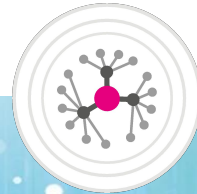




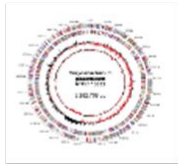
de.NBI & ELIXIR Germany - eine Forschungs- und Serviceinfrastruktur für Bioinformatik in den Lebenswissenschaften



Dr. Nils Hoffmann
de.NBI / ELIXIR-DE
IBG-5
Forschungszentrum Jülich
Außenstelle Bielefeld
c/o Centrum für Biotechnologie
Universität Bielefeld

12 April 2024

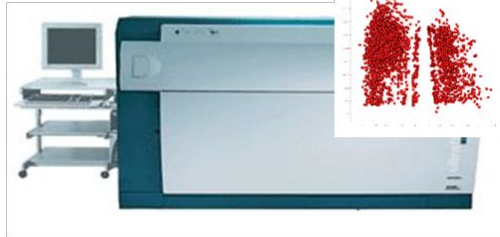
Genomics



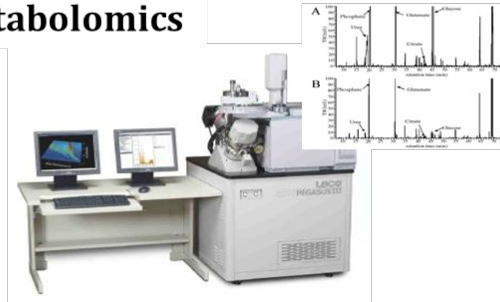
Transcriptomics



Proteomics



Metabolomics



- Life sciences use increasingly complex analytical methods that generate large amounts of data to be evaluated
- This requires extensive knowledge, skills and powerful bioinformatics tools and standards
- de.NBI contributes to solving these challenges by providing tools, services, consulting and training

Big data sets in life sciences are mainly produced by omics technologies

THE GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE – de.NBI



THEMATIC FOCUSES & SERVICE CENTERS:

- HUMAN BIOINFORMATICS
HEIDELBERG CENTER FOR HUMAN BIOINFORMATICS (HD-HuB)
- MICROBIAL BIOINFORMATICS
BIELEFELD-DIESSEN RESOURCE CENTER FOR MICROBIAL BIOINFORMATICS (BIOI)
- PLANT BIOINFORMATICS
GERMAN CROP BIOGENEFORMATICS NETWORK (GCBM)
- RNA BIOINFORMATICS
RNA BIOINFORMATICS CENTER (RBC)
- PROTEOMICS BIOINFORMATICS FOR PROTEOMICS (BioInfra_Prot)
- INTEGRATIVE BIOINFORMATICS CENTER FOR INTEGRATIVE BIOINFORMATICS (CIBI)
- BIODATABASES
CENTER FOR BIOLOGICAL DATA (BioData)
- DATA MANAGEMENT/SYSTEMS BIOLOGY
de.NBI SYSTEMS BIOLOGY SERVICE CENTER (de.NBI-SysBio)

- LOCATIONS OF SERVICE CENTERS
- LOCATIONS OF PARTNERS



SERVICE

- TOOLS
- WORKFLOWS
- DATABASES
- CONSULTING



TRAINING

- TRAINING COURSES
- SUMMER SCHOOLS
- HACKATHONS
- WEBINARS



de.NBI CLOUD

- INFRASTRUCTURE
- PLATFORM AND SOFTWARE AS A SERVICE



INDUSTRIAL FORUM

- INDUSTRY SERVICES
- CONSULTING
- NETWORKING



USER MEETINGS

- CONTINUOUS DEVELOPMENT OF TOOLS
- CONSULTING
- EXCHANGE OF OPINIONS AND EXPECTATIONS
- BOTTOM-UP FEEDBACK

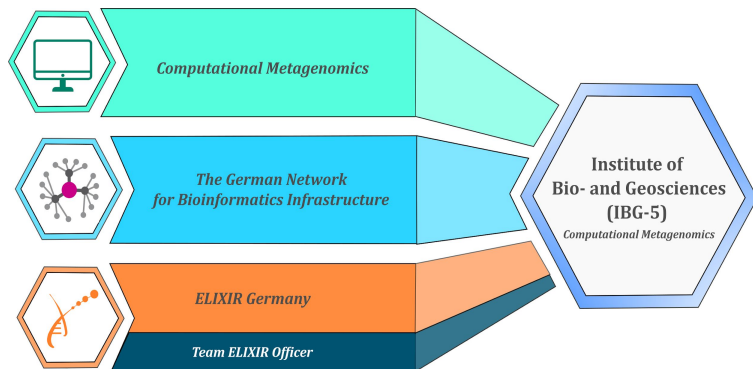
→ With 24 partners, organized in 8 topic-specific service centers, de.NBI **provides comprehensive bioinformatics services** to users in life sciences research from academia, industry and medicine.

→ de.NBI **fosters the cooperation** of the German bioinformatics community with European and international bioinformatics networks.

→ Founded in 2015, funded until Dec. 2021 by BMBF

→ Since 2022, funded via Helmholtz Society & consolidation entrusted to Forschungszentrum Jülich

- Since 2022, **Forschungszentrum Jülich GmbH** as a member of the Helmholtz Association has been entrusted with the continuation of **de.NBI** and **ELIXIR Germany**
- The establishment of the new **Institute for Bio- and Geosciences (IBG-5)** with the name "Computational metagenomics", located at the Bielefeld branch office



Prof. Dr. Alexander Sczyrba
Director of IBG-5
Head of the Computational metagenomics group
HoN ELIXIR DE (shared with O. Kohlbacher)



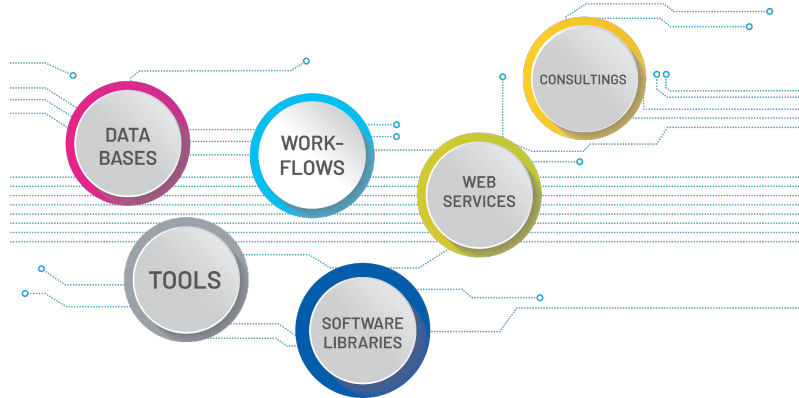
Dr. Tanja Dammann-Kalinowski
Head of the de.NBI administration office



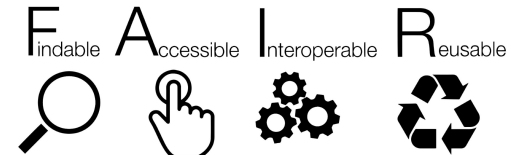
Dr. Irena Maus
Head of the ELIXIR-DE administration office



de.NBI offers > 180 bioinformatics tools and services for the analysis and evaluation of biological data sets



de.NBI Services are committed to apply FAIR data principles



→ The majority referenced, searchable, and findable in the ELIXIR **bio.tools** registry

<https://bio.tools>

<https://www.denbi.de/services>



Training

- de.NBI organizes training events and provides online training materials for a **broad range of topics in bioinformatics**.
- Life scientists learn how to handle and analyze biological big data by applying **tools, standards, and compute services**.
- Current developments in the field of bioinformatics are also addressed in **de.NBI symposia, special workshops** and annual **summer schools**.

Since 2015

- >300 trainers, researchers, bioinformaticians, software developers
- >450 training courses
- >10,000 participants



Examples from de.NBI's training portfolio

- ▣ de.NBI - CeBiTec NANOPORE WORKSHOP (ONLINE) Best Practice and SARS-CoV-2 Applications
- ▣ INTRODUCTION TO THE de.NBI CLOUD (ONLINE) de.NBI Cloud User Meeting
- ▣ 1ST DE.NBI SUMMER SCHOOL 2021 - Analysis and integration of Mass Spectrometry based omics data in Proteomics, Metabolomics and Lipidomics
- ▣ INTRODUCTION TO RNAseq DATA ANALYSIS RNAseq and High Throughput Omics Analysis for Plants
- ▣ ADVANCED METHODS FOR DIFFERENTIAL ANALYSIS OF PROTEOMICS DATA Analysis of Quantitative Proteomics Data Using R
- ▣ BASIC BIOINFORMATICS TRAINING FOR BIOLOGISTS Analysis, Visualization, and Integration of Multi-Level Omics Data
- ▣ ANALYZING SCIENTIFIC DATA USING MACHINE LEARNING ALGORITHMS Using Galaxy

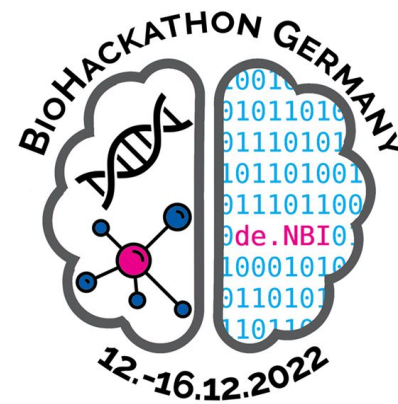
Selected training highlights in 2022/2023

Spring School

[2023] Data management

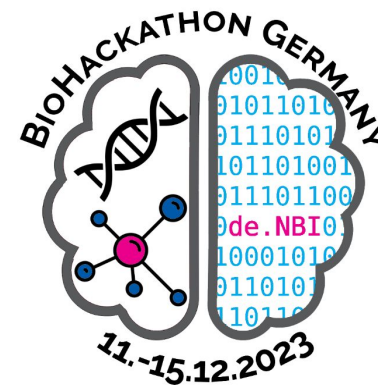


21st European Conference
on Computational Biology
Planetary Health and Biodiversity



Summer School

[2023] Bio Data Analysis
for Life Sciences



German Conference
on Bioinformatics 2022

September 6-8 2022, Halle (Saale)



GCB 2023 Hamburg | 12-14 September



The **de.NBI Cloud** is a fully academic cloud federation providing **compute** and **storage resources** for academia in Germany.

→ Access to fully configured infrastructure for running tools and pipelines in virtual machines



→ Single sign-on via Life Science Login

→ Annual Cloud User Meeting



→ Free of charge for German academia

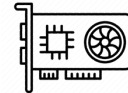
<https://www.denbi.de/cloud>



~70,000
Cores



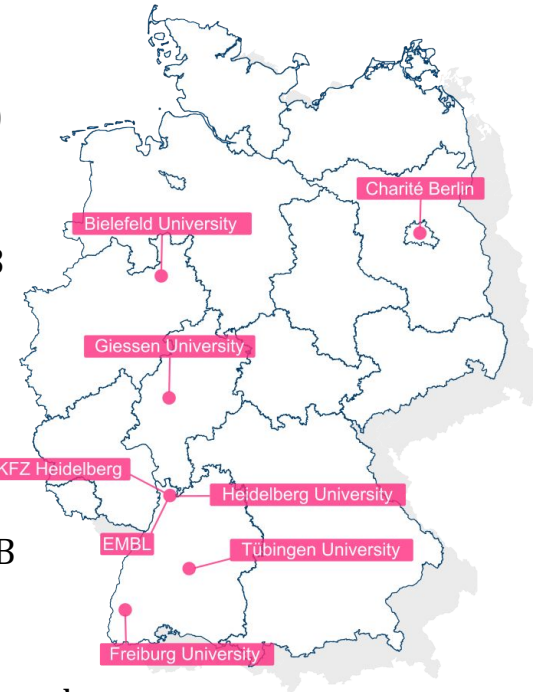
~100 PB
Storage



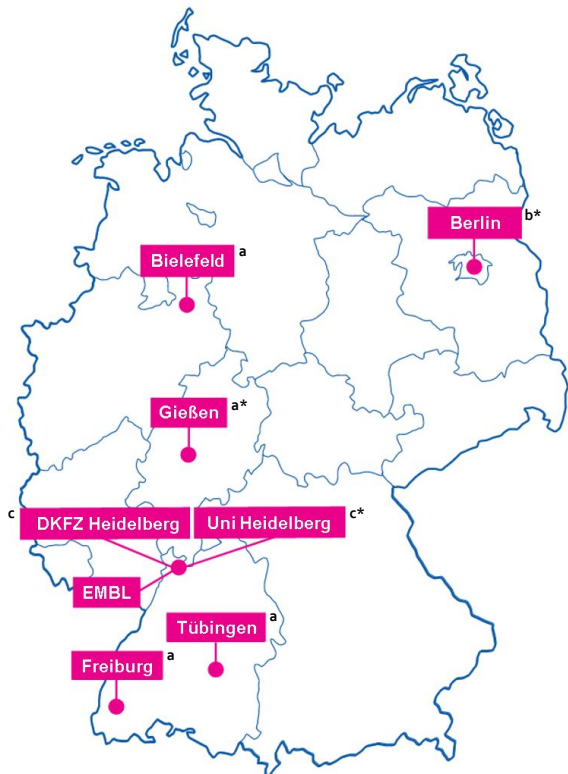
~500
GPU



up to 4 TB
RAM



- > 3000 registered users
- > 1,000 projects
- > 850 publications referencing de.NBI Cloud



DIN EN ISO / IEC 27001 : 2017

Certified Information Security
Management System (ISMS)

Defines technical and organisational measures
and controls for risk assessment and mitigation

- ▶ de.NBI Cloud @ Bielefeld University
- ▶ de.NBI Cloud @ Freiburg University
- ▶ de.NBI Cloud @ Tübingen University
- ▶ de.NBI Cloud @ Giessen University *

BSI IT - Grundschutz

Certified ISMS and Minimum
Required Implementation

Defines **mandatory** technical and
organisational measures and controls for risk
assessment and mitigation

- ▶ BIH / Charite Berlin *
- ▶ DKFZ Heidelberg (TÜV IT Trusted Site
Infrastructure)
- ▶ Heidelberg University (TÜV IT TSI) *

^aISO27001

^bBSI IT Grundschutz

^cTÜV IT TSI

* preparation for certification ongoing

Project application workflow



 LS LOGIN



- **principal investigator of a German university or research institution** applies for cloud resources by proposing a project and describing required resources **through the de.NBI Portal**
- the project is **reviewed by a scientific committee**
- after approval of the application the project is created in the **de.NBI Cloud Portal**
- project resources are **allocated at one of the cloud sites**

LIFE SCIENCE RI
Perun

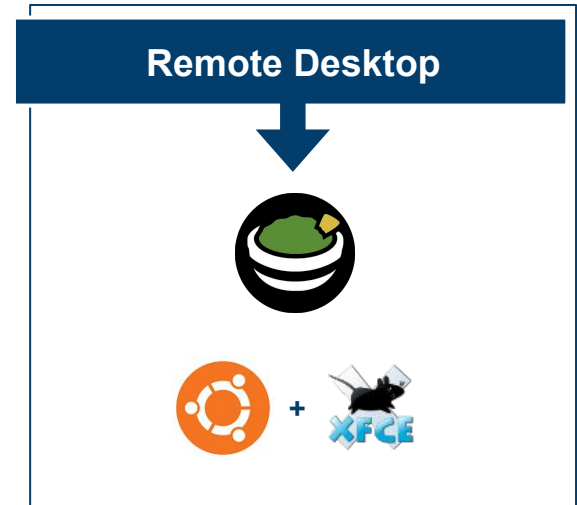
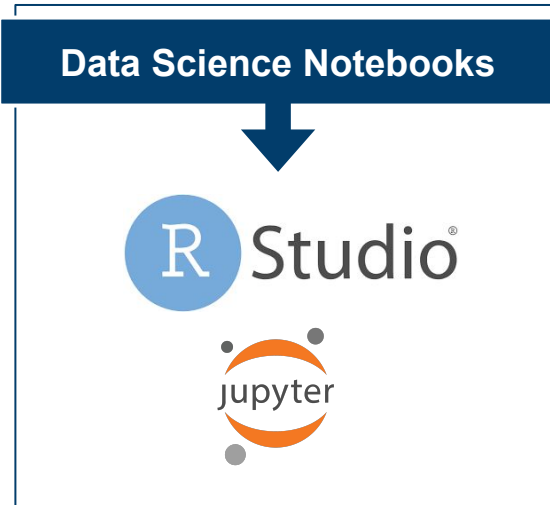


SimpleVM - Simplified Access to Virtual Machines

Self Service VMs and Research Environments



- Beginner-friendly, preconfigured & custom research environments
- Integrated with LifeScience AAI for authentication and authorization
- “One-click” solution for setting up your own cluster
- Browser and SSH access



▶ based on OpenStack, Ansible and Packer, recipes at <https://github.com/deNBI/resenvs>



Cluster

- ▶ Workshop module to define VMs using base images, virtual research environments, or snapshots
- ▶ Add participants with an existing Life Science AAI account to your project
- ▶ Send VM access details to each participant individually
- ▶ Launch VM instances for your participants



Workshop

- ▶ “One-click” solution for starting more complex cluster setups, e.g. for Nextflow
- ▶ Connect multiple machines to one SimpleVM Cluster
- ▶ Job scheduling mechanism (SLURM)
- ▶ Easy automatic scale-up and scale-down
- ▶ Currently in beta development, available on request



SimpleVM

- SimpleVM currently in evaluation for Bielefeld University Data Science Infrastructure Cloud
- For initial projects, please contact Oliver Francke or Matthias Vangerow

IDEs

 THEIA



Visual Studio Code

Data Science Notebooks

 R Studio®



Remote Desktop



- ▶ based on OpenStack, Ansible and Packer, recipes at <https://github.com/deNBI/resenvs>

Training and Community Building

47 Participants



CLOUD
USER
MEETING
BIELEFELD

2019

2018



CLOUD
USER
MEETING²
HEIDELBERG

60 Participants

75 Participants



CLOUD
USER
MEETING
ONLINE 2020

2021

2020



4th CLOUD
USER
MEETING 21
ONLINE

67 Participants

20 Participants



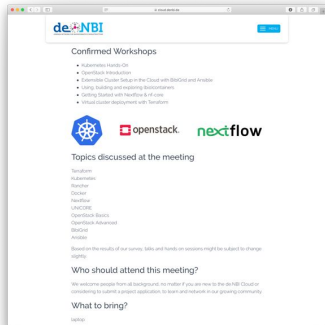
5th
Cloud
User
Meeting 20
22
Jülich

2022

2023



CLOUD
USER
MEETING
23
BIELEFELD



Courses:

Beginners

- Introduction to VM basics: Linux, command line
- Using SimpleVM with Research Environments for applied bioinformatics use-cases

Intermediate

- Bioinformatics workflows: BioConda, BioContainers, CWL, Nextflow, Snakemake

Advanced

- IaaS / PaaS courses for: OpenStack, Kubernetes
- Infrastructure automation: Ansible, Terraform
- Grid computing in the cloud: BiBiGrid / SLURM

Training Announcements at <https://www.denbi.de/de-nbi-events> or @denbiOffice, Registration: <https://events.denbi.de/>

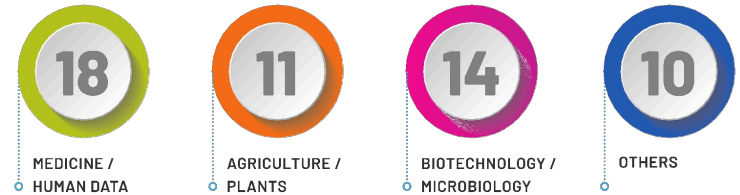
Industrial Forum

- Discussion and networking opportunity for academic and industrial stakeholders
- Access to the **experts** network
- Access to bioinformatics **services, training** and **de.NBI Cloud**
- **Free membership**
- The Industrial Forum currently holds 39 member companies from Austria, Switzerland, Luxembourg, and Germany
- Since 2019: „Annual Meeting of the de.NBI Industrial Forum“



Annual Meeting of the
de.NBI Industrial Forum
Thursday, 23 November 2023

Fields of activity of the member companies of the Industrial Forum



The German Node within ELIXIR Europe



ELIXIR DE

- **ELIXIR** is an intergovernmental organization (ESFRI infrastructure) and brings together life science resources from across Europe
- The ELIXIR consortium consists of **22** member states + EMBL (status 2022)
- **ELIXIR Germany** is the national Node of the European ELIXIR infrastructure since 2016
- The node is operated by members of **de.NBI** - the German Network for Bioinformatics Infrastructure



Participating Institutes



Tools & Services



Training Courses in 2021 with 1,150 Participants

Structure of the ELIXIR Scientific Programme 2024-28

Tiers



SCIENCE



TECHNOLOGY



NODES



PEOPLE

Strategic Priorities Areas



Cellular and molecular research



Biodiversity, food security and pathogens



Human data and translational research



Research data management and knowledge sharing



Reproducible analytics and infrastructure



Federated service delivery



Node operations



National research and open science policy alignment



Integrating Nodes with industry and innovation



Demonstrating impact and supporting sustainability



Users of ELIXIR services



Management and operational staff

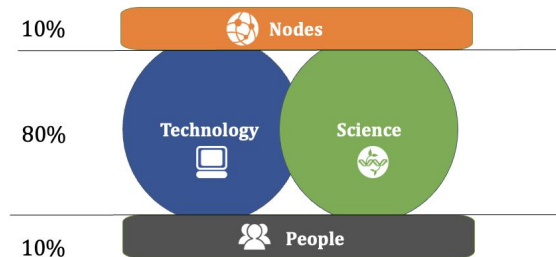


Technical and scientific staff



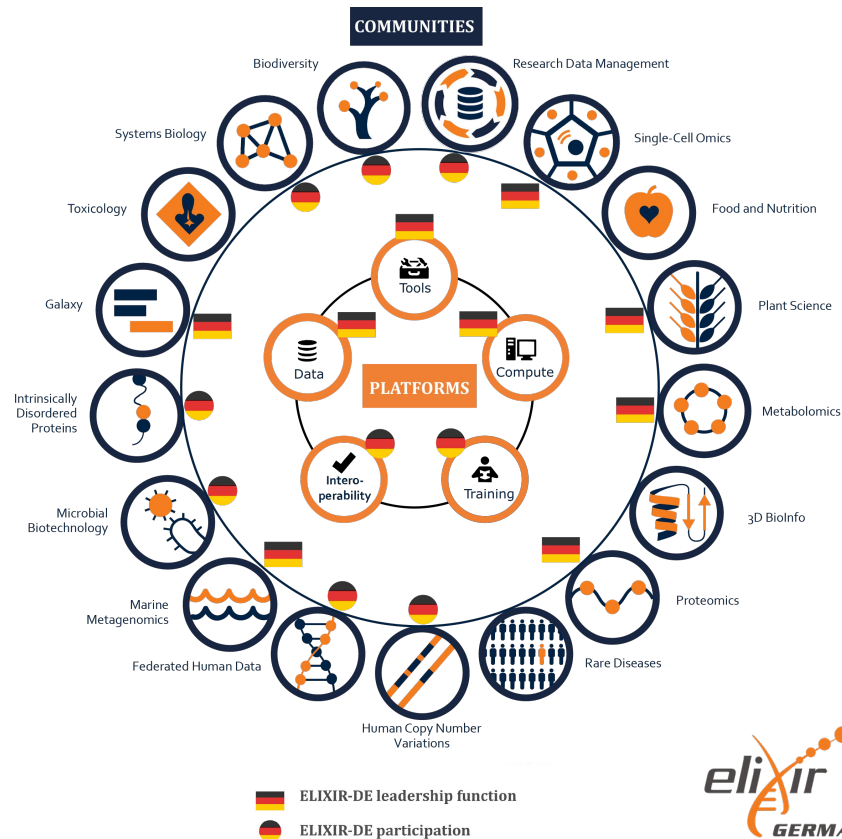
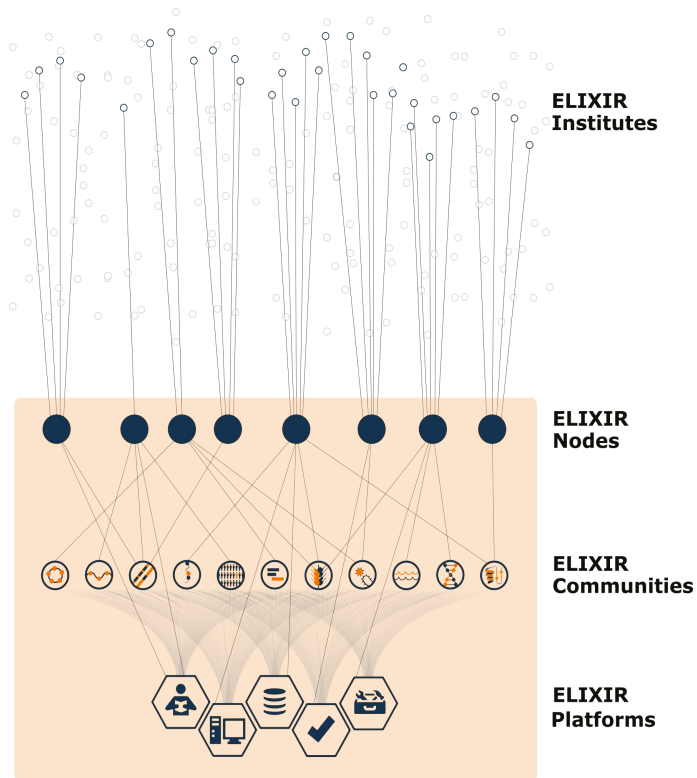
A diverse people network

- **Active involvement** is possible through participation in **ELIXIR Implementation Studies** (Commissioned Services, CoS)
- Participation is **exclusive to official members of a Node**
- Non-official partners may engage with the wider ELIXIR community, they are **not eligible to receive funding through ELIXIR**



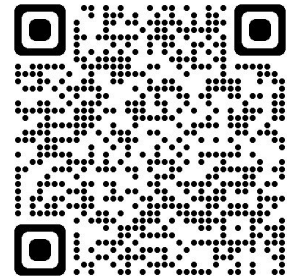
Planned investments 2024-28 by tier/ambition.

ELIXIR Organizational Concept





- Based on “[Ten simple rules for making training materials FAIR](#)” the [Handbook](#) was created.
- The ten chapters provide guidance on how you can FAIRify your training material.



ELIXIR offers **tools and training** through its nodes to the life science community:

- **Software solutions** for RDM problems in complex biological data sets and questions
- Access to **RDM knowledge** and curated **Digital Objects**
- Connection to the national and international **RDM community – access to stakeholders**
- **Guidelines, experts' opinions and best practices**
- **Training** and increasing self-capabilities
- **FAIR** policy and principle are executed
- **Trusted research environments** with the de.NBI Cloud

FAIRsharing.org

FAIRsharing.org
standards, databases, policies

FAIRsharing.org allows you to search for databases and data policies by aspects such as domain, species and country.

TeSS

elixir
TeSS

Search for training courses, webinars, training materials and workflows in TeSS, ELIXIR's training portal.

Step-by-step instructions

FAIRcookbook

The FAIR Cookbook contains step-by-step recipes to accomplish specific data management tasks and to make your data FAIR (Findable, Accessible, Interoperable, Reusable).

Overview of good data management practices

RDMkit

The Research Data Management Kit (RDMkit) guides you through the whole data management life cycle and includes advice specific to your domain, your role and your country.

Data management plan wizard

DSW

The Data Stewardship Wizard (DSW) is an online tool that guides researchers and data stewards through their data management planning.

ELIXIR-DE is active in the ELIXIR RDM efforts.

Involvement of ELIXIR-DE in ELIXIR Compute

ELIXIR Compute Platform

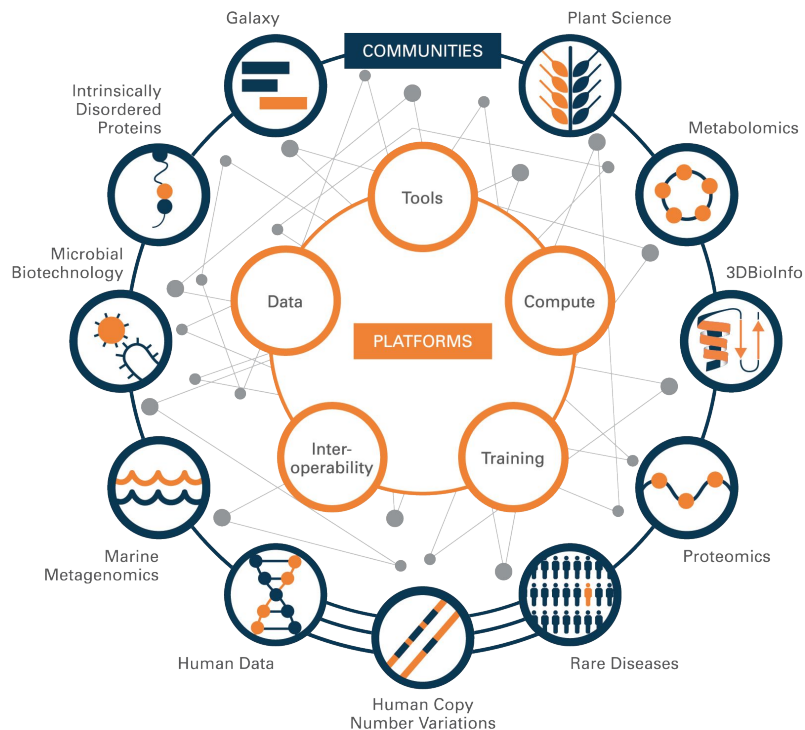
- Identity and access management
- ELIXIR-DE Platform Coordination
- Data Integration for Compute
- ELIXIR Hybrid Cloud Eco-system
- Deploying Reproducible Containers and Workflows across Cloud Environments

ELIXIR Tools Platform

- Packaging, Containerisation and Deployment

EOSC Focus Group

- EOSC Association Technical Interop TF



<https://elixir-europe.org/>

de.NBI Cloud Cooperation & Involvement

International and European Level

International Cooperation with / support of



→ ELIXIR-EXCELERATE



→ EOSC Life (Hybrid Cloud), EOSC Pilot, EOSC Entrust



→ H2020

→ GA4GH



Global Alliance
for Genomics & Health

→ EOSC HuB, EGI, GÉANT, Human Cell Atlas (upcoming)

→ ICGC Mirror, MII, Cancer Genome Collaboratory



HUMAN
CELL
ATLAS

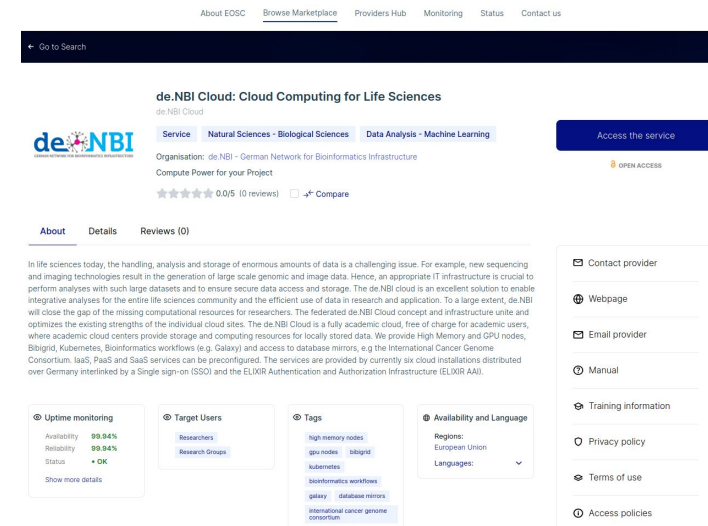
ICGC
International
Cancer Genome
Consortium



Federated German compute and storage cloud ready for all scientific projects in life sciences

→ Central data backbone and catalyst for German and international initiatives with German coordination

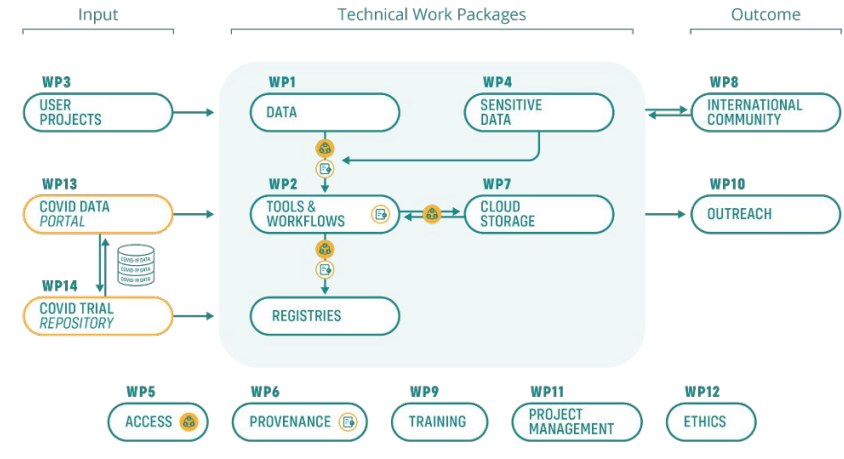
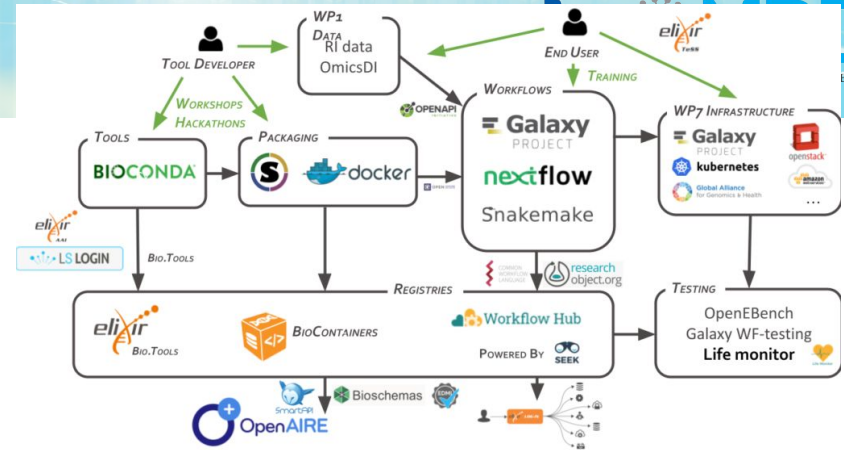
EOSC Marketplace



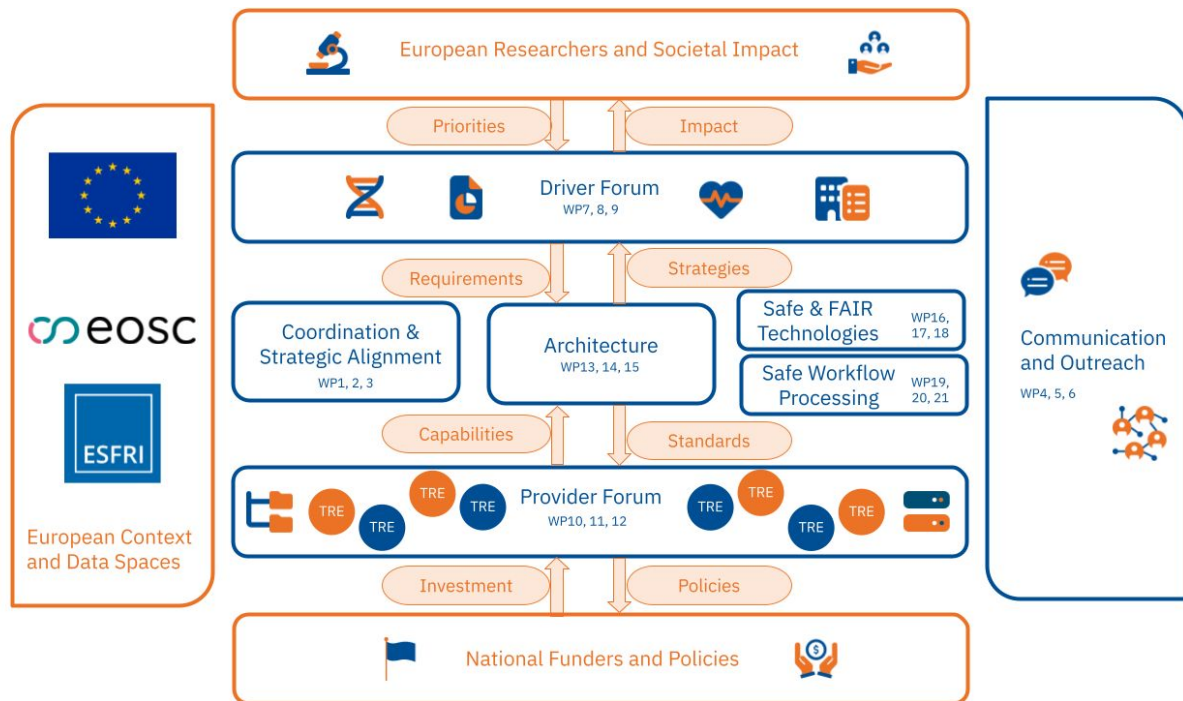
The screenshot shows the EOSC Marketplace interface. At the top, there are navigation links: About EOSC, Browse Marketplace, Providers Hub, Monitoring, Status, and Contact us. Below this is a search bar with a 'Go to Search' button. The main content area displays the service details for 'de.NBI Cloud: Cloud Computing for Life Sciences'. It includes the de.NBI Cloud logo, service categories (Natural Sciences - Biological Sciences, Data Analysis - Machine Learning), and a button to 'Access the service'. The organization is listed as 'de.NBI - German Network for Bioinformatics Infrastructure'. There is a star rating of 0.0/5 (0 reviews) and a 'Compare' button. Below the service details, there are tabs for 'About', 'Details', and 'Reviews (0)'. The 'About' tab is active, showing a description of the service: 'In life sciences today, the handling, analysis and storage of enormous amounts of data is a challenging issue. For example, new sequencing and imaging technologies result in the generation of large scale genomic and image data. Hence, an appropriate IT infrastructure is crucial to perform analyses with such large datasets and to ensure secure data access and storage. The de.NBI cloud is an excellent solution to enable integrative analyses for the entire life sciences community and the efficient use of data in research and application. To a large extent, de.NBI will close the gap of the missing computational resources for researchers. The Federated de.NBI Cloud concept and infrastructure unite and optimizes the existing strengths of the individual cloud sites. The de.NBI Cloud is a fully academic cloud, free of charge for academic users, where academic cloud centers provide storage and computing resources for locally stored data. We provide High Memory and GPU nodes, Ibligid, Kubernetes, Bioinformatics workflows (e.g. Galaxy) and access to database mirrors, e.g. the International Cancer Genome Consortium, IaaS, PaaS and SaaS services can be preconfigured. The services are provided by currently six cloud installations distributed over Germany interlinked by a Single sign-on (SSO) and the ELIXIR Authentication and Authorization Infrastructure (ELIXIR AAI)'. Below the description, there are three panels: 'Uptime monitoring' showing 99.94% availability and reliability, 'Target Users' with a 'Research Groups' button, and 'Tags' including 'high memory nodes', 'gpu nodes', 'ibligid', 'kubernetes', 'bioinformatics workflows', 'galaxy', 'database mirrors', 'international cancer genome consortium', 'iaas', 'paas', and 'saas'. On the right side, there is a sidebar with a 'Contact provider' button and a list of links: Webpage, Email provider, Manual, Training information, Privacy policy, Terms of use, and Access policies.

H2020-INFRAEOS - Implementing the European Open Science Cloud

- EOSC-Life: 13 Biological and Medical Research Infrastructures (RI)
- Biobanking and Biomolecular Resources RI, ELIXIR, EuroBioImaging, INSTRUMENT, ...
- Start: March 2019
- Galaxy workflows (*Grüning, RBC*)
- Tasks in Cloud Deployment Services (WP7 Co-lead: Harald Wagener, *BIH*)
 - ◆ ***Coordination**
 - ◆ ***de.NBI Cloud resources for EOSC projects** (others: EBI, Finland, Czech Republic)
 - ◆ ***Training** (ResOps Training, Appliance)
 - ◆ Container Execution Platform
 - ◆ GA4GH-Compatible Workflow Platform
 - ◆ ***Galaxy Workflow Platform**
 - ◆ ***Cloud Appliances**
 - ◆ Cloud Portals
 - ◆ ***Operations and Support**
- Demonstrator: Metagenomic deployment study (*Sczyrba, Bielefeld*)
- ***tasks with de.NBI Cloud involvement**



- **Goal:** Create a European network of TRUSTed research environments
- Create Blueprints, Best Practice Guidelines and Software Components for sensitive data processing
- 34 European Partners
- Multiple European Infrastructures: CESSDA, ECRIN, HDR UK, ELIXIR, EUDAT
- de.NBI Cloud / ELIXIR-DE participates in TRE Provider Forum (WP10,11,12)



National Research Data Infrastructures (NFDI)

- ▶ GHGA
- ▶ NFDI4Microbiota
- ▶ NFDI4Biodiversity
- ▶ DataPLANT



NFDI Base Services Initiative

- ▶ Federated multi-cloud infrastructure for all NFDIs across scientific domains

GAIA-X

- ▶ HEALTH-X dataLOFT
H. Wagener, BIH/Charité



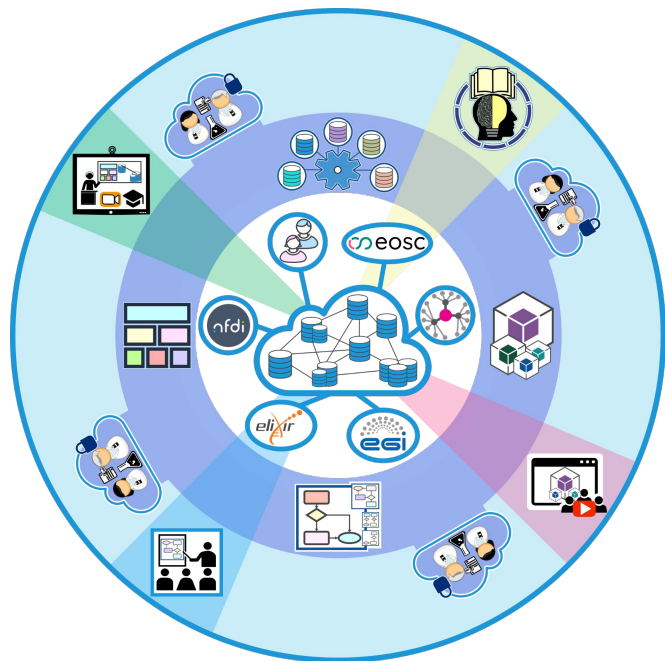
FORSCHUNG

Nationale Forschungsdateninfrastruktur

Forschungsdaten werden oft nur lokal und vorübergehend gespeichert. Dadurch gehen Datensätze verloren. Mit der Nationalen Forschungsdateninfrastruktur möchte das BMBF das ändern: Sie soll ein Wissensspeicher für die ganze Forschungslandschaft sein.

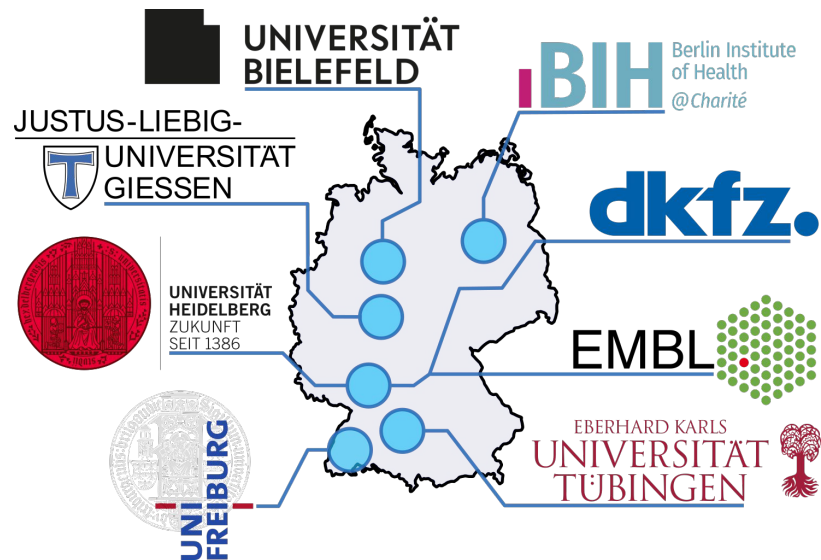


de.KCD - German Competence Center Cloud Technologies for Data Management and Processing



Collaboration with multiple RDM initiatives & NFDIs & other DCCs

Part of the Data competence centers for science German Recovery and Resilience Plan (DARP) and BMBF's Research data action plan
Project Administration: VDI/VDE Innovation und Technik GmbH



<https://datenkompetenz.cloud>

SPONSORED BY THE

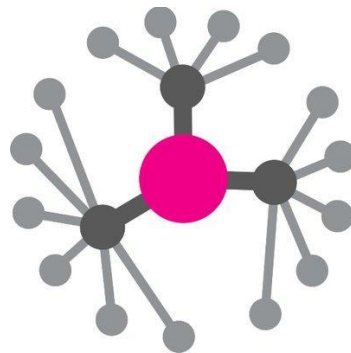
de.NBI Spring & Summer Schools

25-29 September 2023



de.NBI
Summer School
[2023] Bio Data Analysis
for Life Sciences

de.NBI
Spring School
[2024] Bioinformatics for
Microbial Omics



de.NBI Industrial Forum



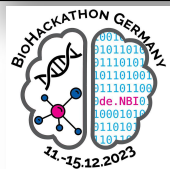
Annual Meeting of the
de.NBI Industrial Forum
Thursday, 23 November 2023

de.NBI Cloud User Meeting



CLOUD
USER
MEETING
23
BIELEFELD

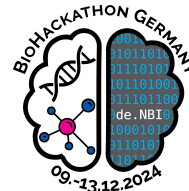
27-29 November 2023



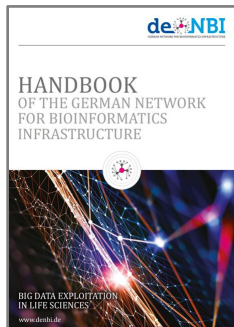
de.NBI Biohackathon Germany



BioHackathon 2024 in
Kassel (9.-13. Dec.)
Applications now open!



de.NBI and ELIXIR-DE Information Material



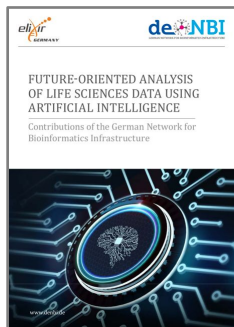
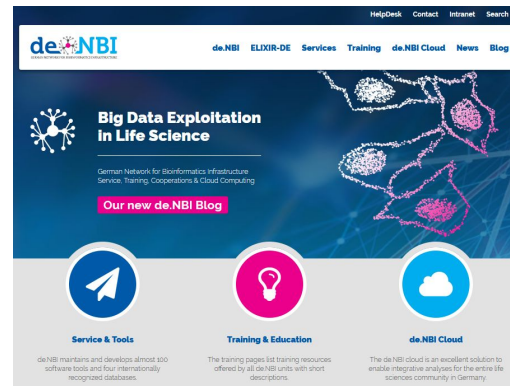
Handbook



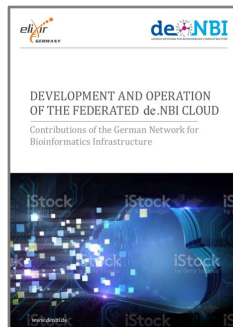
Highlights



COVID-19



AI



de.NBI Cloud



ELIXIR Germany



2nd BioHackathon Germany, 11-15 Dec 2023



Further Material:

- de.NBI Newsletter
- bespoke flyers
- white papers
- de.NBI Blog

<https://www.denbi.de/blog>

<https://www.denbi.de/downloads>

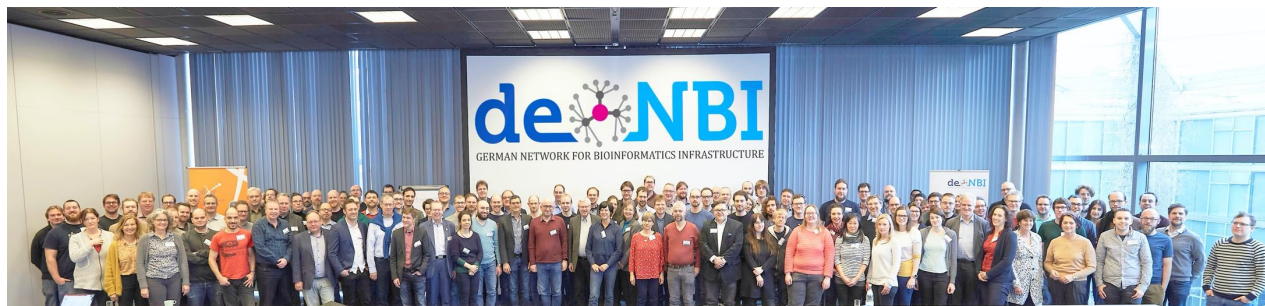
Thanks to:

de.NBI AO

- Tanja Dammann-Kalinowski
- Doris Jording
- Nils-Christian Lübke
- Johanna Nelkner

ELIXIR AO

- Alexander Sczyrba
- Oliver Kohlbacher
- Irena Maus
- Daniel Wibberg
- Helena Schnitzer



The people behind de.NBI

Contact us:

www.denbi.de



@denbiOffice



[linkedin.com/company/de-nbi](https://www.linkedin.com/company/de-nbi)



Cloud & Interoperability
Platform Coordinator

Dr. Nils Hoffmann
+49 (0) 521/106 86780
M3-125
n.hoffmann@fz-juelich.de
ORCID: [0000-0002-6540-6875](https://orcid.org/0000-0002-6540-6875)

SPONSORED BY THE



Federal Ministry
of Education
and Research

