National Research Data Infrastructures (NFDI) Letter of Intent

1 Binding letter of intent as advance notification or non-binding letter of intent

Non-binding letter of intent (anticipated submission in 2021)

2 Formal details

- Planned name of the consortium

  National Research Data Infrastructure for Digital Pathology

- Acronym of the planned consortium

  NFDI4Patho

Applicant institution

- **RWTH Aachen University**

  Univ.-Prof. Dr. rer. nat. Ulrich Rüdiger

  Spokesperson

  - Univ.-Prof. Dr. med. Peter Boor, Ph.D., pboor@ukaachen.de, Institute of Pathology and Department of Nephrology

  Co-Spokespersons

  - Univ.-Prof. Dr. med. Ruth Knüchel-Clarke, rknuechel-clarke@ukaachen.de, Director of the Institute of Pathology
  - Univ.-Prof. Dr. rer. nat. Edgar Dahl, edahl@ukaachen.de, Head of Molecular Oncology Group and Molecular Pathology
  - Univ.-Prof. Dr. med. Rainer Röhrig, rroehrig@ukaachen.de, Institute of Medical Informatics

Co-applicant institutions

- **Charite Universitätsmedizin Berlin**

  Prof. Dr. Axel Radlach Pries

  Co-Spokesperson

- Prof. Dr. Peter Hufnagl, peter.hufnagl@charite.de, Director of the Digital Pathology IT & Coordinator of EMPAIA

- **Technische Universität Dresden**
  Prof. Dr.-Ing. Hans Müller-Steinhagen
- **& German Society of Pathology (DGP)**
  Prof. Dr. med. D. Michael Albrecht

Co-spokesperson

- Prof. Dr. med. Gustavo Baretton, Gustavo.Baretton@uniklinikum-dresden.de, Director of Pathology Dresden and DGP Chairman

Participant

- Joerg Maas, maas@pathologie-dgp.de, DGP Secretary General

- **Georg-August-Universität Göttingen**
  Prof. Dr. Reinhard Jahn

Co-spokesperson

- Prof. Dr. med. Phillip Ströbel, philipp.stroebel@med.uni-goettingen.de, Director of Pathology

- **OFFIS e.V.**
  Prof. Dr.-Ing. Andreas Hein

Co-spokesperson

- Prof. Dr.-Ing. Andreas Hein, Andreas.Hein@offis.de

Participant

- Dr. rer. nat. Marco Eichelberg, marco.eichelberg@offis.de

- **TMF – Technologie und Methodenplattform für die vernetzte medizinische Forschung e.V.**
  Sebastian C. Semler

Co-spokesperson

- Sebastian C. Semler, Sebastian.Semler@TMF-ev.de

- **Justus-Liebig-Universität Gießen**
  Prof. Dr. Joybrato Mukherjee
- **& German Society of Neuropathology and Neuroanatomy (DGNN)**
  Prof. Dr. Till Acker

Co-spokesperson

- Prof. Dr. med. Till Acker, Till.Acker@patho.med.uni-giessen.de, Director of Neuropathology Gießen and DGNN Chairman

- **Medizinische Hochschule Hannover (MHH)**
  Prof. Dr. med. Michael Manns

Co-spokesperson

- Univ.-Prof. Dr. med. Danny D. Jonigk, jonigk.danny@mh-hannover.de, Institute of Pathology

- **Johann Wolfgang Goethe-University Frankfurt**
  Prof. Dr. Birgitta Wolff

Co-spokesperson

- Prof. Dr. med. Peter J. Wild, peter.wild@kgu.de, Director of the Dr. Senckenberg Institute of Pathology

- **Friedrich-Alexander-University (FAU) Erlangen-Nürnberg**
  Prof. Dr. Joachim Hornegger

Co-spokespersons

- Prof. Dr. Kerstin Amann, Kerstin.Amann@uk-erlangen.de, Director of the Department of Nephropathology
- Prof. Dr. Arndt Hartmann, Arndt.Hartmann@uk-erlangen.de, Department of Pathology

Participant

- Prof. Dr. Christoph Daniel, Christoph.Daniel@uk-erlangen.de; Department of Nephropathology
- Dr. Fulvia Ferrazzi, Fulvia.Ferrazzi@uk-erlangen.de, Department of Pathology

- **FIAS (Frankfurt Institute for Advanced Studies)**
  Prof. Dr. Enrico Schleiff, Chairman

  Co-spokespersons
  - Prof. Dr. Volker Lindenstruth, voli@compeng.de

Other participants:

- For the Federal Association of German Pathologists (BDP): Prof. Dr. Gunter Haroske, Head of the Digital Pathology Working group of the BDP
- Prof. Dr. Thomas J. Fuchs (Hasso-Plattner-Institute for Digital Engineering gGmbH (HPI gGmbH)

3 Objectives, work programme and research environment

- Research area of the proposed consortium (according to the DFG classification system:

  205 Medizin (205-06 Pathologie)

- Concise summary of the planned consortium's main objectives and task areas

Very recent technological developments have enabled effective digitalization of pathology. This, in turn, has laid the groundwork for unprecedented possibilities for automated, highly precise and reproducible pathology diagnostics, including the implementation of artificial intelligence, and particularly deep-learning approaches. This new field of pathology, termed **digital pathology,** will predictably transform and advance pathological diagnostics. The high resolution and wealth of visual and sub-visual data within a histological (or ultrastructural) image, the so-called Whole Slide Image (WSI), are unique and by far exceeds any other current imaging approach. Currently, the main bottleneck in the development of new approaches, methods and algorithms in digital pathology (mainly clinical but also preclinical), is the lack of access to relevant, well-curated, heterogeneous (i.e. multicentric) and large datasets (particularly relevant for the "data-hungry" deep learning approaches).

The main objective of NFDI4Patho is to tackle this unmet need by developing a national research data infrastructure making digital pathology data findable, accessible, interoperable, and reusable - following the FAIR data principles.

We have identified eight main areas necessary to achieve this goal (Figure 1).

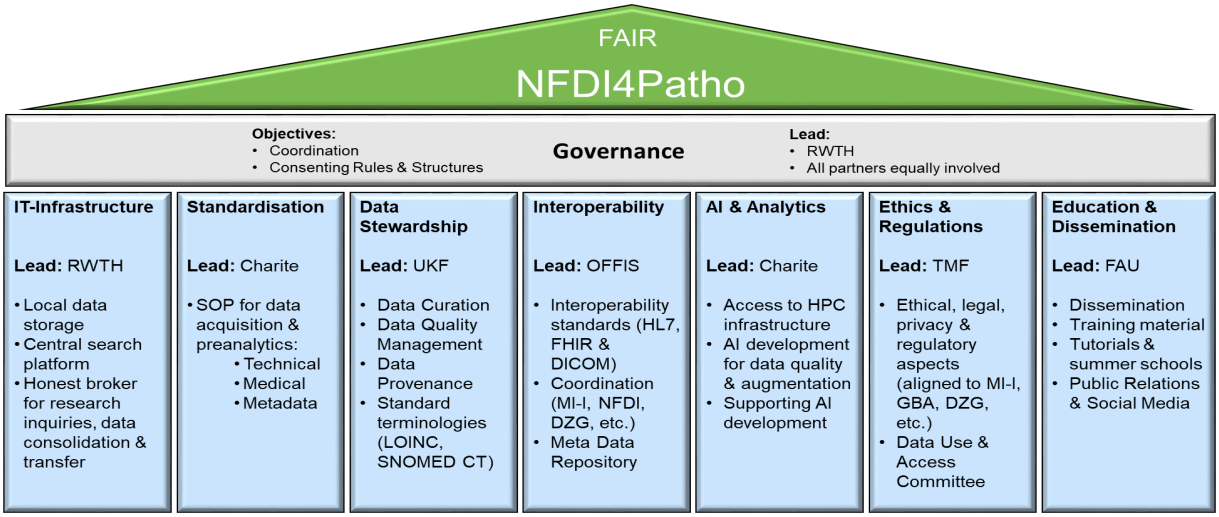

Figure 1: Structure of NFDI4Patho with its objectives within eight main areas.

Governance: Will coordinate the overall progress of the project and, in alignment and contribution of all partners, will develop consensus on rules, processes, sample documents, and standards for operations of NFDI4Patho.

IT-Infrastructure: Will design and implement a decentralized IT architecture with a central platform (office) for data requests & transfer using existing IT systems and components. The platform will handle the research requests, coordinate the review by the Data Use & Access Committee (DUAC), and provide the link to the data owners. After consenting by requesting researcher and data owners, the platform will merge and transfer the data. The processes will be based on experiences and best practices from the BMBF Medical Informatics Initiative (MII), German Biobank Alliance (GBA), AKTIN emergency department registry, the German Center for Lung Research (DZL), and the German Registry of COVID-19 Autopsies (DeRegCOVID).

Standardization: Will develop standard operating procedures (SOPs) for harmonized data acquisition and preanalytics. A strong focus will be put on standardization of various aspects, i.e. technical (stain and slide quality, slide scanners, file formats, etc.), medical (diagnostic accuracy, expert annotations, etc.) and metadata (acquisition, format, quality, etc.). These SOPs will be made available to the whole community, aiming for broad dissemination of standardization.

Data Stewardship: Will develop concepts for standardized and improved data collection and curation in the data generating centers in close cooperation with the previous task. The aim is to improve data availability and quality while at the same time optimizing the data collection effort. All relevant data will be included (clinical data, macroscopy, microscopy, pathology reports, molecular pathology, (multi)omics), etc). The coordinated processes are core elements in setting up data provenance.

Interoperability: Will develop the necessary communication standards DICOM and HL7 FHIR. These will be consented in the corresponding standardization organizations and will support the

standardization effort in the field of digital pathology. These developments will be done in close coordination with other initiatives will, in particular EMPAIA, MII, other NFDI consortia, or the German Centers for Health Research (DZG). The multidimensional image data, the clinical data (phenotyping), and the metadata must be annotated and modeled.
AI & Analytics: Will tackle and optimize access to experts, methods, and infrastructure (High-Performance Computing – HPC) in the field of AI & advanced image analytics in Pathology. The relevant centers with specific competencies will be identified and processes and procedures for access and support, where appropriate based on consented best practice solutions, will be developed.
Ethics & Regulations: Will tackle a variety of ethical and legal requirements when dealing with medical data. Developed solutions will be based on numerous preliminary works end experiences, that can be used in various research projects and will improve the quality and security of research. All relevant regulations will be included, i.e. the WMA Declaration of Helsinki, supplemented by the WMA Declaration of Taipei and the data protection legislation. A major goal of NFDI4Patho is the provision of data for the development of computer-aided decision support systems (CDSS), which fall under the medical device legislation regulated by the EU regulation 2071/745 (MDR). The requirements and solutions for the effective use of NFDI4Patho for the development and approval of CDSS will be tackled here (in close cooperation with EMPAIA).
Education & Dissemination: Will addresses the dissemination of the developed solutions. The implementation will also require the development of competencies in the participating (and other) centers (i.e. education & training concepts). Finally, public relations via press and social media will be tackled here.

- Brief description of the proposed use of existing infrastructures, tools and services that are essential in order to fulfil the planned consortium's objectives

The very specific technical (& medical) aspects of digital pathology are fully addressed by the specific and complementary expertise of consortiums' partners. **NFDI4Patho builds upon existing infrastructures and expertise and close coordination with major research initiatives MII, EMPAIA, DZG, GBA, NFDI consortia** (particularly NFDI4Health and GHGA). NFDI4Patho logic builds on horizontal networking of the individual locations to achieve vertical networking within the university locations through interoperability with the MII data integration centers and other networks. This will make the **data available for broad interdisciplinary research projects** (Figure 2).
The consortium involves **all major German professional societies** involved in digital pathology (with dedicated working groups), i.e. the Association of German Pathologists (BDP), German Society of Pathology (DGP), and the German Society of Neuropathology and Neuroanatomy (DGNN). This will enable a very high (perhaps even complete) coverage and broad implementation of NFDI4Patho (including education). Joining several additional centers is anticipated. The consortium already involves several **centers strongly involved in digital pathology**. Despite that this is still an emerging field, these centers have existing infrastructures

allowing digital pathology workflow. The consortium partners are also involved as Experts in the World Health Organization (WHO) Digital Health Technical Advisory Group (DHTAG), with a focus on digital pathology. NFDI4Patho will also profit from the existing high-performance computing (HPC) clusters on various sites, enabling advanced computational pathology approaches (e.g. JARA, FIAS, BIFOLD, MDC, etc.). NFDI4Patho will also align with the GAIA-X initiative, establishing transparent digital ecosystems on the European level.

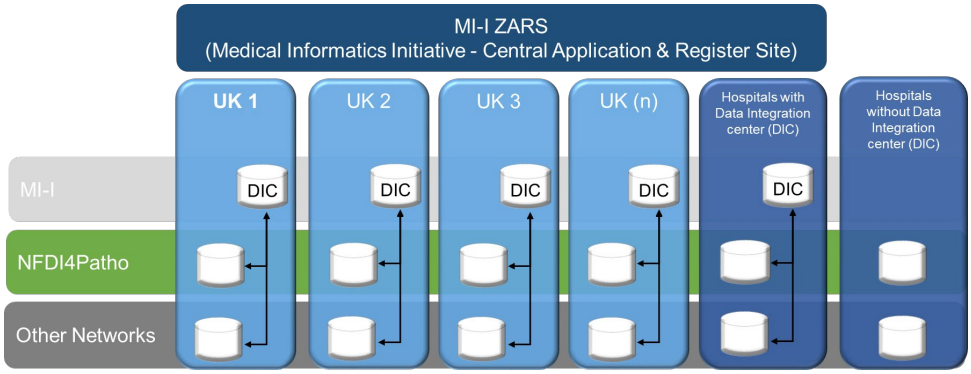

Figure 2: Integration of NFDI4Patho in the Research Data Infrastructure in Germany.

All of the "Pathology" partners and all the above mentioned professional societies are participating in the world-wide first national **COVID-19 Autopsy Registry** (DeRegCOVID, www.DeRegCOVID.ukaachen.de), which is centrally collecting data of all COVID-19 autopsies. More recently, a large consortium (DEFEAT PANDEMIcs), with 29 university centers and >60 institutes, was built around the registry (funded by the Network of University Medicine), among others focusing concepts for digital autopsy registry within the DeRegCOVID. These large initiatives will enable the access to autopsy samples and already provides strong evidence of **highly-collaborative, nation-wide work of the NFDI4Patho partners**.

Essential and close cooperation will be with the project **"Ecosystem for Pathology Diagnostics with AI Assistance" (EMPAIA)**, which has highly complementary aims with NFDI4Patho but lacks the access to research data which will be tackled by NFDI4Patho. This joined effort might uniquely facilitate the development and translation of AI into diagnostic pathology. NFDI4Patho has strong **expertise in Medical Informatics**, e.g. in software engineering, education, and development and maintaining research data infrastructures for various collaborative research projects, including the AKTIN emergency department registry and the DeRegCOVID. The OFFIS an institute of the University of Oldenburg has expertise in the area of **DICOM and interoperability standards**. The Technology and Methods Platform for Networked Medical Research (TMF), as an umbrella organization for medical collaborative research projects and networks in Germany, will provide its extensive expertise in concepts, checklists, and guidelines for NFDI4Patho.

- Interfaces to other proposed NFDI consortia: brief description of existing agreements for collaboration and/or plans for future collaboration

The very specific data sets of NFDI4Patho represent a unique added and complementary value for research data infrastructure in an emerging field, which is not addressed in any other initiative.

This will provide strong added value to other consortia. Although at this point, we do not have developed detailed interactions with the different NFDI consortia, after first contacts we expect to connect to different networks and cross-cutting topics, which will be mutually beneficial. Particularly, engaging with the existing platforms of **GHGA** and **NFDI4Health** will enable the access to the highly complementary pathology datasets (and “pathomics”), which are currently not tackled in any other initiative, in particular the MII. An integrative data-analytics together with the GHGA can be easily envisaged and is already ongoing in the frame of cooperation with partners of the GHGA consortium. We expect to elucidate the possible connections and interactions in detail within the preparation period of this project during the coming year. Similarly, in several sites, we are already engaged in building up concepts of close cooperation with the German Biobank Alliance/German Biobank Nodes - **GBA/GBN**.

4 Cross-cutting topics

- Please identify cross-cutting topics that are relevant for your consortium and that need to be designed and developed by several or all NFDI consortia.

At the current stage, and as outlined already above, we do not have established concrete cooperation with the different NFDI consortia and other cross-cutting topics, but we expect that such interactions are likely to evolve during the project preparation. Particularly expertise in preanalytics and standardization in image data and associated multidimensional metadata, development of standards (e.g. DICOM), expertise in computational pathology, and particularly AI development are some of the various aspects of potential interest for the consortia.