Ontology-driven Data Curation and Knowledge Modeling for Catalyst Layers in Polymer Electrolyte Fuel Cells

Marjan Kohandani[1,2], Mohammad J. Eslamibidgoli[1,2], Michael H. Eikerling[1,2], Kourosh Malek[1,2]

[1] Theory and Computation of Energy Materials (IET-3), Institute of Energy Technologies, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[2] Centre for Advanced Simulation and Analytics (CASA), Simulation and Data Science Lab for Energy Materials (SDL-EM), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Abstract

Research data management in the hydrogen technology field is challenging because large volumes of heterogeneous data are produced [1]. Electrochemical technologies such as fuel cells and electrolyzers are multicomponent devices, with various manufacturing routes being followed and a wide range of characterization and performance measurements applied. The governing phenomena span multiple length and time scales, creating a complex parameter–property space [2], while data are reported with inconsistent standards and formats. In this work, we build a FAIR and searchable knowledge graph for a concrete use case: catalyst layers in polymer electrolyte fuel cells. Our approach consists of three phases. In the first phase, we define the scope of a domain research question and build an initial ontology. In the second phase, we screen the literature and create a PDF corpus, then extract and curate the data into a structured, machine-readable format guided by the ontology, including terminology alignment and unit harmonization for cross-study comparability. In the third phase, we map the curated dataset into a Neo4j knowledge graph and release it as a FAIR resource. Overall, as will be shown at the conference, this workflow enables standardized, traceable, and AI-ready datasets that can be reused across studies to accelerate data-driven discovery and decision-making.

[1] Dreger, M., Eslamibidgoli, M. J., Eikerling, M. H., & Malek, K. (2023). Synergizing ontologies and graph databases for highly flexible materials-to-device workflow representations. Journal of Materials Informatics, 3(1), N-A

[2] Liu, H., Ney, L., Zamel, N., & Li, X. (2022). Effect of catalyst ink and formation process on the multiscale structure of catalyst layers in PEM fuel cells. Applied Sciences, 12(8), 3776.