Memorandum of Understanding between NFDI4Health and GHGA for the second funding period

Authors:
Deschler, K.; Eufinger, J.; Fluck, J.; Kohlbacher, O.; Korbel, J.; Nimptsch, K.; Pigeot, I.; Pischon, T.; Stegle, O.; Winkler E. for the GHGA and NFDI4Health Consortia
Version: 1.0, Date: 17/07/2024, DOI: 10.5281/zenodo.12799244

The Memorandum of Understanding by NFDI4Health and GHGA is an agreement on joint actions and interaction between the undersigned partners on addressing the needs and fostering the developments within the biomedical and genomic research communities.

Introduction

Omics data are an integral component of biomedical research, facilitating the translation of clinical research and standard medical care. In addition, it may be utilised for research into personalised prevention strategies. Genomic sequencing can identify genetic alterations, including those linked to phenotypic differences or disease risk including cancer. Thus, if combined with large population-scale resource omics and phenotypic data, data modalities promise to deliver biomedical utility from personalised diagnostic to prevention. The key for effective use of genomic data is - in alignment with the FAIR principles - (i) the ability to find data in linked repositories and (ii) the proper annotation and linkage of the genomic raw data to the clinical/health data of the patients and study participants. This consolidation and provisioning of data for the clinical and epidemiological, as well as for the genomic research communities is the core mission of NFDI4Health and GHGA, respectively.

GHGA and NFDI4Health will intensify their collaboration to develop new opportunities for data analyses that will advance the scientific exploitation of personal health data and eventually improve population health.

Objectives

The main objectives of the collaboration are the following:

1. Provide legal and ethical frameworks to ease data sharing of personal health data

GDPR regulations, in particular in the context of its interpretation in Germany, cause challenges for researchers and clinicians, who are intending to share personal health data, especially with regards to the consent and re-use of personal data. There is a need for common understandings and legal frameworks, particularly in the context of the German Health Data Usage Act and the European Health Data Space, which will increase the usability of personal health data while at the same time protecting the rights of the data subjects. Alignment of legal and data protection-related standards will be especially important for the involvement of GHGA and NFDI4Health within the German model project GenomSeq, governed by § 64e SGB V, forming the cornerstone for direct translation of personal health data into clinical care. In addition, the contributions and interests of the researchers and clinicians who obtained and are intending to share the data need to be taken into account.
With a focus on the persons providing personal health data (patients/study participants), the consortia will further work on ethical standards for the use and re-use of consented data, as guidelines for the researcher and clinical communities.

2. Consolidate overarching standards for data sharing in biomedicine

While legal and ethical standards provide the basis for data sharing, overarching standards are required for the technical and organisational measures (TOMs) needed for the exchange of personal health data. These TOMs usually differ between each research institution and clinic and cause a huge work overhead for all involved parties. GHGA and NFDI4Health have a track record in coordinating and executing those overarching standards as a genome-phenome archive (GHGA), with standards for data publication and access (GHGA, NFDI4Health), and in implementing the use of DataSHIELD for secure data processing by running federated data analysis without the need of direct access to or transfer of individual

person related health data (NFDI4Health). The consolidation and advocating of overarching standards for the sharing and use of personal health data is a core objective for both initiatives and - as a core strength - will generate a tangible benefit for the genomic research, clinical and epidemiological communities.

3. Align our infrastructure solutions to enable the analysis of comprehensive datasets

In a natural symbiosis, NFDI4Health and GHGA will join forces to align and consolidate their infrastructure services and catalogues. The consortia will work together to create common identifiers and linkage mechanisms, e.g using strategies similar to BioSamples established by the EBI. Intersecting datasets will be mutually linked, demonstrating functional and efficient implementation of data sharing practices. In this context, the consortia will further explore the usage of a joint secure processing environment, which will be developed for the GHGA Cloud Phase, and its connection to federated data analysis using approaches such as DataSHIELD. Alternative strategies will be tested in the context of use cases and datasets relevant to both consortia. The future development and implementation of standardised workflows for cloud based data analysis tailored to the respective target communities will be a long-term goal for the consortia.

4. Establish NFDI4Health and GHGA in the German and European Health Data Space

Using personal health data for biomedical research is a global endeavor with specific national intricacies due to legal and data protection regulations. As such, NFDI4Health and GHGA share the responsibility of bridging national and international communities and of establishing policies and standards for the common use. The consortia aim at taking a pivotal role in influencing national policies and data sharing structures, e.g. within the MV GenomSeq, and by engaging with overarching governance bodies such as the Koordinierungsgruppe Gesundheitsforschungsdateninfrastruktur (GFDI) organised by the Medizin Informatik Initiative (MII), the Medizinischer Fakultätentag (MFT), and the Netzwerk Universitätsmedizin (NUM). The consortia will further pro-actively disseminate existing data sharing policies by their funders for sharing personal health data. The developed and refined concepts will be the foundation for establishing GHGA and NFDI4Health within the European Health Data Space (EHDS) as opinion leaders in the legal, ethical and technical exchange of personal health data across borders.

Measures

To achieve these shared objectives, we will take the following joint measures:

A. Collaborative work within NFDI

Both NFDI4Health and GHGA are actively engaged in several sections and task forces within NFDI. Within this MoU, the initiatives aim at intensifying their joint efforts within the section ELSA, which is representing ethical, legal and social aspects on a national level, by e.g. commenting on health data related legislation. Specifically, the consortia will develop mechanisms to electronically represent and capture different types of consent as legal basis for research use (“ontology for human data consent”). Within the Taskforce Ethics, concepts will be developed on how e.g. dark data can be made available for community use. Furthermore, NFDI4Health and GHGA will collaborate closely, also with other NFDI consortia, to establish common metadata definitions and interoperability standards for the reuse of sensitive personal data.

B. Participating in the corresponding consortia

The current collaborations will be further intensified in the second funding period of the consortia. Partners of each consortium will engage within each other’s project as named participants in the consortium agreement. Personal exchange and alignment on mutual topics will be fostered within joint working group meetings, e.g. on the mapping of metadata models for biomedical research, or the active participation at larger events such as annual meetings by the consortia. Specifically, NFDI4Health representatives will actively shape the discussion on sharing human genome and phenome data in Germany and beyond at the GHGA Annual meeting in October 2024.

C. Joint Use Cases

As a prominent use case, both NFDI4Health and GHGA will support omics data management and data processing aiming at joint analyses of data from large participating cohort studies in Germany, such as the NAKO health study, KORA, EPIC, or DONALD. Prerequisite will be the mutual linkage of the intersecting datasets within the technical structures of the consortia. Within their respective realm, the consortia will compare federated to centralised cloud approaches on data analysis with respect to several aspects relevant to the research

community, such as analysis options, analysis speed, or the legal setup. NFDI4Health will conduct federated analyses based on DataSHIELD already deployed on several cohorts, while GHGA will bring datasets together in a centralised secure space (TRE). Both consortia will furthermore jointly drive the development, deployment, and operation of sustainable data access infrastructures within the European Genomic Data Infrastructure initiative.

D. Organisation of a biomedical task force in the NFDI

The joint planning of a biomedical workshop within the NFDI towards autumn 2024 will be the kick-off for organising and leading a biomedical task force within the NFDI. NFDI4Health and GHGA will coordinate the creation of outputs tailored to biomedical research communities, such as overarching standards for sharing of personal health data or guidelines on the use of archives and compute platforms for biomedical data.

Signatories of the memorandum

For the NFDI4Health

Juliane Fluck (ZB Med, Cologne), Iris Pigeot (BIPS, Bremen), Tobias Pischon (MDC, Berlin)

For GHGA

Oliver Stegle (DKFZ, Heidelberg), Oliver Kohlbacher (University of Tübingen), Jan Korbel (EMBL), Eva Winkler (NCT Heidelberg)

Acknowledgement

GHGA - The German Human Genome-Phenome Archive and *NFDI4Health - Nationale Forschungsdateninfrastruktur für personenbezogene Gesundheitsdaten* are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the NFDI (Nationale Forschungsdateninfrastruktur). Grant Numbers: GHGA: 441914366; NFDI4Health: 442326535.