

## National Research Data Infrastructure for Engineering Sciences (NFDI4Ing)

*Spokesperson: Robert Schmitt, RWTH Aachen University, contact@nfdi4ing.de*

- **DFG research areas 41-45**

41 Mechanical and Industrial Engineering, 42 Thermal Engineering/ Process Engineering, 43 Materials Science and Engineering, 44 Computer Science, Systems and Electrical Engineering, 45 Construction Engineering and Architecture

- **Types of data**

Research, development, and production in engineering is inherently diverse. In computational engineering, for instance, numerical data is generated and evaluated, while other engineering disciplines focus on (data) samples generated in pilot plants, laboratories, or in the environment. Both individuals and teams as part of distributed networks may compile data. The types of data generated and their further use are just as variable as the numerous types of engineering activities. For example, one engineer may work with continuously new versions of the same data set or experiment to create series of measurements or new configurations, while another engineer may modify a data set through many different processing steps – or a data set is passed from one research group to the next to be used and modified, there. The amount of data generated can range from a few bytes to petabytes. In addition, data can be very specific to individual cases (e.g., a calibration curve for a specific apparatus) and very short-lived (e.g., real-time EDR data). In the case of industrial cooperation, research data is also often subject to confidentiality and access must be controlled on different levels of confidentiality.

- **Background, international collaborations, and institutions**

NFDI4Ing brings together pioneers in data FAIRness and in developing research data management (RDM) governance and services, especially for the engineering community. The consortium is composed of 11 (co-)applicant institutions and 55 participants, including the entire TU9, universities of applied sciences, research associations, Clusters of Excellence, Collaborative Research Centres, and Transregio projects. On the international level, NFDI4Ing is collaborating with RDA RDM4Eng IG, IDSA, CESAER, AARC2, DataCite, EOSC, EUDAT CDI, and with several FAIR working groups and initiatives, such as Force11, EOSC FAIR, FAIRsFAIR, and GO FAIR.

The (co-)applicant institutions are RWTH Aachen University, TU Braunschweig, TU Darmstadt, TU Dresden, the German Aerospace Center, Leibniz University Hannover, TIB Hannover, Forschungszentrum Jülich, Karlsruhe Institute of Technology, TU Munich, and University of Stuttgart. Important infrastructures, such as the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities and the High-Performance Computing Center Stuttgart, also contribute to NFDI4Ing.

- **Specific requirements for RDM and how to address them**

NFDI4Ing has opted for a research method-oriented approach to meet the requirements of its community of interest. As heterogeneous as the activities and types of research data of engineers are, so are the demands for RDM. The vast variety of engineering problems addresses specific research (sub-)areas and results in highly specialised, individualised solution approaches. At the same time, strong interdisciplinary similarities can be identified at

the level of the research methods and processes used. NFDI4Ing's archetypes, from ALEX (bespoke experiments) to GOLO (field data), harmonise these commonalities and represent the diversity of RDM needs in engineering research. They are the result of a bottom-up process in consortium formation, serve as structure-giving elements, and represent broad classes of RDM requirements (see figure, below).



The overarching task area Base Services will provide central services for the archetypes and will address the following challenges:

- **Data quality assurance and metrics for FAIR data:** A service to provide a commonly accepted set of quality criteria and approaches including the according data management planning to achieve a quality standard that is accepted by the engineering community.
- **Research software development,** with a focus on supporting the development of sustainable research software prototypes and products for engineering sciences.
- As **harmonised metadata standards and a terminology service for the semantic description of research workflows** are required for most NFDI4Ing archetypes, a service for the development, evaluation, and maintenance of metadata and interoperable specialist vocabularies and ontologies in the engineering sciences will be provided.
- **Data storage, archives and repositories:** Services to support best practices and tools suitable for storage, exchange, publication, and long-term archival of data of varying quality and volume will be provided.
- **Data security and sovereignty:** Services to support best practices for securely providing and accessing (meta-)data in a distributed architecture operated by different scientific communities will be provided.

- **Community-based training on enabling data-driven science and FAIR data:** A service to support common practices to be adapted across engineering disciplines. It will provide discipline-specific trainings tailored towards specific challenges like handling large amounts of data or transitions between empirical and simulation-based research environments.
- **Text and data mining:** As engineering research data is still often ‘hidden’ in scientific literature such as articles and technical reports, a service to support data science in the context of enrichment of unstructured data with structured elements and of enabling text and data mining will be provided.

- **Integration of services into the overall NFDI**

To allow the scalability of services, the NFDI4Ing partners have agreed to follow a set of standards, best practices and requirements to establish the planned services and support them in the long-term. These standards are formulated as Integration Readiness Levels:

- Level 1: The service includes documented best practices, training materials, standards or other community specific features (e.g. metadata standards, specific DMP templates).
- Level 2: The service includes harmonised components and open interfaces.
- Level 3: The service is part of our discipline-specific services or contributes essential components to it. It supports theming, branding, and multi-client capability within NFDI4Ing.
- Level 4: The service is part of the NFDI vision of integrated services & data within the entire German research data landscape (NFDI4X) and within the EOSC.

- **Relevant cross-cutting issues**

NFDI4Ing has signed the Berlin Declaration on NFDI Cross-Cutting Topics and thereby joins the other signatories in the effort of addressing cross-cutting topics in a coordinated fashion. According to point three of the common vision of the Berlin Declaration, all of the listed topics are relevant for NFDI, but not every topic has the same relevance for each consortium. The topics we identified as most relevant for NFDI4Ing are:

<ul style="list-style-type: none"> <li>- Governance</li> <li>- Community integration</li> <li>- Community support</li> <li>- Interaction with already existing RDM services in institutions</li> <li>- Digital twins</li> <li>- Technical infrastructure</li> <li>- Quality assurance in RDM processes and metrics for FAIR data</li> <li>- Research software management</li> <li>- Metadata and terminology services</li> <li>- Repositories and storage</li> </ul>	<ul style="list-style-type: none"> <li>- Persistent identifiers (PID)</li> <li>- Authentication and Authorisation Infrastructure (AAI)</li> <li>- Overall NFDI software architecture – data security and sovereignty</li> <li>- Automated data and knowledge discovery in literature and structured data</li> <li>- Community-based training on enabling data-driven science and FAIR data</li> <li>- Training and dissemination activities comprising education and training materials</li> </ul>
--	--

- **Expectations for the NFDI conference**

We are anticipating an increase in connecting links between consortia throughout the course of the conference. We are particularly interested in contributing to the establishment of NFDI’s structures, one example being NFDI’s content design with respect to cross-cutting issues (i.e., whether the NFDI should finally be established with or without dedicated cross-cutting consortia, how coordination will take place, how responsibilities will be shared, etc.). We are looking forward to this networking event between consortia, the NFDI directorate, and the DFG!

- **Designated members of the consortium**

<i>(Co-)Spokesperson</i>	<i>Participating institution</i>
Verena Anthofer Head of Department Applications, University IT-Service and Computing Centre verena.anthofer@tu-darmstadt.de	TU Darmstadt Karolinenplatz 5 64289 Darmstadt
Sören Auer Director of TIB, Head of Research Group Data Science and Digital Libraries soeren.auer@tib.eu	TIB - Leibniz Information Centre for Science and Technology University Library Welfengarten 1B 30167 Hannover
Torsten Bronger Research Data Management - FZJ Project Leader t.bronger@fz-juelich.de	Forschungszentrum Jülich Wilhelm-Johnen-Straße 52428 Jülich
Bernd Flemisch Coordinator of the Research Area "Discretization, Solvers and Research Software" at the Institute for Modelling Hydraulic and Environmental Systems bernd.flemisch@iws.uni-stuttgart.de	University of Stuttgart Keplerstraße 7 70174 Stuttgart
Regine Gerike Head of Chair of Integrated Transport Planning and Traffic Engineering regine.gerike@tu-dresden.de	TU Dresden 01062 Dresden
Manfred Krafczyk Head of Inst. for Computational Modeling in Civil Engineering / Vice President Digital Affairs & Techn. Transfer kraft@irmb.tu-bs.de	TU Braunschweig Universitätsplatz 2 38106 Braunschweig
Roland Lachmayer Head of Institute of Product Development lachmayer@ipeg.uni-hannover.de	Leibniz University Hannover Welfengarten 1 30167 Hannover
Christian Langenbach Research Data Manager and Deputy Research Information Christian.Langenbach@dlr.de	German Aerospace Center Linder Höhe 51147 Köln
Matthias Müller Director IT Center, Chair for High Performance Computing mueller@itc.rwth-aachen.de	RWTH Aachen University Templergraben 55 52062 Aachen
Britta Nestler Director of the Institute of Applied Materials - Computational Materials Science, KIT and of the Institute of Digital Materials Science at Karlsruhe University of Applied Sciences britta.nestler@kit.edu	Karlsruhe Institute of Technology Kaiserstraße 12 76131 Karlsruhe

Peter F. Pelz Head of Institute for Fluid Systems peter.pelz@fst.tu-darmstadt.de	TU Darmstadt Karolinenplatz 5 64289 Darmstadt
Robert Schmitt Managing Director of Laboratory for Machine Tools and Production Engineerinmueller@itc.rwth-aachen.de	RWTH Aachen University Templergraben 55 52062 Aachen
Annett Schwarz Research Data Management Coordination schwarz@itc.rwth-aachen.de	RWTH Aachen University Templergraben 55 52062 Aachen
Irina Sens Head of Library Operations, Head of Programme Area Collection, Development and Metadata irina.sens@tib.eu	TIB - Leibniz Information Centre for Science and Technology University Library Welfengarten 1B 30167 Hannover
Thomas Stäcker Director of the University and State Library Darmstadt direktion@ulb.tu-darmstadt.de	TU Darmstadt Karolinenplatz 5 64289 Darmstadt
Christian Stemmer Management of the Chair of Aerodynamics and Fluid Mechanics christian.stemmer@tum.de	Technical University of Munich Arcisstraße 21 80333 München
Detlef Stolten Director of the Institute of Energy and Climate Research – Techno-economic Systems Analysis (IEK-3) d.stolten@fz-juelich.de	Forschungszentrum Jülich Wilhelm-Johnen-Straße 52428 Jülich
Achim Streit Director of the Steinbuch Centre for Computing achim.streit@kit.edu	Karlsruhe Institute of Technology Kaiserstraße 12 76131 Karlsruhe