HMC HELMHOLTZ Metadata Collaboration

AWI GEOMAR

HARMONise – Towards the sustainable management of metadata associated with marine molecular sequencing data

Christina Bienhold[1] // Till Bayer[2] // Lars Harms[1] // Stefan Neuhaus[1] // Roland Koppe[1] // Isabell Siebert[1]
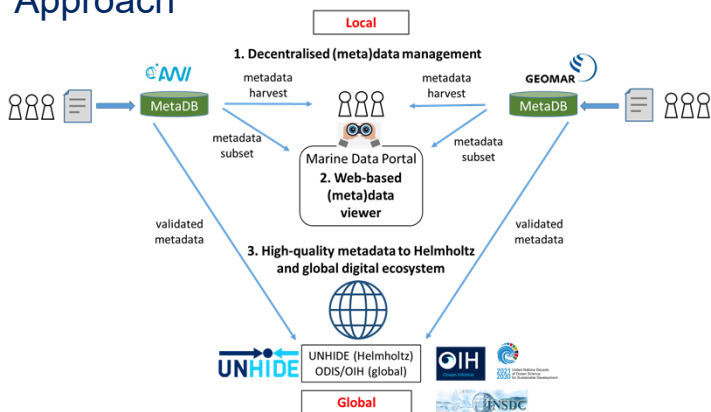
Contact: christina.bienhold@awi.de

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung[1] // [2] GEOMAR Helmholtz-Zentrum für Ozeanforschung

Motivation

Biomolecular data, e.g. DNA and RNA sequences, provides insights into the structure and function of marine communities in space and time. The associated metadata has great internal diversity and complexity, and to date biomolecular (meta)data management is not well integrated across Helmholtz Centres. Enhancing the interoperability of marine biomolecular (meta)data will support FAIR data exchange internally and externally, and promote efficient exploitation of the data.
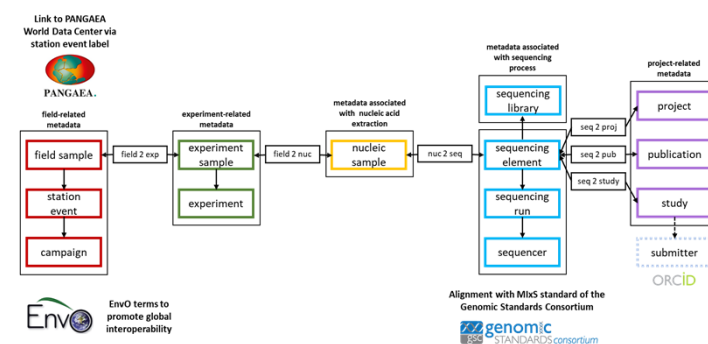
Aims

The project aims to develop sustainable local-to-global workflows that support:

1. **High-quality curation and management of marine biomolecular metadata**
2. **FAIR exchange and publication of (meta)data within HMC systems and globally**
3. **Harmonisation of biomolecular (meta)data with other Helmholtz digital holdings and services**

Approach

Marine Data Portal - https://marine-data.de; HMC unHIDE - unified Helmholtz Information and Data Exchange; ODIS - Ocean Data and Information System; OIH - Ocean Info Hub

Local (meta)data management
- Decentralised (meta)data management allows for local adaptations to promote fitness-for-purpose according to each Center's/group's research mission.
- A web-portal for metadata upload and harvest enables sustainable data stewardship, improves accessibility and supports delivery of high-quality metadata to national and global repositories.

Web-based metadata viewer
- Metadata subsets are harvested by the Marine Data Portal, increasing **findability** across research domains and promoting **reuse** of biomolecular research data.

Helmholtz and global digital ecosystem
- Alignment with community standards and relevant data exchange formats (e.g. UNESCO ODIS-Arch specifications) will support Helmholtz **connectivity** and global **interoperability**.

## Harmonisation and Interoperability

Database schema that aligns metadata with community standards such as the Minimum Information about any (x) sequence (MIxS), and with associated ontology content (ENVO – The Environment Ontology). It also harmonises metadata with existing Helmholtz repositories (e.g. PANGAEA).

HARMONise @Marine Data Portal

The HARMONise Web Portal hosted at AWI supports researchers in sustainably managing metadata associated with DNA and RNA sequences and in making Helmholtz biomolecular (meta)data FAIR (Findable, Accessible, Inter-operable, Reusable).

A metadata subset is harvested by the Marine Data Portal ➡
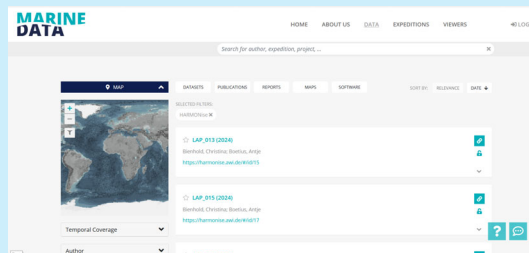
Links to HMC (Hub Earth & Environment)

HARMONise is aligned with the HMC goals. It:

i) **supports** the **development of digital cultures** in the groups/Centers,

ii) **strengthens biomolecular (meta)data exchange**,

iii) **ensures sustainability** through high-quality documentation, and

iv) **enhances Helmholtz-to-global interoperability**

HELMHOLTZ