Czech Academy of Sciences, Institute of Experimental Botany Research Data Management Policy

Preamble

The Institute of Experimental Botany of the Czech Academy of Sciences (IEB CAS) recognizes that the research data are fundamental parts of the research activities, and the management of research data following FAIR principles is a key practice of research integrity. The Institute is committed to ensuring that its research is transparent and that its research outputs are widely accessible, reproducible, and reusable. The Institute is of the view that this enhances the quality and impact of research activities benefiting both the wider research community and individual researchers. Furthermore, this highlights the excellence of the research activities at the institute and enables public engagement with them.

1. Aims

This policy aims to define the basic principles of research data management at all stages of the research data life cycle, mainly: data collection, storage, preservation, and sharing. It is also a statement of the responsibilities of the institute and its researchers concerning the management of research data. This policy further provides information on available institutional supports for researchers to meet the standards outlined in the policy. The Institute recognizes that there are different discipline-specific norms across the diverse spectrum of research. Therefore, the policy does not oblige researchers to use prescribed tools when they work with research data. It aims to promote good practice in research data management and provide guidance to researchers. The policy may be complemented by other policies of the institute.

2. Scope of application

The policy applies to all employees and students at the institute, and affiliated persons and subjects conducting or supporting research activities at the Institute. All research data regardless of their form are covered by this policy.

3. Definitions

- **Research data**. Any information collected, observed, generated, or created to validate or reproduce research findings is described as research data. Research data can include, but is not limited to, documents, spreadsheets, images, audio/video, code, and software.
- **Metadata**. All the information and descriptions about the data. For example: type of the data, title, authors, year, language, keywords, and description about the under which conditions the data are collected.

- **Data Management Plan (DMP)** is a document that outlines mainly how research data will be collected, handled, preserved, and published during and after a certain research project. DMPs are submitted usually with grant proposals and updated with each interim and final report. DMP should briefly define:
    - How the data will be generated and documented?
    - Where and how it will be stored?
    - Who will have access to it?
    - What are the backup strategies and who is responsible for them?
    - How it will be preserved and shared/published?
- **FAIR Principles:** The FAIR Principles describe how research data should be organized to make it more **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable.
- **Repository:** Digital online storage for storing and sharing the results of creative activities (e.g., publications or research data). They can be:
    i. Discipline-specific
    ii. Institutional
    iii. Multidisciplinary
- **Persistent identifier:** Long-lasting reference to a unique entity. For example, Digital Object Identifiers (DOI), handle, ORCID, ResearcherID, ROR, etc.
- **Researcher:** The term "researcher" shall include all members of the Institute, including staff and students, and any affiliated persons who are conducting research at the Institute or on behalf of the Institute.
- **Principal Investigator**: A researcher who has overall responsibility for the research project.

4. Basic principles

Researchers will make all reasonable efforts to manage their research data in accordance with the FAIR principles. The collection, processing, and sharing of research data must not be in breach of intellectual property rights, the General Data Protection Regulation (GDPR), the cyber security legislation and must be in accordance with other legal, institutional, and contractual requirements.

4.1. Data Collection, Storage, and Preservation

- Research data must be stored in a secure location to prevent unauthorized access or loss following the data security guidelines.
- Where applicable, research data should be accompanied by rich metadata using standardized vocabularies. They should also be stored in standard formats to enhance interoperability.
- In the event that researchers are engaged in collaborative research with external partners, it is recommended that they draw up an agreement outlining the rights and responsibilities of the parties involved. This could include matters such as intellectual property rights and licensing, or responsibilities for the management of research data.
- In the event of an exchange of personal data with a third party where the IEB CAS is the data exporter, an agreement must be established to ensure the protection of the data.

- To guarantee the completeness, accuracy, reliability, and security of data, researchers are advised to develop a DMP for their research projects.
- The IEB CAS uses the Fair Wizard instance of the Data Stewardship Wizard (DSW) software as the main tool for documenting and updating DMPs.
- It is recommended that research data that serve as a basis for a publication be preserved for at least 5 years from the day the research results are published, so that the results may be verified. It is recommended that researchers retain their research data for as long as is feasible. If the need arises to delete data at an earlier stage, for instance under contractual obligations, this information should be included in the documentation.
- Research data being preserved must be accompanied by sufficient documentation to ensure wide and diverse audiences can easily interpret them.
- Research data must be preserved in secure, well-recognized, and trustworthy storage systems that

4.2. Data sharing/publication

Researchers are encouraged to make their data available for access and reuse as widely as feasible following the principles outlined below:

- The rich metadata is to be provided with shared data to increase FAIRness and provide sufficient information on the provenance of the research data. The metadata should include references, via persistent identifiers, to other related outputs and entities.
- Research data should be assigned a persistent identifier, and shared in a trusted, well-recognized open repository or an appropriate platform established within the specific research field. The appropriate repositories can be identified in the Registry of Research Data Repositories (re3data.org). In addition to dedicated databases under NCBI and multidisciplinary repositories like Zenodo, and DRYAD, the following systems also can be used:
    I. Image data can be preserved and published in the BioImage Archive and Image Data Resource (IDR).
    II. Proteomic and Mass spectrometry data can be preserved and published in the PRoteomics IDEntifications Database (PRIDE), Mass Spectrometry Interactive Virtual Environment (MassIVE), or other dedicated databases.
    III. DNA/RNA sequencing data can be preserved and published in the European Nucleotide Archive (ENA), Sequence Read Archive (SRA), GenBank, and other dedicated databases.
- The ASEP Institutional Data Repository is primarily recommended for storage, archiving, and publishing. It is also a primary system for such data that is intended to be shared only within the institute.
- To ensure that the conditions for reuse are specified clearly, research data must be licensed in compliance with intellectual property rights, unless funder requirements, statutory or contractual obligations provide otherwise. Open licenses such as Creative Commons Attribution (CC BY) are recommended.

- As far as it is appropriate, the published research articles should include a "Data Availability Statement" that outlines how the associated data may be accessed.

5. Rights and Responsibilities

5.1. Rights and Responsibilities of the Researcher

- Managing scientific publications, research data, and educational resources in accordance with the principles and requirements outlined in this Policy. Preserving machine-readable electronic copies of the scientific publications in the ASEP bibliographic database, if applicable, including their supplementary files, and persistent identifiers of associated research data published in any other repository, archive, or database.
- When new students are enrolled in the research groups, the Principal Investigator, supervisor, or designated member of the group is responsible for introducing them to the basic principles of handling the data they work with.
- Ensuring compliance with the organizational, regulatory, institutional, and other contractual and legal requirements related to the production, curation, management, deposit, and distribution of publications and data if no other agreement with third parties takes precedence.
- Choosing the appropriate type of licensing for their research outputs.
- Adherence to the established citation guidelines when reusing or referencing datasets.
- If researchers encounter difficulties in fulfilling the requirements outlined in this policy due to a lack of appropriate resources or support from the Institute, they are encouraged to seek guidance from the related departments or support teams listed in Chapter 6.

5.2. Rights and Responsibilities of the Institute

- Disseminating information about the obligations of researchers concerning research data management.
- Supporting and empowering the acquisition of skills, and adopting tools related to good practices in research data management. Provide training and guidance to promote best practices in research data management.
- Establishing and maintaining an Institutional Open Access Repository or collaborating with consortia or national repositories containing digital content and providing advanced tools for search, navigation, and Open Access to its content according to international standards.
- Appointing a Data Manager responsible for all data-related matters, including – but not limited to – issues related to the development of Data Management Plans (DMPs), data curation, and compliance with national and European laws.
- Developing and providing services for the storage, safekeeping, registration, deposition, and sharing of data and other records, according to the FAIR principles, as well as their long-term preservation, and providing appropriate guidance to researchers.
- Engaging with the research community to discuss and identify the needs in terms of support and infrastructure.

6. Available supports

Various supports including consultation services on Research Data Management, and Open Science practices are available for all researchers of the Institutes of the Czech Academy of Sciences https://openscience.lib.cas.cz/en/support/.

Moreover, researchers can find support related to research data management in the following departments at the IEB CAS:

1. Data Management Plan & FAIR Wizard tool: caryjeva.a@ueb.cas.cz
2. IT support: nulicek@ueb.cas.cz
3. Department of Project Support and Technology Transfer: hroncova@ueb.cas.cz
4. Institutional data repository (ASEP): caryjeva.a@ueb.cas.cz
5. Institutional library: knihovna@ueb.cas.cz
6. ASEP bibliographic database: klier@ueb.cas.cz

7. Review and Revision

This policy will be reviewed periodically to ensure its continued relevance and effectiveness. Any revisions and amendments will be communicated to the research community within the Institute.