HMC Conference
> Metadata in Action, 28.–30.04.2026, Heidelberg

‹HMC› HELMHOLTZ Metadata Collaboration

Beitrag ID: **78**   Typ: **Talk**

Toward FAIR and Reproducible Data Quality Control: A Use Case–Driven Data-Quality Processing Metadata Schema for Time Series Data

High-quality environmental time series data require transparent, reproducible, and well-documented quality control (QC) workflows that integrate automated procedures and expert judgment. While many QC frameworks offer algorithmic methods, the processing information explaining how data quality decisions are made —including parameterization, flag semantics, and manual interventions —is often not formalized enough to be easily reused, reproduced, or exchanged across infrastructures.

In this talk, we present a metadata schema for time series data that enables FAIR and reproducible data quality processing. The schema is designed to describe QC methods, execution contexts, and resulting quality flags in a machine-actionable and interoperable manner. It employs the OGC SensorThings API data model enhanced by the STAMPLATE schema and the concepts established in the SaQC framework. The schema follows the linked-data approach and aligns with standards such as the W3C Data Quality Vocabulary.

The design of the proposed schema is motivated by concrete use cases for QC of time series data from the TERENO and ACTRIS observation networks. These use cases include detailed analyses of existing automated and manual QC workflows. By comparing and abstracting these practices, we derive common requirements and design patterns for representing QC processing information in a FAIR and reproducible manner. The resulting schema can be used straightforwardly with SensorThings API services and mapped into NetCDF files that align with the Helmholtz Metadata Guidelines for NetCDF. It can also be used with RO-Crates, embedding files in CSV format, for example.

Our metadata schema lays the foundation for a community-driven, FAIR, and reproducible quality control solution. Our goal is to integrate the requirements of other communities and develop a web application that allows users to visually inspect and flag time series data in a manner consistently with our schema.

ONLY WORKSHOPS - Proposed interaction format

Alternative Track

1. Empowering Research Communities: Turning Metadata into Action

ONLY WORKSHOPS - Tentative audience

ONLY WORKSHOPS - Maximum number of participants

ONLY WORKSHOPS - Special technical requirements

Autor: Dr. LOUP, Ulrich (Forschungszentrum Jülich GmbH)

Co-Autoren: Herr LOUISOT, Benjamin (Karlsruhe Institute of Technology); LORENZ, Christof (Karlsruhe Institute of Technology, Karlsruhe, Germany); Herr SCHÄFER, David (Helmholtz Centre for Environmental Research); Dr. GROH, Jannis (Forschungszentrum Jülich GmbH); Herr SORG, Jürgen (Forschungszentrum Jülich GmbH); HANISCH, Marc (GFZ Helmholtz-Zentrum für Geoforschung); Herr INGENBLEEK, Martin (Forschungszentrum Jülich GmbH); Frau BÜTTNER, Nicole (Karlsruhe Institute of Technology); Dr. KUNKEL, Ralf (Forschungszentrum Jülich GmbH); Herr WIESEN, Robert (Helmholtz Centre for Environmental Research); FÖSIG, Romy

Vortragende(r): Dr. LOUP, Ulrich (Forschungszentrum Jülich GmbH)

Sitzung Einordnung: TALK SESSION

Track Klassifizierung: HMC Conference 2026 Track Topics: 3. Ontology-Driven Metadata Harmonization: Closing Semantic Gaps