documentation of other types of data and work, as long as it is practical and applicable.
- (2) Source code generated must be maintained within the MPI-CBG source code repository (Git/GitLab). The GL is responsible for ensuring the completeness and accuracy of this source code.
- (3) Formal work, in the form of technical reports, must also be complete and version controlled, either within the ELN, within the source code repository, or on the fileservers, with appropriate structure.
- (4) The GL is responsible, in all cases, for assuring that the work is replicable from the data and technical reports stored.
- (5) If additional data, stored in repositories or other systems, are required for the work this must be indicated in the reports stored along with the location of this data, using persistent identifiers when they exist.

## GUIDE 3: DATA STORAGE AND ORGANIZATION

### KEYMESSAGE

**MPI-CBG requires that research data, including metadata, are stored on the provided central MPI-CBG storage architecture.**

### PROCEDURE

#### Primary data storage

Data are stored and retained in line with the Good Scientific Practice and legal requirements. Briefly:

- (1) All primary research data must be stored on the MPI-CBG file server, and accessible for the group leader and the institute.
- (2) Fileserver spaces are exclusively for research and other work-related data. The storage of private or non-work-related data on the fileservers is strictly forbidden.

- (3) Documentation of primary data and metadata need to comply, as much as is feasible, with FAIR principles (Glossary). To this end, MPI-CBG provides a recommended organizational structure for fileserver space.
- (4) Researchers are responsible for properly organizing, curating, and describing research data, with appropriate linkage to documentation within the ELN or, in the case of theory groups, to reports stored within Git and Gitlab.
- (5) GLs are responsible for supervising proper organization, curation and description of primary data as outline under (4). GLs are responsible for ensuring that appropriate metadata is stored and correctly associated with primary data.
- (6) Server space and permissions are assigned at the request of the GL to the MPI-CBG Computer Department. A justification for the space may be requested, especially in the case of extensive applications that go beyond the usual scope. Assignment of server space may be subject to availability, as in some cases additional storage may need to be acquired.
- (7) The MPI-CBG leadership reserves the right, under exceptional circumstances, to access all stored data for legal or scientific reasons.
- (8) The ELN is not designed for the storage of large data sets. For these large data sets, it should be indicated within the ELN where the data has been stored.
- (9) Data for publication are to be compiled and stored in a dedicated space which will be provided. This data will become read only and will be kept for a minimum of 10 years starting with the date of publication.
- (10) Re-usability and transfer of (digital and analog) research data generated by leaving staff members is ensured in a defined RDM off-boarding process, whereby research data and organization, curation, description and storage is reviewed. GLs are responsible for ensuring the RDM off-boarding process of their group members, whereas the GLs' off-boarding process is supervised by the responsible institutional representative.

## GUIDE 4: ANALOG DATA

### KEYMESSAGE

**MPI-CBG stipulates that analogue data, just like digital data, are to be secured such that they are indexed, findable, and accessible for re-use.**

### PROCEDURE

The following points need to be observed for reproducibility, and re-use of analog data:

- The origin of the material, compounds, chemicals, cells, DNA/RNA, enzymes, antibodies, blots etc. must be documented in detail.
  - If commercially obtained: company, order number, and batch number must be recorded.
  - If provided by a researcher from another institution (for example plasmids, cells, or antibodies): confirmation of identity (for example by sequencing) and quality (for example by excluding contamination) is compulsory.
- Handing-over analog data is part of the off-boarding process and under responsibility of the group leader (see Guide 7 – RDM offboarding checklist)
- Laboratory notebooks (hard-copy) are institute's property, they must remain in the institute.
  - Leaving scientists may take copies of the documentation of their **own** research in agreement with the group leader.
  - The laboratory notebooks must be orderly labelled and handed over to the Coordinator of GSP & Research Output who will include them in the directory and initiate storage. Storage is in general 10 years, unless a longer storage period is requested.

## GUIDE 5: OPEN ACCESS, FAIR PRINCIPLES, LICENSES, CITEABILITY

### KEYMESSAGE

**MPI-CBG embraces Open Access and FAIR principles and supports its scientists in making research data and software freely accessible and usable, while ensuring that credit is given to authors or creators.**

### PROCEDURE



- (1) Make your data FAIR (findable, accessible, interoperable and re-usable)
- (2) Assign a DOI to your data set.
- (3) Assign the correct license to your data set.
- (4) Include how your datasets shall be cited.
- (5) Contact the library for help and advice on open access publications

## GUIDE 6: RESEARCH DATA AS BASIS FOR A PUBLICATION

### KEYMESSAGE

**MPI-CBG endorses that all primary and processed data associated with a publication must be retained, organized, and described sufficiently in a specific storage location on the MPI-CBG fileserver, which is accessible to the institute management. Linking to data sets which are deposited and accessible in public repositories substitutes for the need to store a copy on the institute's fileserver space**

### PROCEDURE

For research data that are the basis of scientific publications, MPI-CBG adopted the following procedure. It safeguards data integrity and completeness, and simplifies referencing, sharing of data and responding to inquiries.

- (1) If any senior author has an MPI-CBG affiliation, s/he is in charge of ensuring complete and correct storage of all publication-associated data generated at MPI-CBG. If a (co-)author with MPI-CBG affiliation contributed to a publication with a corresponding author from another institution, s/he is obliged to ensure complete and correct storage of the published data resulting from work at MPI-CBG.
- (2) Data that have been generated at MPI-CBG and are the basis for a publication must be deposited on a dedicated fileserver space ("MPI-CBG publications"). This space is accessible to the institute management.
- (3) The recommended folder structure should be followed (see toolbox). Should the recommended structure not be adhered to for a good reason, the responsible scientist must ensure an intuitive structure that enables findability.
- (4) The data will be referenced to the publication and its ID. This copy will be made read only and archived.
- (5) Any primary data linked to the paper should be made public in an appropriate and trustworthy public repository (see Guide 3), supplemented with appropriate metadata and license, unless prohibited by legal requirements. If it is not possible to submit data to an external repository the MPI-CBG will provide appropriate means to share the relevant data.
- (6) Primary data sets hosted in repositories or at MPI-CBG should be unequivocally linked to the publication in our system (e.g., via DOIs or access number)
- (7) Software source code shall be kept in a valid institutional repository, preferably the MPI-CBG Git/GitLab instance (<https://git.mpi-cbg.de>).
- (8) Existent data or software that has been used in the publication must be properly cited (see Guide 5).

## GUIDE 7: RESEARCH INTEGRITY CULTURE

### KEYMESSAGE

**MPI-CBG mandates training in Good Scientific Practice and Research Data Management, and requires proper research data management guided by an RDM on-boarding process, annual RDM reviews, and an RDM off-boarding process.**

### PROCEDURE

- (1) Complete your RDM training as specified on the RDM on-boarding checklist. It is part of the institutional on-boarding process steered by the HR department. Staff members who have already completed onboarding earlier will participate in ongoing RDM training whenever possible.
- (2) Participate in annual RDM checks and institutional offers supporting your RDM skills and knowledge.
- (3) Prepare and complete RDM offboarding in good time.

## GUIDE 8: ENFORCEMENT

### KEYMESSAGE

Compliance with the MPI-CBG Research Data Management policy is obligatory.

### PROCEDURE

#### Annual assessment of research data management

- (1) Group leaders will be provided an annual overview of their research data on the MPI-CBG fileservers, Git/Gitlab and ELN, followed by an internal review. The review will involve the Coordinator for Good Scientific Practice and Research Output, and, as needed, IT experts. They will advise on data storage and organization, and discuss if and what further support is needed.
- (2) Group leaders shall regularly assess the compliance with proper research data management of their group members as a standard part of
  - annual appraisal meetings (for postdocs),
  - Thesis Advisory Committee reports (for predocs), and
  - staff performance interviews (for technicians, and temporary group members such as interns).
- (3) The assessment shall be documented on the respective appraisal forms.
- (4) In case of documented violations, clarification shall be sought by in-depth discussions with the group member. Recommendations on how to improve RDM knowledge, for example by training, should be made. A follow-up review shall be arranged within 6 months.
- (5) Should the situation not improve, it will be brought to the attention of the Coordinator of Good Scientific Conduct and the Managing Director, the latter of which may issue an official warning which is formally documented in the personal file.
- (6) Subsequent violations may result in contract dissolution.

Dresden, 18.04.2024

  
\_\_\_\_\_  
Prof. Anne Grapin-Botton  
Managing Director