Open Energy Metadata: Publishing Energy Data Enriched with Ontology References

Authors: Christian Hofmann[1], Hannah Förster[2], Ludwig Hülk[1], Christoph Muschner[1], Mirjam Stappel[3], Christian Winger[2], Lukas Emele[2], Martin Glauer[4], Adel Memariani[4]

Organizations: 1: Reiner Lemoine Institut, 2: Öko-Institut, 3: Fraunhofer IEE, 4: Otto-von-Guericke-Universität Magdeburg

Modelling climate and energy systems is a challenging endeavour. Scientists with different backgrounds, terminology and interests analyse large, complex, interdependent, and regulated systems. Some studies focus only on the electricity sector, while others will partially or fully include other sectors such as heat and transport in their calculations and analyses. Similar data sets are used that originate from different sources, have different temporal and spatial resolutions, different reference years or even divergent underlying sector concepts. For example, some input data is collected manually, while other data is a product of other models and calculations. In this research field, understandability, reproducibility, and collaboration are challenges for the scientific community. A lot of complex data needs to be published in an accessible and self-explanatory way to meet open science standards for software development and research data management.

A group of scientific institutes from the field of energy research has implemented the idea of creating and using a modular, shared and open research data infrastructure. The Open Energy Family framework is a collection of various tools and information designed to ensure quality, transparency and reproducibility in energy system modelling. It was originally formulated at an early workshop of the openmod initiative and primarily funded by the Federal Ministry for Economic Affairs and Energy (BMWi). The main module is the Open Energy Platform (OEP), a web interface that provides access to a community database. It enables the publication of data sets linked to the corresponding source code and underlying assumptions using a standardized metadata format, the Open Energy Metadata (OEMetadata). The format is publicly developed on GitHub and conforms to the Frictionless Datapackage standard. It is a json file that describes the structure and content of adjacent files or database tables. In its latest version (v1.5.0) we have refined it to allow references to ontologies. An ontology is a well-structured and defined description of reality that includes all the elements of interest and their interactions. It contains a formal naming and a definition of classes, properties and their relationships. An ontology can serve as a reference for concepts, terms and definitions within a particular domain of discourse. The Open Energy Ontology (OEO), a collaborative effort, aims to create a common ontology for energy system modelling.

The poster introduces our approach, combining standardized metadata with references to a domain ontology. We describe the structure and content of the OEMetadata standard and the semantic methods of the OEO. Applying an ontology to data annotations in databases and metadata allows flexible, content-oriented data integration and aggregation. It also offers the possibility of advanced searching functions and logical queries across data sets. With this connection, we can ensure that research data can be annotated in a well-defined and unambiguous way, making the data understandable and reusable.