A road to data liberation in Helmholtz

Christine Lemster* // Constanze Curdt // Sören Lorenz
Helmholtz Metadata Collaboration, HMC Office, GEOMAR Helmholtz Centre for Ocean Research Kiel
* Contact: clemster@geomar.de

About the Helmholtz Metadata Collaboration

The Helmholtz Metadata Collaboration (HMC) is one of five platforms that were set-up to push information and data science in Helmholtz and is funded by the Helmholtz Information & Data Science Incubator. HMC focuses on the mayor challenges derived from making **Helmholtz data treasure visible and FAIR** – especially in the context of interoperability and reusability.

HMC's internal structure is designed to serve the diverse Helmholtz centers by creating a **distributed structure** that allows units to focus on their respective Helmholtz Research Fields (Metadata Hubs) while a central service unit (FAIR Data Commons) implements technical solutions that are the basis for a functioning infrastructure.

Our goal is to establish metadata-centered solutions that improve the FAIRness of the research data. Those solutions are embedded in **national, European and global developments** – as well in the technical as in the research community-specific aspects - and build on existing concepts, tools and processes. This way, their value for larger science community can be assured.

HMC's success is relying on three interdependent aspects that shape our work:

- **Acceptance** within the respective scientific communities within Helmholtz and beyond
- Needs-oriented **technical solutions** that work on a global scale
- Implementation of widely accepted FAIR processes

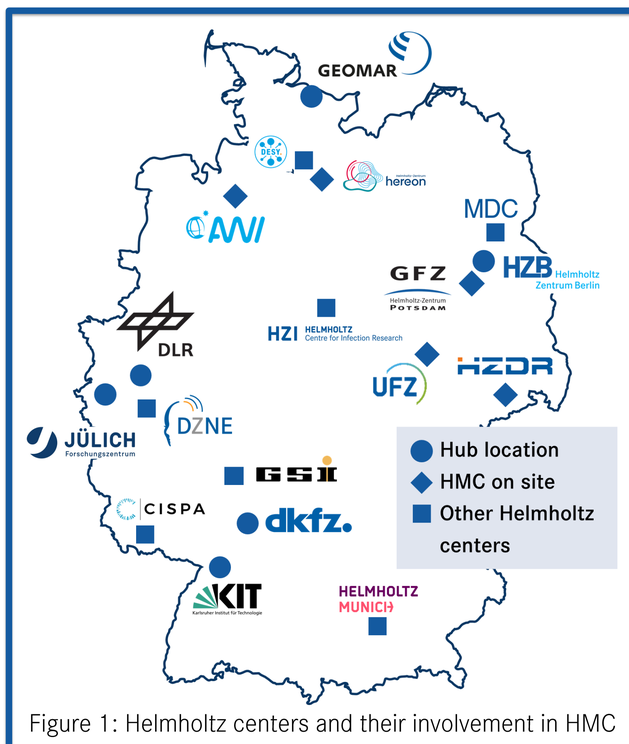

Figure 1: Helmholtz centers and their involvement in HMC

Helmholtz Association

The **Helmholtz Association of German Research Centers** is Germany's framework for federal, large-scale research facilities who contributes to solving major challenges facing society, science, and the economy through top-level scientific achievements in its six Research Fields: **Energy, Earth and Environment, Health, Information, Matter, and Aeronautics, Space, and Transport (AST).**

To tackle the challenges of the data-driven research, a think tank - **The Helmholtz Information & Data Science Incubator** - was establishes that aims at synergizing Helmholtz expertise and creating a future-oriented research environment.

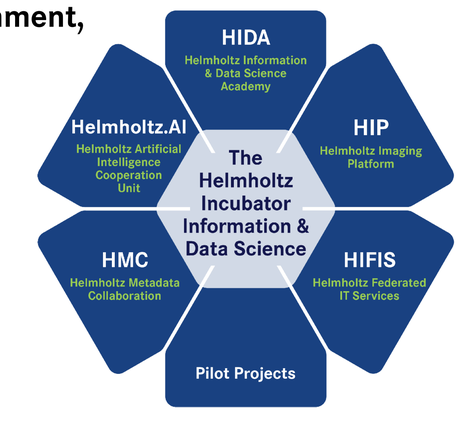

Figure 2: The Helmholtz Incubator platforms

Achievements on our way to a successful implementation

Community Building and Engagement

- Collect and disseminate **information about metadata**
  - Mapping of research data landscape in Helmholtz and beyond through surveys, interviews and data mining
  - Structured, accessible presentation results
  - Drafting of information materials (guidelines, factsheets etc.)
- Conducting **workshops and trainings** on general and domain specific metadata topics
- Provide added value for all people involved within the Helmholtz Centers
- Project-funding for practical developments within the communities: **15 community projects** were funded this far since 2020

Technical Infrastructure and Solutions

- Technical ecosystem for **FAIR Digital Object,** e.g. FDO test bed
- Problem-solving oriented **metadata tools** and catalogues, e.g. Metador or DirSchema
- Automated **FAIR assessment** of data sets based on F-UJI (Tool developed within the FAIRsFAIR project)
- **Customizing technology** for Helmholtz research fields
  - Pilot projects, e.g., FDOs for photovoltaic system
  - FAIR assessments of workflows

FAIR Policies and Processes

- Customizing given recommendations
- From recommendation to practice

HMC Organization

Metadata Hubs

- Scientific expertise
- Community contact
- Technologies
- Training

FAIR Data Commons

- Technical services
- FAIRification

HMC Office

- Management
- Controlling

HMC Community Projects

- Community interaction
- Practical use cases

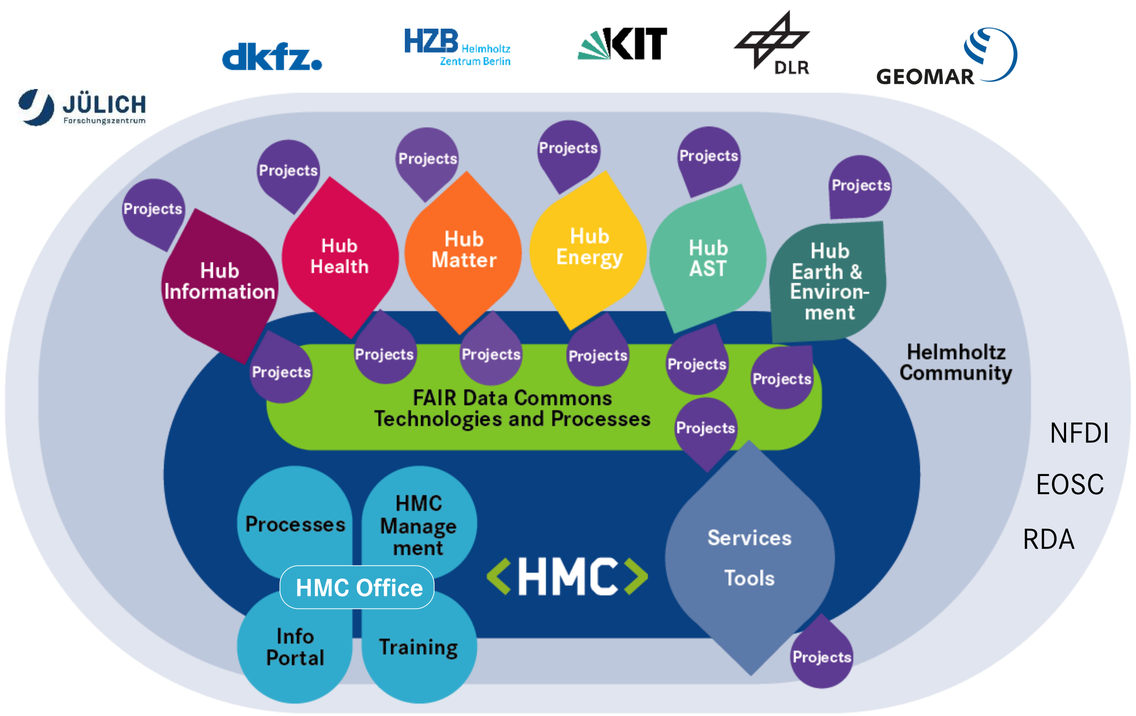

Figure 3: HMC's internal structure

Further information:

- FDO ecosystem: https://github.com/kit-data-manager/
- Metador (Annotating tool for external data sets): https://github.com/Materials-Data-Science-and-Informatics/metador
- DirSchema (Spec and Validator for metadata) https://github.com/Materials-Data-Science-and-Informatics/dirschema
- The Helmholtz Information & Data Science Incubator : https://www.helmholtz.de/en/research/challenges/information-data-science/

You want to know more?

Visit us at www.helmholtz-metadaten.de
Write us at hmc-info@geomar.de
Follow us on Twitter: @helmholtz_hmc