Letter of Intent

nfdi4energy –

Nationale Forschungsdateninfrastruktur

für die interdisziplinäre

Energiesystemforschung

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

X	Binding letter of intent (required as advance notification for proposals in 2021)

2   Formal details

- Planned name of the consortium

 Nationale Forschungsdateninfrastruktur für die interdisziplinäre Energiesystemforschung

- Acronym of the planned consortium

 nfdi4energy

Applicant institution	Spokesperson
Carl von Ossietzky Universität Oldenburg (UOL) Ammerländer Heerstr. 114-118 26129 Oldenburg Head: Prof. Dr. Ralph Bruder	Prof. Dr. Astrid Nieße astrid.niesse@uni-oldenburg.de
Co-applicant institutions	**Spokespersons**
Albert-Ludwigs-Universität Freiburg (ALU) Friedrichstraße 39 79110 Freiburg Head: Prof. Dr. Kerstin Krieglstein	Prof. Dr. Anke Weidlich anke.weidlich@inatech.uni-freiburg.de
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) Schloßplatz 4 91054 Erlangen Head: Prof. Dr. Joachim Hornegger	Prof. Dr. Reinhard German reinhard.german@fau.de
Institute for Advanced Sustainability Studies e.V. (IASS) Berliner Straße 130 14467 Potsdam Head: Prof. Dr. Mark G. Lawrence	Prof. Dr. Dr. Ortwin Renn ortwin.renn@iass-potsdam.de

Karlsruher Institut für Technologie (KIT) Kaiserstraße 89 76131 Karlsruhe Head: Prof. Dr.-Ing. Holger Hanselka	Prof. Dr. Christof Weinhardt christof.weinhardt@kit.edu Prof. Dr. Veit Hagemeyer veit.hagenmeyer@kit.edu
OFFIS e.V. (OFFIS) Escherweg 2 26121 Oldenburg Head: Prof. Dr. Sebastian Lehnhoff	Prof. Dr. Sebastian Lehnhoff sebastian.lehnhoff@offis.de
RWTH Aachen University (RWTH) Templergraben 55 52056 Aachen Head: Prof. Dr. rer. nat. Dr. h.c. mult. Ulrich Rüdiger	Prof. Antonello Monti, Ph.D. amonti@eonerc.rwth-aachen.de
Soziologisches Forschungsinstitut Göttingen (SOFI) e.V. Friedländer Weg 31 37085 Göttingen Head: Prof. Dr. Berthold Vogel, Dr. Martin Kuhlmann	Prof. Dr. Berthold Vogel berthold.vogel@sofi.uni-goettingen.de
Technische Informationsbibliothek (TIB) Welfengarten 1 B, 30167 Hannover Head: Prof. Dr. Sören Auer	Prof. Dr. Sören Auer auer@tib.eu
Participants	
Reiner Lemoine Institut (RLI) Head: Dr. Kathrin Goldammer Contact: Editha Kötter Editha.Koetter@rl-institut.de	
Fraunhofer-Institut für Angewandte Informationstechnik FIT (FhFIT) Head: Prof. Dr. Stefan Decker, stefan.decker@fit.fraunhofer.de	

DFG

3   Objectives, work program and research environment

Research area of the proposed consortium (according to the DFG classification system:
https://www.dfg.de/dfg_profil/gremien/fachkollegien/faecher/index.jsp)

111 Sozialwissenschaften

112 Wirtschaftswissenschaften

407 Systemtechnik

408 Elektrotechnik und Informationstechnik

409 Informatik

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

Energy system research is facing major challenges due to the transformation towards post-fossil systems and the increasing connection between the different energy systems. In this context, the **digitization process towards cyber-physical energy systems (CPES)** alleviates change in many respects: First, it drives and enables new levels of automation, sector-coupling and thus allows to address new optimization potentials towards energy efficiency. Second, the long-envisioned idea of a closer integration of new stakeholders, right down to private households, becomes feasible based on a broad range of new participation opportunities. Third, digitalization creates the conditions for improving energy research itself and fostering a change towards trans-disciplinary energy system development and research. Thus, digitalization in energy systems equally affects technical, social and societal topics, as well as the mode of research of the energy system research community.

Particularly in the area of **research data and research software**, the possibilities of digitization are not yet being fully exploited for several reasons. Additionally, the implications regarding infrastructure projects affecting the society as a whole is not reflected in detail. The whole CPES research and transfer cycle from societal questions to fundamental research up to implementation in the field with industry partners is not yet fully supported with respect to data and software findability, accessibility, interoperability and reusability (FAIR) requirements. On the other hand, some **specific challenges** have to be addressed in the field of CPES research: First, **interdisciplinary research** in a regulated application context has to be realized (including e.g. experts from energy engineering, electrical engineering, economics, computer science, and social sciences). Second, **participation is an obligation** over the whole research and transfer cycle due to the broad societal effects of energy system changes. Third, **transparency and efficient utilization of research results** is needed for all stakeholders, i.e. public, governmental, industrial, and research.

CPES research thus shows several characteristics leading to special requirements regarding research data and research software management.

The goals of nfdi4energy in this context of digitization of energy research are:

- **Goal 1**: Establish research data infrastructure in energy system research, alleviate, and motivate its use in the research community.

- **Goal 2:** Allow traceability, reproducibility and transparency of results for the scientific community as well as for the public.

- **Goal 3**: Facilitate transfer of scientific results into business practice.

- **Goal 4**: Enable and motivate involvement of the public in identifying and addressing relevant research objectives from a social and societal perspective.

- **Goal 5**: Promote better collaboration between scientific research institutes and business partners via research data management, respecting conflicting interests and options regarding openness of data and software.

In **nfdi4energy**, these challenges and goals will be addressed starting with a user-centered requirements analysis for a platform concept that provides appropriate community service within the research, development and transfer cycle. From the beginning of the process, all targeted user groups will be involved. Thus, nfdi4energy will supplement on current research and development within the nfdi initiative: While overarching, less application-specific activities such as NFDI4Ing are primarily devoted to the important topics of interoperability and integrability, e.g., in the area of metadata, nfdi4energy will focus on domain-specific applications for research data management. Within the energy system domain, aspects of transfer are already being worked on, but not in the context of a target group-specific process integration from requirements analysis to transfer e.g. from the integration of the public to the reuse of models. Preliminary work in this direction is currently being carried out in the Zukunftslabor Energy (ZLE), in which an initial platform concept is being developed and which will be built upon here.

The planned services of the research platform represent the implementation of an outlined and to be detailed CPES research, development and transfer cycle support (see Figure 1):

- **Competence:** Building on the platform EnArgus, which gives an overview on energy research projects, an overview of scientific institutes, scientists and commercial enterprises should be created. It will also be used as a foundation and reference point for metadata for software and data in the other categories.

- **Best Practices:** This component will provide an overview of good practices in research projects, list and offer training courses e.g. on software engineering for non-computer scientists, give guidelines for using the platform. Successful examples of cross-
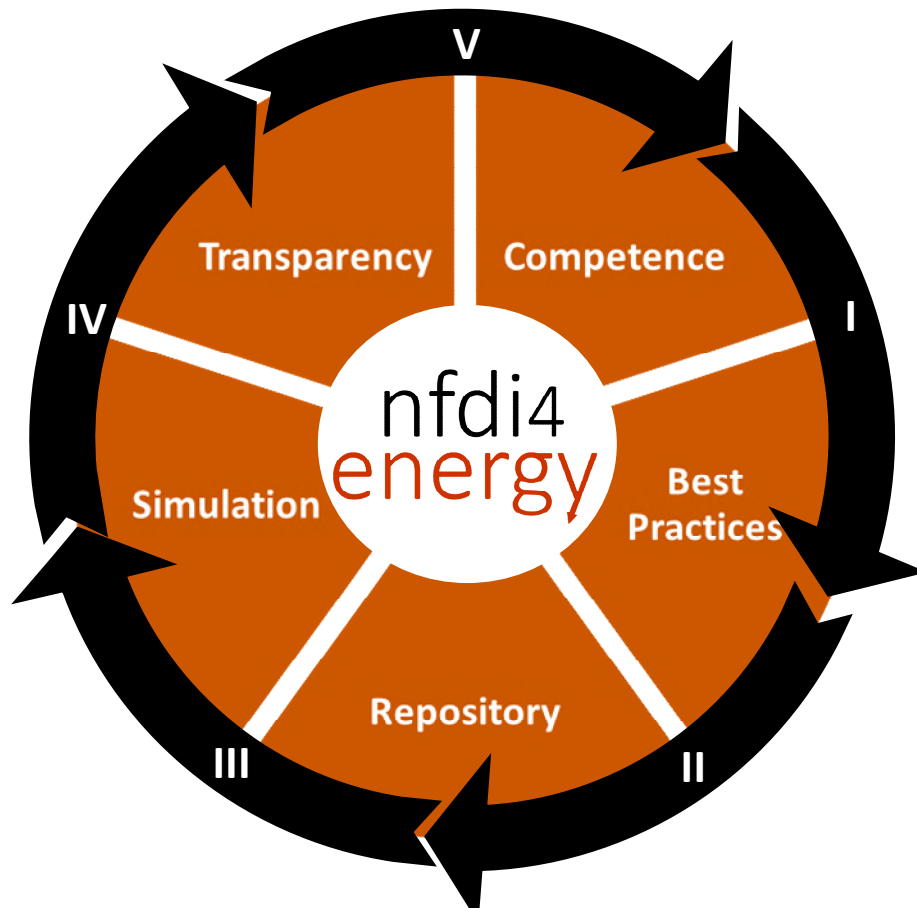
**Figure 1: CPES research, development and transfer cycle based on
nfdi4energy platform components**

institutional research data management provide a low-threshold offer, which is complemented by mapping community standards collected through surveys and information to support the establishment of the FAIR standards through domain-specific guidelines and recommendations.

- **Repository:** A registry for metadata for software and data is provided. It supports the creation and continuous integration of appropriate metadata based on standards and controlled vocabularies like the Open Energy Ontology. Access restrictions for data and software will be considered.

- **Simulation:** The simulation component provides support for co-simulation through community-driven interface harmonization. The creation and use of simulations by non-computer scientists are defined as goals e.g. for the Citizen Science aspects of the nfdi4energy project. Data analytics and visualization aspects will be reflected by the platform as a starting point.

- **Transparency:** The transparency component supports the comparability of scenarios, reflects social and societal aspects and needs, and integrates components from Citizen

Science. Thus, this component plays an important role for the identification of new research questions.

The nfdi4energy consortium will implement the platform reflecting all components as outlined above. A refinement process building on design thinking methods integrating all relevant user groups will promote both process and platform development.

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

nfdi4energy builds on the comprehensive infrastructures, tools, services and data of its members, covering all phases of the CPES research, development and transfer Cycle as depicted above. Community-based integration and extension of existing elements of nfdi4energy will create a powerful infrastructure that supports the research process transformation to collaborative community-driven open science respecting both public and industry interests. Open Science is essential for energy system research, as it makes research results transparent, comprehensible, reproducible, and assessable by politics and society, and reusable for further research. nfdi4energy members are engaged in key community initiatives and projects that advance open science and FAIR principles in energy systems research, and have already contributed to the community in this respect. Exemplary relevant infrastructures from within the consortium include:

In the field of **metadata development, ontologies, and interoperability of research data**, several partners are active in different initiatives from both research and standardization. The development of the Open Energy Platform (OEP) is supported by RLI. The OEP is an open science infrastructure for energy system research that aims to improve transparency, quality and reproducibility in energy system modeling. Other important components of the OEP are the Open Energy Database (OEDB), the Open Energy Modeling Framework as an open source software for energy system modeling, and a domain ontology for energy research, the Open Energy Ontology (OEO). nfdi4energy will further build on the extensive services and repositories of the TIB. These include a terminology service for the collaborative development of terminologies, vocabularies and ontologies, a repository for metadata of research data (Leibniz Data Manager), Open Research Knowledge Graph (ORKG) for FAIR organization and linking of scientific findings in a knowledge graph and a DataCite service for providing persistent identifiers (DOI) for research data, (software) and other artefacts. KIT is an active contributor to the Helmholtz Metadata Collaboration and hosts the HMC Energy Hub. Relevant preliminary work has been performed in the graduate school Energy Status Data. RWTH has worked on

SARGON in this context, while OFFIS and the University of Oldenburg address standardization and integration of simulators using metadata.

Simulation and modeling is an important aspect for CPES research, with several partners contributing to this component, e.g. FAU with the micro traffic simulator Veins, and the lightweight discrete energy simulator MiniDES (MIT). The co-simulation framework mosaik (LPGL 3.0) by OFFIS offers real-time coupling and orchestration of these and other simulators for cross-domain experiments. RWTH provides infrastructure and services for the real-time simulation of (digitalized) electricity grids (DPSIM, DISTAIX) as well as the coupling of laboratory infrastructure across Europe and on an intercontinental level. ALU published simulation models and data w.r.t. carbon emission aspects as well as flexibility modeling research data and software. Data analytics and visualization tools are provided by several partners, e.g. by KIT for several field trials and demonstration sites.

With respect to **Transparency** and **Best Practices**, the KIT holds the KD^{2}Lab: It allows for economic, neurological and psychological experiments to examine human decision processes within controlled laboratory conditions. This laboratory thus supplements work planned by SOFI and IASS in the are of socio-technical aspects of CPES research. OFFIS is building in-house data management and analysis platforms using open source and cloud technologies. The infrastructural foundation of these platforms consists of a Kubernetes cluster running containerized applications in combination with Ceph as a petabyte-scale software-defined storage solution. Scientific end users are presented with tailored JupyterHub instances allowing for project-specific features and data access capabilities.

An extensive discussion on existing infrastructures addressing all components of the envisioned platform will be given in the full proposal.

Interfaces to other funded or proposed NFDI consortia:

The general idea of nfdi4energy is to develop application-specific services and infrastructures addressing the CPES research challenges as depicted above building on results and current work from existing and new nfdi consortia. Thus, while the services and infrastructure have to be provided by nfdi4energy, the underlying methods and concepts will build upon cooperation e.g. with the following consortia:

NFDI4Ing works on FAIR data and open science principles with a special focus on the engineering sciences. With this focus, NFDI4Ing addresses interdisciplinary research as well. A cooperation with this consortium will focus on metadata and terminology services, repositories and storage, as well as automated data and knowledge discovery.

NFDI4DataScience is a consortium with a special focus on data science in academia. Initially, data intense application areas are the core area, like biomedical and social sciences. With the services developed from this consortium though, additional application areas like CPES research are both straightforward and intended. The development and use of knowledge graphs is a key method within this consortium. An intense cooperation in this regard is planned.

The nfdi consortium MaRDI focuses the task of applying the FAIR principles to mathematical research software, models and data. Within CPES research, many approaches apply methods from applied mathematics, especially in the field of multi-criteria optimization. For this, working with well-known benchmark functions is a crucial step in early iterations of the research and development cycle, thus paving the ground for larger system evaluations in later iterations. The MaRDI consortium has shown interest in cooperation; the details will be depicted in the full proposal.

As nfdi4energy addresses socio-technical aspects in CPES research, a cooperation with KonsortSWD is planned as well. Among other initiatives, with RatSWD, this consortium already has a long-term expertise in addressing research data aspects in social, economical and behavioral sciences. It is planned to cooperate with this consortium regarding the social, societal and socio-technical aspectes in nfdi4energy.

With the third funding round in nfdi, base consortia will be funded as well. nfdi4energy explicitly expresses the motivation to cooperate with and reuse results from these consortia as well as serve within the requirement analysis process of these consortia.

4   Cross-cutting topics

nfdi4energy will contribute to the following topics, acknowledging the importance and greater breadth given to this aspect by the common Berlin Declaration on NFDI Cross-Cutting Topics. For this consortium, the most relevant areas of contribution are:

Metadata, Findability

Starting with its Competence and Repository building blocks, nfdi4energy will link its platform with other consortias' platforms to contribute to a federated catalogue. The catalogue will require a shared kernel of common (meta)data schemas, enabling cross-domain findability of input data, research results and research software.

Additionally, nfdi4energy will adopt and push for the adoption of standards (standardized metadata schema) for data lineage (metadata for fair workflows), provenance and licensing information, and accessibility (incl. authentication and authorization schema) in close alignment with the prior art created by the existing consortia.

Training and Education

Building in part on the groundwork done in ZLE, nfdi4energy will contribute both resources and courses through its **Best practices** building block. Training especially on reproducible research is a prerequisite for data reuse and improved scientific practice. The aim is to avoid duplication of materials already present, but instead to build on existing guidelines and recommendations of FAIR principles. Only where gaps are identified, specific materials and guidelines are to be created and disseminated. For increased uptake on the most central topics, training courses will be created, published and conducted (both in presence and online). Possible topics include e.g. research data management for the energy domain, research software testing, revision management and version control, and preparing and conducting a reproducible workflow for later archival of the whole experimental pipeline.

Ethical & Legal Aspects especially for person-related data

Similar to the previous topic, nfdi4energy will contribute knowledge and best practices from the energy domain. Here, the benefits from sharing efforts across domains are most pronounced. nfdi4energy will contribute domain-specific extensions to existing materials. One prominent example of sensitive data are household load profiles. Depending on the depth of analysis, these can be considered person-related data. A second cross-cutting example are proprietary data sources and how to strive towards reproducible research in contexts where confidentiality of data providers must be adhered to.

Quality Management & Assurance

nfdi4energy will participate in the development of community specific quality standards and automated quality checks. Checks can be based of (meta)data ontologies to assure completeness and basic data quality requirements. The common groundwork again lies in standardized data/data formats in energy system research (for data quality and interoperability). Education-wise best practices for data curation before publishing aims to improve the general quality of data packages before submission. A stretch goal this consortium is working towards is enforcing reproducibility of results in published research through reproducible workflows and open data. Although this goal might not be fully achievable within the funding period, significant progress in the ease to achieve this high standard for practicing researchers is to be achieved.

Cultural Change

nfdi4energy contributes to a change in data culture in the community towards open research data and the adoption of FAIR principles, like spearheaded by organizations like the Open Energy

Modelling Initiative. This consortium will support this and other Open Innovation Initiatives through active engagement and outreach in their respective forums. The overall goal: incite scientists to better document data and software, publish both under open licenses and thus realizing the move to a more open research culture.

DFG