Title	Building performance evaluation using OpenMath and Linked Data
Authors(s)	Hu, Shushan, Corry, Edward, Horrigan, Matthew, Hoare, Cathal, Dos Reis, Mathilde, O'Donnell, James
Publication date	2018-09-01
Publication information	Hu, Shushan, Edward Corry, Matthew Horrigan, Cathal Hoare, Mathilde Dos Reis, and James O’Donnell. “Building Performance Evaluation Using OpenMath and Linked Data.” Elsevier, September 1, 2018. https://doi.org/10.1016/j.enbuild.2018.07.007.
Publisher	Elsevier
Item record/more information	http://hdl.handle.net/10197/10999
Publisher's statement	This is the author’s version of a work that was accepted for publication in Energy and Buildings. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Energy and Buildings (174, (2018)) https://doi.org/10.1016/j.enbuild.2018.07.007
Publisher's version (DOI)	10.1016/j.enbuild.2018.07.007

Building performance evaluation using OpenMath and Linked Data

Shushan Hu[a,b], Edward Corry[b], Matthew Horrigan[b], Cathal Hoare[b],
Mathilde Dos Reis[c], James O'Donnell[b,]

[a] *School of Computer Science and Information Engineering, Hubei University, Wuhan, China*
[b] *School of Mechanical and Materials Engineering and UCD Energy Institute, University College Dublin, Belfield, Dublin 4, Ireland*
[c] *Graduate School of Engineering of the University of Nantes, Nantes, France*

Abstract

A pronounced gap often exists between expected and actual building performance. The multi-faceted and cross lifecycle causes of this performance gap are found in design assumptions, construction issues and commissioning and operational compromises. Some important factors are firmly rooted in the lack of interoperability around building information. New solutions to the interoperability challenge offer the potential to leverage and reuse available heterogeneous data in a manner that can significantly assist building performance assessment. Linked data provides an open, modular and extensible solution for the challenge. However, in the buildings domain, the integration of rule-based performance metrics and contextual information has yet to be formally established.

This paper describes an approach to the provision of in-depth building performance assessment through the integration of OpenMath and linked data. An ontology describing performance metrics in RDF is presented, together with an automated metric evaluation solution using multi-silo queries and computer algebra systems, providing a flexible, automated and extensible mechanism for the assessment of building performance. Building managers and engineers can simultaneously analyse time-series building performance

---

Email addresses: `shushan.hu@ucd.ie` (Shushan Hu), `edward.corry@ucd.ie` (Edward Corry), `matthew.horrigan@ucdconnect.ie` (Matthew Horrigan), `cathalhoare@gmail.com` (Cathal Hoare), `mathilde.dos-reis@etu.univ-nantes.fr` (Mathilde Dos Reis), `james.odonnell@ucd.ie` (James O'Donnell)

Preprint submitted to Energy and Buildings *February 5, 2018*

at a range of levels, without burdensome manual intervention such as is the case with traditional solutions. A test implementation on a large university building highlights the potential of this solution.

Keywords: Building performance assessment, data interoperability, Linked Data, OpenMath, performance metric

1. Introduction

Buildings consume a globally significant quantity of energy [1]. European buildings account for 41% of primary energy (27% for residential buildings and 14% for commercial buildings), 22% and 19% in United States respectively [2]. Although buildings consume such a large amount of energy, there is often a discrepancy between a building's actual energy consumption and that predicted at design time [3]. The many factors which feed into this performance gap might be loosely categorised as design, construction or operational issues [4].

One contributory issue, the role of performance data in building management, has been the subject of much recent study. A number of researchers have considered the role of building performance frameworks that decompose performance assessment into a series of measurable components [5, 6, 7]. A key problem associated with the effective implementation of such performance frameworks is the availablity of performance data and the integration of data from disconnected data silos [8, 9]. For example, Building Energy Management Systems (BEMS) typically assess real-time building performance and control building conditions based on advanced communication protocols between various sensing and control nodes [10]. Yet, there is usually little interaction between islands of information such as Building Information Models (BIM) with energy performance predictions for the building and Building Energy Management Systems (BEMS) [11]. Recently, the semantic web has been used to enhance data interaction between BEMS and the other information islands [12, 13].

This paper addresses some of the computer science issues surrounding the integration and interpretation of building performance data. There are three main computer science approaches which could be used: 1) the common data model approach, 2) the adapter approach and 3) the linked data approach.

The common data model approach tries to design a central underlying data model as a data hub with agreement between tool vendors. The data hub is then interpreted by domain-specific applications [14]. It is hard to design a highly flexible and scalable data model and to maintain the model after being extended to a huge size. Rather than attempting to convert all data to a particular data model, the adapter approach leverages adapters to integrate cross-domain data silos, representing different Architecture, Engineering & Construction (AEC) domains [15]. However, this approach can lead to significant data loss and it is not feasible to develop adapters for all data silos [16].

The third approach to the interoperability issue is linked data, which uses existing open protocols and the semantic web for sharing unstructured data [17]. The usage of such linked data technologies aims to overcome the differences among software used in diverse disciplines and to connect various domains of information that have opportunities to identify untapped valuable resources within AEC domains [18]. Linked data has been used on data sources including BIM [19] and sensor information [20].

In practice, quantitative performance metrics are key components for building performance assessment. However, to date researchers and industry have yet to take advantage of multi-silo rule-based performance metrics as such metrics are currently isolated from other available contextual data silos [21]. Therefore, this isolation restricts in-depth building performance assessment.

This paper presents a flexible, automated and extensible solution to the isolation issue for the building performance domain. The approach offers a higher flexibility and a finer spectral granularity for performance assessment purposes. Building managers and engineers can simultaneously analyse time-series building performance at a range of levels, without burdensome manual intervention such as is the case with traditional solutions.

The objective of this paper is to show how insightful building performance assessment can be obtained through a novel integration of RDF and OpenMath. RDF provides an open, modular, extensible schema for data sharing in the AEC domain [22]. OpenMath makes the performance formulae be readable for computers, then calculation of formulae can be automated based on computer algebra systems [23].

Based on previous work [24, 25, 26, 27], this research defines an RDF ontology to present performance metrics and link them to available contextual information in linked dataspaces. Two distinct yet interrelated components are designed to evaluate such metrics: the first is a querying algorithm that extracts datum streams for metrics and the second is a computation algorithm that leverages computer algebra systems to calculate metrics.

The paper is structured as follows: the case for building performance assessment with linked data is made in Section 2. Section 3 describes the integration of performance metrics, the RDF schema and OpenMath. Two algorithms are presented in Section 4 to query datum streams for performance formulae and evaluate performance metrics. Section 5 provides a conceptual implementation of this work, while conclusions are provided in Section 6.

2. Assessing building performance with linked data

Buildings now produce more performance data (e.g. BIM, simulation, occupancy pattern) than ever [28]. Yet, poor interoperability between the many diverse data streams found in modern buildings contributes to inefficient building operation. Interpreting and understanding building data in a cross-domain manner may provide greater opportunity for performance optimisation measures.

Linked data is one possible approach to address the interoperability problem. It provides a flexible, open and modular pathway for exposing, sharing and connecting data in the building domain. For example, Curry et al. [11] proposed a linked dataspace for cross-domain data silos, enabling a shift from the fragmented nature of traditional building performance analysis approaches.

2.1. Improving control strategy

The holistic management and maintenance of building energy can be enhanced through the sharing of different and varied building information. For example, Grzybek et al. [29] has proposed the conceptualization of energy problems through the use of BIM information and semantic web, a linked data technology. The process involved in creating semantic rules to describe the links between these data sets is quite involved and requires significant domain expertise.

In order to reduce the need for such intervention, artificial intelligence algorithms are used to extract rules from historical or simulated building performance data. Yuce et al. [30] presented an Artificial Neural Network/-Genetic Algorithm algorithm, which can identify scheduling rules for the optimised energy management of appliances in the domestic sector. Yuce et al. [31] also proposed a novel ontology for building performance optimisation rules that present mappings between boundaries of environmental variables and control strategies.

These efforts have emerged from advances in communication protocols between the Building Management System (BMS) and the various sensing and control nodes in a building. Building information based on these protocols normally has specific formats.

Modern buildings might have significant amounts of data domains available to the building operator. There is usually little interaction between islands of information, such as a BIM and energy performance predictions for the building, and BMS.

2.2. Improving performance assessment

Augenbroe advocates a rigorous use of building performance indicators to ensure compliance between project specification and performance [32]. Due to the fragmented nature of the AEC industry, many domain specific data models exist, which hampers the rigorous use of performance indicators.

The performance framework using the scenario modelling method [33] followed on from previous performance metric/indicator work by Hitchcock and Augenbroe [34, 32]. O'Donnell et al. [24] presented that the combination of linked data, scenario modelling and complex event processing can provide enhanced performance assessments to building managers. Corry et al. [26] addressed some of the limiting factors of this framework through the formalisation of a performance assessment ontology for buildings. The value of the performance indicator approach is enhanced through recent research developments [6, 35, 36].

Each of these efforts identified how the successful implementation of a performance indicator approach is dependent on access to reliable building information.

2.3. Building information ontology

There are a number of available ontologies that are aimed at sharing and connecting cross-domain data in the building domain [18]. For example, the ifcOWL ontology is defined as an OWL representation of IFC data and serves as an alternative representation of the EXPRESS schema of IFC [37]. Such a file based IFC-to-RDF conversion application has been developed [19]. The Semantic Sensor Network (SSN) ontology is accomplished based on the concept of a stimulus prompting an observation [20]. The SSN enables expressive representation of sensors, their observations, and knowledge of their environment [38].

The Simulation domain Model (SimModel) is developed as an interoperable object-oriented data model which defines all object/attribute/relationship sets used for building performance simulation [39]. This model is made available in an OWL ontology [40]. A new ontology is developed to define and quantify energy-related occupant behaviour in buildings [41]. It improves interoperability between occupant behavior models and building energy modeling programs [42].

Corry et al. [43] proposed an ontology to aid the integration of soft AEC data into the existing linked dataspace for the evaluation of building performance. A hybrid structure was designed to link time-series related building performance and contextual data silos in linked dataspaces [27].

Dibley et al. [44] proposed an ontology framework, which enables real-time building performance monitoring based on semantic building data. The building data is transformed using a building ontology relying on IFC and the Ontosensor ontology. Tomasevic et al. [45] use the common semantics offered by the ontology to improve the energy management of complex infrastructures. A facility ontology was defined to model the static knowledge related to significant energy consumers of target infrastructures.

Abanda et al. [46] developed an OWL ontology (PV-TONS) for photovoltaic (PV) devices in buildings. The ontology was used to construct a system to size and select PV-system components for different types of buildings. A ThinkHome ontology was developed for the representation of home facilities and their energy demand or supply [47]. An agent-based system was developed to autonomously control the smart home in an energy-efficient and comfort-oriented manner.

2.4. Linking performance metric

It is logical to assume that far greater use can be made of linked dataspaces, which integrate cross-domain building information with ontologies. Relationship properties can provide a strong informational backbone to data analysis efforts throughout the operational phase.

While linked dataspaces have been constructed for building performance assessment and optimisation, performance metrics are still isolated from contextual building information in linked dataspaces (Figure 1).

More specifically, an ontology should be defined to represent performance metrics, especially mathematical formulae within them, and link them with other contextual information. In order to reduce burdensome manual intervention such as is the case with hard coded computation approaches, algorithms need to be developed to automatically calculate performance metrics, using computer algebra systems and available linked building information.

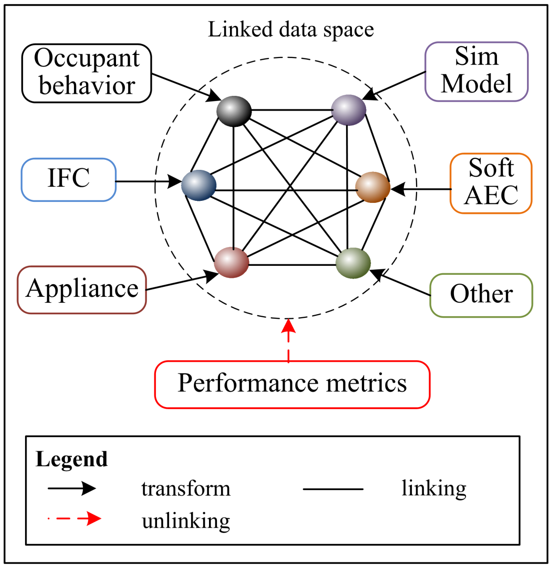


Figure 1: The structure of linked dataspaces addressing the data interoperability issue for building performance assessment.

The remainder of this paper will propose and describe the use of a novel combination of OpenMath/RDF and computer algebra systems for the expression and evaluation of performance metrics.

3. Combining OpenMath and linked data.

OpenMath [23] is an extensible standard for representing mathematical objects through their semantic descriptions. OpenMath exposes the following key tenets:

- *Content Dictionaries* are used to organise and define mathematical symbols [48].
- *OpenMath Objects* are a recursive expression that describes a mathematical structure - expressed in XML - that can be used to both transport and reason about a mathematical expression [49].

These features provides a mechanism through which mathematical expressions can be exposed to computer algebra systems without loss of information [50]. Kohlhase et al. [51] and Lange et al. [52] have used the OpenMath standard in combination with the RDF schema to include mathematical calculations in reasoning systems. It is proposed that OpenMath, in combination with Linked Data, provides a suitable technological platform to enable the linking and evaluating of performance metrics.

Building on previous research [24, 26, 27], a novel approach for deriving insightful building performance assessments is presented. The approach is organised around three key concepts (illustrated in Figure 2):

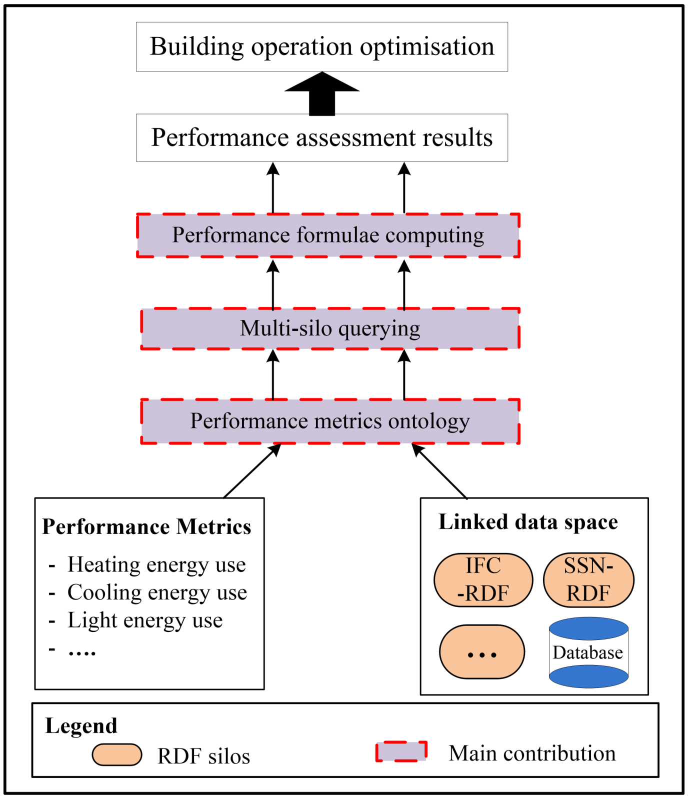


Figure 2: The proposed approach enables multi-silo queries and subsequent evaluation of performance metrics in buildings.

1. A performance metrics ontology - expressed as a combination of OpenMath and RDF schema - which describes performance metrics in an RDF format (Section 3.1).

2. A novel multi-silo query algorithm that uses a combination of SPARQL and SQL to provide datum streams for performance metrics (Section 4.1).
3. An algorithm that is based on computer algebra systems, used to evaluate metrics which are expressed using the performance metrics ontology. (Section 4.2).

The approach described here links cross domain data silos in a flexible and dynamic manner; time-series performance data is decoupled from the calculation process and, as result, can be reused outside of the building management system domain. This flexibility adds value to the data as it can be reused across an organisations processes to provide insightful perspectives on building operation. Furthermore, this system facilities more informed decision making by building managers; previously, such insights would have been unavailable since the evaluation of such fine grained results would often be beyond the expertise and access of typical building managers [7].

3.1. OpenMath derived Performance Metrics Ontology

The chosen approach links categorical contextual information, available in linked data spaces, with performance metrics by defining an OpenMath based [53] RDF ontology that describes standardised performance metrics as mathematical expressions. The ontology consists of two primary components, a performance assessment definition and a performance formulae definition; these are illustrated in Figure 3, where the performance assessment definition is presented in the left panel, while the performance formulae definition is shown in the right panel.

The performance assessment category represents the context of the analysis being undertaken. This context is expressed as a building object; these can include, amongst others, whole buildings, floors, zones, HVAC systems. Each context can have a number of associated qualitative performance objectives that are evaluated through some quantitative performance metric. When these metrics are calculated, they are qualified through additional metadata such as, for example, period of analysis (expressed as the difference of the periods start and end times). The resulting information set is used as input to the performance formulae definition.

The performance formulae definition creates a mathematical representation of the performance metric. It contains both arguments and operators.

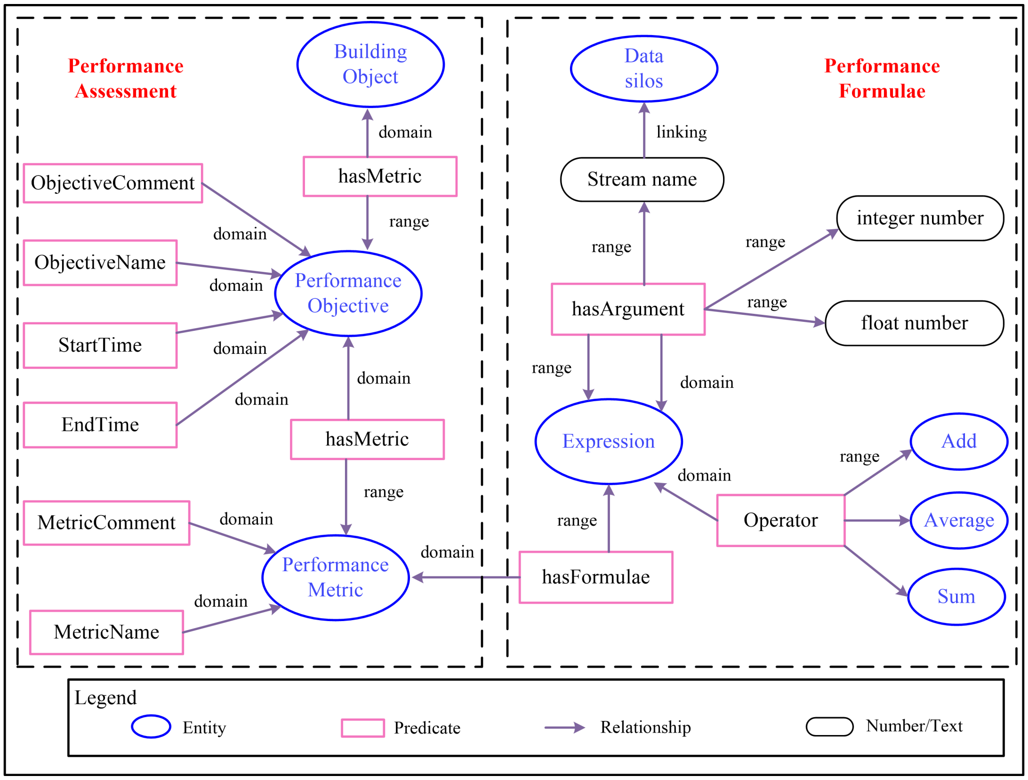


Figure 3: The ontology structure for performance metrics based RDF schema and OpenMath.

Three types of arguments and two types of operator are defined. The argument types include:

1. Sub-expressions that are derived as the expression is decomposed.
2. Real numbers which represent numeric information in both integer and floating point representations.
3. Datum Streams which are expressions of informational concepts, for example temperature and relative humidity found in AEC data silos.

As mentioned a priori, operators are categorised into two forms, regular arithmetic and special. The former category includes the normal operators (addition, subtraction, multiplication and division). The later category includes derivatives of the sum and average operators that are peculiar to building performance calculations.

Rather than express lower and upper bounds of summation, the sum operator operates over the child-parent relationships expressed within linked semantic data; for instance, the sum of the heating energy consumption of a floor is the sum of the consumption for all rooms on the floor. The operation is completed thus; the list of spaces on a floor is identified, and heating consumption for each of these is calculated before being summed to provide a result for the floor. Similarly, the average operator could be applied to calculate the average heating energy consumption for a floor by first applying the sum operator for the floor before dividing this result by the number of spaces on the floor.

Having described a new ontology based on OpenMath and the RDF schema, an automated solution to evaluate performance metrics using multi-silo queries and computer algebra systems is now introduced.

4. Evaluating performance metrics

We now describe how semantic representation of the performance metrics are evaluated. The approach requires two phases:

1. A multi-silo querying algorithm prepares data streams extracted from the linked data spaces to be used as arguments in performance formulae.
2. A performance formulae computation algorithm represents performance metrics as mathematical expressions and, taking the aforementioned data streams as inputs, calculates their value using computer algebra systems.

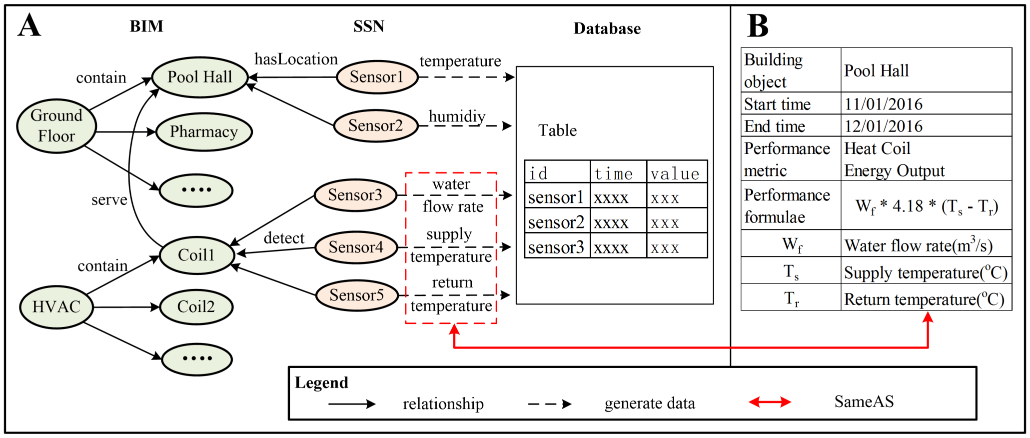
These two phases are now examined in detail.

4.1. Multi-silo Querying for Performance Formulae

The multi-silo querying algorithm prepares datum streams as arguments to performance assessment formulae; implementing this step automates discovery of robust streams.

An example use case is represented in Figure 4A and this is used to explain the operation of the algorithm. The example takes a performance metric - heating coil energy output - and uses it to calculate the energy efficiency of a pool hall in a sports centre. The metric consumes three types of sensing data; W_f represents the water flow rate, T_s the supply temperature and T_r which represents the return temperature.

The algorithm is used to mine the linked data space - represented in Figure 4B - to link building objects to time-series data. This is achieved by mining relationships between building objects in the BIM silo to identify relevant entities, for example, the building storey containing the pool hall, and the HVAC elements serving that space. Sensors associated with each relevant building object are then identified; this permits linking to relevant time-series data retained in relational database tables. Having identified relevant tables, queries are generated to retrieve this data.



id	time	value
sensor1	xxxx	xxx
sensor2	xxxx	xxx
sensor3	xxxx	xxx

Building object	Pool Hall
Start time	11/01/2016
End time	12/01/2016
Performance metric	Heat Coil Energy Output
Performance formulae	$W_f * 4.18 * (T_s - T_r)$
W_f	Water flow rate(m^3/s)
T_s	Supply temperature(°C)
T_r	Return temperature(°C)

Figure 4: A) An Example of possible relationships between data silos and B) A performance metric for a university Pool Hall .

The structure of the algorithm is represented by Figure 5. It operates thus:

A) The time period and building object associated with the performance metric are extracted (step 1). The time period is retained as a parameter to constrain the data streams. The performance formulae is also extracted and parsed to identify its arguments (step 2); these arguments identity what data streams must be prepared as input to the performance formulae.

B) The building object, extracted in step 1, serves - as described a priori - as the starting point for mining the linked space to discover sensors whose data can serve as data streams (steps 3 and 4). These queries are expressed as SPARQL and operate on an RDF representation of

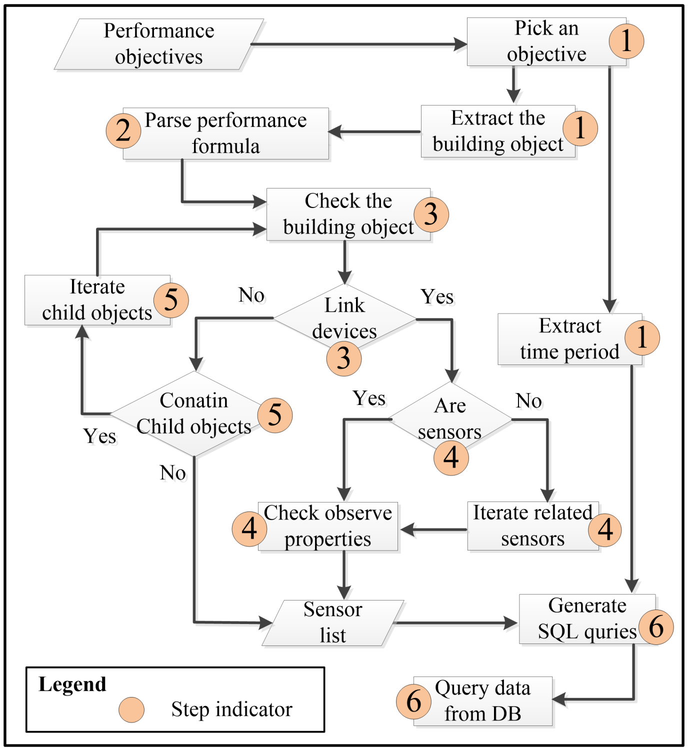


Figure 5: Workflow of the first algorithm for generation and population of data streams.

relationships contained in the BIM information, and between BIM information and sensor information in SSN.

C) If, at this point in the algorithms execution, any data stream remains unfulfilled, steps 3 and 4 are repeated for entities linked to the building object (step 5).

D) Once data sources for all data streams have been identified (or the search for suitable sources has been exhausted), SQL queries are generated (constrained by the time periods extracted in step 1) to populate the streams.

Once suitable data streams have been created, they are used as input to the evaluation of the performance expression. This process is described next.

4.2. Performance Formulae Computation

As described in Section 3, performance formulae are transformed into mathematical representations. Where present, special sum and average operators must be calculated as part of the transformation process; as mentioned earlier, these operate over child-parent relationships expressed within linked semantic data. A second algorithm - illustrated in Figure 6 - handles this transformation and evaluates the result using a computer algebra system to obtain a value for the performance metric.

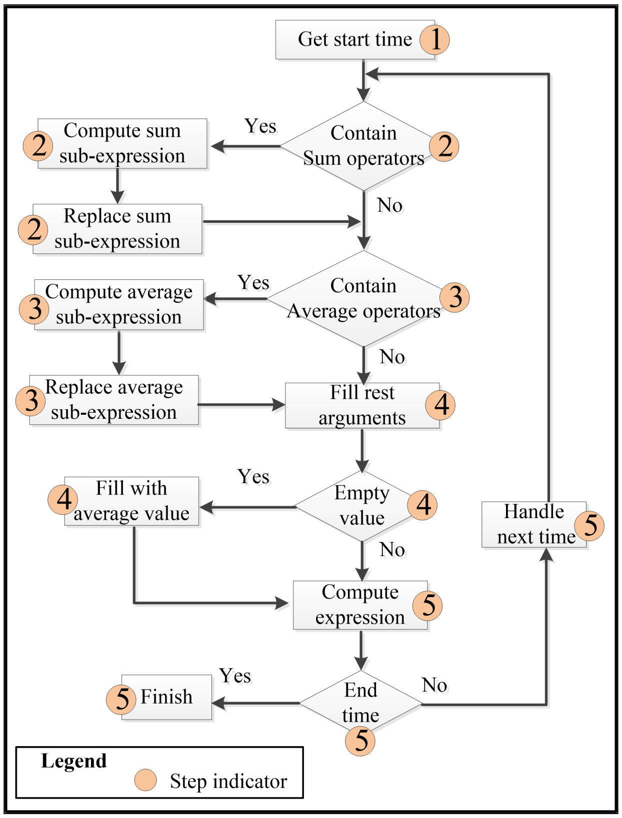


Figure 6: Workflow of the second algorithm for calculation of performance formulae.

The algorithm first extracts the time period over which the metric is to be evaluated. It also checks for the presence of sum operators (step 2 of Figure 6). Where present, these are calculated by summing the values of contained building objects; for example, the energy consumption for a floor is calculated by summing that metric for rooms contained by that floor object (step 2).Then any average operators can then be calculated by taking the result from step 2 and dividing it by the number of contained objects (step 3 of Figure 6).

Step 4 then populates the remaining arguments with the data streams prepared in section 4.1; if any of the data streams remains unfulfilled, it is substituted by taking the average of value at the previous and subsequent values. Having constructed and populated a mathematical expression, this is evaluated using a computer algebra system (step 5 of Figure 6). The process is repeated until the end time is equivalent to the present time.

Having described the implementation of the approach, its application to the *real world* is illustrated with a practical example.

5. Case study: university building performance assessment

In order to investigate and demonstrate the engineering value of the approach in a practical context, a case study has been developed for the sports centre in University College Dublin (UCD), Ireland. This facility extends over 11,000 m^2, spread over a three-storey complex, linking the existing sports and student centres. It includes facilities for student health, debating, drama, societies, media and leisure amenities in addition to a 50m x 25m swimming pool with related ancillary areas, including wellness suite, administration offices and a fitness centre. The building heating and cooling needs are covered by two CHP units, two boilers, a district heating installation and an air cooled chiller. The delivery equipment consists of 8 AHUs, FCUs, underfloor heating and baseboard heaters. The ventilation throughout the building is mechanical.

The building has significant potential for performance optimisation. Since it is a relatively complex building containing a number of different zone types, each with its own conditioning requirements. For example, the temperature and humidity in pool halls should be maintained within acceptable bounds for thermal comfort (30^{o}C +/- .5^{o}C) and reasons of chemical corrosion of the building structure the relative humidity should remain between 40-60%. A BMS and EMS are installed at the facility but the building manager must work with an incomplete representation of environmental and energy performance due to an insufficient number of sensors and costs associated with storing significant volumes of time-series data. Such data inadequacy is representative of data availability in commercial buildings [54]. Due to the complexity of the building and the interactions between the various zones and systems, the building manager has difficulty determining the overall efficacy of the building, its systems and components.

For a deeper understanding of building behaviour, the building manager at the sports centre presently has a significant concern regarding the energy recovered from the Air Handling Units (AHUs) that serve the Pool Hall. These units deliver $21\text{m}^3/\text{s}$ of air to the zone but the economisers are locked at 20% fresh air. The building manager would like indicative feedback on the operational benefits from an enhanced control strategy for the economisers while maintaining zone temperature and relative humidity. Significant reductions in heat recovered by the economisers should result in a sizable reduction in heating coil use from the two AHUs in question. The solution must also be designed in a way that facilitates a continuation of analysis with measured time-series data after any changes have been implemented.

The developed solution includes three components:

- The software framework which implements the newly developed components presented in this paper: these include the ontology structure and the querying and computing algorithms (Section 5.1).
- Selected performance metrics that are organised by a performance assessment method, called Scenario Modelling [39], aligning with the functionality of OpenMath.(Section 5.2).
- Performance assessment results that are evaluated based on Building Energy Performance Simulation (BEPS) of the facility (Section 5.3).

5.1. Software Framework

The software framework assists building managers to obtain insight building performance assessment in an automated and effective manner. It is builded upon an approach taken in a preceding research project, namely the linked dataspace for intelligent energy [35]. The framework is divided into four distinct layers (Figure 7): 1) the source layer, 2) the linked data layer, 3) the support services layer and 4) the interface layer. The remainder of this section will now discuss each of these layers in more detail.

the source layer locates at the bottom of the architecture and provides raw data for upper layers. Selected performance metrics are decomposed into RDF entities using the proposed ontology and stored in an Metric-RDF file. BIM information is transformed from an IFC file into an IFC-RDF file [55]. Sensor information is stored in an SSN-RDF file based on the SSN ontology [56]. A DB-RDF file represents schema information of databases that store time-series data related to building performance based on MySQL [27].

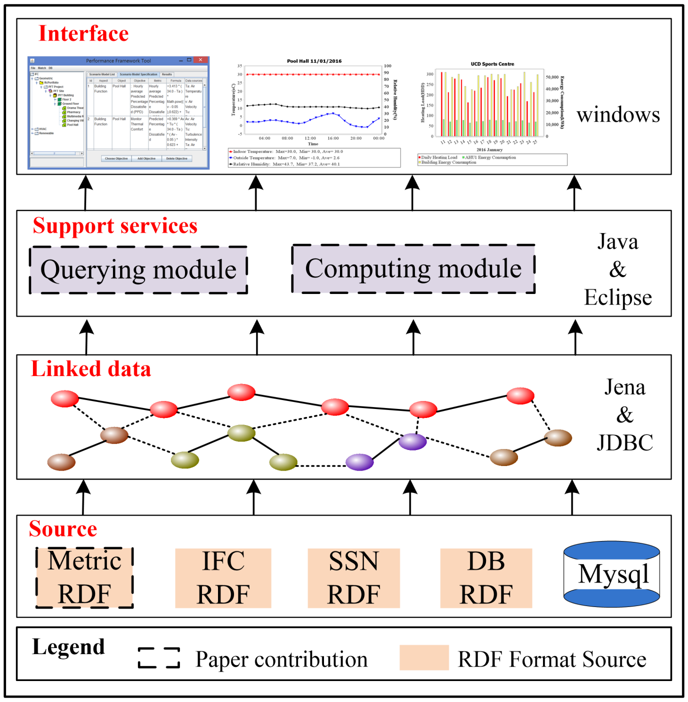


Figure 7: The framework leverages OpenMath, linked data and output from BEPS models to assess building performance.

the linked data layer leverages SPARQL and SQL queries to finish linking of cross-domain silos in the source layer. SPARQL queries are generated using the definition of ontologies, includinghttps://www.sharelatex.com/project/57f60b57834d86bf759469 ifcOWL ontology, SSN ontology, and the proposed ontology. SQL queries are generated with database schema information. The Apache Jena framework and Java database connectivity (JDBC) application programming interface (API) are used to execute SPARQL and SQL queries.

the support services layer is intended to simplify the consumption of linked dataspaces by encapsulating common services for reuse. In the context of this work, two new components, the querying and computing module, were required to implement the contributions developed in this paper. The querying module parses requirements in selected performance metrics and automatically generate SPARQL and SQL queries to obtain data streams. The computing module transfers performance metrics to normal mathematical formulae and leverages computer algebra systems to calculate these formulae for assessment results. Services contained in this layer use code developed in

Java within the Eclipse IDE.

the interface layer locates at the top of the architecture and displays the framework as the graphical user interfaces (GUI) in Windows systems. It visualises and enriches the information coming from the support services layer. Actionable information is now available to the building manger in a manner that was previously unavailable.

5.2. Performance assessment using scenario modelling and OpenMath

Scenario modelling enables the explicit and unambiguous coupling of building functions with other pivotal aspects of building operation through a method that specifically considers the education and technical expertise of building managers [39]. Using this method, building managers have more reliable information that can be communicated to stakeholders at the tactical and strategic levels of organisations. Thus, enabling more informed energy-related decisions by upper management, who require a return on investment for any new strategy or technology [57].

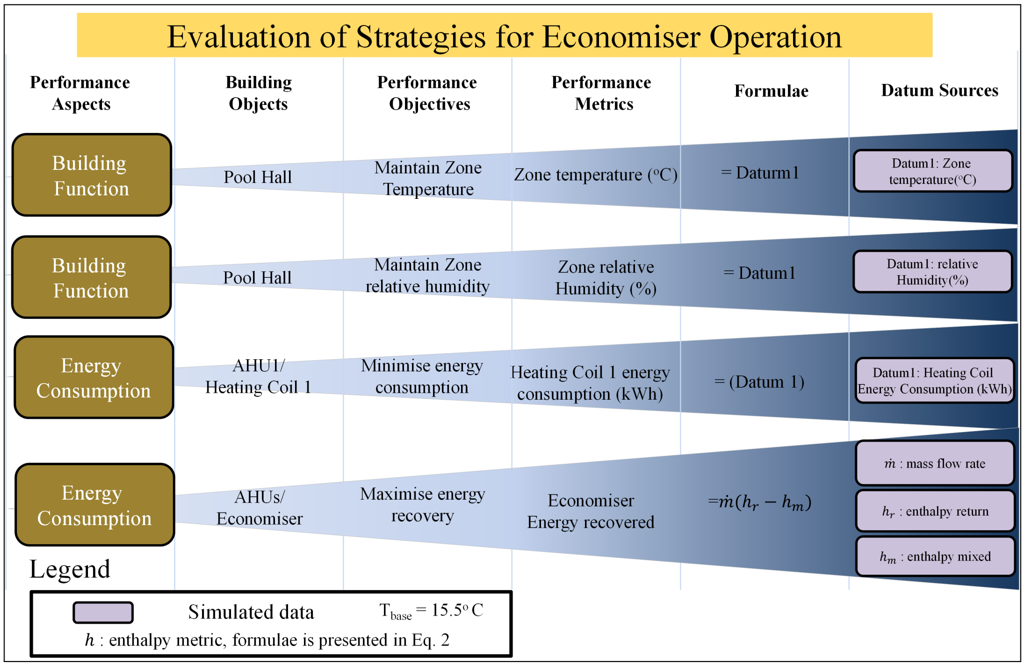


Figure 8: Example scenario model assessing the performance of the Pool Hall and associated energy handling equipment.

In this case, a scenario model assesses the performance of an important zone and associated air handling equipment by addressing building function and energy consumption in parallel (Figure 8). These operational aspects are evaluated using four performance metrics: 1) zone temperature (^{o}C), 2) zone relative humidity (%), 3) heating coil energy consumption (kWh), and 4) economiser energy recovered (kWh). The first two metrics leverage temperature and relative humidity to evaluate occupants' comfort within the pool hall as well as the impact the air borne chemicals may have on the structure if the relative humidity exceeds a certain threshold. The remaining two metrics track energy consumption and recovered energy related to the heating coils and economisers respectively.

The metric that evaluates economiser operation is of particular interest in the context as three interrealted equations are required (Eq. 1 - 4) requires an enthalpy based evaluation of the economiser using the outside air and mixed air conditions [58].

$$h = 1.006t + W(2501 + 1.86t) \tag{1}$$

$$W = 0.62198 \frac{\phi p_{ws}}{p - \phi p_{ws}} \tag{2}$$

$$\begin{aligned}
&\ln p_{ws} = C_8/T + C_9 + C_{10}T + C_{11}T^2 + C_{12}T^3 + C_{13}\ln T \\
&C_8 = -5.8002206E + 03 \\
&C_9 = \quad 1.3914993E + 00 \\
&C_10 = -4.8640239E - 02 \\
&C_11 = \quad 4.1764768E - 05 \\
&C_12 = -1.4452093E - 08 \\
&C_13 = \quad 6.5459673E - 00
\end{aligned} \tag{3}$$

$$T = t + 273.15 \tag{4}$$

Parameter	**Value**	**Unit**
t	Dry-bulb air temperature	(^{o}C)
W	Humidity ratio	(kg/kg)
p	Air pressure	(kPa)
ϕ	Relative humidity	(%)
p_{ws}	Water vapor saturation pressure	(kPa)
T	Absolute temperature	(K)

Extended RDF entity hierarchies of the ecnomoniser energy recovered metric are shown in Figure 9. Serving as one data repository of the proposed ontology, the metric is semantically represented as a recursive hierarchy tree. The root node represents the whole expression and consists of two arguments and one operator. One argument is the mass flow rate and another one is a sub-expression that presents the difference between enthalpy return (h_r) and enthalpy mixed (h_m). Both enthalpies are transformed into sub-expressions in accordance with Eq. 1.

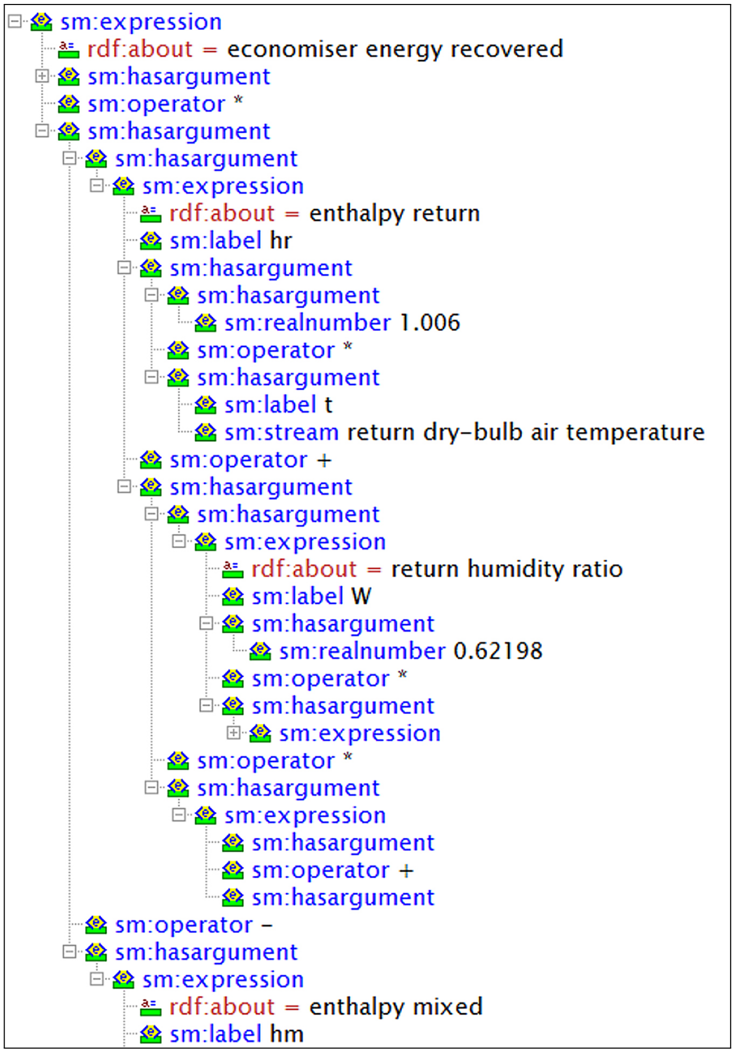
During the evaluation phase, the software framework extracts data streams for this metric, through generating an SPARQL query (Listing 1) and an SQL query (Listing 2). The SPARQL query inferences relevant building objects from the semantic BIM information and then identifies target sensors from the semantic sensor information. The SQL query combines sensor information and a fixed time period to extract sensing records from relational databases.

Listing 1: The SPARQL query extracting sensor entities related to the performance metric.

```
SELECT ?sensor
WHERE{
OPTIONAL{
?device rdf:type ifc:IfcEnergyConversionDevice.
?device rdf:type ifc:IfcAirToAirHeatRecovery.}
?device ifc:IfcRelConnects ?object.
?object rdf:type ifc:IfcZone.
?object ifc:Name "Pool␣Hall".
?device ssn:observedBy ?sensor.
?sensor rdf:type ssn:Sensor.
OPTIONAL{
?sensor ssn:observes ssn:airtemperature.
?sensor ssn:observes ssn:airpressure.
?sensor ssn:observes ssn:ralativehumdity.
?sensor ssn:observes ssn:waterpressure.}}
```

Listing 2: The SQL query extracting performance data from Mysql database.

```
SELECT sensorid,value,time
FROM building_records
WHERE
sensorid IN
```



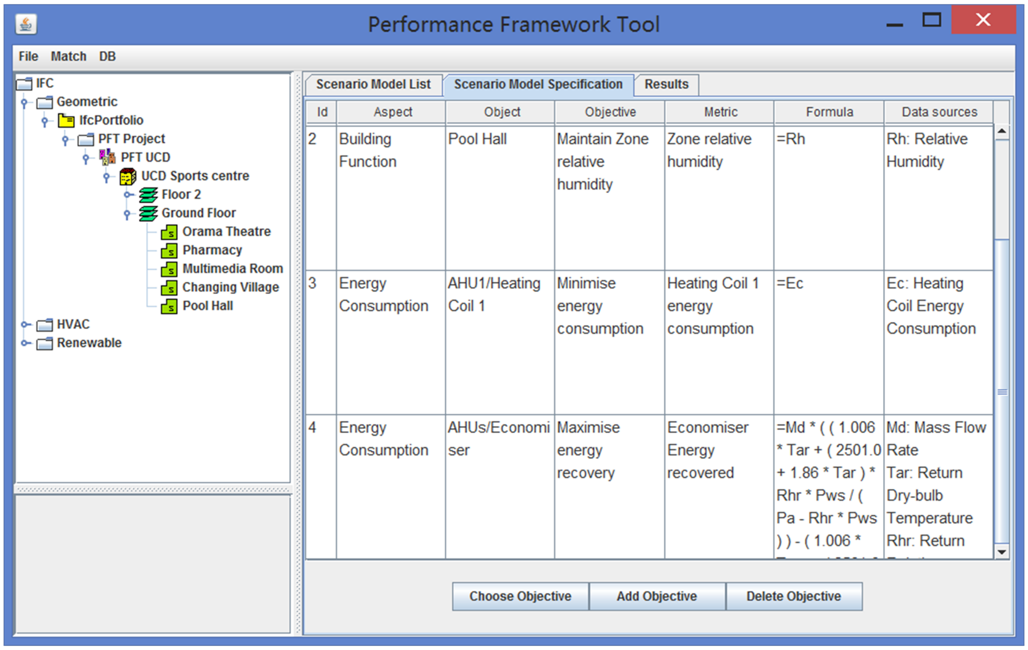
```
sm:expression
  rdf:about = economiser energy recovered
  sm:hasargument
  sm:operator *
  sm:hasargument
    sm:hasargument
      sm:expression
        rdf:about = enthalpy return
        sm:label hr
        sm:hasargument
          sm:hasargument
            sm:realnumber 1.006
          sm:operator *
          sm:hasargument
            sm:label t
            sm:stream return dry-bulb air temperature
        sm:operator +
        sm:hasargument
          sm:hasargument
            sm:expression
              rdf:about = return humidity ratio
              sm:label W
              sm:hasargument
                sm:realnumber 0.62198
              sm:operator *
              sm:hasargument
                sm:expression
          sm:operator *
          sm:hasargument
            sm:expression
              sm:hasargument
              sm:operator +
              sm:hasargument
    sm:operator -
    sm:hasargument
      sm:expression
        rdf:about = enthalpy mixed
        sm:label hm
```

Figure 9: The RDF entity hierarchies of the ecnomoniser energy recovered metric.

```
('sensor01','sensor02',..)
AND
time >= '2016-01-11␣00:00:00'
AND
time <= '2016-01-18␣00:00:00'
GROUP BY sensorid
```

5.3. Results

The scenario model instance is evaluated using the Performance Framework Tool, an in house application. Figure 10 displays three of the performance metrics in the illustrative example in this paper. Most importantly, the Performance Framework tool interface displays the OpenMath representation of each formula in a human readable format.



Id	Aspect	Object	Objective	Metric	Formula	Data sources
2	Building Function	Pool Hall	Maintain Zone relative humidity	Zone relative humidity	=Rh	Rh: Relative Humidity
3	Energy Consumption	AHU1/Heating Coil 1	Minimise energy consumption	Heating Coil 1 energy consumption	=Ec	Ec: Heating Coil Energy Consumption
4	Energy Consumption	AHUs/Economiser	Maximise energy recovery	Economiser Energy recovered	=Md * ((1.006 * Tar + (2501.0 + 1.86 * Tar) * Rhr * Pws / (Pa - Rhr * Pws)) - (1.006 *	Md: Mass Flow Rate Tar: Return Dry-bulb Temperature Rhr: Return

Figure 10: Implementation of OpenMath formulae within the Performance Framework Tool interface.

The scenario model is used to calculate four performance metrics for the base case with a fixed percentage of outdoor air and an alternative case with auto-sized outdoor air (Figures 11 and 12). On closer inspection, the indoor temperature of the pool hall is controlled at a constant level (within 2 degrees of 30^oC) while the outside temperature varies between 3 and 12^oC). The indoor relative humidity is suitable for occupants (around 60%).

The analysis shows that for both cases, building function, as defined by dry bulb temperature and relative humidity is within acceptable limits during the evaluation time period. These results indicate that building function is not affected in a negative manner by a change in economiser control strategy.

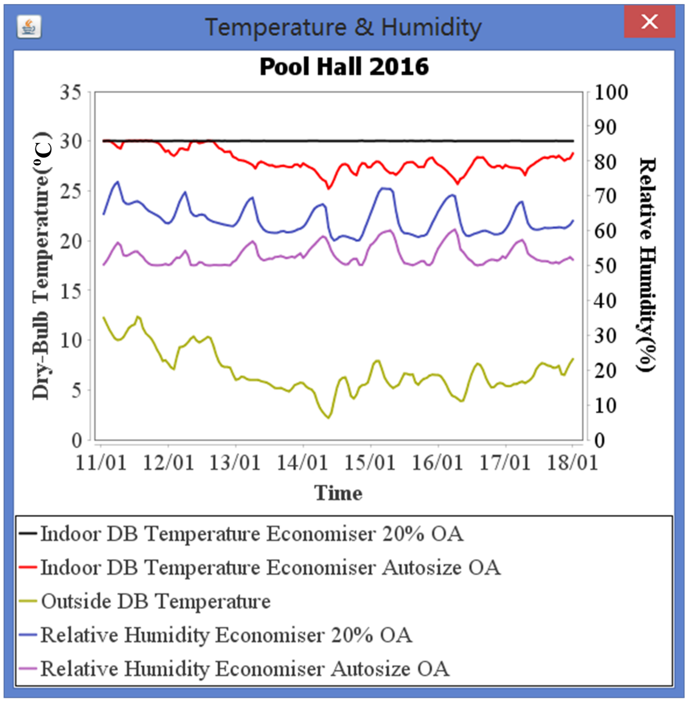


Figure 11: Assessment results of temperature and relative humidity.

The energy consumption of the heating coil is significantly lower in the case with varying modulation of the economiser dampers. This saving is also evident from the additional amount of heat recovered from the economiser (Figure 12). The energy savings are in line with expectations but the fact that Pool Hall humidity levels remain within acceptable bounds is encouraging for the building manager.

5.4. Discussion

There are a number of overarching findings from the demonstration. First of all, an ontology is needed to link building performance metrics and other contextual building information. Linked data provides a low-overhead mechanism to enable meaningful sharing of cross-domain building information. The ontology can deliver significant interactions between available silos that bridges the isolation. An ontology-based framework provides performance insights which are not currently available from isolated BMS systems and thus facilitates performance assessment by building mangers and pertinent stakeholders, without burdensome manual intervention.

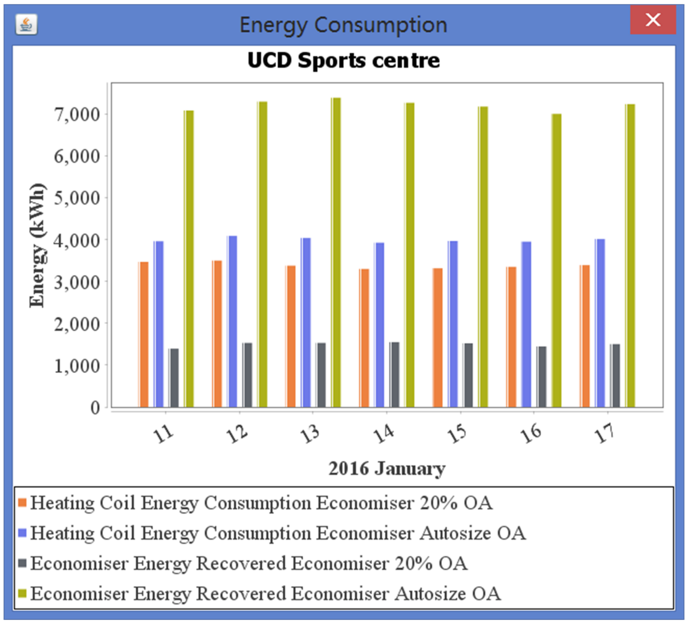


Figure 12: Assessment results of heating coil energy consumption and economiser energy recovered in two control strategies.

Secondly, the ontology can be applied to a wide range of buildings. Performance metrics have a profound impact on building performance analysis commercial buildings. The ontology can quantify impact of metrics such as building function, thermal load and system performance, as well as energy consumption. The framework has the potential to work well in buildings with data that is structured in easily accessible digital formats, e.g. BIM and data in relational databases. The flexibility and modularity of the ontology structure allows the description of solutions for buildings in different climate zones and geographical locations.

Thirdly, there are a number of boundaries to this approach. The evaluation of performance metrics requires datum streams from cross-domain data silos. For the majority of buildings only subsets of performance metrics will be calculable due to the absence of sensor data, e.g. temperature sensors in every zone. Installation of additional sensors may also be required in older buildings in order to obtain specific building information (e.g. occupant behaviour). Transformation algorithms need to be developed for building information that is difficult to access, such as occupant status from video cameras. High quality building information is essential for calculation or inference of accurate assessment results. Engineering adjustments should be made for buildings in different climate zones and geographical locations.

A broader context also arises for the ontology: the approach is applicable for analyses of building performance by pertinent stakeholders and across the entire building life cycle. Building engineers can setup a library of performance metrics that can in turn be shared with other engineers and used to assess building performance in a standardised manner. In the design phase, building architects revise building designs through assessment of their performance based on simulation. During operation, building managers can update and evaluate control strategies by detecting and diagnosing system faults with real-time building performance assessment.

6. Conclusions

Buildings do not operate as intended for a multitude of reasons, including the data interoperability issue between cross-domain building information. Linked data provides an open, modular and extensible solution to the issue by representing previously disconnected heterogeneous data in a homogeneous format (i.e. RDF). However, performance metrics are currently isolated from available contextual data silos.

The objective of this paper was to enable insight and in-depth building performance assessments by linking available computational approaches and traditionally disconnected data sources (i.e. OpenMath and Linked Data). This research defines a new ontology to present performance metrics. Two algorithms now automatically evaluate performance metrics: 1) the multi-silo querying algorithm which uses SPARQL and SQL queries to prepare datum streams for performance metrics, and 2) the computing algorithm which leverages computer algebra systems to calculate metrics. The approach was demonstrated in a sports centre building using a developed tool. Results shows that linking of cross-domain data provides a holistic view of the building performance for building managers.

This paper presents a flexible, automated and extensible solution to data sharing and metric calculating for the building performance assessment domain. Based on the approach, building managers and engineers can simultaneously analyse measured and simulated building performance at a range of levels, without much reinterpretation and conversion compared with traditional solutions. The approach offers a higher flexibility and a finer spectral granularity for building performance assessments. The approach has the potential to be applied in a range of application cases but is most powerful in cases where significant volumes of structured time-series data are present. Using the results of these assessments, building managers can communicate with other stakeholders at the tactical and strategic levels of organisations, and thus enable more informed energy optimisation strategies by upper management.

Future work will focus on extending the range of building information this approach can leverage and investigate integration of multi-criteria decision making theory with the scenario modelling method for enhanced building performance evaluation. The approach could also be extended to urban scale assessments.

7. Acknowledgments

This work is supported by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme, project title SuPerB (Project number 631617).

8. References

[1] M. W. Ahmad, M. Mourshed, D. Mundow, M. Sisinni, Y. Rezgui, Building energy metering and environmental monitoring–a state-of-the-art review and directions for future research, Energy and Buildings 120 (2016) 85–102.

[2] S. Copiello, Building energy efficiency: A research branch made of paradoxes, Renewable and Sustainable Energy Reviews 69 (2017) 1064–1076.

[3] L. La Fleur, B. Moshfegh, P. Rohdin, Measured and predicted energy use and indoor climate before and after a major renovation of an apartment building in sweden, Energy and Buildings 146 (2017) 98–110.

[4] P. De Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction 41 (2014) 40–49.

[5] J. Granderson, Energy information handbook: Applications for energy-efficient building operations, Tech. rep., LAWRENCE BERKELEY

NATIONAL LABORATORY, uRL `https://escholarship.org/uc/item/03z8k1v3` (2013).

[6] C. Neumann, D. Jacob, Results of the project building eq. tools and methods for linking epbd and continuous commissioning, Fraunhofer ISE.

[7] J. T. O'Donnell, Specification of optimum holistic building environmental and energy performance information to support informed decision making, Ph.D. thesis, University College Cork (2009).

[8] C. Fan, F. Xiao, H. Madsen, D. Wang, Temporal knowledge discovery in big bas data for building energy management, Energy and Buildings 109 (2015) 75–89.

[9] G. Costa, L. Madrazo, Connecting building component catalogues with bim models using semantic technologies: an application for precast concrete components, Automation in Construction 57 (2015) 239–248.

[10] D. Lee, C. C. Cheng, Energy savings by energy management systems: A review, Renewable & Sustainable Energy Reviews 56 (2016) 760–777.

[11] E. Curry, J. ODonnell, E. Corry, S. Hasan, M. Keane, S. ORiain, Linking building data in the cloud: Integrating cross-domain building data using linked data, Advanced Engineering Informatics 27 (2) (2013) 206–219.

[12] J. Han, Y. K. Jeong, I. Lee, Efficient building energy management system based on ontology, inference rules, and simulation, in: Proceedings of International Conference on Computer Communication and Management(ICCCM 2011), 2011, pp. 295–299.

[13] L. M. Blanes, M. M. Keane, Ontology-based facility data model for energy management, Advanced Engineering Informatics 29 (4) (2015) 971–984.

[14] P. Pauwels, Reconsidering information system support for architectural design thinking, Ph.D. thesis, Ghent University (2012).

[15] D. Tzovaras, T. Brennan, D1.3 state-of-art and industry analysis, adapt4ee occupant aware, intelligent and adaptive enterprises, Tech. rep., adapt4ee, uRL `http://www.adapt4ee.eu/adapt4ee/files/document/deliverables/Adapt4EE-Deliverable-D1.3.pdf` (2012).

[16] P. Pauwels, Supporting decision-making in the building life-cycle using linked building data, Buildings 4 (3) (2014) 549–579.

[17] T. Berners-Lee, J. Hendler, O. Lassila, et al., The semantic web, Scientific american 284 (5) (2001) 28–37.

[18] P. Pauwels, S. Zhang, Y.-C. Lee, Semantic web technologies in aec industry: A literature overview, Automation in Construction xx (2016) xxx.

[19] P. Pauwels, Ifc-to-rdf conversion service, Tech. rep., UNIVERSITY GENT, uRL `https://github.com/mmlab/IFC-to-RDF-converter.` (2015).

[20] 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, ACM, 2016, p. 7.

[21] Y. Li, J. ODonnell, R. García-Castro, S. Vega-Sánchez, Identifying stakeholders and key performance indicators for district and building energy performance analysis, Energy and Buildings 155 (2017) 1–15.

[22] G. Klyne, J. J. Carroll, Resource description framework (rdf): Concepts and abstract syntax. w3c recommendation, 2004, World Wide Web Consortium, http://w3c. org/TR/rdf-concepts.

[23] S. Buswell, O. Caprotti, D. P. Carlisle, M. C. Dewar, M. Gaetano, M. Kohlhase, The open math standard, Tech. rep., version 2.0. Technical report, The Open Math Society, 2004. http://www. openmath. org/standard/om20 (2004).

[24] J. ODonnell, E. Corry, S. Hasan, M. Keane, E. Curry, Building performance optimization using cross-domain scenario modeling, linked data, and complex event processing, Building and Environment 62 (2013) 102–111.

[25] E. Corry, A semantic web approach to enable the holistic environmental and energy management of buildings, Ph.D. thesis, National University of Ireland, Galway (2014).

[26] E. Corry, P. Pauwels, S. Hu, M. Keane, J. O'Donnell, A performance assessment ontology for the environmental and energy management of buildings, Automation in Construction 57 (2015) 249–259.

[27] S. Hu, E. Corry, E. Curry, W. J. Turner, J. O'Donnell, Building performance optimisation: A hybrid architecture for the integration of contextual information and time-series data, Automation in Construction 70 (2016) 51–61.

[28] Memoori, Big Data for Smart Buildings: Market Research 2015 to 2020, Tech. rep., Memoori (2014).

[29] H. Grzybek, S. Xu, S. Gulliver, V. Fillingham, Considering the feasibility of semantic model design in the built-environment, Buildings 4 (4) (2014) 849–879.

[30] B. Yuce, Y. Rezgui, M. Mourshed, Ann–ga smart appliance scheduling for optimised energy management in the domestic sector, Energy and Buildings 111 (2016) 311–325.

[31] B. Yuce, Y. Rezgui, An ann-ga semantic rule-based system to reduce the gap between predicted and actual energy consumption in buildings, IEEE Transactions on Automation Science and Engineering.

[32] G. Augenbroe, The role of simulation in performance based building, Building performance simulation for design and operation (2011) 15–36.

[33] J. O'Donnell, Simmodel: A domain data model for whole building energy simulation, SimBuild 2011, Sydney, Australia, 11/14/2011-11/16/2011.

[34] R. J. Hitchcock, Standardized building performance metrics final report, Tech. rep., Citeseer (2003).

[35] E. Curry, S. Hasan, S. O'Riain, Enterprise energy management using a linked dataspace for energy intelligence, in: Sustainable Internet and ICT for Sustainability (SustainIT), 2012, IEEE, 2012, pp. 1–6.

[36] V. Corrado, I. Ballarini, L. Madrazo, G. Nemirovskij, Data structuring for the ontological modelling of urban energy systems: The experience of the semanco project, Sustainable Cities and Society 14 (2015) 223–235.

[37] W. Terkaj, A. Šojić, Ontology-based representation of ifc express rules: An enhancement of the ifcowl ontology, Automation in Construction 57 (2015) 188–201.

[38] F. Radulovic, M. Poveda-Villalón, D. Vila-Suero, V. Rodríguez-Doncel, R. García-Castro, A. Gómez-Pérez, Guidelines for linked data generation and publication: An example in building energy consumption, Automation in Construction 57 (2015) 178–187.

[39] J. ODonnell, M. Keane, E. Morrissey, V. Bazjanac, Scenario modelling: A holistic environmental and energy management method for building operation optimisation, Energy and Buildings 62 (2013) 146–157.

[40] P. Pauwels, E. Corry, J. ODonnell, Representing simmodel in the web ontology language, in: American Society of Civil Engineers, 2014, pp. 2271–2278.

[41] T. Hong, S. D'Oca, W. J. Turner, S. C. Taylor-Lange, An ontology to represent energy-related occupant behavior in buildings. part i: Introduction to the dnas framework, Building and Environment 92 (2015) 764–777.

[42] T. Hong, S. D'Oca, S. C. Taylor-Lange, W. J. Turner, Y. Chen, S. P. Corgnati, An ontology to represent energy-related occupant behavior in buildings. part ii: Implementation of the dnas framework using an xml schema, Building and Environment 94 (2015) 196–205.

[43] E. Corry, J. ODonnell, E. Curry, D. Coakley, P. Pauwels, M. Keane, Using semantic web technologies to access soft aec data, Advanced Engineering Informatics 28 (4) (2014) 370–380.

[44] M. Dibley, H. Li, Y. Rezgui, J. Miles, An ontology framework for intelligent sensor-based building monitoring, Automation in Construction 28 (2012) 1–14.

[45] N. M. Tomašević, M. Č. Batić, L. M. Blanes, M. M. Keane, S. Vraneš, Ontology-based facility data model for energy management, Advanced Engineering Informatics 29 (4) (2015) 971–984.

[46] F. H. Abanda, J. H. Tah, D. Duce, Pv-tons: A photovoltaic technology ontology system for the design of pv-systems, Engineering Applications of Artificial Intelligence 26 (4) (2013) 1399–1412.

[47] M. J. Kofler, C. Reinisch, W. Kastner, A semantic representation of energy-related information in future smart homes, Energy and Buildings 47 (2012) 169–179.

[48] K. Wenzel, H. Reinhardt, Mathematical computations for linked data applications with openmath, in: Joint Proceedings of the 24th Workshop on OpenMath and the 7th Workshop on Mathematical User Interfaces (MathUI), 2012, p. 38.

[49] O. Caprotti, D. Carlisle, Openmath and mathml: semantic markup for mathematics, Crossroads 6 (2) (1999) 11–14.

[50] R. Cretney, Editing and reading early modern mathematical texts in the digital age, Historia Mathematica 43 (1) (2016) 87–97.

[51] M. Kohlhase, F. Rabe, Semantics of openmath and mathml3, Mathematics in Computer Science 6 (3) (2012) 235–260.

[52] C. Lange, Towards openmath content dictionaries as linked data, arXiv preprint arXiv:1006.4057.

[53] O. Caprotti, A. M. Cohen, H. Cuypers, H. Sterk, Openmath technology for interactive mathematical documents, in: Multimedia Tools for Communicating Mathematics, Springer, 2002, pp. 51–65.

[54] J. Granderson, G. Lin, M. A. Piette, Energy information systems (EIS): Technology costs, benefit, and best practice uses, Tech. Rep. LBNL-6476E, Lawrence Berkeley National Laboratory, Berkeley, CA (2013). URL `http://eetd.lbl.gov/node/57232`

[55] P. Pauwels, W. Terkaj, T. Krijnen, J. Beetz, Coping with lists in the ifcowl ontology, in: 22nd EG-ICE International Workshop, 2015, pp. 113–122.

[56] L. Lefort, C. Henson, K. Taylor, P. Barnaghi, M. Compton, O. Corcho, R. Garcia-Castro, J. Graybeal, A. Herzog, K. Janowicz, et al., Semantic sensor network xg final report, W3C Incubator Group Report 28.

[57] D. Pati, C.-S. Park, G. Augenbroe, Roles of quantified expressions of building performance assessment in facility procurement and management, Building and Environment 44 (4) (2009) 773–784.

[58] A. Handbook, et al., American society of heating, Refrigeration and Air-Conditioning Engineers, Inc.