UFZ Data Infrastructure

The Helmholtz Centre for Environmental Research (UFZ) operates a set of digital infrastructures covering the research data life cycle, from collection to long-term archiving and publication. Some are general-purpose and used across research groups; others serve specific domains such as biodiversity or ecotoxicology. Below they are grouped by function rather than by owning unit. Detailed, how-to level documentation for these infrastructures is maintained in the UFZ RDM Guidelines.

Spatial Data Infrastructure (SDI)

UFZ's SDI (spatial.IO) combines several components into automated workflows for storing, managing and delivering geospatial (vector, raster, time-series) data:

> Weber, C., Lange, R., Biskamp, B., Panda, M., Schnicke, T., & Bumberger, J. (2026). spatial.IO – An integrated cloud-ready geospatial data management system (v1.0.1). Zenodo. https://doi.org/10.5281/zenodo.10391523

- **S3 storage** – cloud object storage, the common backend also used by time.IO and OMERO
- **GeoServer** – OGC web services (WMS, WCS) including NetCDF support
- **THREDDS Data Server** – OPeNDAP access to gridded/raster data
- **WebGIS** – browser-based map viewer built on the OGC SensorThings API and PyGeoApi
- **GeoNetwork** – the metadata catalogue; also UFZ's entry point for connecting to external metadata federations (see below)

Time Series & Sensor Data Infrastructure

- **time.IO** – time series management: automated sensor data collection, standardized exchange protocols, quality control. Data is ultimately published via spatial.IO; PIDs are handled via the SMS.
- **Sensor Management System (SMS)** – manages sensor metadata and controlled vocabularies; currently the only infrastructure that captures PIDInst identifiers for instruments.
- **SaQC** – quality control toolkit for numerical/time-series data, used across the time series infrastructure.
- **Grafana** – dashboards for time series data.

Sample, Lab & Archiving Infrastructure

- **Data Management Portal (DMP) / Data Investigation Portal (DRP)** – administers logger data, sample and analysis data, and cultivation data; the archive component is UFZ's route to DataCite DOI registration for research data (UFZ has been a member of the TIB DOI Consortium since 2021).
- **Labspace** – electronic lab notebook (ELN) with sample and inventory management, successor to eLabFTW.

Imaging Data Infrastructure

- **OMERO** – stores, visualizes, manages and annotates microscope images and their metadata; uses S3 as its storage backend.

Domain-specific Research Data Platforms

- **BioMe** – modular web platform supporting citizen science projects in biodiversity; several older UFZ biodiversity tools/projects (TMD, BiolFlor, LEGATO, ALARM) now live under it.
- **BEXIS2** – manages biodiversity researchers' data across the full life cycle, from collection to sharing and publication.
- **INTOB** – database and web application for toxicological effects of chemicals on organisms, supporting experiment planning and execution.
- **Chromeleon** – connects lab measuring devices through one interface for chromatography data collection and processing.
- **QualiService** (external, hosted elsewhere) – archives qualitative social science research data for reuse.

Analysis & Workflow Tools

- **KNIME** – low-code data science workflow platform.
- **Galaxy** – UFZ instance of the open-source scientific workflow system for accessible, reproducible computational research.
- **WOMBAT** – Kubernetes-based architecture for running web applications (e.g. tool front-ends built on top of other infrastructures).

Research Infrastructures & Long-term Observatories

Several UFZ-affiliated projects generate and structure data at larger scale and often feed into the infrastructures above:

- **TERENO** – long-term observatory network.
- **eLTER** – (Integrated) European long-term ecosystem research infrastructure.
- **MOSES** – mobile observation platform complementing long-term observatories for investigating highly dynamic or extreme events; includes its own Data Discovery Portal.
- **MOSAIC** – platform for model-based, high-resolution exploration of subsurface structures via minimally invasive methods.

Persistent Identifiers & Metadata Integration

Status across UFZ's infrastructures (per the PID working group inventory, 2026):

- **ORCID** – supported (UFZ employees can link their ORCID to the personnel service; UFZ is an

institutional member of ORCID DE), but registration is voluntary and not systematically enforced across all data infrastructures.
- **DataCite DOI** – used for data publications via the DMP archive component.
- **PIDInst** – captured for instruments via the SMS only.
- **IGSN** (physical samples), **ROR** (organizations) – not yet systematically implemented; identified as gaps.
- No cross-infrastructure policy currently mandates consistent PID capture; responsibility is split across several infrastructures and, in places, unassigned.

Connections Beyond UFZ

- **ODIS-Connect** – ongoing project to connect UFZ's GeoNetwork to the Ocean Data and Information System (ODIS), making UFZ (meta)data findable in a global federation and serving as a blueprint for other GeoNetwork-based Helmholtz centers (e.g. HEREON) that are not yet ODIS-connected.
- **Helmholtz Knowledge Graph / unHIDE** – the ODIS work is explicit preparation for later harvesting of UFZ metadata into the Helmholtz-wide Knowledge Graph, which uses a related architecture.
- **NFDI, HIFIS & Helmholtz Cloud** – UFZ's own infrastructures are complemented by national (NFDI) and Helmholtz-wide (HIFIS/Helmholtz Cloud) services for storage, compute, collaboration and metadata (see the RDM Guidelines' infrastructure overview for the full list).

From:
https://earth-and-environment.helmholtz-metadaten.de/wiki/ - **HMC Earth and Environment Community Wiki**

Permanent link:
https://earth-and-environment.helmholtz-metadaten.de/wiki/doku.php?id=wiki:5_management:3_drafts:ufz_data_infrastructure

Last update: **2026/08/28 09:34**