

# Letter of Intent

## NFDI Small Disciplines

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

*[Please indicate clearly whether your document is a binding letter of intent as advance notification or a non-binding letter of intent.]*

- Binding letter of intent (required as advance notification for proposals in 2019)
- Non-binding letter of intent (anticipated submission in 2020)
- Non-binding letter of intent (anticipated submission in 2021)

### 2. Formal details

**Planned name of the consortium:** NFDI Small Disciplines

**Applicant institution:** Humboldt-Universität zu Berlin, Prof. Dr. Sabine Kunst, Unter den Linden 6, 10099 Berlin

**Spokesperson:**

Prof. Dr. Gerd Graßhoff, E-Mail: [gerd.graßhoff@hu-berlin](mailto:gerd.graßhoff@hu-berlin), Institut für Philosophie, Humboldt-Universität zu Berlin

Prof. Dr. Andreas Degkwitz, E-Mail: [andreas.degkwitz@ub.hu-berlin.de](mailto:andreas.degkwitz@ub.hu-berlin.de), Universitätsbibliothek, Unter den Linden 6, 10099 Berlin

### 3. Objectives, work programme and research environment

**Research area of the proposed consortium (according to the DFG classification system):** 11 Humanities, 12 Social and Behavioural Sciences, 33 Mathematics, 45 Construction, Engineering and Architecture

#### ○ Concise Summary of the planned consortium's main objectives and task areas

The digital transformation is fundamentally changing the way researchers work in all fields. Individual projects become collaboration projects - research becomes faster and more dynamic for all user groups. Data collection, processing, analyses and publications can immediately be used by individuals or research groups and they themselves generate research data as part of the research process. In practice, the individual processing steps do not occur sequentially. The nature, benefits and characteristics of research data are surprisingly complex: Data occur in various media representations and contribute to the information value of the research objects. Research projects regardless of their size, are carried out by their supporting user groups: the researchers themselves, who benefit from the rapid exchange of their publications and research findings. Each future research project is expected to develop an extended Research Management Plan including a Data Management Plan. The services of the projected NFDI can be assumed for this planning. Internal and external research communication, data storage, data usage and data revision as well as full-page archiving up to the final publication of research results and research data form all part of the services offered by the NFDI.

○ **Brief description of the proposed use of existing infrastructures, tools and services**

The *NFDI Small Disciplines* integrates and offers institutional services of dozens of institutional partners in three major fields of activity:

- Collaborative Research Benches in cloud computing environments serve as virtual working environments for a large number of research projects and provide a variety of the best available digital humanities and data analytical tools as a hub of world-wide resources. It will provide super-computing power, big data processing and long-term archiving.
- It provides communications among research communities range from big research groups to individual actors and PhD-students, including the lay public, data providers, large archive holdings, museums and libraries. Projects register directly and quickly their needs and obtain a supportive infrastructure during their research phase. It will allow that research data will be directly published during funding period. It will provide access to large sets of data from each computing bench. Its organization will allow fast data import, linking, processing and archiving according to the FAIR principles. We expect the number of active research units quickly extend in the thousands.
- Scholarly and public communication of the supporting institutions will organize the dialogue between the research groups, special target groups and the lay public.

Typically, each of the *Small Disciplines* are represented by small units at a large number of universities or non-university research institutions. A new mapping initiated by the working group Small Subjects in Mainz counts 151 subjects of which 58 are listed in the StaBu system. In total, the small subjects can be divided into 19 groups, which in turn can be divided into six faculty cultures. These disciplines are the driving forces for innovative research and indispensable for the transdisciplinary transfer of knowledge. The Small Disciplines play a major role in research into civilizations, such as languages, ideas and conceptions or technical innovations, from prehistoric times, in the Ancient World until today. Small Disciplines like those of languages, literature, historical disciplines, material object studies, art history – including those from other faculties -- have a proven and long-standing scientific tradition related to digital research data. Traditionally, their projects are carried out globally; networking is essential. Research activities in the Small Disciplines depend on a number of contacts and exchanges in all regions and upon existing networks. Research projects in the Small Disciplines are conducted in a highly collaborative manner: for example, ancient Egyptian text sources can be collaboratively evaluated and contextualized with the help of philological and linguistic methods as well as with approaches from epigraphy, archaeology, scientific analysis methods in a new and more comprehensive way. The NFDI Small Disciplines will transform the large variety of research data from digital repositories to actively used and created modern digital research data.

- Redundant network of Big Data Centers incl. long term data archiving (NFDI Bridge, Zuse)
- Super-computing centers for parallel data processing and analytics
- Center for data services like Docker lakes, div. data services like OCR, Text mining, language models, Image recognition, 3D Processing
- Publication units in cooperation with publishers and publication services.
- Metadata harvesting
- Legal advisory unit
- Simultaneous quality management and activity monitoring

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

NFDI Small Disciplines has made agreements and plans for collaboration with NFDI4 Memory, NFDI4Objects, NFDI4Objects and BRIDGE4NFDI.

#### **4. Cross-cutting topics**

*NFDI for Small Disciplines* opens up, stores, transforms, connects and analyses research data for humanities research and other disciplines with an innovative Open Science infrastructure. Individual projects become collaboration projects, and therefore research activities becomes faster and more dynamic for all user groups. Data collection, processing, analyses and publications can immediately be used by individuals or research groups and they themselves generate research data as part of the research process. In practice, the individual processing steps do not occur sequentially. The nature, benefits and characteristics of research data are surprisingly complex: Data occur in various media representations and contribute to the information value of the research objects. Research projects regardless of their size, are carried out by their supporting user groups: the researchers themselves, who benefit from the rapid exchange of their publications and research findings. Each future research project is expected to develop an extended Research Management Plan including a Data Management Plan. The services of the projected NFDI can be assumed for this planning. Internal and external research communication, data storage, data usage and data revision as well as full-page archiving up to the final publication of research results and research data form all part of the services offered by the NFDI. Infrastructure units. The NFDI Small Disciplines is able to turn fully operational within 5 months for its broad range of already operating institutional support in the fields:

- Redundant network of Big Data Centers incl. long term data archiving (NFDI Bridge, Zuse)
- Super-computing centers for parallel data processing and analytics
- Center for data services like Docker lakes, div. data services like OCR, Text mining, language models, Image recognition, 3D Processing
- Publication units in cooperation with publishers and publication services.
- Metadata harvesting
- Legal advisory unit
- Simultaneous quality management and activity monitoring

The infrastructure of the NFDI interlinks excellent research achievements of small disciplines with an agile research network. It will be providing sustainable technical, computational, librarian, legal, and long-term archiving support for all researchers of Small Disciplines, according to the considerations of open science. The NFDI infrastructure integrates the latest research tools, including techniques such as cloud computing environments, popular computer libraries such as Scikit-Learn, Numpy, NLTK, distributed large data formats, LaTeX packages, parallel computer architecture, 3D object processing, and advanced deep learning techniques. In these cloud environments, any researcher can log on to workstations and obtain access to computing cloud environments that provide a wide variety of the best digital humanities and data analytic tools. We expect to offer hundreds of cutting edge technologies. Such an infrastructure goes far beyond the accessibility via Internet. Research data is part of the research process: Researchers evaluate, edit and version data and before re-using it for publication. It is not enough to store and archive data or to make it accessible in the long term. Even in the digital age, information must meet the requirements of knowledge review. Knowledge is efficiently gained when generated in a collaborative manner. In

being reusable, it contributes to future research projects and increases the scientific impact of research projects in the Small Disciplines. To this end, the research data infrastructure will become an active part of the digital research cloud, in which text as well as data publications or methods