de.NBI

GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

de.NBI Quarterly Newsletter

This de.NBI Quarterly Newsletter in **February 2021** informs about current developments and recent events in the BMBF project **German Network for Bioinformatics Infrastructure – de.NBI.**

Further information at: https://www.denbi.de/

COVID-19 research within de.NBI and ELIXIR Germany

The coronavirus pandemic is the defining global health crisis of our time. Thus, the German Network for Bioinformatics Infrastructure (de.NBI) and ELIXIR Germany have bundled the bioinformatic forces of our network partners to develop innovative bioinformatics solutions for the analysis of the data related to SARS-CoV-2 and COVID-19 research questions. These efforts are presented in the 110-page brochure entitled *Data Analysis for the COVID-19 Research – Contributions of the German Network for Bioinformatics Infrastructure*. It compiles 15 de.NBI/ELIXIR Germany-related COVID-19 projects and demonstrates how quickly and versatile the established bioinformatics infrastructures provide resources and support for analyzing large amounts of research data.

DATA ANALYSIS FOR THE COVID-19 RESEARCH

Contributions of the German Network for Bioinformatics Infrastructure
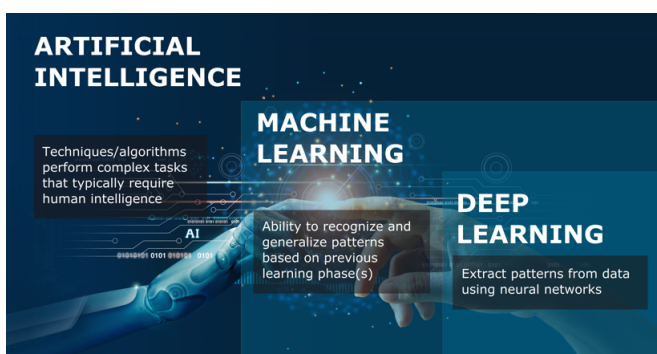
www.denbi.de

Furthermore, the brochure also informs about the network's main tasks, namely service, training, and cloud computing. The de.NBI Cloud, which is available free of charge to all researchers in the life sciences, allows the analysis of even the largest amounts of data. Scientists from Germany who are involved in research projects related to prevention and treatment of COVID-19 disease, or other work in connection with coronaviruses currently receive priority access to de.NBI Cloud resources.

The involvement of de.NBI/ELIXIR Germany in the fight against the pandemic on a national and European level is detailed in an interview with the de.NBI coordinator Alfred Pühler and the Head of the Node of ELIXIR Germany Andreas Tauch.

For more information please visit: https://www.denbi.de/images/Downloads/deNBI_COVID19_brochure.pdf

Involvement of the de.NBI network in artificial intelligence projects

ARTIFICIAL INTELLIGENCE

Techniques/algorithms perform complex tasks that typically require human intelligence

AI

MACHINE LEARNING

Ability to recognize and generalize patterns based on previous learning phase(s)

DEEP LEARNING

Extract patterns from data using neural networks

The use of artificial intelligence (AI) for life sciences has enormous potential to advance topics such as personalized medicine, drug discovery or basic biological research. Researchers within the de.NBI network are well aware of this potential. This is reflected by a large number of AI-related projects, which are carried out by de.NBI members. This collection of currently 20 projects is presented on the de.NBI website.

Moreover, the de.NBI Cloud offers an optimal solution for running AI approaches. Methods of artificial

intelligence can be sped up by the usage of graphical processing units (GPUs) and especially deep learning can benefit from the access to tremendous data amounts. Those benefits are well known within the community as demonstrated by currently 49 projects applying AI approaches that are running in the de.NBI Cloud. Moreover, 19 of those are performed by non-de.NBI members demonstrating the awareness and the benefits of the de.NBI Cloud in the compute and life sciences community.

For more information, please visit: https://www.denbi.de/ai-projects

de.NBI Cloud increasingly deployed for teaching

Since the start of de.NBI Cloud, cloud resources have been used for de.NBI training activities. Recent workshops such as the "Joint de.NBI/SIMBA Training Course on Metagenome anlaysis" and the "de.NBI - CeBiTec Nanopore Workshop 2020 –Best Practice and SARS-CoV-2 applications" used cloud resources to train researchers in Germany recent approaches and tools in bioinformatics.

A new development in the usage of the de.NBI Cloud are university-specific bioinformatics courses for teaching students. This is highlighted by one of our latest testimonials provided by Carl Hermann who is group lead at Health Data Science Unit – Biomedical Genomics Group in Medical Faculty Heidelberg. They use the de.NBI Cloud for various use cases such as the data challenge on cellular deconvolution of tumor samples within the COMETH project. The group is involved in a European network of bioinformatics Master Programs as part of the 4EU+ alliance and uses the de.NBI Cloud services to provide students with the computational infrastructure for running analysis on Hi-C data and domain prediction. Recently, they also offered a full online ChIP-seq and ATAC-seq analysis workshop for members of Frankfurt University. In particular Carl praises the easy setup and the tremendous support by the de.NBI Cloud team.

For more information please visit: https://cloud.denbi.de/testimonial-de-nbi-cloud-in-teaching/

Implementing FAIR data management within de.NBI

In their latest article the SIG 4 Interoperability & Data Management addresses the challenge of heterogeneity of wet lab technologies, data, metadata, software, computational workflows and the levels of implementation and monitoring of FAIR principles within the different bioinformatics sub-disciplines joint in de.NBI. In six use cases the SIG 4 members describe data management and perform a self-assessment of FAIR principles. The use cases cover a wide range of different services within the network:

- Use case 1: Bioinformatics and statistical consulting
- Use case 2: PRIDE upload of proteomics data
- Use case 3: PIA—protein inference algorithms
- Use case 4: Integrated workflow for the handling of NGS data and metadata
- Use case 5: Human genomics data in the cloud
- Use case 6: Data life cycle for high throughput plant phenotyping in controlled environments
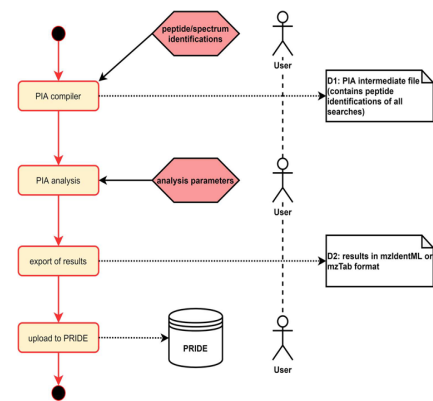
Flowchart diagram of a PIA analysis

Full citation: ***Implementing FAIR data management within the German Network for Bioinformatics Infrastructure (de.NBI) exemplified by selected use cases***, G Mayer et al. (2021), Briefings in Bioinformatics, bbab010, https://doi.org/10.1093/bib/bbab010

Provisionally available at: https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbab010/6135008?guestAccessKey=f03a534c-0b17-4fe1-96ea-20d37f2b4d52

Despite COVID-19: 1,149 training course participants in 2020

In 2020, due to the global coronavirus pandemic, almost all face-to-face training courses in Germany had to be cancelled but the network was able to answer this with the establishment of a wide range of online training courses.

de.NBI started in 2015 with 17 training courses and 329 participants, increasing numbers in 2016 (40 courses with 882 participants), 2017 (69 courses with 1,489 participants), 2018 (77 courses with 1,520 participants) and 2019 (79 courses

with 1,586 participants). However, based on online training, the number and diversity of training activities was still high (40 courses with 1,149 Participants) in 2020. 33 of the courses were held virtually.

The largest course was attended by 210, the smallest one by four scientists. Approx. 75% of the participants were PhD students or Postdocs. The remaining 25% were PIs, undergraduates, and participants from industry. The gender ratio of the students was 45% male and 55% female. We saw a remarkable increase of international and overseas participants in our courses due to the simplified online participation. About 60% of the participants came from outside Germany. In general, the user survey shows high satisfaction values of around 90%.
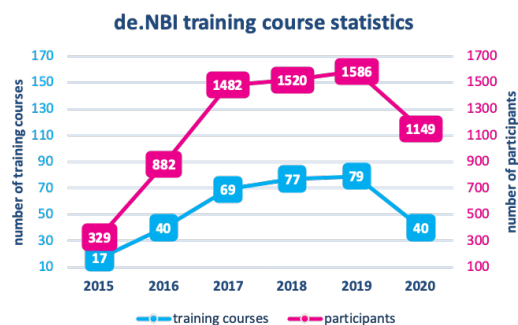
The de.NBI/ELIXIR-DE metaRbolomics Hackathon   by S. Neumann

The programming and statistics environment R has emerged as one of the most popular environments to process and analyze metabolomics datasets. A major benefit of such an environment is the possibility to connect different tools into more complex workflows, and to benefit from the wealth of packages, datasets, and tutorials in statistics in general.

from Stanstup et al. https://doi.org/10.3390/metabo9100200
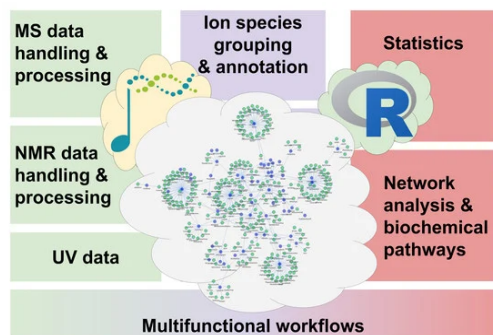
The **#metaRbolomics** community started with a first workshop at the Metabolomics Society conference in Dublin (2016), and then thereafter plan was to write up a short review on how to combine a few R packages into a workflow. After the de.NBI metaRbolomics Hackathon at the venerable Leucorea in Lutherstadt Wittenberg (2019), 19 authors published a review of nearly 300 R packages in Bioconductor, CRAN and elsewhere (10.3390/metabo9100200).

Already back in 2019 we scheduled the **2nd metaRbolomics Hackathon** for 2021, and we are excited that it will take place from Monday till Wednesday **22-24 March 2021 as an online workshop**.

Over the last few years, Bioconductor packages for metabolomics and proteomics data analysis started converging towards a common mass spectrometry infrastructure, and we aim to continue working on improved findability, interoperability and reusability of different packages. The agenda will have a fixed plenary slot each day 9-11 and 15-17 (CET). The online communication channels (videoconference, chat and docs) will remain open all the time, and we will facilitate that breakout groups can meet in between and work on various topics.

We also intend to run an embedded "**Data Clinic**" event on Tuesday 23 March @ 16:00 (CET), which will also be announced to the user communities. The idea is to have a one-hour helpdesk-like slot where users and contributors can show up and meet the experts, ask anything about R code, data, pull requests etc., just like you would have at a face2face conference during the coffee break.

For details, please see https://www.denbi.de/training/1143-de-nbi-elixir-de-metarbolomics-hackathon

de.NBI articles in the first issue of *gesundhyte.de*

In December 2020, the 13th edition of the magazine *systembiologie.de* dedicated to digital health in Germany was published under the new title *gesundhyte.de*. This first edition deals with the broad application potential of current digital methods in the life sciences and medicine. The de.NBI network contributed with three articles to this issue:

- *Ein Motor für die Analyse von COVID-19-Forschungsdaten - Das Deutsche Netzwerk für Bioinformatik-Infrastruktur überzeugt durch Auswerteprogramme und eine Cloud-basierte Rechnerstruktur*, by A. Pühler, I. Maus, V. Ortseifen, and A. Tauch

- *de.NBI-Cloud als akademische Lösung für Lebenswissenschaftler – Cloud Computing bietet flexible skalierbare Rechen- und Speicherressourcen für Big Data Anwendungen*, by C. Lawerenz and A. Sczyrba

- *Bioinformatik: Junge Frauen brauchen Vorbilder*, an interview with Jannie Felden by G. Terstiege

The first article, written by the de.NBI coordinator Alfred Pühler and colleagues of Administration Office, shades the light on de.NBI/ELIXIR Germany-related COVID-19 projects carried out by de.NBI members. 29 different projects have been

outlined with topics ranging from virus genome sequencing, bioinformatics tools for multi-omics data analyses to identification of potential drugs and drug targets against SARS-CoV-2. Without a doubt, the named projects cover the entire COVID-19 research spectrum and demonstrate that the existing de.NBI resources could be used effectively and within the shortest possible time, and that the network is a motor for the analysis of COVID-19 research data. Find all de.NBI/ELIXIR Germany-related COVID-19 research projects mentioned at https://www.denbi.de/covid-19.

The article by C. Lawerenz and A. Sczyrba explains the advantages of the de.NBI Cloud and presents an approach for highly scalable software solutions run on a high-performance infrastructure that is rented dynamically as required. Currently, scientists involved in projects to curb the corona pandemic, prevent and treat COVID-19, or other coronavirus-related work will be given priority access to de.NBI Cloud resources (https://www.denbi.de/cloud).

The third article reports on female group leaders in bioinformatics who are still in the minority. Janine Felden, PI of the Pangaea data publisher and member of de.NBI, is one of them. In an interview with gesundhyte.de, she talked about her professional career and what should change to attract more women to find their way into bioinformatics, an area that still fascinates and inspires Janine today.

This interview has also been published on the website of the German Ministry for Education and Research (BMBF) https://www.bmbf.de/de/dr-janine-felden-im-interview-junge-frauen-brauchen-vorbilder-13239.html.

Find the complete magazine at https://www.systembiologie.de/de/magazin

News from the de.NBI Office

Since January 2021, the de.NBI and ELIXIR-DE Offices have two new employees:

Johanna Nelkner is the new member of the de.NBI Administration Office (AO). Johanna holds a Bachelor's degree in Bioinformatics & Genome Research and has a Master's degree in Genome Based Systems Biology from Bielefeld University. Her research interests are centered around metagenome analyses of microbial communities inhabiting agricultural soils and biogas plants. She is also interested in agile project management topics and likes to take part in mentoring programs and workshops that broaden one's horizon.

Johanna will support the de.NBI AO activities in the SIG 2 Service & Service Monitoring as **Service coordinator**.

Helena Schnitzer is the new member of the ELIXIR Germany Administration Office (AO). Helena holds a Bachelor´s degree in Bioinformatics and Genome Research and has a Master´s degree in Molecular Cell Biology from Bielefeld University. Currently, she is finalizing her PhD in plant biochemistry and physiology. Her research interests comprise proteomics and systems biology.

As a **Training Platform Coordinator**, Helena will support the ELIXIR-DE training activities in the SIG 3 Training & Education and the ELIXIR-CONVERGE WP 2 project.

We welcome Johanna and Helena in our team and look forward to a successful cooperation!

News from the industrial branch of the de.NBI network

The 1st Annual Meeting of the de.NBI Industrial Forum took place on 26 November 2020 as an online event. The topics of the program ranged from COVID-19 over plant bioinformatics to new de.NBI industrial services. Since more than 60 participants were counted, the meeting was considered to be successful.

To get an impression of the performance of the event, we conducted a short anonymous survey. The results are encouraging! Most participants rated the event and the format of the meeting very highly. Over 80% considered the information content as high and would recommend future events to their colleagues. With 90% the speakers' quality was rated highest among all questions asked. Furthermore, we asked how attendees became aware of the event. The vast majority were informed about the event via email, around 20% via the de.NBI Website and around 15% by other channels. Several comments and suggestions as well as criticism were also
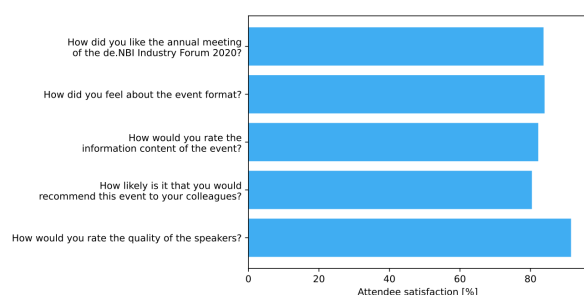
Results of the de.NBI Industry forum survey

mentioned by respondents to the survey. These are listed on the website where we also added answers to the comments: https://www.denbi.de/industrial-forum/1115-de-nbi-annual-industrial-forum-meeting-2020-survey

In addition, we would like to welcome our newest member, KNIME GmbH. The Industrial Forum now counts 35 companies in total from four countries: Germany, Luxembourg, Austria and Switzerland.

Further information at: https://www.denbi.de/industrial-forum

Recent Updates of de.NBI Services

The de.NBI bioinformatics services portfolio 2021 now includes 45 new tools, workflows, web apps and databases. Check our services at https://denbi.de/services.

New features or versions of various de.NBI services and tools have been released in the last months:

- **BRENDA:** The new release 2021.1 is online. Find the service at: https://www.denbi.de/services/346-brenda-the-world-s-largest-and-most-widely-used-information-system-on-all-aspects-of-enzymes. Furthermore a new publication on the Database is out: Chang A *et al.* **BRENDA, the ELIXIR core data resource in 2021: new developments and updates**. Nucleic Acids Res. 2021 Jan 8;49(D1):D498-D508. doi: 10.1093/nar/gkaa1025.

- **Massbank:** New version 2021.02 with many updated records. Find the service at https://www.denbi.de/services/501-massbank-repository-of-mass-spectra-of-small-chemical-compounds

- **Bioinfra.Prot proteomics tools**: Materials of the basic R training course *Advanced analysis of quantitative proteomics data using R* are provided as online media at https://denbi.de/online-training-media-library/1173-advanced-analysis-of-quantitative-proteomics-data-using-r

- **COPASI** 4.29 (Build 228) Released. Find the service at: https://www.denbi.de/services/373-copasi-creating-and-solving-mathematical-models-of-biological-processes
Additionally, CoRC, which provides a high-level API for Copasi's computational backend in the R programming environment has been created. This allows the scripting of sophisticated systems biology workflows in R for reproducible research or for batch-processing of more demanding calculations on compute clusters. a The CoRC interface is designed for both simplicity and flexibility and simulation and analysis results are immediately available in native R data structures, such as tibbles, for further (statistical) analysis or visualization. Find more information on Github (https://jpahle.github.io/CoRC, including tutorials and examples) and in our publication in Bioinformatics (https://doi.org/10.1093/bioinformatics/btab033).

- **SeqAn3** is now also available via macports at https://ports.macports.org/port/seqan3/summary. Note that we renamed the SeqAn port to SeqAn2 to use SeqAn2 and SeqAn3 simultaneously.

Recent Publications by de.NBI partners:

- A Kumar, H Rasche, B Grüning, R Backofen, **Tool recommender system in Galaxy using deep learning**, *GigaScience*, Volume 10, Issue 1, January 2021, giaa152, https://doi.org/10.1093/gigascience/giaa152

- Altuna Akalin, **Computational Genomics with R**, 2020 by Chapman and Hall/CRC, ISBN 9781498781855, Press release by MDC: https://www.mdc-berlin.de/news/news/new-textbook-computational-genomics

- The results of two major international projects on the barley and wheat pan genome with participation of GCBN members have been published. Data are integrated in the GCBN tool suite CATS:

    - S Walkowiak, ...H Gundlach, ...U Scholz, ...KFX Mayer *et al.* **Multiple wheat genomes reveal global variation in modern breeding**. *Nature* **588**breeding. *Nature* **588,** 277–283 (2020). https://doi.org/10.1038/s41586-020-2961-x

    - M Jayakodi, ...H Gundlach, ...KFX Mayer,...U Scholz, *et al.* **The barley pan-genome reveals the hidden legacy of mutation breeding** *Nature* **588,** 284–289 (2020). https://doi.org/10.1038/s41586-020-2947-8

UPCOMING EVENTS

The calendar lists the de.NBI training courses and events scheduled in the next months. All courses are accessible via the de.NBI web page. The network is continuously developing alternative forms of training. Have also a look into our online training library, which is continuously expanded.

Under the actual conditions, please consult the de.NBI training program for updates.

❈ Overview of upcoming de.NBI events

	Date	Event	City	Organizer
🧬	23-26 Feb 2021	Introduction to R – Data manipulation and visualization	Online	BioInfra.Prot
🧬	25 Feb 2021	The February 2021 Papercuts CoFest	Online	RBC
🧬	08-09 Feb 2021	Software Carpentry Workshop 01 2021	Online	Ass. Partner ZB MED
🧬	08-12 Feb 2021	Introduction to working with Chemical Data and KNIME 2021	Online	CIBI
🧬	10 Mar 2021	Processing thousands of datasets simultaneously with Galaxy	Online	RBC
🧬	16 Mar 2021	Successful writing of bioinformatics grant proposals	Online	RBC/ de.STAIR
🧬	17 Mar 2021	Bridging two worlds with Galaxy	Online	RBC
🧬	22-24 Mar 2021	2nd de.NBI/ELIXIR-DE metaRbolomics Hackathon	Online	CIBI/ MASH
🧬	24 Mar 2021	Speed up your data analysis with Galaxy: Features no one knows about	Online	RBC
🧬	12-16 Apr 2021	KNIME Analytics Platform for Data Scientists (Life Science): Advanced - Online	Online	CIBI
🧬	12-16 Apr 2021	Bioinformatics carpentry utilizing Galaxy	Online	RBC/ de.STAIR
🧬	15 Apr 2021	4. ELIXIR-DE CCU Meeting	Online	CAU
🧬	15 Apr 2021	26. CCU Meeting	Online	CAU
🧬	20-22 Apr 2021	Tools for Systems biology modeling and data exchange: COPASI, Cell-NetAnalyzer, SABIO-RK, FAIRDOMHub/SEEK 2021	Online	de.NBI-SysBio
🧬	27 Jun–02 Jul 2021	CSAMA 2021: Statistical Data Analysis for Genome-Scale Biology	Bressanone/ Brixen	HD-HuB
🧬	27 Sep–01 Oct 2021	de.NBI Summer School 2021 - Proteomics & Lipidomics	Online	BioInfra.Prot/ LIFS, CIBI
🧬	05 Oct 2021	LIFS Course - ILS 2021 & LipidomicsForum 2021	Regensburg	BioInfra.Prot/ LIFS
🧬	22-24 Nov 2021	3rd de.NBI/ELIXIR-DE metaRbolomics Hackathon	Wittenberg	CIBI/ MASH

Further information at: http://www.denbi.de/training

IMPRESSUM

SUBSCRIPTION
If you are interested in receiving regular updates on de.NBI – German Network of Bioinformatics Infrastructure, please subscribe to the de.NBI Quarterly Newsletter at: https://www.denbi.de/downloads/de-nbi-newsletter-subscription
The mailing list will exclusively be used to distribute the de.NBI Quarterly Newsletter.

To unsubscribe from the de.NBI Quarterly Newsletter, please use https://www.denbi.de/downloads/de-nbi-newsletter-subscription

The de.NBI Quarterly Newsletter is a service of de.NBI – German Network for Bioinformatics Infrastructure for members, partners and interested public.
All photos are copyright of the de.NBI administration office unless marked otherwise.

de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE