

# NFDI4Microbiota Progress Report

September 2024



Consortia Progress Report Part 1 (B-1)

## Contents

<b>1</b>	<b>General information</b>	<b>1</b>
<b>2</b>	<b>Summary</b>	<b>2</b>
<b>3</b>	<b>Composition of the consortium</b>	<b>4</b>

## 1 General information

### Name of the consortium:

NFDI4Microbiota – National Research Data Infrastructure for Microbiota Research

### Research domains or research methods addressed by the consortium:

Microbiology, as a multidisciplinary field, is linked to a broad spectrum of research areas, encompassing not only the biological sciences but also intersecting with several other disciplines. Due to this NFDI4Microbiota addresses methods and data types from the research fields of microbiology but also several connected research fields covered by the following DFG Research Areas: (1) Biology; (22) Medicine; (23) Agriculture, Forestry, and Veterinary Medicine; (34) Geosciences; (42) Thermal Engineering / Process Engineering; (43) Materials Science, and Engineering; (44) Computer Science, Systems and Electrical Engineering.

### URL of the consortium website and repositories used for publishing output:

- Website: <https://nfdi4microbiota.de/>
- Knowledge base: <https://knowledgebase.nfdi4microbiota.de>
- GitHub repository: <https://github.com/NFDI4Microbiota>
- Zenodo: <https://zenodo.org/communities/nfdi4microbiota>
- Videos: <https://av.tib.eu/publisher/NFDI4Microbiota>

## 2 Summary

**NFDI4Microbiota** aims to support scientists in microbiology to translate their research data into a deep understanding of microbial species by making data, software, analytical workflows, and machine learning models FAIR and open, by providing training, and by developing and disseminating standards. A huge wealth of data is collected in microbiological research, especially through the application of high-throughput technologies known as “omics” approaches. Data are typically stored in an unstructured format and/or with only little metadata, which makes them difficult to use and re-use. NFDI4Microbiota aims to change this by mobilizing, structuring, and linking available data to create a seamless flow of data throughout the research process. As part of the **NFDI**, we are committed to driving the necessary cultural change toward **FAIR principles** and open research practices. Our goal is to make the research process more efficient, promoting transparency, and collaboration to accelerate the development of solutions with societal impact.

The consortium consists of ten partner institutes and over 50 participants. Our tasks are divided into five task areas (TAs): Community, Networking & Training (TA1), Standards & Policies (TA2), Services (TA3), Technical infrastructure (TA4), and Coordination & Communication (TA5). We aim to achieve ten key objectives (KO1 - KO10), as described in our **proposal**. NFDI4Microbiota is supported by five professional microbiological societies (**VAAM**, **GfV**, **DGHM**, **DGP**, and **DGfM**). We attend their annual conferences and organize presentations, information booths, and symposia to promote the FAIR principles (KO1). We reinforce our ties to the international community by attending relevant conferences and workshops. To engage young scientists, we run an ambassador program featuring bimonthly meetings, dedicated training sessions, and first-hand information sharing. Our ambassadors provide valuable feedback and keep us in touch with the needs of their respective institutes. To connect with the broader community, we host events such as our bimonthly coffee talks, expand our reach through social media engagement, and distribute a quarterly newsletter.

Our **web portal** provides a central access point to our ever-growing number of services and hosts our **Knowledge Base**. It also includes a helpdesk that allows the community to ask questions, for example on data management and services, with relevant experts typically responding in about a week. A joint life-science helpdesk is now being developed together with other consortia; this will allow us to share our expertise with other NFDI consortia (KO3) while promoting the “One NFDI” mission to create a cross-disciplinary research data solution for the German research community. Databases such as **StrainInfo** and **VirJenDB** have addressed previously unmet needs by aggregating and linking critical data (KO4). We curate a comprehensive **repository** that compiles metadata standards and recommendations and includes technical specifications and biological

and environmental data guidelines (KO6). Furthermore, we provide the computational infrastructure for data analyses and FAIR data management in the form of the [Aruna](#) data orchestration engine and the Cloud-based Workflow Manager ([CloWM](#)), which offer fast, domain-agnostic data management and scalable workflow execution in a user-friendly web-based environment built on open standards. To standardize data processing and analysis, we provide best-practice computational workflows and the CAMI benchmarking platform as benchmarking environment, founded on the results of the initiative for the Critical Assessment of Metagenome Interpretation (CAMI) (KO7). NFDI4Microbiota aims to equip the community with skills for efficient, data-driven microbial research and to increase overall data literacy (KO2). Our 100+ training courses have been attended by over 1,600 people and cover topics such as research data management and infrastructure, as well as domain-specific topics such as foundational programming and data skills, metadata standards, and FAIR principles, workflow engines, genomics, and metagenomics. Courses are in high demand, with registrations regularly exceeding available places and high satisfaction rates. Our data stewards conduct workshops at participating institutions and provide assistance with data management issues, promoting high-quality research data management (KO5).

To adapt to future needs (KO10) and to connect to the community, NFDI4Microbiota has established Use Cases relating to data generation, analysis, and FAIR deposition together with consortium partners and other community members, which link the developed service portfolio to community applications. To close gaps in the portfolio of provided solutions, Strategy Fund and Flex Fund projects were introduced, in which grants support relevant work by the co-applicant institutes and other community members, respectively.

NFDI4Microbiota is active across all sections and most of their working groups, task forces, and the communication and management circle. Members have contributed to several joint publications. Currently, the spokesperson of NFDI4Microbiota is a member of the NFDI Scientific Senate and is involved in developing NFDI's strategy for the years ahead. Following the "One NFDI" spirit, NFDI4Microbiota collaborates closely with other life science consortia, such as [NFDI4Biodiversity](#), [NFDI4Health](#), [DataPlant](#), [NFDI4BIOIMAGE](#), [NFDI4Immuno](#), and [GHGA](#), organizing joint events such as the cross-consortia hackathon. Furthermore, NFDI4Microbiota is involved in several [Base4NFDI](#) projects ([KG14NFDI](#), [DMP4NFDI](#), [PID4NFDI](#), [MC4NFDI](#), [RDMTraining4NFDI](#)), and members of the consortium are the leading institutions in some of these projects.

While the consortium has successfully achieved all the planned milestones up to this point (excluding those that were necessarily canceled due to financial cuts), several tasks remain, including the integration of further stakeholders and the promotion of services to the community. NFDI4Microbiota will address these challenges and continue providing benefits not only to the microbiology community, but also to the broader academic environment.

### 3 Composition of the consortium

Applicant institution	Location	Duration	ROR
ZB MED	Cologne	10/2021 -	<a href="#">0259fwx54</a>

Spokesperson	Institution, Location	Duration	ORCID
Konrad Föstner	ZB MED, Cologne	10/2021 -	<a href="#">0000-0002-1481-2996</a>

Co-Applicant institution	Location	Duration	ROR
Helmholtz Centre for Infection Research (HZI)	Braunschweig	10/2021 -	<a href="#">03d0p2685</a>
European Molecular Biology Laboratory (EMBL)	Heidelberg	10/2021 -	<a href="#">01yr73893</a>
Bielefeld University (BIBI)	Bielefeld	10/2021 -	<a href="#">02hpadn98</a>
Friedrich Schiller University Jena (FSU)	Jena	10/2021 -	<a href="#">05qpz1x62</a>
Helmholtz Centre for Environmental Research (UFZ)	Leipzig	10/2021 -	<a href="#">000h6jb29</a>
Justus Liebig University Gießen (JLU)	Gießen	10/2021 -	<a href="#">033eqas34</a>
Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures (DSMZ)	Braunschweig	10/2021 -	<a href="#">02tyer376</a>
Philipps University of Marburg (UMR)	Marburg	10/2021 -	<a href="#">01rdrb571</a>
RWTH Aachen University (RWTH)	Aachen	10/2021 -	<a href="#">04xfq0f34</a>

Co-Spokesperson	Institution, Location	Task areas	Duration	ORCID
Alice McHardy	HZI, Braunschweig	TA1 - TA3, TA5	10/2021 -	<a href="#">0000-0003-2370-3430</a>
Alexander Sczyrba	BIBI, Bielefeld	TA1 - TA4	10/2021 -	<a href="#">0000-0002-4405-3847</a>
Jens Stoye	BIBI, Bielefeld	TA1 - TA4	10/2021 -	<a href="#">0000-0002-4656-7155</a>
Jörg Overmann	DSMZ, Braunschweig	TA1 - TA3	10/2021 -	<a href="#">0000-0003-3909-7201</a>
Manja Marz	FSU, Jena	TA1 - TA3	10/2021 -	<a href="#">0000-0003-4783-8823</a>
Peer Bork	EMBL, Heidelberg	TA1 - TA4	10/2021 -	<a href="#">0000-0002-2627-833X</a>
Ulisses Nunes da Rocha	UFZ, Leipzig	TA1 - TA4	10/2021 -	<a href="#">0000-0001-6972-6692</a>
Alexander Goessmann	JLU, Giessen	TA2 - TA3	10/2021 -	<a href="#">0000-0002-7086-2568</a>

Co-Spokesperson	Institution, Location	Task areas	Duration	ORCID
Thomas Clavel	RWTH, Aachen	TA1, TA3	10/2021 -	<a href="https://orcid.org/0000-0002-7229-5595">0000-0002-7229-5595</a>
Anke Becker	UMR, Marburg	TA1 - TA4	10/2021 -	<a href="https://orcid.org/0000-0003-4561-9184">0000-0003-4561-9184</a>

Participating Institutions	Location	Duration	ROR
Eberhard Karls University Tübingen	Tübingen	08/2024 -	<a href="https://ror.org/03a1kwz48">03a1kwz48</a>
Georg-August-Universität Göttingen	Göttingen	10/2021 -	<a href="https://ror.org/01y9bpm73">01y9bpm73</a>
Heinrich-Heine-University Düsseldorf	Düsseldorf	10/2021 -	<a href="https://ror.org/024z2rq82">024z2rq82</a>
Helmholtz Munich (German Research Center for Environmental Health)	Munich	10/2021 -	<a href="https://ror.org/00cfam450">00cfam450</a>
Julius Maximilian University Würzburg	Würzburg	10/2021 -	<a href="https://ror.org/00fbnyb24">00fbnyb24</a>
Karlsruher Institut for Technology	Karlsruhe	10/2021 -	<a href="https://ror.org/04t3en479">04t3en479</a>
Leibniz Institute for Analytical Sciences - ISAS	Dortmund	03/2023 -	<a href="https://ror.org/02jhqqg57">02jhqqg57</a>
Leibniz Institute of Freshwater Ecology and Inland Fisheries	Berlin	10/2021 -	<a href="https://ror.org/01nftxb06">01nftxb06</a>
Max Delbrück Center (MDC)	Berlin	10/2021 -	<a href="https://ror.org/04p5ggc03">04p5ggc03</a>
Max Planck Institute for Dynamics of Complex Technical Systems	Magdeburg	10/2021 -	<a href="https://ror.org/030h7k016">030h7k016</a>
Max Planck Institute for Marine Microbiology	Bremen	10/2021 -	<a href="https://ror.org/02385fa51">02385fa51</a>
Robert Koch-Institute (RKI)	Berlin	10/2021 -	<a href="https://ror.org/01k5qnb77">01k5qnb77</a>
Ruhr-University Bochum	Bochum	10/2021 -	<a href="https://ror.org/04tsk2644">04tsk2644</a>
German National Library of Science and Technology (TIB)	Hannover	10/2021 -	<a href="https://ror.org/04aj4c181">04aj4c181</a>
Technical University of Darmstadt	Darmstadt	10/2021 -	<a href="https://ror.org/05n911h24">05n911h24</a>
Technical University of Munich	Munich	10/2021 -	<a href="https://ror.org/02kkvpp62">02kkvpp62</a>
University of Greifswald	Greifswald	10/2021 -	<a href="https://ror.org/00r1edq15">00r1edq15</a>
University of Hohenheim	Hohenheim	03/2023 -	<a href="https://ror.org/00b1c9541">00b1c9541</a>
Kiel University	Kiel	10/2021 -	<a href="https://ror.org/04v76ef78">04v76ef78</a>
de.NBI e.V	Heidelberg	12/2022 -	
German Aerospace Center (DLR)	Cologne	10/2021 -	<a href="https://ror.org/04bwf3e34">04bwf3e34</a>
Jülich Research Centre	Jülich	10/2021 -	<a href="https://ror.org/02nv7yv05">02nv7yv05</a>

Participating Institutions	Location	Duration	ROR
Fraunhofer Society	Munich	10/2021 -	<a href="#">05hkkdn48</a>
Fraunhofer Institute for Molecular Biology and Applied Ecology	Schmallenberg	10/2021 -	<a href="#">03j85fc72</a>
University of Erlangen-Nuremberg (FAU)	Erlangen	10/2021 -	<a href="#">00f7hpc57</a>
GEOMAR Helmholtz Centre for Ocean Research Kiel	Kiel	10/2021 -	<a href="#">02h2x0161</a>
Goethe University Frankfurt	Frankfurt	10/2021 -	<a href="#">04cvxnb49</a>
Leibniz Institute for Natural Product Research and Infection Biology (HKI)	Jena	10/2021 -	<a href="#">055s37c97</a>
Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences	Potsdam	10/2021 -	<a href="#">04z8jg394</a>
Anhalt University of Applied Sciences	Köthen	03/2023 -	<a href="#">0076zct58</a>
Leibniz Institute on Aging - Fritz Lipmann Institute (FLI)	Jena	10/2021 -	<a href="#">039a53269</a>
Ludwig Maximilian University of Munich	Munich	10/2021 -	<a href="#">05591te55</a>
Martin Luther University Halle-Wittenberg	Halle	10/2021 -	<a href="#">05gqaka33</a>
Max Planck Society (MPG)	Munich	10/2021 -	<a href="#">01hhn8329</a>
Max Planck Institute of Biophysics	Frankfurt am Main	10/2021 -	<a href="#">02panr271</a>
Max Planck Institute for Plant Breeding Research	Cologne	10/2021 -	<a href="#">044g3zk14</a>
Otto-von-Guericke University Magdeburg	Magdeburg	10/2021 -	<a href="#">00ggpsq73</a>
University of Bonn	Bonn	10/2021 -	<a href="#">041nas322</a>
TH Köln - University of Applied Sciences	Cologne	10/2021 -	<a href="#">014nnvj65</a>
University of Duisburg-Essen	Essen	10/2021 -	<a href="#">04mz5ra38</a>
Hamburg University	Hamburg	10/2021 -	<a href="#">00g30e956</a>
Ulm University	Ulm	10/2021 -	<a href="#">032000t02</a>
University of Cologne	Cologne	10/2021 -	<a href="#">00rcxh774</a>