

NFDI 2021

Letter of Intent

1 Binding letter of intent as advance notification of a full proposal

<input checked="" type="checkbox"/>	Binding letter of intent (required as advance notification for proposals in 2021)
-------------------------------------	---

2 Formal details

- **Planned name of the consortium**
National Research Data Infrastructure for Microscopy, Biophotonics and Bioimage Analysis
- **Acronym of the planned consortium**
NFDI4BIOIMAGE
- **Applicant institution**
Heinrich Heine University Düsseldorf
Universitätsstraße 1
40225 Düsseldorf
Head: Prof. Dr. Anja Steinbeck
- **Spokesperson**
Dr. Stefanie Weidtkamp-Peters
stefanie.weidtkamp-peters@hhu.de
Center for Advanced Imaging
- **Co-applicant institution**
German Cancer Research Center (DKFZ)
Im Neuenheimer Feld 280
69120 Heidelberg
Head: Prof. Dr. Michael Baumann
- **Co-spokesperson**
Dr. Elisa May, Prof. (apl.)
elisa.may@uni-konstanz.de
(at DKFZ from Sep. 1st, 2021)

- **Co-spokesperson**
Dr. Jan-Philipp Mallm
j.mallm@dkfz-heidelberg.de
DKFZ Single-cell Open Lab

- **Co-applicant institution**
European Molecular Biology Laboratory (EMBL)
Meyerhofstr. 1
69117 Heidelberg
Head: Prof. Edith Heard, Ph.D.

- **Co-spokesperson**
Dr. Anna Kreshuk
anna.kreshuk@embl.de
Cell Biology and Biophysics Unit

- **Co-applicant institution**
University of Freiburg
Friedrichstraße 39
79098 Freiburg
Head: Prof. Dr. Kerstin Krieglstein

- **Co-spokesperson**
Dr. Björn Grüning
gruening@informatik.uni-freiburg.de
Freiburg Galaxy Team

- **Co-applicant institution**
Osnabrück University
Neuer Graben 29 / Schloss
49074 Osnabrück
Head: Prof. Dr. Susanne Menzel-Riedl

- **Co-spokesperson**
Dipl. Inf. Susanne Kunis
sukunis@uni-osnabrueck.de
CellNanOS, Department of Biology

- **Co-applicant institution**
University of Münster
Schlossplatz 2
48149 Münster
Head: Prof. Dr. Johannes Wessels
- **Co-spokesperson**
Dr. Thomas Zobel
thomas.zobel@uni-muenster.de
Imaging Network
- **Co-spokesperson**
Dr. Markus Blank-Burian
blankburian@uni-muenster.de
eScience WWU, Research Data Infrastructure

- **Co-applicant institution**
Leibniz Institute for Natural Product Research and Infection Biology,
Hans Knöll Institute (HKI)
Beutenbergstr. 11 a
07745 Jena
Head: Prof. Dr. Axel A. Brakhage
- **Co-spokesperson**
Prof. Dr. Marc Thilo Figge
thilo.figge@leibniz-hki.de
Applied System Biology

- **Co-applicant institution**
Leibniz Institute for Neurobiology (LIN)
Brenneckerstr. 6
39118 Magdeburg
Head: Prof. Dr. Stefan Remy
- **Co-spokesperson**
Pavol Bauer, Ph.D.
pavol.bauer@lin-magdeburg.de
Department Cellular Neuroscience

- **Co-applicant institution**
Technische Universität Dresden
01062 Dresden
Head: Prof. Dr. Ursula M. Staudinger
- **Co-spokesperson**
Dr. Robert Haase
robert.haase@tu-dresden.de
Bio-image Analysis Technology Development
DFG Cluster of Excellence “Physics of Life”

- **Co-applicant institution**
German Biolmaging - Gesellschaft für Mikroskopie und Bildanalyse e.V.
c/o University of Konstanz
78547 Konstanz
Head: Dr. Stefanie Weidtkamp-Peters
- **Co-spokesperson**
Dr. Stefanie Weidtkamp-Peters
stefanie.weidtkamp-peters@hhu.de
Chair of Board

- **Participant**
Dr. Astrid Schauss, University of Cologne
- **Participant**
Petra Hätscher, University of Konstanz
- **Participant**
Prof. Dr. Paul Czodrowski, TU Dortmund University
- **Participant**
Dr. Markus Becker, Leibniz Institute for Plasma Science and Technology, Greifswald
- **Participant**
PD Dr. Thomas Bocklitz, Leibniz Institute of Photonic Technology (IPHT), Jena
- **Participant**
PD Dr. Claire Chalopin, Leipzig University
- **Participant**
Prof. Dr. Jan Huisken, University of Göttingen

- **Participant**
Dr. Jan Bumberger, Helmholtz-Center for Environmental Research – UFZ, Leipzig
- **Participant**
Dr. Timo Dickscheid, Forschungszentrum Jülich
- **Participant**
Prof. Dr. Jan Lohmann, Prof. Dr. Vincent Heuveline, Heidelberg University
- **Participant**
Jianxu Chen, Ph.D., Leibniz Institut für Analytische Wissenschaften – ISAS, Dortmund
- **Participant**
Dr. Carsten Fortmann-Grote, Max Planck Institute for Evolutionary Biology, Plön
- **Participant**
Dr. Jean-Marie Burel, Open Microscopy Environment Consortium, University of Dundee, UK

3 Objectives, work programme and research environment

- **Research area of the proposed consortium** (according to the DFG classification system: https://www.dfg.de/dfg_profil/gremien/fachkollegien/faecher/index.jsp)
 - 201 Basic Research in Biology and Medicine
 - 202 Plant Sciences
 - 203 Zoology
 - 204 Microbiology, Virology & Immunology
 - 206 Neurosciences
 - 308 Optics, Quantum Optics and Physics of Atoms, Molecules and Plasmas
 - 409 Computer Science

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

NFDI4BIOIMAGE is a primarily method-centric consortium, addressing research data management (RDM) for bioimaging and biophotonics, which constitute key enabling technologies in the biomedical and the life sciences. The main focus is on the field of advanced imaging techniques, mostly light microscopy. The consortium moreover covers the RDM needs of bioimage informatics as a discipline-centric part. NFDI4BIOIMAGE aims to become the central reference entity for bioimaging data across all disciplines in the NFDI. The main objective is to enable researchers to produce FAIR imaging data in their everyday work. We aim

to support all steps of the bioimage data cycle, from image acquisition to bioimage analysis to publication and archiving. Furthermore, NFDI4BIOIMAGE intends to serve as a gateway for international bioimaging communities to the NFDI. To this end, it will build upon well-established international collaborations to align the work programme with global activities for bioimage data management. The consortium consists of experts from all relevant areas: i.e., software engineers for data management platforms, data formats, and metadata annotation; providers of IT infrastructure services for bioimage data handling and analysis; researchers specializing in multimodal data integration; and core facility leaders with substantial experience in user training and community integration. We designed our Task Area structure to optimally meet all requirements for establishing an integrative bioimaging data management ecosystem in Germany. The measures include developing generic solutions but also testing and adopting them in discrete community use cases.

The objectives in the Task Area **Image (meta)data formats and standardization** include establishing data acquisition standards and their representation in metadata, as well as minimal required metadata for bioimaging experiments along with controlled vocabularies and ontologies. These measures are addressed in collaboration with international partners. Another focus lies on the co-development of highly demanded cloud-compatible next-generation file formats suitable for object storage and dynamic access of large image data.

In the Task Area **Data linking and multimodal data integration**, interoperability of bioimage (meta)data and data management platforms with other data types and relevant databases is addressed. Discipline-specific data models (e.g., Brain Imaging Data Structure, BIDS, or Annotated Research Context, ARC) are explored as exchange formats for integration with image data management platforms and will be tested in discrete use case scenarios. The integration with electronic lab notebooks and laboratory information management systems is also a topic for this Task Area.

The discipline-centric focus of our work programme is set in Task Area **Bioimage informatics / bioimage analysis**. The main objective is to increase the reproducibility of bioimage analysis workflows both in interactive software tools and in highly automated (AI-based) analysis pipelines by developing accompanying data management best practices. These will be distributed as guidelines for FAIR bioimage analysis. Facilitated access to remote bioimage analysis computing environments, e.g., in Galaxy Imaging, and increased interoperability of different software tools will be established. The close collaboration with the Task Area **Technical infrastructure, access to archives and compute resources** will allow to set up the technical infrastructure required for testing and deploying cloud-based “software as a service” for bioimage data management and bioimage analysis. Guidelines for IT staff will be developed

to aid in setting up appropriate local computational resources. Seamless connection with public image archives and scenarios for long-term storage of large image datasets are addressed.

The work programme measures are designed to meet the bioimaging RDM needs of the community. In Task Area **Training and community integration**, we will foster increased exchange within the community by leveraging and expanding existing activities. We will build on established exchange forums and improve community outreach channels via websites and social media. Community use cases are collected to test and refine the deliverables and services of our consortium continuously. An open portfolio of educational resources on bioimage data handling and bioimage-analysis-associated RDM will be established, including community training and train-the-trainer workshops. Importantly, NFDI4BIOIMAGE aims to set up a team of image data management specialists for bioimaging-specific data stewardship. The consortium will commission these experts for direct support of community use cases.

Project management, governance, and continued evaluation of the project's progress are dealt with in the Task Area **Coordination and networking**. Collaboration with other NFDI consortia and aligning our work with international bioimaging RDM activities will ensure building an efficient and sustainable bioimaging data management ecosystem well-integrated within the international landscape. This Task Area will coordinate the exchange on cross-cutting topics with the sections of the NFDI.

- **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 NFDI4BIOIMAGE consortium builds upon German BioImaging – Gesellschaft für Mikroskopie und Bildanalyse (GerBI-GMB) and its network of >50 imaging core facilities spread across Germany. During more than ten years of activity, GerBI-GMB has created a well-appreciated platform for knowledge exchange for imaging scientists, has spearheaded novel training concepts like job shadowing for core facility staff, and has organized renowned scientific events (e.g., Trends in Microscopy Conferences). Due to the large community of users in imaging core facilities and in research laboratories, a significant amount of bioimaging data in Germany is produced in connection with GerBI-GMB. Therefore, NFDI4BIOIMAGE is in an excellent position to represent the large community of core facility users and providers and address specific needs of the bioimaging community in Germany. On the other hand, our consortium also fulfils the prerequisite to become a powerful distribution grid for bioimage RDM services and deliverables within the whole NFDI across disciplines. At the international level,

NFDI4BIOIMAGE leverages the well-established network of collaborators of GerBI-GMB including the Open Microscopy Environment (OME)¹ consortium, Euro-BioImaging², Global BioImaging³, Bioimaging North America (BINA)⁴ and more. The long-established working group of GerBI-GMB on image data management and analysis has recently been extended by the Research Data Management for Microscopy (RDM4mic) group. This group focuses on bioimaging data management based on OMERO (OME Remote Objects) in close collaboration with members from the OME Team at the University of Dundee. The RDM4mic group is open to the wider community beyond GerBI-GMB and is aiming to coordinate all activities around OMERO in Germany. NFDI4BIOIMAGE builds on connections to the IDR and BIA, two public endpoint repositories for image data hosted at EMBL-EBI, for which we aim to provide tools and services to ensure interoperability with local storage solutions in Germany. Members of GerBI-GMB have initiated a smaller-scale, OMERO-centered RDM infrastructure project, I3D:bio (DFG funding program: Information Infrastructure for Research Data), which will start approx. one year in advance of the envisioned NFDI4BIOIMAGE consortium. This project will closely interact with and complement the planned NFDI consortium. Members in NFDI4BIOIMAGE contribute proven experience in research management software engineering, metadata standardization, and file format development. For deployment of bioimage analysis tools and data management support, we will build on existing infrastructures like Galaxy, the OpenStack-based WWU cloud, and de.NBI, and extend established community exchange platforms (e.g., biii.eu in collaboration with members of NEUBIAS, or Elixir RDM-Kit in collaboration with OME). Our co-applicants and participants contribute substantial prior expertise on (image) data infrastructures (e.g., Galaxy Europe, Euro-BioImaging, or the novel EMBL Imaging Centre). Several are members in CRC-INF projects and Clusters of Excellence dealing with the topic of (image) RDM. Members of NFDI4BIOIMAGE are also involved in generic research data management projects, in part with leading roles, e.g., in the project bw2FDM, which produced forschungsdaten.info, the central German platform for research data management.

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

NFDI4BIOIMAGE provides, on the one hand, generic solutions that will benefit many disciplines in the biomedical and the life sciences. On the other hand, it will integrate discipline-specific aspects by collaborating closely with discipline-centered consortia within the NFDI.

¹ <https://www.openmicroscopy.org/>

² <https://eurobioimaging.eu/>

³ <https://globalbioimaging.org/>

⁴ <https://www.bioimagingna.org/>

Existing agreements for collaboration are:

- DataPLANT: Bridge the Annotated Research Context with bioimaging data management platforms to link (meta)data between local instances and archives. Align the development of metadata annotation tools (SWATE, MDEmic) for interoperability. Collaboration on data stewardship and virtual desktop infrastructure (University of Freiburg; CEPLAS Cluster of Excellence, Düsseldorf, and Cologne).
- GHGA: Spatial omics is an emerging transformative approach for linking nucleic acid sequencing or mass spectrometry data to image data providing localization in tissues or cells. Data linkage and interoperability of (meta)data between archives and management systems is elaborated in close collaboration with GHGA at our co-applicant institution DKFZ (applicant institution in GHGA).
- NFDI-Neuro: Cellular neuroscience depends on a number of microscopy techniques. We will adopt the established BIDS standard to capture microscopy data including all relevant information from sample preparation to image archiving in collaboration with FZ Jülich and LIN Magdeburg (co-spokesperson: Pavol Bauer).
- NFDI4Chem: Integration of chemistry data(bases) with bioimaging data (e.g., integration of spectroscopy data into image data management platforms). Adoption and interoperability of ELNs.
- InnoMatSafety: Define imaging metadata requirements for material safety testing, identify common vs. application-specific metadata and shared ontologies. Elaborate re-usable ELN modules for new and established (standard) operating procedures. Identify shared training needs and establish collaborative workshops.
- NFDI4Immuno: Immunological research combines various methods, e.g., measurements of T cell reactivities, antibody serum titers, and in-vivo imaging of immunological tissues. We work on shared data management workflows to link relevant data types across local management systems and federated repositories.
- DeBioData: Develop reproducible workflows and guidelines for handling high content imaging data with multi-feature annotation in pre-clinical drug development.
- NFDI4Health, NFDI4Patho: Assess connections between clinical imaging and biological/biomedical imaging: Patient data handling and record linkage, interoperability between medical imaging standards (e.g., DICOM) and bioimaging data formats (TIFF, ZARR).
- NFDI4Phys: Collaboration on physical image data acquisition quality and quality assessment, including their representation in metadata. Common interests are also issues of proprietary data formats and image data processing workflows. We aim to connect

(meta)data standards and platforms in different physics domains, including plasma science, biological and biomedical physics, with bioimaging-specific solutions.

- **NFDI4Microbiota**: Collaboration on microbiology-specific requirements for bioimage data and shared community training.
- **NFDI4Biodiversity**: Tools and services for image data handling in biodiversity, ecology and evolution studies in relevant use cases (e.g., UFZ Leipzig, University of Konstanz).

We envision further collaborations during the project phase that will develop based on use cases and interactions on cross-cutting topics.

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.**

NFDI4BIOIMAGE aims to support numerous research areas for bioimaging research data management and seeks to collaborate with other NFDI consortia on these aspects as well as on cross-cutting topics. Several topics collected in the Leipzig-Berlin-Declaration on Cross-Cutting Topics⁵, which are now being addressed in sections by the NFDI association⁶, are of particular interest for NFDI4BIOIMAGE, e.g.:

- (Meta)data formats and standardization
- Data provenance
- Research data commons, infrastructure long-term sustainability
- Vision and strategy of the NFDI and its integration in the scientific landscape
- Access management
- General and person-related legal and ethical aspects (e.g., GDPR)
- Cultural change, novel job profiles, and personnel recruitment

- **Please indicate which of these cross-cutting topics your consortium could contribute to and how.**

(Meta)data formats and standardization: Due to the complexity of image data and data sets generated with modern bioimaging techniques, classical file formats now constitute bottlenecks in bioimage data usage that must be addressed. Next-generation file formats (NGFF) enable new, cloud-compatible information structuring as is required for object storage, parallel access, and range requests. The expertise of NFDI4BIOIMAGE in co-developing NGFF holds potential

⁵ Bierwirth M. et al.: Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung, Zenodo. <https://doi.org/10.5281/zenodo.3895209>

⁶ Sure-Vetter, Y.: NFDI-Verein: Übersicht und Perspektiven, Plenarvortrag, NFDI-Konferenz 2021, 08.07.2021

for widespread use of NGFF throughout the NFDI. Additionally, expertise in metadata annotation tools and metadata formats exists at both a bioimaging-specific level and a generic level (e.g., state initiatives and working groups on metadata), and will be contributed to the NFDI as a whole.

Data provenance tracking relies on persistent identification, software versioning, documentation and reliable metadata interoperability, and data linkage. These aspects are an integral part of the work planned in our consortium, and we will collaborate on these topics with the respective NFDI section.

Our consortium includes academic IT infrastructure providers with substantial expertise in cloud engineering and SaaS deployment. Establishing a joint strategy for **sustainable infrastructures** (regarding both hardware and software) in the NFDI is an important topic to which our consortium could contribute. We would also like to emphasize the importance of (imaging) core facilities in the German (and international) research landscape. Central integration of these research support infrastructures into the long-term **strategy and vision of the NFDI** represents a topic to which NFDI4BIOIMAGE could deliver significant input. Similarly, data management and data stewardship expertise have to be anchored firmly in both generic research support teams and scientific core facilities. Topics like **access management** are at the heart of any scientific core facility. Finally, German imaging core facilities have been very active regarding the definition of **new job profiles** associated with scientific infrastructures, again a cross-cutting topic of high interest for the NFDI as a whole.