

IAMO Research Data Management (RDM) guideline

endorsed by the IAMO Board of Directors, Feb 16, 2023

ToC:

- 1 Motivations for and basic principles of RDM at IAMO
- 2 Fields of action in detail
 - 2.1 Long-term preservation of all research data
 - 2.1.1 Scope of data dealt with
 - 2.1.2 FAIR data publishing in repositories
 - 2.1.3 Safe archiving: RADAR cloud or in-house
 - 2.2 Internal list of all IAMO research data
 - 2.3 Provision of the Research Data Management Organizer tool
 - 2.4 Assistance and training in RDM

Appendix: IAMO Workflow to adequately share/ preserve research data (Flowchart)

1 Motivations for and basic principles of RDM at IAMO

1. Compliance with external requirements (e.g., DFG-Kodex)
 - a. Contribution to open data /FAIR data initiatives: facilitate re-use of data
 - b. Keep all research data accessible for ten years (to check methodological correctness)
2. Compliance with data privacy requirements (GDPR/DSGVO)
3. Research excellence and successful applications for third-party funds via sound RDM
4. Property/use-rights/intellectual property rights: clarifications, agreements to avoid frictions among involved persons/parties
5. Achieve the highest possible level of transparency of the IAMO research data pool. Efficient use of (financial and time) resources in generating a research data pool

Further explanations: *This guideline explains RDM workflows to be followed by all IAMO researchers and those colleagues involved in implementation, particularly in support (RDM-coordinators¹, IT). It focuses on data management issues *beyond* what is required for high-quality research (e.g., sound survey design, item construction, and correct data processing).*

2 Fields of action in detail

2.1 Long-term preservation of all research data

IAMO researchers are responsible for ensuring that all research data they generate and/or use are safely preserved for at least ten years, preferably published along [FAIR principles](#) (2.1.2) or reliably archived with sufficient documentation (2.1.3) otherwise. Preservation becomes mandatory in either

¹ Each scientific department has one contact person for RDM issues, the “RDM coordinator”.

of three situations: a) when any research results based on the data are published, b) when the researcher who created the data leaves IAMO, or c) when a project (e.g., a Ph.D. thesis) that has been initially prepared at IAMO is completed after the researcher has left IAMO. Datasets used in IAMO research that are reliably long-term available elsewhere (e.g., in a data repository registered in re3data.org) do not have to be mirrored. All published and archived datasets must be enlisted and sufficiently described in the IAMO research data directory (2.2) implemented in the IAMO Research Information System (FIS).

What necessarily forms part of a research data package to be preserved varies by motivation (section 1) and cannot be specified exactly for all cases. First, IAMO encourages storing “raw data” (at least those that cannot be re-retrieved from other sources if needed). Second, software code (or detailed description of all processing steps) specifying the path from raw data to an analysis dataset should be stored or, alternatively, the analysis dataset itself. This includes software generated sets of simulated data with no or little raw data input. Third, to facilitate quality and correctness checks, documentation of the path from an analysis dataset to the results presented in a scientific publication must be preserved. This may be software code or detailed description in methodological documents to supplement any description of data processing steps in the publication where necessary.

2.1.1 Scope of data dealt with

Data types/formats:

- Quantitative, qualitative, photos, video, audio
- collected and generated data; big data (e.g., web harvested, remote sensing),
- software: scripts for data processing, data analyses, and simulations, applications

Data origins:

- obtained from external sources (e.g., purchased)
- generated/collected by IAMO researchers alone
- generated in multi-institute research consortia

2.1.2 FAIR data publishing in repositories

- IAMO encourages the publication of IAMO research data wherever possible.
- Preferably in registered² discipline-specific and/or data-type-specific curated repositories (e.g., SowiDataNet, Pangaea.de, qualiservice.org)
- The IAMO researcher’s responsibility is to choose the repository in consultation with their department’s RDM-coordinator³, register and upload the data and metadata according to the respective repository’s organisational, format, and documentation requirements.
- In the specific case of the [GESIS SowiDataNet](#) repository for quantitative social science data, the RDM coordinators on the department level are involved in the role of IAMO curators.

2.1.3 Safe archiving: RADAR cloud or in-house

Data not FAIR-published must be archived to facilitate accessibility for ten years. This serves two purposes: a) checking results (motivation 1b mentioned in section 1) b) potential re-use at IAMO

² Registry of Research Data Repositories, re3data.org

³ See footnote 1.

(motivation 5). The researcher is responsible for initiating and completing data archiving (including metadata). The respective department's RD coordinator is available for consultations in this process.

- The preferred location for archiving research data is the [RADAR repository](#) of Fiz Karlsruhe.
- For archiving data which cannot be archived with RADAR (for legal or technical reasons), IAMO will provide storage space (disk devices or long-life-DVD) upon request. In this case, metadata will be only stored in the IAMO research data directory (implemented in IAMO's FIS, see section 2.2).

2.2 Internal list of all IAMO research data

All IAMO research data, i.e., data published (2.1.2), archived (2.1.3), or stored in other locations, must be referenced in the IAMO internal *research data directory* which is implemented in the Research Information System (FIS) accessible via IAMO's internal webpage. This part of the FIS database facilitates findability of all IAMO research data for result checks and re-use. Additionally, it allows monitoring and reporting this type of IAMO research output. Once the FIS research data module has been technically implemented, entering metadata (including information on the location of files) will be required. Metadata can be entered describing single standing data sets or in case of publications as part of registering a journal publication.

Colleagues leaving IAMO have to sign a statement indicating that they have saved all their IAMO research data and registered them in the IAMO research data directory or that they will do so after finishing the work on their IAMO projects.

The figure in the appendix visualizes the procedure described, i.e. the workflow for archiving/publishing and documenting research data at IAMO.

2.3 Provision of the Research Data Management Organizer tool

IAMO provides its own instance of the *Research Data Management Organizer* tool (accessible from the IAMO network under rdmo.iamo.de) and encourages its use. The platform for managing information on research data management assists in research data management in general and in the preparation of data management plans (as frequently required in research funding proposals) in particular. Using the tool helps ensure that all relevant RDM issues are taken care of.

2.4 Assistance and training in RDM

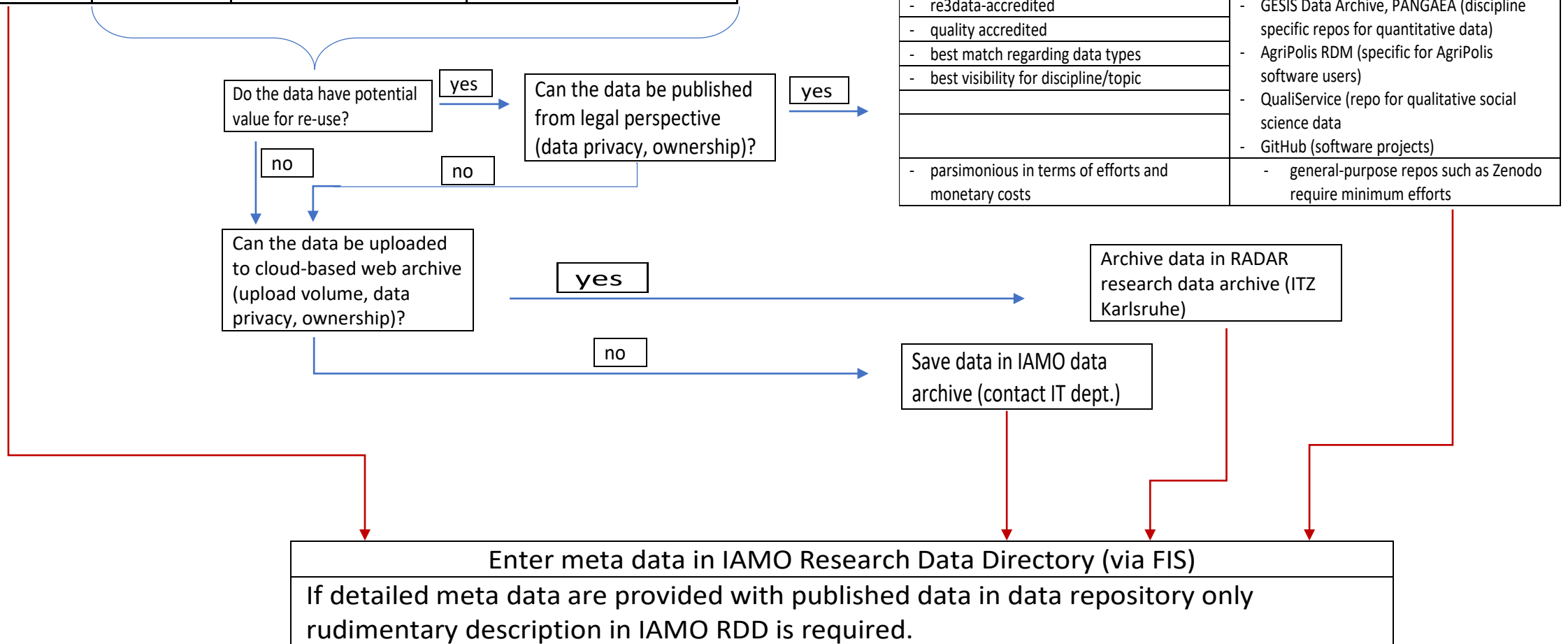
IAMO offers support to researchers in RDM issues (e.g., in preparing research data management plans, generally by using RDMO) or helps find experts for remaining unclear matters. The first person to approach is the RD coordinator on the department level. In-house archiving of data (2.1.3) is facilitated by the IAMO library. The RD coordinators also act as IAMO reviewers/curators for data sets submitted for publication in SowiDataNet or for archiving in RADAR.

The RDM working group initiates, organizes, and promotes trainings in RDM topics (e.g., data organisation, processing documentation (incl. versioning), metadata preparation, data publication, and data management plan preparation).

RD coordinators and the contact person in the IAMO library for in-house archiving are appointed by the department heads.

Appendix: IAMO Workflow to adequately share / preserve research data

Raw data		self-generated (from survey / web search / web crawl / computer simulation)	Processed data
obtained from external source			Either the processed data themselves or the material (description, software code) required for reproducing the processed data from raw data
Reliably long-term available from external source	Not reliably long-term available from external source		



Publish in appropriate research data repository	
Criteria for repository choice:	Examples
- re3data-accredited	- GESIS Data Archive, PANGAEA (discipline specific repos for quantitative data)
- quality accredited	- AgriPolis RDM (specific for AgriPolis software users)
- best match regarding data types	- QualiService (repo for qualitative social science data)
- best visibility for discipline/topic	- GitHub (software projects)
- parsimonious in terms of efforts and monetary costs	- general-purpose repos such as Zenodo require minimum efforts