Harmonizing Energy System Scenario Data Through Ontological Annotation: A Case Study Within the NFDI4Energy Ecosystem

Finn Niklas Peters[1,*] (ID), Aamir Muhammad[2] (ID), Felix Engel[2] (ID),
Mirko Schäfer[1] (ID), and Mirjam Stappel[3] (ID)

[1]INATECH, University of Freiburg, Germany

[2]TIB Leibniz Information Centre for Science and Technology, Germany

[3]Universität Osnabrück, Germany

*Correspondence: Finn Niklas Peters, finn.peters@inatech.uni-freiburg.de

Abstract. Long-term energy scenario studies based on Energy System Optimization Models (ESOMs) play a central role in shaping the transition of the German energy system. However, the diversity of model structures and variable naming conventions complicate cross-study comparison and limits reproducibility. This work presents an approach for improving the interoperability of scenario data through the use of ontologically annotated datasets. Within the NFDI4Energy framework, we conducted a case study that tested a set of semantic tools and workflows on data from the Ariadne project, which employs the models REMod, REMIND, and PyPSA-DE [1]. Using the Open Energy Ontology (OEO), the TIB Terminology Service, and the NFDI4Energy Annotator, we developed and evaluated a practical procedure for annotating energy scenario data in the IAMC format [2], [3]. The results demonstrate that automated suggestion tools substantially reduce manual effort while maintaining semantic quality after expert review. The documented workflow and derived recommendations aim to support scientists in performing transparent, reproducible, and machine-readable energy scenario analysis.

Keywords: Energy System Models, Ontology, FAIR Data, Semantic Annotation, Scenario Comparison

1. Introduction

Energy transition research relies heavily on scenario modeling to explore possible futures for national and sectoral energy systems. Studies employing ESOMs produce diverse datasets describing energy mix, emissions, capacities, and costs. To identify robust policy decisions, results from multiple studies should be compared [4]. Yet differences in variable definitions and model structures often hinder direct comparison and code availability also reduce the transparency of model outcomes [5]. To overcome these barriers, interoperability between datasets is essential. This requires a common vocabulary and consistent semantics across research outputs. The NFDI4Energy initiative seeks to provide infrastructure and standards that enhance data findability, accessibility, interoperability, and reusability (FAIR) for energy research.

2. Research Objective

Our main objective is to streamline and simplify the creation of inter-scenario and inter-model comparisons through ontological annotation of scenario data. By annotating variables with terms from a shared ontology, datasets can be queried, aligned, and compared more systematically. However, semantic annotation requires both domain expertise and familiarity with ontological tools—resources that are often scarce. Accordingly, we aimed to:

- Evaluate the effectiveness of existing NFDI4Energy services for ontology-based annotation,
- identify obstacles and potential improvements in the annotation workflow and the ontology, and
- formulate best-practice recommendations for the research community.

3. Methodology

We selected a large scenario dataset from the Ariadne project with variables formatted according to the IAMC template [3]. The dataset integrates outputs from the models REMod, REMIND, and PyPSA-DE, offering an ideal testbed for interoperability efforts. The following NFDI4Energy tools and services were integrated into our workflow:

- Open Energy Ontology (OEO) – the controlled vocabulary for energy domain concepts,
- TIB Terminology Service – providing ontology search and management functions,
- NFDI4Energy Annotator – offering automated term suggestions via a REST API.

The process involved (i) generating automated annotation suggestions, (ii) manual review and correction of proposed terms, and (iii) submission of new term proposals to the OEO when required. We documented the entire pipeline to ensure reproducibility.

4. Results and Lessons Learned

Automated annotation using the NFDI4Energy Annotator significantly reduced manual workload, providing rapid first-pass mappings of IAMC variables to OEO terms. Nonetheless, human validation remained vital to ensure semantic precision, especially for ambiguous or compound variables. Practical challenges included differing levels of granularity between model terminology and ontology classes, missing concepts in existing vocabularies, and partial overlaps between scenarios. The use of the TIB Terminology Service allowed refinement of the term mapping, while feedback to the OEO development team helped extend ontology coverage. It also showed that long-term scenario data features specific data that should not be covered by the scope of the OEO and instead be built into a new ontology that inherits parts of the OEO, but also adds more specific sectoral terms.

5. Discussion and Outlook

The case study demonstrates that ontology-based annotation can enhance the transparency and comparability of energy scenario data, moving toward machine-readable and queryable datasets. Future work will focus on:

- Integrating automated annotation directly into modeling pipelines;
- Developing community guidelines for FAIR data publication;
- Instead of extending ontology scope to cover all sectors, creating a new and modular ontology that reuses parts of the OEO and has sector modules.

Expanded collaboration within NFDI4Energy and beyond will be key to achieving sustainable data interoperability in the energy modeling community.

Data Availability Statement

The annotated dataset and the documented annotation workflow are derived from the Ariadne project scenario data [1]. The NFDI4Energy Annotator, TIB Terminology Service, and Open Energy Ontology [2] are publicly accessible.

Underlying and Related Material

The annotation workflow and tooling used in this study are based on the NFDI4Energy infrastructure, including the Open Energy Ontology [2] and the IAMC data format [3].

Author Contributions

Conceptualization F.P., M.Sch., M.St.; methodology F.N.P., F.E., A.M., M.St.; software F.N.P., F.E., A.M.; data curation F.N.P.; writing – original draft F.N.P.; writing – review & editing M.Sch.; supervision M.Sch.

Competing Interests

The authors declare that they have no competing interests.

Funding

References

[1] G. Luderer et al., *A cost-efficient energy transition: Scenarios for climate neutrality 2045–summary*, Ariadne-Report, 2025.

[2] M. Booshehri et al., "Introducing the open energy ontology: Enhancing data interpretation and interfacing in energy systems analysis", *Energy and AI*, vol. 5, p. 100 074, 2021. DOI: https://doi.org/10.1016/j.egyai.2021.100074.

[3] International Institute for Applied Systems Analysis. "IAMC Data Format Repository", Accessed: Dec. 8, 2025. [Online]. Available: https://github.com/IAMconsortium/common-definitions.

[4] M. G. Prina, B. Nastasi, D. Groppi, S. Misconel, D. A. Garcia, and W. Sparber, "Comparison methods of energy system frameworks, models and scenario results", *Renewable and Sustainable Energy Reviews*, vol. 167, p. 112 719, 2022.

[5] R. Morrison, "Energy system modeling: Public transparency, scientific reproducibility, and open development", *Energy Strategy Reviews*, vol. 20, pp. 49–63, 2018.