HELMHOLTZ

DataHub
Research Field Earth & Environment

The DataHub Community – fostering FAIR and sustainable research data management across our research field

Sabine Barthlott & Christof Lorenz for the DataHub community

DataHub Meeting Potsdam, January 2023

The DataHub of the Research Field Earth and Environment is a joint initiative of all centers of the Helmholtz Association participating in the research program '*Changing Earth*'.
Within the DataHub, *data management tools* are developed and made available, and *data products* are offered in *thematic viewers*. The continuous and long-term development process of these solutions is the DataHub's task.

DataHub Intitative

The **DataHub acts as the central competence center** of the research field with regard to the development and operation of sustainable data and information infrastructures.

Central goals of the DataHub:

- *connectivity* of distributed infrastructures,
- homogenization of *data descriptions* and *data flows*
- homogenization of *portal* and *viewer technologies*
- *joint data products* for research, politics and society.

DataHub: internal and external network

Internal:

- Community with more than 100 members of the whole research field
- Thematic working groups ("AK"s) for Viewer & Portal Technologies, Metadata & PIDs, Architecture

External:

The activities are further embedded in the context of building the **German National Research Data Infrastructure (NFDI)** which also serves as the German contribution to the **European Open Science Cloud (EOSC).**
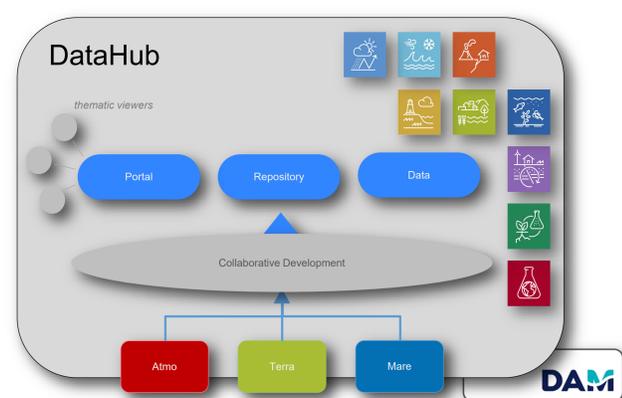

Technological Approach

- Integration of existing research data management tools
- Implementation of new tools and overarching technology

driven by following principles:

(1) **co-design approach** to maximize usability

(2) the architecture captures a **distributed system**, the components of which are operated by the partners

(3) central **Identity and Access Management** (IAM) provided by the Helmholtz Digital Services for Science

(4) **Modular** software for better re-usability

Essential components

- domain oriented **data repositories**
- **thematic viewers** focusing on specific stakeholders
- central **Earth Data Portal**
- **Spatial Data Infrastructures** operated by the partners
- integrated data flows „**from sensor to archive**",
- modules for **Quality Assurance** and **Quality Control**
- domain specific **data explorers** as a tool for researchers to browse data via visual exploration