

# Data Centers

## About the Data Center Network

The **ten Data Centers listed below** are infrastructure partners in the GFBio broker network. They entail seven Data Centers located at natural history collections throughout Germany and three Data Centers specialised in plant, nucleotide and environmental data. The Data Centers are departments of science informatics and data infrastructures at recognised science institutions devoted to manage, store, archive and publish various types of bio- and geodiversity data. Data submitted through GFBio are transmitted to and curated by data curators at a matching GFBio Data Center, based on the profiles below. Data are available via various facets of the [GFBio search portal](#), as well as on several other portals dependent on the Data Centers' scope, such as GBIF.

The core group of **ten GFBio data centers** has agreed on a number of consensus documents, tools, data pipelines and standards and technical formats for interoperability.

Wondering if your data is suitable for archival & publication in one of the data centers? [Get to know the Data Centers](#) below !

Ready to archive your data? **Submit your data** via the [GFBio submission service](#) ↗

Questions regarding your data? [Contact our Helpdesk!](#) ↗



## Data Centers

## Data Centers specialized on Plant, Nucleotide and Environmental Data

**e!DAL-PGP** – Plant Genomics and Phenomics Research Data Repository



e!DAL-PGP archives, curates and publishes cross-domain, plant-related research data that exceeds existing repositories due to their size or scope. This includes for example:

- image collections from plant phenotyping and microscopy
- unfinished genomes
- genotyping data
- visualizations of morphological plant models
- data from mass spectrometry
- software & documents

[Learn more about the data center e!DAL-PGP](#)

**ENA** – European Nucleotide Archive



The [European Nucleotide Archive](#) archives, curates and publishes nucleotide sequence data and associated metadata.

The GFBio Brokerage Service provides the timely, standards-compliant deposition of all molecular sequence data into the public repositories of the INSDC ([DDBJ](#), [EMBL-EBI](#) and [NCBI](#)). The key components of the service include: (a) Support for metadata standardization, curation and quality control, (b) negotiation of embargo periods, including communication with INSDC, (c) parallel submission of environmental metadata to [PANGAEA](#) and other [GFBio data centers](#), (d) cross-linking sequence and environmental data (e.g. [PANGAEA](#)) or other contextual or related data via accession number and DOI.

[Learn more about the data center ENA](#)

**PANGAEA** – Data Publisher for Earth & Environmental Science



PANGAEA archives, curates and publishes multidisciplinary (e.g. geochemical, biological observational and occurrence) data from marine and terrestrial environments. Curation includes user support, definition of data set granularity, quality control, archival format transformation, metadata description and control. Supported data types are tabular data but also binary data, e.g. multimedia.

[Learn more about the data center PANGAEA](#)

## Data Centers at Natural Science Collections

**BGBM** – Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin

**DSMZ** – German Collection of Microorganisms

The BGBM has a strong interest in collection and observational data as well as taxon-level data management covering the entire taxonomic workflow, in particular in the field of Botany. The taxonomic expertise and interest is that of the BGBM (see <http://www.bgbm.org/en/research>). Taxonomic research, through which plants are identified, described, named and classified, is a central element of our daily work.

*Data archived:*

- Botanical specimen data; Referenced multimedia objects
- Botanical observational data
- Botanical systematics and monographic works
- Experimental raw data (data sets and/or data packages) only if well documented and in formats and structures appropriate for long-term archiving

[Learn more about the data center BGBM](#) 

The DSMZ data center focusses on providing organism-linked information covering the multifarious aspects of bacterial biodiversity.

Data archiving for research projects:

- Data accompanied by the deposit of a biological resource (bacterial or cell culture lines) within the DSMZ collections ([Deposit in the DSMZ](#))
- Data describing microbial diversity (e.g. taxonomic classification, morphology, physiology, cultivation, origin, natural habitat)

[Learn more about the data center DSMZ](#) 

**LIB** – Leibniz Institute for the Analysis of Biodiversity Change

**MfN** – Leibniz Institute for Research on Evolution and Biodiversity, Berlin

blocked URL



The [Leibniz Institute for the Analysis of Biodiversity Change](#) is a foundation under public law. The [Biodiversity Data Center](#) as part of the LIB is aimed at hosting, archiving, publishing and distributing data from biodiversity research and zoological collections.

The Biodiversity Data Center handles and curates data on:

- The specimens of the institutes collection, including provenance, distribution, habitat, and taxonomic data.
- Observations, recordings and measurements from field research, monitoring and ecological inventories.
- Morphological measurements, descriptions on specimens, as well as
- Genetic barcode libraries, and
- Genetic and molecular research data associated with specimens or environmental samples.

[Learn more about the data center LIB](#) 

MfN has a strong interest in collection and observational data, taxonomic data, and trait data. MfN curates the following data regularly:

1. MfN collection related data, such as collection specimen data, DNA (DNA collection - exceeding MfN physical collection), collection metadata, and media representations of MfN collection specimens including additional metadata
2. Research project related data (MfN involvement, leadership or partnership, national or international), such as observational data (e.g. ATBI)

[Learn more about the data center MfN](#) 

**SGN** – Senckenberg Gesellschaft für Naturforschung

**NSNB** – Staatliche Naturwissenschaftliche Sammlungen Bayerns

# SENCKENBERG

world of biodiversity



SGN has various interests in different data domains, including: Collection data, molecular data, observational data, environmental data, series of measurements and trait data. Data archiving for research projects is focusing on botanical, zoological and anthropological data, especially collection data in combination with the deposit of

- physical objects
- referenced multimedia objects.

[Learn more about the data center SGN](#) 

The [Staatliche Naturwissenschaftliche Sammlungen Bayerns](#) (SNSB) is a natural history collection facility in Bavaria. The [SNSB IT Center](#) as part of the SNSB is its institutional repository primarily for scientific bio- and geodiversity data of the natural history collections belonging to the SNSB.

The mission comprises research activities in the field of biodiversity informatics and data science. Software is mainly designed and set up following the concepts of the [Diversity Workbench](#) (DWB). DWB software tools are registered in [bio.tools](#), a service of [ELIXIR Europe](#).

The SNSB IT Center as [GFBio Data Center](#) supports scientists and institutions by offering DWB support and workshops. Additional services are provided on a case-by-case basis. They might include the sustainable DWB management of data from its generation up to persistent storage, archiving and publication of approved, quality-controlled, standardized and well-structured, i.e. FAIR

- occurrence and provenance data (from specimens, biological samples and observations)
- taxonomic and checklist data
- trait data (e.g., morphological, anatomical, chemical and molecular descriptions)

[Learn more about the data center SNSB](#) 

**SMNS** – State Museum of Natural History Stuttgart

**NATURKUNDE  
MUSEUM  
STUTTGART**



Research at the SMNS is based on our scientific collections of fossils, plants, insects, molluscs and vertebrates. Taxonomic research, through which organisms are identified, described, named and classified, is a central element of our daily work. The SMNS has a strong interest in collection and observational data as well as in genomic research data.

Data archiving for research projects focusing on botanical, zoological and paleontological data:

- Collection data, together with the deposit of physical objects, referenced multimedia objects.
- Observation and occurrence data, species monitoring projects, referenced multimedia objects.
- Taxon reference list data and checklist data.

[Learn more about the data center SMNS](#) 



**Do you have questions, feedback or need help?**

[Contact our Helpdesk](#) for direct support.

