

Letter of Intent

NFDI4MobilTech – Nationale Forschungsdateninfrastruktur für Mobilitätsdaten



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

<input type="checkbox"/>	Binding letter of intent (required as advance notification for proposals in 2020)
<input checked="" type="checkbox"/>	Non-binding letter of intent (anticipated submission in 2021)

2 Formal details

- **Planned name of the consortium**
National Research Data Infrastructure for Mobility Technology
- **Acronym of the planned consortium**
NFDI4MobilTech
- **Applicant institution**
Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe
Prof. Dr.-Ing. Holger Hanselka
- **Spokesperson**
Prof. Dr. J. Marius Zöllner, marinus.zoellner@kit.edu,
KIT, Institute of Applied Informatics and Formal Description Methods (AIFB)
- **Co-spokespersons**
Prof. Dr. rer. nat. Frank Gauterin, frank.gauterin@kit.edu, KIT, Institute of Vehicle System Technology
Prof. Dr. Anne Kozirolek, anne.kozirolek@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-spokespersons**
Prof. Dr. Sahin Albayrak, sahin.albayrak@dai-labor.de, TUB, Distributed Artificial Intelligence Laboratory
Prof. Dr. Frank Straube, straube@logistik.tu-berlin.de, TUB, Institute of Technology and Management

- **Co-applicant institution**
FIZ Karlsruhe – Leibniz Institute for Information Infrastructure (FIZ), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen
Sabine Brünger-Weilandt
- **Co-spokesperson**
Prof. Dr. Harald Sack, harald.sack@fiz-karlsruhe.de, FIZ Karlsruhe – Leibniz Institute for Information Infrastructure
- **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**
Prof. Dr. J. Marius Zöllner, zoellner@fzi.de, FZI, Technical cognitive systems
- **Co-applicant institution**
German Aerospace Centre (DLR) with Institute of Transportation Systems, Lilienthalplatz 7, 38108 Braunschweig and Institute of Transport Research, Rudower Chaussee 7, 12489 Berlin
Prof. Dr.-Ing. Katharina Seifert and Prof. Dr. rer. nat. Barbara Lenz
- **Co-spokesperson**
Prof. Dr. Katharina Seifert, katharina.seifert@dlr.de, DLR, Institute of Transportation Systems
- **Co-applicant institution**
Gottfried Wilhelm Leibniz Universität Hannover (LUH), Nienburger Straße 1, 30167 Hannover
Prof. Dr. Volker Epping

- **Co-spokesperson**
Prof. Dr. Christian Heipke, heipkeipi@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

- **Participant**
Dr. Klaus Estenfeld, 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

The list of (Co-)spokespersons and participants reflects the current state of planning and is subject to change during the course of the preparation of the proposal for 2021

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 NFDI4MobilTech 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, unified access to prior distributed comprehensive data for mobility research.
- Integrate continuous data acquisition from physical test fields /trial sites in mobility research.
- Enable search, analysis, data processing and refinement across multiple repositories.
- Contribute to standardization of metadata schemas and data exchange interfaces in mobility.
- Allow privacy-safe search for complex patterns and scenes in data on different levels.
- Ensure the availability and longevity of data on the NFDI4MobilTech platform.

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



- Support the reproducibility of research in the domain of mobility technology.
- Maximize the impact of the platform by reaching a large community.
- Provide training and active support to the community.
- Improve the global visibility of German mobility research.

Task Areas (TA)

Architecture: The overall architecture of the platform shall be modelled to coordinate the different activities and tools around the data lifecycle. Making architecturally significant design decisions is a central crosscutting process where all TA will contribute. 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.

Metadata Schema: This TA builds up an inventory of data sources covering existing and expected future data sources. Based on this inventory, one or several metadata schemes will be assessed with respect to its applicability for the NFDI4MobilTech platform and – if necessary- adjusted including user feedback.

Data Acquisition: This TA will design, evaluate and update processes to gain access to new and updated, relevant data sources. Processes to establish contact with new data providers and as well as standardization and refinement of the data acquisition process are the main goals.

Data Extraction & Aggregation: The TA will develop and implement a hierarchical database infrastructure and data management system that allows to represent, (pre-)process, and filter heterogeneous data from different sources on different abstraction levels in a unified way. It creates a connected, interpreted and privacy-safe assembly of the available research data.

Data Analysis: NFDI4MobilTech handles data coming from a variety of domains and sources with numerous analysis methods. This TA comprises measures to define a common reference framework for data analysis, discovery, composition and presentation techniques for analysis methods, visualization techniques for analysis results and the incorporation of user feedback.

Use & Reuse: The Goal of the TA is the implementation of an operating model for the NFDI4MobilTech platform. Therefore, it is necessary to elicit and analyse stakeholders, usage scenarios, quality requirements, and constraints and overcome the current barrier between available data and information needs of potential users. Important aspects include data privacy, data quality, and accessibility for documented use, scientific evaluations, or use in teaching.

Storage & Archiving: To find, share and reuse data inside the community the research data is stored together with necessary metadata. Access is regulated in an organizational as well as technical way. To ensure and maintain sufficient data quality, suitable data curation strategies are devised, implemented, and - if necessary - enforced. Long term storage and accessibility of MobilTech research data is planned and implemented according to its specific needs.

Security & Access Control: This TA will select appropriate security measures to be used in system design and develop tailored solutions where standard measures cannot fulfil all requirements to protect confidential data as well as prevent unwanted data modification.

Consistency & Data Integration: During the design and operation of mobility systems, large amounts of heterogeneous data have to be handled. This TA combines methods and technologies from software engineering, database research and domain-specific methods from other disciplines to create unified approaches for managing consistency in mobility data.

User Involvement & User Orientation: This TA aims to maximize the relevancy and the usability of the NFDI4MobilTech platform by reaching out to various stakeholders. The goals are to integrate user feedback and requirements, ensure usability and accessibility, provide guidance and support and organise community integration activities such as challenges or surveys.

Networking: This TA promotes and increases the visibility of the platform to increase its adoption in the community. This involves measures like annual conferences, involvement in standardization organisation, cooperation with other NFDIs, involvement in cross-cutting activities, online portals and other dissemination measures.

Governance & Coordination: The TA Governance & Coordination deals with the set-up of all relevant rules, mechanisms and bodies to govern and coordinate the build-up and operation of NFDI4MobilTech.

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), Application Platform for Intelligent Mobility – AIM (DLR), Test Bed Lower Saxony, (DLR), smartphone mobility data MovingLab (DLR), Test Area Autonomous Driving Baden-Württemberg - TAF (FZI, KIT), Karlsruhe Applied Machine Learning Living Lab – KarAMEL (FZI), Living Lab Automotive & Living Lab smartMobility (FZI), digital test field SAFARI, Berlin (FUB), test road infrastructure for autonomous driving DIGINET-PS (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, TU Berlin). *Approx. 15 Automated vehicles:* (DLR, IOSB, FUB, FZI, KIT, TU Berlin, TUBS, TUDA, UU). *Repositories:* Clearing House for Transport Data (DLR), Data sets (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: AIM Backend (DLR), Repository for archival and publication of research data, data privacy and anonymization RADAR (FIZ), 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 tool FROST (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 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 NFDI4Mobiltech 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 intend to collaborate in the definition of metadata standards and data encodings for spatial data. In addition, NFDI4Earth could benefit from NFDI4MobilTech data such as temporal and spatial distribution of humidity or particle emissions on roads collected by vehicle and road infrastructure sensors to validate weather forecast or particle transport models or vehicle camera data for detailed environment topography recognition. NFDI4MobilTech could improve

predictive autonomous drive control by microscopic weather models and data as well as by detailed maps of ground topography.

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.

Forum X: The future of mobility involves more digitalization, more connectivity, and higher automation. While these seem all technological, an even more important and cross-cutting dimension in mobility transformation involves social and economic factors, which are critical in the acceptance and adoption of new technologies. In this context, the joint work of ForumX, which addresses experimental data, and NFDI4MobilTech will jointly focus on ensuring a comprehensive coverage of research data and related practices from both technological and social dimensions. This will particularly involve the development and adoption of compatible metadata standards and data exchange interfaces across the two consortia.

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 NFDI4MobilTech 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. Do 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 has 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.

Architecture for Distributed Research Data Infrastructure: Aiming for the third round, we intend to pick up on generic architecture developments already started in other consortia. Domain specific challenges in mobility will require certain architectural developments and adaptations. Separation of generic and specific aspects of the architecture enables the reuse of generic parts in other consortia with similar boundary conditions. An example might be a framework for the handling and processing large datasets from numerous data sources (sensors, cameras etc.) with heterogeneous and varying quality, which is characteristic for mobility data resulting from long time observation in test fields etc. We plan to deal intensively with the whole data lifecycle from creation to deletion according to standards defined in the community and to generate transferrable methodological results. We will seek for a permanent exchange and coordination with other NFDI consortia on architectural questions in order to identify and use synergies.

Data Privacy and Anonymisation: The goal is the best compromise between a maximum of information preserved within the data for research and the interest to protect individual privacy. The constant handling of big data streams requires highly automated procedures. Example are ML-based methods for automatic anonymisation of image data by blurring or removing license plates and faces or for finding and removing characteristic patterns of movement or specific combination of attributes to avoid conclusions on specific persons. All measures need to include an estimation of effectiveness. Another objective is to develop a concept for handling all input data that cannot be anonymised sufficiently, either because technical means of anonymisation are not yet available, or because raw data is intended to be kept for later processing steps. Those developments for privacy-safe datasets preserving a maximum of information will likely be adaptable in other science disciplines and hence be offered to other consortia for reuse.

Metadata and Standards: It is planned to coordinate decisions and adjustments or developments of metadata formats and standards with other NFDI consortia and the metadata community. The work on the metadata concept will be constantly reported and synchronized against existing standard initiatives. In case that existing national and international standards will be used or adapted to specific needs of NFDI4MobilTech coordination with corresponding standards' organisations will be organised. This aims at a close NFDI4MobilTech cooperation with standards' organisation. They will be informed about the application and adjustment of corresponding standards and its combination with other domains' metadata standards. ASAM e.V., one of the main drivers of standardisation concerning mobility and automation data, is part of the consortium to ensure the closest exchange possible and integrate their experience in successfully defining industry relevant standards into NFDI4MobilTech.