National Research Data Infrastructure for Earth System Sciences (NFDI4Earth)

Spokesperson: Prof. Dr. Lars Bernard, Technische Universität Dresden,
lars.bernard@tu-dresden

1. THE NFDI4EARTH CONSORTIUM

Research Domain addressed by NFDI4Earth

NFDI4Earth (www.nfdi4earth.de) addresses digital needs of researchers in Earth System Sciences (ESS). ESS comprises numerous sub-disciplines with the overarching aim to understand the functioning of all subsystems of the Earth System and quantify their interactions in space and time. ESS responds to today's pressing global change challenges by intensifying national, European, international and interdisciplinary efforts for the creation of global, sustainable solutions. NFDI4Earth is a building block of this strategy.

Types of Data and Methods addressed by NFDI4Earth

NFDI4Earth has to deal with a broad variety of data types due to the range of subjects and subdisciplines in ESS. Data stems from a plethora of observations, simulations and analysis in different combinations, forms and formats (numerical, textual, and graphical). Many different open and standardized as well as proprietary data formats and interfaces are used. Relevant standardization stems for instance from the World Wide Web Consortium (W3C), the Open Geospatial Consortium (OGC), ISO/TC 211 as well as from legal frameworks, e.g. the EU INSPIRE directive. Some specific examples are: XML data that are widely used for domain specific schemata (e.g. GeoSciML, WaterML, BorholeML etc.), data formats like netCDF/HDF5 and GeoTIFF that are most commonly used to exchange meteorological, hydrological and remote sensing data. Examples for geospatial data are GML, GeoJSON or ESRI's Shape format. Several approaches exist that address common ontologies. However, it still lacks in providing (semantic) data mapping and transformation paths and data integration remains tedious. Due to the extensive thematic range of NFDI4Earth, a large number of data exploitation methods are required - ranging from numerical and statistical modelling, over machine learning to semi-automated and human interpretation. A common characteristic of ESS data are their spatio-temporal variability and a wide range of space-time reference systems ranging from the smallest to the largest scales. Example use cases range from the analysis of data in Earth System data labs or large model comparison studies to field studies and stakeholder interviews in the investigation of local adaptations to global changes. Often complex digital processes are established, which mostly require specific scientific expertise. ESS also often faces a data paradox: on the one hand large (dense) data streams have to be dealt with (from, e.g., remote sensing sensors, models or social media), while on the other hand many data streams yield data that are spatially, temporally and thematically sparsely distributed and incomplete (e.g. certain types of observations or survey results).

Planned measures and services for RDM in ESS

In a user-driven process NFDI4Earth will provide researchers with FAIR, easy, coherent, efficient, open, and – whenever possible – unrestricted access to all relevant Earth System data, scientific data management tools and data analysis services. The design and implementation of the NFDI4Earth will follow a progressive and iterative process. Here, we list measures and services that are core to NFDI4Earth:

- ESS Pilots to engage the NFDI4Earth community. Pilots will stem from different ESS domains, indicating the researchers' needs and are the community's contribution to an

agile development of NFDI4Earth.

- Services are offered via a one-stop frontend and include a support structure that provides an easy route for all ESS researchers to access research data as well as analytical and computing environments. In addition, researchers will receive support for data integration, to publish data and software according to agreed openness and quality criteria, and in archiving and preserving research data.
- Requirements driven technology design – keeping the users and providers of ESS research data and software connected and in continuous open dialogue.
- Education and Training for ESS research data management and ESS data science
- Services supporting the exploitation of new technologies and that (help to) create next generation solutions for research data management and research data analytics for ESS.
- Services to support the establishment of common agreements and regulations for ESS researchers and infrastructure providers towards FAIR and open research data and software management.
- Measures to fully link and embed NFDI4Earth services and achievements into national and international developments.

Specific requirements on RDM in ESS and the NFDI4Earth way to address them

A multitude of observations and model data in very high spatial, temporal and thematic resolutions lead to rapidly increasing data volumes. Thus, describing and assessing Earth System processes, their dependencies, and their changes urgently require efficient workflows and extremely powerful data analytics frameworks. Today, various services to support RDM in ESS exist. However, they are scattered and heterogeneous as well as mostly project based. Consequently, they are lacking long-term perspectives. The above listed measures and services do therefore target the provision of sustainable and coherent solutions that are findable via a one-stop framework, provision of common data analytics frameworks (e.g. 4D Data Cubes) and a one community approach to Openness, FAIRness and sustainable RDM in ESS.

Relevant experiences of the NFDI4Earth consortium

The initial NFDI4Earth consortium has been formed by a large, representative group of the German ESS community and has adopted an open and inclusive membership strategy for continuous growth. It covers all domains of the Earth System from solid Earth to ocean, atmosphere, hydrosphere, cryosphere, and terrestrial systems. Co-applicants either operate major research data infrastructures related to ESS or act as research and education hubs in this field. The participants strongly support the core team of applicant and co-applicants with domain-specific expertise, additional repositories, and as multipliers into the ESS community. Co-applicants and participants often act in a double role as providers and as users of Earth System data. Data provided by NFDI4Earth partners are often used in a much wider international context. Since many of the participating repositories have been operational for many years, NFDI4Earth will not so much create new data repositories and archives, but instead primarily lead to a consolidation and harmonization of existing research data infrastructures with the clear goal to bring all of them to the highest FAIRness levels and to advance interoperability.

Networking relevant partners and Infrastructures

The NFDI4Earth consortium has been created in a bottom-up process and comprises currently 55 members from universities, research institutions, infrastructure providers, public authorities and different research organizations. A community process to screen the landscape of existing infrastructures, services and collaboration tools supporting ESS resulted in a list of more than 100 platforms and tools. Many of the screened platforms are

hosted, co-operated or co-developed by one or several members of the NFDI4Earth consortium. Prominent examples are PANGAEA (AWI), GFZ Data Services and re3data (GFZ), Earth System Grid Federation (DKRZ). NFDI4Earth can further capitalize on a series of consolidation efforts, e.g. the DataHub started by the Helmholtz Research Field Earth and Environment. NFDI4Earth will connect to international efforts like the climate data store (ECMWF), the European Space Agency (ESA) data cubes, the European Open Science Cloud (EOSC) to develop novel avenues to make scalable data analytics available to a wide user community in the cloud. In this context, NFDI4Earth represents a number of the German centres for scientific computing and has close linkages and access to HPC and storage resources (e.g. DKRZ, KIT, FZJ, TUD-ZIH).

NFDI4Earth strongly links into international initiatives and networks and will ensure international embedding of NFDI4Earth into for instance ENVRI, ICOS, EPOS, EOSC, ISO, OGC, WDS, RDA, and WMO.

NFDI4Earth Interfaces to the overall NFDI

Additional to the joint work and interfaces addressed as cross cutting topics (below), some of the NFDI4Earth consortium members do simultaneously participate in other NFDI consortia, thus acting as direct interfaces. Moreover, NFDI4Earth plans for cross-consortium pilots jointly with other NFDI consortia (e.g. NFDI4BioDiversity, KonsortSWD) to demonstrate the NFDI benefit for interdisciplinary research.

2. CROSS-CUTTING TOPICS

Relevant NFDI Cross-Cutting Topics from an NFDI4Earth perspective

NFDI4Earth co-spokespersons co-drafted and signed the *Berlin Declaration and the Leipzig-Berlin Declaration on NFDI Cross-Cutting Topics* (https://zenodo.org/record/3895209) to define a core set of NFDI cross-cutting topics. The NFDI4Earth fully shares the idea of the Research Data Commons (RDC) as the NFDI-wide agreed and developed set of basic and common NFDI standards, technological approaches and infrastructure components and will streamline its design and developments with this concept.

Potential NFDI4Earth Contributions to NFDI cross-cutting topics

The NFDI4Earth work plan contains several measures to contribute to the development of common NFDI services. The NFDI4Earth consortium offers these specific NFDI contributions:

- NFDI can capitalize on NFDI4Earth long-standing active participation in various international networks, focusing on RDM infrastructures, standardization, legislative and organizational frameworks.
- NFDI4Earth will actively participate in the NFDI harmonization efforts. Here, NFDI4Earth builds on widespread and valuable experiences in contributing to international standards and agreements, most prominently in the area of metadata standards and data encodings for spatial data. NFDI4Earth profits from a well-established culture and skills concerning community software developments, which will be contribute into the NFDI.
- NFDI4Earth would serve as the NFDI knowledge hub for dealing with Earth related spatio-temporal data
- Having one focus on the development of criteria frameworks to assess services and offerings from research data infrastructures, NFDI4Earth would derive generic results and findings and provide these into the NFDI development.

3. WHAT WE EXPECT FROM PARTICIPATION IN THE NFDI CONFERENCE

We are keen to hear about latest developments of the NFDI-Direktorat and the NFDI-Verein, the planned process for 2nd round of NFDI DFG-applications, to learn about the approaches of other NFDI consortia, to network and to discuss about the NFDI cross-cutting topics.

Planned members of the NFDI4Earth Consortium (Co-spokesperson and further Participants)

Co-spokesperson	Institution
Prof. Dr. Stephan Frickenhaus	Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI) Am Handelshafen 12, 27570 Bremerhaven
Prof. Dr. Andreas Wytzisk-Arens	Bochum University of Applied Sciences (Fachhochschule Bochum) Lennershofstr. 140,44801 Bochum
Hannes Thiemann	German Climate Computing Center (Deutsches Klimarechenzentrum) Bundesstraße 45a, 20146 Hamburg
Prof. Dr. Monika Sester	Gottfried Wilhelm Leibniz University Hannover Welfengarten 1, 30167 Hannover
Roland Bertelmann	Helmholtz Centre Potsdam – GFZ German Research Centre for Geosciences Telegrafenberg, 14473 Potsdam
Dr. Hildegard Gödde	Helmholtz Centre Potsdam – GFZ German Research Centre for Geosciences / Geo.X Telegrafenberg, 14473 Potsdam
Prof. Dr. Peter Braesicke	Karlsruhe Institute of Technology (KIT) Kaiserstraße 12 76131 Karlsruhe
Prof. Dr. Markus Reichstein	Max-Planck-Institute for Biogeochemistry (Max-Planck-Institut für Biogeochemie) 07745 Jena
Prof. Dr. Miguel Mahecha	Leipzig University (Universität Leipzig) 04104 Leipzig
Prof. Dr. Karsten Wesche	Senckenberg Society for Nature Research Senckenberganlage 25 60325 Frankfurt

Participants	Institution
Prof. Dr. Albert Remke	52° North GmbH Martin-Luther-King-Weg 24, 48155 Münster
Dr. Konstantin Ntageretzis	Geoverbund ABC/J c/o Forschungszentrum Jülich GmbH 52425 Jülich
Tino Langos	Federal Institute for Geosciences and Natural Resources (BGR), Stilleweg 2, 30655 Hannover
Dr. Ines Heine	Bundesamt für Kartographie und Geodäsie Geodateninfrastruktur Deutschland (GDI-DE) Richard-Strauss-Allee 11, 60598 Frankfurt/M.

Dr. Claudia Börner	Brandenburg University of Technology Cottbus – Senftenberg, Platz der Deutschen Einheit 1, 03046 Cottbus
Dr. Gauvain Wiemer	Deutsche Allianz für Meeresforschung e.V. Wissenschaftsforum | Markgrafenstraße 37, 10117 Berlin
Dr. Jutta Graf	German Aerospace Center (DLR) Linder Höhe,51147 Köln
Dr. Lorenz Reimer	Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Inhoffenstraße 7b, 38124 Braunschweig
Dr. Frank Kaspar	Deutscher Wetterdienst Frankfurter Straße 135,63067 Offenbach
Prof. Dr. Uwe Ulbrich	Freie Universität Berlin, Department of Earth Sciences Kaiserswerther Str. 16/18, 14195 Berlin
Sören Lorenz, Dr. Emanuel Söding	GEOMAR Helmholtz Centre for Ocean Research Kiel Wischhofstr. 1-3, 24148 Kiel
Prof. Dr. Patrick Hostert	Humboldt-Universität Berlin, Department of Geography Unter den Linden 6,10099 Berlin
Prof. Dr. Kay-Christian Emeis	Helmholtz-Zentrum Geesthacht – Centre for Material and Coastal Research, Max-Planck-Straße 1, 21502 Geesthacht
Dr. Jörn Quedenau	Institute for Advanced Sustainability Studies e.V. (IASS) Berliner Straße 130, 14467 Potsdam
Prof. Dr. Rita Adrian Prof. Dr. Hans-Peter Grossart	Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin
Dr. Gotthard Meinel	Leibniz Institute of Ecological Urban and Regional Development, Weberplatz 1, 01217 Dresden
Dr. Thorsten Agemar Prof. Dr. Inga Moeck	Leibniz Institute for Applied Geophysics Stilleweg 2, 30655 Hannover
Prof. Dr. Heiner Igel	Ludwig-Maximillians-Universität München, Department of Earth and Environmental Sciences, Geschwister-Scholl-Platz 1, 80539 München
Dr. Stephan Hachinger	Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Boltzmannstraße 1, 85748 Garching
Falko Glöckler	Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, 10115 Berlin
Dr. Jan-Philipp Dietrich	Potsdam Institute for Climate Impact Research (PIK) Telegraphenberg A 31, 14473 Potsdam
Prof. Dr. Susanne Buiter Dr. Julia Kowalski	RWTH Aachen University, Division of Earth Sciences and Geography, Templergraben 55, 52062 Aachen
Dr. Markus Schmalzl	Generaldirektion der Staatlichen Archive Bayerns Schönfeldstr. 5, 80539 München
Dr. Angelina Kraft	TIB – Leibniz Information Centre for Science and Technology University Library, Welfengarten 1 B, 30167 Hannover
Dr. Patric Seifert	Leibniz Institute for Troposheric Research (TROPOS) Permoserstraße 15, 04318 Leipzig
Prof. Dr. Ingo Kowarik Dr. Marco Otto	Technische Universität Berlin, Institute of Ecology Straße des 17. Juni 135, 10623 Berlin

Prof. Dr. Jörg Benndorf Prof. Dr. Jörg Matschullat	Technische Universität Bergakademie Freiberg Institute of Mine Surveying and Geodesy Interdisciplinary Environmental Research Centre Akademiestraße 6,09599 Freiberg
Dr. Kristian Bär	Technische Universität Darmstadt, Schnittspahnstr. 9, 64287 Darmstadt
Prof. Dr. Nguyen Xuan Thinh	Technische Universität Dortmund, Department of Spatial Management and Modelling, August-Schmidt-Straße 4, 44227 Dortmund
Prof. Dr. Sabine Attinger	Helmholtz Centre for Environmental Research – UFZ Permoserstraße 15, 04318 Leipzig
Prof. Dr. Christian Borgemeister Prof. Dr. Klaus Greve	Rheinische Friedrich-Wilhelms-Universität Bonn Center for Development Research Center for Remote Sensing of Landsurface Regina-Pacis-Weg 3, 53113 Bonn
Prof. Dr. Horst Marschall	Goethe-Universität Frankfurt am Main Theodor-W.-Adorno-Platz 1, 60323 Frankfurt am Main
Prof. Dr. Ramin Yahyapour Prof. Dr. Wolfram Horstmann	Georg-August-Universität Göttingen Göttingen State and University Library, Wilhelmsplatz 1, 37073 Göttingen
Prof. Dr. Michal Kucera	Universität Bremen Bibliothekstraße 1, 28359 Bremen
Prof. Dr. Detlev Stammer Remon Sadikni	Universität Hamburg, Center for Earth System Research and Sustainability (CEN), Mittelweg 177, 20148 Hamburg
Thilo Paul-Stüve	Christian-Albrechts-Universität zu Kiel, University Computing Centre, Christian-Albrechts-Platz 4, 24118 Kiel
Prof. Dr. Peter Dannenberg PD. Dr. Dominik Hezel	University of Cologne, Department of Geosciences, Albertus-Magnus-Platz, 50923 Köln
Prof. Dr. Jürgen Haase Prof. Dr. Michael Vohland	Universität Leipzig, Faculty of Physiscs and Earth Sciences Ritterstraße 26, 04109 Leipzig
Prof. Dr. Edzer Pebesma	Universität Münster, Institute for Geoinformatics Schlossplatz 2, 48149 Münster
Prof. Dr. Martin Kühn Dr. Detlev Heinemann	Carl von Ossietzky Universität Oldenburg, ForWind – Center for Wind Energy Research, Ammerländer Heerstraße 114-118, 26129 Oldenburg
Prof. Dr. Bodo Bookhagen Dr. Peter Kostädt	University of Potsdam, ‘Earth and Environmental Sciences’, University Library and Center for IT and Media Management, Am Neuen Palais 10, 14469 Potsdam
Prof. Dr. Thomas Walter	Eberhard Karls Universität Tübingen Center for Data Processing Geschwister-Scholl-Platz, 72074 Tübingen
Prof. Dr. Peter Krebs	Water Science Alliance e.V. c/o Technische Universität Dresden, 01062 Dresden
Prof. Dr. Gunnar Lischeid	Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Straße 84, 15374 Müncheberg