

Development of recommendations for the implementation of semantic artifacts in HMC Earth and Environment

Dorothee Kottmeier¹ // Emanuel Söding² // Andrea Pörsch³ // Stanislav Malinovschii² // Yousef Razeghi⁴ // Sören Lorenz² 0000-0002-4467-642X (D) 0000-0002-0007-630X D000-0002-4263-4234 (D) 0009-0002-9792-6768 i> 0000-0001-8577-6614 0000-0003-4502-6223

Introduction

The HMC Earth and Environment Hub is dedicated to establishing a framework for the semantic interoperability between the involved research data infrastructures in the field. Standardizing and semantically annotating metadata and harmonizing existing semantic resources are crucial for bridging the gap across diverse and complex data sets. Our goal is to work with the community to develop recommendations for implementing semantic concepts that meet the community's needs.

Establishment of the "Metadata – Semantics" working group

- Regular meetings with interested members from the Helmholtz community.
 - Work package 1: Instruments & methods
 - Work package 2: Observable properties
- Detailed analysis of challenges and common development of solutions and recommendations
- Communication through E-mail and the HMC community portal.
- Documentation of results in the HMC E&E Wiki.

Home / Working groups / AK MD Semantics

AK Metadaten - Semantic Concepts & Implementation

WG AK MD Semantics

Chats & Discussions Documen

Clearly representing, digitising, and coordinating knowledge is key to implementing distributed dig coordinated semantic resources will allow human and machine agents to understand and act on He ready efficiency. This group will define and coordinate the implementation of semantic concepts an digital Infrastructures. It will evaluate resources deployed by infrastructures for quality, fitness for pu

The live meeting protocols are collaborately written in a DESY hedgedoc: https://notes.desy.de/hoNMZ5PHRkaMmX5YiV4ozg?view

Fig. 2: Screenshots of HMC Community portal and the HMC Community Wiki

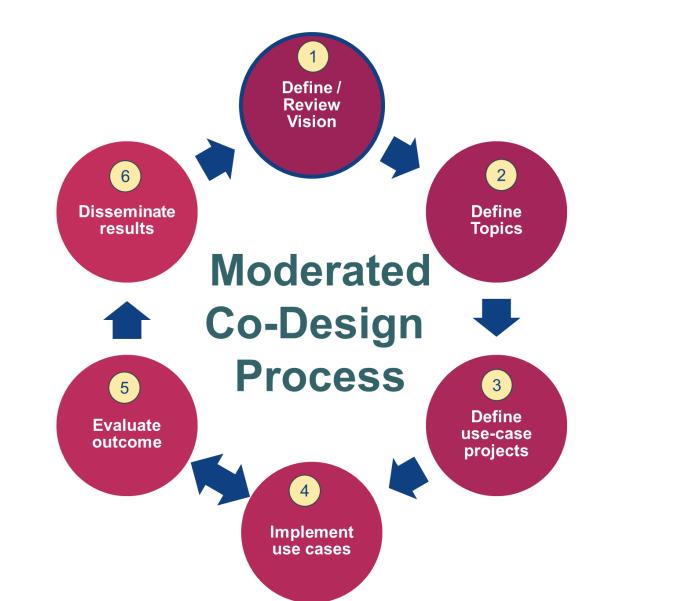


Fig. 1: Working procedure of the working group





DRAFT INTRODUCTION

Data Management Recommendations and Pra

Many people and stakeholder groups produce data rele In order to create a harmonized interoperable and ultin to be aligned across our data systems and implemente

In this Wiki, HMC and its partners will formulate recor concrete suggestions for the implementation and supp and the another the stand for any detail and interface of .

Identified challenges

• 0. Basic Recommendations and

M0 Use PID Metadata

M1.0 use ORCID

M1.1 Centers

encourage

M1 2 Employees

B0 Which data to we want to

Trace: • start

Introduction

HMC Vision and Mission

Recommendation Tree

Assumptions

harmonize

- Despite the wide range of controlled vocabularies (CVs) in use, there is no consensus on the most suitable options. Many CVs are inadequate to meet all requirements.

- More support is needed to evaluate existing semantic resources in terms of their "FAIRness," sustainability, quality, relevance to stakeholders, utility, and user-friendliness etc.

 Metadata typically consists of diverse components, necessitating a detailed examination of which elements should be included in data descriptions and annotations (cf. I-ADOPT framework).

- Metadata classification is often ambiguous and prone to errors.

– The integration of data across the existing data infrastructures in the field is hindered by insufficient transparency in the use of CVs. User guidance such as in interfaces or metadata naming (e.g.,

predefined syntax)—is still partially lacking.

Conclusions

- The "Metadata Semantics" working group unites Helmholtz Earth and Environment community members to tackle semantic interoperability challenges.
- Prioritizing metadata standardization and fair controlled vocabularies is key to enhancing semantic linking
- Ongoing collaboration on a common vocabulary and clear metadata guidelines is vital for better data integration

Recommended solutions

- vocabulary.

¹ Helmholtz Centre for Polar and Marine Research - AWI ² GEOMAR Helmholtz Centre for Ocean Research Kiel ³ Helmholtz Centre Potsdam GFZ ⁴ Helmholtz Centre for Environmental Research – UFZ

1. A list of recommended vocabularies and a guideline that provides advice on the evaluation of existing semantic resources should be created.

2. Metadata should be decomposed into components using a consistent approach, with each component mapped individually. A common syntax should guide naming conventions, and fair controlled vocabularies should be employed for semantic linking whenever possible.

3. If multiple external vocabularies can be used for mapping metadata, it should be clearly defined which vocabulary should be used with priority. Even if specific needs aren't fully met, the focus should be on establishing a common framework.

4. Internal vocabularies should be published and contextualized through mapping to external fair vocabularies. If needed, the research community should collaborate to develop a common



HELMHOLTZ