Letter of Intent

NFDI4Mobility – Nationale Forschungsdateninfrastruktur für Mobilitätsdaten

Deutsche Forschungsgemeinschaft
Kennedyallee 40 · 53175 Bonn, Germany · Postal address: 53170 Bonn, Germany
Tel.: + 49 228 885-1 · Fax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de

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

☒	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 Mobility Technology

- **Acronym of the planned consortium**

  NFDI4Mobility

- **Applicant institution**

  Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe

  Prof. Dr.-Ing. Holger Hanselka

- **Spokesperson**

  Prof. Dr. J. Marius Zöllner, marius.zoellner@kit.edu,

  KIT, Institute of Applied Informatics and Formal Description Methods (AIFB)

- **Co-spokesperson**

  Prof. Dr. rer. nat. Frank Gauterin, frank.gauterin@kit.edu,

  KIT, Institute of Vehicle System Technology

  Prof. Dr. Anne Koziolek, anne.koziolek@kit.edu,

  KIT, Institute for Program Structures and Data Organization

  Prof. Dr. Jörn Müller-Quade, crypto-info@iti.kit.edu,

  KIT, Institute of Theoretical Informatics

  Prof. Dr. Ralf Reussner, ralf.reussner@kit.edu,

  KIT, Institute for Program Structures and Data Organization

- **Co-applicant institution**

  Technische Universität Berlin (TUB), Straße des 17. Juni 135, 10623 Berlin

  Prof. Dr. Christian Thomsen

- **Co-spokesperson**

  Prof. Dr. Sahin Albayrak, sahin.albayrak@dai-labor.de,

  TUB, Distributed Artifical Intelligence Laboratory

  Prof. Dr. Ingo Weber, ingo.weber@tu-berlin.de,

  TUB, Chair of Software and Business Engineering

  Prof. Dr. Frank Straube, frank.straube@tu-berlin.de,

  TUB, Chair of Logistics

  Prof. Dr.-Ing. Stefan Tai, tai@tu-berlin.de,

TUB, Chair of Information Systems Engineering

- **Co-applicant institution**

  Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Fraunhoferstraße 1, 76131 Karlsruhe

  Prof. Dr.-Ing. habil. Jürgen Beyerer
- **Co-spokesperson**

  Prof. Dr.-Ing. Jürgen Beyerer, Juergen.Beyerer@iosb.fraunhofer.de,

  Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB

- **Co-applicant institution**

  FZI Research Center for Information Technology (FZI), Haid-und-Neu-Str 10-14, 76133 Karlsruhe

  Prof. Dr. Andreas Oberweis, Jan Wiesenberger, Prof. Dr. J. Marius Zöllner
- **Co-spokesperson**

  Dr. Alexander Viehl, Viehl@fzi.de,

  FZI, Technical cognitive systems

- **Co-applicant institution**

  German Aerospace Centre (DLR) Institute of Transport Research, Rudower Chaussee 7, 12489 Berlin

  Prof. Dr. rer. nat. Barbara Lenz
- **Co-spokesperson**

  Prof. Dr. Meike Jipp, Meike.Jipp@dlr.de,

  DLR, Institute of Transport Research

- **Co-applicant institution**

  Gottfried Wilhelm Leibniz Universität Hannover (LUH), Nienburger Straße 1, 30167 Hannover

  Prof. Dr. Volker Epping
- **Co-spokesperson**

  Prof. Dr. Christian Heipke, heipke@ipi.uni-hannover.de,

  LUH, Institute of Photogrammetry and GeoInformation

- **Co-applicant institution**

Technische Universität Braunschweig (TUBS), Mühlenpfordtstr. 23, 38106 Braunschweig

Prof. Anke Kaysser-Pyzalla

- **Co-spokesperson**

  Prof. Dr. Ina Schaefer, i.schaefer@tu-braunschweig.de,

  TUBS, Institute of Software Engineering and Automotive Informatics

- **Co-applicant institution**

  Technische Universität Darmstadt (TUDA), Otto-Berndt-Straße 2, 64287 Darmstadt

  Prof. Dr. Tanja Brühl

- **Co-spokesperson**

  Prof. Dr. Hermann Winner, winner@fzd.tu-darmstadt.de,

  TUDA, Institute of Automotive Engineering

- **Co-applicant institution**

  Technische Universität München (TUM), Arcisstr. 21, D-80333 Munich

  Prof. Dr. Thomas F. Hofmann

- **Co-spokesperson**

  Prof. Dr. Klaus Bogenberger, klaus.bogenberger@tum.de,

  TUM, Chair of Traffic Engineering and Control

- **Co-applicant institution**

  University of Cologne (UoC), Albertus-Magnus-Platz, 50923 Köln

  Prof. Dr. Dr. h.c. Axel Freimuth

- **Co-spokesperson**

  Prof. Dr. Andreas Vogelsang, vogelsang@cs.uni-koeln.de,

  UoC, Chair of Software and Systems Engineering

- **Co-applicant institution**

  Goethe University Frankfurt, Theodor-W.-Adorno-Platz 1

  60323 Frankfurt am Main

  Prof. Dr. Enrico Schleiff

- **Co-spokesperson**

  Prof. Dr. Indra Spiecker genannt Döhmann, LL.M. (Georgetown), spiecker@jur.uni-frankfurt.de,

Chair in Administrative and Constitutional Law, Environmental Law, Information Law and Legal Theory

- **Co-applicant institution**
  Leibniz Information Centre for Science and Technology (TIB), Welfengarten 1 B, 30167 Hannover
  Prof. Dr. Sören Auer
- **Co-spokesperson**
  Dr. Irina Sens, irina.sens@tib.eu,
  Head of Library Operations

- **Participant**
  Armin Rupalla, ASAM e.V.
- **Participant**
  Dr. Achim Bonte and Dr. Irina Sens, FID move, Specialised Information Service Mobility and Transport Research
- **Participant**
  Prof. Dr. Klaus Dietmayer, Ulm University (UU)
- **Participant**
  Prof. Dr. Franz Quint, University of Applied Sciences, Karlsruhe
- **Participant**
  Prof. Dr. Daniel Göhring, Freie Universität Berlin (FUB), Berlin
- **Participant**
  Dr.-Ing. Gotthard Meinel, Leibniz Institute of Ecological Urban and Regional Development (IOER), Dresden

3. Objectives, work programme and research environment

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

407-04 Verkehrs- und Transportsysteme, Intelligenter und automatisierter Verkehr
409-06 Informationssysteme, Prozess- und Wissensmanagement

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

Main Objectives:

The NFDI4Mobility initiative aims to create data management structures for research on ground-based mobility with detailed consideration of technical systems. A sound data repository for scientific use will provide combined access to data generated by research vehicles, test fields, digital infrastructures, mobile devices, etc. of research institutions in Germany and Europe. The

platform will allow for search and analysis at different levels of abstraction, from raw data to high-level descriptions of scenes and scenarios to support the development of technical systems as well as understanding mobility behaviour and systemic integration.

The most important objectives of the consortium are:

- Provide easy access to comprehensive data for mobility research
- Integrate physical test fields / trial sites related to mobility research, which will provide continuous data flow to the platform
- Realize compatibility of data sources, which will allow search, analysis, and data processing across multiple repositories
- Improve the compatibility of metadata schemas and data exchange interfaces in the mobility domain by contributing to the relevant standardization efforts
- Allow search for complex patterns and scenes in data on different levels from raw to highly aggregated data without violation of personal rights
- Ensure the availability and longevity of data on the NFDI4Mobility platform
- Support the reproducibility of research in the domain of mobility technology
- Maximize the impact of the platform by reaching a large community of users and stakeholders in the domain of mobility technology
- Provide training and active support to the community for maximizing the usability of the platform and accessibility of research data
- Improve the global visibility of German researchers and research institutes in the domain of mobility research

Task Areas (TA)

The task area **Architecture** will moderate the architecture design process. The architecture of the NFDI4Mobility platform shall be modelled and documented explicitly to coordinate the different activities and tools around the data lifecycle and the development of the platform. Making architecturally significant design decisions is a central crosscutting process where all task areas will contribute. The measures in this TA will moderate the process, make suggestions, evaluate alternatives and document the decisions. As a result, the architecture will define interfaces, component specifications, their dependencies and variations.

The purpose of the task area **Metadata Schema** is the identification of a metadata schema for the description of data sets relevant for the given NFDI4Mobility data context, that is "… ground-based transport and its travel behaviour aspects with detailed consideration of technical systems". For this, an inventory of data sources is needed. This inventory will cover both i.) already available and known data sources and ii.) expected future data sources. Based on this inventory

investigation, one or several metadata schemes will be assessed with respect to its applicability for the NFDI4Mobility platform.

The task area **Data Acquisition and Use Cases** will design, evaluate and update processes that allow an ongoing growth of the NFDI4Mobility platform. For this goal processes to establish contact with new data providers as well as standardization and refinement of the data acquisition process are done in this TA. Additionally, use cases and requirements from various stakeholders are also part of this TA to ensure maximal relevancy and usability of these datasets for various user groups and an expanding community.

The purpose of the task area **Data Preprocessing** is to process data streams from a single platform (a vehicle, a stationary traffic sensor system) upon upload, with the option to reprocess occasionally on demand if new requirements or new capabilities are introduced to the NFDI platform. The processing aims to assist cross-platform fusion performed in TA Consistency, and in particular the handling of user queries in TA Data Analysis, by precomputing relevant features. It also anonymizes privacy-critical data within the stream.

The NFDI4Mobility considers data coming from a variety of domains and sources, such as autonomous cars, simulation environments, real-world testbeds amongst others. For each of these origins, there are domain-specific analysis methods for varying user needs and applications. The challenge we are addressing in the task area **Data Analysis** is to determine the best suitable analysis method for a specific task in the mobile technology application domain and then to integrate these methods and their results to build human understandable analysis workflows beyond specific tasks across application (sub-)domains.

The goal of the task area **Data Use and Reuse** is to develop a sustainable operating model for the NFDI4Mobility platform. The measures in the task area are structured along the dimensions of Ecosystem Quality, Ecosystem Health, Ecosystem Governance, Ecosystem Storage, and Ecosystem Usage. In an additional measure, we will collect the results of the different dimensions and develop a sustainable operating model for the platform.

The purpose of the task area **Privacy** is to ensure that all data shared with the NFDI4Mobility platform appropriately respects data privacy. This concerns especially the privacy of all persons about whom information might be contained in the research data and who might not have given explicit consent for this data usage and where explicitly capturing consent is impractical. We will investigate how data can be captured and processed in a way that no information about an individual can be extracted from that data and how these measures can be applied in NFDI4Mobility.

During the design and operation of mobility systems, large amounts of heterogeneous planning and run-time data have to be handled. This data is created, collected, and manipulated independently of each other, so integration and consistency management are crucial. In the task

area **Consistency and Data Integration**, we will combine approaches from software engineering, database research as well as domain-specific methods from other engineering disciplines to manage consistency in mobility data.
Promoting and increasing the visibility of the platform and thereby increasing the adoption of the platform is one of the main objectives of the task area **Networking**. We will focus on seven different measures to achieve this goal. All project members will be actively involved in this task.
The task area **Governance & Coordination** deals with the set-up of all relevant rules, mechanisms and bodies to govern and coordinate the build-up and operation of NFDI4Mobility.

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

Infrastructures

Research Data Sources: *Test Areas:* Digital Test Field A9 (TUM), Test Area Autonomous Driving Baden-Württemberg – TAF-BW (FZI, KIT), Living Lab Future Mobility (FZI), digital test field SAFARI, Berlin (TUB), test road infrastructure for autonomous driving DIGINET-PS (TUB), extension of DIGINET-PS by BeIntelli-Project (TUB), AutoNV_OPR Ostprignitz-Ruppin (TUB), automated city driving environment "Stadtpilot" (TUBS), August Euler Air Field for automated vehicle and tram test (TUDA), Information and Communication Technology Infrastructure for Connected and Automated Road Transport - ICT4CART (UU). *Test Stands:* X-in-the-loop test stand network XiL-BW-e (KIT, UU), Vehicle test stands (KIT, TUB). Approx. 15 *Automated vehicles*: (IOSB, FUB, FZI, KIT, TU Berlin, TUBS, TUDA, UU). *Repositories:* Clearing House for Transport Data (DLR), German Mobility Panel MOP (KIT), KITTI Vision Benchmark Suite (KIT), TUBS Road User Dataset (TUBS), DriveU Traffic Light Dataset DTLD (UU) etc.
Computing resources for user access, data management services, and interlink of distributed repositories: Repository for archival and publication of research data, data privacy and anonymization RADAR (KIT), FORDATIS (IOSB), Steinbuch Centre for Computing (SCC, KIT), high performance computing cluster in Berlin etc.
Tools open to the community or as blueprint for NFDI tools: Helmholtz Analytics Toolkit – HeAT (DLR, KIT), sensor data management tool FROST-Server (IOSB), video and lidar data labelling and visualisation tool (KIT), labeling tool for object annotation in laser scans and videos (TUBS), full software stack for autonomous driving (FUB), data management plan creation tool RDMO (TUBS), Mobile Edge Computing based object recognition - MEC-View (UU) etc.
Services open to the community or as blueprint for NFDI services: Smart Data Innovation Lab provides hard- and software for the whole data lifecycle – SDIL (KIT), virtual machines - bwCloud SCOPE (KIT), research data management services - RDM@KIT (KIT), TUBS.researchdata

competence center offers data consultation, training and project support, and training material (TUBS) etc.

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

NFDI4Ing: The agreement for shared tasks includes annual workshops, joint working groups and coordination of community activities in the subject area transport and mobility. The focus of NFDI4Mobility will be on fine-grained, decidedly domain specific challenges of ground-based transport and its travel behaviour aspects, while NFDI4Ing will enable the integration of engineering-specific requirements in the mobility domain, including air-borne transport, with other engineering disciplines based on their common denominators.

MaRDI: We will jointly specify an interface to provide traffic and mobility data for use in their mathematical models. The data formats of the interface will cover all modes (public and private transport) with their supply data (road and rail networks, traffic control data etc.) and their demand data (traffic and passenger counts, origin-destination-matrices, etc.). The interface will provide functionality to search specific meta data according to classifications and ontologies suitable for the MaRDI domain context. It will be a good prototype for interfaces to other domain communities.

NFDI4Earth: The consortia of NFDI4Earth and NFDI4Mobility intend to collaborate in the agreements on Open and FAIR scientific data sharing as well as in the definition of Metadata profiles, Data Encodings and common vocabularies to allow easy usage of data for the respective communities. NFDI4Mobility could improve the research for predictive autonomous drive control by weather models and weather data as well as by detailed topography maps.

FAIRmat: The consortia will collaborate on the concepts and implementation of federated storage and archiving, data compression, search engines, and AI tools. We see a significant win-win situation by advancing the materials encyclopaedia and analytics toolkit based on vehicle and road infrastructure data in order to identify advanced materials that exhibit reduced ageing and/or tailored mechanical properties, improved sensor properties, or more efficient battery materials.

In addition to the above-mentioned initiatives, with which agreements have already been reached, consultations are planned with the NFDIxCS and NFDI4DataScience initiatives.

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.

Architecture for Distributed Research Data Infrastructure: A platform architecture to operate central tasks and to interlink local repositories of the partners will be necessary for the

NFDI4Mobility consortia. Other consortia may have similar needs which would be a basis for collaboration.

Data Privacy and Anonymisation: All research areas, which use data with possible relation to persons have the need to deal with data privacy issues and protect individual rights or persons effected by the data collection. Special challenges occur in relation to image and video data collected in the public domain but all other data sources with the possibility to reconstruct a relation to a person remain relevant. In general, research on all systems including the interaction with humans will need to deal with privacy questions. Research with such data requires tools and methods to ensure data privacy while maintaining the value of the data to address specific research questions. Anonymization is widely used as first measure but is only one out of several approaches.

Metadata and Standards: In recent years, a variety of general and domain specific metadata formats have been developed to store background information about respective data. But to combine multiple distributed data sources with varying uncertainties etc. in a research data infrastructure and keep the information accessible throughout the data processing pipeline, further development is still needed. The challenge is to find one or the smallest possible amount of existing metadata standards for the representation and description of the largest possible amount of data sets to have a lean and still expressive metadata scheme. A general subset of metadata describing e.g. the origin and terms of use of data could be very supportive for synergetic cross-domain data usage.

As a starting point, in many domains data standards and formats have been established, which often include definitions or suggestions for a subset of possible metadata.

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

The mobility sector generates a wide heterogeneous variety of data at different levels of abstraction, from simple source-destination relations to complex overall models of traffic situations resulting from sensor fusion. To a large extent, these data are not generated and collected in a controlled experimental environment, but in real traffic situations with high variance and hardly existing reproducibility. Therefore, to achieve reliable research results, the largest possible data base from wide-ranging recordings is usually used. However, this methodology results in data sets of fluctuating quality, and hinders comparability and general observability. The consortium therefore plans to deal intensively with the entire life cycle of the data from generation, processing and interpretation to deletion according to standards defined in the community and to generate methodological results that can also be transferred to other domains.

ASAM e.V. is one of the leading drivers of standards in the mobility sector. Therefore, the topic of standards development can be discussed and advanced from a purely scientific perspective as well as from the perspective of an organization responsible for the establishment and further development of standards.
Since mobility of persons in public space inevitably always describes personal data, safeguarding the personal rights of affected persons and protecting the data while at the same time maintaining its usefulness for research is a major challenge to which the consortium will devote the necessary attention. The approaches developed in this context are transferable to other domains that research of socio-technical systems.