EGU25-6484, updated on 19 Nov 2025
https://doi.org/10.5194/egusphere-egu25-6484
EGU General Assembly 2025

Harmonizing NetCDF Metadata Workflows: A Collaborative Initiative for Enhanced Data Integration and Reusability

Björn L. Saß[1], Romy Fösig[2], Emanuel Söding[3], Christof Lorenz[2], Sabine Barthlott[2], Philipp S. Sommer[1], Elke M. I. Meyer[6], Beate Geyer[6], Ulrich Loup[4], Corinna Rebmann[2], Katharina Loewe[2], Sibylle K. Haßler[2], Florian Obersteiner[2], Tobias Kerzenmacher[2], Angela Schäfer[5], Dorothee Kottmeier[5], Klaus Getzlaff[7], and Andrea Pörsch[8]

[1]Helmholtz Coastal Data Center, Helmholtz-Zentrum Hereon, Geesthacht, Germany (bjoern.sass@hereon.de)
[2]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Ocean Research Technology Centre, GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany
[4]Institute of Bio- and Geosciences, Forschungszentrum Jülich GmbH, Jülich, Germany
[5]Alfred Wegener Institute, Helmholtz Center for Polar- and Marine Research, Bremerhaven, Germany
[6]Institute of Coastal Systems - Analysis and Modeling, Helmholtz-Zentrum Hereon, Geesthacht, Germany
[7]Ocean Circulation and Climate Dynamics, GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany
[8]GFZ Helmholtz Centre for Geosciences, Potsdam, Germany

In the pursuit of making data FAIR (Findable, Accessible, Interoperable, Reusable)[1], datasets need to be well-described to enable both human and machine users to maximize the benefits towards knowledge discovery and innovation. Self-describing data enriched with metadata are essential to facilitate interoperability and reusability. Within the Earth System Science (ESS) community, the NetCDF data format has become the quasi-standard for storing multidimensional data, supported by general metadata standards such as the CF conventions (Climate and Forecast)[2], the Attribute Convention for Data Discovery (ACDD)[3] for global attributes and more specific ones like the AtMoDat project[4] for atmospheric modeling.

NetCDF files can be self-describing if they contain all relevant metadata. In practice, NetCDF metadata is frequently incompatible because they are either non-standardized or not as detailed as required for repositories or data portals. We aim to address these discrepancies by establishing a standardized NetCDF data workflow that ensures the seamless integration of NetCDF data into downstream processes and enables the extraction of metadata by downstream applications.

The NetCDF Metadata Guideline Initiative is a collaborative effort from researchers, research data management, research data infrastructure and metadata experts from several research centers across Germany. We are supported by the Helmholtz DataHub, the central data infrastructure of the Research Field Earth and Environment within the Helmholtz Association. This initiative aims to develop harmonized data handling guidelines to unify the diverse sub-communities within both observational and modeling fields, building towards a unified and consistent infrastructure for environmental research data.

Our approach includes examining existing guidelines, followed by their integration and expansion to address diverse needs within Earth and Environment disciplines. We will produce a set of comprehensive guidelines designed to enhance data interoperability and reusability, along with tools to facilitate their adoption.

Key milestones include:

- Reconciling attributes in former guidelines.
- Implementation of a collaborative and public guidelines document.
- Development of machine-readable templates and validation tools.
- Provision of highly user-friendly tools that support scientists when entering metadata profiles based on the guidelines.
- Integration of our enhanced NetCDF-profiles into selected downstream clients like the Earth Data Portal (EDP)[5].

By harmonizing metadata practices, we aim to enhance the interoperability and accessibility of geoscientific data, facilitating more efficient data sharing and utilization across research domains.

The implementation of standardized metadata practices for NetCDF across the ESS communities, would enable data repositories such as PANGAEA® Data Publisher and the World Data Center for Climate (WDCC) to present metadata in compliance with established norms and integrate it into their specific schemas.

This presentation will outline the key challenges identified, the proposed solutions, and the anticipated impact on the geoscientific community. With this presentation we call for participation in this evolving initiative to create a common NetCDF metadata foundation.

[1] *Wilkinson et al., 2016: https://doi.org/10.1038/sdata.2016.18*
[2] *https://cfconventions.org/*
[3] *https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3*
[4] *https://www.atmodat.de/*
[5] *https://earth-data.de*