Introduction to the
Helmholtz Metadata Collaboration (HMC)

Constanze Curdt
HMC Office, GEOMAR

Information Event | HMC Project Call 2022 | 16 May 2022
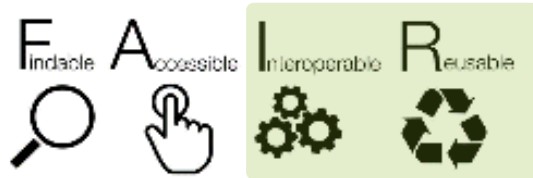
www.helmholtz-metadata.de

HMC Mission and Challenges

- Make Helmholtz data treasures visible and **FAIR**

- Support researchers in (automatically) describing their data by means of a suitable, standard-compliant description with metadata.

- Allow researchers to reuse Helmholtz research data for advanced methods of data processing and analysis.

- Provide comprehensive & sustainable services, consulting, information & tools for efficient metadata handling as a distributed & shared facility.

- Jointly develop, share & consolidate community-expertise for metadata of the six Helmholtz research areas.

- **Turning FAIR into reality* on all levels to enable data reuse**

- **This makes HMC a research infrastructure platform!**

* From: Turning FAIR into Reality, Final Report and Action Plan from the European Commission Expert Group on FAIR Data, doi: 10.2777/1524

HELMHOLTZ

HMC Locations

HMC Host Centres

FZJ – Information & FAIR Data
Forschungszentrum Jülich

DKFZ – Health
German Cancer Research Center

HZB – Matter
Helmholtz-Zentrum Berlin für Materialien und Energie

KIT – Energy & FAIR Data
Karlsruhe Institute of Technology

DLR – Aeronautics, Space and Transport
German Aerospace Center

GEOMAR – Earth & Environment and HMC-Office
GEOMAR Helmholtz Center for Ocean Research Kiel

HMC Satellites

HZDR – Energy
Helmholtz Center Dresden Rossendorf

AWI – Earth & Environment
Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung

GFZ – Earth & Environment
Helmholtz Center Potsdam German Research Center for Geosciences

UFZ – Earth & Environment
Helmholtz Center for Environmental Research

HMC Structure

1. Metadata Hubs
Community expertise, training, technologies

2. FAIR Data Commons
Technical services and FAIRification

3. HMC Projects (1.2M€/y)
Community and use cases

4. HMC Office
Management and controlling

● Community Interfaces

Communication / Exchange about Cross-cutting topics (CCT)

- Hubs a starting from different levels and with different velocities of Research Data Management development.

- Groups working on cross-cutting topics help to organize the work across the research field specific hubs.
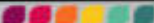
Cross-cutting topics

Mapping
Training
FAIR Concepts
From Dev2Deploy
Community Services
Internal & External Communication
HMC Glossary / Semantics

Current activities

➢ Surveys

➢ HMC wide programme

➢ FAIR recommendations

➢ Evaluation of community tools

➢ Community helpdesk and market place

➢ Enable community dialogue

➢ Roadmap towards a Helmholtz knowledge graph

	Objective	Year 1	Year 2		Year 3	Year 4	Year 5
Central components HMC Office	**HMC Office**						
	Management, communication and PR and project office	Develop and set up communication structures			Ongoing coordination, communication and information		
	Building, providing and disseminating knowledge, training and consulting	Mapping expertise; developing training concepts			Regular services and further development of formats		
	Coordinating FAIR metadata and technology across hubs	Developing requirements catalog and guidelines with hubs and bringing them to application			Continuous provision and further development of established solutions		
Decentral components HMC Office	**FAIR Data Commons**						
	Implementing FAIR Data Object	Implementation of Fair Data Object Model; services for using catalogs/registries			Continued developments in coop. with EOSC, RDA, NFDI; integration of results from projects		
	Technical implementation of FAIR principles	Survey, test and further develop tools			Connect, disseminate and evolve together with hubs		
Domain-specific services	**Metadata Hubs**						
	Metadata expertise and its community in each domain	Integrate expertise and establish exchange formats			Ongoing coordination and management of activities and communication channels		
	Mapping, building up and disseminating methodological knowledge; training and consulting	Establishment of domain-specific training and interaction formats			Ongoing development of formats offered online and F2F		
	Processes, tools, and services for indexing research data collections	Collection, requirements analysis and exemplary application			Usability, adaptation and further development, FAIRness of data can be recorded continuously		
Projects	**HMC Projects**						
	Call for projects, consulting for applicants, monitoring	Develop calls, establish processes			Monitoring of further calls		
	Support and networking of the projects, integration in HMC	Support initial projects, create networking opportunities			Continuous monitoring of the projects, integration of the results in HMC		

https://www.helmholtz.de/fileadmin/user_upload/HIDA/2_HMC-Proposal-April2019.pdf

HMC is paving the ground of the road to FAIR for Helmholtz

- **Community Building**
  - collect and disseminate **information about metadata**
  - design and conduct **workshops and trainings**
  - involve community through **annual HMC project call**
  - provide added value for all people involved within the Helmholtz Centres

- **FAIR Policies and Processes**
  - customizing given **recommendations**
  - from recommendation to practice

- **Technical Developments**
  - **Metadata tools** and catalogues
  - **FAIR Digital Object tools** and **cook books**
 - Customizing technology for Helmholtz research fields

- **National & international Networking**
  - seamlessly **bridging of activities and initiatives** across domains and stakeholders

Community Building
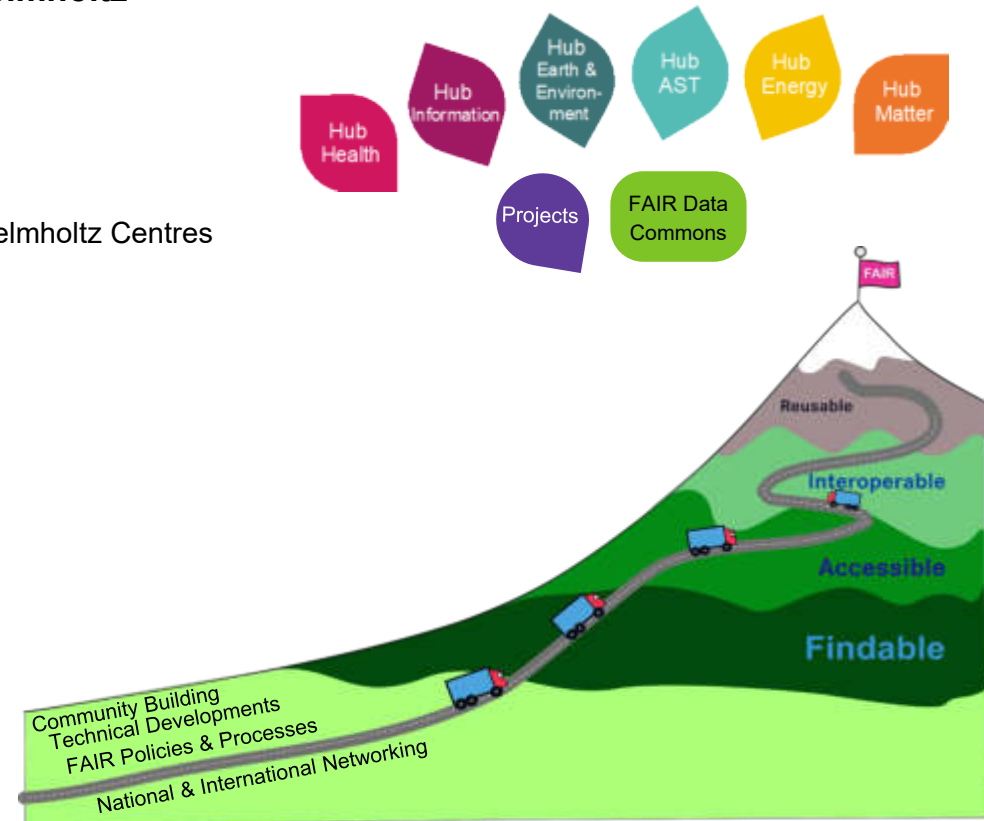
- **Gathering Intelligence: „Metadata Landscape" Mapping & Survey**
  - HGF-wide surveys for communities and a pilot survey on infrastructures in Hub E&E
  - mapping of recommendations, repositories, vocabularies, metadata standards ,…
  - (meta)analysis of communities, infrastructures & services

- **Community Engagement: Workshops, Trainings, Materials**
  - Several hub-specific trainings & workshops have been conducted on metadata and FAIR
  - Seminar Series HMC FAIR Friday
  - Guidance material, documentation (e.g. Fact Sheets)

- **Community Involvement: HMC Projects**
  - 15 HMC projects funded from the inaugural HMC proposal call 2020 (9 projects, 1,8M€ INF-Budget) and call 2021 (6 projects, 1,2M€ INF-Budget), distributed amongst 13 Helmholtz Centres

HMC COMMUNITY SURVEY 2021
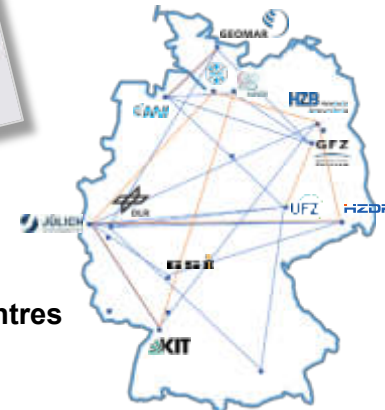
Interconnections and distribution amongst Centres
- Projects Call 2020
- Projects Call 2021
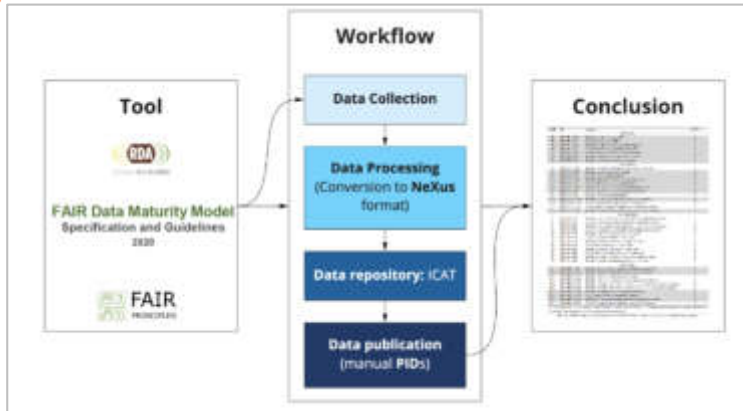
Further info on events

Further info on projects

HELMHOLTZ

FAIR Policies and Processes ... examples from the hubs
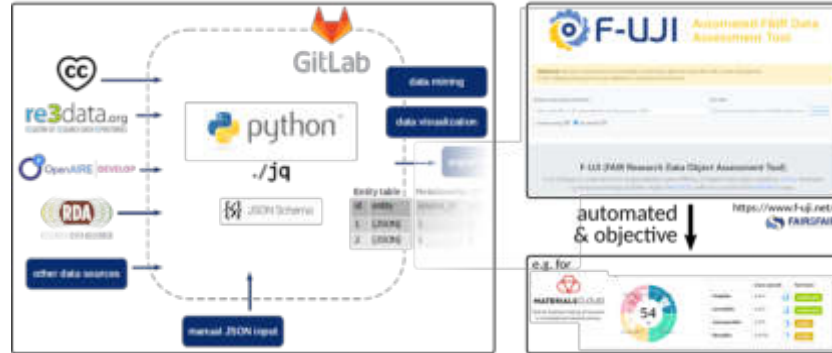
Objective and automated FAIR assessment of data publications & repositories

Manual FAIR assessment of a prototypical beamline instrument

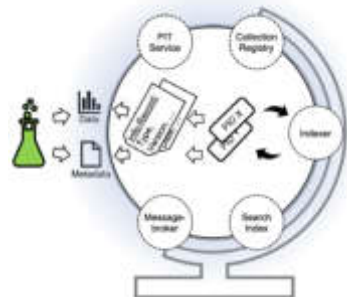**Co-Design Implementation Process** for FAIR infrastructures
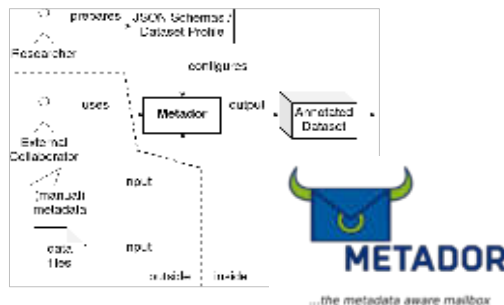
HELMHOLTZ

Technical Developments

FAIR Data Commons

FAIR Digital Object (FDO) concept and testbed

FAIR Data Commons

End-user oriented tool development (Metador)

METADOR

...the metadata aware mailbox

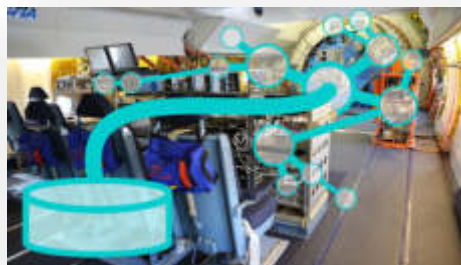| Service / Tool	Short Description
	FAIR Data Commons
Collection Registry	Creation and management of collections of digital resources. It allows to create virtual collections of heterogeneous research data located in distributed data resources.
MetaStore	The metadata management system consisting of metadata repository and metadata schema registry allows to master huge collections of heterogeneous metadata documents.
Metadata Standards Catalog	Metadata catalog for registering and discovering existing metadata standards, see text below.
Web Annotation Protocol Server	Implementation of the W3C Web Annotation Data Protocol standard for preserving and managing annotations of research data.
FAIR DO Testbed	The FAIR DO Testbed facilitates the implementation of FAIR DOs using various services offered by WP 2.
Generic Metadata Editor	The Generic Metadata Editor allows to auto-generate modern user interfaces for creating and updating metadata documents.
FAIR DO Cookbook	The FAIR DO Cookbook aims to provide support and step-by-step instructions on how to approach different challenges while realizing FAIR DOs.
Metador	*Metadata enrichment and transmission assistance for digital objects in research* - is a structured submission interface for research data linked to metadata. Metadata requirement are to be supplied by completing a configurable form for each upload file. This tool is intended for deployment in research groups and designed for quick and easy integration into existing scientific workflows.
DirSchema	Directory Schema – is a specification and validation tool that allows enforcing structural and metadata requirements in local datasets. This tool is intended to be deployed to research groups and used during dataset preparation in order to harmonize and enrich datasets across groups. This will increase machine interpretability and reusability, e.g. ease (automatic) data analysis or metadata harvesting pipelines.

Hub Energy

Application of **FDO concept** for improving FAIRness of information in a photovoltaic system

Hub AST

Development of an **ontologically-aware** database for instruments and sources

HELMHOLTZ

- **Highly distributed structure** -> gradual development from a stable internal structure and perspective to an open network in HGF and beyond
- HMC partners in advancing FAIRness of research data
- **Networking** on various levels
- On **national** level involvement in subject-specific **NFDI**-consortia & cross-sections
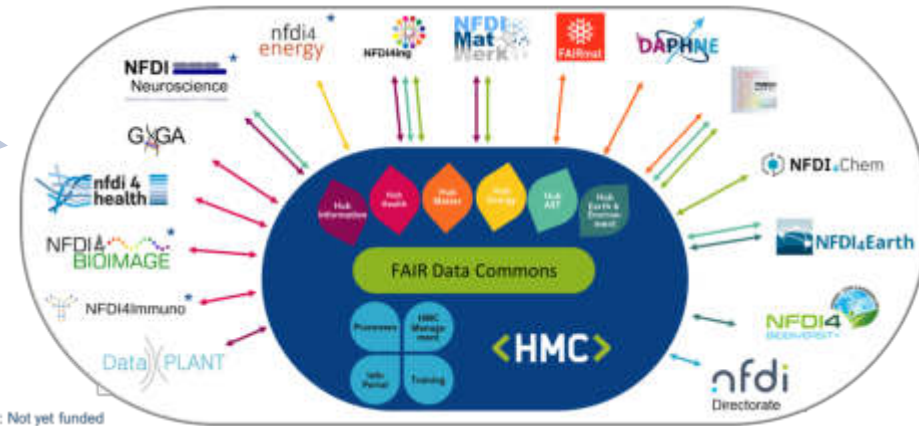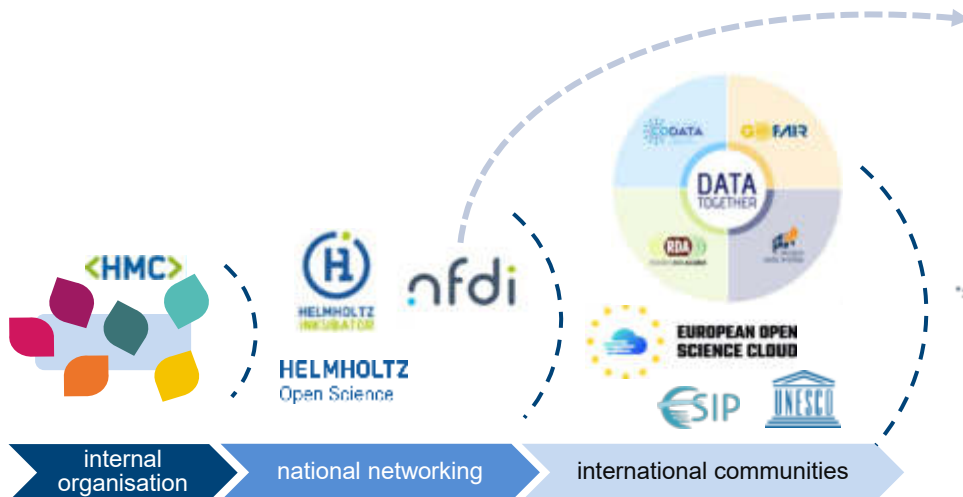
HMC Participation in NFDI Consortia

internal organisation | national networking | international communities

Role of projects in HMC ..

- are an important instrument to **connect HMC and community experts**

- can help to **identify community demands**

- work on **practical challenges** in the field of metadata

- should work in **close cooperation** with the HMC hubs

- **results** should be integrate into HMC and HGF Community in the long term

- will be integrated into HMC activities as **collaboration** partners

HMC offers ..

- **Integration of project results** into HMC

- **Knowledge exchange** with HMC Family

- Participation in HMC organized **trainings, seminar** etc.

- **Workshop** participation for exchange with HMC community, HMC projects etc

- Usage of **HMC Infrastructure** and Services

- … to become **part of the HMC Family!**

- **…**

HELMHOLTZ

Questions ?

Questions?

Questions?

Contact:
HMC Office
Dr. Constanze Curdt
ccurdt@geomar.de

Information Event | HMC Project Call 2022 | 16 May 2022

www.helmholtz-metadata.de