Guidelines on the Handling of Research Data in Biodiversity Research

Purpose:

This document specifies the DFG's expectations for research data management in its funded projects within the scope of biodiversity research. It is a recommendation meant to inform applicants as well as reviewers and the involved DFG committees.

Plans for research data management are to be described adequately in proposals and should be considered in reviews. Interim and final reports should give account of the actual research data management in the project.

Motivation:

Research data management comprises the planning, acquisition, processing and storage of data collected and generated within the scope of research projects. It assures the quality of and access to data and enables reproducibility and re-use.

The DFG aims to "enhance the long-term archiving and curation of research data" and therefore "funds projects that seek to achieve an efficient re-use of research data" (see e.g. section 2.4 of the Proposal Preparation Instructions for Project Proposals). Consequently, it expects corresponding statements in project proposals. Data represent a valuable resource whose loss cannot be compensated. The utmost attention should be paid to their administration, securing and provision. This should be done in a way that assures, as much as possible, the reproducibility of results as well as the re-use of data, also for purposes other than those for which they were initially collected. In order for this to succeed, data management issues should already be taken into account when applying and consistently be considered in the course of the project.

The following aspects should be taken into consideration:

1. Enabling free public access to data deriving from DFG-funded research should be the norm. Restrictions due to legal, copyright or ethical aspects will be approved after corresponding justification.
2. In order to actually enable re-use, stored data should be quality-assured and adequately described.
3. All research projects/proposals should include a data management plan. The plan should — to the extent applicable — comprise the following information:
    a) whether, and if so, with what effort, the data are reproducible (onetime observations, repeatable experiments);
    b) kind (individual, tissue, etc.) and type of data (picture, audio, text, source code, numbers);
    c) how/with which tools the data will be gathered and evaluated/processed;
    d) file formats; the use of open or openly documented formats is recommended; if data are only legible with special software, the software has to be documented or included in the database (if permitted under copyright);
    e) documentation and description of the data (context of the investigations, methods used, etc.); these should be aligned with standards;
    f) how the data will be administrated, stored and secured while the project is in progress;
    g) how quality assurance of the data will be implemented;
    h) the connection to research objects (e.g. voucher specimen or soil samples) and other referenced data;
    i) who, besides the applicants, will be responsible for research data management;
    j) how, where and for what period the data will be made available for re-use; how it will be ensured that the data are findable, accessible and re-usable; alternatively, an explicit explanation as to why the data are not suitable for re-use, or why re-use is only possible with restrictions;
    k) when the data will be made available for re-use;

l) on which conditions re-use will be rendered possible and, if applicable, who will decide this in the long term;
m) costs of research data management during the project and of the provision of data for re-use, and how these costs will be funded.

4. All collaborative research projects should establish specific data guidelines that regulate within the project the provision, transfer and re-use of data, as well as the approach to and consequences of non-compliance.

Authors: Members of the Working Group on Data Management of the DFG Senate Commission on Biodiversity Research