

Hinweise

Hinweise zu diesem MoU

- *Neue Arbeitsversion, um eine zweite Version des MoU mit weiteren Konsortien aus der 3. Runde zu erstellen*
Zeitplan: abgestimmter Text bis Ende des Jahres
- *Unterschriften vor Zwischenbericht (3. Runde)*

v2.0

Memorandum of Understanding (v2)

The speakers of the consortia DataPLANT, NFDI4Microbiota, NFDI4Biodiversity, FAIRagro, NFDI4Objects, and NFDI4BIOIMAGE agree on the following areas of cooperation in the area of research data management for the biological sciences.

About the consortia

DataPLANT is a consortium for the plant science community, providing a state-of-the-art tool-box for managing, processing and publishing data packages early on in the research process following an open contribution and participatory model. Its overarching goal is to provide robust, easily accessible and user-oriented research data management practices, tools, and infrastructure to support collaborative plant biology research.

NFDI4Microbiota offers a broad set of solutions and services to the microbiological research community to make their data and other research output FAIR and open. This includes among others training, data storage solutions, data analysis services and RDM consulting.

NFDI4Biodiversity is a broad resource provider and expert network with a variety of services for biodiversity, ecological and environmental data. The consortium develops a common infrastructure and has intensive efforts in data mobilisation according to the FAIR principles, with focus on occurrence and molecular sequencing data (metabarcoding), as well as enabling the different method and data oriented subcommunities.

FAIRagro is developing a research data infrastructure for agrosystem sciences. Quality-assured and legally compliant datasets are connected and made discoverable, reusable and machine-actionable in accordance with FAIR principles. Support for data analysis and research data management is provided through a coordinated network of trainers, data stewards, and a central helpdesk.

NFDI4Objects develops solutions for research data management related to the material heritage of approximately three million years of human and environmental history. The consortium is aimed at individuals of all experience and career levels who wish to collaboratively shape the digital transformation of research data infrastructure within the involved disciplines in a sustainable manner.

Image data is central to biological research because it captures spatial, temporal, and functional relationships in high resolution, e.g., in disease research. **NFDI4BIOIMAGE** develops new formats, metadata standards, analysis methods, and processing environments for optimal use, e.g., in AI-supported pattern recognition.

Goals and areas of cooperation

Despite different subject areas and research contexts (field/lab), the target communities of the six consortia use similar methods and similar data types are relevant for their research. Based on their relevance for the community, the consortia's investments in solutions for certain data types and methods in the first three years of operation were quite complementary. This provides us with a solid basis to leverage resources.

Table 1: Complementary investment in major biological data types

The focus of investment of DataPLANT, NFDI4Microbiota, NFDI4Biodiversity, FAIRagro, NFDI4Objects, and NFDI4BIOIMAGE in community services, guidelines, blueprints, use cases, training, etc. for different biological data types were highly complementary in the first three years of operation.

Data type	DataPLANT	Microbiota	Biodiversity	FAIRagro	Objects	BIOIMAGE
Type 1: Collection data, species occurrence data	-	-	+++	+	+++	-
Type 2: Taxon data (catalogue, checklist or so-called red list)	-	+	+++	+	+	-
Type 3: biological and ecological study data including functional and phylogenetic trait data	+	++	+	+++	+	++
Type 4: non-molecular analysis data f.e. in laboratory research contexts	+++	+++	+	++	++	+++
Type 5: molecular (sequence) data	+++	+++	++	+++	+	+

Legend: +++ high; ++ solid; + not much in focus; - no priority so far

We therefore intend to use biological data, along with related software, tools, workflows, models, and other digital infrastructure, as the connecting element to harmonise research data management strategies for different research contexts, thereby facilitating access to services and other data-related resources in the field of biology, and developing synergies across consortia. The cooperation will be open to include other consortia with biological data.

Synergies may include but are not restricted to:

- Education and training
- Outreach and communication
- Service portfolio (reuse and co-development of services)
- User support
- Activities in the UN policy arena regarding digital sequence information

Living the collaboration: Activities

The consortia will implement a Biodata Interest Group within NFDI to govern joint activities and to include other consortia over time. The interest group shall review training materials and service developments to highlight synergies and complementarity, so that operational goals in the consortia can be aligned and co-marketing activities can be planned. It will develop cross-consortia collaborations on essential services and advocate for sustained funding of biodata services. The interest group shall be integrated in the work programmes and coordination shall alternate between the consortia.

- 1) **Joint partners and staff:** As all involved consortia have started or are approaching the second funding period, they will strengthen their partnership through shared co-applicants, participants and staff. Shared partners will allow the co-funding of staff positions for core activities in service development, marketing, and in the policy sector.
- 2) **Joint Helpdesk and User Support:** The consortia will coordinate their user support activities by aligning helpdesk structures and establishing shared triage and referral processes. This collaboration will be developed in close connection with the Base4NFDI proposal Support4RDM, which explores the networking of helpdesks across the NFDI.
- 3) **Joint infrastructure:** The consortia rely on collaboration with service centres and cloud infrastructure such as provided by de.NBI. They have a common interest to stimulate the further development of such infrastructure to meet the needs of their target communities and will provide new use cases encouraging the sharpening of portfolios for research data infrastructure providers in the life sciences. This will be facilitated through the service and cloud providers who are shared members of the consortia. There is also a joint interest in multicloud technologies and the Research Data Commons concept. This will be promoted jointly within the current negotiations about the NFDI architecture and basic services.

Additionally, as part of the joint infrastructure strategy, the consortia will explore the establishment of a shared data hub as a long-term, cross-consortia platform for collaborative data curation and publication. The institutions will determine and assume responsibilities for hosting, operation, and maintenance as the collaboration progresses.

- 4) **Joint programming for the uptake of services and standards:** The consortia will develop dedicated cross-consortia flexfunds projects and use cases, for example regarding the use of **FAIR Digital Objects**, e.g. ARC (the Annotated Research Context) or OME-Zarr; **workflows**, e.g., SciWin (Scientific Workflow Infrastructure), or CloWM (Cloud-based Workflow Manager); **benchmarking tools** (e.g., the CAMI benchmarking portal); **storage infrastructure** (e.g., Aruna); an **overarching search service**; **ontology portal** (e.g.; BiodivPortal and Bioportal),

and eventually TS4NFDI; the joint development of tools for the collection of service **usage metrics**; and the aligned implementation of **schemas** (e.g., bioschemas.org) by data and service providers in the consortia. These projects and use cases will be used to promote services across subject areas and to plan cross-consortia workshops and seasonal schools. For services marketed across consortia, joint user groups will be set up.

- 5) **Joint scientific conferences:** The consortia are organizing a joint conference. This will promote scientific exchange between the various consortia and involve the user communities. Possible main topics are data engineering, data management, interoperability concepts, data stewardship, data science, reference data sets and use cases. Such a conference should be organized alternately by at least one consortium and focus on the subject area of the respective consortium. This series and other meetings enable the exchange of ideas and discussion, allowing the consortia to refine their approaches, gain new perspectives, and build networks with the communities. They also inspire interdisciplinary learning and spark collaborations that lead to joint publications, project proposals, and data sharing initiatives.
- 6) **Joint publication activities:** The consortia will pursue coordinated publication activities, including the preparation of special issues in relevant peer-reviewed journals (e.g. Journal of Integrative Bioinformatics (JIB)) to disseminate outcomes of cross-consortia projects, conferences, and community initiatives.
- 7) **Pooling know-how for outreach purposes:** The consortia use branded knowledge bases to reach their target audiences and market their expertise and services. The consortia strive to pool capacities for the development and update of such resources, in order to achieve economies of scale and avoid duplication of efforts. Care will be taken to enable marketing in each consortium's community of interest, while providing easy access to comprehensive biodata resources for biology in general. Collaboration with existing knowledge brokers in the field, e.g. ELIXIR, forschungsdaten.info, WiNoDa, and GFBio e.V., will be pursued as part of this activity.
- 8) **Joint training:** The consortia will harmonize their training offerings and will try to complement each other's training portfolio. They will conduct co-hosted training events and will share and co-create training material.

Transparency

The partners provide information on the collaboration on their websites.

Metadata

Metadaten für die Publikation:

Reserved DOI: 10.5281/zenodo.13880944, <https://doi.org/10.5281/zenodo.13880944>

Autoren/Creator:

- Konrad Förstner (ZB MED - Information Centre for Life Sciences)
ORCID: [0000-0002-1481-2996](https://orcid.org/0000-0002-1481-2996)
- Frank-Oliver Glöckner (Alfred-Institut für Polar- und Meeresforschung & Universität Bremen – MARUM)
ORCID: [0000-0001-8528-9023](https://orcid.org/0000-0001-8528-9023)
- Dirk von Suchodoletz (University of Freiburg)
ORCID: [0000-0002-4382-5104](https://orcid.org/0000-0002-4382-5104)
- Barbara Ebert (German federation for biological data (GFBio))
ORCID: [0000-0003-3328-6693](https://orcid.org/0000-0003-3328-6693)
- Alice Carolyn McHardy (HZI)
ORCID: [0000-0003-2370-3430](https://orcid.org/0000-0003-2370-3430)
- Timo Mühlhaus (Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau)
ORCID: [0000-0003-3925-6778](https://orcid.org/0000-0003-3925-6778)
- Stefanie Weidtkamp-Peters (Heinrich Heine University Düsseldorf)
ORCID: [0000-0001-7734-3771](https://orcid.org/0000-0001-7734-3771)
- Contributor:
 - Barbara Götz (ZB MED) ORCID: [0000-0002-6382-7211](https://orcid.org/0000-0002-6382-7211)
 - Josh Moore (GerBI e.V.) ORCID: [0000-0003-4028-811X](https://orcid.org/0000-0003-4028-811X)
 - Matthias Lange (IPK Gatersleben) ORCID: [0000-0002-4316-078X](https://orcid.org/0000-0002-4316-078X)
 - Uwe Scholz (IPK Gatersleben) ORCID: [0000-0001-6113-3518](https://orcid.org/0000-0001-6113-3518)
 - Maja Magel (University Hospital of RWTH Aachen) ORCID: [0009-0004-2517-0791](https://orcid.org/0009-0004-2517-0791)

Resource type: Publication/Other (Kein MoU oder Agreement auswählbar)

Title: Memorandum of Understanding between NFDI4Biodiversity, NFDI4Microbiota and DataPlant

Licenses: Creative Commons Attribution 4.0 International

Description:

In a Memorandum of Understanding, DataPLANT, NFDI4Biodiversity and NFDI4Microbiota agree to work more closely together in order to better align their offerings for the management of biological data.

Keywords: NFDI, Memorandum of understanding, research data management

Language: englisch

Publication date: Date of publication

Communities: NFDI, DataPlant, NFDI4Biodiversity, NFDI4Microbiota,

Call 2024.07.24

Call 9. Juli 2024; 8:00 - 8:50

Teilnehmende:

- Frank-Oliver Glöckner
- Barbara Ebert
- Dirk von Suchodoletz
- Barbara Götz
- Konrad Förstner

Gemeinsame Finanzierung:

- Während Projektlaufzeit: ARUNA: jeweils 1 volle Stelle aus 4Microbiota und 4Biodiv, mit unterschiedlichem Arbeitsfokus:
 - N4B - Service-to-Service
 - N4M - Service-to-User
- --> Joint User group of N4M and BioDiv would be useful to align expectations and understand potential target groups better
- Im Folgeantrag: DSI: (insgesamt 0,5 Stelle bei DSMZ) halb/halb von BioDiv und N4M finanziert:
 - 2026: 0.5 N4Biodiv
 - 2027: 0.25/0.25
 - 2028: 0.25/0.25
 - 2029: 0.25/0.25
 - 2030: 0.25/0.25
 - 2031: 0.5 N4Microbiota

Zukünftige Services:

- Workflow Management System CloWM, bereits von N4M und teilweise BioDiv genutzt (Use Cases)
- Galaxy: Kompetenzen bei DataPLANT und de.NBI partnern, ebenfalls von Interesse für die beiden anderen Communities
- Knowledge Bases: Bündelung der Kräfte beim Schreiben von Artikeln zu allgemeinen FDM-Themen bzw. Best practices für einzelne Datentypen => ggf. in einigen Jahren Zusammenlegung prüfen, je nach Bedarf

Weitere Vernetzung:

- FAIRAgro und Bioimage ansprechen, MoU in 2025 um diese beiden Partner erweitern; weiterer Kandidat könnte 4Objects sein (Community cluster 10: Biosciences = Paläontologie)
- NFDI4Microbiota plant (unabhängiges) MoU mit den biomedizinischen Konsortien GHGA, 4Immuno => Größere MoU der Biomedizin angehen wenn Finanzierung/Zukunft von NFDI steht - 2029 - 2030;
 - NFDI4Health, GHGA, 4Immuno, 4Objects, => vielleicht dann da auch genug für eine Base4NFDI-Antrag

- Biomedizinische Interessensgruppe (Health, Immuno, etc.) vs. Biodata Interessensgruppe (BioDiv, Plant, FAIRAgro, etc.) => gemeinsame Klammer finden (tragen die gemeinsamen Datentypen?)

Zeitplanung MoU:

- Finalisierung im August nach Submission des Folgeantrags
- Im September VÖ auf Websites

Action items:

- Barbara:
 - proofread introduction
 - User group ARUNA --> Giessen aktivieren
 - Koordination Finalisierung MoU mit Judith Engel im August

Participants etc.

Note to ourselves (i.e. not a part of the Mou): joint partners

Relations between consortia at partner level

Institution	DataPLANT	Microbiota	Biodiversity	FAIRagro	Objects	BIOIMAGE
de.NBI e.V.		Participant	Participant			
DKFZ						Co-applicant
DSMZ		Co-applicant	Co-applicant			
EMBL		Co-applicant				Co-applicant
FZ Jülich	Co-applicant	Participant		Co-applicant		Participant
GerBI e.V.						Co-applicant
GFBio e.V.	Participant		Co-applicant			
HHU	Participant	Participant				Co-applicant
HKI		Participant				Co-applicant
HZI		Co-applicant				
IGB		Participant	Participant			
INP						Participant
IPHT						Participant
IPK	Co-applicant		Co-applicant	Co-applicant		
LIN						Co-applicant
Heidelberg U						Participant
MPIEB						Participant
RPTU	Co-applicant		Participant	Participant		
Ruhr-Uni Bochum	Participant		Participant			
RWTH Aachen		Co-Applicant	Participant			
Senckenberg				Co-applicant		
TU Dresden						Participant
U Bielefeld	Participant	Co-applicant	Co-applicant	Participant		
U Dortmund				Participant		
U Duisburg-Essen	Participant	Participant	Participant			
U Freiburg	Applicant					Co-applicant
U Gießen		Co-applicant	Co-applicant	Associated Partner		

U Göttingen	Participant	Participant	Participant	Participant		Participant
U Kiel		Participant				
U Köln	Participant	Participant				Participant
U Konstanz	Participant					Co-applicant
U Mainz						Participant
U Marburg		Co-applicant				
U Münster						Co-applicant
U Osnabrück						Co-applicant
U Tübingen	Participant	Participant	Participant			
U Jena		Co-applicant	Co-applicant			
U Leipzig						Co-applicant
UFZ		Co-applicant		Participant		Participant
ZB MED		Co-applicant		Co-applicant		

FEDA

Beispieltext FEdA

Umsetzung

Für die Unterstützung der FEdA-Projekte entlang des Datenlebenszyklus werden die etablierten Helpdesk-Prozesse des Konsortialpartners GFBio e.V. angeboten (www.gfbio.org). Das Personal der FEdA-Koordinierungsstelle erhält nach einer Einführung in geeigneter Weise Zugang zum Helpdesk und begleitet die FEdA-Projekte zusammen mit den fachlichen Datenkurator:innen aus dem NFDI4Biodiversity-Netzwerk (Front Office/Back Office-Modell).

Die FEdA-Koordinierungsstelle sichert über eine Datenpolicy die Akzeptanz der NFDI-Kooperation in den Projekten und die Übermittlung von Metadaten zu den projektbezogenen Datensätzen an das Portal, soweit diese in Datenzentren von NFDI4Biodiversity-Partnern hinterlegt werden. Die Entscheidung über die Veröffentlichung von Datensätzen bleibt im Einzelfall den Forschenden überlassen. Auch besteht kein Zwang, NFDI-Dienste zu nutzen.

Soweit Forschende die Dienste des GFBio-Helpdesk in Anspruch nehmen, gelten die auf der Website einsehbaren Nutzungsbedingungen. Bei sehr voluminösen oder aufwändigen Datenlieferungen ist eine Kostenbeteiligung an der Archivierung nicht ausgeschlossen. Diese Verhandlungen sind von den Projektbeteiligten zu führen und nicht Sache der FEdA-Koordinierungsstelle.

Für die Lieferung der Metadaten wird eine Entwicklungszusammenarbeit geplant. Die Terminierung und Durchführung von Informationsveranstaltungen und Schulungen erfolgt nach Absprache zwischen den Koordinierungsstellen von FEdA und NFDI4Biodiversity. Die FEdA-Koordinierungsstelle erhält bei der Erstellung von Informationsmaterialien zum Forschungsdatenmanagement Unterstützung aus dem Partnernetzwerk von NFDI4Biodiversity.

Die Koordination von NFDI4Biodiversity sichert zu, Fachleute aus anderen NFDI-Konsortien anzufragen, wo Methoden und Datentypen der FEdA-Projekte dies erfordern. Sollte eine Mitwirkung von Datenzentren außerhalb des Konsortiums dauerhaft benötigt werden, so wird die FEdA-Koordinierungsstelle mit diesen direkt Vereinbarungen treffen.