NFDI4Memory – Letter of Intent

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

☐	Binding letter of intent (required as advance notification for proposals in 2019)
X	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**
  NFDI4Memory – historically engaged humanities

- **Acronym of the planned consortium**
  NFDI4Memory

- **Applicant institution**
  Leibniz Institute of European History
  Alte Universitätsstraße 19, 55116 Mainz

- **Spokesperson**
  Prof. Dr. Johannes Paulmann
  Director, Leibniz Institute of European History

3   Objectives, work programme and research environment

- **Research area of the proposed consortium**
  Humanities

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

As an innovative digital working environment, NFDI4Memory establishes a systematic and sustainable link between historical research, memory institutions (archives, libraries, museum, collections) and information infrastructures. The consortium addresses the disciplines that use historical methods or rely on data that need historical contextualization. This pertains to the historical disciplines, but also other disciplines in which the use of historical data is or has been part of the methodology, such as economics, social sciences, geography, and area studies. It assures the quality of research data in general, thereby safeguarding the historically critical function of the humanities as academic disciplines, a topic highly relevant to modern society. Source criticism adapted to digital data therefore is the principle methodological approach of NFDI4Memory.

The consortium connects the requirements of digital research data with established standards of analog research traditions, thus opening up new research paths while anchoring digital research firmly in the relevant disciplines. By this approach, it contributes to deepen and develop productively communication links existing between historically engaged research and Digital Humanities so that established methodological and ethical standards remain guaranteed in the digital age. NFDI4Memory aims at the integration of historical data of very different data formats resulting from specific application contexts and requirements. These includes, amongst others, texts ranging from antiquity to the modern era, images, photos, audio and video recordings, statistics, structured data, metadata, ontologies, or hypertexts. NFDI4Memory also focuses on personal data, spatial structures, and changes in classification systems, categories, and classifications over time. In terms of quality management and in order to have results acknowledged as scholarly output, the consortium also aims at establishing a specific system for the evaluation

of digital research in the form of digital reviews based on the high standard of peer review. NFDI4Memory is also particularly qualified through the participating disciplines to reflect on the methodological and epistemological dimensions of the digital revolution. As digital literacy is not widespread in the field of historically engaged disciplines, we need a structured way to enable historical basic research in the future.

NFDI4Memory pursues six main objectives that reflect major needs of historically working disciplines:

1) **Securing Quality Standards:** NFDI4Memory establishes quality standards and retrievability for connecting the different kinds of research data with information on origin, context, usage, rights and licenses. Research libraries, archives and other memory institutions as well as universities and non-university research institutions have already developed an expertise with regard to this challenging task. The consortium builds on projects and long-term funded initiatives and platforms which connect an increasing number of digitized manuscripts with metadata and relevant research literature. In order to address specific issues internationally across institutes as well as disciplines our consortium has been developing a task force on norm data that will guarantee interoperability within the NFDI as a whole.

2) **Implementing Data Linkage:** The systematic preparation and combination of research data and source data on the one hand, and the possibility, on the other hand, of examining and verifying research data of all kinds on the basis of analog, digitalized as well as and digital-born sources entails a great potential for innovative research. Such a systematic interlinking requires the development, efficient implementation and maintenance of hypermedia structures. These structures must be created as a new academic-led digital knowledge system based on ontologies that will enable a fundamental orientation in the digital knowledge space: different from commercial providers, an academic led ordering of knowledge will not only make "hits" visible but also bring to light gaps, i.e. material that has hitherto been accessible only to a limited extent, as well as lost and nonexistent information. The new digital ordering needs to be explicit, that means it must be able to be interrogated and analysed.

3) **Providing Digital Services:** The consortium, in close cooperation with the users, identifies necessary tools and services to support institutions in research data management and the establishment of corresponding technical and organizational structures. These are based on the requirements along the data life cycle and the FAIR principles. At the same time, requirements relating to the interoperability of services must be taken into account. Existing open and documented standards are the starting point (e.g. the International Image Interoperability Framework – IIIF). Therefore, the consortium will adapt existing services and tools and close existing gaps in the portfolio through new and further developed solutions. Important topics include persistent identification, rights management, semantic technologies as well as data analysis, annotation and enrichment. By providing best practice guidelines, interface specifications and software architectures templates, the consortium will also address the issue of the sustainability of tools and services and contribute significantly to the formation of structures in the digital humanities.

4) **Developing Digital Literacy:** NFDI4Memory develops specialized methodological innovation and activities in the field of digital literacy. The consortium critically reflects the relation between data, norm data, metadata, and knowledge with the aim to build on, develop and safeguard the specialized competencies of historical analysis.

5) **Creating a New Data Culture:** We address the ongoing need among communities working in the area of historical research to raise awareness of the relevant FAIR principles across all areas and phases of the research process. NFDI4Memory therefore creates a national node for providing support and material through training formats both for the scholarly user community and for societal transfer. The services we will provide are based on the principles of "teaching the teachers" and "peer-to-peer training". Continuing education and advanced training will also be offered to (junior) researchers to equip them with relevant skills and to ensure that adequate research data management measures can be implemented independently. Last but not least, it will also develop DH-curricula for universities. NFDI4Memory will

support university departments in embedding digital research and infrastructures in their training for students, PhDs, and postdocs and support the development of new career paths in the digital humanities. This is closely linked with the dissemination of best-practice examples for advanced training in the area of data literacy.

6) **Fostering International Co-operation and Data Integration:** NFDI4Memroy deepens the international networks of all consortium partners and expands these networks through new collaborations. They may build on networks established through Horizon2020, for example the European Research Infrastructure for Heritage Science (2016-2025), European Open Science Cloud, hypotheses.org, Europeana and Europeana pro as well as the Consortium of European Social Science Data Archives (CEESDA). Plans have also been made to integrate data from NFDI4Memory into the European Open Science Cloud (EOSC) thereby broadening the user base for history research with international contexts. Furthermore, there will be made connections with the large international initiatives RDA and Go FAIR.

These generic objectives will be the base line for creating the concrete work program in an iterative process with both the participants involved and the communities of scholars. The resulting task areas will include not only specific measures such as services and standards, but also tasks related to management, governance, and cross cutting aspects.

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

Metadata management services in infrastructure facilities already exist. This is not the case, however, for software development, data modeling of scientific issues, data curation, standards for data publication, research data cultures and corresponding services and tools which need to be updated and extended to fully cover the requirements of FAIR data management. Above all, there is a lack of established standards, norm data and classification models for data maintenance and description as well as the possibility of creating data transparency even for analogue data. Another particular challenge for research data management in historically engaged communities lies in the great diversity and structural variance of the underlying data models which require modular and flexible solutions that currently only exist in specific areas. Until now, there have only been isolated and primarily project-centered solutions to these challenges. As a result, there is still little interoperability of the research data generated and no commonly established consensus on how such data should be produced, standardized, integrated, re-used, stored, or published. Thus, the field requires a robust agreement on interoperable standards for quality-assured documentation of research processes.

In order to tackle this problem, NFDI4Memory will build on existing services. Furthermore, it will develop and implement automated tools for annotation using quality-assured measures like controlled vocabularies that have been agreed on both nationally and internationally. Organizational and technical solutions and services (for example the DIMAG network) already exist for the archiving of genuine digital archive material as well as research data repositories such as RADAR does for archiving secondary research data. As memory institutions, archives represent a central component in the historically engaged research landscape. They offer a wide range of interdisciplinary relevant research data in the form of analogue, digitized and born-digital archival material including related indexing information. This content is already freely accessible to research via various channels such as specialized information systems or comprehensive portals. Of particular importance is the Archivportal-D within the German Digital Library acting as "Data-Hub" and providing a central platform for primary research data that can be linked to secondary scientific research data (e.g. prosopographic analyses). Thus archives have been advising and accompanying researchers for centuries in their search for and evaluation of research data. At the same time, they have a broad knowledge and specific experience regarding the long-term preservation of these data while ensuring the authenticity of the sources - because it can only be guaranteed by archives. Archival content, services and expertise can therefore be usefully brought together in NFDI4Memory.

As for collections, in an advanced stage of the consortium, NFDI4Memory will be able to utilize the tools and information services developed by KultSam, currently under application

within the national research infrastructure programme. KultSam – Cultural Collections as Digital Repositories of Knowledge for Research, Teaching and Public Communication - is planned to digitally catalogue, preserve and network collection-related knowledge for a wide range of user groups in the cultural sciences. Within KultSam, special research centres are planned, aiming for consistently user-focused information services collaboratively bundled, catalogued and presented in a virtual research environment. NFDI4Memory and the NFDI as a whole may benefit from these. Comprehensive search portals of Specialised Information Services like Historicum.net or osmikon provide manifold options to integrate research data created within NFDI4Memory into the search space. Grown structures provided by library publication services will facilitate the integration of research data with traditional publication media. NFDI4Memory strives for interactively usable and dynamically combinable forms of publications that will enable researchers to make full use of the heuristic potential of digital media. Solutions for the long-term preservation of research data and new forms of publications can build on the well-established infrastructures as well as on the preservation related network Nestor.

User participation and involvement in NFDI4Memory will have two interrelated tiers in the form of use cases and in the provision of data literacy. Selected use cases will confront the challenges of highly heterogeneous research data, and will allow to map a broad cross-section of exemplary and, at the same time, innovative applications together with the needs that are arising in the historical disciplines with regard to digital practices. As part of the process of building NFDI4Memory, these cases will also be linked with other initiatives to develop work packages suitable for formulating and implementing generic and innovative solutions for research data management. For the second tier, data literacy and evaluation is a vital requirement in the humanities for the user communities of researchers, research data managers and research data users alike. Data literacy is also an important element in societal transfer with regard to reliability and trust in scholarly research.

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

NFDI4Memory will establish a network of all institutions and projects that work digitally in the area of the historical disciplines and bringing them into contact with one another in a subject-related context (e.g. collaboration with the Standardisation committee of the DNB, GESIS, Ost-Data, Manuscripta mediaevalia etc.). As documented in the Memorandum of understanding, NFDI4Memory will closely collaborate with three other initiatives, namely NFDI4Culture, NFDI4Text+, and NFDI4Objects. Here, we will coordinate our activities with project objectives, work stages, and milestones in both national and international contexts. NFDI4Memory will actively participate in the coordination group and strategic meetings already in 2019. There is also an agreement with the originally planned consortium ECONSOCHIST to integrate part of their objectives into NFDI4Memory. This initiative will serve as a strategic node for the cooperation with consortia in the social sciences.

4   Cross-cutting topics

The infrastructure partners of NFDI4Memory have long and extensive experience as well as high level of expertise in areas such as the indexing, processing, and authenticity of data, metadata, data authorship. They also possess profound skills in the development and operation of technical architecture for the purposes of hosting, retrieval, access provision, and long-term storage. In addition, they have relevant knowledge in legal and ethical matters of data management. This collective knowledge and expertise will have a synergetic effect for the further development of a dynamic and future-oriented digital research environment. There are, however, cross-cutting topics that are not only relevant for NFDI4Memory but also need to be designed and developed together with other NFDI consortia:

a) Standardization: It is only within the initiative of the NFDI that it is possible to establish standard data and controlled vocabulary through intensive cooperation with other consortia (e.g. in the context of the Authority Data Task Force). The same applies for the creation of a

digital knowledge system, which must be promoted and driven forward as a joint task and developed in close consultation and coordination with the specialist disciplines.

b) *Interoperability:* The consortium aims to ensure that interoperability is achieved through the use of open standards and structured metadata. Appropriate processes are to be established in the relevant specialist communities and further developed in exchange with other consortia. The consortium wants to ensure traceability and comprehensibility around the genesis of research data through a thorough documentation of research processes and, from this, to develop generic research data management plans for the different project scenarios (from source-based, individual research to interdisciplinary collaborative research). A further goal is to proactively support the setting up of repositories that provide research data and knowledge as well as to support the description of data through metadata enrichment, controlled vocabularies and ontologies with a view to facilitating the efficient exchange and effective reuse of research data and ensuring that it can be located and interpreted in the long term.

c) *Data ownership and data protection:* The widest possible accessibility and reusability of data, tools and matrices needs to be actively supported in accordance with data ownership and data security. In line with this, NFDI4Memory aims to develop processes that guarantee easy and sustainable access to research data based on the permanent availability of standardized research data. NFDI4Memory further aims to support the citability and thereby also the circulation of historical knowledge through the allocation of permanent identifiers (for example DOI, EPIC-Handle, ARK, URN). Lastly, NFDI4Memory also plans, to provide or pass on successful solutions to the target communities and to organize a structured system of exchange whereby experiences and ideas about risks and best-practice approaches in research data management can be shared.

d) *Authority data management:* Descriptive and provenance metadata brings research data of all kind to its full potential. It enables discoverability, interoperability, verifiability and trust. Moreover, metadata enables the exploitation of innovative tools and services. Metadata should make use of authority controls and ontologies as well as unique identifiers. To achieve these goals efficiently, NFDI4Memory in close cooperation with other consortia has to agree upon standardized metadata formats, which support machine-understandable semantics in order to enable large scale data integration ("Taskforce Norm Data"). Thereby, the consortium will be able to publish and consequently share multidimensional data according to the specific needs of historical research. Data provenance has to be considered and documented for human creators of data and metadata as well as for algorithms and services to be able to determine their reliability.