Linking Data Model and Formula to Automate KPI Calculation for Building Performance Benchmarking

Yun-Yi Zhang[1], Zhen-Zhong Hu[1,2], Jia-Rui Lin[1,3,*], Jian-Ping Zhang[1,3]
[1] Department of Civil Engineering, Tsinghua University, Beijing 100084, China
[2] Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China
[3] Tsinghua University – Glodon Joint Research Center for Building Information Model (RCBIM), Tsinghua University, Beijing 100084, China
[*] corresponding author: lin611@tsinghua.edu.cn, jiarui_lin@foxmail.com

ABSTRACT Buildings consume a large proportion of global primary energy and building performance management requires massive data inputs. Key Performance Indicator (KPI) is a tool used for comparing different buildings while avoiding problems caused by heterogeneous data sources. However, silos of building and energy consumption data are separate, and the linkages between a KPI formula and different data sets are often non-existent. This paper develops an ontology-based approach for automatically calculating the KPI to support building energy evaluation. The proposed approach integrates building information from BIM and energy and environmental information collected by sensor networks. A KPI ontology is developed to establish a KPI formula, thereby linking static and dynamic data generated in the building operation phase. Each KPI can be defined by inputs, a formula and outputs, and the formula consists of parameters and operators. The parameters can be linked to building data or transformed into a SPARQL query. A case study is investigated based on the proposed approach, and the KPIs for energy and environment are calculated for a real building project. The result shows that this approach relates the KPI formula to the data generated in the building operation phase and can automatically give the result after defining the space and time of interest, thus supporting building performance benchmarking with massive data sets at different levels of details. This research proposes a novel approach to integrating the KPI formula and linked building data from a semantic perspective, and other researchers can use this approach as a foundation for linking data from different sources and computational methods such as formula created for building performance evaluation.

KEY WORDS Automation, Building Performance, KPI, Linked Data, Ontology, Sensor Network

I. INTRODUCTION

Buildings consume 40% of global primary energy[1], and are the most energy consumers in many countries and areas including the European Union[2], the USA[3] and China[4]. Researchers have found that systematic building energy management can help reduce energy use by 5% to 30%[1]. Building performance benchmarking (BPB) is an important approach utilized in building performance management as it compares performance metrics to best industry practices.

By measuring performance with specific indicators, or key performance indicators (KPIs), BPB gives building managers a clear view of the targets they need to meet to achieve efficient building energy consumption. For example, electricity, gas and water consumption per square meter, per person served or per guest room, and output/input ratio of building equipment are usually used in comparing a facility's performance with others'. To formulate the calculation process of KPIs, formula or equations are always used. A well-defined KPI provides a quantitative metric to compare buildings under different conditions regardless of building-specific factors and compare buildings under different conditions. that is, KPIs make it transparent so the building managers could find what's going well and what to do to further improve performance of a facility[5].

Meanwhile, it also compiles multiple data sources to calculate KPIs of a facility[6]. Not only performance sensing data based on Internet of Things (IoT) technology, but also properties and topological connections of spaces and building service systems are required[7]. Collected sensing data is usually modeled as time series and persisted in database, while data related to spaces and service systems is stored in as-design or as-built models based on CAD or BIM these days[8]. However, these data usually reside in different systems (or sources) and heterogeneous formats. Lacking of interoperability between different data silos hinders fully utilization of building energy data. Plenty of efforts have been invested in the integration of different data models to explode the value of big data. For example, Curry et al.[9] combined scenario modeling and linked data from different data silos to make assessments. Corry et al.[10] extended the use of linked data to establish the mapping between existing databases to aid assessments. Based on these works, it could be concluded that current works mainly focus on transforming one model to another or establishing semantic links between different data models. Nevertheless, few attentions were paid to semantic links between parameters of KPI formula and properties of data models. In this manner, process of calculating a specific KPI consists of 1) extract data from different data sources, 2) manually map extracted data to parameters of KPI formula, 3) calculate KPI based on mathematic formula automatically. This is time-consuming, tedious and error-prone. Even though automatic calculation of KPI is investigated previously, equations are embedded or hard coded in applications. Which means, there lacks the flexibility in creating or updating KPIs in accordance with clients' favor.

To solve the problem of the lack of data linkage, semantic web technologies provide an opportunity to represent information in structured graphs and integrate information from different sources. Semantic web technologies can be used to improve data interoperability, linkages across domains and logical inference methods[11]. An ontology can be developed with semantic web technology and defined as an explicit and shared conceptualization of a given domain that provides explicit logical assertions about information, which aids in converting human knowledge into a computer-understandable format[12,13]. In recent years, ontology and semantic technology have been widely used in the construction industry[11].

However, most research did not focus on linking input with KPI formula and the process is accomplished manually, which causes much rework Some ontologies have been established to increase the operability of building related data, but these research focuses on the data itself, not on the linkage between data and KPI formula, leaving alone the process of obtaining and loading data into the formula. To solve this problem, an otology is developed to link building energy data and the KPI formula, and an approach to calculate key performance indices using data integrated from heterogeneous data sources is proposed.

The remainder of the paper is organized as follows. In section II, a brief literature review is given, and the research gap is identified. In section III, an ontology to describe building energy consumption is developed, and a methodology for KPI calculation is proposed and introduced in detail. In section IV, a case study is provided to identify possible application scenarios and validate the feasibility of the proposed approach. Section V summarizes the conclusions, limitations and future research.

II. RELATED WORKS

A. Sensor and Building Information Integration

BIM (Building Information Modeling) is a process supported by various tools and technologies involving the generation and management of digital representations of the physical and functional characteristics of places[14]. Information generated throughout the design, construction and operation stages can be integrated to form nD models in a unified framework[15]. Therefore, the integration of building performance data into BIM has a promising future for the precise management of building operations[16].

The major data source of a building operation phase is the sensor system. Sensors in buildings are utilized to constantly collect real-time information, including (1) energy consumption (electricity, water, gas consumption, etc.), (2) device operation (the operation statuses of air conditioners, heaters, etc.), and (3) environmental quality (temperature, humidity, concentrations of toxic gases, etc.) information.

Many buildings are equipped with building management systems, which consist of monitoring and control parts. These systems are utilized to monitor environmental quality and energy consumption throughout the building operation phase and control each device according to predefined strategies [17].

Several studies have attempted to integrate sensor data with BIM. For instance, Arslan et al. developed a prototype system called the "real-time environmental monitoring, visualization, and notification system" using BIM and wireless sensor networks (WSNs)[18]. Riaz et al. proposed a BIM- and sensor-based data management system for automating the management of health and safety issues at construction sites[19]. Natephra et al. proposed the integration of thermal and environmental data provided by sensors with BIM to assess the thermal performance of the building envelope[20]. Suprabhas and Dib discussed the feasibility of using sensor data combined with BIM for maintenance-based facility management[21].

In these studies, monitoring data were integrated with BIM in many scenarios; however, energy consumption monitoring data have seldom been collected and integrated with BIM in a standardized data model for further utilization and analysis.

B. Building Related Ontologies

There are a number of available ontologies that are aimed at sharing and connecting cross-domain data in the building domain[11]. For example, the ifcOWL ontology is defined as an OWL (web ontology language) representation of IFC (industrial foundation classes) data and serves as an alternative representation of the EXPRESS schema of IFC[22]. A corresponding file-based IFC-to-RDF (resource description framework) conversion application has been developed[23]. The semantic sensor network (SSN) ontology is based on the concept of a stimulus prompting an observation[24]. The SSN includes sensors, their observations, and knowledge of their environment[25]. The BOT (building topology ontology) is a minimalist ontology for describing the core topological concepts of a building[26]. BOT deletes the unnecessary details of ifcOWL in the scope of the geometric and topological representations of a building in specific cases.

One primary example of an energy simulation model is SimModel, which was devised as an interoperable data model for the exchange of simulation data between energy simulation tools. This model is available in an OWL ontology[27] and can be used to generate RDF graphs of model data[28]. By exporting the data into an RDF data model, they can be easily combined with other RDF data. However, the actual combination and management of IFC and SimModel graphs remain topics of discussion. Sørensen et al. reviewed the existing ontologies relevant to creating digital links between virtual models and physical components in the construction process to improve information handling and sharing in construction and building operation management[29]. Corry et al. proposed a semantic-based approach to integrating heterogeneous building data[30]. Semantic web technologies have been used in environmental monitoring to facilitate knowledge encoding and data integration outside the construction environment[31]. Metal et al. used an ontology to integrate air quality and 3D city models[32]. Opera described an ontology for air pollution analysis and control and applied the ontology in expert and multiagent systems[33]. Reitzes and Snyder developed an ontology for real-world indoor environmental quality monitoring and control[34]. Stocker et al. devised an ontology-based environmental monitoring system to measure and compute mean hourly PM2.5 concentrations[35]. Pundt et al. described the use of ontologies via the Internet on the basis of an example involving GIS (Geographic Information System) supported environmental monitoring in the field[36]. Dibley et al. proposed an ontological framework for intelligent sensor-based building monitoring with a focus on the ontology development process to deliver an intelligent multiagent software framework that supports real-time building monitoring[37]. Moreover, there are also a few attempts to devise a comprehensive ontology to express the linkage between monitor data and building itself. Balaii et al devised a uniform schema for representing metadata in buildings called Brick, linking location, equipment and measurement[38]. Mahdavi et al. devised an otology for building monitoring, linking building environment, inhabitant and control systems and devices[39-41]. Yehong Li et al. developed an ontology called EM-KPI focusing on energy management in district and building levels, with a reference to MathML to express the definition of KPIs and input parameters can be extracted[42].

These existing ontologies mainly focus on the environment at the urban level, and the information related to building environmental monitoring is not effectively organized, and recent attempts to express building monitoring data still leaves it open to establish a method to explore the data. An integrated semantic modeling approach for the KPI of building performance would be beneficial to comprehensively understanding building performance.

C. Other Recommendations

Several studies have developed applications to support building performance analysis by combining building information and energy consumption data[11,43]

For instance, Curry et al.[9] combined scenario modeling and linked data to support decisions in building design and operation stages. Curry et al.[44] and O'Donnell et al.[45] further extended the use of linked data combined with diverse cross-domain building data to support operational building management. Corry et al.[10] discussed using semantic web technologies to aid the integration of AEC data into an existing building performance framework for evaluating building performance in the operational phase. Corry et al.[46] also developed a performance assessment application based on a corresponding ontology. Shushan Hu et al.[47] attempted to combine linked data with OpenMath to retrieve information from separate multibases and describe building performance metrics. Botao Zhong et al.[17] developed an ontology for building environment compliance assessment.

Tomasevic et al.[48] focused on the operational phase and discussed the use of an ontology-based building performance analysis method to provide feedback to facility managers. Furthermore, Dibley et al.[37,49] proposed an OntoFM system to support real-time building monitoring with a multiagent system and access to semantic building data. The building data can be tracked by the OntoSensor ontology[50] and a general-purpose ontology, the Suggested Upper Merged Ontology (SUMO[51]).

These studies suggest that ontological methods are promising for the integration of relevant data generated by different sources and can support data inference. However, previous research focused on building data integration and succeeded in integrating static building data, sensor networks, energy consumption data, etc.; however, a common method of processing integrated data does not exist. The performance evaluation process is separate from data collection, and data are first retrieved and analyzed with predefined formulas. Little attention has been paid to linking the semantics of the formula parameters with the data to support automatic and iterable building performance evaluations. Therefore, this research sheds light on the use of the KPI formula with linked data and the identification of the relationships between formulas and data to achieve automatic and iterable building performance evaluations.

III. METHODOLOGY

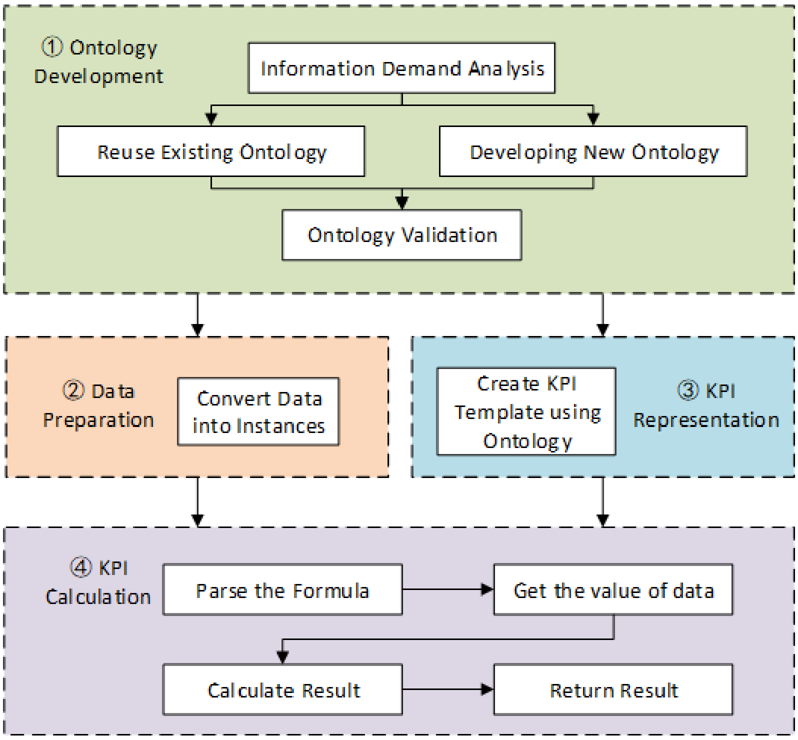

FIGURE 1. Overall methodology for ontology-based KPI Calculation

The overall methodology for KPI calculations based on the proposed ontology is shown in Fig. 1. The entire process requires linkages between building, sensor and observation data, so ontologies representing each discipline are developed. In this process, the information required for KPI calculations is first analyzed. The relevant existing ontologies related to building and sensor networks are also reviewed and used in the research. A KPI ontology linking these data is developed to form an integrated fusion model that uses the KPI formula and supports automatic KPI calculations. The ontologies are then created and validated with the help of Protégé, an open-source ontology editor.

Static and dynamic data are then converted to instances according to the ontologies established. In this study, BIM and the monitoring platform are the major data sources. The relevant data are extracted and converted into instances and stored in an RDF file with the help of the open-source library dotNetRDF in C#.

A KPI formula is first developed and then applied using the established ontology. A program is developed for users to create a KPI formula template in a graphical user interface and query calculation results in a certain space and time period.

IV. IMPLEMENTATION

A. Ontology development

1) Information Requirement

Building energy consumption monitoring and evaluation mainly require information from various data sources, including information about buildings, sensors, and energy consumption. These data can be divided into two main categories: static data and dynamic data. Static data are those that do not change with time or that remain constant over a relatively long period of time. Dynamic data are those that vary over time and are often in the form of streaming data. Static information and dynamic information must be linked to each other to yield an accurate KPI result, as shown in Fig. 2.

Static information includes (1) basic building information describing the building nature, such as building identifier, type, usage, and completion time information. This type of data generally requires manual inputs; however, BIM and some existing building monitoring systems may provide some related information[52]. (2) Building geometric & topological data describing the space and its distribution, including the building area, adjacent relations, and the hierarchy of spaces, are required. This type of data can be obtained from BIM. (3) A sensor network with known sensor type, position, accuracy, and collection frequency information is necessary. This type of data can be obtained from a sensor platform; however, some manual work may also be required.

Dynamic information is automatically and periodically collected by sensor networks and is often stored on monitoring platforms. These data include (1) energy consumption data describing electricity, water and gas consumption in different spaces and from different sources and (2) environmental information describing the environmental status, including temperature, humidity, CO_2 concentration, and other data.

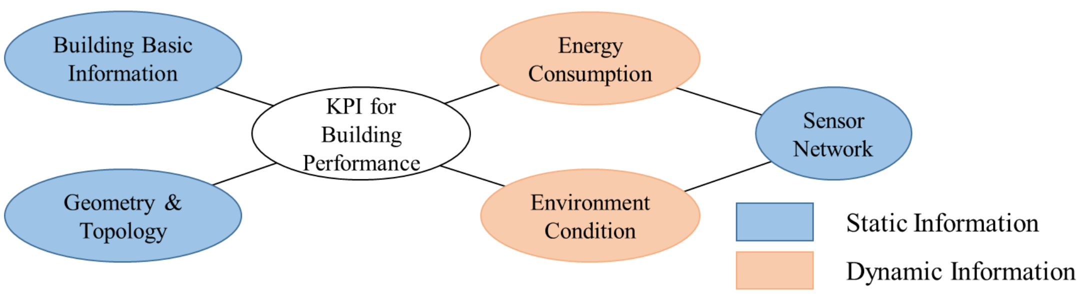

FIGURE 2. Information requirement for energy consumption evaluation

According to the information requirements above, ontologies describing the building information and sensor networks are selected and designed to represent the relevant information and linked data.

2) Building information ontology

There are some available ontologies that can be used to share and connect cross-domain data in the building domain. For example, the ifcOWL ontology was developed as an OWL representation of IFC data and serves as an alternative representation of the EXPRESS schema of IFC. ifcOWL is equivalent to IFC; hence, it includes literally everything in the IFC schema. However, ifcOWL is too complicated for building information representation in this research. As an alternative, BOT is a minimalist ontology that reflects the core topological concepts of a building. This approach was developed by the W3C Community Group. Considering the underlying problem, BOT is suitable for geometric and topological representations.

The major structure of BOT ontology is as Fig 3. The classes of the BOT ontology have the prefix “bot” and mainly include bot:Zone and bot:Element and are related with bot:hasElement. These instances represent the geometric and topological characteristics of a building. Class bot:Zone is divided into four layers: site, building, storey and space, making a parent-child reference relationship with bot:containsZone to form a tree structure. Class bot:Element describes roofs, doors, windows and exits in this research because the amount and form of them affect energy consumption.

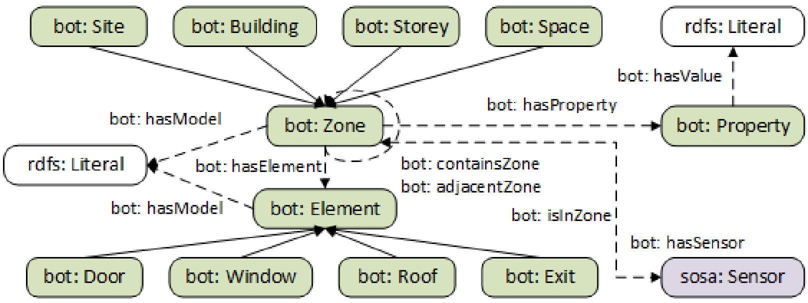

FIGURE 3. Building information ontology

3) Sensor information ontology

The SSN ontology was developed by the World Wide Web Consortium (W3C) Semantic Sensor Networks Incubator Group and is an ontology that describes sensors and their observations. The SSN includes a lightweight and self-contained core ontology called SOSA (Sensor, Observation, Sample, and Actuator) that encompasses the basic classes and properties. This ontological framework can describe sensors, observations and other related information as in Fig 4.

Each building may have one or more sensor platform represented by sosa:Platform, and they contain several sensors (sosa:Sensor). Each sensor locates in some certain space which is related to bot:Space, however, it must be made clear that the position of the sensor is not necessarily equal to the feature of interest. The core class in sensor ontology is sosa:Observation, linking the observation value with sosa:hasSimpleResult, and linking the observation time with sosa:resultTime. Moreover, each observation needs to be related to the space or equipment that the sensor monitors represented by sosa:featureOfInterest, which is an equivalence to bot:Space.

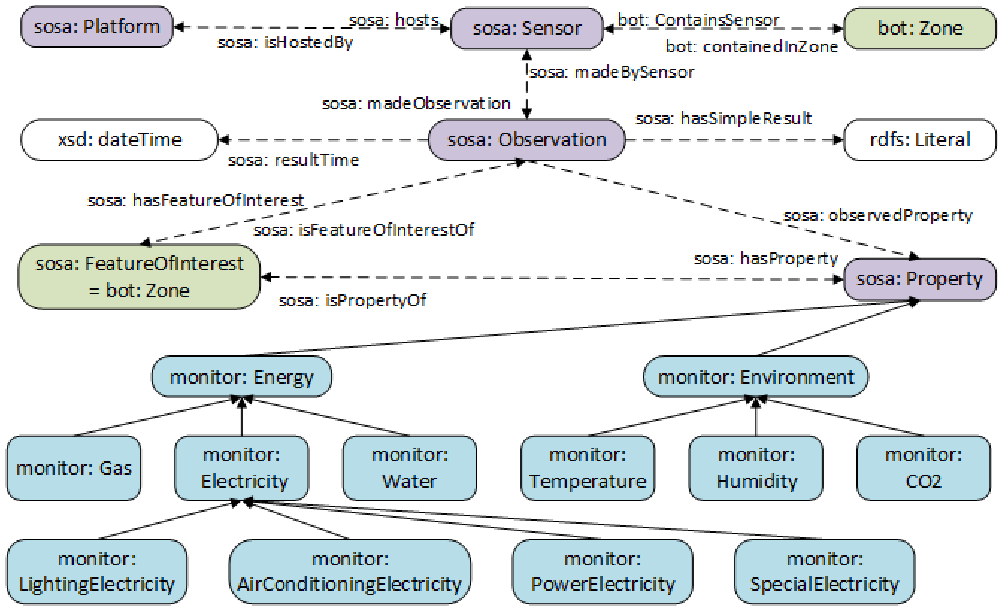

FIGURE 4. Sensor information ontology

4) KPI Ontology

The calculation of KPI requires an understanding of the linkages among different data sources, however, all the information needed is generated, stored and represented in different data formats in various information systems. This approach poses a considerable challenge for supporting energy consumption analysis. Fig. 5 shows an example of an energy consumption efficiency KPI, and breaking data storage barriers is crucial.

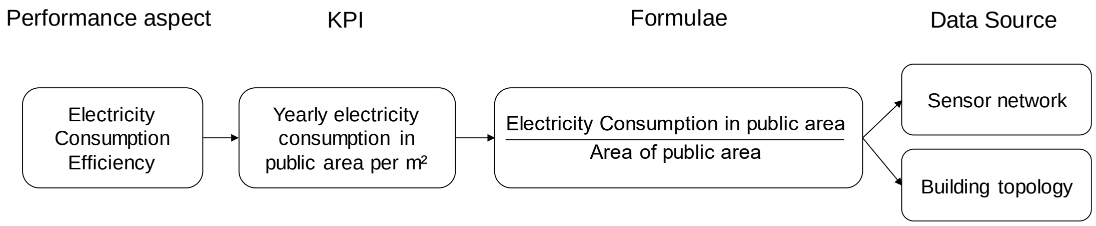

FIGURE 5. Information requirement for energy consumption evaluation

This research establishes a KPI ontology to recognize the linkage between different data sources. Each KPI indicates the energy consumption efficiency or building performance of a space or the entire building in a certain period from a specific perspective. Hence, the formula often requires lists of data in a time period, and the data source is often the sensor network, which is related to the building geometry and topology.

The calculation process associated with a KPI formula includes (1) computations with a constant, e.g., unit conversion, (2) aggregation operations on a list of data, e.g., averaging and summing, (3) normalizing the value or comparisons at different scales, and (4) evaluating the output differences before and after a period. Each calculation process can be represented by a tree, where both arithmetic calculations and aggregation operations are utilized for data lists.

The KPI ontology links data related to the building topology & geometry and sensor network, as well as collected data. This ontology aims at representing a KPI semantically and supports the retrieval of relevant data and automatic calculations. This ontology mainly consists of three components: input, output and process components, with an indicator class in the center.

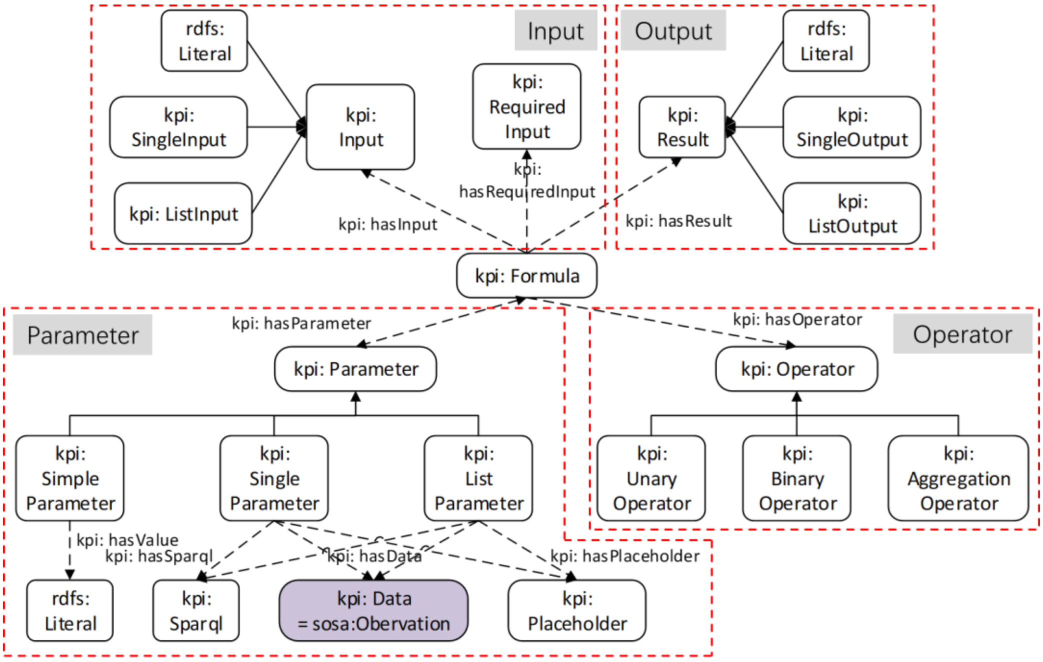

FIGURE 6. KPI ontology

The input component of the ontology describes the external parameter of the KPI. These inputs constrain the context of the KPI assessment, e.g., the space, building elements, MEP equipment, and time period. These data, once retrieved, can be treated as required inputs for the indicator.

The output component of the ontology describes the KPI results. After the calculation is performed, the results can be saved as linked data, making it possible to retrieve the results without reperforming the calculation.

The process component describes the calculation process of the KPI and is the core of the ontology. Each indicator can be calculated through a mathematical expression, which can be broken down into trees consisting of parameters and operators. The parameter type can generally be divided into 4 categories:

(1) a single numeric parameter;

(2) a single parameter that can be retrieved from the ontology;

(3) a list of parameters that can be retrieved from the ontology;

(4) a subexpression. The parameters that are hidden in the ontology are further linked with the class kpi:Data, which is equivalent to sosa:Observation. The kpi:SPARQL and kpi:Placeholder classes are linked to locate the exact instance.

The operators mainly consist of 3 types:

(1) unary operators, including -, sqrt, etc.;

(2) binary operators, including +, -, *, /, ^, max, min, etc.;

(3) aggregation operators, including sum, average, standard deviation, max, min, etc.

B. Data Preparation

On the basis of the ontologies established above, the required information can be represented as ontological instances for energy usage evaluation. Most of the static data, including basic building information, the topological and geometric properties of buildings and the positions of sensors, are included in BIM. Exemplified by Autodesk Revit, the information can be extracted through embedded functions using the Revit API. Once the building information model is created to the level of detail of the required information, data can be extracted and then converted into ontology instances. In this research, an open-source library called dotNetRDF developed in C# language is used to process this conversion. By using this library, the structured data can be read and converted to ontological instances on the basis of the predefined ontological structure, and an adRDF/XML file is generated as the output.

Dynamic data are those collected by sensors and are typically stored on a monitoring platform in a structured format. An interface is developed to read the .csv data generated, which is convert into ontological instances. The linkage between the static data and dynamic data is based on sensor identification. Each data set includes information about the sensor that made the observation, and for the static data, each sensor is included in the space, thus supporting analyses in both space and time.

C. Formula Representation

A KPI is often expressed as a mathematical formula. This index is linked to the formula and various inputs. The formula can be used to calculate the index values; for example, the total electricity use per day equals the difference in the electricity meter reading between the start and end of the day divided by total time in a day. Prior to extracting information and performing calculations, the relevant inputs, including those related to the time period and space of interest, must be specified. These nodes are related to the index by kpi:Input. The formula linked to the index consists of parameters and operators. A SPARQL query statement with placeholders is linked to extract the corresponding data, where a placeholder is a proxy that is replaced by external parameters. The aim of utilizing KPI is eliminating the difference in area or scale to facilitate the comparison between spaces or buildings and give a benchmark for building energy performance, and the KPI could also serve as normalized input for data mining to grasp the energy consumption pattern of different buildings in a district. There are several commonly used KPIs as listed in Table 1 to evaluate the overall energy consumption, energy consumption of specific usages, energy consumption of specific spaces and effectiveness of MEP systems and equipment.

TABLE 1. Examples of commonly used KPIs

Genre	KPI	time period
KPI for overall energy consumption	total electricity consumption per area/capita	per day/ per year
	total fuel consumption per area/capita	per day/ per year
	total water consumption per area/capita	per day/ per year
KPI for energy consumption of specific usages	electricity consumption for cooling per area	per day/ per year
	electricity consumption for heating per area	per day/ per year
	electricity consumption for lighting per area	per day/ per year
	electricity consumption for ventilator per area	per day/ per year
	electricity consumption for elevator per area/capita	per day/ per year
KPI for energy consumption of specific spaces	electricity consumption of public space per area	per day/ per year
	electricity consumption of rental space per area	per day/ per year
	electricity consumption of restaurant per area/capita	per day/ per year
	electricity consumption of guest rooms per available room	per day/ per year
	PUE (Power Usage Effectiveness) of information center	per day/ per year
KPI for MEP	electricity consumption of cooling station per area	per day/ for cool

systems and equipment	electricity consumption of air conditional terminal per area	supply season per day/ for cool supply season
	EER (Energy Efficiency Ratio) of water chiller	for cool supply season
	EER of cold/heat source equipment	for cool/heat supply season
	EER of cooling station equipment	for cool supply season

D. Indicator Calculation

The structure of the algorithm used to calculate a KPI is presented in Fig. 7, and the following steps are required.

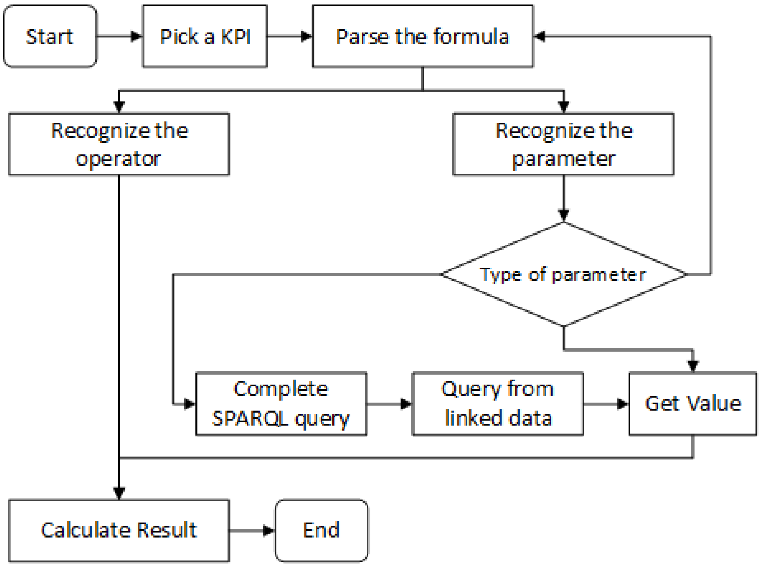

FIGURE 7. Algorithm to calculate KPI

The user first selects a KPI and specifies the required inputs, including the time period and spatial domain. The instances representing the inputs are extracted, and the formula is then parsed into an operator and parameters. If the parameter is a sub-formula, the parsing process is repeated. If the parameter is an explicit value, the value can be directly established. If the parameter needs to be retrieved from the ontology, a SPARQL query is first performed, and the placeholder is replaced with the corresponding inputs. Then, the query is performed to obtain the required values. The parameters and operator finally form a simple formula that is used to calculate the result.

V. CASE STUDY

A. Project Information

To validate the KPI calculation approach presented in section III and IV, we chose two office buildings both located in Shanghai, China as an application case study (Figure 8). One building is Xinzhuang Comprehensive Building (hereinafter building 1), which measures 22 meters in height, with 7 floors aboveground, and 1 floor underground, totaling an area of 9,992 square meters. The second building is Shanghai Jianke Building (hereinafter building 2), which measures 102 meters in height, with 24 floors aboveground, and 2 floors underground, totaling an area of 38,189 square meters. For these two buildings, sensors are installed on each floor to collect electricity consumption in total and by item, including lighting, socket, and air-conditioning over a 15-minute interval. These two buildings differ in the scales but they are both functioning as office buildings and are in the same city. Therefore, they are comparable only if the factor of building scale can be eliminated, so taking advantage of KPI could facilitate the comparison of building energy performance and could also give a benchmark for other buildings in the same area. Another difficulty to compare is that the sensors and sensing platforms of the buildings are from different companies, and the dynamic data is isolated from other data silos at present, making it difficult to interpret information hidden in the data and to perform comparisons among similar buildings. To facilitate energy data analysis, we propose the use of linked data to connect data silos and calculate KPIs for evaluating the energy consumption efficiency.

FIGURE 8. Photo of Xinzhuang Comprehensive Building and Shanghai Jianke Building

As stated in section III, energy KPI calculations require an understanding of the correlations among different data sets, specifically, building topology and geometry, sensor network and observation data sets, to form a group of linked data.

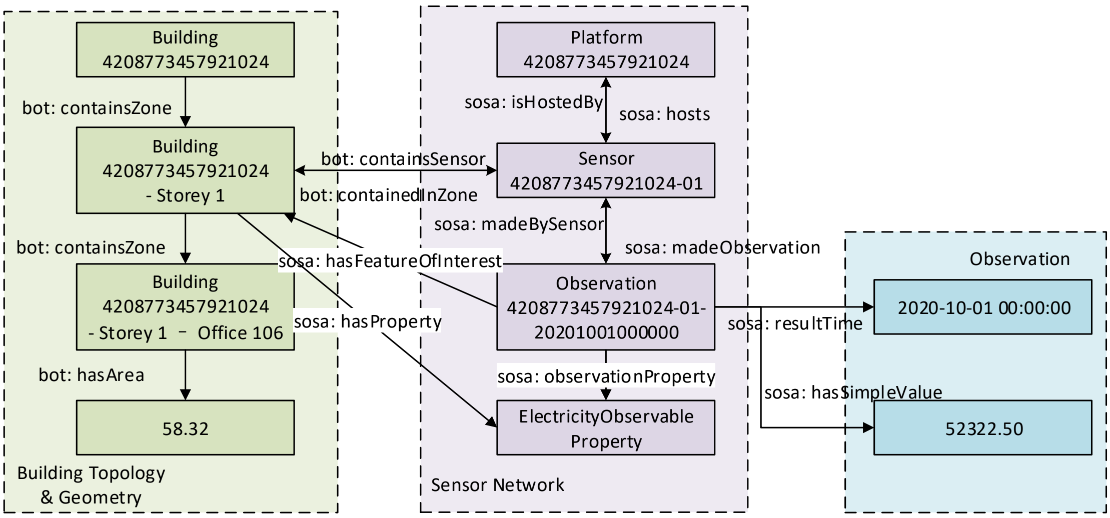

FIGURE 9. RDF graph of ontology instances of building 1

Building topology and geometry data are extracted from the BIM file of the project. The sensor network information is extracted from the energy monitoring platform and manually linked to the corresponding space of the building topology. Observation data, including electricity usage and temperature data from 2018 to 2020, are exported in CSV format. All the data above are read using the dotNetRDF library, converted into ontological instances and stored in RDF format. As an example, the RDF graph of ontological instances of building 1 is shown in Fig. 9.

B. Define a KPI Formula

The electricity consumption per area per day is a basic KPI used to evaluate the electricity usage intensity. This indicator eliminates the influence of area and thus can be used to compare different spaces inside a building as well as different buildings of a similar type. To calculate this KPI, the linkages between the building, sensor network and observations are necessary. The ontological instances for electricity consumption per area per day are shown in Fig. 10 and Fig. 11.

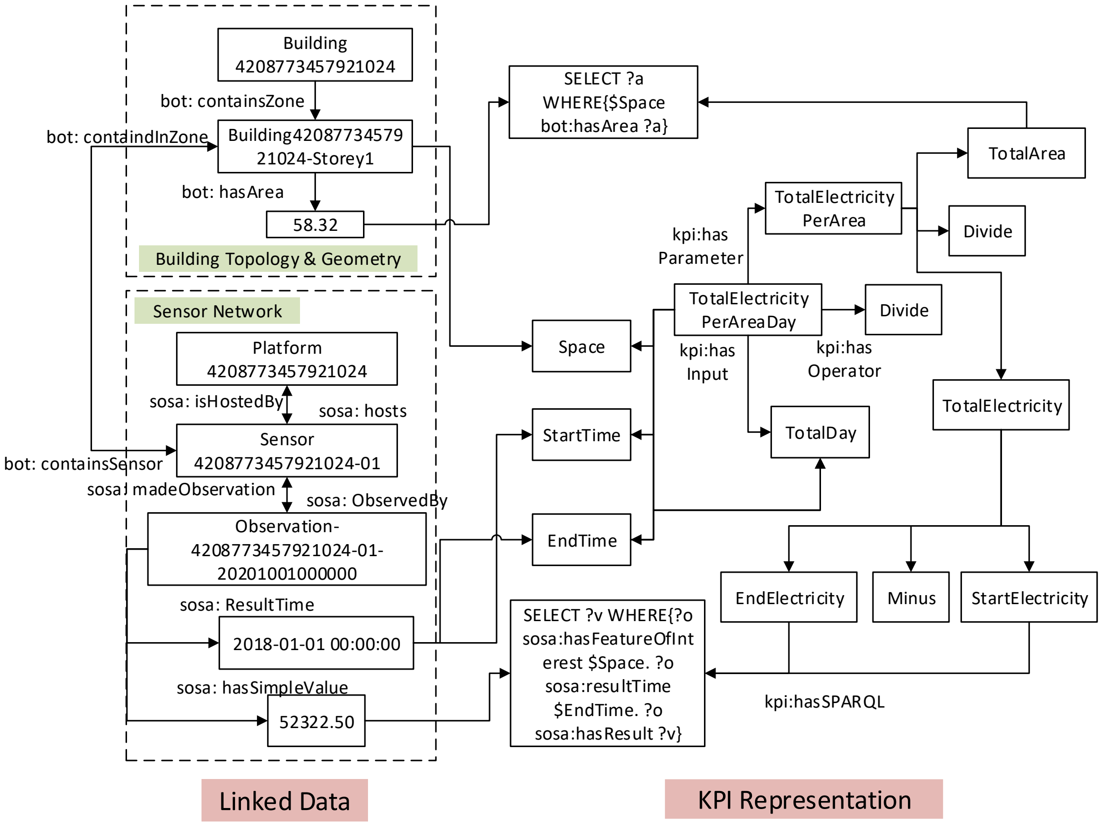

FIGURE 10. Representation for electricity consumption per area per day

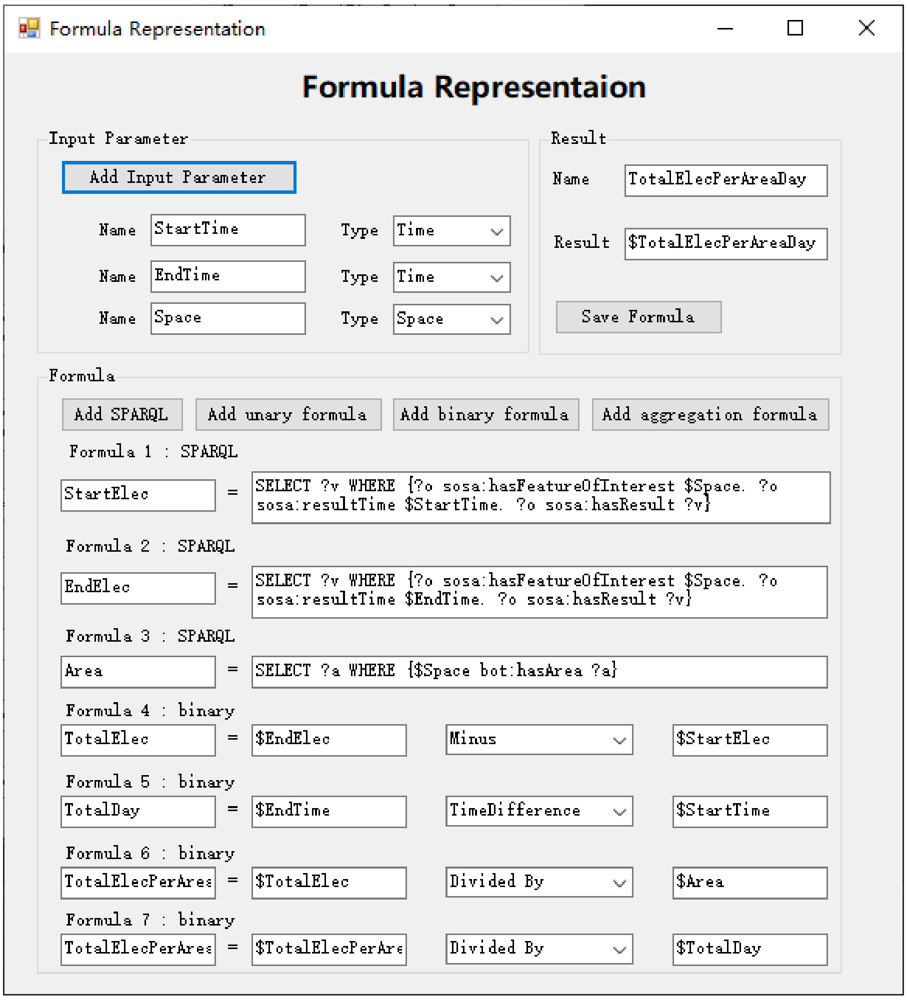

FIGURE 11. Representation for electricity consumption per area per day

C. KPI Calculation

This KPI also requires space, start time and end time inputs. The formula first involves the total electricity per area divided by time, where the time can be calculated through the time difference at the start and end of the selected period. The total electricity per area is the total electricity divided by total area. The total area can be queried through SPARQL with the input space as the placeholder. Total electricity is the difference in electricity meter observations between two endpoints. Each observation can be queried by SPARQL with space and time placeholders.

In this case, the electricity consumption per day per area of the two buildings in July 2020 are derived and shown in Fig 12. As can be seen in the graph, the electricity consumption of both buildings fluctuates on a weekly basis, with higher consumption on weekdays and lower on weekends, which agrees with the function of office buildings. The energy consumptions in both buildings show an upward trend, and the reason may be the consumption of air conditioning increases as the outdoor temperature rises to keep the indoor environment steady. In the comparison between the two building, building 1 earned green building label three stars and it outperforms in the overall energy efficiency according to the result of KPI.

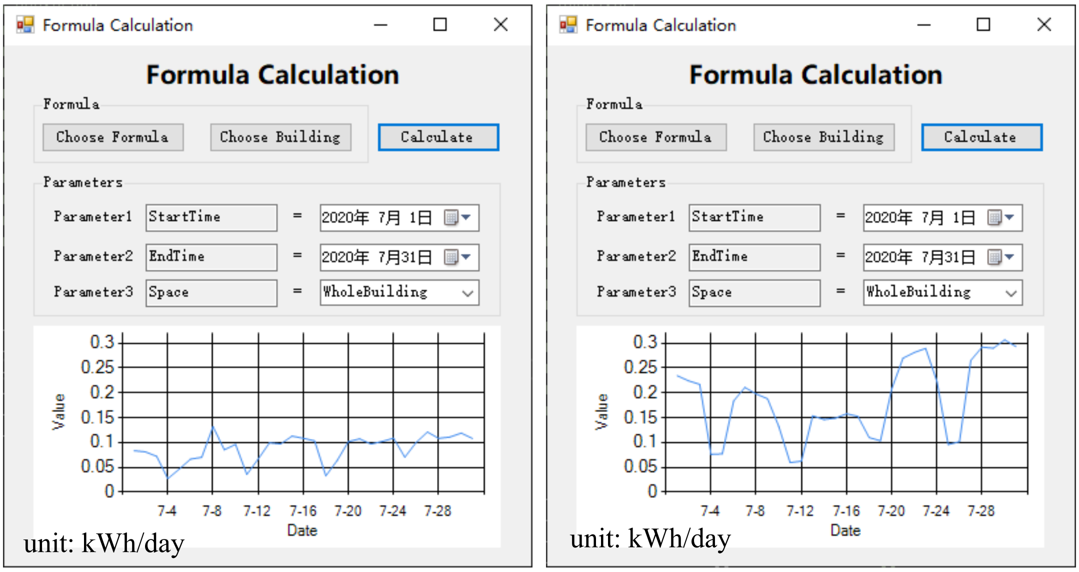

FIGURE 12. Comparison of electricity consumption per area per day of the two buildings

To achieve a further insight of this difference, we could further analysis the energy consumption for different usage, e.g., lighting and air conditioning. These buildings located in subtropics and air conditioning is the major consumption of electricity. Fig 13 shows the Comparison of electricity consumption for lighting (left) and air conditioning (right) per area per day of the two buildings. The electricity consumptions of both usages share the pattern of total consumption, however, air conditioning contributes larger so it should be the focus to save energy.

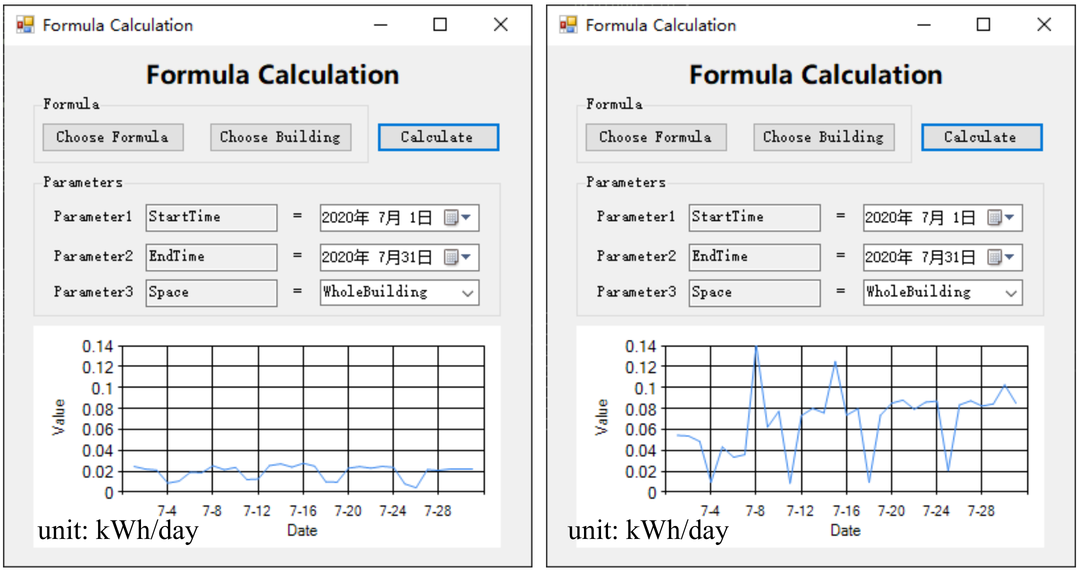

FIGURE 13. Comparison of electricity consumption for lighting (left) and air conditioning (right) per area per day of the two buildings

Besides the comparison between different time, building and item, the difference between different spaces inside a building can also be derived easily as the relationship between data silos are identified in the linked data and the KPI formula are the same regardless of the space. Manager may be interested if there is any space that are wasting electricity or the pattern disagrees with other spaces. Taking building 1 as an example, the total electricity consumption per area per day of different storeys and we found no significant different between these storeys, and user can easily browse any space in the building as described in the linked data.

D. Discussion

In the scenarios above, a KPI formula is demonstrated with real-life data. In this case, the formula can be represented according to the established ontologies using a program with a user interface, input parameters, and specific calculation process, which yields the relevant results. In terms of calculations, the user selects the space and time of interest to obtain information from the linked data, and the calculation process is automatically implemented. This process can also be iterated based on a list of spaces or times to compare the corresponding KPI values.

Traditional methods require queries of raw data from different data silos, and the relationships between the formula parameters and data must be manually determined in many cases. Moreover, manual assessment is required throughout much of the analysis process, making it difficult to automate or iterate. This research integrates data from different sources with related ontologies and links the data with the corresponding KPI formula to achieve automatic KPI definition and calculations. The link between the two ontologies are realized with SPARQL, when parameters in a formula needs to be modified, user can revise the query statement to amend the linkage, and this approach can also deal with difference in data structures of data silos.

Moreover, ontology-based approach also benefits to handle tree-structured data. Building spaces and MEP systems can both be expressed in the form of a tree. Sensors linked to the children of a node are all contribute to the parent sensor semantically. Ontology-based approach can handle the case with the help of the inference function, which is complicated to accomplish by querying into the database.

In this case study, we found existing ontologies could contribute to the scenario, however, each ontology requires extension according to specific requirement. For instance, SOSA ontology models observation for all data collected, nevertheless, in the process of building performance evaluation, specific indices have to be defined explicitly to achieve a comprehensive evaluation. Hence, in this research, classes in the existing ontologies are further extended for actual utilization. Similar method could be also applied to MEP system analysis, including making extension to MEP system entity and establish the linkage between MEP equipment with corresponding spaces thus supporting further evaluation and analysis.

It should be noted that this research focuses on validate the feasibility of the proposed method to automatize the process of retrieving data from different data silos and derive the result of a typical KPI, and only a few simple KPIs are taken as examples to illustrate the process. Even though the utilized KPI is a simple approximation of the performance, comparing the same KPI of two similar buildings could help us understand why a building performs better than the other one, and then building managers could make better decisions to improve building performance. However, it has to be clarified that an accurate assessment of building performance requires a comprehensive framework where a set of indicators are needed to fully describe the real energy performance of a building because the performance depends on multiple variables. And an assessment framework which could give an accurate and rigorous depiction of the energy performance is open for future research.

VI. CONCLUSION

Buildings consume a large proportion of global primary energy and building performance management requires massive data input. KPI is a common means to evaluate building performance, but the data silos of building information and energy consumption are separated and heterogeneous, and the linkage between KPI formula and data are not recognized. This paper develops an ontology-based approach to calculate key performance index automatically to support building energy evaluation among different information systems. A case study is conducted in which the approach is applied, and the KPIs for energy consumption and building environment is calculated in a real building project after selecting the space and time period of interest. The results validate the feasibility and effectiveness of the proposed approach.

This research attempts to shed light on involving formula into building linked data semantically and other researchers can use the approach as a step stone to further involving data from different sources to enhance building performance evaluation. Linking KPI formula into building linked data semantically benefits to make use of the advantage of ontology that inference can be accomplished automatically. Heterogeneous data can be processed in a standardized approach and relativity in tree-structured data can be understood.

Formula in this research can be further extended to more complex computation modules. For instance, existing simulation models require data from building linked data. In this case, formula in this research can be extended to simulation models, where input and result can be provided in a similar way, and the computation process will be accomplished by external programs. External factors such as climate conditions and human behaviors could be considered in more complex models, extending the utilization of KPIs. Thus it is promising to connect to data analytic tools by retrieving corresponding data input, supporting further data utilization.

Further work may also involve extending existing ontologies for specific scenarios. In this research, observation class is extended to specific subclasses of particular items. Building class is also extended to specific types of buildings. In further research, more extension can be made aiming at various scenarios. The framework of this approach also simulates researchers to contribute to the building related ontologies.

Buildings generate dynamic data continuously, therefore, the linked data could be too large to store in one single server. A distributed data storage approach might be able to solve the problem. Additionally, in the era of big data, the principle of moving computation closer to data can solve the pressure of data transmission. How to achieve moving code to data is also an open question to solve in the future.

KPI ontology is useful as it opens the way to define, calculate, and analyze different and valuable energy KPIs, however, smart metering systems and the current management tools of facility managers also offer opportunities and a more accurate way to help FM in their daily energy management in a certain building, while KPIs offer an opportunity to management buildings in a district as a whole. Future work also involves utilizing the result of the comparison KPIs and sharing energy saving approaches with other buildings in the same district, combining the energy management in building level and district level.

ACKNOWLEDGMENT

This work was funded by the National Key R&D Program of China (Grant No. 2017YFC0704200). Dr. Lin was also supported by the National Natural Science Foundation of

China (No. 51908323) and the Tsinghua University Initiative Scientific Research Program (No. 2019Z02UOT).

This work emerged from the IBPSA Project 1, an international project conducted under the umbrella of the International Building Performance Simulation Association (IBPSA). Project 1 will develop and demonstrate a BIM/GIS and Modelica Framework for building and community energy system design and operation.

REFERENCES

[1] A. Costa, M. M. Keane, J. I. Torrens et al., "Building operation and energy performance: Monitoring, analysis and optimisation toolkit" Appl Energ, vol. 101, pp.310-316, Jan. 2013, 10.1016/j.apenergy.2011.10.037.

[2] M.W. Ahmad , M. Mourshed , D. Mundow et al., "Building energy metering and environmental monitoring–a state-of-the-art review and directions for future research" Energy Builds, vol. 120, pp. 85–102, May 2016, 10.1016/j.enbuild.2016.03.059.

[3] S. Copiello, "Building energy efficiency: a research branch made of paradoxes" Renew Sust Energ Rev, vol. 69, pp. 1064–1076, March 2017, 10.1016/j.rser.2016.09.094.

[4] Z. Tong, Y. Chen, A. Malkawi et al., "Energy saving potential of natural ventilation in China: The impact of ambient air pollution" Appl Energ, vol. 179, pp. 660-668, Oct. 2016, 10.1016/j.apenergy.2016.07.019.

[5] J.R. Lin, J. Cao, J.P. Zhang et al., "Visualization of Indoor Thermal Environment on Mobile Devices based on Augmented Reality and Computational Fluid Dynamics" Automat Constr, vol. 103, pp. 26-40, March 2019, 10.1016/j.autcon.2019.02.007.

[6] K. Kang, J.R. Lin, J. Zhang, "BIM- and IoT-based monitoring framework for building performance management" Journal of Structural Integrity and Maintenance, vol. 3, no. 4, pp. 254-261, Nov 2018, 10.1080/24705314.2018.1536318.

[7] Y.W. Zhou, Z.Z. Hu, J.R. Lin et al., "A Review on 3D Spatial Data Analytics for Building Information Models" Arch Computat Methods Eng, vol. 27, pp. 1449–1463, Nov. 2020, 10.1007/s11831-019-09356-6.

[8] Y.Y. Zhang, K. Kang, J.R. Lin et al., "Building information modeling–based cyber-physical platform for building performance monitoring" Int J Distrib Sens N, vol. 16, no. 2, pp. 1-21, Jan. 2020, 10.1177/1550147720908170.

[9] E. Curry, J. O'Donnell, E. Corry, "Building optimisation using scenario modeling and linked data", presented at the 1st International Workshop on Linked Data in Architecture and Construction (LDAC), Ghent, BE, 2012.

[10] E. Corry, J. O'Donnell, E. Curry, et al., "Using semantic web technologies to access soft AEC data", Adv Eng Inform, vol. 28, no. 4, pp. 370-380, Oct. 2014, 10.1016/j.aei.2014.05.002.

[11] P. Pauwels, S. Zhang S, Y. C.Lee, "Semantic web technologies in AEC industry: A literature overview", Automat Constr, vol. 73, pp. 145-165, Jan. 2017, 10.1016/j.autcon.2016.10.003.

[12] T.R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?", Int J Hum Comput Stud, vol. 43, pp. 907–928, Nov. 1995, 10.1006/ijhc.1995.1081.

[13] D. Richards, S.J. Simoff, "Design ontology in context - a situated cognition approach to conceptual modelling", Artif Intell Eng, vol. 15, no. 2, pp. 121–136, April 2001, 10.1016/s0954-1810(01)00010-3.

[14] R. Sacks, C. Eastman, G. Lee G et al., "BIM handbook: a guide to building information

modeling for owners, designers, engineers, contractors, and facility managers”. Hoboken:John Wiley & Sons, 2018.

[15] J. Zhang, Q. Liu, Z. Hu et al., “A multi-server information-sharing environment for cross-party collaboration on a private cloud” Automat Constr, vol. 81, pp. 180-195, Sept. 2017, 10.1016/j.autcon.2017.06.021.

[16] J.R. Lin, Z.Z. Hu, J.P. Zhang et al., “A natural-language-based approach to intelligent data retrieval and representation for cloud BIM” Comput Aided Civ Infrastruc. Eng, vol. 31, no. 1, pp. 18-33, Jan. 2016, 10.1111/mice.12151.

[17] B. Zhong, C. Gan , H Luo et al., “Ontology-based framework for building environmental monitoring and compliance checking under BIM environment”, Build Environ, vol. 141, pp. 127-142, August 2018, 10.1016/j.buildenv.2018.05.046.

[18] M. Arslan, Z. Riaz, A.K. Kiani et al., “Real-time environmental monitoring, visualization and notification system for construction H&S management”, Electron J Inf Technol Construct, vol. 19, no. 11, pp. 72-91, 2014.

[19] Z. Riaz, D.J. Edwards, E.A. Parn, et al., “BIM and sensor-based data management system for construction safety monitoring”, J Eng Des Technol, vol. 15, no. 6, pp. 738-753, 2017, 10.1108/jedt-03-2017-0017.

[20] W. Natephra, A. Motamedi, N. Yabuki et al., “Integrating 4D thermal information with BIM for building envelope thermal performance analysis and thermal comfort evaluation in naturally ventilated environments”, Build Environ, vol. 124, pp. 194–208, 2017, 10.1016/j.buildenv.2017.08.004.

[21] K. Suprabhas, H.N. Dib, “Integration of BIM and utility sensor data for facilities management”, in Computing in Civil Engineering 2017, Seattle, WA, USA, 2017, pp. 26–33.

[22] W Terkaj, A Šojić , “Ontology-based representation of ifc express rules: an enhancement of the ifcowl ontology”, Automat Constr, vol. 57, pp. 188–201, Sept. 2015, 10.1016/j.autcon.2015.04.010.

[23] P. Pauwels, IFC-to-RDF Conversion Service, (2015) . [Online]. Available: https://github.com/mmlab/IFC- to- RDF- converter.

[24] H. Issa, L. van Elst, A. Dengel, “Using smartphones for prototyping semantic sensor analysis systems”, in Proceedings of the International Workshop on Semantic Big Data, San Francisco, CA, 2016, pp. 7.

[25] F. Radulovic , M. Poveda-Villalón , D. Vila-Suero et al., “Guidelines for linked data generation and publication: an example in building energy consumption”, Automat Constr, vol. 57, pp. 178–187, Sept. 2015, 10.1016/j.autcon.2015.04.002.

[26] W3C Community Group, Building Topology Ontology, (2019). [Online]. Available: https://w3c-lbd-cg.github.io/bot.

[27] P. Pauwels, E. Corry, J. O’Donnell, Representing SimModel in the web ontology language, in: R. Issa, I. Flood (Eds.), Computing in Civil and Building Engineering, Orlando, Florida, USA, 2014, pp. 2271–2278. http://dx.doi.org/10.1061/9780784413616.282.

[28] P. Pauwels, E. Corry, J. O’Donnell, “Making SimModel information available as RDF graphs”, in Proceedings of the 10th European Conference on Product and Process Modelling (ECPPM), Vienna, Austria, 2014, pp. 439–445.

[29] K.B. Sørensen, P. Christiansson, K. Svidt, “Ontologies to support RFID-based link between virtual models and construction components”, Comput Aided Civ Infrastruc Eng, vol. 25, pp. 285–302, March 2010, 10.1111/j.1467-8667.2009.00638.x.

[30] E. Corry, P. Pauwels, S. Hu et al., “A performance assessment ontology for the environmental and energy management of buildings”, Automat Constr, vol. 57, pp. 249–259, Sept. 2015, 10.1016/j.autcon.2015.05.002.

[31] E.W. Patton, P. Seyed, P. Wang et al., “SemantEco: a semantically powered modular architecture for integrating distributed environmental and ecological data”, Future Generat Comput Syst, vol. 36, no. 3, pp. 430–440, July 2014, 10.1016/j.future.2013.09.017.

[32] C. Métral, G. Falquet, K. Karatzas, “Ontologies for the integration of air quality models and 3D city models”, in Proceedings of 2nd Workshop of COST Action C21- towntology, 2007, pp. 18–33.

[33] M.M. Oprea, “AIR_POLLUTION_Onto: an ontology for air pollution analysis and control”, in IFIP International Federation for Information Processing, Boston, MA, US, 2009, pp.135-143

[34] J. Adeleke, D. Moodley, “An ontology for proactive indoor environmental quality monitoring and control”, in Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists. Stellenbosch, South Africa, 2015, pp. 2.

[35] M. Stocker, M. Rönkkö, M. Kolehmainen, “Towards an ontology for situation assessment in environmental monitoring”, in International Congress on Environmental Modelling and Software, San Diego, CA, US, 2014, pp. 1280-1288.

[36] H. Pundt, Y. Bishr, “Domain ontologies for data sharing-an example from environmental monitoring using field GIS”, Comput. Geosci, vol. 28, no. 1, pp. 95-102, Feb. 2002, 10.1016/S0098-3004(01)00018-8.

[37] M. Dibley, H. Li, Y. Rezgui et al., “An ontology framework for intelligent sensor-based building monitoring”, Autom. Constr, vol. 28, pp. 4-14, Dec. 2012, 10.1016/j.autcon.2012.05.018.

[38] B. Balaji, et al., Brick: Towards a Unified Metadata Schema For Buildings. Conference: the 3rd ACM International Conference, 2016.

[39] A. Mahdavi, M. Taheri, “An ontology for building monitoring” J Build Perform Simu, vol. 10, no. 5-6, pp. 499-508, Oct. 2016, 10.1080/19401493.2016.1243730.

[40] A. Mahdavi, M. Taheri, “A building performance indicator ontology” In Book: eWork and eBusiness in Architecture, Engineering and Construction, 2018. 6 Pages. eBook ISBN9780429506215

[41] M. Taheri, A. Mahdavi, “Structured representation of monitored occupancy data” Bauphysik, vol. 40, no. 6, pp. 434-440, Dec. 2018, 10.1002/bapi.201800024.

[42] Y.Li, R. García-Castro, N. Mihindukulasooriya et al., “Enhancing energy management at district and building levels via an EM-KPI ontology” Automat Constr, vol. 99, pp. 152-167, March 2019, 10.1016/j.autcon.2018.12.010.

[43] Y. Peng, J.R. Lin, J.P. Zhang et al., “A hybrid data mining approach on BIM-based building operation and maintenance, Build Environ, vol. 126, pp. 483-495, Dec. 2017, 10.1016/j.buildenv.2017.09.030.

[44] E. Curry, J. O’Donnell, E. Corry et al., “Linking building data in the cloud: integrating cross-domain building data using linked data”, Adv Eng Inform, vol. 27, no. 2, pp. 206–219, April 2013, 10.1016/j.aei.2012.10.003.

[45] O’D. James, E. Corry, S. Hasan et al., “Building performance optimization using cross-domain scenario modeling, linked data, and complex event processing”, Build Environ, vol. 62, pp. 102-111, April 2013, 10.1016/j.buildenv.2013.01.019.

[46] E. Corry, P. Pauwels, S. Hu, et al., “A performance assessment ontology for the environmental

and energy management of buildings", Autom. Constr, vol. 57, pp. 249–259, Setp. 2015, 10.1016/j. Autcon.2015.05.002.

[47] S. Hu, E. Corry, H. Horrigan et al., "Building performance evaluation using OpenMath and Linked Data." Energ Build, vol. 174, pp. 484-494, Sept. 2018, 10.1016/j.enbuild.2018.07.007.

[48] N.M. Tomasevic, M.C. Batic, L.M. Blanes, et al., "Ontology-based facility data model for energy management", Adv Eng Inform, vol. 29, no. 4, pp. 971-984, Oct. 2015, 10.1016/j.aei.2015.09.003.

[49] M. Dibley, H. Li, J. Miles et al., "Towards intelligent agent based software for building related decision support", Adv. Eng. Inform, vol. 25, no. 2, pp. 311-329, April 2011, 10.1016/j.aei.2010.11.002.

[50] D. Russomanno, C. Kothari, O. Thomas, "Building a sensor ontology: a practical approach leveraging ISO and open geospatial consortium (OGC) models", in Proceedings of the International Conference on Artificial Intelligence, Las Vegas, NV, US, 2005, pp. 637–643.

[51] I. Niles, A. Pease, "Towards a standard upper ontology", in Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS), Ogunquit, MA, US, 2001, pp. 2–9.

[52] J.R. Lin, Y.C. Zhou "Semantic classification and hash code accelerated detection of design changes in BIM models" Automat Constr, vol. 115, pp. 103212, July 2020, 10.1016/j.autcon.2020.103212.