NFDI 4 BIODIVERSITY

www.nfdi4biodiversity.org
www.linkedin.com/company/nfdi4biodiv/

Expectations and realities of building a national research data infrastructure in Germany

Dr. Barbara Ebert

Hanover, February 2024

DATEN VIELFALT FÜR ARTEN VIELFALT

20.02.2024

Vision of NFDI

2023

Data as a common good for excellent research, organised by the scientific community

20.02.2024

Origin of research data

publicly funded for example ...

... at over 400 universities and technical colleges

... at the non-university research institutions and the scientific academies

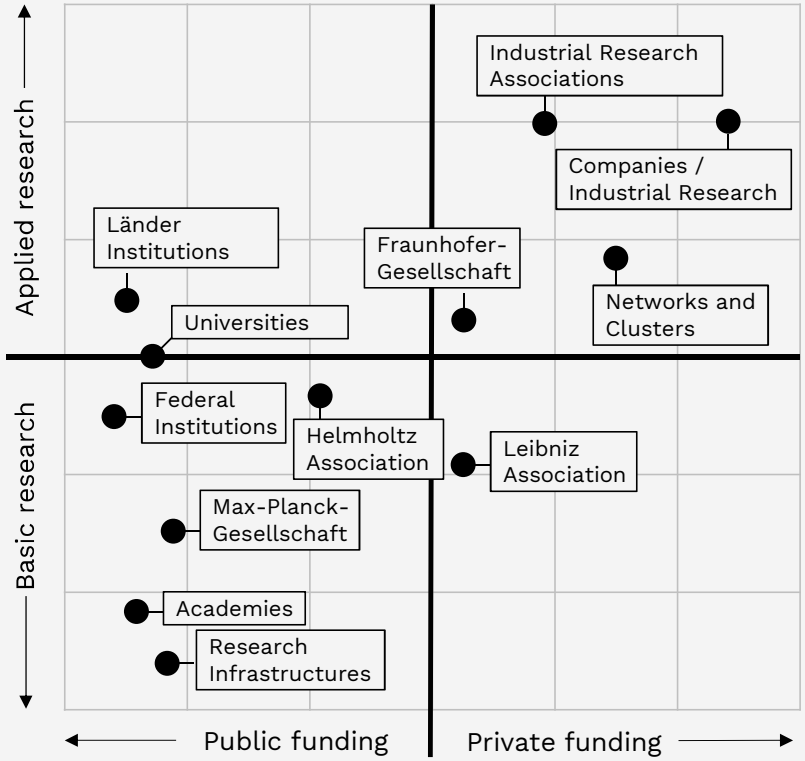
... in large, jointly funded research infrastructures

Area of activity of the Joint Science Conference of the federal government and federal states (Article 91b of the Basic Law)

20.02.2024

German R&D under the microscope

A self-ranked assessment of how public and private research organizations in Germany are funded and their research priorities.



https://www.nature.com/articles/d41586-019-00910-7
Source: Federal Ministry of Education and Research

Acknowledging the need for coordinated action

The rising tide of data – nearly as many digital bits as there are stars in the universe.

Source: IDC's Digital Universe Study 2014
https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.html

20.02.2024

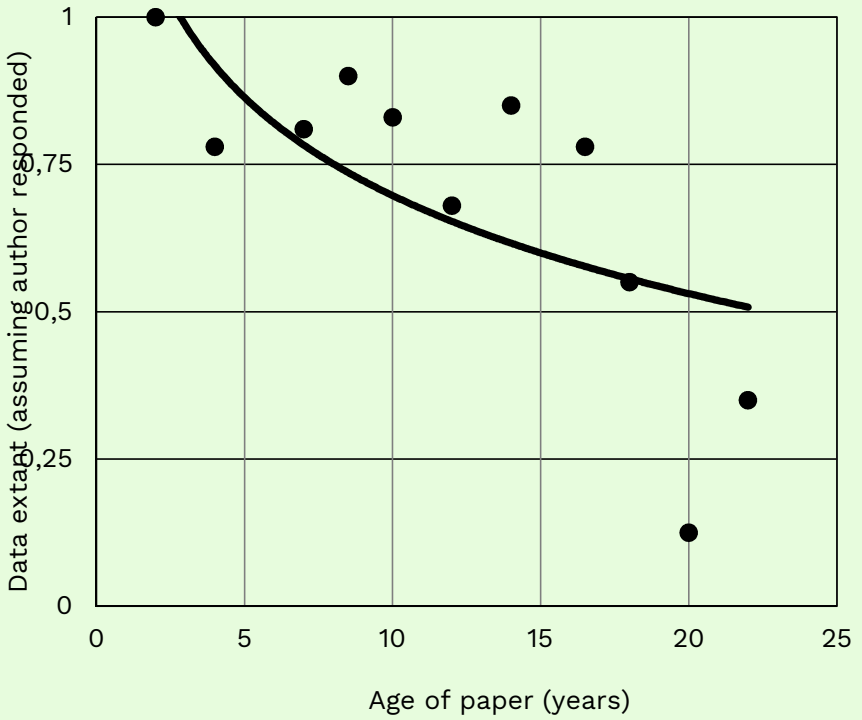
Availability of Research Data with Time

- Data being lost are estimated to increase by **17% in every** year after publication.
- Find a working e-mail address for the first, last, or corresponding author fell by **7% per year**.
- Overall, we only received 19.5% of the requested data sets, and only 11% for articles published before 2000.

20.02.2024

Missing data

As research articles age, the odds of their raw data being extant drop dramatically.



Vines, Timothy H. et al. Current Biology, 2014, Volume 24, Issue 1, 94-97

Make or buy?
The bigger picture

- Data as a special form of scientific knowledge: Several large scientific publishers are expanding their spectrum
  - but: the trust is currently impaired
- Data and scientific methods are closely entwined. There are many arguments in favour of self-organised solutions

20.02.2024

Accepted community-driven federated initiatives in DE already exist

Three examples	
Since 2003: The German Astrophysical Virtual Observatory (GAVO) connects German research projects to the global VO	Funding ended 2017 now part of NFDI
Since 2004: The German Data Forum RatSWD develops a network of research data centres at public authorities and scientific institutions	Funding ended 2020 now part of NFDI
Since 2013: German Federation for Biological Data forms a network to host data from research projects on biodiversity	Funding ended 2021 now part of NFDI

Setup of NFDI

- Joint funding by Federal government and the *Länder* (§91b GG)
- Founding of an association "NFDI e.V", to organise the data landscape
- Funding of up to 30 domain-related "consortia" (science-driven process)
- Scientific institutions join the NFDI association and support its goals

Source: DFG-Förderatlas 2021, Fig. 2.7

20.02.2024

Facts and figures

Nationale Forschungsdaten Infrastruktur

Initial funding period 2019-2028, 90 Mio EUR per year		
27 **Funded consortia** broadly covering research domains in Germany + basic services	**Several hundred organisations teaming up as suppliers and in decision processes** **270** organisations with formal membership in the NFDI association	**> 1,000** **Professional staff**

Consortia organise collaborative and domain-specific data infrastructures

Nationale Forschungsdaten Infrastruktur

Life Sciences	Humanities & Social Sciences	Natural sciences	Engineering Sciences
Plant, Agrosystem, Immunology, Human Genome-Phenome, Biodiversity, Ecology & Environment, Bioimaging, Health, Microbiota	Business & Economic, Social, Educational Behavioural Science, Cultural Heritage, History, Language & Text	Physics: Photon/Neutron, Condensed Matter, Chemical physics; Astronomy Catalysis, Mathematics, Chemistry, Earth System Science	Data Science & AI, Energy Systems, Engineering, Material Science, Computer Science

Work on cross-cutting topics

Section-metadata — **(Meta)data, Terminologies, Provenance**

Section-infra — **Common Infrastructures**

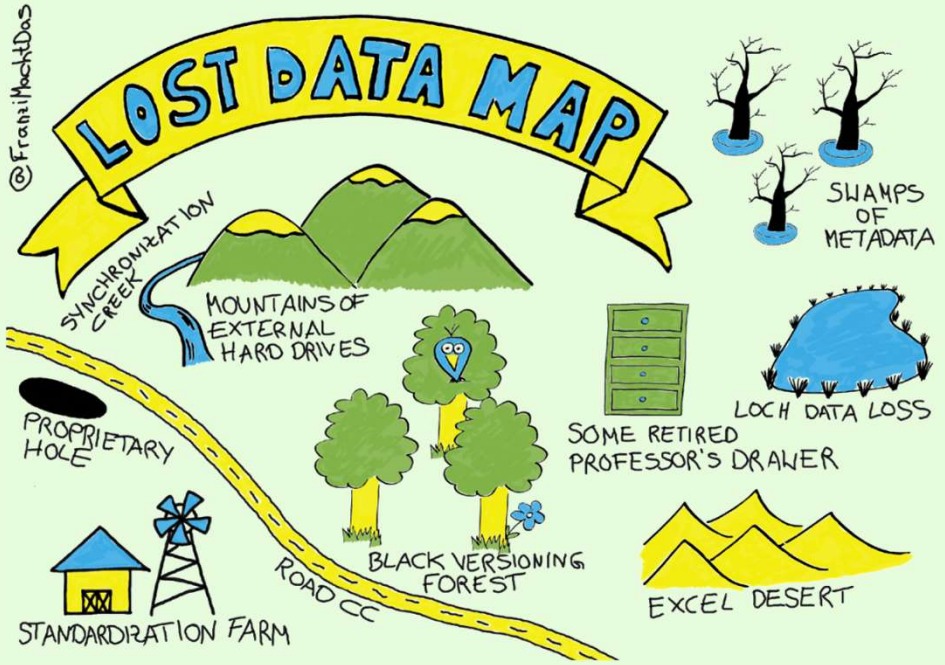
Section-edutrain — **Training & Education**

Section-ELSA — **Ethical, Legal and Social Aspects**

Section-industry — **Industry Engagement**

Our Goal: Making Research Data FAIR

F	FINDABLE
A	ACCESSIBLE
I	INTEROPERABLE
R	REUSABLE



Source: F. Mau (2019)
DOI 10.5281/zenodo.4388671

20.02.2024

Okay, but...

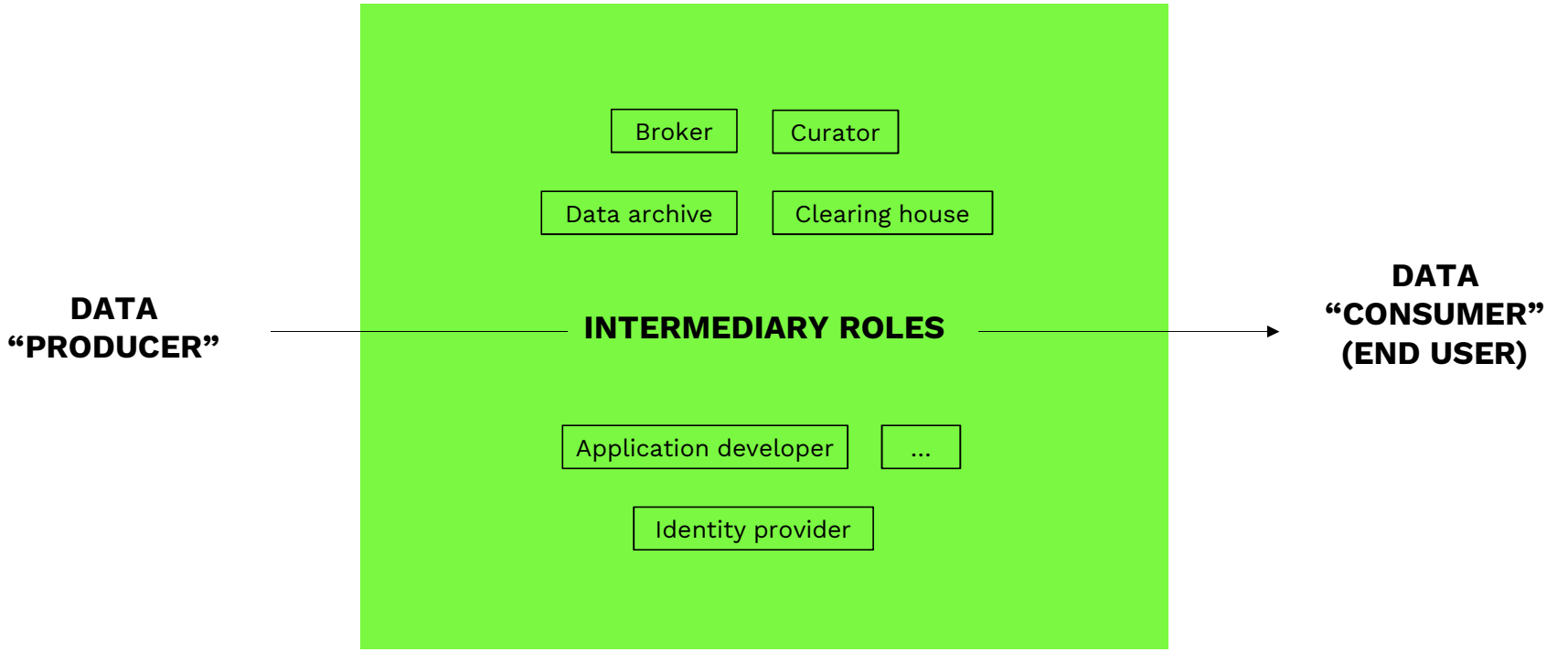
World wildlife Photography awards

20.02.2024

What does NFDI do, essentially?

In NFDI we organise
the roles between data producers and
data consumers in science

Organising intermediary roles in NFDI



Rolemodel adapted from https://internationaldataspaces.org/use/reference-architecture/

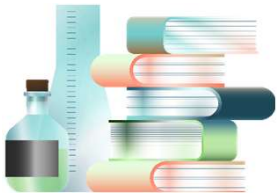
NFDI 4
BIODIVERSITY
BIODIVERSITY, ECOLOGY & ENVIRONMENTAL DATA

How a community can organise in the long term

20.02.2024


Our Mission
Providing reliable FAIR data to enable better science, decision making and conservation of biodiversity

Our consortium



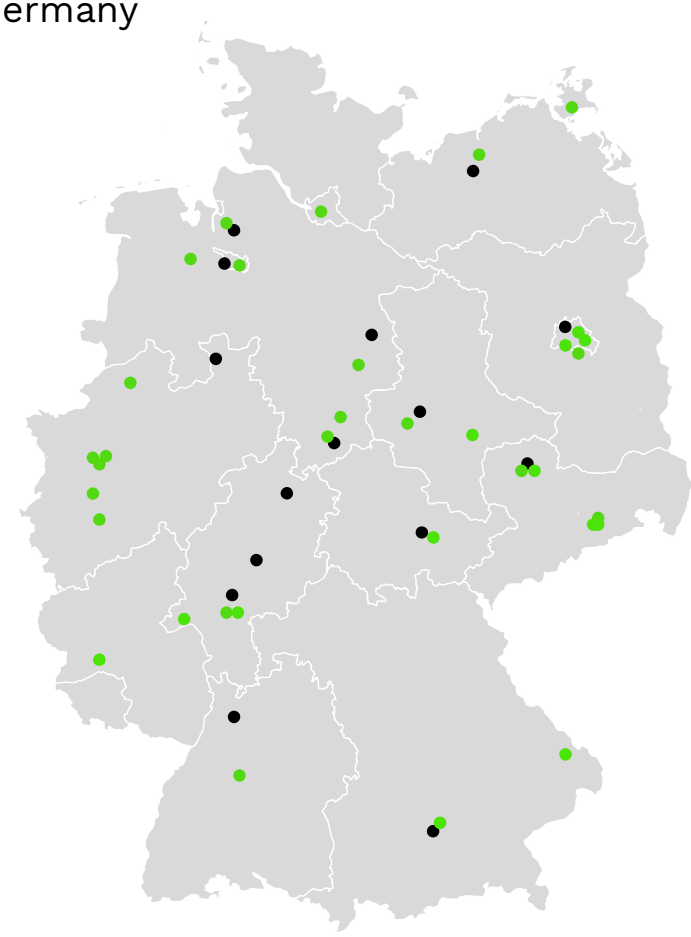
Research groups and institutes **17**

Expert and citizen science, government authorities **14**

Natural science collections **9**

Scientific IT centres **11**

51 **Co-applicants** and **Participants** across Germany



Our roots

German Federation for Biological Data

Established 2013 to facilitate data sharing for biological and environmental research

Funded by DFG - German Research Foundation, project number 229241684 and 408180549.

20.02.2024

Our Brains

Foto: Anne Barth, All Hands Conference 2022

20.02.2024

Portfolio Development & Governance

20.02.2024

Our solutions-based approach

Use cases set the priorities
for
Service portfolio development
and
implementation of the **Research Data Commons concept**

"Leistung aus Vielfalt" (RfII 2016)

Managing a diverse method- and data-oriented partner group driven by a mission	
	Strong coordination and staff pooled at the speaker organisation and GFBio e.V.
	Flexfunds to explore opportunities and support improvements
	Productive relations with national and international initiatives
	Rich portfolio of services (initially **80 services** provided by 23 partners, and rising)

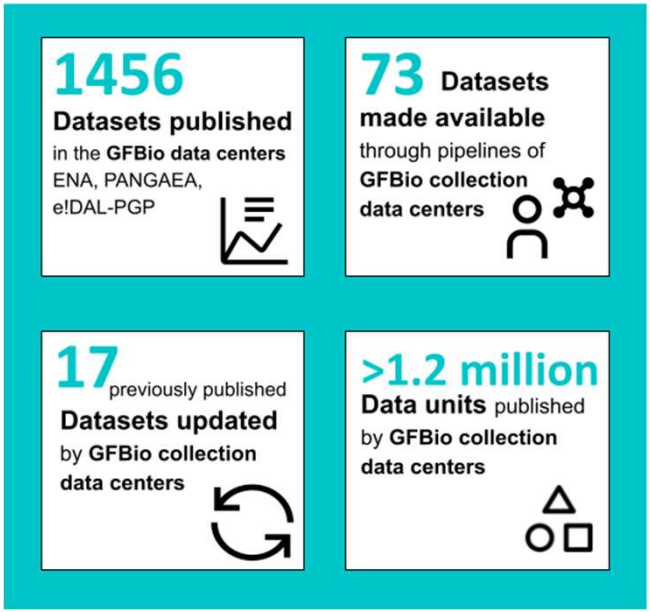
Benefits for the Community

20.02.2024

Maximilian Koppernock

1. Making data FAIR

Curated, archived and published data sets from our distributed archiving network



BioCASE

GBIF
Global Biodiversity
Information Facility

Biodiversity
Information
Standards
TDWG

gfbio
GERMAN FEDERATION
FOR BIOLOGICAL DATA

All data are findable under www.search.gfbio.org

20.02.2024

Which stakeholders are mobilising data with us?

Postdoc Paul

(and other individual researchers)

Nationalpark Bayerischer Wald

DeutschlandFlora

Arachnologische Gesellschaft

Institut für Ostseeforschung Warnemünde

Dachverband dt. Avifaunisten

Artenfinder

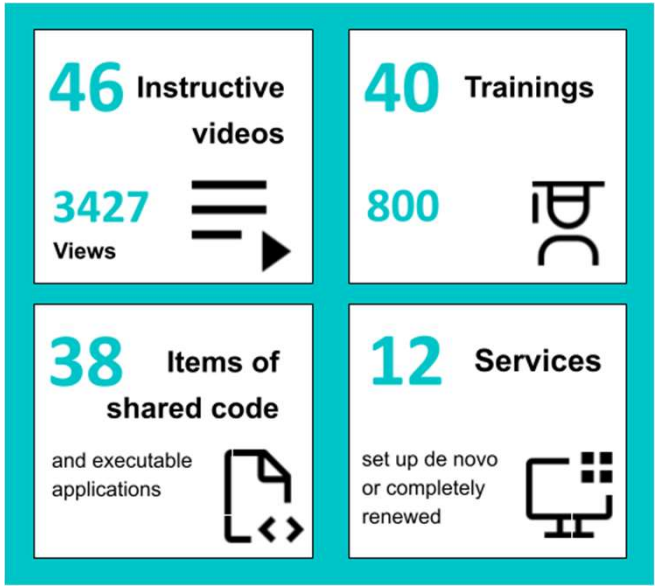
Staatliche Archive Bayerns

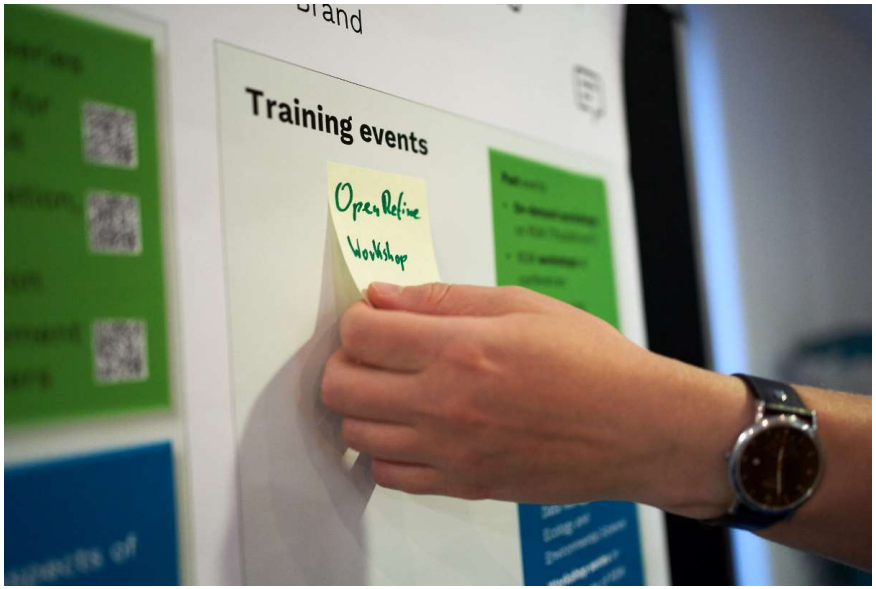
Rote Liste Zentrum

https://nfdi4biodiversity.org/en/use-cases-overview/

20.02.2024

2. Producing tangible community resources





Available through our website www.nfdi4biodiversity.org

3. Building interoperable services

Steps towards a Research Data Commons

- Key components implemented in the Cloud
- Ready for pilot applications to be connected to the RDC architecture

Dai:si Data Set Search

Metadata harvester

gfbio e.V.

BiodivPortal

https://kb.gfbio.org/display/KB/BiodivPortal

ARUNA

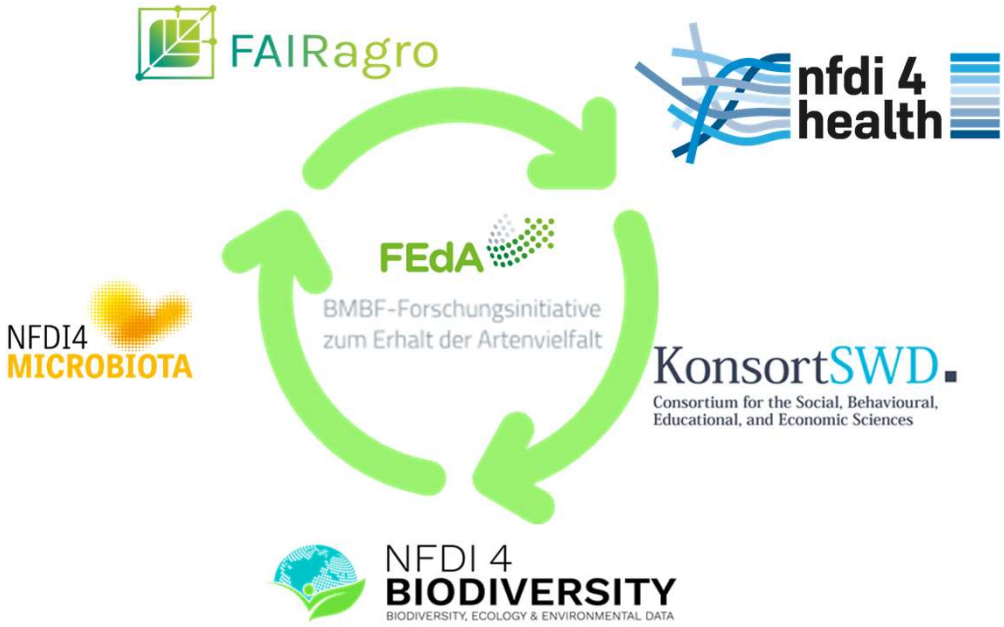
Your distributed scientific data mesh

https://kb.gfbio.org/display/KB/Aruna+Object+Storage

4. Serving interdisciplinary research

The case of FEdA

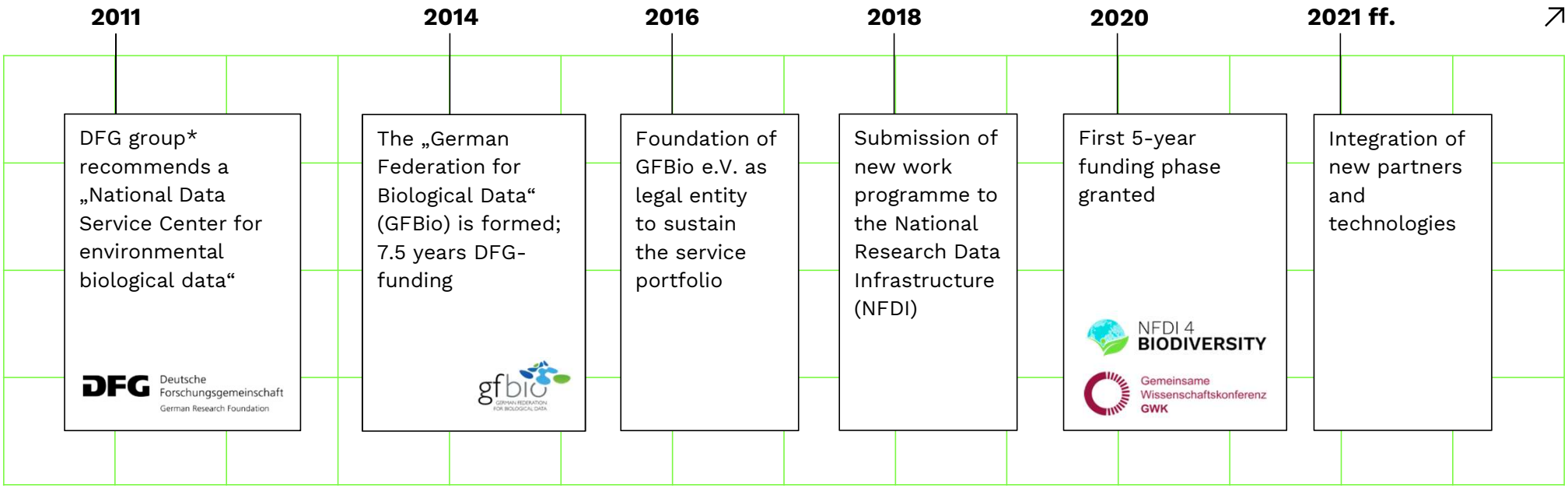
- BMBF funding initiative for biodiversity conservation, started 2020
- several cohorts of socio-ecological research projects with interdisciplinary research methods
- RDM support organised across NFDI-consortia (five consortia on board since December 2023, for cohorts 1-3)




Reflections
20.02.2024

Twelve years of organising joint services for environmental & biological data in Germany

*AG Biodiversitätsdaten der Senatskommission für Biodiversitätsforschung

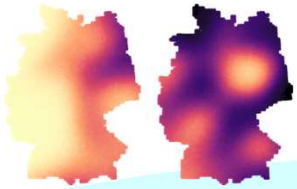


20.02.2024

Ultimate goal: Data serves knowledge production

occurrence data

research institutes, associations, foundations, citizen science



distributions

state & federal agencies

ecosystem data

national parks

experiments

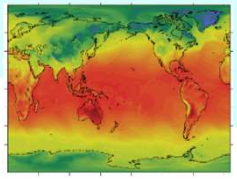
research institutes

NFDI 4 BIODIVERSITY

BIODIVERSITY, ECOLOGY & ENVIRONMENTAL DATA

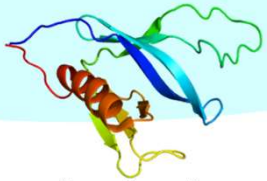
satellite data

research institutes & agencies



model output

research institutes



molecular data

research institutes

collection data

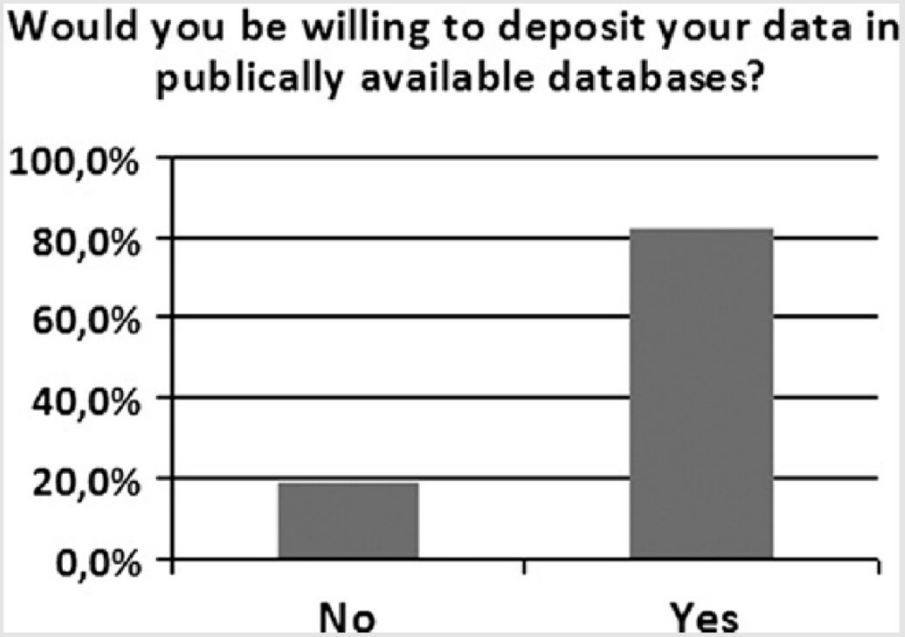
museums & collections

functional traits

research institutes

Adapted from Christian Wirth, iDiv

Data availability is key – groundwork by DFG Senate Commission on Biodiversity Research WG (2010 ff.)



Contents lists available at SciVerse ScienceDirect
Ecological Informatics
journal homepage: www.elsevier.com/locate/ecolinf

A comparative evaluation of technical solutions for long-term data repositories in integrative biodiversity research

Kerstin Bach [a,*], Daniel Schäfer [b], Neela Enke [c], Bernhard Seeger [b], Birgit Gemeinholzer [d], Jörg Bendix [a]

Doi:10.1016/j.ecoinf.2011.11.008

Contents lists available at SciVerse ScienceDirect
Ecological Informatics
journal homepage: www.elsevier.com/locate/ecolinf

The user's view on biodiversity data sharing — Investigating facts of acceptance and requirements to realize a sustainable use of research data —

Neela Enke [a,*], Anne Thessen [b], Kerstin Bach [c], Jörg Bendix [c], Bernhard Seeger [d], Birgit Gemeinholzer [e,*]

[a] Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Str. 6–8, 14195 Berlin, Germany
[b] Center for Library and Informatics, Marine Biological Laboratory, 7MBL Street, Woods Hole, MA 02556 USA
[c] Department of Geography, Philipps-University Marburg, Deutschhausstr. 10, 35032 Marburg, Germany
[d] Department of Mathematics and Computer Science, Philipps-University Marburg Hans-Meerwein-Str., 35032 Marburg, Germany
[e] Department of Systematic Botany, Justus Liebig University, Heinrich-Buff-Ring 38, 35392 Giessen; Germany

doi: 10.1016/j.ecoinf.2012.03.004

Important step: becoming part of the NFDI

27 Consortia

> 260 member organisations

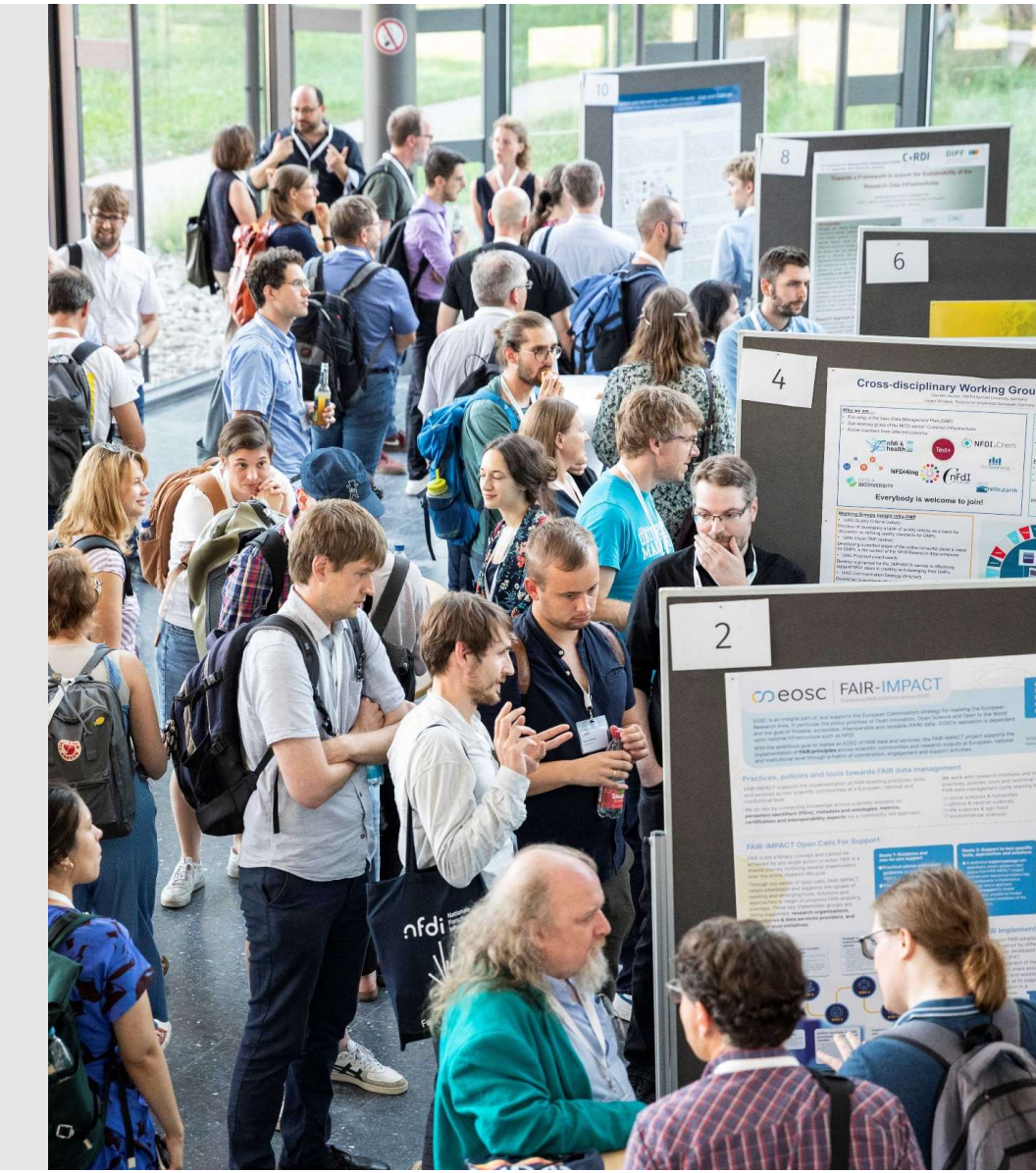
www.nfdi.de

NFDI e.V. Association - Directorate -				
Data Plant	GHGA	NFDI4Biodiversity	NFDI4Cat	NFDI4Chem
NFDI4Culture	NFDI4Health	NFDI4Ing	KonsortSWD	BERD@NFDI
DAPHNE4NFDI	FAIRmat	MaRDI	NFDI4DataScience	NFDI4Earth
NFDI4Microbiota	NFDI4MatWerk	PUNCH4NFDI	Text+	NFDI4Memory
NFDI4Objects	NFDI4BIOIMAGE	NFDI4Energy	NFDI4Immuno	FAIRAgro
NFDIxCS	Base4NFDI			

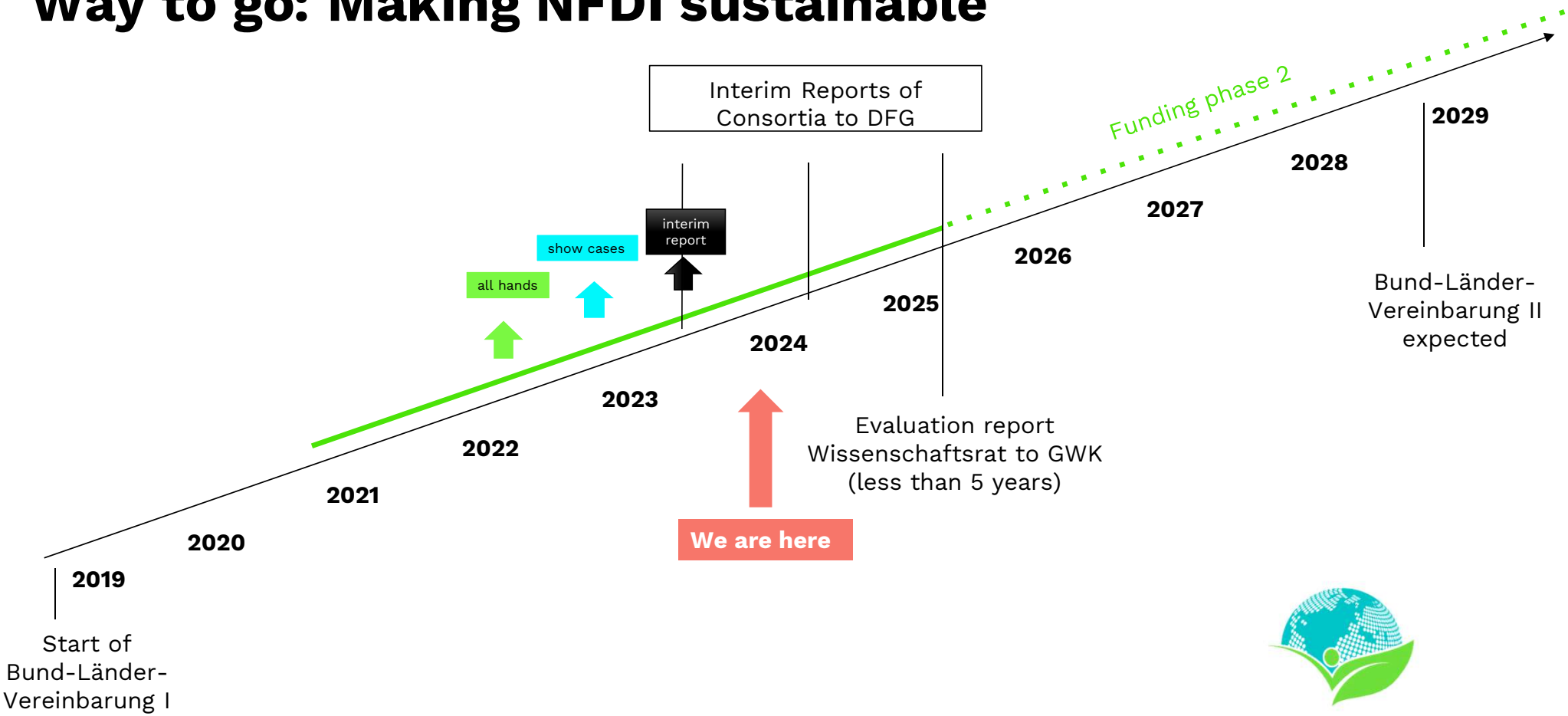
What can we say after three years of NFDI?

- NFDI has helped to **stabilise** our core set of services and **grow the network of motivated data providers**
- Collaborating **with other NFDI consortia** has helped enormously to cover community needs
- Consolidation of services is a **long-term issue** and we are glad that some needs can be tackled across consortia
- We **underestimated the communication efforts** necessary to a common undestanding, both with domain stakeholders and fellow consortia

20.02.2024



Way to go: Making NFDI sustainable



20.02.2024

NFDI as the ticket to sustainability

- Helping with concepts and white papers
  - In NFDI sections
  - In the consortia assembly
  - In task forces and circles
- Helping to make added value of NFDI activities known
  - In meetings with executive staff of authorities, institutions etc.
  - At scientific conferences and in events with scientists

Being part of NFDI can be quite overwhelming

Don´t get disheartened

20.02.2024

Motivation

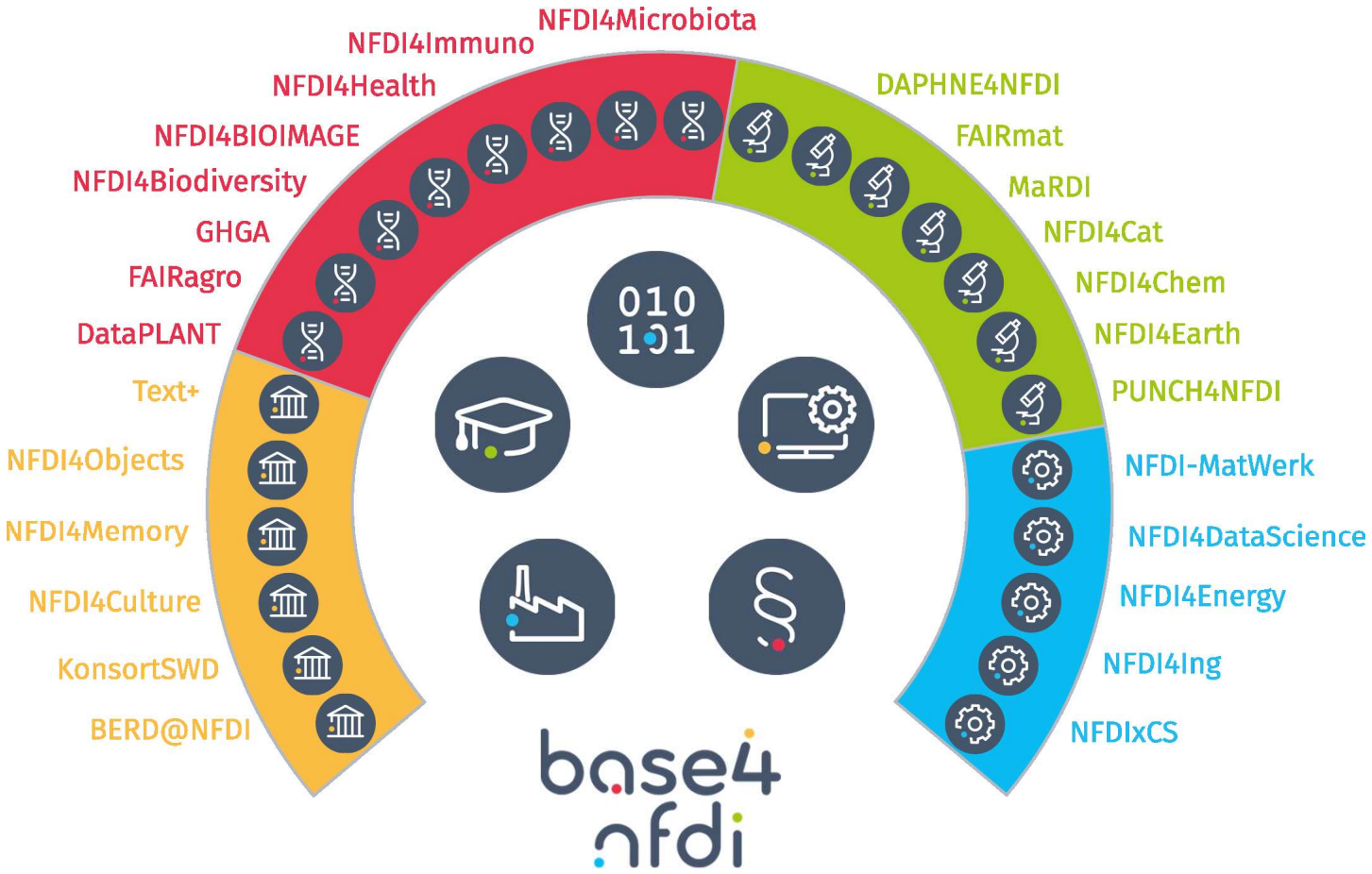
„Data as a common good for excellent research, organised by the scientific community“

Vision NFDI

Part of ONE NFDI

NFDI 4 BIODIVERSITY

Funded by the DFG, project number 442032008



20.02.2024

Bonus material

20.02.2024

Nationalpark Bayerischer Wald

NFDI 4 BIODIVERSITY

Funded by the DFG, project number 442032008

20.02.2024

The road towards NFDI

2011-2015:
Structured exploration
Basic recommendations on how to adapt the organisation of scientific information infrastructures in the digital turn

2016-2019
Realisation phase
Joint Science Minister Conference adapts RfII recommendation; DFG is tasked with the funding process
https://www.gwk-bonn.de/NFDI.pdf

KII 2011

WR 2012

RfII 2014

2016

GWK 2016-2018

DFG 2019

2020

2021

2022

2028 ff.

Operational phase of 10 years
3 rounds of funding decisions
Founding of the NFDI e.V.

Parallel initiatives on the international level

20.02.2024

Our Community

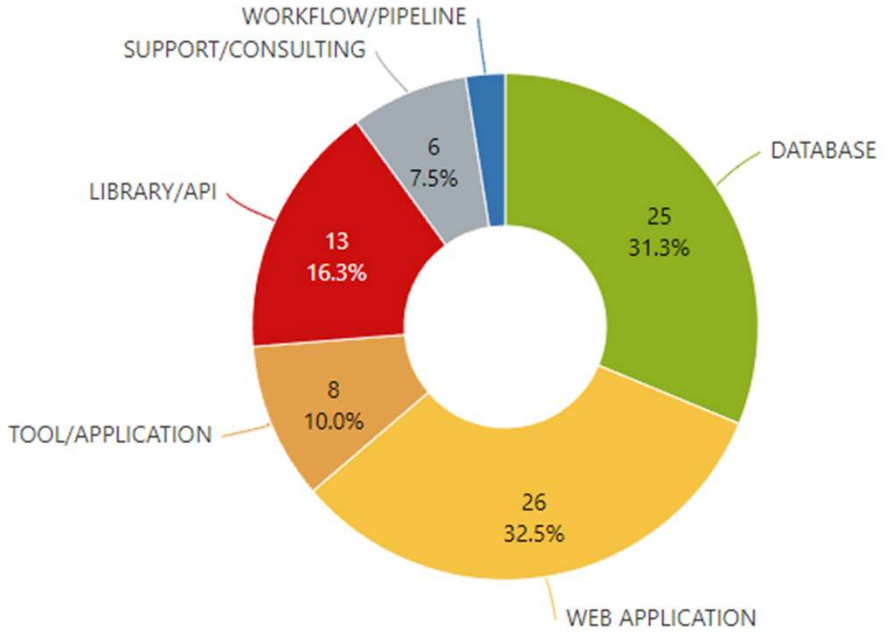
Stakeholder networks who generate and use biodiversity data to address global challenges

Science	↗ Improving knowledge about biodiversity changes and losses
Citizen Science	↗ Recording regional species information
Natural science collections	↗ Preserving knowledge and biodiversity
Authorities	↗ Monitoring, measures and policies
Politics	↗ Framing the National Biodiversity Strategy as part of the Global Biodiversity Framework

20.02.2024

Portfolio of services in the consortium

Over 80 databases, web applications, API etc. available through partners



Source: Service Portfolio Report, doi: 10.5281/zenodo.8327366

Our vision: A Research Data Commons