

The following pages are an excerpt from the proposal of the NFDI4Health consortium submitted on 15.10.2019 to the DFG for the first funding period. They are intended for an open, transparent inter-consortia discussion of cross-cutting topics.

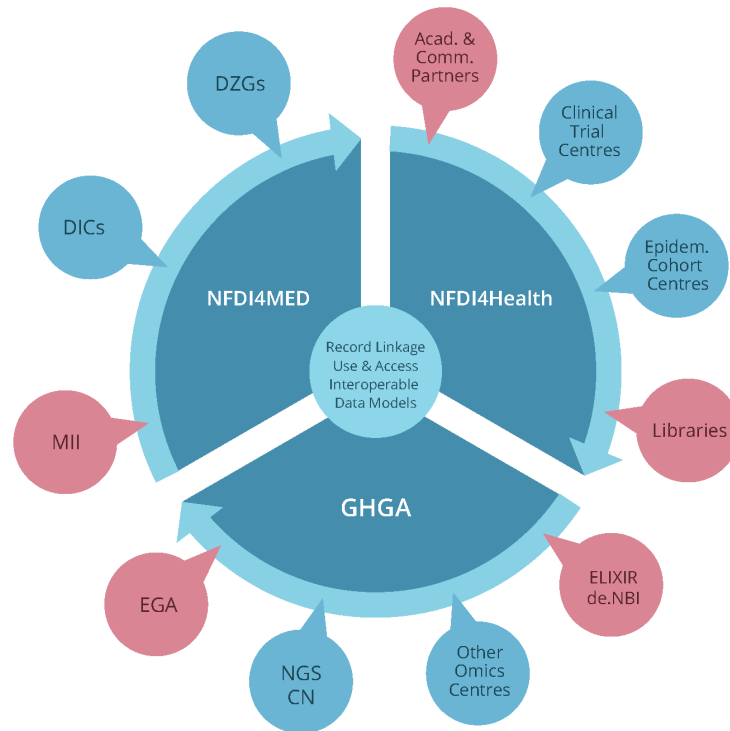
The consortium within the NFDI

In autumn 2017, ZB MED initiated the creation of NFDI4Life as an overarching consortium representing all disciplines in the life sciences. This broad approach turned out to be incompatible with the formal national agreement on the establishment of NFDI in Germany. Therefore, NFDI4Life was split up into several individual consortia, including NFDI4Agri, NFDI4BioDiversity, NFDI4Health, NFDI4MED, NFDI4Microbiota, and a newly structured NFDI4Life Umbrella at the end of 2018/beginning of 2019. During the subsequent elaboration process, the two epidemiology-orientated consortia NFDI4Health and NFDI4NutEpi merged to form NFDI4Health, while NFDI4Medicine and DZG4NFDI merged to form NFDI4MED.

NFDI4Health seeks cooperation with all other life science NFDI consortia and with consortia dealing with person-related data beyond the health domain. Particularly intense consultations have taken place with the other two consortia from the medical domain, namely NFDI4MED and GHGA. Since NFDI4Health, **NFDI4MED** and **GHGA** target different types and provenances of health-related data bodies requiring different types of expertise, infrastructure and methodology, they have agreed on a close partnership in addressing processes of patient identification and record linkage, in providing metadata standards and common interoperable data models and in establishing common services building on the framework provided by the MII. This is facilitated by the fact that several NFDI4Health co-applicants and participants are also active in the MII and **NFDI4MED**. The three consortia provide complementary infrastructure components: bridging storage and management of medical research data from epidemiological and clinical studies (NFDI4Health), patient-oriented clinical routine care, clinical and basic research data (**NFDI4MED**), and omics raw data (**GHGA**) as illustrated in Figure 2.

Mutual consultations have led to coordinated working programmes of the three consortia and an explicit commitment to an intensive collaboration. While the different data bodies result in a clear division of labour, central tasks will be worked on jointly by all three consortia. Ethical, legal and societal issues are, e.g., synergistic cross-sectional topics to which these consortia will contribute. In detail, the following joint working activities are planned: **Collaboration data search and request brokering**: NFDI4Health and **NFDI4MED** will build central data search and access services for their communities. They will coordinate issues of data discovery and governance of data use and access processes and by this will ensure interoperability. This collaborative activity is also open to **GHGA**. **Collaboration core data set**: Like **NFDI4MED**, NFDI4Health will use the MII core data set as a starting point for, e.g., metadata standardisation which will be further specified in a joint effort. **Collaboration data publishing/archiving**: There is general agreement that the guidelines, services, etc.

developed by one consortium may be used by the other consortia. For example, NFDI4Health plans to develop publication guidelines that can be used by **NFDI4MED**. **Collaboration on distributed analyses of data:** In particular, both, NFDI4Health and **NFDI4MED** will develop complementary add-ons for the DataSHIELD framework which will be accessible by the user communities. In addition, NFDI4Health will extend the Personal Health Train (PHT) framework that will be useful for **NFDI4MED**.



Interactions and responsibilities of three NFDI consortia in the medical field (blue circles: data sources, red circles: method development & transfer; DZGs: German Centres for Health Research, DICs: data integration centres, EGA: European Genome-phenome Archive, NGS CN: next-generation sequencing competence network)

NFDI4Health will also coordinate its efforts together with **KonsortSWD** in addressing challenges that are related to making sensitive cohort and survey data reusable and interoperable. BIPS has recently joined the EcoSoc Implementation Network in order to cooperate directly under the auspices of the GO FAIR initiative.

NFDI-Neuro and NFDI4Health will share standardisation policies and processes to develop common standards for (meta-)data, record linkage and interfaces. For standardisation of chemical components such as medication, dietary factors or metabolome data, NFDI4Health will explore the possibilities of standardisation and cross-sectional mapping together with **NFDI4Chem** and **NFDI4Microbiota**. NFDI4Health will further explore opportunities for data linkage with environmental data. In this respect, NFDI4Health is looking forward to a close collaboration with, e.g., **NFDI4Agri**, **NFDI4Earth**, **NFDI4BioDiversity** and

NFDI4NanoSafety. All cross-cutting topics (see below) and standardisation with further life science consortia will be organised in concert with **NFDI4Life Umbrella**.

The above collaborations are fostered by joint memberships of (co-)applicants of NFDI4Health in NFDI4Agri (ZB MED), NFDI4BioDiversity (HITS), NFDI4Crime (UAS Mittweida), NFDI4Life Umbrella (almost all NFDI4Health (co-)applicants, planned for 2020), NFDI4MED (Charité/BIH, TMF), NFDI4Microbiota (ZB MED, planned for 2020), NFDI-Neuro (Fraunhofer MEVIS, planned for 2020). Further steps have been taken to strengthen collaboration among NFDI consortia. During a meeting in Berlin, August 2019, eleven consortia agreed on the so-called **Berlin Declaration** (<https://doi.org/10.5281/zenodo.3457213>) that describes a common vision in particular on cross-cutting topics and that is supported by further NFDI consortia. NFDI4Health will contribute to all cross-cutting topics that are relevant for person-related health data accounting for the specific requirements of our user communities and data holders as, e.g., overarching data standards and interoperability. Of specific importance are the following cross-cutting topics of the **Berlin Declaration**:

Technical infrastructure and concepts

Standardisation: (Co-)applicants of NFDI4Health have leading roles in relevant domain-specific (inter-)national standardisation bodies (e.g., ISO¹⁶, IEC¹⁷, CEN/CENELEC¹⁸, DIN¹⁹, HL²⁰), as well as meta initiatives such as the European network EU-STANDS4PM⁸ and the European COST action CHARME²¹, which will be beneficial for many NFDI consortia, especially the ones related to health and life sciences.

Monitoring of data quality and federated data analysis: Quality of data to be shared has to be closely monitored according to predefined criteria. Here, NFDI4Health will build on the DFG-funded project Standards and tools for data monitoring in complex epidemiological studies where a broad epidemiological community was involved under the lead of UM Greifswald. In addition, NFDI4Health will provide experience in the development and implementation of federated data analysis infrastructures to all other NFDI consortia.

Data management tools and data sharing models: NFDI4Health will develop tools, e.g., to create data management plans with a specific catalogue of relevant questions for data types represented by NFDI4Health, and set up processes and models for data sharing of sensitive person-related data. Appropriate models and metadata sets will be developed, e.g., by extending the generic metadata standard of DataCite. Furthermore, requirements for archiving metadata and data sets will be formulated. These tools and models can be adopted by other domains handling sensitive data, e.g., by KonsortSWD.

Legal and ethical aspects

Data privacy, data protection laws and record linkage: Given the specific requirements with respect to data privacy, NFDI4Health will provide key knowledge and expertise regarding individual data protection and its implementation for the exchange of study data.

NFDI4Health will benefit from a strong expertise provided by a legal expert with focus on data protection and long-standing experience in ethical requirements of research studies involving humans (U Bremen). NFDI4Health will develop solutions to enable sharing of personalised data in Germany and will give recommendations for necessary revisions of German law, in particular with respect to record linkage.

Community (user) involvement

Training and education: NFDI4Health will provide specific training in good practice of health data collection, data management, data access constraints and correct use/analysis of data. This does not only concern primary data collection but also the secondary use of existing databases and registries. Here, NFDI4Health will build on the strong expertise in “Good Practice of Secondary Data Analysis” and “Good Practice Data Linkage” of several co-applicants and participants, in particular U Magdeburg and PMV Forschungsgruppe. Moreover, together with NFDI4Earth, NFDI4BioDiversity, the city and state of Bremen and U Bremen, NFDI4Health has started to establish a graduate education programme on research data management and data science. Here, also KonsortSWD and NFDI-Neuro have expressed their interest to actively contribute to the development of the curriculum and its teaching modules.

From NFDI4Health perspective, the last two topics above need most urgent support from the NFDI’s collaborative framework. Changing the legislation with respect to data sharing and record linkage will only be possible as a consolidated action of the whole NFDI. Also changing the curricula of education programmes will be facilitated by a concerted action involving all federal states. Since suitable governance structures are key to ensure sustainable operations within NFDI, NFDI4Health in addition urgently asks for an appropriate legal entity to serve the interests of the consortia.