

EGU21-8144

<https://doi.org/10.5194/egusphere-egu21-8144>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



## A Standard for the FAIR publication of Atmospheric Model Data developed by the AtMoDat Project

**Andrea Lammert**<sup>1</sup>, Anette Ganske<sup>2</sup>, Amandine Kaiser<sup>1</sup>, and Angelina Kraft<sup>2</sup>

<sup>1</sup>Deutsches Klimarechenzentrum GmbH, Data management, Hamburg, Germany (lammert@dkrz.de)

<sup>2</sup>Technische Informationsbibliothek (TIB), Welfengarten 1 B, 30167 Hannover, Germany

Due to the increasing amount of data produced in science, concepts for data reusability are of immense importance. One aspect is the publication of data in a way that ensures that it is findable, reusable, traceable and comparable (FAIR<sup>1</sup> principles). However, putting these principles into practice often causes significant difficulties for researchers. Therefore some repositories accept datasets described only with the minimum metadata required for DOI allocation. Unfortunately, this contains not enough information to conform to the FAIR principles - many research data cannot be reused despite having a DOI. In contrast, other repositories aid the researchers by providing advice and strictly controlling the data and their metadata. To simplify the process of defining the needed amount of metadata and of controlling the data and metadata, the AtMoDat<sup>2</sup> (Atmospheric Model Data) project developed a detailed standard for the FAIR publication of atmospheric model data.

For this purpose we have developed a concept for the “ideal” description of atmospheric model data. A prerequisite for this is the data publication with a DataCite DOI. The ATMODAT standard<sup>3</sup> was developed to implement this concept. The standard defines the data format as NetCDF, mandatory metadata (for DOI, landing page and data header), and naming conventions used in climate research - the Climate and Forecast conventions (CF-conventions<sup>4</sup>). However, many variable names used in urban climate research, for example, are not part of the CF-conventions. For this, standard names have to be defined together with the community and the inclusion in the list of CF-conventions has to be requested. Furthermore we developed and published Python routines which allow data producers as well as repositories to check model output data against the standard.

The ATMODAT standard will first be applied by the project partners of the two participating universities (University of Hamburg and Leipzig). Here, climate model data are processed with a post-processor in preparation for publication. Subsequently, the files including the specified metadata for the DataCite metadata schema will be published by the World Data Center for Climate<sup>5</sup> (WDCC). Data fulfilling the AtMoDat standard will be marked at the landing page by a special EASYDAB<sup>6</sup> (Earth System Data Branding) logo. EASYDAB is a currently developed branding for FAIR and open data from the Earth System Sciences. This indicates to future data users that the dataset is a verified dataset that can be easily reused. The standardization of the data and the

further steps are easily transferable to data from other disciplines.

1 Wilkinson, M., Dumontier, M., Aalbersberg, I. et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

2 <https://www.atmodat.de/>

3 [https://cera-www.dkrz.de/WDCC/ui/cersearch/entry?acronym=atmodat\\_standard\\_en\\_v3\\_0](https://cera-www.dkrz.de/WDCC/ui/cersearch/entry?acronym=atmodat_standard_en_v3_0)

4 <https://cfconventions.org/>

5 <https://cera-www.dkrz.de/WDCC/ui/cersearch/>

6 <https://www.easydab.de/>