The Journal of Open Source Software

Flexible Metadata Annotation for Agent-based Energy System Models

Christoph Schimeczek [1¶], **Felix Nitsch** [1,2], **and Ulrich Frey** [1,3]

1 German Aerospace Center (DLR), Institute of Networked Energy Systems, Curiestr. 4, 70563 Stuttgart, Germany **2** BOKU University, Institute of Sustainable Economic Development, Feistmantelstraße 4, 1180 Vienna, Austria **3** University of Graz, Department of Environmental Systems Sciences, Merangasse 18, 8010 Graz, Austria ¶ Corresponding author

DOI: 10.21105/joss.09541

Software

- Review
- Repository
- Archive

Editor: Frauke Wiese

Reviewers:

- @lkstrp
- @sstroemer

Submitted: 01 August 2025
Published: 11 August 2026

Summary

We introduce a comprehensive and flexible way of metadata annotation for agent-based energy system models using FAME-Io (Nitsch et al., 2023). This approach allows users to annotate modelling components, model instances, and model data with metadata. Provided metadata are used to automatically generate metadata summaries for model output. Our implementation is agnostic to the metadata structure and can thus be applied following any standard, e.g., the OEMetadata template (Hülk et al., 2025) or the CodeMeta standard (Jones et al., 2023). It applies to both the metadata annotations and the generated metadata summary files.

FAME-Io itself is part of the modelling framework FAME. FAME-Io reads human-readable input files and compiles a binary representation to be forwarded to the simulation execution software FAME-Core (Schimeczek, Deissenroth-Uhrig, et al., 2023). Human-readable results can be extracted using FAME-Io again from the simulation output binaries. Hence, the original simulation input can be recovered. Each step of the workflow shown in Fig. 1 can now be enriched with metadata, thus enabling maximum transparency of modelling with the FAME framework.

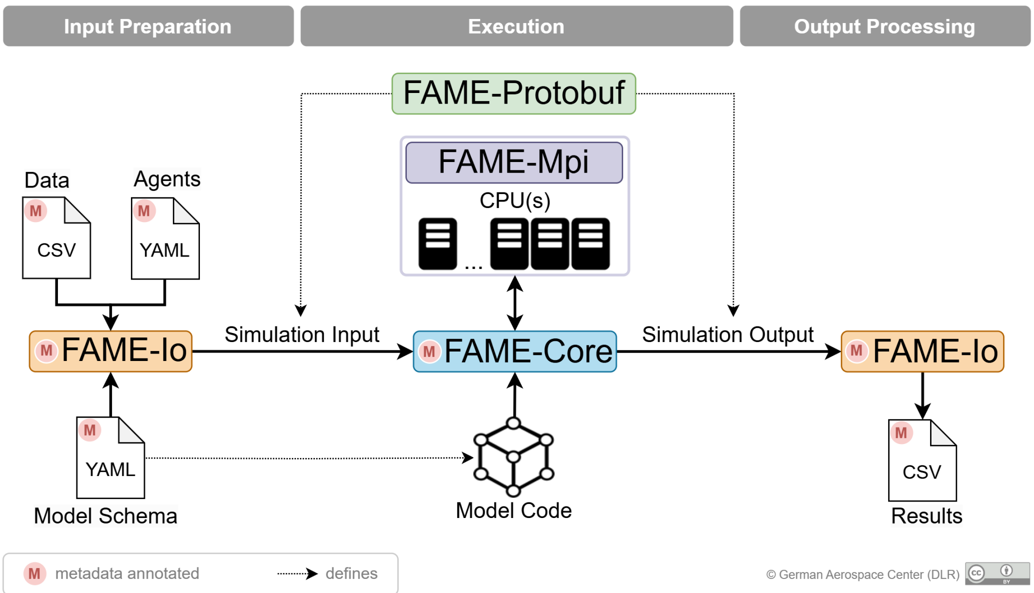

Figure 1: Metadata in the FAME workflow.

Statement of need

Metadata are critical for the long-term organization, retrieval, and interpretation of data across diverse domains (Leipzig et al., 2021). They provide essential context for researchers by describing the attributes, structure, and provenance of data. By standardizing information, metadata ensure interoperability and enable effective data management practices. As a result, usability and data integrity are improved (Diercks et al., 2023).

This is especially important for the reproducibility of complex scientific workflows (Villamar et al., 2025) and modelling results, yet significant barriers remain (Ivie & Thain, 2019). Metadata annotation of scientific models, however, has increasingly gained traction, especially after various reproducibility crises (Leipzig et al., 2021). First standardisation attempts have been made (Druskat et al., 2021; Jones et al., 2023), however, these metadata do not suffice to fully describe models and their workflows. In the energy systems domain, only a small fraction of the models provide even these basic metadata. At the time of writing, 19 out of 117 open source models registered on the Open Energy Platform provide a `CITATION.cff` file, and only two feature a `codemeta.json` file. To the best of our knowledge, no model includes a complete set of metadata annotations of inputs, outputs, and modelling components like we demonstrate below.

The presented approach enables modellers to comprehensively annotate FAME-based models, their components, and model configurations with metadata and to reuse these metadata when describing model results. This allows preserving model-specific information within the generated data sets and thus strengthens modelling transparency and result explainability. In particular, our flexible approach does not put any constraints on the metadata structure, allowing to satisfy domain-specific needs on a per-model basis, important for transferability (Villamar et al., 2025). While the presented work is not directly applicable to other modelling frameworks, we provide an instructive example and template on how metadata can be easily integrated in modelling workflows. This template adheres to the OEMetadata standard (Hülk et al., 2025) and thus follows suggestions by Schlögl et al. (2026) to use common and open data formats and to adopt ontologies. With it, we hope to pave the way of comprehensive metadata annotation for other scientific models and tools.

Metadata annotation

With FAME-Io all aspects of models and inputs are described using YAML files. YAML allows creating complex data structures from lists, key-value dictionaries, and scalar values. Since release 3.5 of FAME-Io every model aspect can be annotated with metadata. To this end, FAME-Io reserves the keyword `metadata` that can be added to any element that is a dictionary (see Fig. 2). All content below the metadata keyword is then automatically associated with the dictionary. To annotate a scalar, it is first transformed into a dictionary by associating the scalar with the reserved keyword `value`. Then, the metadata keyword can be used to add metadata to the scalar. In a similar way, lists can be annotated by embedding them in a dictionary, this time associating the list elements with the reserved keyword `values`. To increase robustness, FAME-Io uses case-insensitive keyword detection.

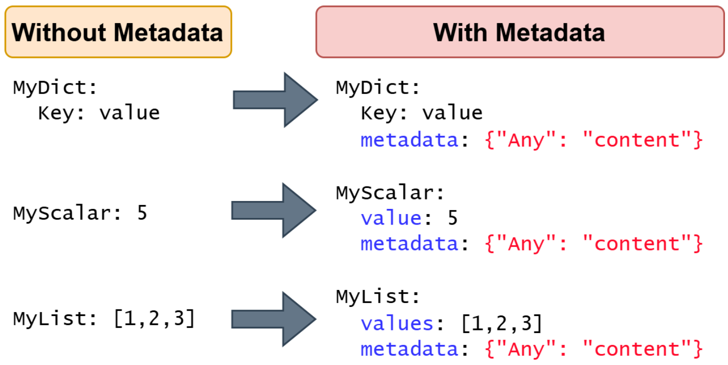

Figure 2: Example on how to add metadata annotations to mappings, scalars, and lists.

In order to extract and store metadata annotations from every model configuration element, FAME-Io classes and data structures were enhanced. All data structures inherit an abstract base class Metadata that automatically extracts input under the keyword `metadata` and injects the stored metadata when the data structures are exported. This enables metadata description of agent classes, their attributes, and output columns on the model level. In addition, each scenario, each agent instance, and each input value can be annotated on the execution level. The structure of all FAME-Io input files is fully described in its documentation.

Since no metadata standard for the annotation of models in the energy systems domain exists yet, we decided not to restrict the structure of the metadata. Thus, any content can be placed below the metadata keyword. Once standards evolve, implementing tools to check the compatibility of metadata annotations will be considered. Adapting to a standard should then be easy, given that just the terms below metadata sections have to be refactored.

Automatic metadata summary

FAME-Io creates a `metadata.json` file describing the extracted simulation outputs. Previously provided input and model metadata are used in this process. Again, no standard for metadata description of model results has been established so far, but the OEMetadata template is congruent in that respect. Therefore, FAME-Io provides a template compatible with the OEMetadata specifications. FAME-Io's templating engine, however, allows defining the structures for the created metadata file — and which sections of metadata to use from the model metadata.

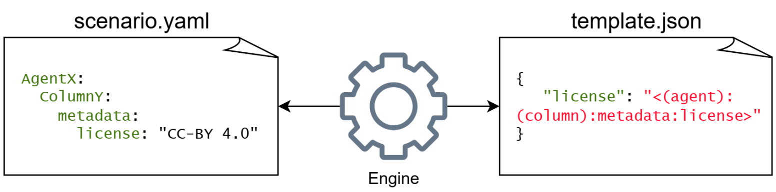

Figure 3: Relation between model metadata, metadata template, and output metadata files.

As illustrated in Fig. 3, the templating engine will search for replacement statements (indicated by angle brackets) in the template and then try to locate the information in the original model or input metadata. Since FAME-Io creates an individual output file per type of agent, agent-specific metadata can be matched to these files using the keyword `(agent)`. Similarly, metadata specific to a column within an output file can be matched with the column section using the keyword `(column)`. Iteration statements in the template enable covering any number of agents and output columns without redundancy in the template. This allows a slim and generalisable definition of model output metadata templates without requiring knowledge of the actual model structure.

The FAME-Io documentation covers in detail which metadata are used by the provided OEMetadata template and how own templates can be defined.

Example application

We provide a fully-fledged example of metadata annotation for the electricity market model AMIRIS (Schimeczek, Nienhaus, et al., 2023) in the repository AMIRIS-Examples. Model components have been fully annotated regarding their inputs, outputs, and capabilities. Wherever available, the closest concepts in the Open Energy Ontology, see Booshehri et al. (2021), are provided. Every output of the model is fully described by an accompanying metadata file. Extracting the metadata and compiling the output metadata file has a negligible impact on the total runtime of the FAME-Io routines. The additional data amount to a few kilobytes for our example scenarios. This demonstrates that comprehensive metadata annotation can be integrated into existing modelling workflows with low computational overhead, while substantially improving the documentation and interpretability of simulation input and output.

AI usage disclosure

No generative AI was used in the creation of this software, its documentation, or this paper.

Funding

This project was funded by the German Aerospace Center (DLR) as part of the Helmholtz Metadata Collaboration (HMC).

Acknowledgments

The authors would like to thank Ludwig Hülk, Carsten Hoyer-Klick, and Eugenio Salvador Arellano Ruiz for their advice and comments. Also, the authors express their gratitude to Lukas Trippe and Stefan Strömer for their careful review and helpful comments.

References

Booshehri, M., Emele, L., Flügel, S., Förster, H., Frey, J., Frey, U., Glauer, M., Hastings, J., Hofmann, C., Hoyer-Klick, C., Hülk, L., Kleinau, A., Knosala, K., Kotzur, L., Kuckertz, P., Mossakowski, T., Muschner, C., Neuhaus, F., Pehl, M., … Stappel, M. (2021). Introducing the open energy ontology: Enhancing data interpretation and interfacing in energy systems analysis. *Energy and AI*, *5*, 100074. https://doi.org/10.1016/j.egyai.2021.100074

Diercks, P., Gläser, D., Lünsdorf, O., Selzer, M., Flemisch, B., & Unger, J. F. (2023). Evaluation of tools for describing, reproducing and reusing scientific workflows. *Ing.grid*, *1*. https://doi.org/10.48694/inggrid.3726

Druskat, S., Spaaks, J. H., Chue Hong, N., Haines, R., Baker, J., Bliven, S., Willighagen, E., Pérez-Suárez, D., & Konovalov, O. (2021). *Citation File Format*. Zenodo. https://doi.org/10.5281/zenodo.5171937

Hülk, L., Huber, J., Hofmann, C., & Muschner, C. (2025). *Open Energy Family - Open Energy Metadata (OEMetadata)*. GitHub. https://github.com/OpenEnergyPlatform/oemetadata

Ivie, P., & Thain, D. (2019). Reproducibility in scientific computing. *ACM Computing Surveys*, *51*(3), 1–36. https://doi.org/10.1145/3186266

Jones, M. B., Boettiger, C., Cabunoc Mayes, A., Smith, A., Gruenpeter, M., Lorentz, V., Morrell, T., Garijo, D., Slaughter, P., Niemeyer, K., Gil, Y., Fenner, M., Nowak, K., Hahnel, M., Coy, L., Allen, A., Crosas, M., Sands, A., Chue Hong, N., … Ross, N. (2023). *CodeMeta: An exchange schema for software metadata. Version 3.0*. GitHub. https://github.com/codemeta/codemeta

Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K., & Greenberg, J. (2021). The role of metadata in reproducible computational research. *Patterns*, *2*(9). https://doi.org/10.1016/j.patter.2021.100322

Nitsch, F., Schimeczek, C., Frey, U., & Fuchs, B. (2023). FAME-Io: Configuration tools for complex agent-based simulations. *Journal of Open Source Software*. https://doi.org/10.21105/joss.04958

Schimeczek, C., Deissenroth-Uhrig, M., Frey, U., Fuchs, B., El Ghazi, A. A., Wetzel, M., & Nienhaus, K. (2023). FAME-Core: An open framework for distributed agent-based modelling of energy systems. *Journal of Open Source Software*. https://doi.org/10.21105/joss.05087

Schimeczek, C., Nienhaus, K., Frey, U., Sperber, E., Sarfarazi, S., Nitsch, F., Kochems, J., & El Ghazi, A. A. (2023). AMIRIS: Agent-based market model for the investigation of renewable and integrated energy systems. *Journal of Open Source Software*. https://doi.org/10.21105/joss.05041

Schlögl, M., Waltersdorfer, L., Regner, P., Siposova, A., & Brenning, A. (2026). Overcoming barriers to reproducibility in geoscientific data analysis: Challenges and practical implementation strategies. *Environmental Modelling & Software*, *200*, 106962. https://doi.org/10.1016/j.envsoft.2026.106962

Villamar, J., Kelbling, M., More, H. L., Denker, M., Tetzlaff, T., Senk, J., & Thober, S. (2025). Metadata practices for simulation workflows. *Scientific Data*, *12*, 942. https://doi.org/10.1038/s41597-025-05126-1