VizBrick: A GUI-based Interactive Tool for Authoring Semantic Metadata for Building Datasets

Sangkeun Lee[1], Borui Cui[1], Mahabir Bhandari[1], Na Luo[2], and Piljae Im[1]

[1] Oak Ridge National Laboratory, Oak Ridge TN 31380, USA
{lees4, cuib, bhandarims,impl}@ornl.gov
[2] Lawrence Berkeley National Laboratory, Berkely CA 94720, USA
nlulo@lbl.gov

Abstract. Brick ontology is a unified semantic metadata schema to address the standardization problem of buildings' physical, logical, and virtual assets and the relationships between them. Creating a Brick model for a building dataset means that the dataset's contents are semantically described using the standard terms defined in the Brick ontology. It will enable the benefits of data standardization, without having to recollect or reorganize the data and opens the possibility of automation leveraging the machine readability of the semantic metadata. The problem is that authoring Brick models for building datasets often requires knowledge of semantic technology (e.g., ontology declarations and RDF syntax) and leads to repeated manual trial and error processes, which can be time-consuming and challenging to do without an interactive visual representation of the data. We developed VizBrick, a tool with a graphical user interface that can assist users in creating Brick models visually and interactively without having to understand the Resource Description Framework (RDF) syntax. VizBrick provides handy capabilities such as keyword search for easy find of relevant brick concepts and relations to their data columns and automatic suggestions of concept mapping. In this demonstration, we present a use-case of VizBrick to showcase how a Brick model can be created for a real-world building dataset.

Type of Submission: Demo
Keywords: Interactive tool, Brick ontology, Data standardization

Acknowledgement

Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a non-exclusive, paid up, irrevocable, world-wide license to publish or reproduce the published form of the manuscript, or allow others to do so, for U.S. Government purposes. The DOE will provide public access to these results in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

1 Introduction

Datasets collected from smart buildings are critical assets for research to improve human interactions in built environments. One of the biggest challenges of managing smart building data assets is that they are not standardized in many cases due to being collected from many sources by various entities (e.g., different research organizations, etc.) for different reasons. Such lack of a standardization disables the reusability of data assets and cause redundancy of data, which inherently increases the cost significantly. Standardizing building datasets. Brick ontology [1] is the state-of-the-art open-source unified schema that consists of extensible concepts and relationships that can semantically describe physical, logical, and virtual building dataset assets [2]. Creating a Brick model for a building dataset means that the contents of the datasets are semantically described using the standard terms defined in the Brick ontology. It will enable the benefits of data standardization, without having to recollect or reorganize the data. Moreover, semantic description of a dataset enables machine-readability which allows potential automation of data processing and utilization.

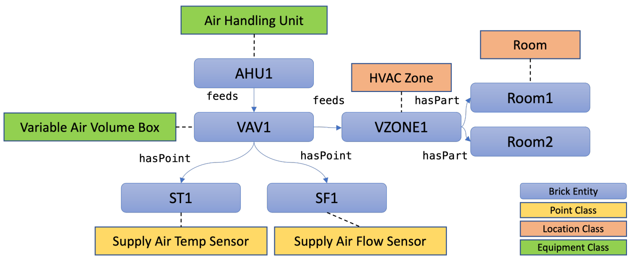

Fig. 1. A visual representation of a Brick model created for an example building dataset

Fig. 1 shows an example of a Brick model created for an example synthetic building dataset. From the figure, we can understand what equipment, locations, and points (entities) are related to this dataset, and more importantly we can also understand the relationship between the entities.

A general workflow of Brick model creation for a building dataset is as follows. Reviewing available non-semantic metadata (e.g., spreadsheet) for the dataset is the first step of the workflow. In this step, we aim to semantically understand the building structure, equipment, and measurements. Then, we define entities such as Location, Zone, Equipment, and Point. Then, we create relationships across the created entities, for instance, where entities are located, what entities are part of other entities, and so on. During the process, proper Brick ontology concepts and relationships need to be identified and mapped to the instances. Next, validating the created Brick model both in syntax and semantics is necessary. Steps in this workflow may be repeated with trial and error before finalizing the created model, and it can be time-consuming and challenging to do without an interactive visual representation of the data. Another challenge is that building scientists often have limited knowledge in semantic technology such as class and relationship definitions in Brick ontology and RDF syntax.

2 VizBrick Implementation and Capabilities

VizBrick provides three main capabilities as follows (Fig. 2.). Firstly, it provides an interface to review an existing tabular formatted non-Semantic metadata for a building dataset that user targeted to create a Brick model for. VizBrick imports a CSV (comma-separated values) file format that contains the list of data labels (data points) and their description. Users can review and modify the metadata before starting the entity and relationship creation process. Secondly, it provides an interactive entity and relationship creation interface. Users can perform keyword search against the Brick ontology and manually select specific class and relationship of Brick ontology to create entity and relationship instances. VizBrick also provides a suggestion capability that provides considerably properly matching class in the Brick Ontology for each data points in the original metadata. For keyword search and suggestion, VizBrick uses the TF-IDF (term frequency-inverse document frequency) technique [3] on the datapoint and Brick ontology descriptions.

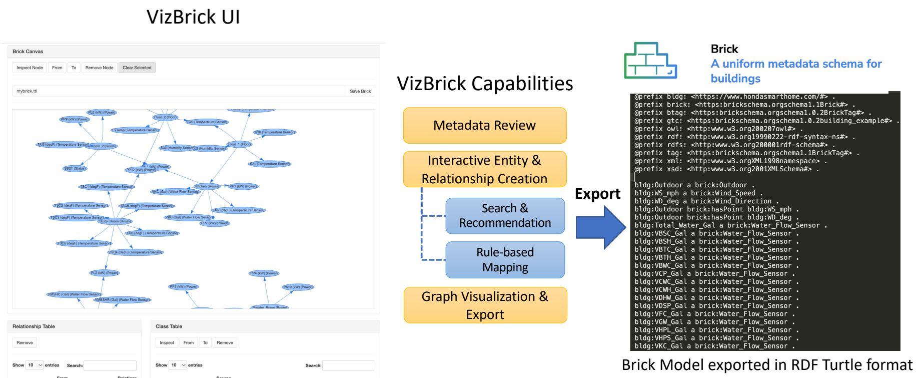

Fig. 2. Overview of the VizBrick

The recommendation result is provided as a ranking list with score and users can provide additional keywords for better matching (Fig. 3).

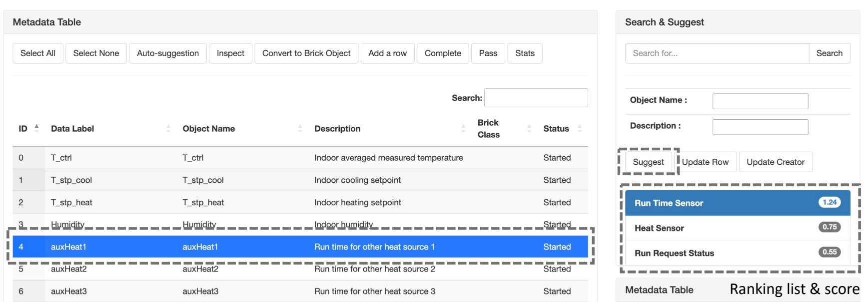

Fig. 3. VizBrick provides a list of Brick class suggestions that most likely to be matching with the selected data label.

In addition, VizBrick provides a rule-based mapping between data labels and Brick classes and edge creations using user-defined rules based on string patterns in data labels (e.g., data labels related to 'temperature senso' start with 'T_') (Fig. 4). Lastly, users can select all or sub parts of the created Brick entities and relationships to visualize the currently editing Brick model and export it as a standard Resource Description Framework (RDF) data model represented in a Terse RDF Triple Language (Turtle) format. Exported RDF files can be used with any other standard Brick related tools [4]. VizBrick has been released as an open-source software and registered in the US Department of Energy (DOE) Office of Scientific and Technical Information (OSTI) [6]. The tool is publicly available for testing and a demonstration video is available.

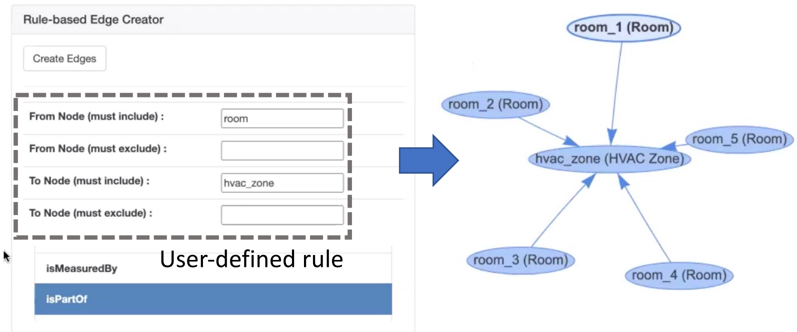

Fig. 4. User can create multiple relationships between entities based on user-defined rules.

3 Demonstration Plan and Conclusion

For demonstration, we will show how VizBrick can be used to create a Brick model for Ecobee Donate Your Data (DYD) that comprises data collected by Ecobee thermostats from 1,000 single family homes in four US states. Starting with a simple tabular CSV-formatted metadata file that describes the dataset, we will demonstrate how a semantic metadata file formatted in TTL file can be generated and used in the other standard Brick tools. The demonstration will take the following steps.

1. Importing, reviewing, and editing the original metadata
2. Creating Brick class and relationship instances using VizBrick search and suggestion toolbox
3. Visually interact with the work in progress Brick model
4. Exporting and validating the Brick model

This demo will showcase how VizBrick tool can significantly reduce the level of efforts to create a Brick model for a building data asset, which is crucial for data standardization.

References

1. Brick: A uniform metadata schema for buildings, https://brickschema.org/, last accessed 2022/6/30

2. Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., Gluck, J., Hong, D: Brick: Towards a unified metadata schema for buildings. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments, pp. 41-50 (2016).
3. Aizawa, A. :An information-theoretic perspective of tf–idf measures. Information Processing & Management, 39(1), 45-65. (2003).
4. Brickly: A block based editor for SPARQL queries with Brick, https://github.com/ezrichards/brickly, last accessed 2022/6/30
5. Lee, Sangkeun, Im, Piljae, Bhandari, Mahabir, & Cui, Borui., VizBrick. [Computer software]. https://github.com/liza183/vizbrick. https://doi.org/10.11578/dc.20220609.1.Ecobee Donate Your Data (DYD) dataset, https://bbd.labworks.org/ds/bbd/ecobee, last accessed 2022/6/30