An open endpoint and framework for the development of linked data for building energy systems

James Allan [1], Jong Gwan An[1], Reto Frickerr [1] , Sascha Stoller [1] and Philipp Heer [1]

[1] *Urban Energy Systems Laboratory, Empa, Ueberlandstrasse 129, 8600 Dübendorf, Switzerland*

Abstract

The NEST building is a research and innovation building in Switzerland. Data collected by the NEST building is available for researchers and organisations to develop models and test applications. This data resource includes an endpoint to an RDF knowledge graph containing linked data about the building and its energy systems (https://graphdb.nestcloud.ch/).

Our primary motivation is to improve the energy performance of the building, which includes both simulation and control of the energy systems. We have diverse static datasets containing information about the building and its energy systems. The datasets include BIM models, sensor metadata, engineering schematics and 3D city models. There are also real-time data streams and historical data from sensors and actuators installed in the building. The BMS records and stores approximately 10,000 measurements every minute. Semantic web technologies offer a promising solution to improve the discoverability and interoperability of this data. They also provide a flexible and extensible framework for representing and querying complex data. Despite these benefits, RDF knowledge graphs are challenging to create and maintain. A lack of strictness introduces scope for errors and inconsistencies in the data; engineers have raised concerns that models will differ between creators, even if they follow the same ontologies. Standards such as SHACL help achieve consistency at the expense of flexibility. We have also found that converting all data into RDF is impractical. There is also an abundance of ontologies to represent building and energy systems concepts, which contain many overlapping classes and properties, e.g. BOT, BRICK, ifcOWL, SSN/SOSA, SAREF, RealEstateCore. Despite this, there are limited examples of knowledge graph instances of the same building according to different ontologies, which makes it difficult to compare data modelling approaches.

We have recognised a need for a traceable and repeatable framework for developing knowledge graphs for building energy applications. Such a framework will enable different approaches to be replicated and compared. We propose an approach to maintain the connection between the knowledge graph and the raw input datasets. We achieve this through version-controlled repositories for processing scripts and data, e.g. GitHub and Zenodo. The connection between the raw input data and knowledge graph should be maintained so that if one is updated, the other is synchronised. We aim to provide a framework for the continual evolution of pipelines to generate knowledge graphs for buildings and their energy systems. We will present our initial pipeline and the resulting knowledge graph. We will give an overview of the challenges faced and an outline for future research.

Keywords

Data Integration, Knowledge graphs, Use case

Proceedings LDAC2023 – 11th Linked Data in Architecture and Construction, June 15–16, 2023, Matera, Italy
EMAIL: james.allan@empa.ch (J. Allan); jonggwan.an@empa.ch (J.G. An); reto.fricker@empa.ch (R. Frickerr); sascha.stoller@empa.ch (S. Stoller); philipp.heer@empa.ch (P. Heer)
ORCID: 0000-0001-8763-5831 (J. Allan); 0000-0002-6027-1027 (J.G. An); 0000-0003-2999-5753 (P. Heer)