IOW Data Policy

The Leibniz Institute for Baltic Sea Research Warnemünde (IOW) [1] supports and fosters the essential process of providing **research data** comprehensively, openly and long-term available, to the greatest extent practicable. The aim is to ensure access, interpretability and reproducibility of the research outcomes regardless of **data creator** presence, e.g. due to project durations, thus making valuable research findings transparently visible and (re)usable in the form of **data** and **metadata**. For this purpose, IOW has adopted the current Data Policy which covers aspects of the internal handling, provision and publication of different types of research data and points out corresponding responsibilities as well as options for data access, licensing and external data visibility. This policy applies to all researchers and data contributors at IOW (e.g. also lab staff, project leads, project data managers) and can be specified and extended by further policies within research fields or research projects. The IOW Data Policy will be reviewed as necessary to ensure it keeps pace with scientific requirements and **research data management** best practices.

1 Preamble

As part of the Leibniz Association, IOW supports Open Access and Open Science with its possibilities and limitations as depicted in the "Open Access policy of the Leibniz Association" [2] and the "Leibniz Open Science Policy" [3], and considers the "Guidelines on the Handling of Research Data within the Leibniz Association" [4]. Likewise, IOW adheres to the "Guidelines for Good Scientific Practice in the Leibniz Association" [5] as well as to the "Guidelines for Safeguarding Good Research Practice" by the German Research Foundation (DFG) [6] in their recent version.

All newly collected and produced research data, whether it is observed, automatically measured, computer-simulated, manually gathered or for example audiovisually recorded, must be handled in accordance with the FAIR Data Principles [7], hence making **datasets** or **data products** in general Findable, Accessible, Interoperable and Reusable. This applies to all **analyzed, quality-controlled** or published data and, where applicable and necessary, to unprocessed **raw data**. In order to foster the implementation of the FAIR Data Principles, the **IOW research data management infrastructure (IOW RDMI)** provides concrete guidelines for data handling and publication [8].

2 Key Principles

At IOW we distinguish between four types of **research data** which all cover **raw data**, **analyzed/processed data** and **final data** or **data products**: (a) Environmental Data, (b) Experimental Data, (c) Modeling Data as well as (d) Software, Documentations and Surveys.

(a) *Environmental Data* is obtained by observation and sampling of the natural world and can always be georeferenced (which also applies to georeferenced samples that are analyzed in a laboratory).

(b) *Experimental Data* is generated from experiments that provide an idealized representation of certain aspects of the natural world under controlled conditions without a data-relevant georeference.

(c) *Modeling Data* is derived from a mathematical, abstract representation of the natural world in the form of equations, numerical algorithms, or obtained from Artificial Intelligence applications (e.g. neural networks, machine learning). It may or may not be georeferenced.

(d) *Software, Documentations and Surveys* are any kind of source code (e.g. of numerical models, analysis scripts) or other textual and/or tabular information (such as protocols, device specifications, opinion surveys or cruise/study reports, also from social science studies).

For optimal sharing, reproducing, publishing and reusing of research data of all types, it is necessary to carry out a thorough discipline-specific documentation of the scientific work from the outset, at best even before data collection or data production. Such documentation includes early considerations on data storage, backup, sharing options, archiving, selection of a suitable data repository as well as correct acquisition and citation of own or third-party data used. Research data should always be accompanied by comprehensive **metadata** which precisely describes the context of the data. Metadata includes, but is not limited to, for instance information about

- data capture (like used devices and conditions, applied analysis methods and/or software, the context or project framework of the data creation, geographic and temporal references, contact or responsible persons),
- data content (such as structure, size and format of the data, parameter names, variables, units etc.),
- data level (e.g. raw, processed, **quality-controlled data** or third-party data by other institutions),

- conditions for reuse (stated by a usage license) and/or potential access restrictions and
- related resources.

Depending on the data type, community-specific FAIR metadata standards may apply, adherence to which may require the provision of additional metadata.

Data producers and creators at IOW – together with their project or working group leaders and the supervising PIs – are responsible for their data and metadata. The data producer/creator has first access to the data if it belongs to a PhD or a postdoc project. Should problems arise here, the DFG guidelines on good research practice apply in principle. Data producers and creators are required to use structuring tools such as a Data Management Plan (DMP) for a clear, organized management and documentation of their data both when working alone and as part of a team. A DMP should be created prior to data generation, ideally during the planning phase of any research project or research initiative, and should be considered a "living" document that is updated with the relevant metadata as needed during the course of the research. The DMP provides added value, especially in research collaborations, if it contains, for instance, a tabular overview that is actively filled in by the scientists with detailed, up-to-date metadata on the data to be collected, the data collected and its preliminary internal availability, as well as access information to the finally published data. When preparing data and metadata, it is part of everyone's good scientific practice to use community standards such as commonly used formats, vocabularies or units. All data should be prepared in an openly accessible format, easily machine-readable and discoverable according to the FAIR Data Principles.

3 Data Storage and Archiving

Strategies for **research data** storage and backup must be outlined in the DMP prior to and updated during data collection and processing. Throughout the entire scientific work process, **data producers and creators** at IOW must aim for a safe interim storage and backup of their research data. Valid storage solutions include, e.g., local storage media, IOW data servers and IOW cloud solutions, which further enable collaborative work and exchange with both internal colleagues and external partners. Furthermore, the implementation of a reliable backup strategy includes redundancy, possibly encryption and security measures. Automated backups, regular versioning and storing data in multiple locations are crucial to minimize the risk of data loss.

Long-term archiving of all types of research data may be independent of publication of the data and may happen at any point during the scientific work process, i.e. at any stage of data processing. Data archiving requires the preparation of a **finalized dataset** enriched with comprehensive **metadata** and preferably in an open source data format to ensure reusability. The **IOW RDMI** offers specialized centralized archiving solutions, e.g. **IOW_DATA**.

4 Data Publication, Access and Licensing

Finalized datasets must be securely archived and published in a citable manner, in a suitable trustworthy, type- or discipline-specific long-term data repository with persistent identifiers (IDs) like DOI [9] or Handle [10] (for documents, data, journal publications) and ORCID [11] (identifier for easy identification of authors). For data belonging to the long-term monitoring program of the IOW, the **IOW RDMI** with secure long-term storage has to be used. **Research data** should be archived with the least possible temporal delay, ideally before writing an accompanying textual publication and before the end of a project or a contract period. Archived data may be put under embargo for a maximum of two years unless specified otherwise in the DMP, so that related work can be completed and associated scientific journal articles can be written exclusively in accordance with this Data Policy. Furthermore, a license must be provided. The data output of all research conducted at IOW will be made publicly visible as far as possible regardless of the type of funding under which the data has been created in order to strengthen both the authors' and the institute's profile in the international research community.

Any type of resource (datasets, software, documentations, scientific journal articles etc.) that are related to or build on each other must be referenced by using persistent IDs in order to ensure unambiguous linking. This can often also be done retrospectively by adjusting the **metadata** of publications. The provision of machine-readable detailed, meaningful metadata with the published datasets is essential in order to make the data clearly understandable, interpretable, reusable as well as harvestable and hence accessible via national and international data portals. Data portals offer centralized access to decentrally stored data and can create a connection to data held at IOW, thereby increasing its public visibility. This requires that all metadata associated with research data must be submitted to the **IOW RDMI** for long-term availability and dissemination via the IOW metadata catalogue **IOWMETA** [12].

5 Access to Research Data

In conclusion, IOW grants access to its **research data** in the following form and under the following conditions:

1) IOW research data held by the **IOW RDMI** will generally be made openly available to any person or any organization who requests them. If it is proposed to restrict access to any data, we will give legitimate reasons for doing so.

2) To facilitate the international exchange of data, all IOW **metadata** will be made available in compliance with internationally established metadata standards, such as ISO 191** standards for georeferenced metadata.

3) To protect the research process, IOW will allow its researchers exclusive access to IOW research data for a limited period of time for scientific analysis and publication. Metadata will be made publicly available immediately after their submission to the IOW RDMI.

4) In general, IOW recommends that all IOW research data will be accompanied by an international Creative Commons (CC) data license of type CC BY 4.0 [13]. This license allows the **data consumer** to copy and redistribute IOW research data in any medium or format and to remix, transform and build upon them for any purpose, even commercially. It requires, however, that IOW researchers are given appropriate credit. Therefore, the following reference is mandatory: *"**Data Owner**: Leibniz Institute for Baltic Sea Research Warnemünde (IOW)"*. Any changes to the data must be indicated in a reasonable manner which does not give the impression that the licensor (in this case IOW) endorses the user or the use of the data.

5) Data originally provided to IOW by a third party may have their own access and license conditions, which restrict how this data can be distributed and used. For such data, the user is referred to the respective data owner.

6) IOW research data held by the IOW RDMI will be supplied free of charge with the possible exception of large and complex requests where costs of supply may be charged, or where third-party license conditions either prevent such free supply or require us to make specific charges.

7) In case the research data includes highly sensitive data (as in social science studies), usually only public access to the metadata must be provided.

Appendix:

    A1) Glossary
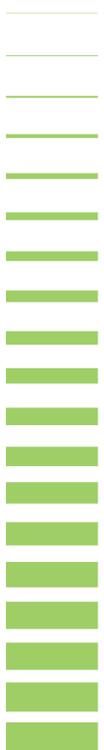   **A2) Listing of cross-references**

Warnemünde, 29.10.2025
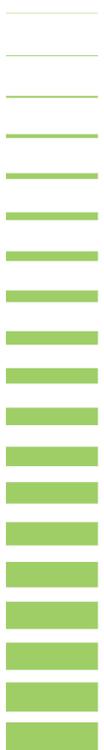
Appendix

A1) Glossary:

- **Data/Dataset/Research data:** All kind of data generated during the research process or resulting from research (numbers/variables linked with units; texts; audiovisual recordings etc.). Data originate for instance from observations, collections, numerical modeling or derivations and can be validated, processed, further processed, analyzed and finally archived.

- **Raw data:** Unprocessed data collected during initial data acquisition, e.g. during expeditions, experiments or numerical modeling.

- **Analyzed/Processed data:** Prepared, edited, cleaned, modified raw data.

- **Quality-controlled data:** Analyzed/processed data, which has been validated to meet data type and application-specific quality criteria, and which ideally does not require further modifications for error correction. Raw data and quality-controlled data may contain identical values with the distinction that a data validation step has been passed.

- **Finalized dataset:** Analyzed/processed data, which is released, i.e. published via a trustworthy data repository, and hence citable. In contrast to the previously mentioned kinds of data, this data can be reused by persons who were not involved in data acquisition and data management, and ideally also by machines. Metadata of finalized datasets must be made available via IOWMETA.

- **Data product:** Finalized dataset or data visualization which can be recreated from quality-controlled data and source code. Metadata of data products must be made available via IOWMETA.

- **Data producer and creator:** A person or team of people who generates or collects data during the research process and prepares/processes it for final archiving and/or publication.

- **Data owner:** Institution responsible which assures the implementation of standards and guidelines as well as the quality of data. The data owner manages the intellectual property rights of the research data.

- **Data consumer:** A user of data (a person or also a machine) who uses and processes data (content) for further purposes.

- **Metadata:** Descriptive information on research data about data content (such as measured variables or units), time period/frame and location of data capture, data acquisition and its purpose, responsible or contact persons.

- **Research data management:** Sustainable, responsible, open and FAIR handling and managing of research data which includes everyday working tasks of researchers such as data capture, processing, analysis, storage, archiving, publishing and (re-)use. A proper, effective and sophisticated handling of research data creates added value both for the data creator itself and for the entire research environment.

- **IOW research data management infrastructure (IOW RDMI)**: The IOW RDMI consists of resources and staff involved in research data management at IOW. It provides central systems for data and metadata storage and archiving.

- **IOWMETA**: A metadata information system for the collection and distribution of metadata via standardized web services (ISO 191**, optionally INSPIRE). The acquisition of metadata can be done automatically or manually for specific finalized datasets or data products. Connection to the Marine Data Infrastructure Germany (MDI-DE [14]) / Spatial Data Infrastructure Germany (GDI-DE [15]), the German Marine Data Portal [16] coordinated by the German Marine Research Alliance (DAM) [17], the portal umwelt.info [18] and other data portals is enabled via the above-mentioned web services. IOWMETA is limited by the ISO-standard for georeferenced metadata and does not offer regular options for experimental data without content-relevant georeference.

- **IOW_DATA**: A central, secure file system on which primary data, unvalidated data and validated data is stored "as is". In conjunction with the metadata in IOWMETA, this data can be searched and retrieved.

A2) Listing of cross-references:

1. Leibniz Institute for Baltic Sea Research Warnemünde (IOW). https://www.iow.de/. Last accessed in October 2025
2. Open Access policy of the Leibniz Association (2016). https://www.leibniz-gemeinschaft.de/open-access-policy. Last accessed in October 2025
3. Leibniz Open Science Policy (2022). https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/Forschung/Open_Science/Open_Science_Policy.pdf. Last accessed in October 2025
4. Guidelines on the Handling of Research Data within the Leibniz Association (2018). https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/Forschung/Open_Science/Leitlinie_Forschungsdaten_2018_EN.pdf. Last accessed in October 2025
5. Guidelines for Good Scientific Practice in the Leibniz Association (2019). https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/%C3%9Cber_uns/Integrit%C3%A4t/Guidelines_for_Good_Scientific_Practice_in_the_Leibniz_Association.pdf. Last accessed in October 2025
6. Deutsche Forschungsgemeinschaft (2025). Guidelines for Safeguarding Good Research Practice. Code of Conduct. https://doi.org/10.5281/zenodo.14281892
7. FAIR Principles. https://www.go-fair.org/fair-principles/. Last accessed in October 2025
8. IOW Data Portal. https://www.iow.de/data-portal.html. Last accessed in October 2025
9. DOI Foundation. https://www.doi.org/. Last accessed in October 2025
10. Handle System. https://en.wikipedia.org/wiki/Handle_System. Last accessed in October 2025
11. ORCID (Open Researcher and Contributor ID). https://orcid.org/. Last accessed in October 2025
12. IOW Metadata Catalogue (IOWMETA). https://iowmeta.iow.de/. Last accessed in October 2025
13. Creative Commons License CC BY 4.0. https://creativecommons.org/licenses/by/4.0/. Last accessed in October 2025
14. Marine Dateninfrastruktur Deutschland (MDI-DE). https://www.mdi-de.org/. Last accessed in October 2025
15. Geodateninfrastruktur Deutschland (GDI-DE). https://www.gdi-de.org/en. Last accessed in October 2025
16. Marine Data Portal. https://marine-data.de/. Last accessed in October 2025
17. Deutsche Allianz Meeresforschung (DAM). https://www.allianz-meeresforschung.de/en/. Last accessed in October 2025
18. Portal umwelt.info. https://umwelt.info/en. Last accessed in October 2025