

Incubator Platform Helmholtz Metadata Collaboration (HMC)

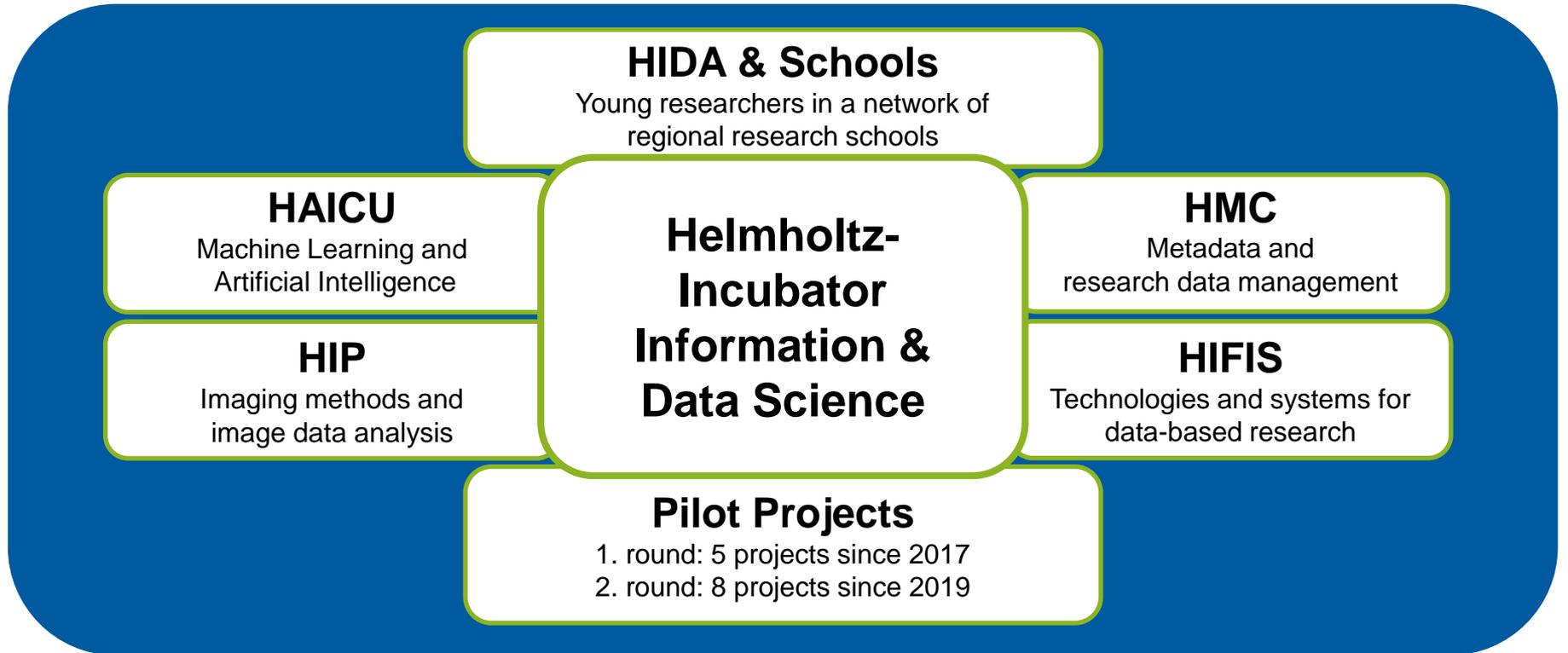
Dr. Wolfgang Süß
Karlsruher Institute of Technology
Institute for Automation and Applied Informatics

- 1) HMC Mission
- 2) Structure & Goals



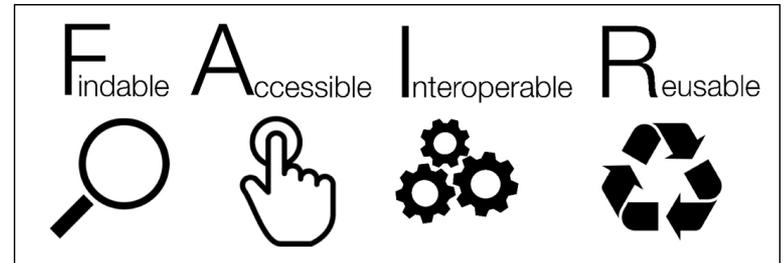
Properly documenting a data set is the key to preserving its usefulness through time.

Bundling all activities in a powerful framework

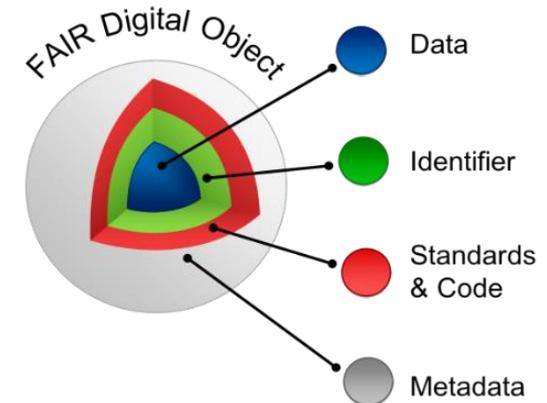


The mission of the Helmholtz Metadata Collaboration (HMC) is to **support researchers to find, access, machine evaluate and reuse** research data from the Helmholtz Association *for advanced methods of data processing*.

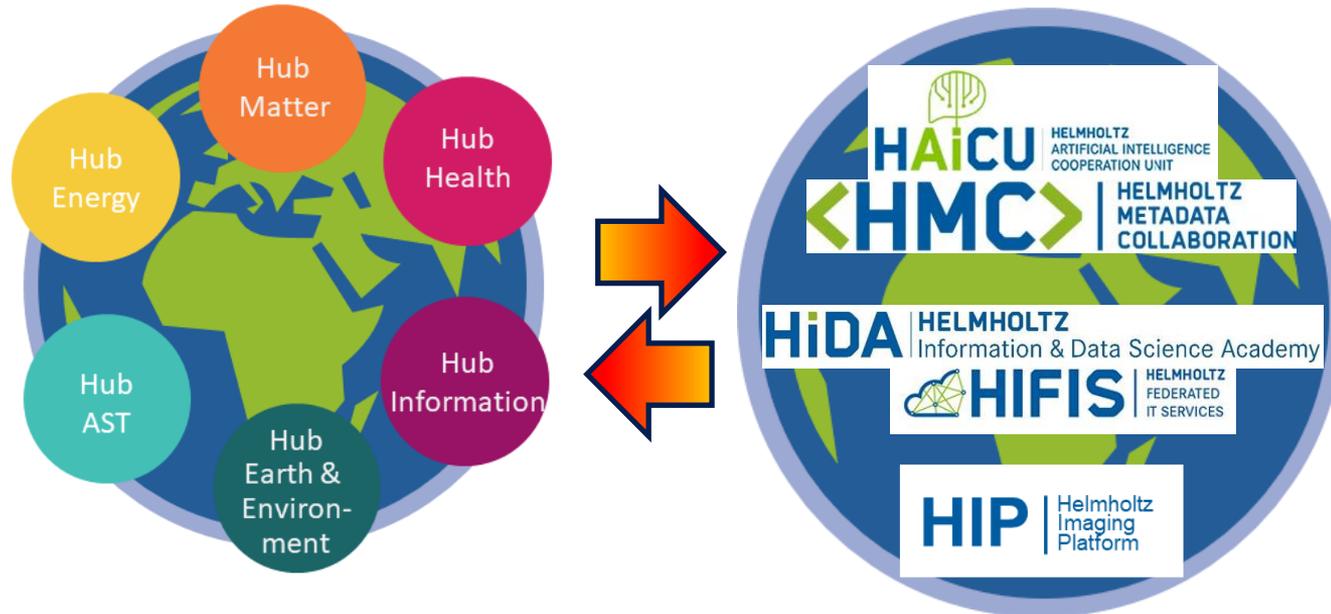
To achieve this, concepts and services are being developed and established, **enabling the enrichment of research data** during the various phases of their creation by means of a suitable, standard-compliant description with metadata.



- Make Helmholtz' Data **FAIR** - findable, accessible, interoperable and reusable
- Provide comprehensive and **sustainable services**, consulting, information and tools for efficient metadata handling as a **distributed shared facility**.
- Jointly develop, share and **consolidate community-expertise** for metadata of the six Helmholtz research areas.

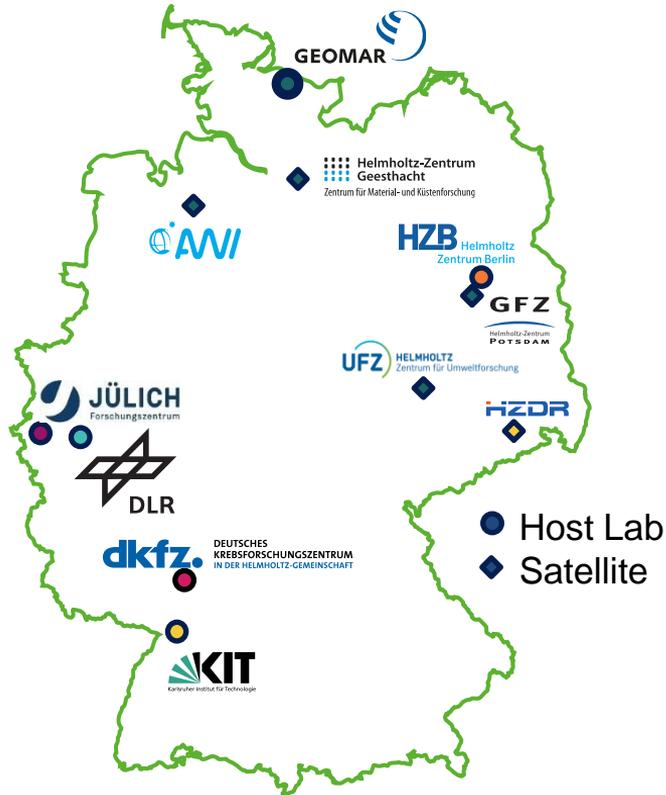


two sides of the same medal



**Helmholtz
Disciplinary
World**

**Helmholtz
Digital
World**



Berlin
Helmholtz-Zentrum Berlin für Materialien und Energie (HZB)

Bremerhaven
Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI)

Dresden
Helmholtz-Zentrum Dresden-Rossendorf (HZDR)

Geesthacht
**Helmholtz-Zentrum Geesthacht
Zentrum für Material- und Küstenforschung (HZG)**

Heidelberg
Deutsches Krebsforschungszentrum (DKFZ)

Jülich
Forschungszentrum Jülich

Karlsruhe
Karlsruher Institut für Technologie (KIT)

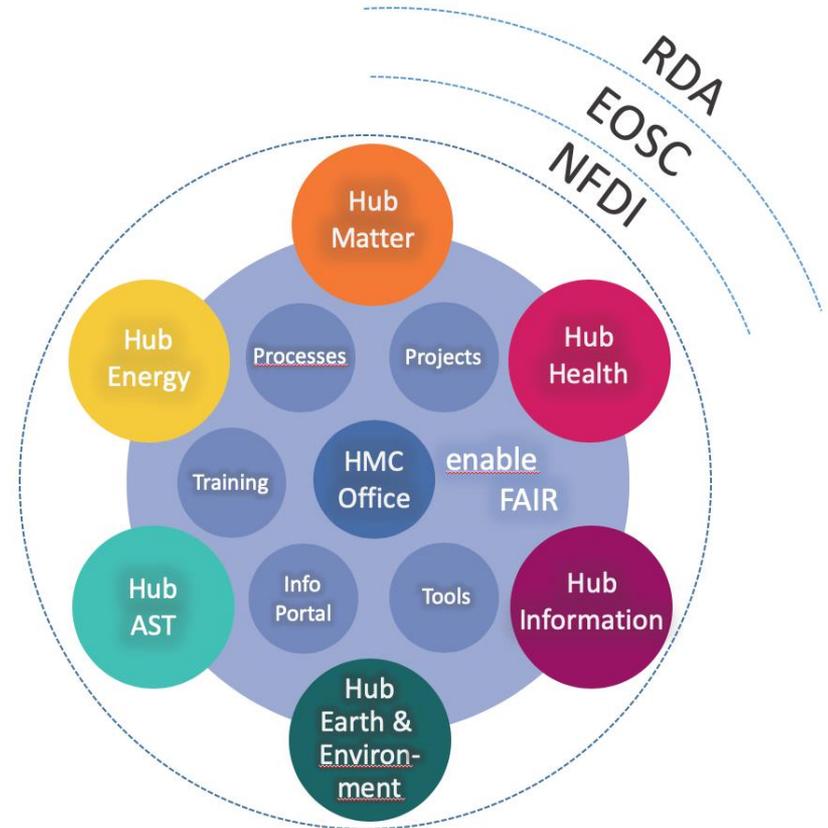
Kiel
**GEOMAR Helmholtz-Zentrum für Ozeanforschung
Kiel**

Köln
Deutsches Zentrum für Luft- und Raumfahrt (DLR)

Leipzig
Helmholtz-Zentrum für Umweltforschung – UFZ

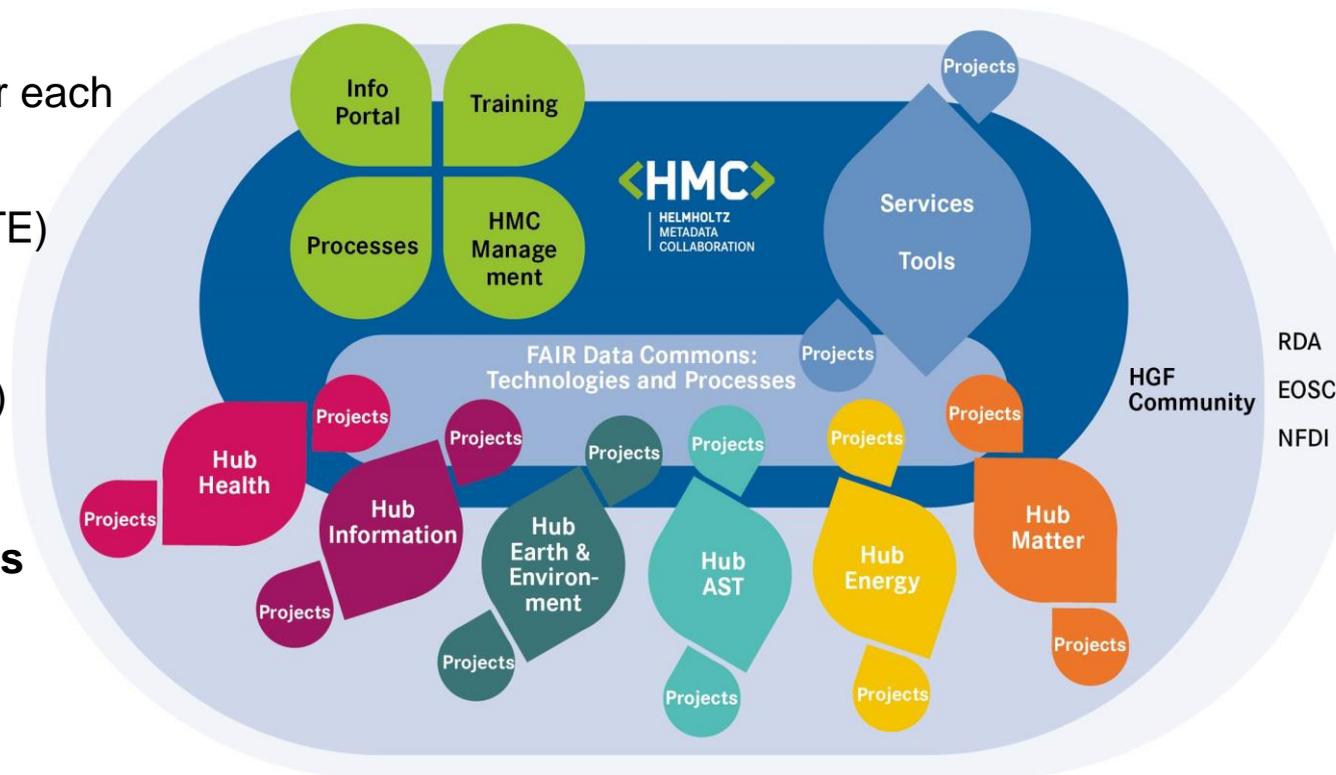
Potsdam
**Helmholtz-Zentrum Potsdam –
Deutsches GeoForschungsZentrum GFZ**

1. Establish **domain specific hubs** for each research field and determine demand.
2. Provide **standards, best practices, processes and tools** for researchers
3. Support researchers in the **organization of research data** and the use of available tools and (infra-)structures
4. Set up a framework to **integrate all developments** of different research fields



Structural Components

1. **domain specific hubs** for each research field (6x5 FTE)
2. **HMC service office** (4 FTE)
3. **HMC enable FAIR** (8FTE)
4. **HMC community projects** (1.2M€/y)



Structural components – central services

HMC Office

GEOMAR

Management of the strategic and operational processes of the platform development, incl. annual project calls and reporting.

Basic technical services and tools

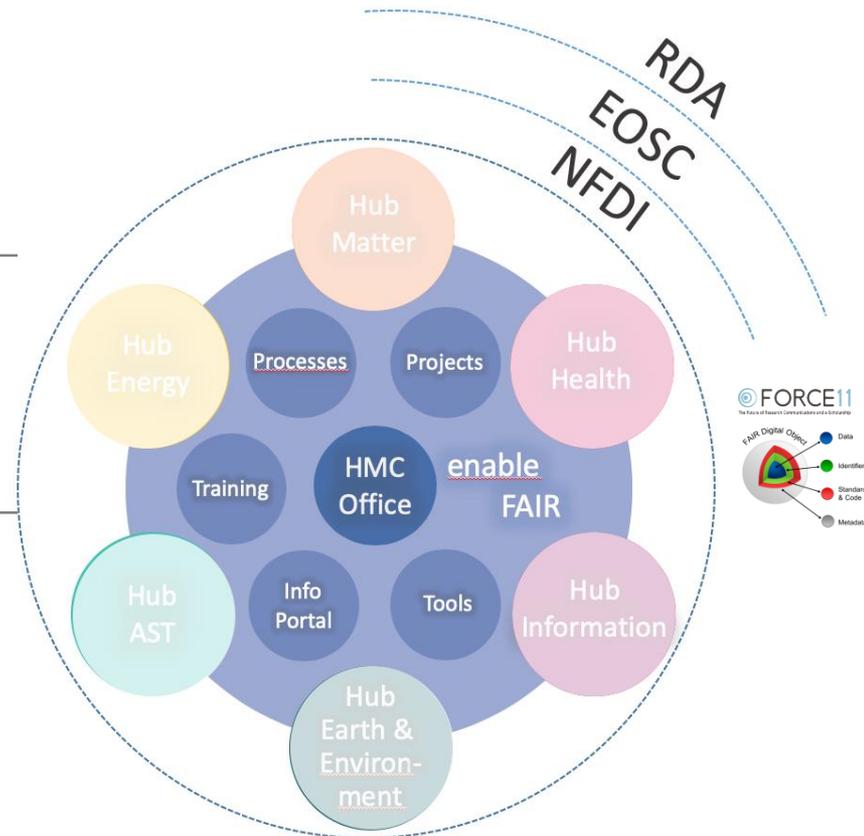
KIT

Provides HMC-wide technical services and tools, incl. **interoperability** to national and international infrastructures

Technically enabling FAIR

FZJ

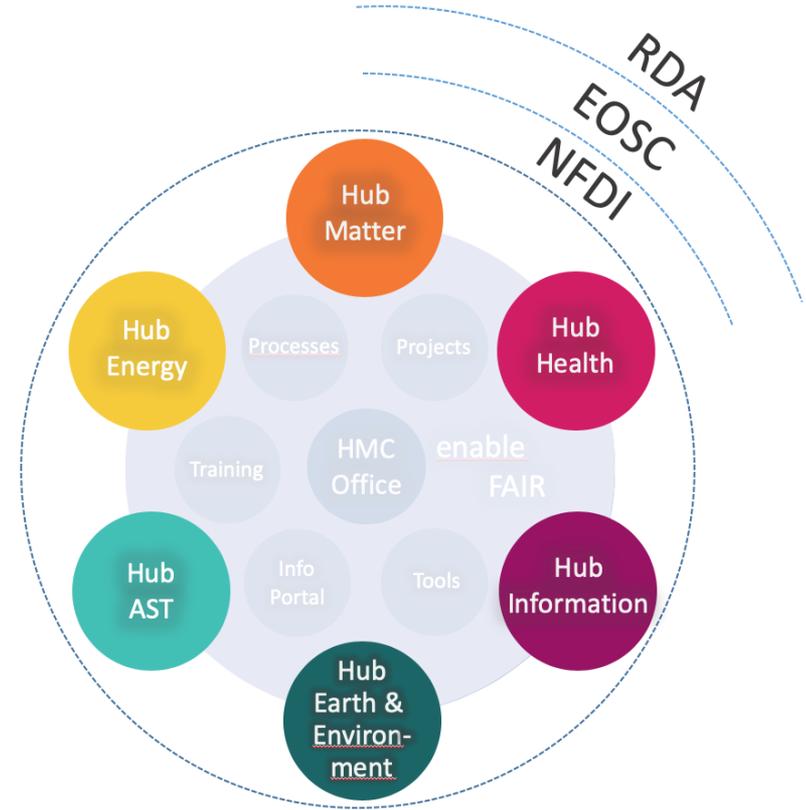
Customizing of services for hubs and researchers along the FAIR principles, incl. consulting, workflow and process management.



Domain specific Metadata Hubs

The research areas contribute competences, ideas and requirements of their domains via six Metadata Hubs.

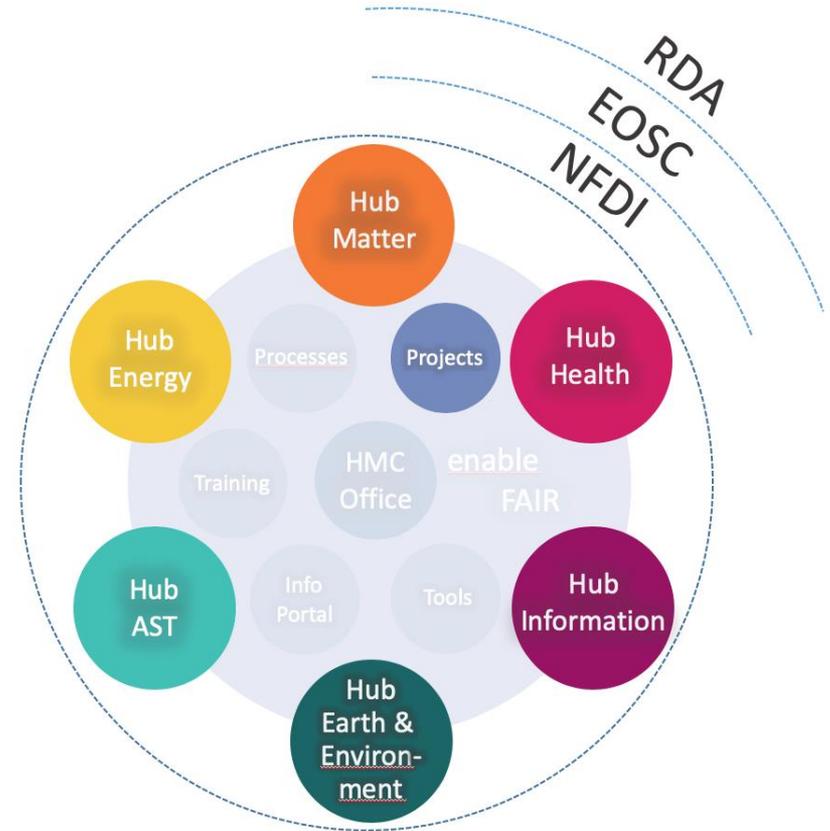
Energy	KIT
Earth & Environment	GEOMAR
Health	DKFZ
Aeronautics, Space and Transport	DLR
Matter	HZB
Key Technologies / Information	FZJ



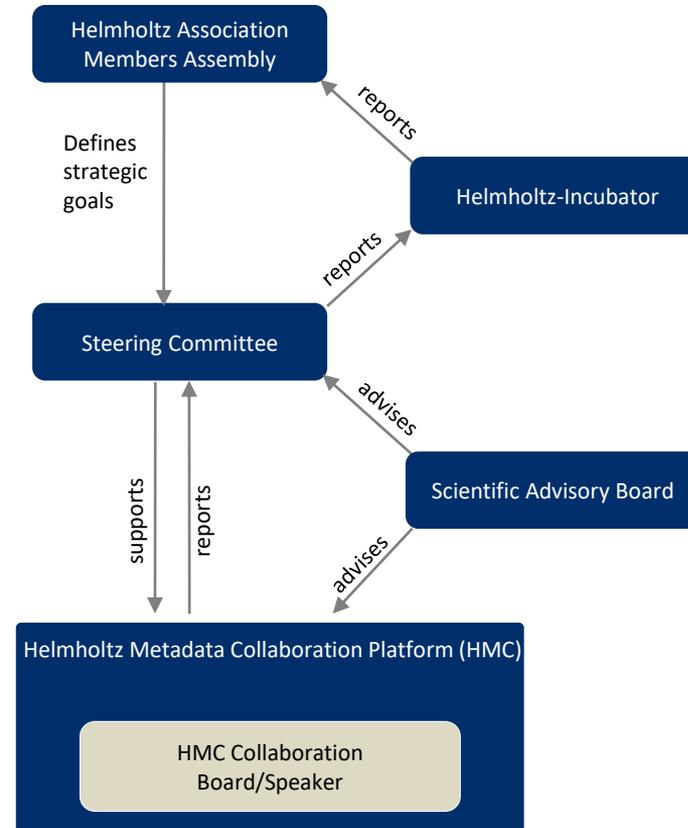
- **WP 3.1: Coordination & management**
- **WP 3.2: Knowledge & transfer**
 - information base about ontologies, vocabularities and standards
 - landscape of research data expertise
 - domain specific methods
 - consulting and training
- **WP 3.3: Components & processes**
 - Processes, tools and services for meta data repositories
 - Ingest: automation of metadata
 - Tools for metadata standards
 - Access
 - Creation of FAIRness
 - Provenance

HMC Projects (IVF)

- **practical solutions** in the context of metadata, e.g. specific tools or workflows
- Solutions shall be of **operational value** for the HMC service portfolio
- Two calls per year
- Funding periods from 12-24 months
- 1.2 mio funding via *Impuls- und Vernetzungsfonds (IVF)* plus matching by participating centres
- HMC office organizes call criteria and funding, based on advice of steering group and SAB



- **Steering Board** supervises the HMC platform and approves the reporting
- **Scientific Advisory Board** governs the project calls and review process
- **Collaboration Board** operational, implements the HMC working plan



Research Field	Members	Centre
Energy	Veit Hagenmeyer	KIT
Earth & Environment	Frank Oliver Glöckner	AWI
Health	Wolfgang zu Castell	HMGU
Aeronautics, Space, and Transport	Carina Haupt	DLR
Matter	Michael Bussmann	HZDR
Key Technologies / Information	Gisbert Breitbach	HZG
Open Science Office	Roland Bertelmann	GFZ

one member per research field appointed by HGF + Open Science office

Research Field	Members	Centre
Energy	Rainer Stotzka, Wolfgang Süß	KIT
Earth & Environment	Sören Lorenz, Emanuel Söding	GEOMAR
Health	Frank Ückert, Barbara Port	DKFZ
Aeronautics, Space and Transport	Christian Langenbach	DLR
Matter	Ants Finke, Heike Görzig	HZB
Key Technologies / Information	Stefan Sandfeld, Volker Hofmann	FZJ
<i>HGF Liaison</i>	<i>Sünje Dallmeier-Thiessen</i>	<i>HGF Office</i>

HMC interface to incubator platforms

HiDA | HELMHOLTZ
Information & Data Science Academy

HAiCU | HELMHOLTZ
ARTIFICIAL INTELLIGENCE
COOPERATION UNIT

provides academy
Network, personnel

offers courses and
teaching formats on
metadata knowledge

<HMC> | HELMHOLTZ
METADATA
COLLABORATION

offers metadata
standards and tools

both provide
scientific use cases
and good practice

offers
metadata tools
& services

provides backbone
cloud platform
& software hub

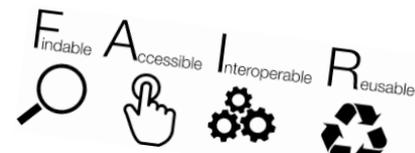
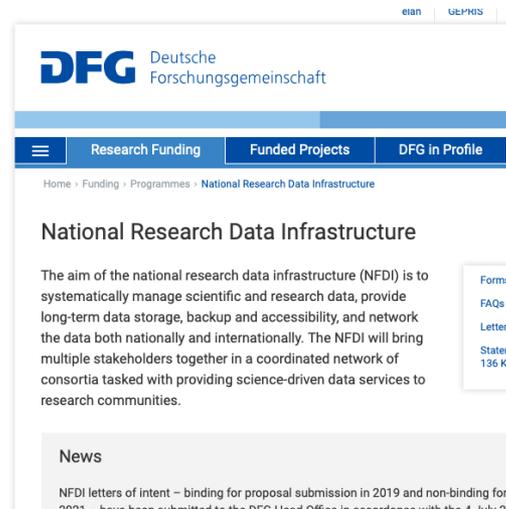
HIFIS | HELMHOLTZ
FEDERATED
IT SERVICES

HMC Office

- will link to the NFDI directorate, EOSC and FAIR offices and RDA Germany and Europe

Each metadata hub

- will link to community-specific consortia, EOSC pilots, FAIR-pilots and RDA groups

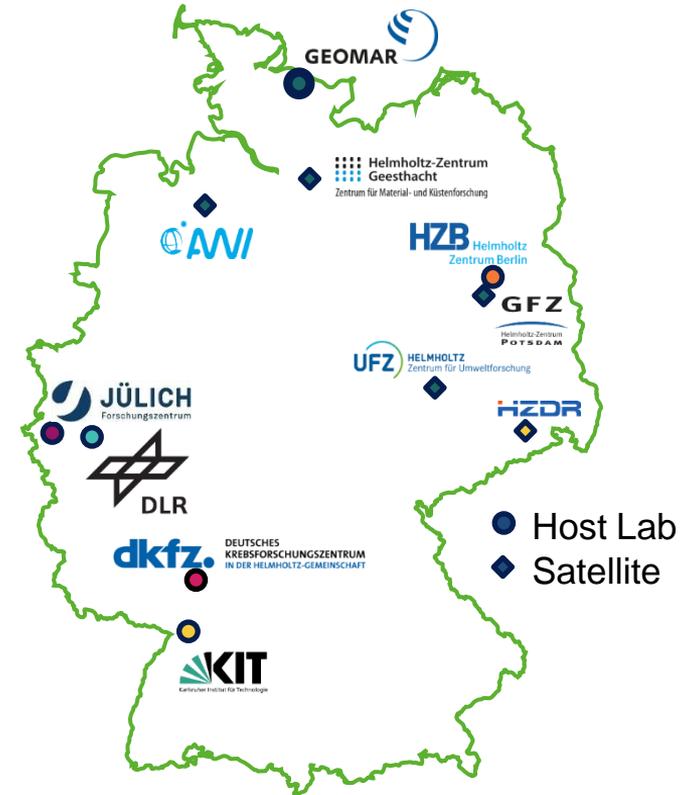


Challenges

1. **Strong community involvement** to speed up implementation
2. **Rapid development** of example tools and data methods as proof-of-concept
3. Implementing a **transparent provenance track of data products** demonstrating their Helmholtz Origin

Success Factors

1. **Long term funding** leads to sustainable (Helmholtz supported) infrastructures
2. **Decoupling the structural development** of tools and data by defining FDOs
3. **Creation of incentives** by connecting tools with data



Thank you for your attention!

Contact: wolfgang.suess@kit.edu