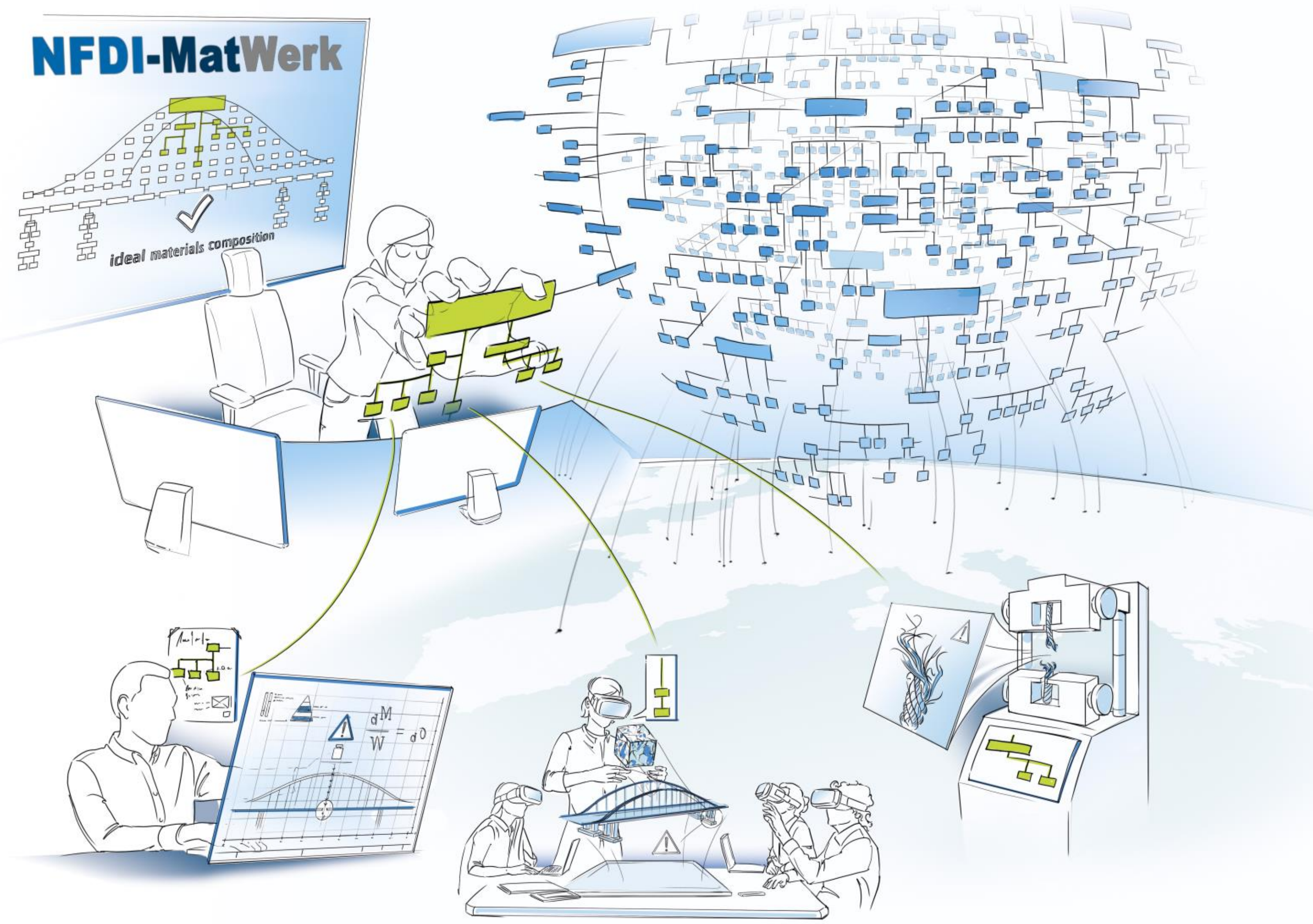

1 Binding letter of intent as advance notification or non-binding letter of intent

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

2 Formal details

- Planned name of the consortium:

  DE: Nationale Forschungsdateninfrastruktur für Materialwissenschaft & Werkstofftechnik

  EN: National Research Data Infrastructure for Materials Science & Engineering

- Acronym of the planned consortium

  NFDI-MatWerk

- Applicant institution

  Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
  Hansastraße 27 c
  80686 München
  Head of the Institution: Prof. Dr.-Ing. Reimund Neugebauer, President

- Spokesperson

  Prof. Dr. Christoph Eberl
  chris.eberl@iwm.fraunhofer.de
  Deputy Director Fraunhofer Institute for Mechanics of Materials, IWM
  Wöhlerstraße 11,
  79108 Freiburg

- Co-applicant institutions and Spokespersons

Friedrich-Alexander-Universität Erlangen-Nürnberg
Prof. Dr. Erik Bitzek
Martensstr. 5a
91058 Erlangen

Karlsruher Institut für Technologie
Prof. Dr. Peter Gumbsch
Straße am Forum 7
76131 Karlsruhe

Max-Planck-Institut für Eisenforschung GmbH, MPIE
Dr. Tilmann Hickel
Max-Planck-Straße 1
40237 Düsseldorf

Universität des Saarlandes
Prof. Dr. Frank Mücklich
Campus D33
66123 Saarbrücken

RWTH Aachen
Prof. Dr. Matthias S. Müller
Templergraben 55
52062 Aachen

FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH
Prof. Dr. Harald Sack
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen

Technische Universität Bergakademie Freiberg
Prof. Dr. Stefan Sandfeld
Lampadiusstrasse 4
09599 Freiberg

Forschungszentrum Jülich GmbH
Prof. Dr. Ruth Schwaiger
Wilhelm-Johnen-Straße
52428 Jülich

Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI
Prof. Dr. Philipp Slusallek
Stuhlsatzenhausweg 3
D-66123 Saarbrücken

Karlsruher Institut für Technologie
Prof. Dr. Achim Streit
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen

Fraunhofer-Institut für Werkstoff- und Strahltechnik
Prof. Dr. Martina Zimmermann
Winterbergstr. 28
01277 Dresden

3 Objectives, work program and research environment

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

https://www.dfg.de/download/pdf/dfg_im_profil/gremien/fachkollegien/amtsperiode_2020_2023/fachsystematik_fachkollegien_amtsperiode_2020-2023.pdf)

This Consortium applies for funds in the research areas

405 – Werkstofftechnik and

406 – Materialwissenschaft

according to the DFG classification system.

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

Since the Stone Age the mastery of materials has always played a key role for the progress of our society and economy. Today the study of materials lies at the heart of the relatively young **research discipline Materials Science & Engineering (MSE)**. The discipline aims to develop novel materials, characterize, optimize and model them, develop processes to manufacture such materials, evaluate their lifetime and develop strategies to maximize the reusability at the end-of-life. The aforementioned aspects relate to a wide spectrum of experimental and in silico data. Such data is available in various formats, originating from highly heterogeneous sources and, most often, the data and its properties depend on the full history of the material. One of the challenges concerning the plethora of data in MSE is their inherent multiscale character with various chemical and physical phenomena on different time and length scales. This is reflected by the strongly heterogeneous microstructures present in virtually all materials, which determine the mechanical properties and a substantial fraction of the functional performance. The "microstructure" constitutes the local chemical composition and the local spatial arrangement of different phases and crystalline areas, which typically contain various crystal defects. Realistic configurations range from defects at the atomistic level, through micron scale phases up to macroscale pores. Furthermore, any process (e.g., thermal treatment or mechanical loading) applied to a sample or a specimen changes the materials microstructure and, thereby, the related material performance. This strong path dependence necessitates the formal description, the experimental access and the capturing of data along the full material history. Only then, the reliable interpretation, more importantly, the robust prediction of the material behavior can potentially be achieved.

Therefore, the materials-specific digital infrastructure, as anticipated by the NFDI-MatWerk, has to be able to capture the various states of the material on different time and length scales. Tracking the evolution of the material data over the course of its lifetime defines another challenge. Furthermore, the data space has to be sufficiently rich to track the various highly complex dependencies between materials properties, path dependence and the local microstructure (e.g. regions of different chemical composition or crystal structure). The ultimate objective is the prediction of the correlation of the initial microstructure and the processes applied to the material with the final state of the material and the resulting multi-physical properties. Therefore, the NFDI-MatWerk consortium proposes to set up a digital platform, which will take this complex hierarchical character of materials into account, enabling synergies and reducing the technological barriers within the community. To date, NFDI-MatWerk has identified **five task areas** for proceeding with this goal. Each task area (TA) is coordinated by a minimum of **one field expert and one materials scientist** from the group of co-applicants. The task areas and their objectives are listed below.

- **Task Area Strategy Development**

1. Investigate and coordinate the cultural change within the materials community.
2. Develop jointly with the materials community new incentive structures and aid in their implementation to foster new scientific insights through data sharing.
3. Maintain NFDI-MatWerk's ties to other strategic participant projects and relevant authorities and organizations, both nationally and internationally.

- **Task Area Ontologies for Materials Science**

4. Develop and implement a unified materials ontology. This enables interoperability between heterogeneous materials data, hierarchical microstructure information, analytics tools and materials models in a flexible and seamless manner.
5. Develop environments enabling simple implementation of 'private' materials ontologies. These allow the integration of 'homegrown' tools and digital workflows. The ontology developing environment needs to supported a stringent version control and the possibility to suggest or share these ontology branches to partners and the greater community.

- **Task Area Materials Workflows and Software Development**

6. Develop and distribute best practice digital workflows for experimental and simulation environments that allow a representation of the multiphysical character of materials.
7. Enhance tools for elaborated programming of workflows, integrating measurement devices as well as materials data analytics. The workflow environment needs to reproducibly connect publishable figures with the underlying raw data.
8. Provide advanced data curation methodologies and automated quality checks. Implement appropriate machine-learning approaches to ensure that materials data meets pre-defined community standards.

- **Task Area Materials Data Infrastructure**

9. Provide an environment to share and access materials raw data and metadata, based on the FAIR principles[1], which contributes to a comprehensive digital representation of materials and material classes.
10. Set up straightforward data sharing across institutions and disciplines for collaborative projects.

[1] Data ought to be "**F**indable, **A**ccessible, **I**nteroperable and **R**eusable (FAIR) both for machines and for people", as defined by Wilkinson et al, 2016 [LINK].

11. Quickly establish local data spaces and their guidelines in accordance with NFDI-MatWerk's infrastructure, with integrated data sovereignty (e.g. access management) and a secure authorship (e.g. through cryptography and certification of workflows and data).

- **Task Area Community Interaction**

12. Offer workshops, hackathons, summer schools and professional trainings together with the Task Areas on topics related to NFDI-MatWerk for different levels of qualification.

13. Gather input from the MSE community on NFDI-related topics.

14. Create information and course material to help with onboarding of national and international new students and scientists as well as implement it into their teaching.

15. Get experts involved through MSE societies, e.g. DGM, GAMM and DVM.

16. Define and establish standards for data curation.

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

On a technical side, the consortium NFDI-MatWerk builds upon existing infrastructure and services. E.g. the data centers of RWTH Aachen and KIT provide profound experience with large scale FAIR data storage, the DFKI the experience in software and KI development, the MPIE and IWM and others digital materials workflows in connection with well-established databases, the FIZ-Karlsruhe ontology development, and our societies DGM, DVM, GAMM and others provide deeply rooted community interactions. Furthermore, infrastructure use-cases will be established between related NFDI's (e.g. NFDI4Ing, FAIRmat, NFDI4Chem, MaRDI), ensuring effective knowledge transfer and dissemination. Technical aspects all NFDI-consortia will have to implement will require harmonization and be addressed in "4 Cross-cutting topics".

The **requirements as well as other existing infrastructure, tools, and workflows from our scientific community** will be incorporated through direct interaction in form of workshops, annual meetings, hackathons and online seminars. Via our so-called **participant projects** representative and partially complex scientific situations have been selected to derive the requirements for the digital materials infrastructure (e.g. DIWAN in Dresden, Defect phases in structural materials CRC 1394 in Aachen, BMBF MaterialDigital, SimTech in Stuttgart, *liv*MatS in Freiburg, and many more).

Following the suggestions by the expert commission, a **transparent Continuous Integration / Continuous Delivery (CI/CD) cycle has been implemented** to define the work packages within NFDI-MatWerk: Based on the community requirements (derived from questionnaires, workshops, conferences and participant projects) usage profiles are derived and translated into infrastructure use-cases. These are separated into more general work packages to develop infrastructure services or products. Based on the necessary effort and their benefit across the use-cases, the work packages are prioritized by the consortium and then implemented by expert teams. After implementation, the **infrastructure use-cases are used to query** the validity of the infrastructure implementation. The use-case related participants then **beta test** the developed infrastructure in close interaction with developers. After beta testing, the new features of the digital materials infrastructure are rolled out to the greater community and the cooperating NFDI consortia and the process starts again. All steps within the development cycle are openly communicated and documented so that the materials community can participate.

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

The consortium NFDI-MatWerk has already identified infrastructure use cases within the NFDI to initiate collaborations in engineering and natural sciences.
Regarding **engineering sciences**, **NFDI4Ing** received a funding admission to implement common interfaces between the different disciplines of engineering sciences. Thus, specific consortia for the larger disciplines must be set up that address and implement their specific needs and directly connect to their communities. Therefore, NFDI-MatWerk **aims to develop a digital infrastructure that allows for the representation of materials microstructure** and its path dependence. **The two initiatives NFDI4Ing and NFDI-MatWerk are therefore highly complementary.** Accordingly, NFDI4Ing and NFDI-MatWerk jointly initiated MSE-specific working groups within a workshop in July 2020 and identified infrastructure use-cases concerning joint development of software infrastructure, an ontological interface between materials information, as well as Industry 4.0 standards describing the related machines and processes. In the closely related **scientific discipline of computer and data sciences**, NFDI-MatWerk has already started discussions with **NFDIxCS, NFDI4DataScience** and **NFDI4HPC**. Since NFDI-MatWerk is relying on large-scale data and computational resources, the collaboration with dedicated expert consortia will be highly valuable in the ontology field as well as reproducible use of Artificial Intelligence and the related training data.
In the **field of natural sciences,** NFDI-MatWerk has started collaborations with different consortia from Physics (FAIRmat), and Chemistry (NFDI4Chem). In Physics, we aim to collaborate closely with **FAIRmat, NFDI4Phys** as well as **DAPHNE**. Here, the interface to **FAIRmat** is most relevant for NFDI-MatWerk, since condensed-matter physics is a complementary perspective on solid-state phenomena. FAIRmat focuses their effort on fundamental chemo-physical effects. In contrast, NFDI-MatWerk addresses the relevance of the microstructure for engineering materials and their applications. At the same time, **NFDI4Phys** aims to harmonize the semantic description of physical units and quantities, which will be closely related and relevant to the ontologies of NFDI-MatWerk. Finally, **DAPHNE** intends to provide interfaces to large experimental facilities, both ontologically and in regards of software, for which a collaboration with NFDI-MatWerk will also be **mutually beneficial**.
Based on **NFDI4Chem's** efforts on electronic lab-books, NFDI-MatWerk would like to team up with them to apply a similar solution to the materials community's needs and share our tools, and efforts on the developing materials ontology.
In life science, NFDI-MatWerk has initiated a collaboration with the approved consortium **DataPlant**. DataPlant has vast experience in chaining and docking of individual software modules for workflow integration, which is something NFDI-MatWerk intends to implement, too.
Finally, NFDI-MatWerk has initiated a collaboration with the **MaRDI** consortium for mathematics. NFDI-MatWerk plans to implement machine-readable representations for mathematical expressions and interdependencies developed by MaRDI. Therefore, a shared use-case has been discussed and will be established with the start of the consortia.

4 Cross-cutting topics

- Please identify crosscutting topics that are relevant for your consortium and that need to be designed and developed by several or all NFDI consortia.

There are several obvious common challenges in setting up a scientific digital infrastructure. These have been pointed out within the NFDI conference by many consortia and summarized by the NFDI-chairman York Sure-Vetter. We will be happy to aid into this effort and would be happy to participate also in the orchestration of the other upcoming tasks.
Many communities face the challenge to set up a FAIR research data infrastructure incorporating highly heterogeneous data in a structured form. Most consortia aim to solve this through the usage of ontologies, relying on harmonization in context of their architecture, interfaces, as well as GUIs for developing them. Furthermore, issues of data sovereignty, including IT-security, as well as a common authentication and authorization infrastructures (AAI), are calling for a unified approach. Generally, interfaces (APIs) for software accessibility to different units will be another reoccurring challenge. Furthermore, processes or systems for data curation will require growing attention for each consortium, as the absolute amount of involved data will be mounting.
Finally, there are also "soft" or non-technical aspects that the different NFDI-consortia must regard. Here, the legal setup and framework for running a shared infrastructure will require a lot of joint attention, since the environment will touch legal aspects such as ownership, data protection and national security. Furthermore, every community will need to implement new working models that incentivize sharing knowledge and data, which will require a joint culture change throughout all scientific disciplines.

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

Due to the large number of currently isolated digitalization projects, the community of NFDI-MatWerk has already started to gather experience that might be relevant to other consortia and to the NFDI as a whole. Here, the existing solutions for consistent data workflow management connecting experimental and simulation data on different hierarchical levels in combination with data driven modelling approaches enable seamless transformations and analyses of highly heterogeneous large data sets. Since dealing with big-data analyses will become a focus of many consortia, NFDI-MatWerk's approaches might be valuable to others, too.
Furthermore, the consortium of the NFDI-MatWerk could already collect experience on the development of materials ontologies, which are perhaps the most prominent method towards meeting the FAIR-principles. Here, NFDI-MatWerk and its members will be able to provide this experience as well as help linking different knowledge graphs.
Finally, the exchange and servicing of (decentral) data is common business in the field of MSE. Therefore, the consortium and its members have already significant experience regarding legal aspects, e.g. on contract configurations, pitfalls and sovereignty issues.