

# Helmholtz Metadata Collaboration (HMC) – Integrating Large-Infrastructure Data

Emanuel Söding<sup>1</sup> // Martin Weinelt<sup>1</sup> // Andrea Pörsch<sup>2</sup> // Helen Kollai<sup>3</sup> // Pier Luigi Buttigieg<sup>4</sup>

## About HMC

The Helmholtz Metadata Collaboration is tasked to connect and integrate the Helmholtz Association's data products into the ongoing global activities, like the EOSC and other coordinated programs. HMC is working on a concept to address interoperability and reusability to make Helmholtz' data FAIR.

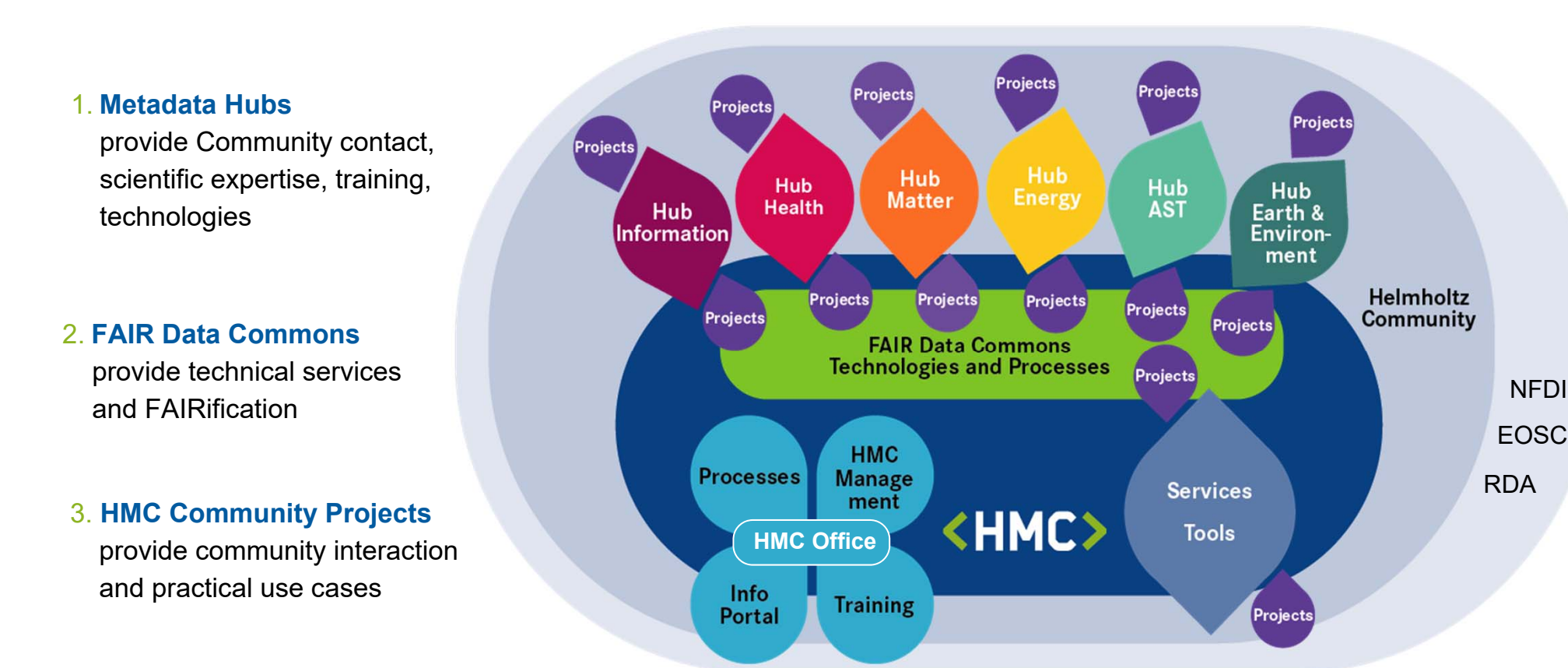
To achieve this HMC develops a Helmholtz wide knowledge framework, upgrades our data infrastructures with consistent semantic concepts and implements technical concepts like FAIR Digital Objects. We thus strive to establish practical use cases **turning "FAIR in to reality"** at Helmholtz.

## HMC Facts

- **Budget:** ~5M€/yr
- **Personnel:** 42 permanent FTEs, ~15 project FTEs
- **Scope:** Six research fields (Energy, Earth and Environment, Health, Information, Matter, and Aeronautics, Space, and Transport.)
- **Implementation:** 10 involved Centres
- **Funding:** 5 years set-up phase (2019-2024), permanent funding subject to review

## HMC Organization

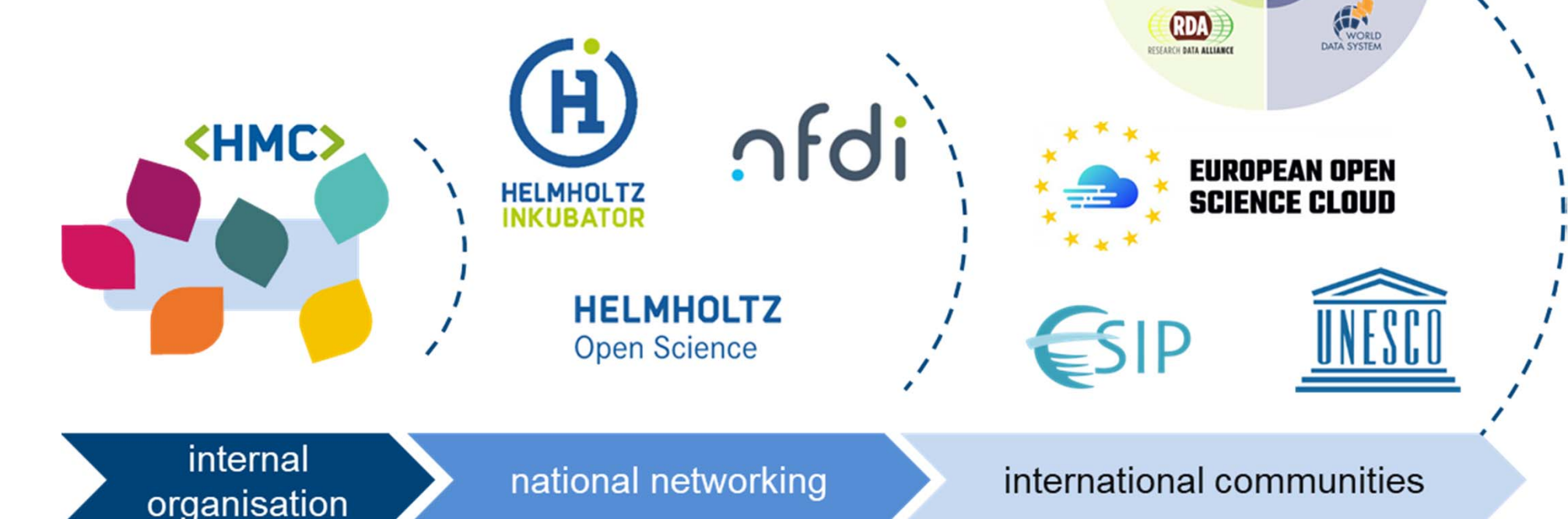
**Bridge tech and practice**, i.e. science community demands and current RDM implementation scenarios.



## HMC Integration

Define **common practices** with partners.

- **Networking with partners** on all levels through a gradual development of an **open network perspective** within Helmholtz and beyond
- **Connectability to internal developments** on all levels by implementing common standards, recommendations, interfaces



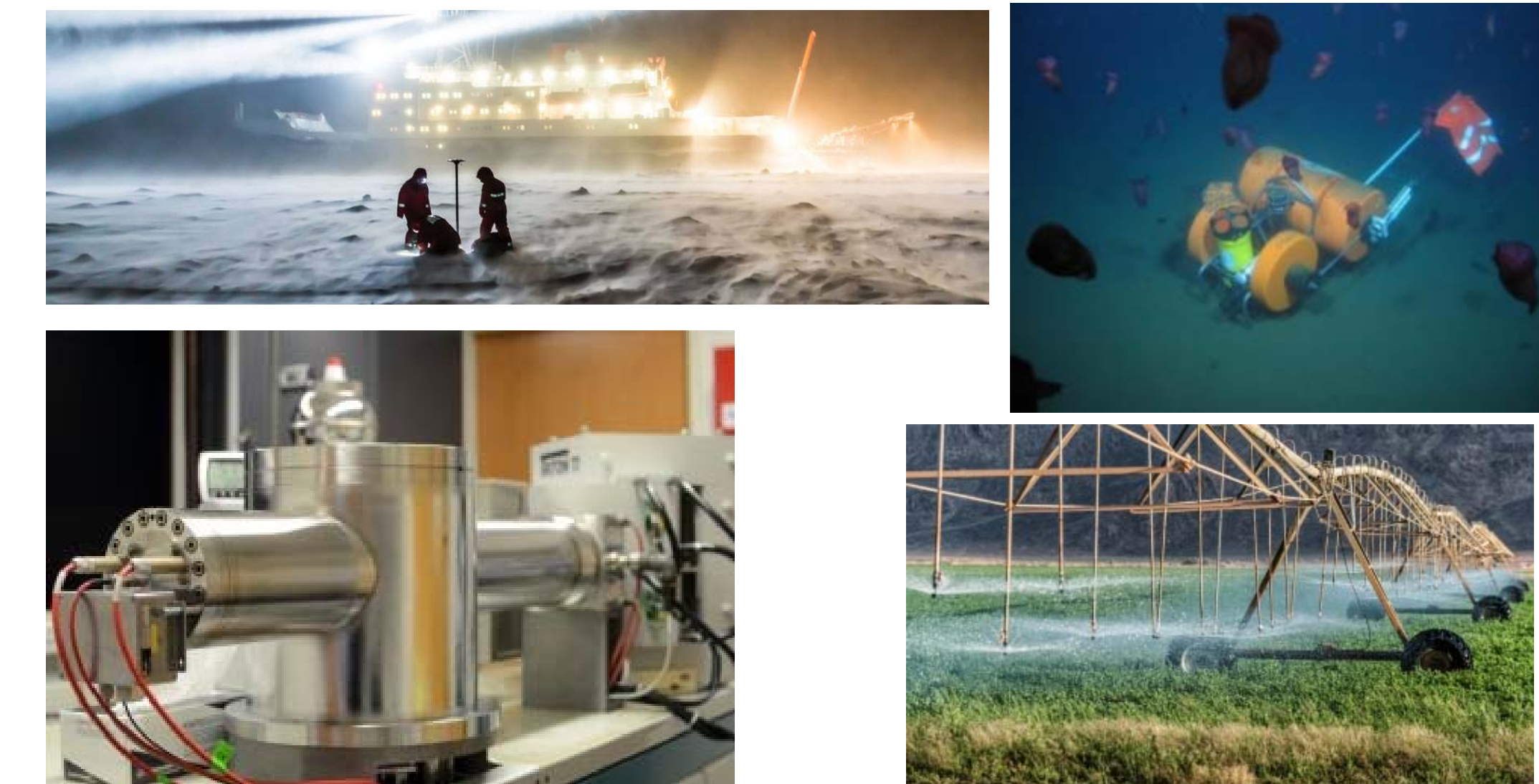
## Current Activities

- Setting up **communication structures**
- **Harmonizing activities** across research areas
- **Determine status** of Helmholtz structures
  - Mapping of „Metadata Landscape“
  - Evaluate infrastructures
  - Assess technical capabilities
- Define **vision**
- Decide on long-term **implementation concept**

## Main Challenges

- Strong **community involvement** to speed up implementation
- Rapid development of **example tools and data** methods as proof-of-concepts
- Implementing a transparent provenance track of data products demonstrating their Helmholtz origin to funders

## Hub Earth and Environment



### Scope of Data Products:

- Wide Range** of data sources and topics
- covers Earth-, Ocean- and Atmospheric Sciences, Biodiversity, Environmental Sciences
  - from expeditions and field campaigns, laboratories, mobile and stationary sensors, time series, modelling

## ESIP and the HMC

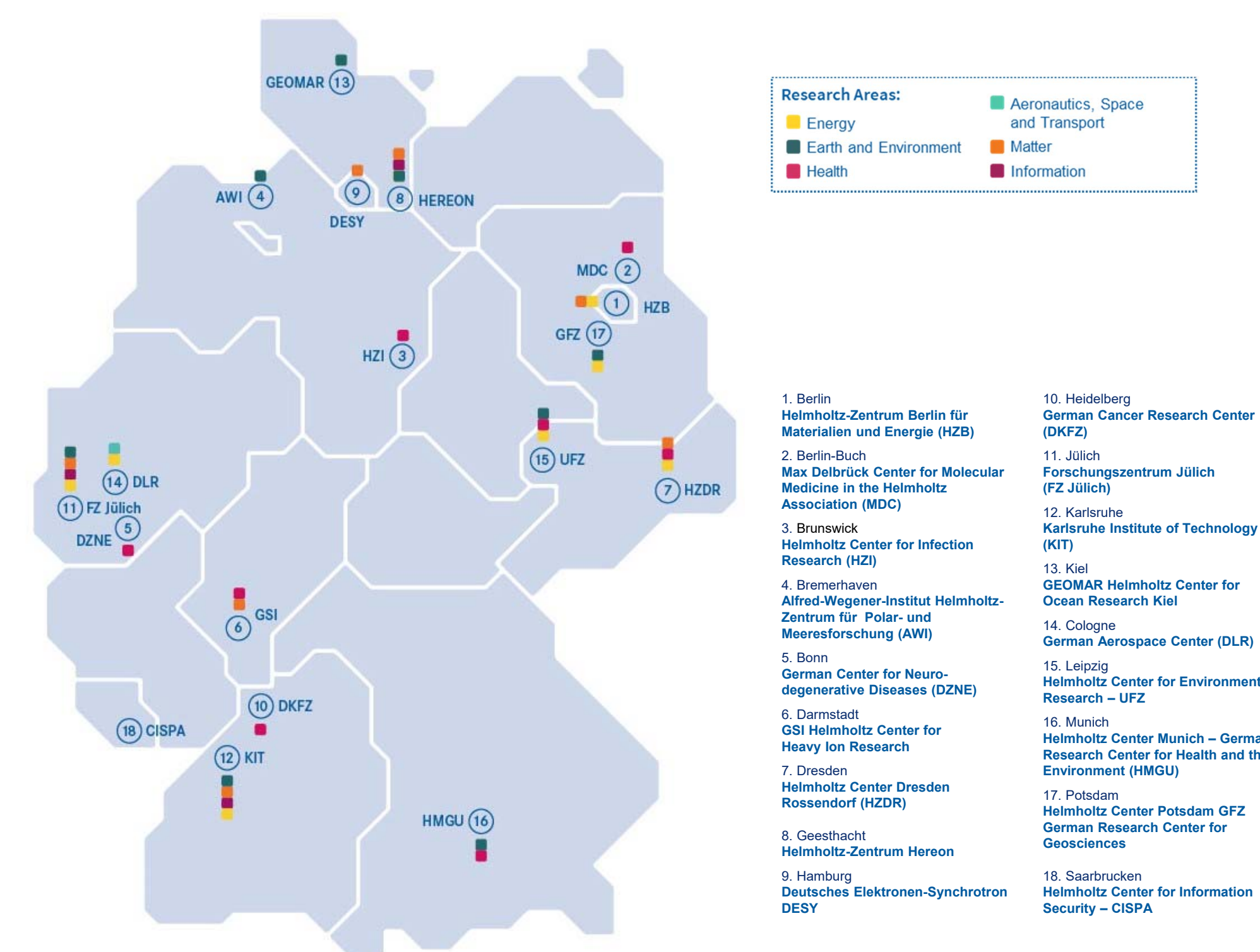
The HMC is happy to support and expand the existing links its personnel have developed with ESIP. Clusters and committees that we are engaged with include the:

- SeREEn Federation
- Semantic Technology Committee
- Semantic Harmonization Cluster
- Science on Schema.org
- Soil Ontology Cluster

**We hope to align and synchronise the Helmholtz digital ecosystem to the broader Earth and Environment community** by sharing our capacities, experiences, and challenges with ESIP, in service of better data science for the planet!

**HMC Contact:**  
Emanuel Soeding, [esoeding@geomar.de](mailto:esoeding@geomar.de)

## Helmholtz Association and Research Centres



The **Helmholtz Association of German Research Centers** is Germany's framework for federal, large-scale infrastructures, like polar programs, traffic and aerospace, energy research, particle accelerators and others.

Helmholtz contributes to solving major challenges facing society, science, and the economy through top-level scientific achievements in six Research Fields: **Energy, Earth and Environment, Health, Information, Matter, and Aeronautics, Space, and Transport.**

## Infrastructures and Use-Cases



See more at: <https://helmholtz-metadaten.de/de/erde-und-umwelt/use-cases>