

MIRRI-ERIC Policy on Data Management

To overcome the current situation, the MIRRI Information System (MIRRI-ERIC IS), as part of the Collaborative Working Environment (MIRRI-ERIC CWE), will be established by deploying an integrated, high-quality, automatically validated, manually annotated, semantic-rich, non-redundant micro-biological resource database which provides all relevant information and associated contextual data (metadata) about a particular biological resource. MIRRI-IS will be designed as the central entry point for users, curators and developers that need access to the integrated knowledge of mBRCs and selected third-party databases while assuring that the specific competences of partner mBRCs remain transparent. The aim is to establish a trademark for high quality data and expertise, which enhances the reputation of participating mBRCs. The MIRRI-ERIC policy on data management is a commitment to a FAIR (Findable, Accessible, Interoperable and Reusable) provision of data and information.

Requirements for mBRCs to comply with the MIRRI-ERIC policy on data management:

To allow MIRRI-IS to be operational, MIRRI-ERIC partners need to comply with:

1. Machine-readable mBRCs catalogs.
2. In case information is not digitally available yet, proper digitalization of key information needs to be undertaken.
3. Provision of accurate data.
4. The MIRRI Minimum Data Set (MIRRI MDS) of descriptors include 1) Strain Number, 2) Other Strain Number, 3) Present Name, 4) Organism Type, 5) Restrictions, 6) Status, 7) History of Deposit, 8) Growth conditions, 9) Form of supply, 10) Geographic Origin and 11) additional accession number(s) to link the data to the International Nucleotide Sequence Database Collaboration (INSDC), in case this is available. Besides these fundamental fields, specific “data packages” and additional subfields will be added over time to enrich the MDS. This will be extended towards a recommended data set (RDS) and finally full data set (FDS).

5. The content of the fields is expected to follow the guidelines, data model, controlled vocabularies and ontologies specified by the MIRRI-ERIC consortium.
6. The final set of fields, including their expected content, will be consolidated in the Minimum Information about Biological Resources (MiaBRe) data standard and checklist.
7. Curation level and quality of data needs to be assured by unified Standard Operating Procedures in mBRCs.
8. Provision of the data in a structured electronically available format.
9. For each biological resource, data needs to be made available in machine-readable format and in regular time intervals. Over time, each mBRC in MIRRI-ERIC should provide their data by Web Services in an XML based exchange language, e.g. based on the Microbiological Common Language (MCL) and its extensions.