**ZALF DMP Template adjusted to fit the DFG requirements**

Version of 15th June 2022

Based upon the DFG “Checklist Regarding the Handling of Research Data” published 21.12.2021. <https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/>

User note: the following DMP already contains basic information on ZALF data management, serves as a template and must be adapted to the respective requirements arising from the project. In particular, the notes marked in yellow must be supplemented and adapted. This DMP will be provided to ZALF WG RDM if funding is approved.

**0. Project and contact information**

Project name: \_\_\_\_\_please add\_\_\_\_\_\_

Project ID/Funding ID: \_\_\_\_\_please add if known, otherwise omit\_\_\_\_\_\_

Projectleader: \_\_\_\_\_please add\_\_\_\_\_\_

ID like ORCID: \_\_\_\_\_please add\_\_\_\_\_\_

Contact: Tel: \_\_\_\_\_please add\_\_\_\_\_\_

Project description: \_\_\_\_\_please add or reference\_\_\_\_\_\_

Research funding agency: German Research Foundation

Funding Program: \_\_\_\_\_please add\_\_\_\_\_\_

**1. Data description:**

*How does your project generate new data?*

*Is existing data reused?*

*Which data types (in terms of data formats like image data, text data or measurement data) arise in your project and in what way are they further processed?*

*To what extent do these arise or what is the anticipated data volume?*

The following data (parameters, columns with values) are collected, processed and published:

* Experimental data: Measured values, ...
* Simulations: complete input files, result files, ...
* Interview data: Question, target group, ...
* Secondary data: Contents, sources, associated processing scripts.

The following third-party data will be reused in the project:

* ...
* …

Further processing of the data is mainly done with the usage of \_\_\_please list the current methods of data processing used here\_\_\_\_\_\_. Scripts used are part of the research work and will be published together with the data according to good scientific practice.

The aforementioned data will be stored and processed in the following formats: \_\_\_\_\_\_\_please name the intended formats\_\_\_\_\_\_\_. These formats are compatible with the BonaRes Repository and there are no barriers to publish the data and thus to reuse it. We assume that the volume of data collected in the project is limited to \_\_\_\_\_XXTB\_\_\_\_ and the volume of published data is limited to \_\_\_\_\_\_\_XXTB\_\_\_\_\_\_.

**2. Documentation and data quality:**

*What approaches are being taken to describe the data in a comprehensible manner
(such as the use of available metadata, documentation standards or ontologies)?*

*What measures are being adopted to ensure high data quality?*

*Are quality controls in place and if so, how do they operate?*

*Which digital methods and tools (e.g. software) are required to use the data?*

All data are stored in a structured manner after collection and described with standardized metadata according to the BonaRes metadata schema (<https://doi.org/10.20387/bonares-5pgg-8yrp>). The metadata incorporate the DataCite (<https://schema.datacite.org/meta/kernel-4.1>) and INSPIRE (<https://inspire.ec.europa.eu/metadata/6541>) standards and are enhanced with discipline-specific metadata elements to foster scientific reuse. The metadata incorporate the [AGROVOC](https://www.fao.org/agrovoc/) (FAO) and [GEMET](https://www.eionet.europa.eu/gemet/en/themes/) (EEA) controlled vocabularies to ensure international interoperability. The entire history of data generation, processing, provision, and potential subsequent changes (updates) will be captured as "lineage" in the metadata.

Data quality is monitored and ensured during the course of data collection and processing through the following methods: \_\_\_\_\_please add\_\_\_\_\_\_

In the course of publication in the BonaRes Repository, formal aspects of data quality (structure, format) and the metadata are checked and, if necessary, corrected. The expertise available in the BonaRes Repository in the field of curation of agricultural data is used to check the content of the data for plausibility. Data are provided in common, non-proprietary formats (CSV, TXT, XLSX, GDB) for ease of reuse; no special tools or software are required (\_\_\_\_\_ check if this is guaranteed\_\_\_\_\_\_).

**3.Storage and technical archiving the project:**

*How is the data to be stored and archived throughout the project duration?*

*What is in place to secure sensitive data throughout the project duration (access and usage rights)?*

The data generated by this project are stored at ZALF and are subject to German law. ZALF's data storage infrastructure is based on a multi-level approach. Active data is stored on high-performance flash disks (ZALF Storage), which provide immediate and fast access. Primary data is backed up several times a day by so-called snaphots, which enable fast recovery. Primary data is replicated by mirroring in a physically separated data center (on the ZALF site) for data recovery purposes. In the event of a failure, this enables continued operation via the mirror data center. The maximum data loss is up to < 5 min (Recovery Point Objectives: RPO); the time required for recovery is < 2h during working hours (Recovery Time Objectives: RTO). If required, encryption of the data can provide further data security (\_\_\_\_\_ check whether this is necessary and, if so, coordinate with ZALF IT beforehand. Select alternative and insert: a) The encryption of the data offered by ZALF for further security is not necessary for the project. Insert justification if necessary; b) In the project, sensitive data is additionally secured by encryption. \_\_\_\_\_\_).

In addition, daily incremental and weekly full backups are saved to magnetic storage (hard disk), with a retention period of four weeks. In addition, a full backup is saved to magnetic storage (tape) once a month. The storage media are kept off-line in a separate location and are therefore not directly accessible. The RPO for this backup is 12 months.

Data access to all data stored at ZALF is safeguarded by an extensive and fine-grained user administration.

**4. Legal obligations and conditions:**

*What are the legal specifics associated with the handling of research data in your project?*

*Do you anticipate any implications or restrictions regarding subsequent publication or accessibility?*

*What is in place to consider aspects of use and copyright law as well as ownership issues? Are there any significant research codes or professional standards to be taken into account?*

The ZALF Data Policy (<https://doi.org/10.5281/zenodo.2567723>) defines relevant terms, rights and licenses, the handling of research data by a detailed data management plan and responsibilities. The ZALF data protection officer is involved in this process (datenschutz@zalf.de).

Legal specifics for handling and publishing data are not available \_\_\_\_\_\_please check\_\_\_\_\_\_\_

Strong restrictions on the subsequent use of data do not exist \_\_\_\_\_\_please check\_\_\_\_\_\_\_

All data are described with extensive metadata. Part of this metadata regulates the conditions of reuse (access and usage rights). The ZALF approach is to provide data as freely as possible. Therefore, the CC-BY license of the Creative Commons is assigned by default. It is an option for ZALF authors to embargo and thus restrict access to the data for a limited period of time. This takes into account the requirements for the right of first use.

**5. Data exchange and long-term data accessibility:**

*Which data sets are especially suitable for use in other contexts?*

*Which criteria are used to select research data to make it available for subsequent use by others?*

*Are you planning to archive your data in a suitable infrastructure? If so, how and where? Are there any retention periods?*

*When is the research data available for use by third parties?*

The data described in (1.) can be reused for the following purposes (examples): \_\_\_\_\_\_\_\_\_\_please specify in detail further possible uses\_\_\_\_\_\_\_\_\_

We follow the principle of making the acquired data available as extensively as possible for reuse. The publication takes place within 24 months after the collection and quality control at least before the end of the project. Further specifications are defined in the Project Data Policy \_\_\_\_\_ check and adapt\_\_\_\_\_. The most important criteria for publication is to publish only quality-checked data (see (2.)). If the data is sensitive, it will be sufficiently anonymized before publication. Unverified raw data will not be published. All data published to the BonaRes Repository will be permanently available for a long time. The database underlying the BonaRes Repository is totally integrated into the ZALF storage infrastructure (see (3.))

**6. Responsibilities and resources:**

*Who is responsible for adequate handling of the research data (description of roles and
responsibilities within the project)?*

*Which resources (costs; time or other) are required to implement adequate handling of research data within the project?*

*Who is responsible for curating the data once the project has ended?*

In the project, the designated Datasteward (0.5 FTE \_\_\_\_\_please adapt\_\_\_\_\_\_) is responsible for all internal research data management issues, data publication and contact to ZALF data service.

Datasteward: \_\_\_\_\_please add\_\_\_\_\_\_

ID like ORCID: \_\_\_\_\_please add\_\_\_\_\_\_

Contact: Tel: \_\_\_\_\_please add\_\_\_\_\_\_

\_\_\_\_\_\_\_\_Please describe further roles in the project\_\_\_\_\_\_\_\_\_

*Help concercing the different types of roles can be found here (page. 20):*

[*https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel\_v4.4.pdf*](https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel_v4.4.pdf)

In the project, the following additional resources are provided to ensure professional data management with the goal of data publication (e.g., deliverable: data policy; each author's share of data management; staff training): \_\_\_\_\_\_explain in detail\_\_\_\_\_\_\_\_

The project is supported in data management by the ZALF service facilities RDM and DIS. They will also be responsible for data curation of the published data after the end of the project.

*Information for your proposoal: Costs for data preparation and following up works concerning data should be calculated within the project. Further information can be found here: https://www.dfg.de/foerderung/grundlagen\_rahmenbedingungen/forschungsdaten/beantragbare\_mittel/index.html*