

Instructions and Template

for Consortia Progress Reports

National Research Data Infrastructure (NFDI)

**Achtung: Textbeiträge und referenzierbare Outputs
müssen bis 15.8.2023 fertig sein**

Meldestelle für Outputs:

<https://kb.gfbio.org/display/NFDI/Project+Outputs>

A General Instructions

Progress reports of NFDI consortia must be submitted to the DFG Head Office in PDF format via elan, the DFG's submission portal (elan.dfg.de). They consist of two PDF documents: The first part, section B-1, will be published on the DFG homepage and the NFDI community page on Zenodo¹; the second part, section B-2, will be made available only to the NFDI Expert Committee, the DFG Head Office and the NFDI Directorate. Note that PDF documents must not be larger than 10 MB or include password protection or access restrictions for reading, copying or printing. Please create a clickable table of contents for any PDF document you submit.

The following template provides a mandatory structure and a number of tables to be completed. Please adhere to this format and follow the instructions printed in grey (please delete the grey

¹ <https://zenodo.org/communities/nfdi>

text before submission). Please note that section B-2 (all subheadings excluding the appendices) should not exceed 25 pages.

Progress reports must be written in English, font Arial, 11 pt, 1.5 line spacing

The DFG privacy policy applies, see www.dfg.de/privacy_policy.

B-1 Progress Report Template Part 1, [for publication](#)

1 General Information

- Name of the consortium

Up to 300 characters, including spaces, no special characters.

[NFDI4Biodiversity: Biodiversity, Ecology and Environmental Data](#)

- Research domains² or research methods addressed by the consortium

NFDI consortia define their scope based on the research domains or research methods they address. Please summarise research domains or research methods your consortium addresses in no more than 1,000 characters, including spaces.

[NFDI4Biodiversity](#) covers methods and data types from organismal biological and ecosystem research (DFG review boards 202 Plant Sciences, 203 Zoology; plus subject areas from: 201 Basic Research in Biology and Medicine, 204 Microbiology, Virology and Immunology, 207 Agriculture, Forestry and Veterinary Medicine, 304 Geology and Palaeontology and others). Data typically include species observations recorded in the form of tables, photos, videos, specimens and audio files, often with additional information (e.g. molecular data, experimental measurements, ecological information) and others. Specimens are also frequently collected and preserved for further analysis and description (collection management). Resources also cover molecular data collected by -omics methodology (genomics, metabolomics, glycomics, and transcriptomics), data from Earth System Sciences (f.e. environmental conditions, information in land use or colonisation) and data streams from continuous monitoring of environmental parameters by sensors/satellites. In addition to laboratory analyses, in vivo/in vitro experiments and computational methods are used (bioinformatics, environmental modelling and machine learning). Interdisciplinary studies on nature conservation and biodiversity monitoring efforts (socio-ecological research) also use social sciences methodology.

- URL of the consortium website and repositories used for publishing output

² Please provide the respective names and numbers of the DFG review boards when listing the research domains.

www.nfdi4biodiversity.org
<https://zenodo.org/communities/nfdi4biodiv>
<https://www.youtube.com/@NFDI4Biodiv>

2 Summary

The summary describes in no more than two pages the main results and notable success stories. While the summary is a retrospective one, you may also briefly outline challenges still to be addressed. Since this summary is for publication on widely used platforms, it should be informative for a broad academic public, not only the communities served by the consortium.

3 Composition of the consortium

Please list all (co-)applicant institutions, (co-)spokesperson(s) and participants, each with the duration of their involvement (mm/yy - mm/yy). Please indicate ORCID-IDs, if available.

▪ Applicant institution

Applicant institution	Location	Duration
MARUM – Center for Marine Environmental Sciences, University Bremen	Bremen, Germany	10/20 - today

▪ Spokesperson

Spokesperson	Institution, location	Duration
Prof. Dr. Frank Oliver Glöckner [0000-0001-8528-9023]	MARUM, University Bremen	10/20 - today

▪ Co-applicant institutions

Co-applicant institutions	Location	Duration
Freie Universität Berlin, Botanic Garden and Botanical Museum Berlin (BGBM)	Berlin	10/20 - today
Universität Bielefeld, Institut für Bioinformatik-Infrastruktur (BIBI)	Bielefeld	10/20 - today
Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)	Braunschweig	10/20 - today
Research Institute for Farm Animal Biology (FBN)	Dummerstorf	10/20 - today

German Federation For Biological Data (GFBio e.V.)	Bremen	10/20 - today
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)	Göttingen	10/20 - today
Heidelberg Institute for Theoretical Studies gGmbH (HITS)	Heidelberg	10/20 - today
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)	Gatersleben	10/20 - today
Bavarian Natural History Collections (SNSB)	München	10/20 - today

▪ Co-spokespersons

Co-spokespersons	Institution, location	Task area(s)	Duration
Anton Güntsch [0000-0002-4325-4030]	BGBM, Berlin	TA2	10/20 - today
Prof. Dr. Jens Stoye [0000-0002-4656-7155]	BIBI, Bielefeld	TA2	10/20 - today
Prof. Dr. Jörg Overmann [0000-0003-3909-7201]	DSMZ, Braunschweig	TA5	10/20 - today
Prof. Dr. Klaus Wimmers [0000-0002-9523-6790]	FBN, Dummerstorf	TA1	10/20 - today
Dr. Barbara Ebert [0000-0003-3328-6693]	GFBio e.V., Bremen	TA5	01/21 - today
Dr. Michael Diepenbroek [0000-0003-3096-6829]	GFBio e.V., Bremen	TA4	10/20 - today
Dr. Ivaylo Kostadinov [0000-0003-4476-6764]	GFBio e.V., Bremen	TA5	10/20 - today
Prof. Dr. Ramin Yahyapour [0000-0002-9057-4395]	GWDG, Göttingen	TA2	10/20 - today
Dr. Wolfgang Müller [0000-0002-4980-3512]	HITS, Heidelberg	TA4	10/20 - today
Dr. Uwe Scholz [0000-0001-6113-3518]	IPK, Gatersleben	TA4	10/20 - today
Dr. Dagmar Triebel [0000-0003-1980-3148]	SNSB, München	TA3	10/20 - today
Dr. Mark Frenzel [0000-0003-1068-2394]	UFZ, Leipzig	TA1	10/20 - today
Prof. Dr. Alexander Goesmann [0000-0002-7086-2568]	JLU, Gießen	TA4	10/20 - today

Prof. Dr. Birgitta König-Ries [0000-0002-2382-97229]	iDiv, Jena	TA3	10/20 - today
Prof. Dr. Aletta Bonn [0000-0002-8345-4600]	iDiv, Jena	TA1	10/20 - today
Prof. Dr. Bernhard Seeger [0000-0002-9362-153X]	Philipps University Marburg, Marburg	TA4	10/20 - today
Prof. Dr. Birgit Gemeinholzer [0000-0002-9145-9284]	University of Kassel, Kassel	TA1	10/20 - today

▪ **Participants**

Please list the participants of the consortium in one of the two tables below, differentiating between participation as an institution or as an individual.

Participating institutions	Location	Duration
Institution 1	City 1	
Institution 2	City 2	
...		

Add lines as required. Delete this table if you have no participating institutions.

Participating individuals	Institution, location	Duration
Dr. Hubert Höfer [0000-0003-3962-151X]	State Museum of Natural History Karlsruhe (SMNK), Arachnological Society (AraGes), Putbus	10/20 - today
Prof. Dr. Stephan Frickenhaus [0000-0002-0356-9791]	Alfred Wegener Institute – Helmholtz Centre for Polar and Marine Research (AWI), Bremerhaven	10/20 - today
Dr. Janine Felden [0000-0002-4177-6644]	Alfred Wegener Institute – Helmholtz Centre for Polar and Marine Research (AWI), Bremerhaven	10/20 - today
Dr. Jonas Zimmermann [0000-0002-0522-0569]	Freie Universität Berlin, Botanic Garden and Botanical Museum Berlin, Research Group Diatoms (BGBM-DI), Berlin	10/20 - today
Dr. Gerwin Kasperek [0000-0001-6528-9720]	Specialised Information Service Biodiversity Research (BIOfid), University Library Johann Christian Senckenberg,	10/20 - today

	Goethe University Frankfurt, Frankfurt am Main	
Gregor von Halem	Federal Maritime and Hydrographic Agency (BSH), Hamburg	10/20 - today
Silvia Bender	German Federation for the Environment and Nature Conservation (BUND), Berlin	10/20 - 05/20
Magnus Wessel	German Federation for the Environment and Nature Conservation (BUND), Berlin	05/20 - today
Sina-Victoria Barysch	Deutsches Netzwerk für Bioinformatik-Infrastruktur e.V. (de.NBI), Bremen	07/22 - today
Prof. Dr. Florian Leese [0000-0002-5465-913X]	University of Duisburg-Essen, Aquatic Ecosystem Research (DNAquaNet), Essen	10/20 - today
Johannes Schwarz	Federation of German Avifaunists (DDA), Münster	10/20 - today
Dr. Christoph Sudfeldt	Federation of German Avifaunists (DDA), Münster	10/20 - today
Klaus-Jürgen Conze	Association of German-speaking odonatologists (GdO), Essen	10/20 - today
Prof. Dr. Heiko Brunken [0000-0002-3745-5629]	German Ichthyological Society (GfI), Düsseldorf	10/20 - today
Prof. Dr. Christian Ammer	Ecological Society of Germany, Austria and Switzerland (GfÖ), Berlin	10/20 - today
Prof. Dr. Helmut Hillebrand [0000-0001-7449-1613]	Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Oldenburg	10/20 - today
Prof. Dr. Thomas Brey [0000-0002-6345-2851]	Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Oldenburg	10/20 - today
Stephan Karger	Hessian Agency for Nature Conservation, Environment and Geology (HLNUG), Wiesbaden	10/20 - today
Dr. Jana Moser	Leibniz Institute for Regional Geography (IfL), Leipzig	10/20 - today
Prof. Dr. Hans-Peter Grossart [0000-0002-9141-0325]	Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin	10/20 - today
Prof. Dr. Adrian Paschke	Institute for Applied Informatics association (InfAI), Leipzig	10/20 - today
Dr. Gotthard Meinel [0000-0002-9201-7664]	Leibniz Institute of Ecological Urban and Regional Development (IÖR), Dresden	10/20 - today
Dr. Holger Beer	Julius Kühn Institute (JKI) Federal Research Centre for Cultivated Plants, Quedlinburg	10/20 - today
Dr. Kai Gedeon	State Office for Environmental Protection Saxony-Anhalt (LAU), Halle (Saale)	10/20 - today

Holger Lueg	Saxon State Office for Environment, Agriculture and Geology (LfULG), Dresden	10/20 - today
Dr. Christian Bölling	Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science (MfN), Berlin	10/20 - today
Dr. Jana Hoffmann [0000-0001-6214-661X]	Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science (MfN), Berlin	10/20 - today
Stefan Munzinger	naturgucker.de non-profit registered cooperative, Northeim	10/20 - today
Prof. Dr. Florian Jansen [0000-0002-0331-5185]	Network Phytodiversity Germany (NetPhyD), Rostock	10/20 - today
Prof. Dr. Jörg Müller	Bavarian Forest National Park, Grafenau	10/20 - today
Jan Rommelfanger	Hunsrück-Hochwald National Park Office, Birkenfeld	10/20 - today
Dr. Margit Ksoll-Marcon	General Directorate of the Archives of the Bavarian State (GDA), München	10/20 - 08/22
Dr. Markus Schmalzl	General Directorate of the Archives of the Bavarian State (GDA), München	09/22 - today
Prof. Dr. Thomas Hickler [0000-0002-4668-7552]	Senckenberg – Leibniz Institution for Biodiversity and Earth System Research (SGN), Frankfurt am Main	10/20 - today
Dr. Joachim Holstein [0000-0002-1541-7170]	Stuttgart State Museum of Natural History (SMNS), Stuttgart	10/20 - today
Prof. Dr. Wolfram Horstmann [0000-0001-8673-6104]	Göttingen State and University Library (SUB), Göttingen	10/20 - today
Florian Hoedt [0000-0002-6068-1659]	Johann Heinrich von Thünen Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Braunschweig	10/20 - today
Prof. Dr. Thomas Kneib [0000-0003-3390-0972]	Georg-August-University Göttingen, Campus Institute Data Science (CIDAS), Göttingen	10/20 - today
Prof. Dr. Christian Wirth [0000-0003-2604-8056]	Leipzig University, Leipzig	10/20 - today
Prof. Dr. Miguel D. Mahecha [0000-0003-3031-613X]	Leipzig University, Leipzig	10/20 - today
Dr. Peter Grobe [0000-0003-4991-5781]	Zoological Research Museum Alexander Koenig (ZFMK), Bonn	10/20 - today
Dr. Nicolas Dittert [0000-0002-1143-3405]	Leibniz Centre for Tropical Research (ZMT), Bremen	10/20 - today

Add lines as required. Delete this table if you have no participating individuals.

B-2 Progress Report Template Part 2, for internal use only

Remember the page limit of 25 pages (excluding the appendices).

Please indicate at suitable paragraphs, how you reacted to recommendations of the NFDI Expert Committee you may have received (final section of the document *Hinweise aus Begutachtung und Beratung* that you received at the same time as the funding letter).

1 Consortium

The following subheadings correspond closely to the funding criteria on “Readiness and Relevance of the Consortium” and “Consortium Structure and Viability”. Consult the Guidance Notes on Funding Criteria (DFG form NFDI 120) for more information on the specific subheadings and the NFDI Guidelines (DFG form NFDI 100) for general information on the consortia.

If one or more members of your consortium are participating in other NFDI consortia, please list these multiple participations here, e.g.:

Member 1 is also participating in consortium X and Y

Member 2 is also participating in consortium Z

...

Tentative list

(authoritative list is being collected here: [Multiple participations of consortium members](#))

Prof. Dr. Frank Oliver Glöckner, Uni Bremen is also participating in NFDI4Earth, NFDI4DataScience

Prof. Dr. Jens Stoye, Uni Bielefeld/BIBI is also participating in NFDI4Microbiota

Prof. Dr. Alexander Goesmann, UniGI is also participating in NFDI4Microbiota

Prof. Dr. Jörg Overmann, DSMZ is also participating in NFDI4Microbiota

Prof. Dr. Ramin Yahyapour, GWDG is also participating in NFDIxCS

Dr. Wolfgang Müller, HITS is also participating in NFDI4Health

Prof. Dr. Stephan Frickenhaus, AWI is also participating in NFDI4Earth

Dr. Gotthard Meinel, IÖR is also participating in NFDI4Earth

Dr. Jana Hoffmann, MfN Berlin is also participating in NFDI4Earth, NFDI4Objects

<i>Prof. Dr. Miguel Mahecha, U Leipzig, is also participating in</i>	<i>NFDI4Earth</i>
<i>Markus Schmalzl, SAB is also participating in</i>	<i>NFDI4Earth, NFDI4Objects, FAIRAgro, NFDI4Memory</i>
<i>Wolfram Horstmann, SUB, is also participating in</i>	<i>NFDI4Culture,</i>
<i>NFDI4Memory</i>	
<i>Florian Hoedt, Thünen Institute is also participating in</i>	<i>FAIRagro</i>
....	

1.1 Composition of the consortium and its embedding in the community of interest

▪ Composition of the consortium

Please describe and explain changes to the consortium with respect to the partners (e.g. new partners, partners that have left the consortium) both for (co-)applicants and participating institutions and persons. How have these developments affected the work of the consortium?

The composition of the NFDI4Biodiversity consortium proved to be quite stable. The move of two co-applicants to other institutions during the starting phase in 2020 had little effect on the actual work. As a result, a) Kassel University joined as a co-applicant institution in 2021 and b) the participation of the Max-Planck-Institute for Biogeochemistry (MPI-BCG) was not pursued further. MPI-BCG renewed its interest in participation through another PI and joined in 2023. In 2022, the newly founded de.NBI Association (e.V.) was accepted as a new participant to the Consortium. This serves to strengthen relations with the de.NBI Service Centres and the de.NBI Cloud infrastructure, both are important to provide bioinformatics expertise and computing infrastructure for the Biodiversity community.

- Integration of communities of interest,³ relevance for the research system: By which procedures have communities of interest been given an active role in the consortium? How have the needs of the communities been identified and how has the consortium reacted to changing needs? Please describe qualitatively, and whenever possible also by means of quantitative indicators, which benefits the consortium has generated for the communities that it addresses as well as for other disciplines and the research system at large.

³ Please refer to the "White Paper: Interim Report Reference" to describe the different ways in which the consortium interacts with communities. Cf. <https://doi.org/10.5281/zenodo.7688729>

The community of interest for NFDI4Biodiversity is a data- and method-oriented stakeholder group in biodiversity research and conservation (B1, p. 12ff.). The community is quite diverse in terms of training, institution type and sector, from academic research groups to natural history museums, public authorities and citizen science initiatives. Members have a joint motivation to make meaningful contributions to nature preservation and to slow-down mass extinction. A lack of proper data management and sharing is a central issue of concern, although there are already several standardisation initiatives and research infrastructures available for sub-communities in bioinformatics, taxonomic and ecological research, as well as biodiversity monitoring. The transdisciplinary nature of the community requires to integrate not only diverse biodiversity information, but also different technical designs and stakeholder networks [B2]. NFDI4Biodiversity aims to achieve this through a large (n=36) participant group that represents the different sub-communities as well as important national and international infrastructure and standardisation initiatives [O1].

In this way, we are able to match individual data-related needs with existing solutions in the Community. There is a rather large proportion of use case projects (26) where we support partners practically with mobilisation of data and/or roll-out of existing, tried-and-tested software solutions. A team of use-case managers facilitates this process. Through the central helpdesk [O2], this support is also available for the general Community ("Community as recipient"). Information is delivered to the Community mainly via events, workshops and presentations.

*[insert quantitative information on Community requests for
collaboration - helpdesk, contact inbox and LOI]*

The 26 use case projects and their needs set the priorities for the Service Portfolio Management (Task Area 3) and the Research Data Commons development (Task Area 4). Initial interviews with use case partners conducted in Task Area 1 were analysed for mentioning of tools and services by staff in Task Area 3 [REF??]. Use case partners were approached as test users for the new cloud-based Core Storage and data integration workflows. In 2022 and 2023 Flexfunds calls were issued which enabled data and service providers to enhance their products in targeted and very productive projects. Resource provider workshops on legal issues, certification, data and software licensing were important steps to establish NFDI4Biodiversity as an expert forum for safe and competent handling of data for broad and responsible use ("Community as participatory network").

*[insert statistics on communication activities and social media
campaigning]*

All 50 partners have signed the “Mittelweiterleitungs- und Kooperationsvertrag”, which provides us with a coherent basis for collaboration. The mission statement ratified in 2022 states a value proposition and further principles for the joint work [O3].

At the level of governance, a Strategic Advisory Board (SAB) ensures interconnectivity with the community of interest. The SAB composition has 10 members from science, policy and society, taking into account the larger mission of NFDI4Biodiversity as part of the National Research Data Infrastructure NFDI in Germany. Five positions are held by representatives nominated by the Federal Ministry for Education and Research, the Federal Ministry for environmental affairs, the academic and scientific societies and a related NFDI consortium. All other positions are filled ad personam (“Community as governance”).

Notes from retreat

- o *Point of reference: chapter 2.2 in the application (pages!)*
- o *composition specific to biodiversity community - associations, authorities etc.*
 - *complex due to different “work style” and timelines*
- o *common methods and data types needed for biodiversity monitoring and research “big picture”*
- o *How community needs have been identified (and how we have reacted)*
 - *TA1: structured survey use cases and needs/ workplan*
 - *Helpdesk request => Jimena, analysis*
 - *Direct contact*
 - *Inbox contact form (6-10 online meetings, with minutes) => Barbara*
 - *LOI for project proposals => statistics*
 - *Community presentations at events => statistics*
 - *Community workshops => statistics of events auswerten*
 - *Exhibition booths => statistics*
 - *Social media campaign => Follower growth; other interactions (tweets with most impressions?)*
- o *How we give active role*
 - *General approach: mainly “community as recipient”*
 - *Participatory elements: Flexfunds call in consortium*
- o *Community engagement*
 - *Helpdesk as central platform => reference eScience-Tage*
 - *Software trainings*
- o *Community activation*
 - *Rollout of ABCD consensus element to more data providers (LAND use case);*
 - *teaching RDM skills*
 - *LOI for project proposals => app. 17, people approach the Speaker*
 - *Help with implementation of tools (TA1M2, HR Pool are available to use case partners => blueprints, best practices (“incentivise and enable the community”))*
 - *MoU with FEdA research initiative for data management*
 - *prototyping RDMO for Biodiversity Exploratories*
- o *Further interconnectivity*
 - *NMZB vice versa / avoid duplicate work*
 - *Strategic advisory board: community of interest represented as well as other consortia, NMZB and GfÖ (each reserved 1 position)*
 - *DSI expert network and working group in NFDI*
 - *RDMO network*

1.2 The consortium within the NFDI / General Contribution to the development of the NFDI

How has the consortium contributed to the building of the NFDI as a whole? How have collaborations with related consortia developed? Which cross-cutting topics have been specifically addressed by this consortium? What has been achieved with respect to these cross-cutting topics and with respect to cross-cutting topics in general? How does the consortium interact with Base4NFDI and which basic services⁴ is the consortium participating in? Do you see need for improvement regarding both cross-cutting topics and basic services? How have you participated in and contributed to the governance structure of the NFDI?

In NFDI4Biodiversity we are strongly committed to collaborate across consortia and contribute to all major activity lines. Efforts to consolidate and unite activities on cross-cutting topics, which were started by NFDI4Biodiversity members in 2019/20 with the Berlin and the Berlin-Leipzig declarations [B3, B4], were continued in 2021 in NFDI Strategy Workshops [O4]. NFDI4Biodiversity project staff was instrumental in the conceptualisation of the Section Common Infrastructures [O5]. Co-applicant Michael Diepenbroek serves as first elected speaker of the Section (2021-2023). This is in line with the NFDI4Biodiversity vision of a Research Data Commons (RDC) as part of a larger NFDI architecture. Consequently, senior staff members were drivers in uniting the NFDI Consortia Assembly behind the idea of a joint Basic Service Initiative, from the initial position paper [B5] to the application phase. In 2023, we proposed aData mesh concept [B6] as a framework for a NFDI-wide architecture.

*[insert information on partners' participation in basic services
proposals for the calls in 2023]*

We have project delegates in each NFDI section as well as several working groups and Task Forces, in order to ensure alignment with our work and flow of information (for outputs see O6 to OZ). The delegates report in Task Area 2 and in the quarterly project forum events. This system worked well in the initial phase of the sections, but with over 45 Working Groups and Task Forces now [**REF joint NFDI report], an organised participation gets more and more difficult to handle.

On the governance level, the Executive Secretary Barbara Ebert represents NFDI4Biodiversity in the Consortia Assembly and was elected to the NFDI Scientific Senate for its first term

⁴ "Basic services" means *Basisdienste* as defined and discussed by the expert committee in their statements in 2020 (https://www.dfg.de/download/pdf/foerderung/programme/nfdi/stellungnahme_nfdi_201112_de.pdf) and 2021 (https://www.dfg.de/download/pdf/foerderung/programme/nfdi/stellungnahme_nfdi_basisdienste.pdf).

(2021/2022). In 2022/23, she was part of the writing team for the Vision and Mission of NFDI [REF??].

On the level of consortia and their work programmes, common interests with other consortia have been pursued, both single domain and cross-domain. Within the life sciences domain and the environmental sciences, NFDI4Biodiversity shares sub-communities with DataPLANT and NFDI4Microbiota (Omics methods and data, de.NBI service infrastructure of joint interest) as well as FAIRagro and NFDI4Earth (environmental factors, satellite imaging, sensor data, cross-community data services of joint interest). These common interests have resulted in a number of joint workshops and community events. In terms of frequency and duration, two collaborations can be highlighted: 1) We have formed an informal interest group on policy development regarding the issue of Digital Sequence Information in the Global Biodiversity Framework (involving four consortia). The group organised two awareness-raising InfraTalks in the NFDI and contributes to lobbying activities of the Alliance of University and Non-University Biodiversity Research in Germany [O7]. 2) NFDI4Biodiversity and KonsortSWD joined forces to provide interdisciplinary data management support to a cohort of projects within the German Research Initiative on Biodiversity Preservation (FEEdA, O8). This collaboration has been ongoing since 2021 and involves joint helpdesk support and workshops for the researchers.

Based on this experience we feel that joint data management and archiving support would be a cross-community activity that would be of immediate benefit for the research community, but requires some initial organisational effort and mutual learning [O9]. These and occasional other collaborations have been collected in the xxx [joint publication]

Data Train erwähnen als Beispiel für Cross-domain mit NFDI4Health und KonsortSWD (NFDI4Ing, NFDI-MatWerk) <https://www.bremen-research.de/aktivitaeten/forschungsdaten/nfdi>

[check if this matches the picture of the list under B.2 Chapter 1]

Notizen Schreibklausur

- o *Point of reference: Chapter 2.3 of the proposal (p. 34ff.)*
 - *Entwicklung anders als gedacht bzgl. fast aller dort gelisteten Konsortien, mit Ausnahme von 4Earth*
 - *Entwicklung konsortienübergreifender Dienste und Verständigungen zu Zertifizierung etc. (S. 36f.): Sektionen und Base4NFDI, aber auch Potenzial für multilaterale Kooperation*
- o *Quantitative view: analyse table, in how many of the listed activities are we active, and how intensely? f.e. low intensity in ELSA (listening), medium in Edutrain (active participation), high in Infra (strategic lead)*
 - https://docs.google.com/spreadsheets/d/1zQOMdQpI_CKEhVCtBdDGPEwKsbMWc9ii/edit?usp=sharing&ouid=118125260863039608675&rtopof=true&sd=true => cross community coordinators

- o *Single domain collaboration (life sciences)*
 - *DataPlant, NFDI4Microbiota, interest in BiolImage, NFDI4Agro (common partners)*
 - *with 4Microbiota: joint interest in RDC, Hackathon*
 - *DataPLANT, 4Microbiota, FAIRagro:DSI group*
- o *Cross Domain Collaboration*
 - *with 4Earth: joint use cases, strategic linking with their pilot projects; important data in the domain, Hackathon*
 - *with KonsortSWD: Data management for FEdA projects => Poster as reference*
 - *with 4Culture: Hackathon*

1.3 International networking

How have you implemented and further developed your strategy for embedding your activities in the international context? What forms of international cooperation does the consortium pursue? How do you perceive the services you have established with a view to international activities and services in this area?

NFDI4Biodiversity is settled into an extensive landscape of initiatives, projects, and infrastructures at the national and international level. They organise method-oriented sub-communities in the field, negotiate standards, and develop important international research infrastructures. Technical and organisational integration within this landscape is a cornerstone of consortium activities. A selection of mature tools and data standards from these networks is used and disseminated in the NFDI4Biodiversity project in order to mobilise and publish data collected by national stakeholders [O10]. In the bioinformatics subcommunity, use of the Biotools registry [B7] is encouraged. The option to register own services in the EOSC marketplace has been earmarked as an appropriate next priority in TA3.

Regarding representation, we aim for maximal cross-network coherency rather than maximising involvement. Efforts are coordinated in Task Area 2 ("2connect"). An assessment was performed to identify the top ten candidates for engagement in the initial project phase of NFDI4Biodiversity out of ~60 initiatives (for results, see Table X). The actions taken range from actual cooperation (cross-fertilisation through coordinated engagement in both networks, as in the case of GBIF, TDWG, bioschemas.org and ELIXIR) to liaison (active reporting on activities through delegates, as in the case of EOSC and RDA). With the exception of bioschemas.org we were able to build on existing consortium contacts and activities. Cooperation with ELIXIR (the European Research Infrastructure for the Life Sciences and its German node de.NBI), as well as bioschemas.org also involves multilateral cross-consortia activities (workshops, seasonal schools).

Activities have been compiled in two internal strategic reports (2021 and 2023), which serve to inform the Steering Committee and the Strategic Advisory Board [O11, O12]. The

documentation of relevant national and international players is continuously maintained and reprioritised.

Notizen aus Schreibklausur etc.

- o Point of reference: Chapter 2.4 of the proposal
 - Participation in selected networks has developed differently from what was proposed. Partly due to how the project evolved, partly due to lack of resources. Results of assessment regarding the international networks named in the proposal (<https://kb.gfbio.org/display/NFDI/Initiatives+Network>):
 - GEO BON / Europa BON: medium high priority /needs strategy. Activity mainly through Aletta (informs about developments)
 - eLTER: medium high priority. Activity mainly through Use Case project, one presentation of 4Biodiv in eLTER Meeting 2021 (Mark). No alignment with policy
 - NEON: low priority, no activity
 - ELIXIR - high priority / with de.NBI, multiple activity, also in training area
 - CETAF: low priority
 - EOSC: high priority, few activities (somehow unsure how to connect)
 - RDA: high priority, selected activities, benefits from bioschema and other WG
 - DataONE: low priority, no activity
 - TDWG: high priority, regular activity (conference contributions)
 - ESIP: not rated
 - ENVRI: low priority, ENVRI Plus in eLTER RI & LifeWatch?? To do
- o Our work profits from established international activities which provide accepted tools and standards: The Global Biodiversity Information Facility (GBIF) with its strong stakeholder group of collection data centres and related initiatives on the European level; and the bioinformatics domain with the international nucleotide archives and European research infrastructure initiative ELIXIR (German node: de.NBI).
 - Grundlage: Reports of TA2, <https://kb.gfbio.org/display/NFDI/Reports+of+TA2>
- o Strategy for embedding..

“within Task Area 2 we will team up to coordinate international implementations and harmonise access to existing services” (proposal p. 11)

 - First strategic report as (internal) source for text esp. assessment of current involvement,
 - The Global Biodiversity Information Facility - GBIF
 - very well embedded: German node institutions are partners, B. Geminholzer is chair Scientific Board
 - GBIF hosted portal
 - Biodiversity Information Standards - TDWG
 - very well embedded: WG leads are partner, Board position for TDWG president
 - Workshop and paper on national data infrastructures (##reference)
 - Research Data Alliance - RDA
 - watcher /no implementation of standards in first half of project (except if bioschemas.org is an RDA recommendation?)
 - European Open Science Cloud - EOSC
 - indirectly through ERICS in the field
 - Which services are listed in EOSC catalogue?
Recherche nach Partnereinrichtungen, Ergebnis April 2022
(<https://helpdesk.gfbio.org/browse/NFDI4BIO-433>):
 - o GWDG - SSH Open Marketplace; B2Share
 - o de.NBI - NBI Cloud
 - o DSMZ - BacDive
 - o Heidelberg Institute for Theoretical Studies (HITS) - kein konkreter Service
 - o AWI in cooperation with MARUM - kein konkreter Service
 - desweiteren ggf. die Beteiligung und Gespräche mit FAIRCORE4EOSC und FAIR-IMPACT
 - participation in <https://eosc.eu/eosc-task-forces> - müsste man erfragen
 - Bioschemas - Bioschemas.org
 - focus of work in TA2M3 in the first phase: two workshops 2023, participation in user group (?)
 - (Atlas of Living Australia - ALA; framework not established, but useful tools -
 - currently Lebendiger Atlas - Natur Deutschland established as GBIF hosted portal; GBIF is closely collaborating with ALA, and there are developments for aligning tools)
- o Service strategy: avoid the implementation of parallel structures - establish portfolio as part of the international landscape. Examples (Reference: ##Service Portfolio report):

- *GFBio Terminology Service: moved inhouse solution to Biportal, engagement in Ontoport Alliance*
- *The German GBIF nodes' capacity to support data providers in Germany with BioCASE tools is enhanced, leading to mobilisation of additional datasets for GBIF and the public domain*
- *introduce bioschemas.org as standard*
- *GBIF hosted portal service used for Lebendiger Atlas*
- *IAM: Life Sciences Login option for all services established*
- *pending: Assess how the recently launched ELIXIR RDM kit and data life cycle relate to the GFBio RDM framework and data lifecycle that is currently being used in the consortium*
- *activities w/o international scope: DMP Tool: national developer network RDMO;
RDC components - internat. frameworks??*

1.4 Organisational structure and viability / sustainability

Have you implemented the governance structure as foreseen in the proposal, or have you adapted and/or modified the internal structure? To what extent has the spending of funds deviated from the proposal's financial plans, especially with respect to the distribution among partners? How reliable and sustainable are infrastructures/structures that you are integrating / building on?

The governance structure was implemented with minor adjustments:

a) The concept of a "Scientific Advisory Board" was replaced by a "Strategic Advisory Board", with representatives from science, society and policy to include a broader range of resource provider and user perspectives (as proposed for milestone M1.1.1).

b) The independent "Board of Governors", conceptualised as an "Aufsichtsrat" and active link between the NFDI Governance and the Consortium, became obsolete once the governance of the NFDI Association and the Consortia funding through DFG became clear. Positions foreseen for funders were included in the Strategic Advisory Board (also see 1.1).

The disbursement of funds was adjusted after the initial budget cut of 27 % (see 4.2): The support staff foreseen for Task Area Leads (0.5 FTE) was dropped, except in the case of TA1 with its large stakeholder group. Staff resources available to Co-Applicant institutions were reduced to max. 1 FTE (except for the staff pooled in the joint project office). A central budget management was established to deal with underspending due to vacant staff positions and sickness rates (see 4.2). Underfunding in combination with the inflexible annual usage regime for the grant is one of the risk factors the consortium has to deal with. Few of the partners have mandates to run national services, but rather open their tried-and-tested infrastructure and services for external users. This means that central infrastructure (like the collection data centres and PANGAEA) is organisationally stable, but the operational staff for Community services (data curators, developers etc.) is mainly project-funded (see also 4.4).

Major international infrastructures at GBIF and ELIXIR appear rather stable in terms of organisational continuity, depending on budget negotiations with their member states. Sustainability of their matching German nodes is less clear: Securing the future of the de.NBI Cloud and its community-based service centres (part of the German ELIXIR node) has been a

major challenge during the runtime of the NFDI4Biodiversity project, and funding for the German GBIF nodes ran out in 2014.

At the consortium level, the NFDI4Biodiversity Helpdesk and back office team and the academic cloud services the consortium uses will need to be stabilised beyond the current cooperation agreement. The managing partner GFBio Association is building competencies to manage additional agreements with providers and is the long-term structure for this task.

Notes from Schreibklausur

- o Point of reference: Chapter 2.5 of the proposal (p. 42f.)
- o Co-applicants: no staff for TA leads
- o SIG Grants not implemented, SC decides directly on recommendation by the speakers / TA5M1 - gab nichts zu verteilen
- o Underspending year 1, asymmetric budget cuts (see 4.2); calculated Mittelübertrag in 2022;
- o Disbursement of funds/financial support: the distribution of funds between partners has evolved; regular annual adaptation of financial plans by SC - Regeln für die zentrale Mittelbewirtschaftung, see also 4.2
 - staff resources 1 per co-spokesperson (except A-01 and C-01)
 - spontaneous flexfunds in 2022, VERY productive
 - invest small amounts in optimisation of tools, for example API Checklist Red List Centre
- o Decision making
 - Scientific AB became strategic AB (merged with ssfb)
 - Board of Governors not implemented due to NFDI involvement
 - GFBio e.V. as managing partner
- o Viability
 - Surveys All Hands Conference
- o Means of communication: much more online due to covid
- o Risk management / risk list continuously managed, executive report
- o Sustainability of infrastructure/structures we are building on
 - rather stable in terms of organisational continuity: Collection data centres, ENA, PANGAEA (reference: GFBio data centres)
 - rather stable in terms of organisational continuity: GBIF and ELIXIR structures/tools
 - not stably funded: German GBIF nodes
 - stable in the context of the consortium cooperation contract: Helpdesk and the distributed team for second level support (partner contributions, mainly funded by NFDI)
 - stable in the context of the consortium cooperation contract: de.NBI and academic cloud resources

1.5 Operating model

- Mode of operation and financial model

Has your mode of operation and your financial model worked well for both providers and users? Can you share some of your experience in this regard? With a view to the perpetuation of the NFDI: What parts of the financial model have proved of value, which have not? Have you applied or observed user fees at all? If so, please share your experience in quantitative terms. Were partners' "own contributions" (in-kind and in-cash) sufficient overall?

Our mode of operation is that of a distributed resource provider and expert network: Data providers can team up with scientific IT centres and developers to implement data management solutions and/or mobilise data for scientific use. IT development capacities (both central and on-site) are provided to develop or improve software for data handling. Individual researchers

receive support through the central helpdesk at the GFBio Association. So far, this mode of operation works sufficiently well to improve overall data availability and increase competencies in the community, while allowing for different speeds. The pace is often slower than expected, and the network will need to grow further to include other major stakeholders. Community feedback has been positive, while stressing a desire for additional networking opportunities.

Pillars of the financial model are the DFG grant and own contributions. 90% of the budget are used to cover staff costs. Apart from coordination and management, co-applicants usually have one FTE (12 person months per year) to support with their tasks in the work program, plus in some cases additional person months for specific use case projects. Financial support for participants is between 6 and 36 person months during the course of the project. Fixed minor budget allocations for use case partners (6-18 person months) turned out to be hardly manageable and would better be allocated through annual flexfunds calls. The 1 FTE project staff in co-applicant institutions is sub-critical in cases where services are in high demand or complex software is developed de novo. For other providers, the additional FTE are satisfactory and help bring their resources to NFDI effectively. For all co-applicants and several participants, the communication overhead in the multiple layered networking structure of NFDI is quite high and time consuming.

Own contributions of partners are often their own infrastructure and services, which are opened for external users (see 4.1). Additional workload is in part or fully shouldered by project staff. Support with tools or curation can also be in-kind, f.e. with services that are designed for public use and attract users through NFDI4Biodiversity. So far, the infrastructure and storage has been sufficient, due to the fact that much of the data mobilisation takes place at the provider institutions themselves. However, when the RDC reference implementation becomes operational, re-negotiations may become necessary.

Notes from Schreibklausur

- o *Point of reference: Chapter 2.6 (page 51)*
- o *Fiscal year problem*
- o *Of value: Flexfunds (see SC meeting June 2023)*
- o *Experience with finance model*
 - *Co-Applicants: financial model generally okay (see <https://kb.gfbio.org/display/NFDI/2023-03-02+TA5M2+Meeting>),*
 - *own contributions sufficient? no formal assessment*
 - *Participants with low budgets (use case partners): mixed experience; better not assign these small badges in advance, rather plan implementations as flexfunds projects. In some cases a small Finanzspritze at the right time can work wonders)*
 - *Need flexible funding to stimulate quality management in a provider network: providers can sometimes not finance the necessary improvements of services diagnosed in annual service performance reviews (risk identified in SIG Service Monitoring, ##reference Service Portfolio reports)*
- o *User fees: no (and not planned, too much administrative overhead, liability, taxes)*

- Data storage and computing power

Which infrastructures not funded by the NFDI with data storage and computing power have contributed in a relevant manner to the services you have developed? Please describe in quantitative terms.

Notes

- o de.NBI Cloud at University of Gießen (Core storage reference implementation)
- o de.NBI Cloud at University of Bielefeld (implementation of data analysis pipelines)
- o GWDG cloud (website)
- o MARUM (collaborative working platform, helpdesk - Jira and Confluence)
- o Marburg - ?
- o AWI (data archiving)

It is important to note that in the early phase of NFDI4Biodiversity, the cloud services were conceptualised as a federated resource (multi cloud). Partners had earmarked de.NBI cloud resources for NFDI services. Since the transfer of de.NBI to Forschungszentrum Jülich, the future of the de.NBI Cloud and the method-oriented service centres is not fully clear yet / under negotiation. Currently partners Giessen and BIBI provide Cloud infrastructure as own contribution

2 Research Data Management Strategy

The following subheadings correspond closely to the funding criteria on the Research Data Management Strategy. Consult the Guidance Notes on Funding Criteria (DFG form NFDI 120) for more information on the specific subheadings and the NFDI Guidelines (DFG form NFDI 100) for general information. If you refer to existing information infrastructures, data repositories or reusable software, explain how they are (or were) integrated. List references in the appendix under Bibliography.

2.1 Scientific relevance and quality of the measures

How successful has the implementation of the planned measures been (with respect to the description in the proposal)? Have you achieved the aims of your task areas? How has progress related to the timeline sketched in the proposal? What deviations from or reorganisation of the work programme became necessary, possibly also due to contributions to the NFDI as a whole? What relevance do the implemented measures have for the community they address and/or for the individual members of this community? How, i.e. based on what criteria, would you measure the relevance and quality of your RDM strategy?

The RDM strategy involves a **significant engagement of researchers, public authorities, expert organisations and citizen science**, tackled in Task Area 1 (2involve), with 26 use case projects and several support and training measures. Interviews were conducted with use case partners and a support strategy devised (TA1M1, TA1M2). The GFBio Helpdesk infrastructure has been successfully restructured for purposes of the NFDI4Biodiversity consortium (TA1M3, [O2]). Researchers are supported through the established GFBio service portfolio, structured

along the Research Data Life Cycle [B8]. Training and education activities were developed on-demand with partners, including a series of teaching videos [O13, O14]; a five-day Winter School with the GfÖ - Gesellschaft für Ökologie [O15-O17], two workshops on legal issues and **nine** RDM trainings, including **six** at Universities and research institutions (see O18 Training and Education Report). The establishment of a training platform has been postponed until opportunities for joint infrastructures become clearer in the NFDI section EduTrain. Training materials are published on the NFDI4Biodiversity Zenodo Community and YouTube Channel. In the use case projects, several of the anticipated risk scenarios emerged. Shifts in requirements, different speeds of partners and capacity shortages had to be dealt with. After rescheduling part of the work program in early 2022, some conflicting interdependencies between measures could be resolved. Despite COVID-19 and a complete shift to remote work in the first 24 months of the project, considerable progress was made towards actual mobilisation of data from use case partners and implementation of tools.

Highlights:

- IÖR Data integration in GdO Demonstrator Portal (UC6, UC15)
In a joint effort, data from participants in Use Case #6 (Land use/covering) and Use Case #15 (occurrence data from the German Dragonfly Society) were combined with other environmental data sets in the cloud-enabled geovisualisation tool Geo Engine [B9], to demonstrate potential portal components for partner organisations [O19, O20], similar to the VAT tool in the GFBio Portal [B10]
- Enabling FAIR data management practices at NP Bayerischer Wald (UC12):
In a partnership with Uni Jena and the general directorate of the state archives of Bavaria (SAB/GDA), data from the National Park can be managed in a hosted BEXIS solution at the Biodiversity Exploratories Data Center, with new workflows to ensure the required long-term archiving at the Bavarian State Archives (work in progress)
- Mobilising of historical species occurrence and forestry data in Bavaria through the general directorate of the state archives of Bavaria (GDA)). Together with the Leibniz Institute for Ecological Spatial Development (IÖR), a workflow for the automated extraction of metadata from analogue map series, e.g. peatland maps, was developed [[Use Case Staatliche Archive Bayerns - NFDI4Biodiversity](#)] (UC 24)
- A data analysis exercise with data mobilised for iDiv PlantHub (UC 22) and publicly available data from iNaturalist yielded surprising results and was published in Nature Ecology & Evolution [O22]
- Mobilising freshwater fish occurrence data + geovisualisation/new data publication format (UC19 IfL + iDiv/FSU): Data story Red list freshwater fish.....

- Launching Critterbase Data Warehouse for marine organisms as new data product and service for marine organism research community (UC5) - [OX-OY]....
- Launching Lebendiger Atlas Natur Deutschland (UC23): new data product and service for the German biodiversity data provider community (GBIF hosted portal, [O21])
- Waterlink R package as new service for aquatic biodiversity research community ? (UC26): cross-community use case with NFDI4Earth [O23]
-

The combination of active support in terms of expertise and resources resonates well with the partners. Funding for use case and flexfunds projects helped to go “the last mile” for improving data resources and tools so they can be shared with others, as in the case of Hydrographr, [O21]). The training activities were received very well - the Winter School had more than 120 applications for 20 seats. There is obviously a need for in-depth training. Effective co-operation with regional and local programs like DataTrain in Bremen and the Hessian Research Data Initiative HeFDI is fruitful and efficient in terms of resource management. Self-learning materials are a new instrument with a high potential to be used in undergraduate learning. Their actual uptake will be monitored closely.

The second pillar of the RDM Strategy is a **joint cloud-based infrastructure** called Research Data Commons (RDC, [O22]). The architecture of RDC is a self-service platform with intermediary services for data providers and users. The necessary components are being developed in Task Area 4 (see 2.4). A number of service components, like the Aruna Object Storage [O23, O24] the Terminology Service [OX] and the search index [OZ] are already used for the development of first demonstrators, but still require extensive testing until their final release. A full design and implementation of the architecture is currently under discussion across consortia, moderated by the speakers of the NFDI section Common Infrastructure.

In project year 2 and 3, we managed to move several partner services to the cloud and implement pilot applications with data from different use case partners. We use Cloud services of the German Bioinformatics Network de.NBI (the German ELIXIR node) and of the GWDG - Gesellschaft für wissenschaftliche Datenverarbeitung (see 1.5).

- Several GFBio Core Services (e.g. DMP Tool, Data Search, GFBio Portal [B8]) were deployed on the de.NBI cloud infrastructure designated for the initial implementation of RDC.
- The metadata harvesting infrastructure, which collects and indexes metadata from ten data centres to power the GFBio Data Search, was adapted for and re-deployed on the de.NBI cloud. In addition, a showcase was developed to support the mobilisation of data from Use Case #18 “Fischartenatlas” of Gfl within a demonstrator of the NFDI Data Discovery Portal (Milestone 4.4.2, [O27]).

- As part of measure TA1M4 (training and education), a scalable JupyterHub infrastructure for collaborative analysis and training has been implemented on the RDC infrastructure and made its debut in the one-week GfÖ Winter School on Data Management in December 2022 [OY].
- A Cloud-based training environment for the data management system Diversity Workbench [B11] has been set up in the GWDG Academic Cloud and is currently piloted with selected users [OZ].

An anticipated risk scenario emerging in TA4 was the difficulty to retain and recruit IT staff. Several work packages were outsourced to external developers in 2022 to keep timelines.

All in all, the integration of resources and needs of over 50 partners is ambitious. Out of the 24 milestones due in the first two years, 15 were completed in time and three have delivered results, but are formally not completed because the necessary technical documentation is not finished. Two milestones due in project month 30 were finished early. This is much better than the results of year 1 (with only 6 out of 18 milestones reached in time). However, there is still the necessity for more conservative planning, considering the scarce staff resources in TA4. Despite the adaptations in the work program, various measures keep being delayed because capacities were overstretched and processes between stakeholders are slower than expected.

Notes from Schreibklausur

- *Point of Reference: Chapter 3 of the proposal (p. 58 ff.)⁴*
- *TA1M1: Use Cases*
- *TA1M2: support with tools*
 - *uptake in year 1 placed back due to slower onboarding of use cases; now established:*
 - *BEXIS data information system for National Park (Uni Jena)*
 - *Support with BioCASE provider software - several partners (BGBM)*
 - *Onboarding of services & organisation of support: RightField, BIIGLE...*
 - *User trainings: DWB, will become more intense in year 4+5 as more tools have been onboarded*
- *TA1M3: Front-office back-office model (Pilot: FEdA; plus others in the pipeline)*
- *TA1M4 / embedding in university teaching; good progress in milestones*
 - *Lehre BSc Marine Geoscience (Bremen, based on materials adapted from GFBio)*
 - *Lehre MSc MarMic Blockkurs RDM based on materials adapted from GFBio)*
 - *Data Train Interdisziplinär, Cross-NFDI, MSc, Phd, Postdoc (Bremen + bundesweit)*
 - *XY RDM trainings delivered at Universities, for research initiatives, at GfÖ Annual Conference and in Winter School*
- *RDC (TA4) - components: test integration in cloud-based storage?*
- *New data products/data streams community: TA3M3*

2.2 Metadata standards and reliable services

To which metadata standards or other RDM standards did your consortium contribute?
Which reliable and sustainable services have you established? Please explain each item in some detail.

The NFDI4Biodiversity Community works with a portfolio of accepted metadata standards for data and digital objects in the field of taxonomics, natural history collections and genomics [see B12]. Several consortium partners are engaged in standardisation activities, f.e. TDWG [B13], the Genomics Standards Consortium/MIxS [B14] or MIAPPE [B15]. Partners mobilising occurrence data for GBIF and the Atlas of Living Germany [O19] are encouraged to deliver the ABCD consensus elements agreed in the GFBio project [B16]. Within the project, implementation of Bioschemas [B17] was prioritised as the most promising standard to achieve convergence of the consortium in terms of discoverability of data resources, and increase potential for trans-disciplinarity (activity of TA2M3). An exploratory workshop with data providers was conducted at the All Hands Conference in 2022, followed by two hackathon-type workshops in 2023, for example to improve findability of partner resources in search engines. These activities were supported by experts from DataPLANT and ELIXIR. The ambitious work plan for TA2M3 was reduced after the initial budget cuts. Instead of the five milestones originally foreseen (including an extensive landscape analysis), a prioritisation exercise is conducted (starting from the compilation in [B12], due in project month 29), and existing studies from recent years will be brought together and documented in a “meta analysis” (project month 42). The project consortium will not conduct its own standards ratification activities. If necessary, separate funding will have to be raised.

In terms of semantic integration, progress has been made to expand the GFBio terminology service [B18] into a centralised repository and service for semantic applications in NFDI4Biodiversity and beyond (TA4M2). Caching mechanisms to centrally store essential decentralised resources have been implemented as ETL (Extract Transform Load) processes. In particular, resources useful for the portal semantic search (work in progress: M4.2.4 adaptable semantic integration) are being prioritised.

We carry on linking our efforts with European and international initiatives and infrastructures. A close cooperation with the Ontoportal Alliance (<https://ontoportal.org/>) has been initiated. The backbone of the GFBio Terminology Service is being replaced by an OntoPortal installation called BiodivPortal⁵. Beside contributing to an international collaborative effort, we are involved at an early stage in standardisation efforts for terminology evolution⁶ as well as the adoption of standards for terminology mapping⁷. This will reinforce our work on terminology mapping and linking capabilities [B24,B25] and versioning and synchronisation mechanisms [B26].

Partners involved in TA4M2 continue developing semantic metadata exchange formats based on internationally agreed semantic standards that are required by the RDC Semantic Storage (TA4M3) for data linkage and schema transformation. The initial catalogue of metadata models

⁵ <https://github.com/BiodivPortal>

⁶ <https://w3id.org/kgcl>

⁷ <https://w3id.org/sssom>

has been expanded with semantic core models including Schema.org, ABCD 3.0, Dublin Core and Datacite. The adoption and extension of such models is done in close cooperation with the community. In particular, the need for a more fine grained description of observable properties led to the adoption of the RDA I-Adopt⁸ (Interoperable Descriptions of Observable Property Terminology) semantic framework and its integration to the metadata model of Schema.org. A slice of I-Adopt terminologies is made available via BiodivPortal and an automatic annotation is provided through the service to enable an automatic generation of I-Adopt metadata annotations [REF CoRDI?].

Notes from Schreibklausur

- *##Reference: Service portfolio report and/or Wiki articles (TA4 Workspace / Knowledge base)*
- *Point of reference: p. 68 of the proposal (chapter 3.1)*
- *We are already able to work with accepted community standards in the fields of taxonomy, collections, genomics*
 - *regarding project activities: decision to implement bioschema.org in our data resources as most useful activity at present, two workshops (TA2M3)*
 - *consortium partners are regularly engaged in standardisation committees (TDWG, GBIF, MiXS) - bring know-how as planned*
- *landscape analysis was cancelled due to budget cuts*
- *metadata-related services:*
 - *GFBio TS;*
 - *onboarding of RightField (pending),*
 - *BioCASE and ABCD consensus element rollout...?*

2.3 Implementation of the FAIR principles and data quality assurance

To what degree and in what way have you achieved implementation of the FAIR principles?⁹

Have your procedures and regulations for access to and use of data proved of value?

How have you implemented protection of personal data? If applicable: What is the consortium's contribution to reducing uncertainties regarding General Data Protection Regulation-compliant handling of research data?

The consortium is committed to make existing data discoverable and reusable in accordance with the FAIR principles (see values in the mission statement [O3]). Know-how is disseminated through training and education activities [O18] and Use Case projects. Best practices already available in the community have been applied specifically in the mobilisation of species occurrence data from Use Case partners:

- The **findability** of data is being improved by publishing datasets in a suitable GFBio Data Centre with supportive curation staff. Data is then pushed to the Global Biodiversity

⁸ <http://ceur-ws.org/Vol-2969/paper10-s4biodiv.pdf>

⁹ Please refer to the "White Paper: Interim Report Reference" to describe to what degree you have achieved the implementation of the FAIR principles. Cf. <https://doi.org/10.5281/zenodo.7688729>

Information Facility (GBIF) and to the GFBio portal, with their respective community-oriented search indexes.

- The deposit in data archives also improves **accessibility**: The data archives issue persistent identifiers and ensure that access rights are specified by a clear license.
- To increase **interoperability** of the species occurrence data, providers are encouraged to deliver the ABCD consensus elements with the metadata [B16] and to use standardised ontologies offered as part of RDC. In particular, the IT infrastructure of data archives ensure machine-actionability of the metadata.
- To increase **reusability**, we encourage data providers to use the Creative Commons licenses CC0 and CC-BY. Parts of the community prefer to use CC-BY-NC licences to exercise more control over their data. A discourse is underway to replace this practice by using CC-BY-SA, which is less prohibitive. We also recommend considering data publications as good practice to document provenance and scope of the data. The GFBio Data Centers also offer distributed archiving as a service to deposit outputs of multi-method projects in dedicated archives, while maintaining relationships between them.

In this way, implementation of FAIR data practices has been supported for the Red list of freshwater fish [OX] and data from the fish atlas of the German Ichthyological Society (GfI e.V.) [OY]. Partners who were able to improve FAIR data practices through flexfunds projects are

- AraGes
- Partner IPK mobilised biodiversity sequencing data regarding the FAIR guidelines described in <https://faircookbook.elixir-europe.org/content/recipes/reusability/miappe.html>. In detail 2,500 records have been registered at EMB-BioSamples. Furthermore, 12 biodiversity projects have been registered at EBML-ENA. In total an amount of 39 Terabyte sequence raw read files have been transferred.
-
- DDA
- ...

One of the GFBio Data Centers, the data publisher PANGAEA, has been confirmed to have a high degree of FAIRness according to the Fuji-Tool (FAIRsFAIR, [B19]). The other GFBio Centers are assessing if a FAIR assessment exercise is feasible for them at this point.

Starting point: p. 69 of the proposal (chapter 3.2)

- massive downsizing (for an introduction see White Paper Interim Report Reference)
- explain major obstacles to FAIRness in our Community (Datenschutz randständig) => eher inkonsistente Liefer- und Verwertungsketten

White Paper: "an outline of how your consortium has been working to achieve implementation of the FAIR data principles and what has been done in order to monitor progress should be given."

Actual focus of work in project year 1-3:

- early data mobilisation (p. 64 of proposal)
 - Fische als Beispiel für Entwirren von Verwertungsketten
 - Mobilisierte Datensätze (KPI oder Outputs)
- Supporting tools (TA1M2 philosophy; describe role of "TA4M1 brokering framework")
- iAdopt framework (Naouel Karam)
- Checkliste Editor API..?
- FAIR principles in data centres:
 - PANGAEA: high degree of FAIRness according to the Fuji-Tool (FAIRsFAIR)
 - <https://www.f-ujii.net/>
 - <https://op.europa.eu/en/publication-detail/-/publication/03b5562d-6a35-11ed-b14f-01aa75ed71a1/language-en>
 - => use case for other data centres?
 - Was ist mit ENA? Gibt es ein Assessment?
- Educational activities - FAIR principles regular building block in training curricula

2.4 Services provided by the consortium

In describing the services currently provided by the consortium, distinguish clearly between services that consortium members provide as part of their institutional mission (*Grundaufgaben*) based on existing funding, and new services that have been established within the NFDI framework.

NFDI4Biodiversity has a strong backbone of **Data archives**, where data can be stored, published and discovered, subject to the Data Centers' individual missions. The GFBio Data Centers [B21] already form a group with a joint submission service and helpdesk as well as a Data platform [B22]. Within NFDI4Biodiversity, the group organised admission of further Data archives under a joint mandate, to form an archiving network and close gaps in terms of scope and target communities. Project funding is used to cover additional costs, where archiving exceeds institutional missions (see [OX - service report], p.X).

The Research Data Commons [O22] are the architectural framework for a comprehensive **data platform**, exceeding the scope of the present GFBio portal [B22] by additional end-user services for data transformation, discovery and analysis as well as a central storage solution. Components are pre-existing solutions like the GFBio Terminology service [B18], which are being enhanced within the project, as well as the ARUNA Core Storage infrastructure [O25,26], which was developed by the University of Gießen with funding from NFDI4Biodiversity, NFDI4Microbiota and the FAIR DataSpaces project [B23]. NFDI4Biodiversity aims to provide a reference implementation of RDC and first use cases demonstrating the advantage of using its tools. Operation and sustainability of a productive RDC data platform are under negotiation across consortia within NFDI.

The central **Consulting service** structure provided through the engagement of several NFDI4Biodiversity partners is a helpdesk system with an attached ticket infrastructure, operated by GFBio e.V. with UNiBremen/MARUM as technical provider (see [OX - service report], p.X).

Individual consulting services offered by partners can plug into this system by training staff for back-office operations (TA1M3). Staff available for Helpdesk support is mainly project-funded.

Overall, partners bring 60+ individual services and installations to the project, which are in part for internal data management tools, and in part publicly available community resources. Task Area 3 organises collection and profiling of these services in an initial service catalogue. From the results so far, the largest group are **Databases (X)**, followed by **Web applications (X)** and **software** that can be downloaded and executed locally on the users' hardware(X). Of these, XY are supported as community services through the NFDI4Biodiversity Helpdesk [OX/ Service Portfolio Report and OZ/Website NFDI4Biodiv-Services]. A quality management structure has been proposed by the SIG Service Monitoring and FAIRness, which will support service providers with proper documentation and technical upgrades. Further partner services will be onboarded through a structured process, which is currently piloted in Task Area 3.

Project funding has been used to provide additional community resources: A hosted Jupyter Notebook was set up in the de.NBI cloud for training purposes [OX], and an API was programmed to get machine-actionable reference data from the Checklist Editor of the Rote Liste Zentrum [OZ]. The GBIF hosted portal infrastructure was used to set up a "Living Atlas of Germany", featuring regional GBIF data and data providers from Germany [O19]. The Visualisation, Analysis and Transformation (VAT) component in the GFBio portal was upgraded technically to the latest Geo Engine technology [B9] and enhanced with customisation features to allow for easier integration with other partners' data portals [OZZ].

Table of Community services

Service types and partner roles are named according to the White Paper "interim report reference" provided by NFDI. References are Outputs (O) or listed in the Bibliography (B).

Type	Name	Partners and roles	References
Consulting	Central Helpdesk	GFBio e.v. - Operator Uni Bremen - Technical provider	
Consulting	- Data management helpdesk	GFBio e.v. + GFBio Data Centers - Consulting	
Consulting	- Certification helpdesk	GFBio Data Centers - Consulting	
Consulting	- Technical services helpdesk	Service provider staff - Consulting	
Web application	Data management planning tool	GFBio e.V. - operator, consulting GFBio Data Centers - consulting	
Web application	Data submission service	GFBio e.V. - operator GFBio Data Centers - Curation	
Web application	BIIGLE Image and Video Annotation Service	BIBI / Uni Bielefeld - Operator	

Web application	GFBio Data Portal - data discovery service	GFBio e.V. - Operator GFBio Data Centers - data provider	
Web application	Living Atlas of Germany (data discovery)	BGBM - Operator GBIF - technical provider	O19
Web application	Critterbase	HIFMB - Operator —	
Web application	Jupyter Notebook 4Biodiversity (data analysis)	?	OX
Tool/ application	BioCASE provider software	BGBM - consulting BioCASE consortium? - Technical prov	
Databases	PANGAEA	UniBremen/MARUM and AWI- Operator AWI - technical provider	
Databases	German Barcode of Life Reference Library	LIB - Operator	

Notes from Schreibklausur

- Starting point: p. 65 of the proposal (chapter 3.3): which or how many of the listed services have been implemented?
- "Startkapital" and extension of GFBio services
- Grund- and Projektaufgaben:
 - project funding for new services
 - project funding for use of existing institutional services by others (f.e. additional PANGAEA data manager)
- Results from initial sustainability assessment of the engagement yielded first indications on activities which would need to be stabilised by future funding
 - "die nationale Aufgabe" = Sammelauftrag für Biodiversitätsforschung Deutschland; wir erfüllen den gewissermassen als Kollektiv, aber lückenhaft (z.B. Data centres mandates are institutional, Collection Data centres have narrow mandates and are very selective)
 - backoffice institutions
 - no Grundfinanzierung for GFBio services

##Reference: Service portfolio report - results of TA5M2 activities (Hintergrundgespräch, evtl. sustainability assessment)

2.5 Impact of changes of external conditions / constraints

How have developments in the research and infrastructure landscape impacted on your work? What developments or changes were (most) relevant? How did you react to these? Please only name aspects you have not already described above.

NFDI4Biodiversity was designed as a consortium with broad stakeholder representation in order to work on more efficient data management and data sharing processes, and achieve consolidation of existing initiatives. An unexpectedly high workload was necessary, however, to

build relations with new initiatives arising in parallel, like the FEdA research initiative [B27, O8], the National Monitoring Centre NMZB [B28], the umwelt.info Portal [29], GAIA-X [B30] and the Data Competence Centre Call of BMBF in 2022 [B31], who have in part similar missions and need to be included in consolidation efforts to avoid duplicate work. Relations were actively built by Memorandum of understanding (O8), engaging in advisory boards (NMZB, umwelt.info), project participation of partners (FAIR Data spaces, [B23]) and letters of intent for collaboration (four data competence centres). Requests for collaborations were also brought to us by stakeholders and resulted in letters of support for project proposals (n=12) or bilateral consultations and matchmaking with partners (n=13). In addition, there are large funding programmes coming up in the nature conservation sector [B32], and action is necessary to lobby for proper data management mandates, which are unfortunately not common in these types of funding programmes yet. Partners were also excessively engaged to save parts of the de.NBI service network built as part of the German ELIXIR node, when a major change in finance and governance was decided in 2021 [B33]. This intense relationship management required roles and resources that were originally not covered in the work program.

Notes from Schreibklausur

- *Kleinteiligkeit der Landschaft hat kaum abgenommen,*
 - *Basisdienste Ausschreibung, Konsolidierung in Base4NFDI*
 - *DKZ Konsolidierung*
 - *KI Ausschreibung fehlende Verknüpfung zu NFDI*
 - *GAIA-X noch Beteiligung gewuppt (FAIR Data Spaces), aber beim Dateninstitut war kaum Beteiligung zu mobilisieren*
- *de.NBI unklare Zukunft als cloud provider, viele PIs der Konsortien involviert - eigentlich gedacht als konsolidierte INfrastruktur, stattdessen hat eine nur in Teilen erfolgreiche stabilisierung viel Energie gebunden*
- *Neue große Förderprogramme im Naturschutz, Sensibilisierung BMUV für die Chancen der NFDI, Gespräche*
 - *AKN, BFN Datenverfügbarkeit nicht in den Förderbedingungen*
- *wenig impact: neue Regeln wie UIG und Open Data Gesetz*

3 Additional Aspects

- Equal opportunities and diversity are of general concern in research and not limited to individual criteria as listed above

Please explain your measures and achievements either in the preceding sections, e.g. composition of the consortium, integration of user communities, governance, quality assurance etc., and refer here to the respective paragraphs, or address the topic in full in this section.

Regarding gender diversity we have a balanced co-applicant consortium and aim for balanced composition in boards and committees as well as panels and speakers at events. Our German

style guide addresses the use of gender appropriate language. To foster participation of persons with family and child-rearing responsibilities, we consider family times when scheduling meetings and limited travel options, f.e. by streaming the Annual Conference and other events. To enable persons with different language skills in German or English, we generally use English on slides in German-speaking events and explicitly encourage switching to either English or German in discussions, if people are more fluent in the one or other language.

Notes

- *Grundlage: im SC präsentierte Maßnahmen*
(<https://kb.gfbio.org/display/NFDI/2022-11-25+SC-23>)

▪ Reflexions and comments

What adjustments at the overarching NFDI association level, by the DFG and/or in the structural framework do you think would be useful? What need for change do you see in your consortium?

The collaboration in the NFDI Association with its 45 working groups, task forces and committees is challenging, but often also fruitful, because of mutual learning effects: There is an open and cooperative atmosphere which enables us to share solutions to our challenges with others and profit from theirs vice versa. We would encourage to review from time to time, if the topics discussed in NFDI working groups are not already being covered elsewhere in a similar way, in order to avoid duplicate work. Workforces of consortia and Directorate have joined nicely. However, we would encourage that the Directorate is staffed adequately to support section speakers with organisational support and potentially also a strategic fund to run strategic projects within NFDI. Our difficulty to integrate Base4NFDI into the governance shows that an organisational level for joint infrastructure is missing. In public relations, the concept of the NFDI is generally received well, but we often meet with scepticism because of the project-based funding. The commitment of Bund and Länder needs to be communicated more clearly.

The management of the funding by DFG is professional and guided by an in-depth understanding of the complex organisational challenges the consortia face and the commitment to accomplish a successful initialisation of the NFDI together with all stakeholders. The internal financial management of the consortia by the Steering Committee and the speaker organisation is extremely useful to make best use of the budget throughout the five-year funding phase, allowing to organise adaptations according to needs. However, we would have appreciated the opportunity to re-calculate the budget after initial budget cuts (see 4.2). **Regarding software, we encountered the problem that professional, project-specific software support could not be purchased due to ...**

Regarding the structure and progress of the consortium, we struggle with the complexity of the partner network and too many interdependencies between measures in the work program (f.e. use cases projects to RDC development). In preparing for a Fortsetzungsantrag, we would restructure in a way that allows a) measures to be implemented by small groups and independently of each other, b) use case projects to start in their own time and pace, by allocating flexfunds budgets, c) consolidation of results and integration of partners in participatory formats such as Interest Groups, structured review exercises and Resource Provider Workshops. Regarding infrastructure and service development, we feel that joined measures across consortia could be useful to pool staff resources and avoid duplicate work. The Aruna Object Storage development is already of joint interest for NFDI4Biodiv, NFDI4Microbiota and FAIR Data Spaces project. Joining forces would also be feasible for the central RDM Helpdesk backoffice and a life sciences knowledge base for data management.

Notes from Schreibklausur and SIG Strategy

- *NFDI association*
 - *coordinators for the section, allocated to directorate to ensure knowledge transfer and compliance with NFDI procedures as chairpersons continue to change*
 - *Dreistufigkeit ermöglicht auch Lastenverteilung*
 - *Lerneffekte*
 - *workforces from consortia and directorate have joined nicely, we are content with the developments*
 - *we have mobilised strategic capability, it is important to further develop this in the coming years as we build the working foundations for the federated network*
 - *NFDI e.V. : useful and challenging but time consuming*
 - *Koordinationsstellen für die Sektionen wären eine Hilfe, um auf Dauer solide Outputs zu bringen*
 - *Das Direktorat braucht ein Budget um Strukturen für die Gemeinschaft bereitzustellen*
 - *"Außenwirkung ist gut"*
 - *"konkret würde ich mir wünschen, dass in Zukunft stärker betrachtet wird, welche Themen nicht erfolgreich auf nationaler Ebene bearbeitet werden können, weil sie fachspezifisch und international auf dem Weg sind. Das betrifft aber weniger die Struktur und mehr die Inhalte."*
 - *Zweifel, ob da eine "nationale FDI" entsteht, die den Namen verdient - die Kleinteiligkeit ist ein Problem*
- *DFG*
 - *instrumental and reliable partner in developing the NFDI*
 - *flexibility is appreciated -*
 - *high commitment for NFDI and the consortia/partners*
 - *drop fiscal year*
 - *recalculation of the Zuwendung -> 4.2*
 - *Software licensing and support*
- *Structural framework*
 - *Eigene Mittelbewirtschaftung innerhalb des Konsortiums*
 - *Konsortien nach BLV und nach Satzung ist herausfordernd*
 - *Bund-Länder-Vereinbarung ist vertrauensfördernd, trotz projektförmiger Finanzierung*
 - *Fortsetzung muss früh kommuniziert werden, sonst Vertrauensverlust*
 - *Flexfond ausbauen*
- *Consortium*
 - *koordinierte Entwicklung des Service-Portfolios mit thematisch nahen Communities (Plant, Microbiota)*
 - *"Gemeinsame Infrastruktur: Poolen von Ressourcen auf der Sektionsebene (zum Beispiel Aruna Core-Storage nicht allein in NFDI4Biodiv weiterentwickeln)"*
 - *Use case policy: better integration in workplan, maybe through flexfunds*
 - *partner and project timing often not compatible, critically low budgets*
 - *Projektstruktur überdenken?*
 - *Angeregt in Resource Provider Workshop: Orientierung an de.NBI mit seinen Servicezentren für Teilcommunities*

- *Auflösen/neuer Zuschnitt der WPs -?*
 - *“Das 4bio Programm leidet (wie auch viele EU Projekte) darunter, dass die jeweils knappen Ressourcen der Partners auf unzählige Measures verteilt sind. Diese (teilweise) Mikroressourcen sind nicht sinnvoll einzusetzen und führen zu einem schädlichen Verwaltungs- und Steuerungs- und Dokumentationsaufwand. Aus meiner Sicht wäre es wünschenswert, wenn wir in Zukunft klare Fokus-Topics definieren würden, in denen jeweils weniger Partner auskömmlich finanziert werden, so dass sie der Verantwortung gerecht werden können. Kleinere Themen können auch flexibel über externe Projekte (s. TOP3) abgearbeitet werden.”*
 - *“grow 'member' base in coordinated fashion with scalable solutions; use flexpool for concrete deliverables(tools/data products) as well as onboarding for further collaborations - 4Bio sollte Plattform darstellen für Zusammenarbeit in weiteren Projekten”*
 - *SC muss als Klammer fungieren, mit einem gemeinsamen Verständnis zu RDC Informationen aus TA dort besser zusammenbringen*
- Are there any other comments or observations from your work in the consortium that you wish to share with the expert committee?

This section is for comments or observations that do not seem to fit elsewhere.

- Tbd

***** END OF DRAFT FOR STEERING COMMITTEE MEETING IN JUNE*****

4 Spending of Funds

Please complete the following table to show the actual spending of funds broken down by institution. List the name of each institution, adding additional lines as necessary. Add additional explanations as required.

Overall Spending by Institution

	year 1 (Oct-Dec)	year 2	year 3	year 4	year 5	year 6 (Jan-Sep)	Total in €
Institution	Totals in €						
Applicant institution							
Co-applicant institution 1							
Co-applicant institution 2							
Co-applicant institution 3							
...							
Total spending per year							

4.1 Description and Summary of Contributions by (Co-) Applicants

Research data management is an intrinsic part of the research process, and research communities rely on professional management here. (Co-) Applicant institutions are therefore expected to make appropriate financial and in-kind contributions to this area. Please describe the contributions made by the applicant and co-applicant institutions. Your listing may include contributions towards services provided by institutions as part of their institutional mission (*Grundaufgaben*) that can be used by the consortium, but be sure to distinguish them from your own contributions towards realising the work programme.

4.2 Description of the cash flow developments

Please describe how the consortium reacted to the budget cuts. How were the funds reallocated/redistributed? Which adjustments in the annual finance plans were necessary, e.g. due to changes in the composition of the consortium? Please also briefly describe any other adjustments that affected the cash flow.

Reaction to budget cuts: Description available in first Executive Report

The original budget was calculated to provide for central coordination and software development teams, to create supporting positions in each of the 15 co-applicant institutions (1.5 full-time equivalents (FTE) to support their duties as task area leads) and to build a centrally managed flex fund providing participants and their use case projects with 3-18 person months of support. The budget cut required some quite painful adjustments for the partners.

After the official start of the project in October 2020, the freshly formed Steering Group (SC) adopted a new financial plan, based on the following adjustments:

- Positions for coordination and software development were reduced by 20% (minus 1 FTE each team),
- Staff for co-applicants was reduced by 30% (i.e. by 0.5 FTE, allocating max. 1 FTE each).
- Direct support for use case projects was left intact, as it was already quite low in some cases.

In terms of workload, the SC adopted an adapted milestone plan, based on adjustments like downsizing or deleting milestones (see annex A.4). Some activities were postponed to a later starting date - f.e. the promotion of “research data management as an integral part of biodiversity research” (objective 1 of the grant application), since similar activities of the partners were still ongoing in the expiring GFBio project.

Care was taken to leave the two core activities intact: The 23 use case projects, which serve to integrate a substantial number of new partners and resource types into a common infrastructure; and the cloud-based service platform called “research data commons” (RDC). The financial restrictions were added to the risk assessment (see section 3).

Additional thoughts:

- Flexfunds for accession of new use cases reduced to zero => reduced dynamics
- Struggled with finance plan, due to schematic cuts in the annual budgets => did not fit staff resource planning
- Cuts reduced the dynamic of the consortium to react to new developments

Other financial adjustments

(use material from [Executive Reports](#), risks assessment section)

- Adjustments regularly and structured each year (describe zentrale Mittelbewirtschaftung and re-introduction of flexfunds in year 2+3
 - => re-filled annually from underspend budgets, not plannable,
 - instrument is valuable needs more reliability (strong effects with little money)
- Redistribution of event budget in 2023

Appendix

The appendix may only include the following information and documents:

1 List of outputs produced by the consortium

Please list the different kinds of outputs (in the broad sense of the classes of the FRBR-aligned Bibliographic Ontology, FaBiO¹⁰) that you have produced and published briefly, but including at least the following data: title, year, and persistent identifier / web link.

O1. Weber, Judith; Ebert, Barbara; Diepenbroek, Michael; Kostadinov, Ivaylo; Glöckner, Frank Oliver (2021): NFDI4BioDiversity - NFDI-Konsortium für Biodiversitäts-, Ökologische und Umweltdaten. In: Bausteine Forschungsdatenmanagement Nr. 2 (Juli); <https://doi.org/10.17192/bfdm.2021.2.8334>

O2. Linares, Ebert, Engel (2023): E-science Tage (*in print*)

O3. NFDI4Biodiversity Mission statement: <https://nfdi4biodiversity.org/en/mission-statement/>

¹⁰ Cf. for the Classes <https://sparontologies.github.io/fabio/current/fabio.html#classes>.

- O4. NFDI Cross-cutting Topics Workshop Report. Ebert, Barbara; Fluck, Juliane; Glöckner, Frank Oliver; Koepler, Oliver; Miller, Bernhard; Schmitt, Robert; Schrader, Torsten; Stegle, Oliver; Steinbeck, Christoph; von Suchodoletz, Dirk; Wagemann, Kurt; Knebes, Jennifer; Kraft, Sophie; Seitz-Moskaliuk, Hendrik; Sure-Vetter, York; Wössner, Elena. Zenodo March 11, 2021, p . <https://doi.org/10.5281/zenodo.7825374>.
- O5. Sektionskonzept Common Infrastructures zur Einrichtung einer Sektion im Verein nationale Forschungsdateninfrastruktur (NFDI) e.v. Diepenbroek, Michael; Schimmler, Sonja; Ebert, Barbara. Zenodo October 28, 2021, p . <https://doi.org/10.5281/zenodo.5607490>.
- O6. White Paper: Interim Report Reference. Amelung, Lisa; Anthofer, Verena; Danabalan, Renita; Demandt, Évariste; Ebert, Barbara; Elschner, Elisabeth; Espinoza, Sara; Eufinger, Jan; Fuchsloch, Stefanie; Götz, Barbara; Henzen, Christin; Hunold, Johannes; Idda, Theresa; Jansen, Lukas; Krieger, Ulrich; Rodrigues, Cristina M.; Meister, Maria; Miller, Bernhard; Pitroff, Sarah; Popp, Christian; Pushkina, Aleksandra; Schatlowski, Nicole; Schimmler, Sonja; Schneide, Christiane; Schörner, Thomas; Schwetje, Thorsten; Seegert, Jörg; Trippel, Thorsten; Weber, Judith; Wössner, Elena; Zinke, Wolf.. Zenodo March 1, 2023, p . <https://doi.org/10.5281/zenodo.7688729>.
- O7. Bündnis der universitären und außeruniversitären Biodiversitätsforschung Deutschlands - Alliance of University and Non-University Biodiversity Research in Germany (2023): Submission of views on issues for further consideration for digital sequence information on genetic resources to the Convention on Biological Diversity; Reference SCBD/NPU/DC/KG/CGA/90785 (2023-003), online: <https://chm.cbd.int/database/record?documentID=264056>
- O8. FEDa Cooperation <https://nfdi4biodiversity.org/de/kooperation-fed/>
- O9. When Data Crosses Borders - Join Forces! Ebert, Domisch, Henzen et al, CoRDI Conference 2023 (*submitted*)
- O10. Connecting National and International Data Infrastructures in Biodiversity Research: The Case of NFDI4Biodiversity, a German Consortium for Biodiversity, Ecology and Environmental Data. Ebert et al. CoRDI Conference 2023 (*accepted*)
- O11. First strategic report on international involvement. Güntsch et al (2023), Zenodo (*in production*)
- O12. Second strategic report on international involvement. Güntsch et al (2023), Zenodo (*in production*)
- O.13 Data Science for Ecologists in R, 15 part series -. Gerber, Nina; Fitjer, Sören (2021) <https://www.youtube.com/playlist?list=PLE9wx7mxT8gsjR8lKEa43lk7VtXiYgPDy>, Accessed: 24.08.2022
- O14. Linear models 2 - Introduction to the Linear Model. Fitjer, Sören (2022), <https://www.youtube.com/watch?v=7W9zxcwRjig&list=PLE9wx7mxT8gsjR8lKEa43lk7VtXiYgPDy&index=13>, Accessed: 24.05.2023

- O15. Winter School Lectures, 6 part series on YouTube. Kostadinov, Ivaylo, Linares, Jimena, Kipker, Dennis-Kenji, Grenié, Matthias, Müller, Wolfgang, & Oellermann, Michael. (2022). <https://www.youtube.com/playlist?list=PL06Unzn1hDrhZ4mfsdfeypJajuGsuzdfm>, Accessed: 15.05.2023
- O16. Lectures of the NFDI4Biodiversity & GfÖ Winter School on Data Management in Ecology and Environmental Science, 4 parts. NFDI4Biodiversity (ed.) (2022). Zenodo [https://zenodo.org/communities/nfdi4biodiv/search?page=1&size=20&q=title:\(%22Legal%22\)%20OR%20\(keywords:\(%22winterschool2022%22%20OR%20%22WinterSchool2022%22%20OR%20%22Winterschool2022%22\)%20AND%20\(%22nfdi4biodiversity%22%20OR%20%22NFDI4Biodiversity%22\)\)](https://zenodo.org/communities/nfdi4biodiv/search?page=1&size=20&q=title:(%22Legal%22)%20OR%20(keywords:(%22winterschool2022%22%20OR%20%22WinterSchool2022%22%20OR%20%22Winterschool2022%22)%20AND%20(%22nfdi4biodiversity%22%20OR%20%22NFDI4Biodiversity%22)))
- O17. Template for the NFDI4Biodiversity & GfÖ Winter School on Data Management in Ecology and Environmental Science (*in preparation*)
- O18. [NFDI4Biodiversity Training and Education Report](#). Daniel Tschink et al. (*in preparation*).
- O19. Use Case Gesellschaft deutschsprachiger Odonatologen e.V. (GdO) <https://nfdi4biodiversity.org/de/was-wir-tun/use-case-gdo/>
- O20. Use Case IÖR Monitor, <https://nfdi4biodiversity.org/de/was-wir-tun/use-case-i%C3%B6r/>
- O21. Lebendiger Atlas Natur Deutschland (2023, im Aufbau) - <https://land.hp.gbif-staging.org/>
- O22. Citizen science plant observations encode global trait patterns. Wolf, S., Mahecha, M.D., Sabatini, F.M. et al. Nat Ecol Evol (2022). <https://doi.org/10.1038/s41559-022-01904-x>
- O23. Hydrographr R package (2023) - <https://github.com/flowabio/hydrographr>
Website: <https://flowabio.org/project/hydrographr/>
- O24. Towards a Research Data Commons in the German National Research Data Infrastructure NFDI : Vision, Governance, Architecture. Michael Diepenbroek et al. (2023), 1st International Conference on Research Data Infrastructures, Karlsruhe (*submitted*). https://drive.google.com/file/d/18XILCNPpKXZn3pPqdx32rhMYSZHqoCIK/view?usp=share_link
- O25. Aruna Core Storage Infrastructure. Universität Gießen (2022). Gitlab <https://github.com/ArunaStorage> , Website: <https://aruna-storage.org/>
- O26. The Aruna Object Storage - A distributed multi cloud object storage system for scientific data management. Marius Dieckmann, Sebastian Beyvers, Jannis Hochmuth, Anna Rehm, Frank Förster, Alexander Goesmann (2023), 1st International Conference on Research Data Infrastructures, Karlsruhe (*submitted*) <https://drive.google.com/file/d/1FAwUkypRjLMCEyhdY7gLQn2DDsmqcc7o/view?usp=sharing>
- O27. Platzhalter - Data portal
- O28.

- *Publikationsliste:* <https://kb.gfbio.org/pages/viewpage.action?pageId=33260334>
- *Outputs in Planung (evtl. unter Angabe der Finanzierung, gerade bei Service-Doks)*
 - *technische Dokumentationen RDC Komponenten (im Wiki/Knowledge base)*
 - *“best practice zum Nachlesen” aus den Use Cases (TA1)*
 - *peer-reviewed paper w. international “role of national infrastructures” (lead: Anton)*
 - *CoRDI Einreichung RDC + evtl. peer-reviewed paper in 2023/24*

2 Bibliography

Please list here the references cited throughout section B.2., except your own results, which appear in the section above.

- B1. Glöckner, Frank Oliver, Diepenbroek, Michael, Felden, Janine, Güntsch, Anton, Stoye, Jens, Overmann, Jörg, Wimmers, Klaus, Kostadinov, Ivaylo, Yahyapour, Ramin, Müller, Wolfgang, Scholz, Uwe, Triebel, Dagmar, Frenzel, Mark, Gemeinholzer, Birgit, Goesmann, Alexander, König-Ries, Birgitta, Bonn, Aletta, & Seeger, Bernhard. (2020). NFDI4BioDiversity - A Consortium for the National Research Data Infrastructure (NFDI). Zenodo. <https://doi.org/10.5281/zenodo.3943645>
- B2. Effective Biodiversity Monitoring Needs a Culture of Integration. Kühl, Hjalmar S.; Bowler, Diana E.; Bösch, Lukas; Bruelheide, Helge; Dauber, Jens; Eichenberg, David. et al. (2020): In: One Earth 3 (4), S. 462–474. DOI: 10.1016/j.oneear.2020.09.010.
- B3. Berlin Declaration on NFDI Cross-Cutting Topics (Version 1). Glöckner, Frank Oliver, Diepenbroek, Michael, Felden, Janine, Overmann, Jörg, Bonn, Aletta, Gemeinholzer, Birgit, Güntsch, Anton, König-Ries, Birgitta, Seeger, Bernhard, Pollex-Krüger, Annette, Fluck, Juliane, Pigeot, Iris, Kirsten Toralf, Mühlhaus, Timo, Wolf, Christof, Heinrich, Uwe, Steinbeck, Christoph, Koepler, Oliver, Stegle, Oliver, ... Bernard, Lars. (2019). Zenodo. <https://doi.org/10.5281/zenodo.3457213>
- B4. Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung. Maik Bierwirth, Frank Oliver Glöckner, Christian Grimm, Sonja Schimmler, Franziska Boehm, Christian Busse, Andreas Degkwitz, Oliver Koepler, & Heike Neuroth. (2020). Zenodo. <https://doi.org/10.5281/zenodo.3895209>
- B5. Stellungnahme der NFDI-Konsortien zu Basisdiensten. Konsortialversammlung des Vereins Nationale Forschungsdateninfrastruktur (NFDI) e.V. Zenodo February 15, 2022, p . <https://doi.org/10.5281/zenodo.6091657>
- B6. Data Mesh Principles and Logical Architecture. Dehghani Z (2020), <https://martinfowler.com/articles/data-mesh-principles.html> (2023-04-24)

- B7. Biotoools - Essential scientific and technical information about software tools, databases and services for bioinformatics and the life sciences (website). <https://bio.tools/>
- B8. GFBio Research Data Services (website/web applications) - <https://www.gfbio.org/services/>
- B9. Geo Engine Analysis platform (web application) - <https://www.geoengine.de/>
- B10. GFBio VAT - The Visualization, Analysis and Transformation System (web application) - <https://vat.gfbio.org/#/>
- B11. Diversity Workbench (= DWB) virtual research environment (database system) - https://diversityworkbench.net/Portal/Diversity_Workbench
- B12. Data exchange standards, protocols and formats relevant for the collection data domain within the GFBio network - website, https://gfbio.biowikifarm.net/wiki/Data_exchange_standards_protocols_and_formats_relevant_for_the_collection_data_domain_within_the_GFBio_network
- B13. TDWG Biodiversity Information Standards - www.tdwg.org
- B14. MixS - Minimal Information about any Sequence Standard <https://genomicsstandardsconsortium.github.io/mixs/introduction/background/>
- B15. MIAPPE - Minimum Information About Plant Phenotyping Experiments <https://www.miappe.org/>
- B16. ABCD for data publication (Wiki article) - https://gfbio.biowikifarm.net/wiki/ABCD_for_data_publication
- B17. <https://bioschemas.org/>
- B18. GFBio Terminology Service (website) - <https://www.gfbio.org/terminologyservice/>
- B19. Gfl Fish Atlas (web application) - <https://biodiv-atlas.de/fische/#!/home>
- B20. <https://www.f-uji.net/>
<https://op.europa.eu/en/publication-detail/-/publication/03b5562d-6a35-11ed-b14f-01aa75ed71a1/language-en>
- B21. GFBio Data Centers (website) - <https://www.gfbio.org/data-centers/>
- B22. GFBio Data Portal (website) - www.gfbio.org
- B23. FAIR Data Spaces project (website) - <https://www.nfdi.de/fair-data-spaces/?lang=en>
- B24. Alsayed Algergawy, Naouel Karam, Amir Laadhar, Franck Michel: Too big to match: a strategy around matching tasks for large taxonomies. Proceedings of the 17th International Workshop on Ontology Matching (OM 2022) co-located with the 21th International Semantic Web Conference (ISWC 2022), Hangzhou, China, October 23, 2022. CEUR Workshop Proceedings 3324, CEUR-WS.org 2023

- B25. Mina Abd Nikooie Pour, Alsayed Algergawy, Patrice Buche, Leyla Jael Castro, Jiaoyan Chen, Hang Dong, Omaira Fallatah, Daniel Faria, Irini Fundulaki, Sven Hertling, Yuan He, Ian Horrocks, Martin Huschka, Liliana Ibanescu, Ernesto Jiménez-Ruiz, Naouel Karam, Amir Laadhar, Patrick Lambrix, Huanyu Li, Ying Li, Franck Michel, Engy Nasr, Heiko Paulheim, Catia Pesquita, Tzanina Saveta, Pavel Shvaiko, Cássia Trojahn, Chantelle Verhey, Mingfang Wu, Beyza Yaman, Ondrej Zamazal, Lu Zhou: Results of the Ontology Alignment Evaluation Initiative 2022. Proceedings of the 17th International Workshop on Ontology Matching (OM 2022) co-located with the 21th International Semantic Web Conference (ISWC 2022), Hangzhou, China, October 23, 2022. CEUR Workshop Proceedings 3324, CEUR-WS.org 2023
- B26. Andreas Kohlbecker, Naouel Karam, Adrian Paschke, Anton Güntsch: Preserving Taxonomic Change and Subsequent Taxon Relationships over Time. Proceedings of the Joint Ontology Workshops (JOWO) 2021 Episode VII: The Bolzano Summer of Knowledge, Bolzano, Italy, September 11-18, 2021. CEUR Workshop Proceedings 2969, CEUR-WS.org 2021
- B27. FEdA - Forschungsinitiative zum Erhalt der Artenvielfalt - www.feda.bio
- B28. NMZB - Nationales Monitoringzentrum zur Biodiversität - <https://www.monitoringzentrum.de/>
- B29. umwelt.info -Portal für Umwelt- und Naturschutzinformationen - <https://www.umweltbundesamt.de/umwelt-info>
- B30. GAIA-X Ökosystem für Daten und Infrastruktur
https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/gaia-x/gaia-x_node.html
- B31. Richtlinie zur Förderung von Projekten zum Aufbau von Datenkompetenzzentren in der Wissenschaft, Bundesanzeiger vom 21.06.2022 - <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2022/06/2022-06-21-Bekanntmachung-Datenkompetenzzentren.html>
- B32. Aktionsprogramm natürlicher Klimaschutz
<https://www.bundesregierung.de/breg-de/aktuelles/natuerlicher-klimaschutz-2182120>