



NFDI4Earth

Auf dem Weg zu einer Nationalen Forschungsdateninfrastruktur – das Beispiel NFDI4Earth

DFNS 2022

Lars Bernard (TU Dresden) für das NFDI4Earth Konsortium

Research Data Management ?!

Take on idealized research process:

- Includes all aspects of acquiring, evaluating, analyzing, sharing, and publishing research data *and* research methods/software
- Research data is the foundation, intermediate, and final product of research

Overarching Goal:
Exchange and reproducibility of research results



Research Data Management

Tools



...

Initiatives



...

Repositories



...


Standards & Best Practices



...

Research Data Management

- Where do we stand typically/ideally...

 **PANGAEA.**
Data Publisher for Earth & Environmental Science

Not logged in

SEARCH SUBMIT ABOUT CONTACT

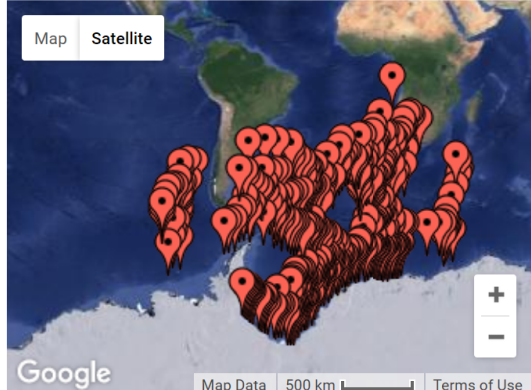
Citation: **Petschick, Rainer; Kuhn, Gerhard; Gingele, Franz (1996):** Grain size distribution of surface sediments of the South Atlantic. *PANGAEA*, doi <https://doi.org/10.1594/PANGAEA.51464>,
In supplement to: Petschick, R et al. (1996): Clay mineral distribution in surface sediments of the South Atlantic: sources, transport, and relation to oceanography. *Marine Geology*, **130(3-4)**, 203-229, doi [https://doi.org/10.1016/0025-3227\(95\)00148-4](https://doi.org/10.1016/0025-3227(95)00148-4)


Always quote above citation when using data! You can download the citation in several formats below.

[RIS Citation](#) [BisTeX Citation](#) [Text Citation](#) [Facebook](#) [Twitter](#) [Google+](#) [Show Map](#) [Google Earth](#)

Project(s): **Paleoenvironmental Reconstructions from Marine Sediments @ AWI** (AWI_Paleo) 🔍

Coverage: *Median Latitude:* -64.064892 * *Median Longitude:* -23.222025 * *South-bound Latitude:* -78.254000 * *West-bound Longitude:* -98.001000 * *North-bound Latitude:* -20.119000 * *East-bound Longitude:* 40.868333
Date/Time Start: 1983-02-13T00:00:00 * *Date/Time End:* 1995-05-10T21:07:00



 **DataCite**
Resources Contact

Search

Sort by

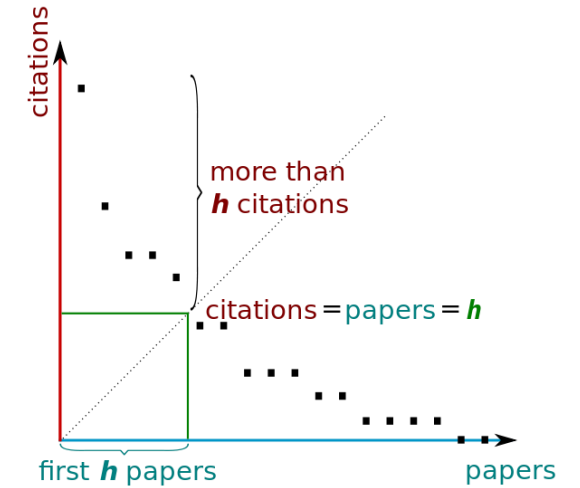
Scientific and statistical data formats

research Centre (ERC) in the Environmental
monitoring and assessment activities associated with

Plain text Structured text

Research Data Management – Challenges

- User driven !
- Support to researchers !
- Research data management as part of curricula !
- Reputation mechanisms !
 - It's still the paper that counts, not the data/software publication
- Sovereignty and sustainability
- Curation and quality assurance
- Balancing data protection and access
- Long-term archiving for software and data



<https://en.wikipedia.org/wiki/H-index>



Cases of Cholera Victims in London
(John Snow 1854)

Research Data Management – Challenges



“If we build it, they will come”

Nelson, B. Data sharing: Empty archives. *Nature* **461**, 160–163 (2009). <https://doi.org/10.1038/461160a>

Towards the National Research Data Infrastructure (NFDI)

RfII recommended to establish a German National Research Data Infrastructure (2016*):



<http://www.rfii.de/>

- Novel to implement the FAIR principles
(Findable, Accessible, Interoperable, and Re-usable).
- **Sustainable** (not only project solutions !)
- Step-by-step - **taking into account the different digital cultures** within disciplines
- Internationally integrated
- Driven by **scientific users and consortia of users and providers**

** RfII (2016): Enhancing Research Data Management: Performance through Diversity.*

Towards the National Research Data Infrastructure

„...systematize scientific data holdings with a national research data infrastructure and ensure sustainable access..."

Coalition Agreement, March 2018

<https://www.bundesregierung.de/Content/DE/StatistischeSeiten/Breg/koalitionsvertrag-inhaltsverzeichnis.html>

„...We want to further develop the National Research Data Infrastructure..."

Coalition Agreement, March 2021

<https://www.bundesregierung.de/breg-de/service/gesetzesvorhaben/koalitionsvertrag-2021-1990800>



picture alliance / dpa

Towards the National Research Data Infrastructure

„...systematize scientific data holdings with a national research data infrastructure and ensure sustainable access...”

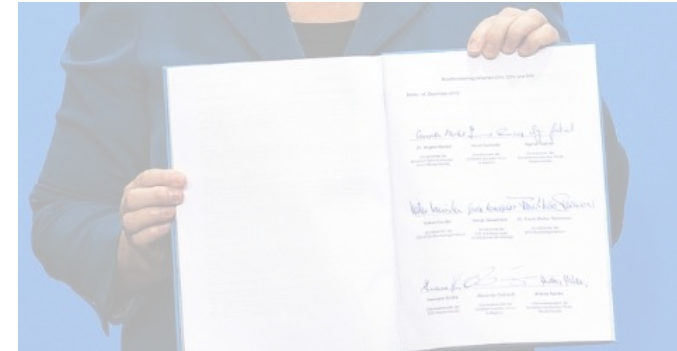
Coalition Agreement, March 2018

<https://www.bundesregierung.de/Content/DE/StatischeSeiten/Breg/koalitionsvertrag-inhaltsverzeichnis.html>

„...We want to further develop the National Research Data Infrastructure...”

Coalition Agreement, March 2021

<https://www.bundesregierung.de/breg-de/service/gesetzesvorhaben/koalitionsvertrag-2021-1990800>



picture alliance / dpa

The federal-state program to create the NFDI (11/2018)

- Funding over 10 years with € 90 million p.a.
- Funding of up to 30 consortia in 3 rounds of calls –
Managed by the German Science Foundation (DFG)
- Foundation of the **NFDI Association** (NFDI e.V.)
and the **NFDI Directorate** in Karlsruhe (10/2020)



<https://dfg.de/nfdi>

nfdi
<https://www.nfdi.de/>

NFDI Consortia – after 2 from 3 calls

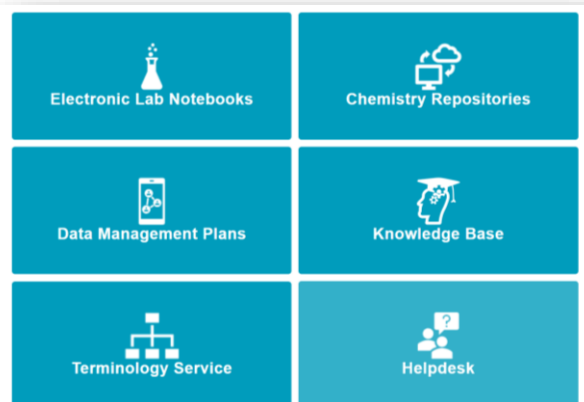
1st Round 2020

1	DataPLANT	DataPLANT - Data in Plant research	Biology
2	GHGA	GHGA - German Human Genome Archive	Medicine
3	KonsortSWD	KonsortSWD - Consortium for the Social, Behavioural, Educational, and Economic Sciences	Social Sciences
4	NFDI4BioDiversity	NFDI4BioDiversität - Biodiversity, Ecology & Envi-ronmental Data	Biology
5	NFDI4Cat	NFDI4Cat - NFDI for Catalysis-Related Sciences	Chemistry
6	NFDI4Chem	NFDI4Chem - Chemistry Consortium in the NFDI	Chemistry
7	NFDI4Culture	NFDI4Culture - Consortium for research data on material and immaterial cultural heritage	Humanities
8	NFDI4Health	NFDI4Health - National Research Data Infrastructure for Personal Health Data	Medicine
9	NFDI4Ing	NFDI4Ing - National Research Data Infrastructure for Engineering Sciences	Engineering Sciences

2nd Round 2021

1	BERD@NFDI	NFDI for Business, Economic and Related Data	Social and Behavioural Sciences
2	DAPHNE4NFDI	DAta from PHoton and Neutron Experiments for NFDI	Physics
3	FAIRmat	FAIR Data Infrastructure for Condensed-Matter Physics and the Chemical Physics of Solids	Physics
4	MaRDI	Mathematical Research Data Initiative	Mathematics
5	NFDI4DataScience	NFDI for Data Science and Artificial Intelligence	Computer Science
6	NFDI4Earth	NFDI Consortium Earth System Sciences	Geosciences
7	NFDI4Microbiota -	National Research Data Infrastructure for Microbiota Research	Life Sciences
8	NFDI-MatWerk	National Research Data Infrastructure for Materials Science & Engineering	Materials Science and Engineering
9	PUNCH4NFDI	Particles, Universe, NuClei and Hadrons for the NFDI	Physics
10	Text+	Language and Text Based Research Data Infrastructure	Humanities

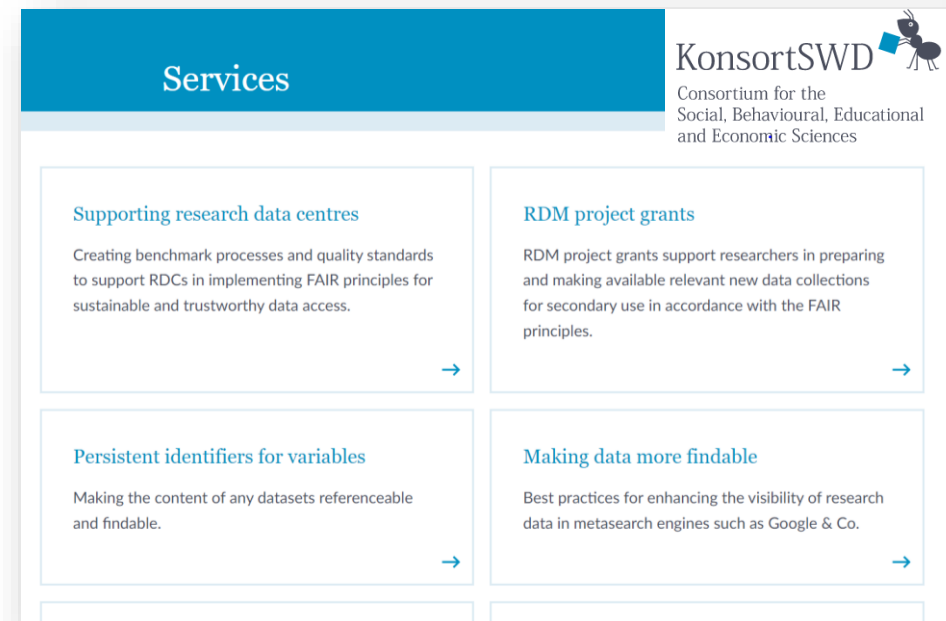
Snapshots from the NFDI Landscape



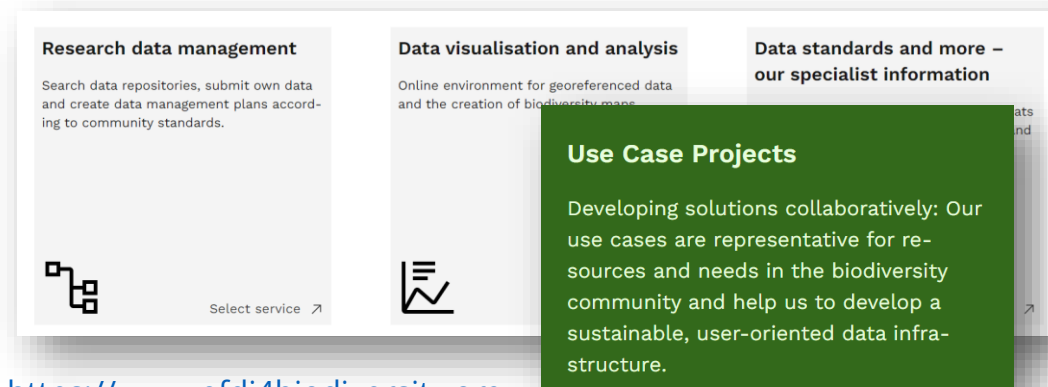
<https://www.nfdi4chem.de>



<https://nfdi4ing.de/>

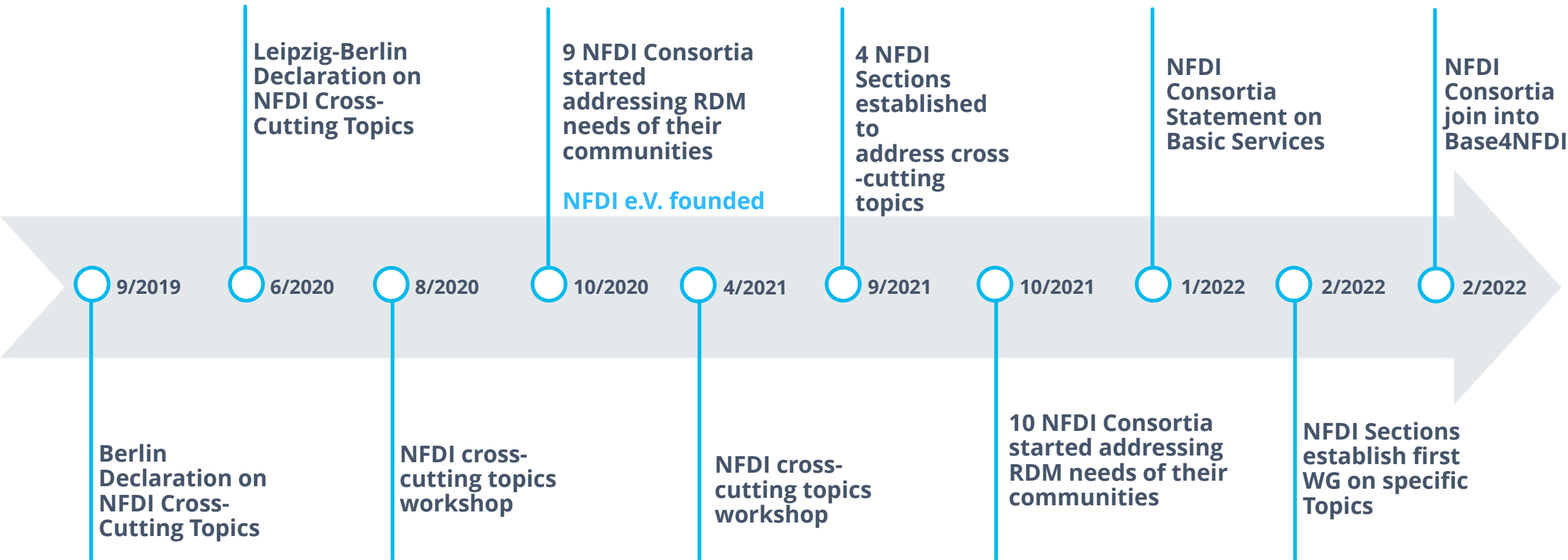


<https://www.konsortswd.de>



<https://www.nfdi4biodiversity.org>

Cooperation of NFDI Consortia – joining into *one* NFDI...

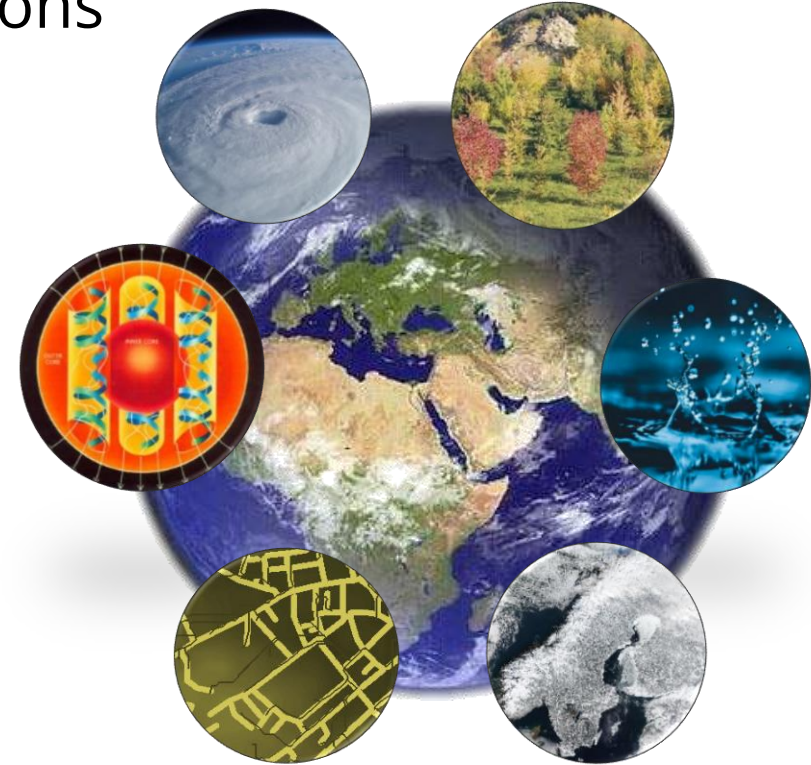




NFDI4Earth

**Research Data Infrastructure
for Earth System Sciences**

- Geosphere, Atmosphere, Biosphere, Hydrosphere, Cryosphere, Anthroposphere and all their interactions
- From Local Processes to Global Challenges
e.g. earth systems dynamics, natural hazards, climate change, environmental pollution, water scarcity, land use change, scarcity of raw materials
- Observing, measuring, modelling, analyzing, predicting the Earth System
- In international and interdisciplinary settings using spatio-temporal data as the common reference

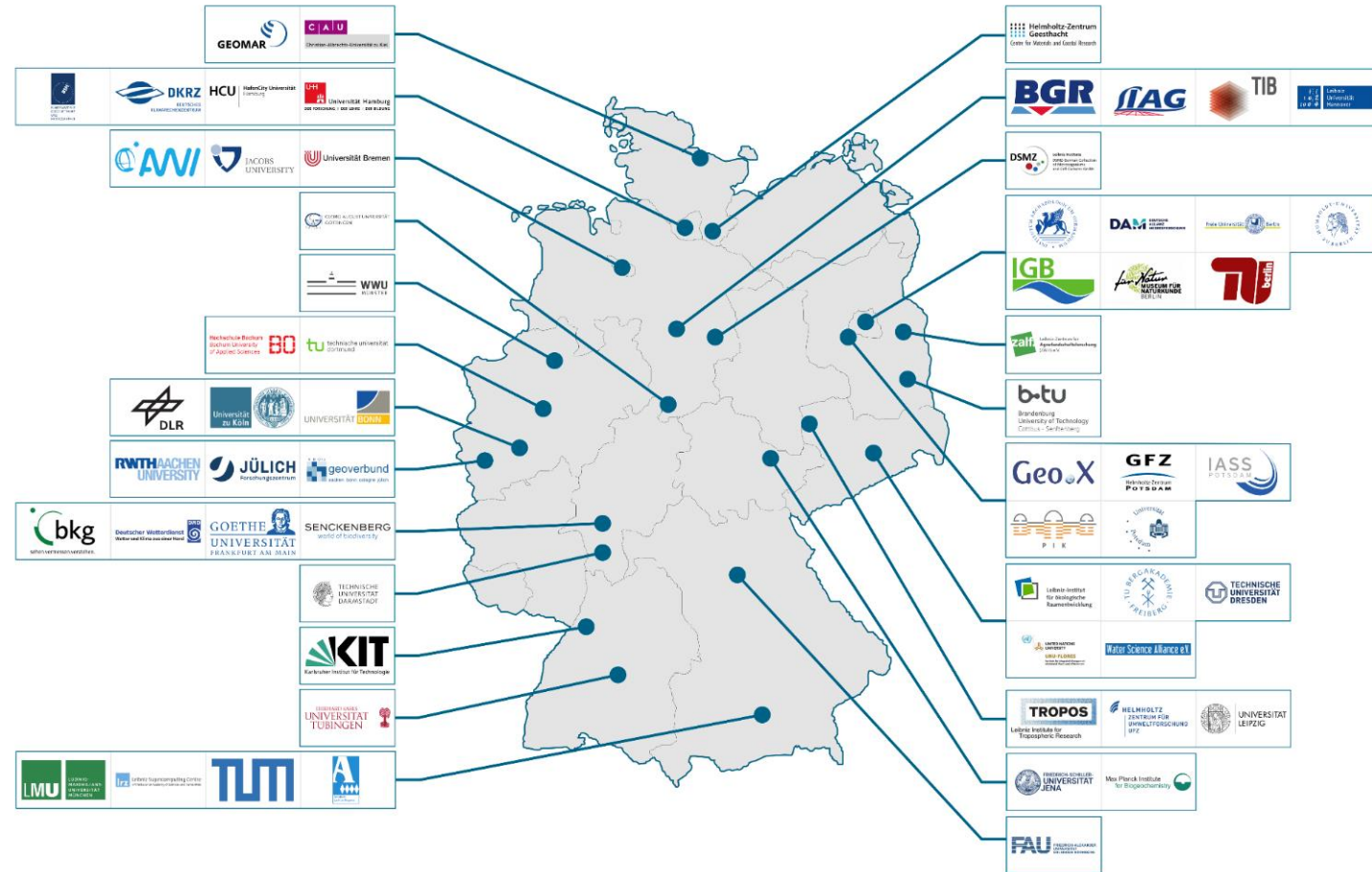


Who we are

- 58 institutions covering the breadth of Earth System Sciences (ESS)

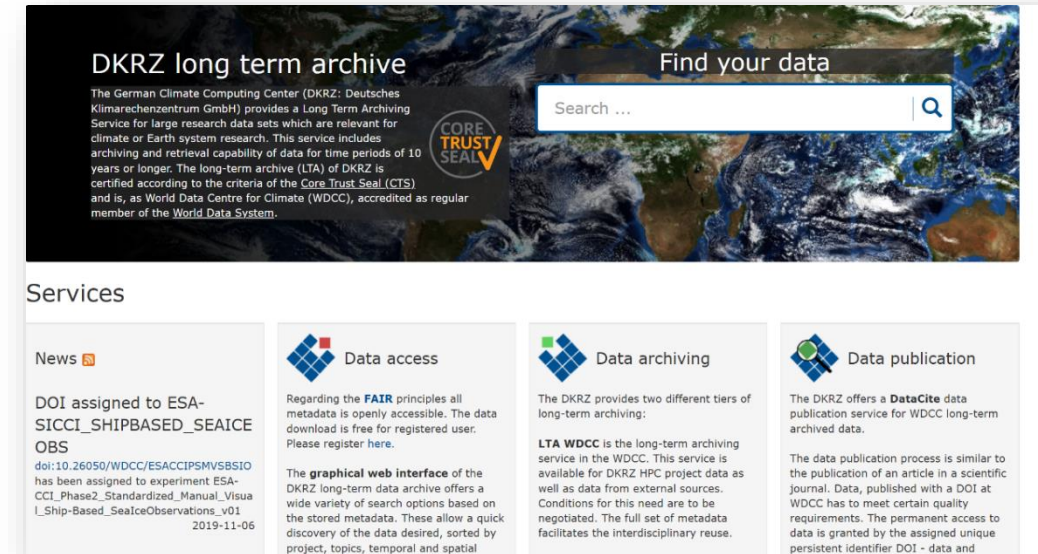
- Universities
- Research Organizations
- Infrastructure Providers
- Governmental Institutions
- Scientific Associations & Networks

- Established 2018 as **Open Consortium** and the ESS branch in the NFDI – **DFG funded since 10/2021**



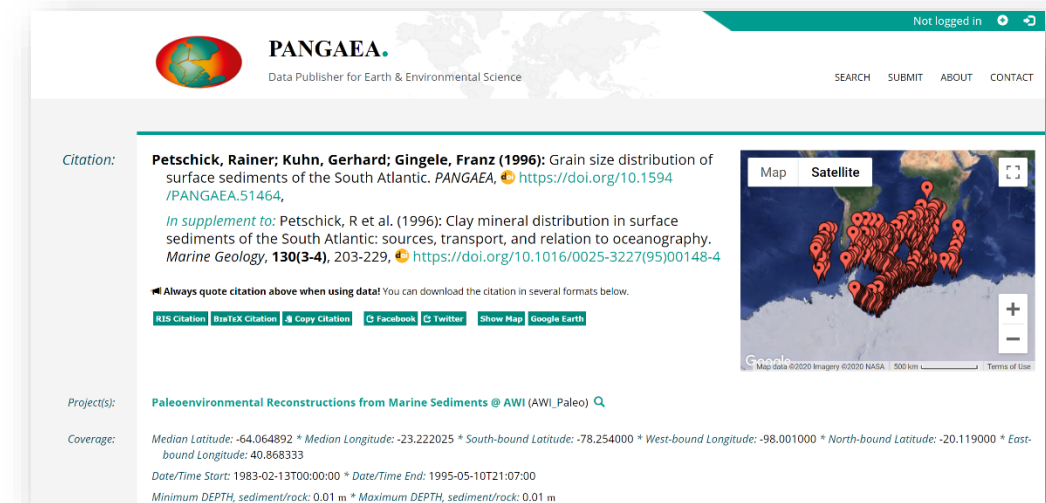
Where do we start in Earth System Sciences

- Data from different sources, growing openness
- Often demanding models and data intensive applications
- Positive attitude to Openness, FAIR, and Collaborative Development
- Established data repositories
- Several standards for interoperability



The screenshot shows the DKRZ long term archive website. The header features a search bar with the text "Find your data" and a search icon. Below the header, there is a section titled "Services" with four columns: "News", "Data access", "Data archiving", and "Data publication". The "News" column contains a link to a DOI assigned to ESA-SICCI_SHIPBASED_SEAICE OBS. The "Data access" column discusses the FAIR principles and provides a link to register. The "Data archiving" column describes the LTA WDCC service. The "Data publication" column explains the DataCite data publication service.

<https://cera-www.dkrz.de/>

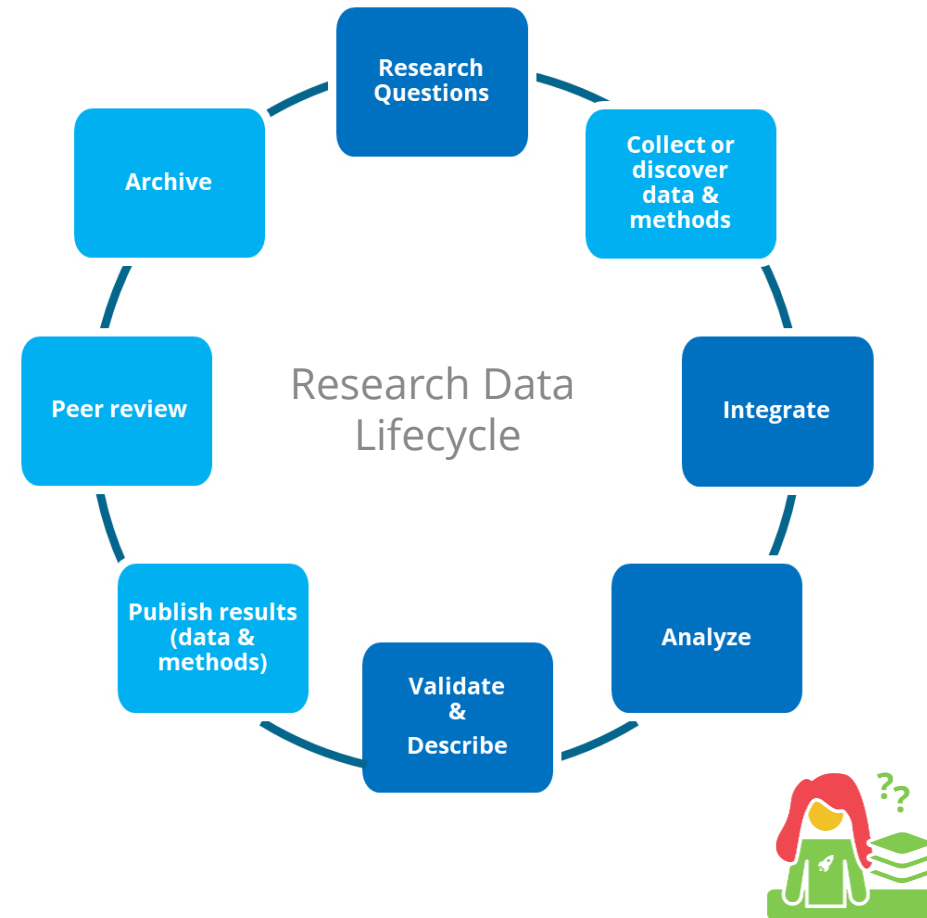


The screenshot shows the PANGAEA website. The header includes the PANGAEA logo and the text "Data Publisher for Earth & Environmental Science". The main content area displays a citation for "Petschick, Rainer; Kuhn, Gerhard; Ginge, Franz (1996): Grain size distribution of surface sediments of the South Atlantic. PANGAEA, <https://doi.org/10.1594/PANGAEA.51464>". Below the citation, there is a map showing the location of the data in the South Atlantic. The map is titled "Map" and "Satellite" and shows a distribution of red dots representing data points. The bottom section of the page provides project information, including the project name "Paleoenvironmental Reconstructions from Marine Sediments @ AWI (AWI, Paleo)" and coverage details such as median latitude, longitude, and date/time.

<https://doi.pangaea.de/10.1594/PANGAEA.51464>

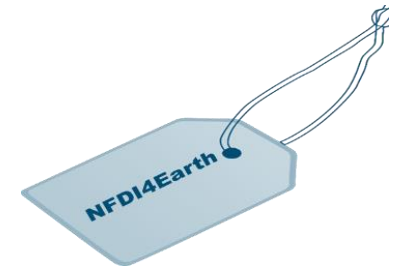
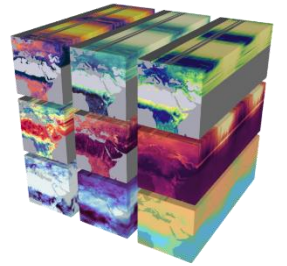
Research Data Management - Our Issues

- Large, diverse number of data services (>> 100) and activities related to ESS...only a few sustainable
- Data heterogeneity
quality, curation levels, licenses, semantics, sizes, repeatability, scale, spatio-temporal resolutions,...
- Different data cultures and velocities to FAIR
from data streams to long tail data; RDM experts and novices
- Incomplete support along the data lifecycle
- Lack of support (platforms, tools) for joint and collaborative interpretation of heterogeneous and decentralized data

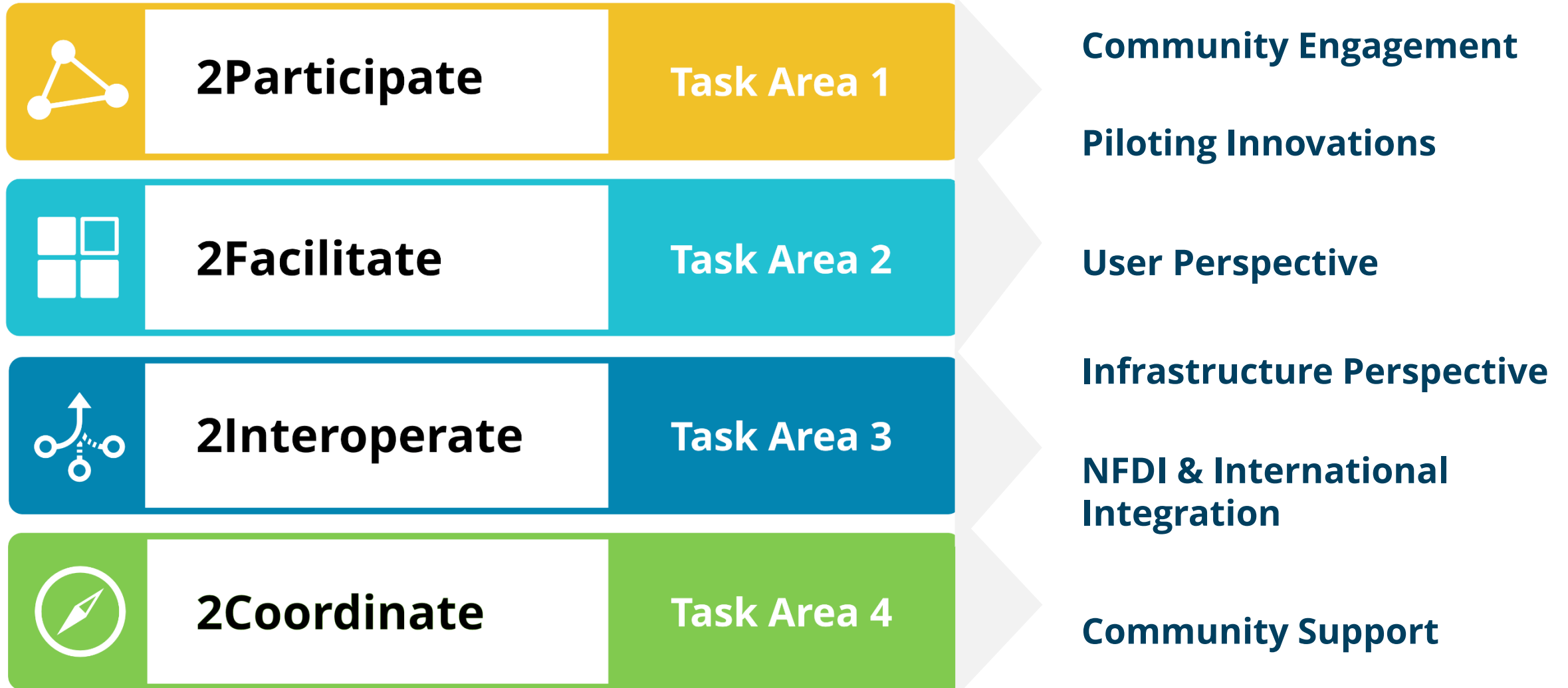


Key Goals

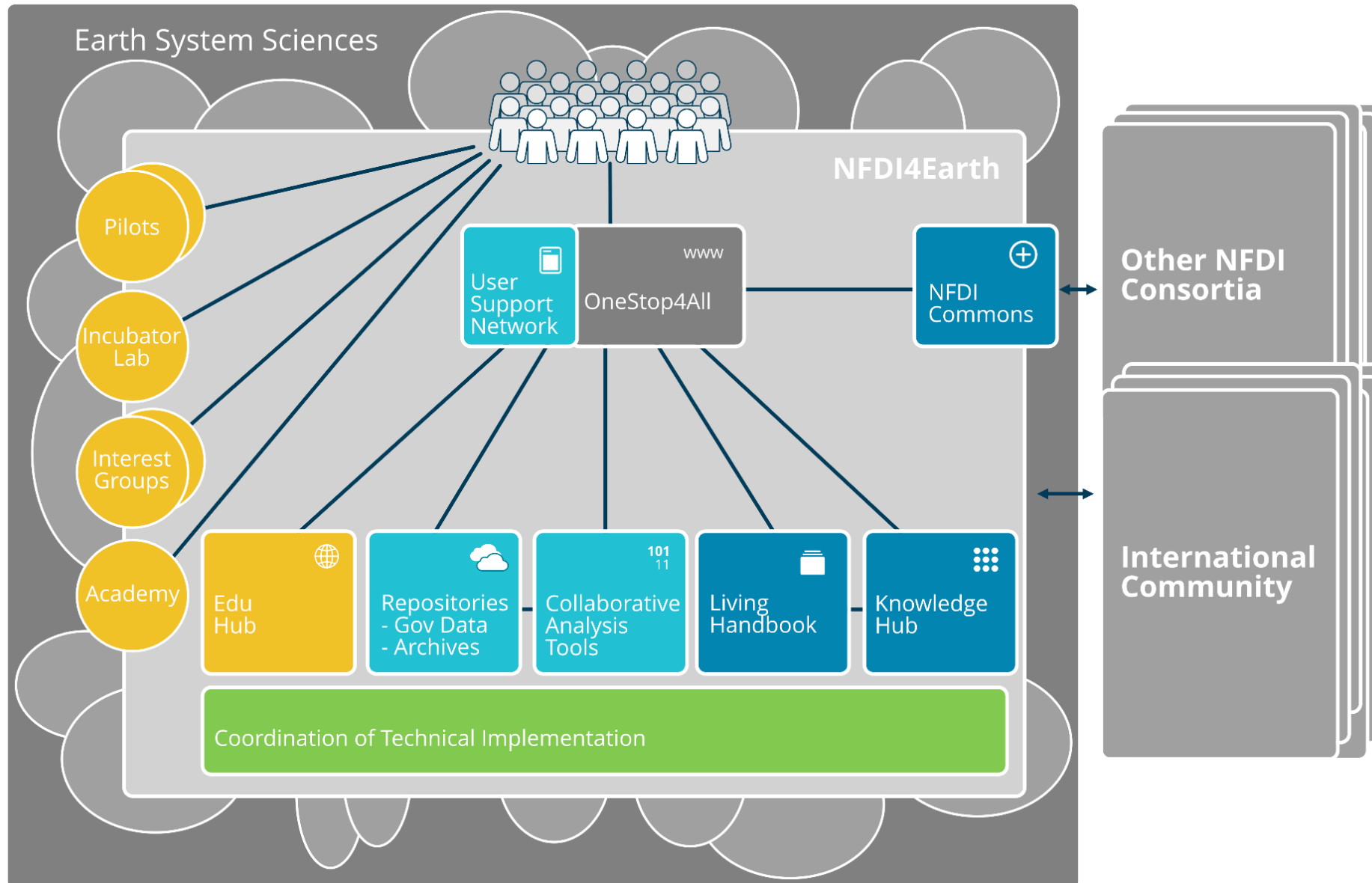
- *One Community* approach to sustainable, Open and FAIR Research Data Management in Earth System Sciences (ESS)
- Community driven agile development of innovative platforms for data integration and collaborative data analysis
- Qualification for people, data, tools and services as a basis for FAIR RDM and viability
- *OneStop4All* and *User Support Network* for ESS RDM as integral part of NFDI and International Infrastructures
- Key-driver of the build-up and operation of the NFDI



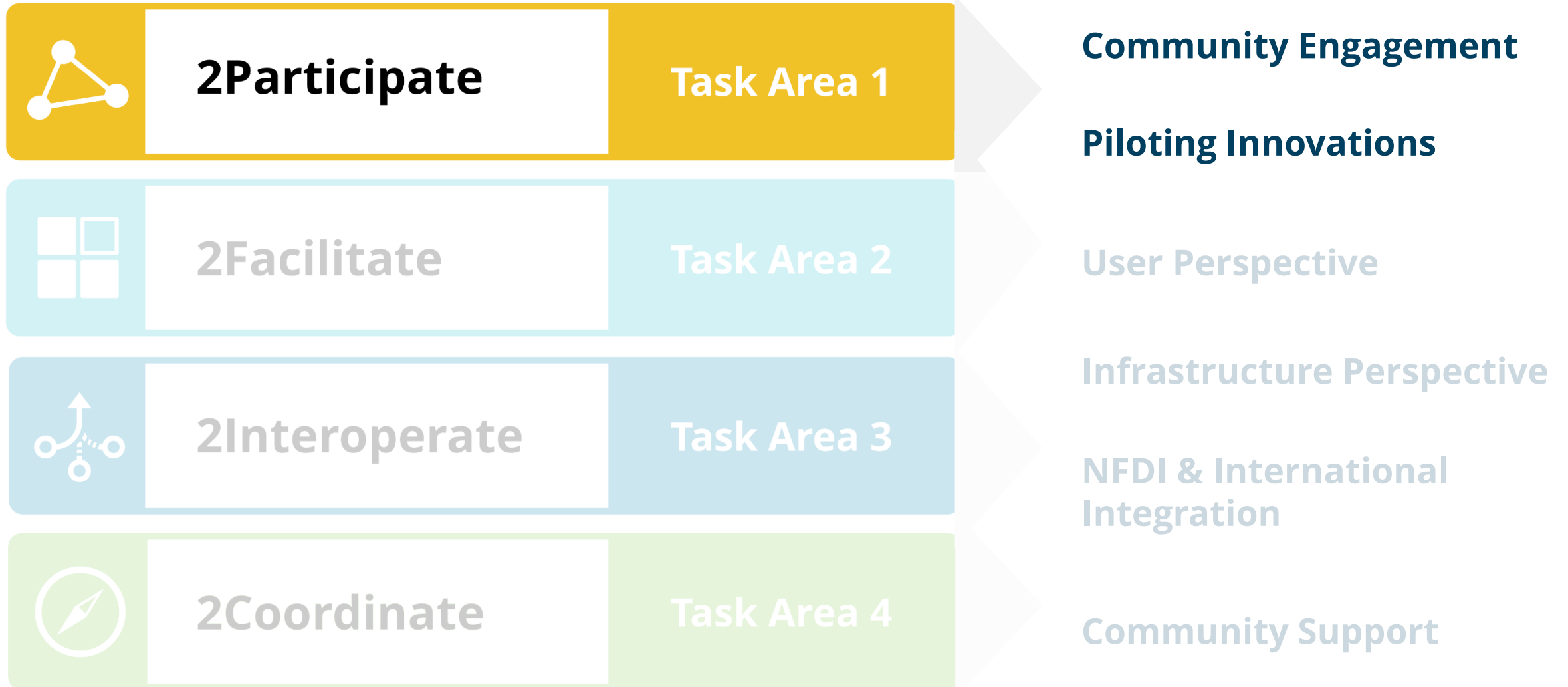
NFDI4Earth – Strategy 2021-26



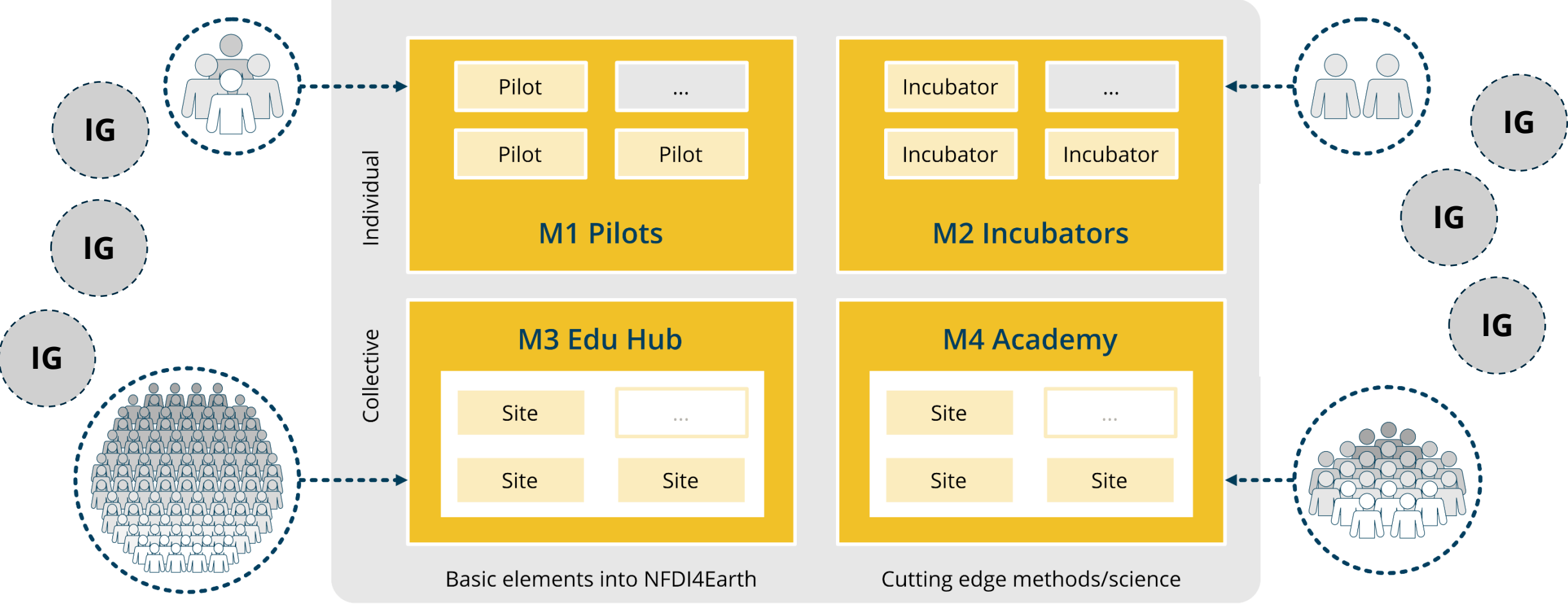
Implementing the NFDI4Earth Strategy



NFDI4Earth – Strategy 2021-26

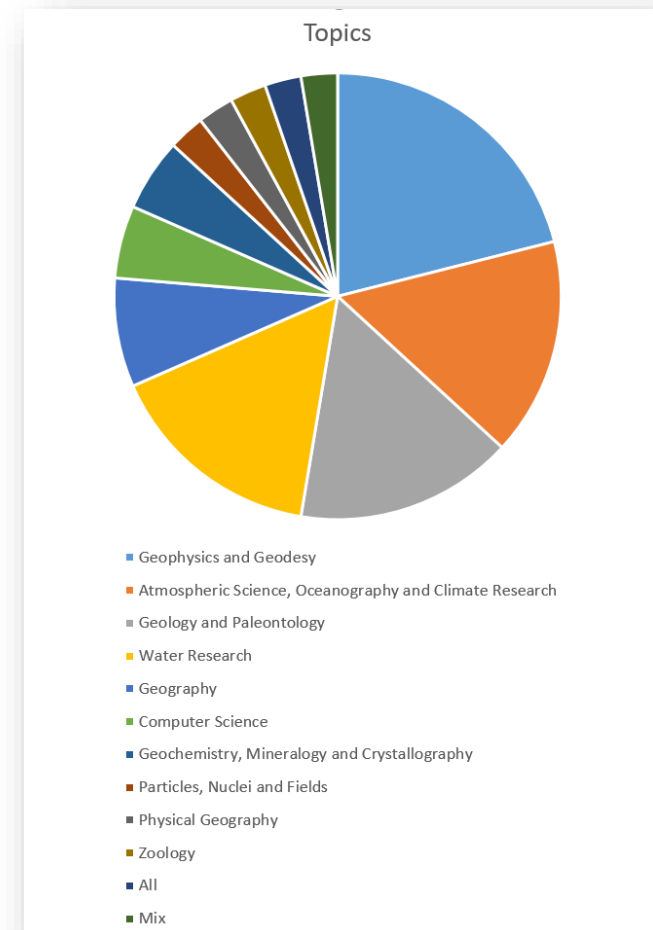
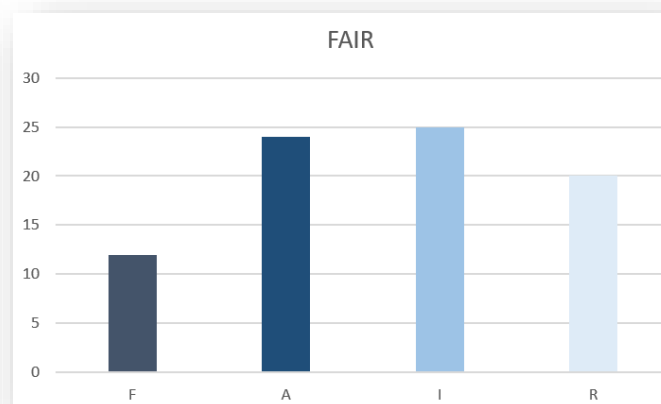
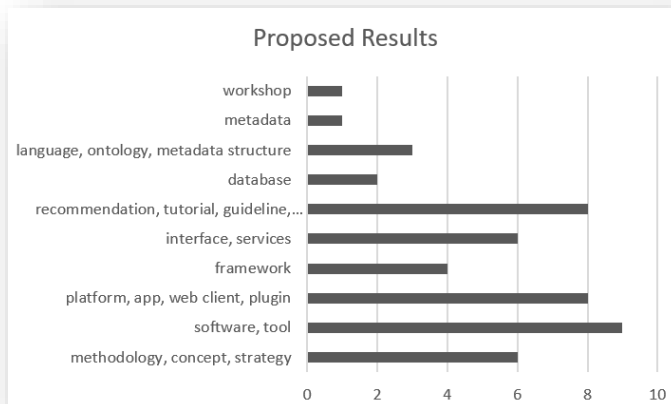


Engage with the community, foster innovation

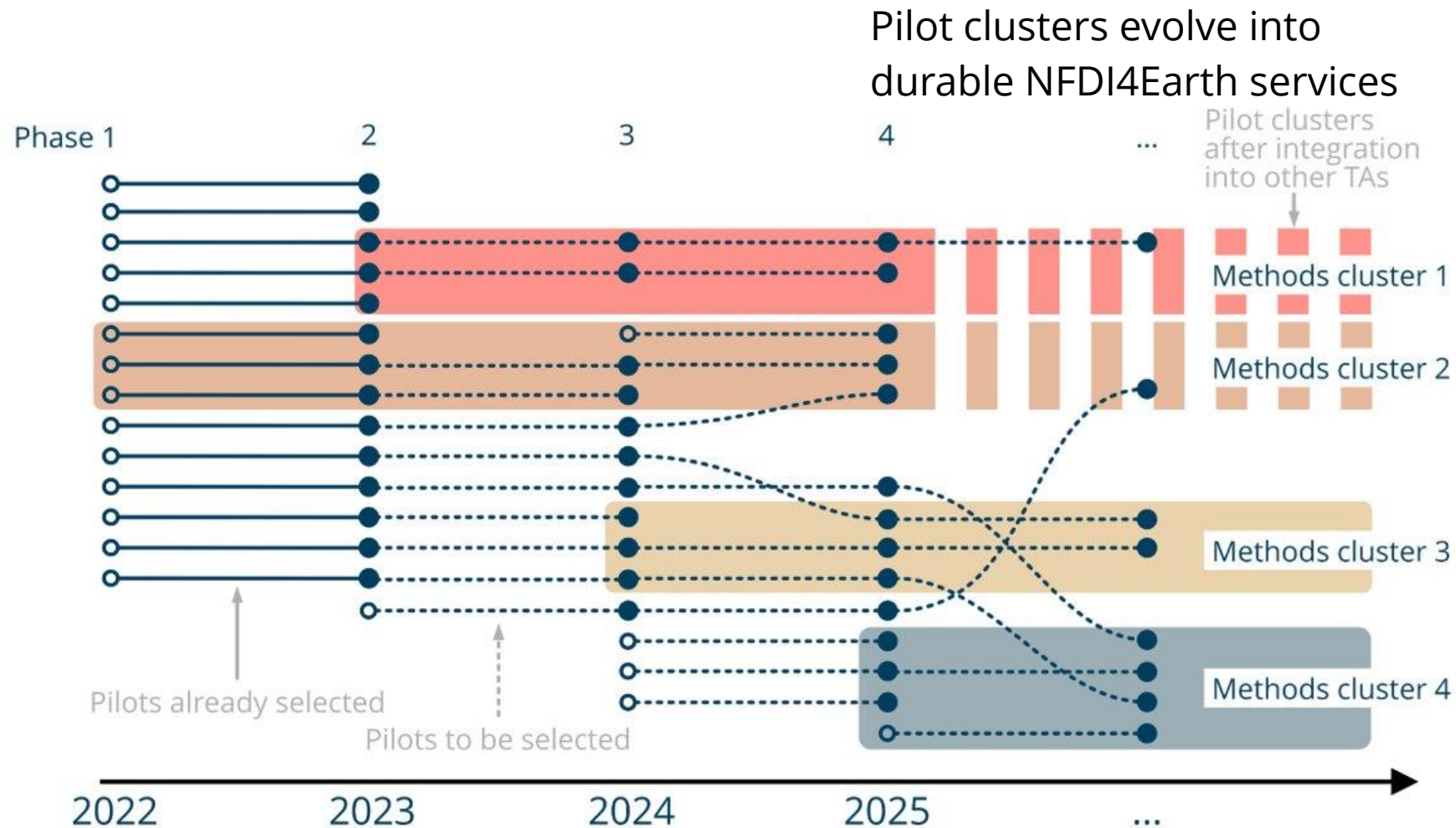


1st Open Call for Earth System Science Pilots (2020)

- Received 38 applications for one year pilots
- 14 pilots selected in a transparent, criteria-based review process to start 6/2022
- Future open calls, evaluations and funding recommendations are prepared: ensuring coverage over all objectives, all related communities and all types of data repositories)



Pilots: from individuals to clusters ...



Future pilot calls can be directed, if gaps are identified.

One example pilot cluster *Data Cubes*

**Statistical Learning
on data cubes.**

Winkler et al.



**Model evaluation in
data cubes.** Eyring et al.



**EO4Glaciers data
cube.** Braun et al.



■ **Challenges:**

- Bridging different data models
- Enabling FAIR principles
- Ready for the future (e.g. clouds)

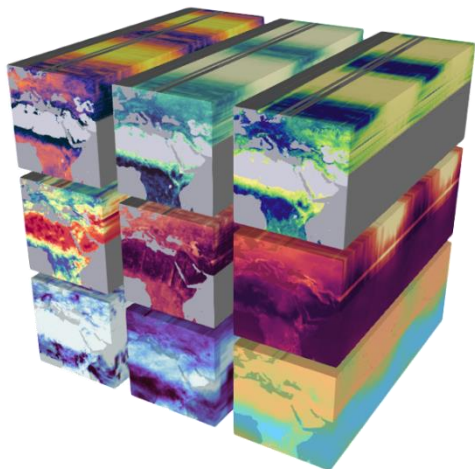
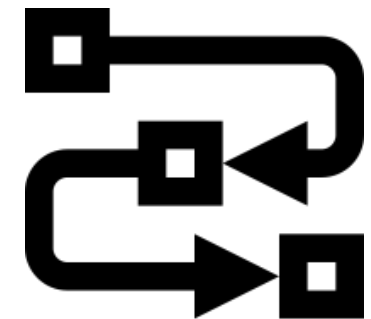
**Socioeconomic
data cubes.**

Kraemer et al.

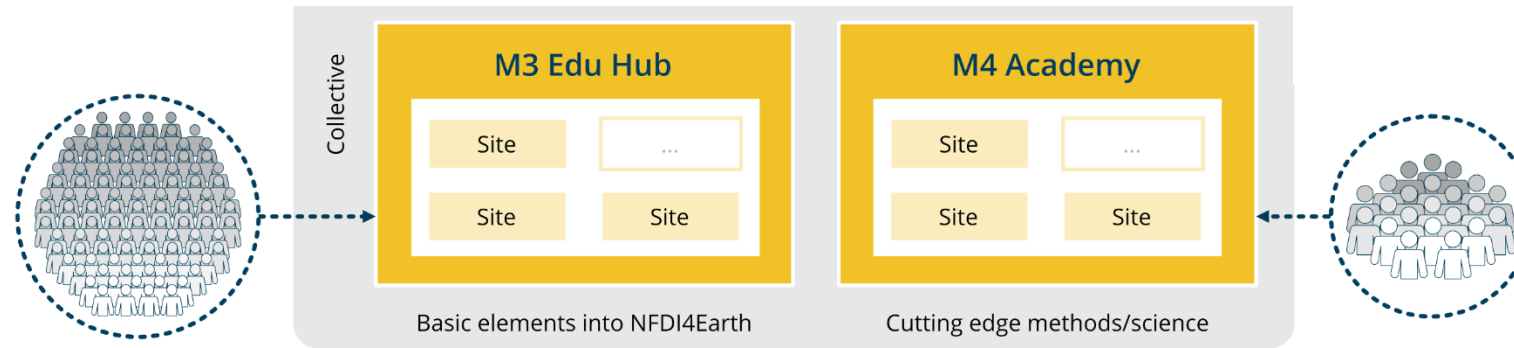


**Processing and
Visualization.**

Unnithan et al.

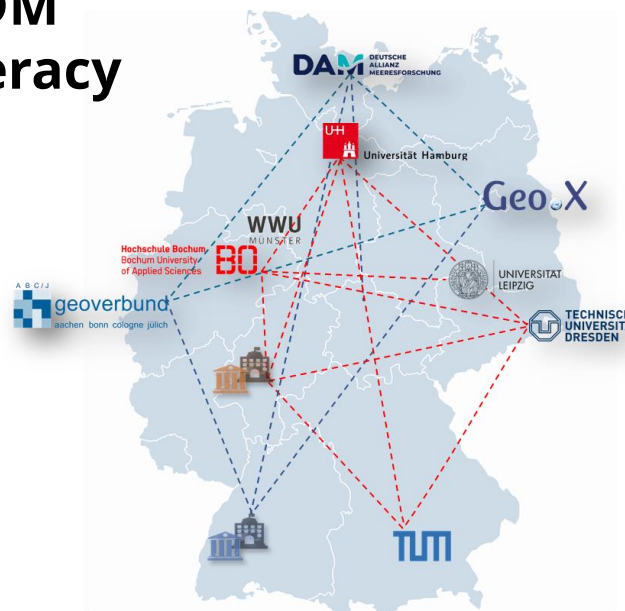


Education at all Levels



Edu Hub: Develops competences in RDM and spatiotemporal data literacy

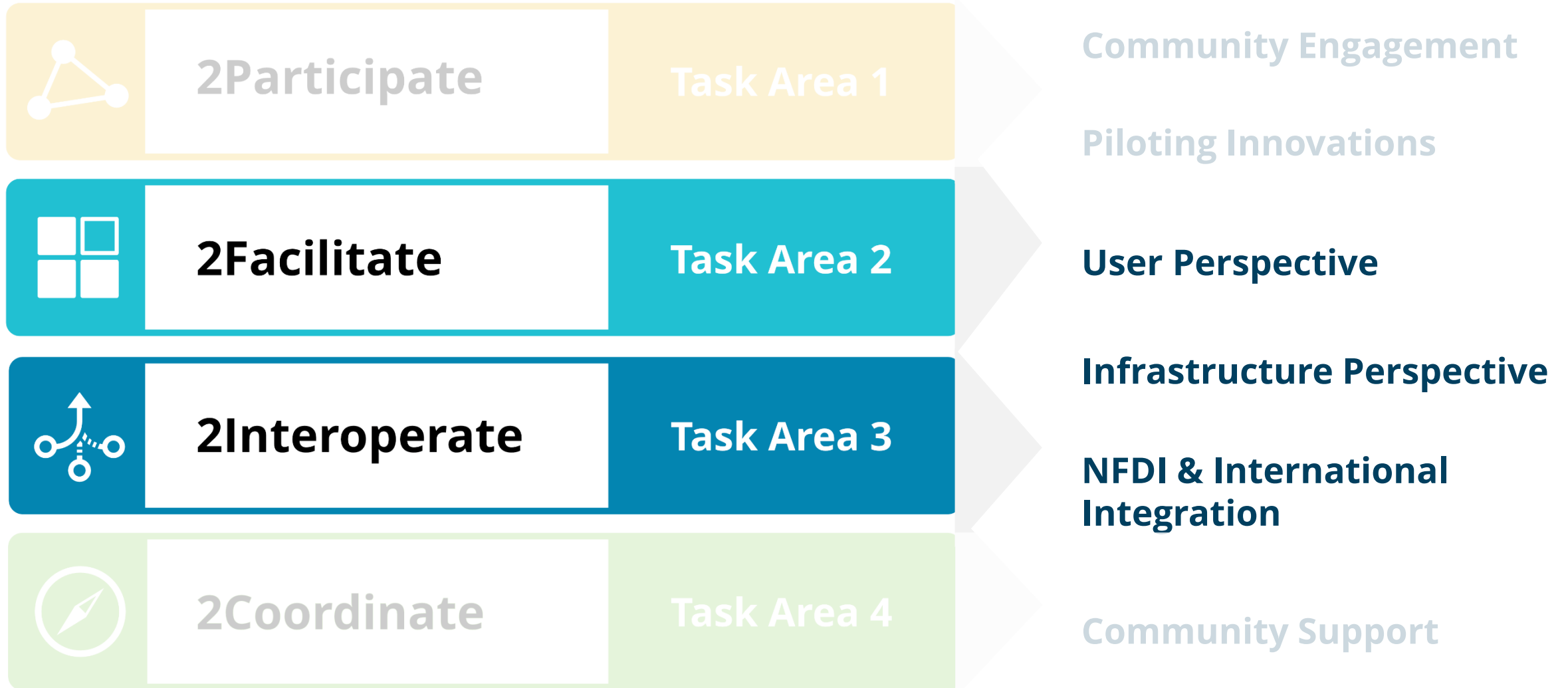
- Wide range of ready to use open educational resources
- New Qualification Profiles (e.g. *research data curator*)
- Network Education & Training Sites



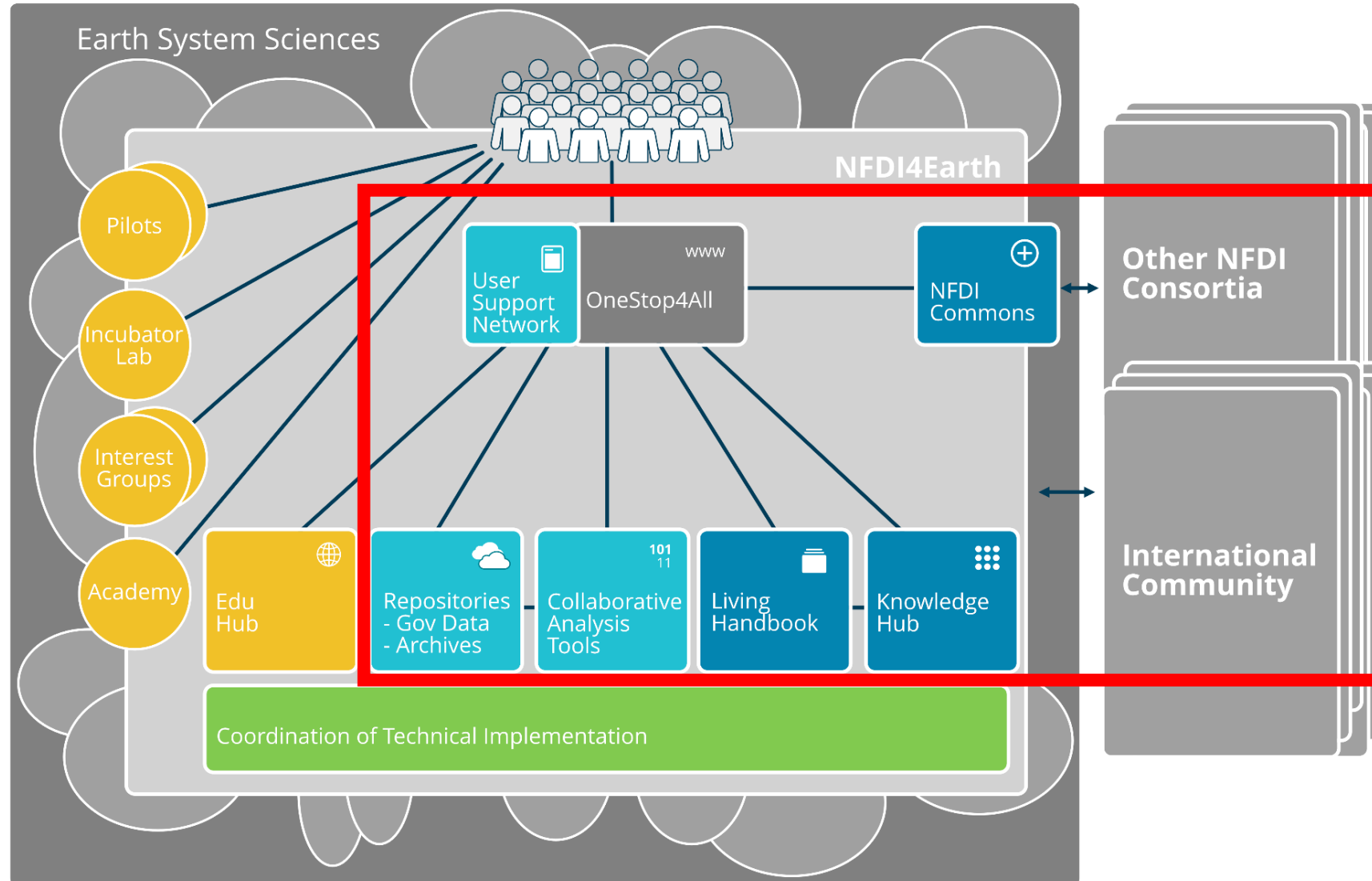
Academy: Explore novel approaches

- Interconnect junior researchers in data driven applications as users of NFDI4Earth resources
- Trigger proofs of concept
- Network Academy Sites

NFDI4Earth – Strategy 2021-26



Implementing the NFDI4Earth Strategy



Two perspectives

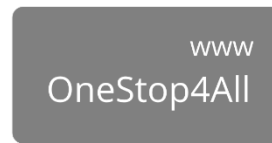
1. User
2. Provider / Infrastructure

Individual User Perspective on RDM Support



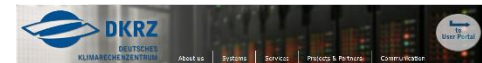
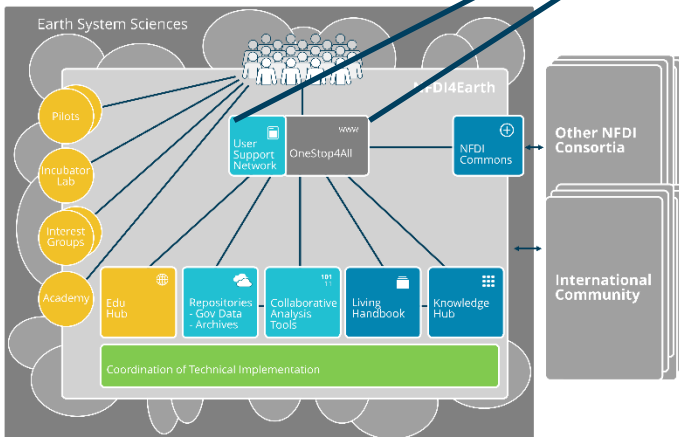
ESS researcher, student,...

Developed, operated and maintained by NFDI4Earth

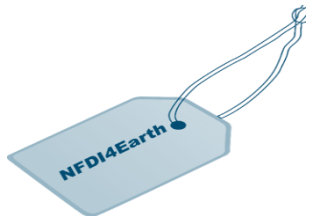
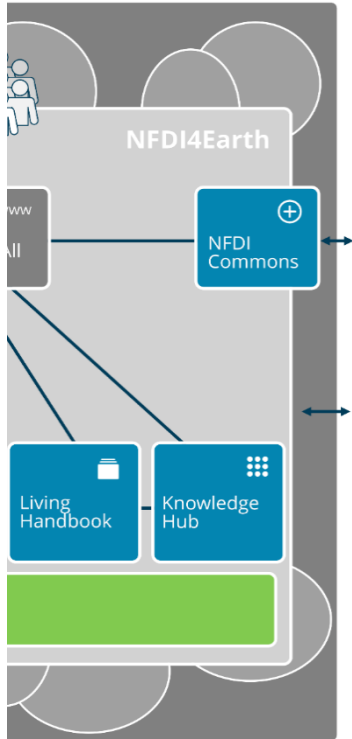


Enabled by NFDI4Earth

Contributed by NFDI4Earth participants,



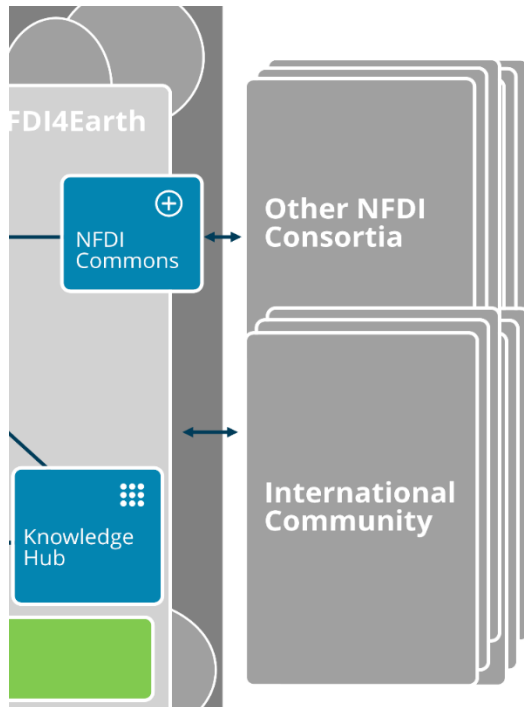
Provider/Infrastructure Perspective: Interoperability and Integration



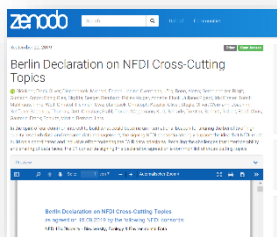
- **Sustainable Architecture to interoperable combine and link services** of NFDI4Earth partners
 - Identify (core) components and services
 - Define a set of **Common Standards for FAIR ESS RDM**
 - Continuously contribute to / align with NFDI
- Establish **NFDI4Earth Knowledge** Hub (on services, standards, ontologies/vocabularies, etc.) as a common backbone
- Establish **self-evaluation process for ESS RDM service offerings** (building on EOSC/FAIR Metrics, RISE...)
- Develop **NFDI4Earth Label** as key driver and indicator for interoperability of services

Contributing to the NFDI Process

- NFDI4Earth specific contributions:



- Knowledge hub for accessing and **dealing with geo-spatio-temporal data**
- Long-standing active participation in various **relevant international networks** (AGU, EGU, RDA, EOSC,...OGC, WMO...)
- Collaboration with **governmental agencies** and other national stakeholders (GDI-DE, BGR, BKG, DWD,...)
- Widespread experiences in **contributing to international standards for geodata and services** (Metadata, Data Encodings, APIs)
- Well-established culture and skills concerning **community software developments**
- Representing a number of the German HPC centres **contributing to the envisioned scientific computing cloud**

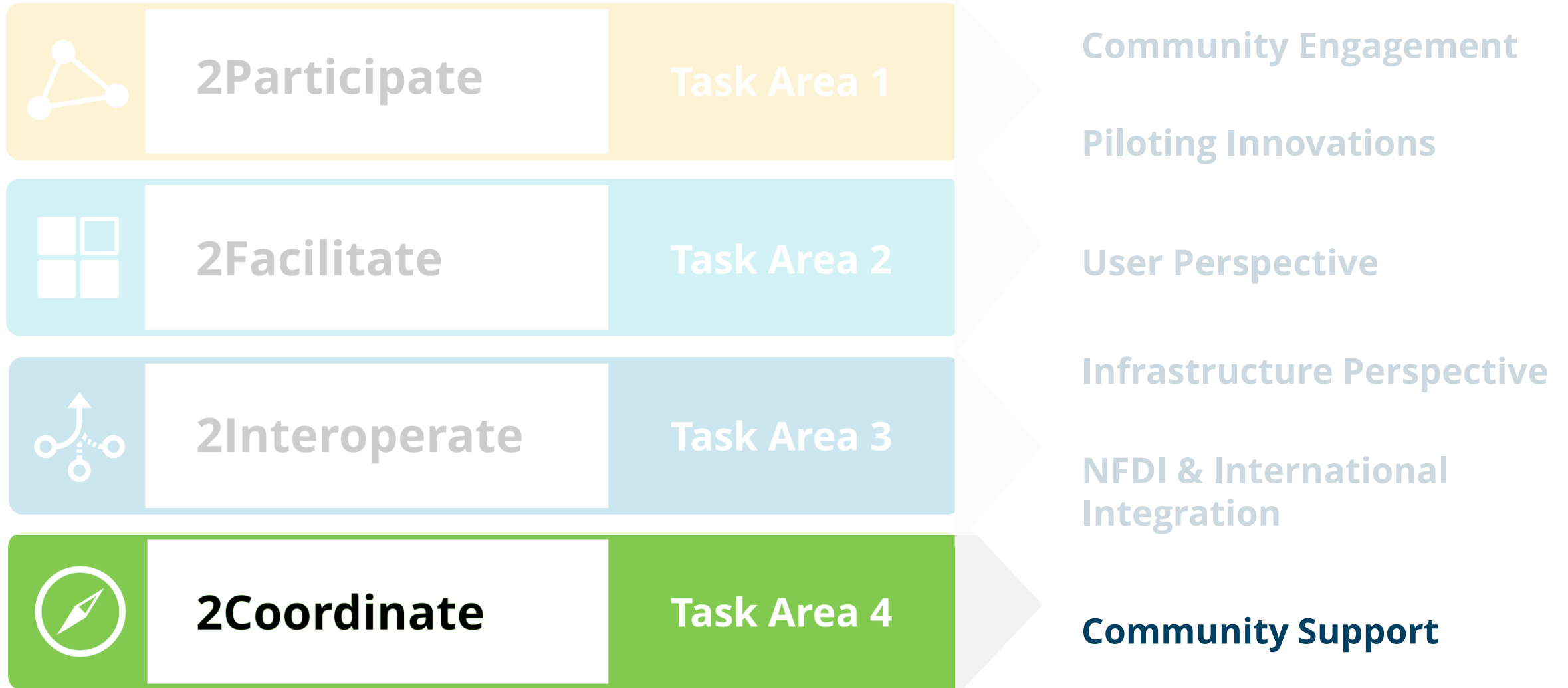


<https://doi.org/10.5281/zenodo.3457213>

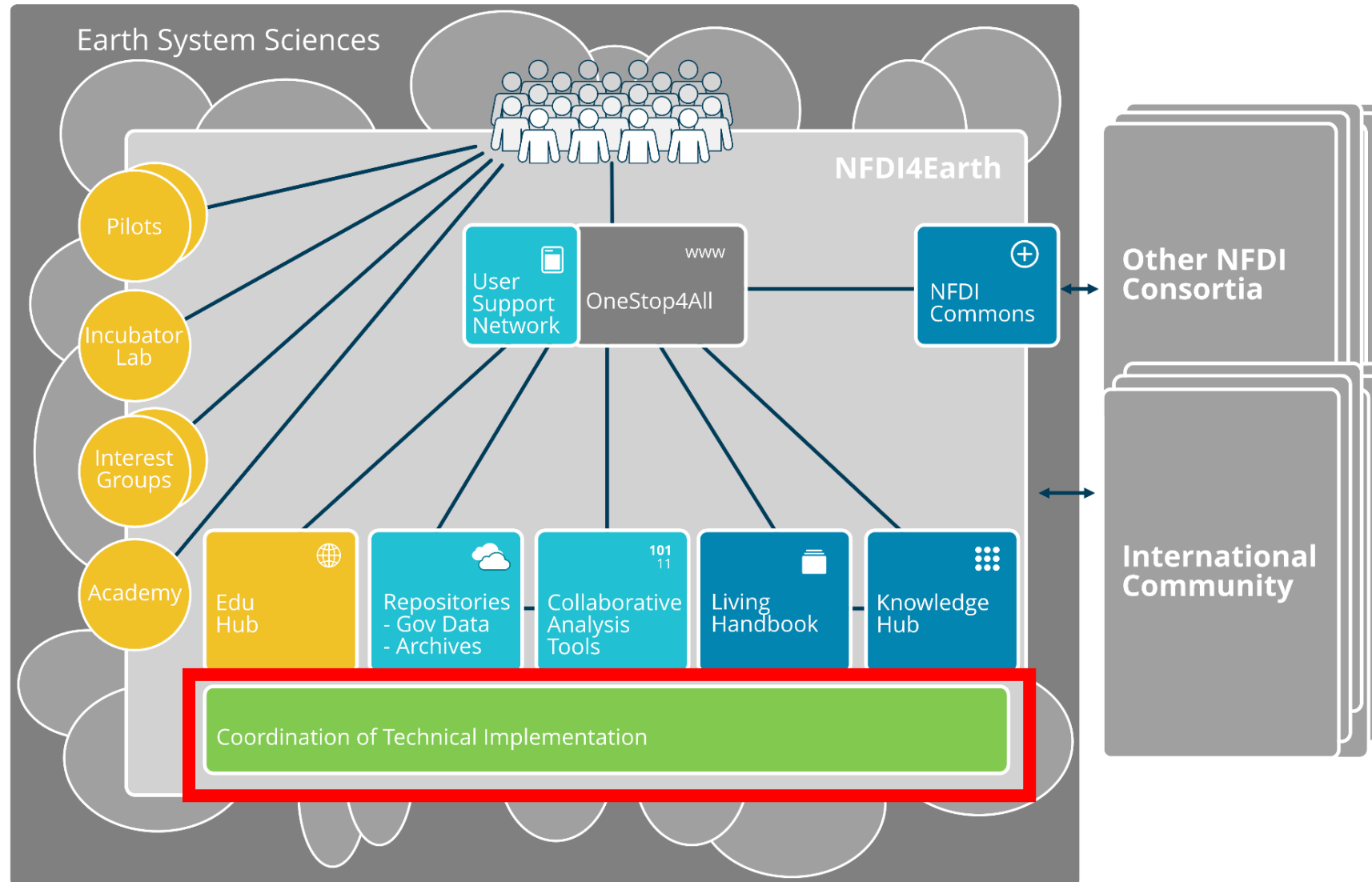


<https://zenodo.org/record/3895209>

NFDI4Earth – Strategy 2021-26



Central Support Services for the federated NFDI4Earth




- Consolidate and support implementation for OneStop4All, Knowledge Hub, Living Handbook
- Provide Virtualization Services, Ticketing System, EduHub platform

- **NFDI4Earth FAIRness and Openness Commitment**
 - Engage all ESS stakeholders (researchers, societies, infrastructures, publishers,...)
 - Foster incentives for Open and FAIR RDM in ESS and **trigger cultural change**
- Proceed as **one national ESS Community Effort**
 - Link into and benefit from related initiatives (e.g. Helmholtz Research Field Earth and Environment DataHub, GDI-DE,...)
 - Become **integral part** of NFDI, EOSC and international infrastructures
- Develop a **common long term operation model for NFDI4Earth services** building on partners' contributions
- Provide an **inspiring and vibrant environment for novel Earth System Data Science approaches**



Get Involved !

- Join the Interest Groups (**always open**)
- Launch a new Interest Group (**always open**)
- Join the Consortium (**from 6/2022 onwards**)
- Join the annual
NFDI4Earth Plenary Meetings
- Apply for an NFDI4Earth Incubator
1st Call just closed – 2nd call in 2023
- Apply for an NFDI4Earth Pilot **2nd Call will be launched 2023**



Become part of the NFDI4Earth activities!

NFDI4Earth is designed as an open consortium. If you are interested in participating and shaping NFDI4Earth, we invite you to initiate an Interest Group or join an existing one:

[Read the guidelines for detailed information](#)

The following interest groups are endorsed by the Steering Group of NFDI4Earth (as at April 2022)

- [Geology & Geophysics](#)
- [High-performance Computing in Earth System Sciences](#)
- [Machine Learning](#)

and others are being prepared or currently under review

- [\(Geo\)Research Software and Reproducibility](#)
- [FAIR Data](#)
- [Standardisation of geochemical Analysis Data](#)

Further ideas were considered at an early stage during the 4th plenary in Potsdam, but have not yet been implemented

- > [Data Management Plan](#)
- > [Industry & Company Data](#)

If you are interested in one of the topics and interest group resp. mentioned above please click on the corresponding IG!

Thank You !



<https://www.nfdi4earth.de/>