Enhancing energy management at district and building levels via an EM-KPI ontology

Yehong Li[a b], Raúl García-Castro[c], Nandana Mihindukulasooriya[c], James O´Donnell[d], Sergio Vega-Sánchez[b]

[a]Department of Building Energy Efficiency Technology, Guangzhou Institute of Building Science CO.,LTD., GuangDong, China

[b]Department of Construction and Building Technology, School of Architecture, Universidad Politécnica de Madrid, Avda. Juan de Herrera 4, 28040 Madrid, Spain

[c]Ontology Engineering Group, ETSI Informáticos, Universidad Politécnica de Madrid, Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Spain

[d]School of Mechanical and Materials Engineering, ERC & UCD Energy Institute, University College Dublin, Belfield, Dublin 4, Ireland

yehong.li@alumnos.upm.es, rgarcia@fi.upm.es, nmihindu@fi.upm.es, james.odonnell@ucd.ie, sergio.vega@sdeurope.org

Corresponding Author:

Yehong Li

Institute of Building Energy Efficiency Technology

GuangZhou Institute of Building Science CO.,LTD.

BaiYun DaDao Bei 833, BaiYun District, 510440 GuangZhou, GuangDong, China

Tel: +86 13268341086

Email: Yehong.li@alumnos.upm.es

Abstract

The use of information and communication technologies facilitates energy management (EM) at both district and building levels but also generates a considerable amount of data. To gain insights into such data, it is essential to resolve the cross-domain data interoperability problem and determine an approach to exchange performance information and insightful data amongst various stakeholders. This paper developed an EM-KPI (key performance indicator) ontology to exchange key performance information and data for districts and buildings. The ontology contains two components: namely KPIs and EM master data; these, respectively, represent multi-level performance information for energy performance tracking and the key data for data exploitation. Through a demonstration, a sample linked dataset generated using the data correlation predefined in the ontology is presented. The linked data analysis proves the feasibility of the ontology for exchanging data among different stakeholders and for exploring insights in relation to performance improvements.

Key words:

District; building; energy management; stakeholders; ontology; linked data.

1. Introduction

Buildings account for approximately 40% of the total final energy use in EU countries [1]. However, a large portion of existing buildings are either designed or operated inefficiently [2]. Energy management (EM) is a measure adopted to improve energy efficiency in buildings. Furthermore, there is an increasing need to manage energy not only in a single building, but also on a district scale [3]. Since the implementation of smart cities involves increasing distributed electricity generation such as solar panels in energy distribution networks, EM at a district level, for the purpose of combining the electricity supply and demand of buildings, is pivotal [4,5]. The use of information and communication technologies (ICTs) facilitates the realisation of joint EM that integrates the energy supply and demand sides.

Meanwhile, the use of ICTs also generates a massive amount of data and information, which could provide new analysis possibilities for data-driven decision support and offer insights in relation to potential performance improvement [6]. According to the National Institute of Standards and Technology (NIST) in the United States, it could save up to $2 trillion in energy costs by 2030, through exploiting the data from smart grids [7]. Although the expansion of data presents great opportunities for energy performance improvement, there are still challenges faced in the effort to make sense of this data. The problem is twofold. Firstly, there is an interoperability problem between the cross-domain heterogeneous data. Secondly, the solution requires the extraction of insightful data in order to avoid unnecessary analysis.

The extraction of core, insightful data is the primary challenge encountered when seeking to access a large amount of data. Data exploitation is valuable only if they address the issues related to the stakeholders. It is important to focus on data that is worth collecting, analyses which are worth sharing and problems which are worth solving [8]. Master data offers a way to represent key data that provides the most valuable information in an organisation [9]. In this case, master data refers to the critical data objects that need to be shared across or beyond an organisation which support decision-making. Master data was initially used for enterprise data management due to the large volumes of data generated during business processes [10]. In the context of energy management, a similar situation is encountered. Introducing the concept of master data into the energy field can help make the large amount of energy-related data actionable, thus bringing additional insight and value through improved decision-making.

The master data involved in EM should be shared among different stakeholders; therefore, it is essential to support their performance concerns. In our previous study, we defined stakeholders as those who have an interest in, who have influence in and who are impacted by the actions of energy management; a detailed methodology was developed for selecting the KPIs (key performance indicators) that underpin stakeholders' performance goals; additionally, the use of KPIs for master data identification was proposed [11]. Using KPIs to identify the master data can ensure that the data also supports stakeholders' concerns. KPIs offer a means not only to measure the progress made towards stakeholders' goals, but also to condense a large amount of data into a critical piece of performance information [12]. If the stakeholders can gain easy access to both the performance information represented by KPIs and their related master data, these stakeholders can obtain a better understanding of performance and areas that requiring improvement. Therefore, it is crucial to develop an approach to facilitate the interchange of key performance information and insightful master data among the stakeholders.

The cross-domain data, however, are usually stored in different information islands; there is little interaction between each other for effective data sharing and exchange [13]. One of the main barriers is due to the interoperability of heterogeneous data, since the data in EM usually includes multiple domains. To resolve the data interoperability problem, the semantic web provides a possible solution. Linked data harnesses the ethos and infrastructure of the Web to enable data sharing and reuse on a massive scale [14]. In recent years, linked data has been the subject of growing interest and applications in the building and energy fields. For example, Corry et al. [15] used linked data to access AEC (architecture, engineering and construction) data for building performance analysis. The EU FP7-funded project SEMANCO used ontological modelling to access widely dispersed energy-related data pertaining to cities for the purpose of improved energy analysis [16]. Undergoing projects such as NewTrend [17] and OptEEmAL [18] create ontology-based district information models for building and district retrofitting. Furthermore, the project READY4SmartCities [19] presented a set of guidelines for generating linked data in the energy domain in order to support the interoperability and exploitation of data. In addition, a range of ontologies pertaining to smart cities, energy and other related fields have been collected [20]. However, an ontology aiming to integrate both the multi-level key performance information and the multi-domain master data has not been developed to date.

In this paper, an EM-KPI ontology is developed to enable the exchange of master data and key performance information for energy management at district and building levels, only the energy type of electricity is considered. The ontology integrates multi-level KPIs, their calculation and the master data domains, in order to provide the basis for both performance tracking and the exploration of insights for performance improvement. Stakeholders are involved in the data exchange in order to promote engagement and enhance multi-level EM. In Section 2, a review of the ontologies for EM on district and building scales is conducted. Section 3, meanwhile, illustrates the development process of the targeted EM-KPI ontology. Section 4 presents a demonstration of the ontology adopted to generate linked data. Finally, Section 5 concludes the study and presents recommendations for future work.

2. Related Work

Ontologies are the foundation of linked data; an ontology represents the concepts and relationships within a specific domain in a well-defined and unambiguous manner [21]. We reviewed existing ontologies in the field of EM at district and building scales. Currently, the application of ontologies in EM targets system control rather than the purpose of generating linked data. Most of the research is focused on using ontologies for smart buildings/homes or smart grids, only a limited number of studies have been carried out to integrate both buildings/homes and grids to enable integrated, multi-level energy management.

2.1. Ontology for System Control and Operations

Previous studies used ontologies for system control and operations in building EM. For example, Grassi et al. [22] focused an ontology framework for device and energy description in order to achieve intelligent home management and energy saving. Kofler et al. [23] defined a Smart Home Ontology covering domains such as buildings, processes, exterior conditions, and energy and resource information. In addition, Wicaksono et al. [24], Han et al. [25] and Caffarel et al. [26] used ontology-driven approaches for building energy management. However, all of these studies are limited to the single building/home scale.

Energy utilities must also be considered in order to extend energy management to the district scale. The related studies addressing the use of ontologies to optimise EM in smart grids or microgrids include that of Rohjans et al. [27], who used semantic web services to realise information exchange for transmission and distribution grids. Additionally, Neumann et al. [28] presented an ontology for system integration in power systems, Salameh et al. [29] dealt with microgrid management, and Macek et al. [30] developed an ontology-based energy monitoring and control system for smart energy grids. These studies mostly focus on the energy supply side, while paying little attention to the demand side of buildings.

For district-scale EM, the supply side and demand side are of the same importance. However, there are still limited studies about the integration of building/home and microgrid. Previous studies include that of Anvari-moghaddam et al. [31], who developed an ontology-driven control system for integrated building and microgrid management. Meanwhile, The EU-funded research projects ENERsip [32] and DIMMER [33], respectively, proposed an ontology to model smart grid neighbourhoods and created a virtual district information model (DIM) for energy management in a smart city. While other studies dealt with ontological approaches to building and district EM using artificial intelligence [34].

2.2. Ontology for Knowledge Management and Information Integration

The studies described above all target system operations. However, the use of ontologies to generate linked data for data reuse and exploitation is another issue. Corry et al. [35] presented a data-driven approach to the structured performance assessment of buildings utilising semantic web technologies and linked data. Similar studies also include the research of Curry et al. [13], who proposed integrating cross-domain building data using linked data for managing a building. These studies introduce the role of linked data for operations in a single building. Additionally, Shah et al. [36] devised an ontology covering general classifications of domestic appliances for the home energy management domain, but without consideration of any other energy-related data, such as energy use data and building data.

Regarding the application of linked data in power grids, Zhou et al. [37] presented a semantic information model, comprising electrical equipment, organisation, infrastructure, weather, spatial and temporal ontologies. Simmhan et al. [38], meanwhile, used an integrated ontology for load optimisation and advanced analytics in smart grids. In addition, Gillani et al. [39] developed a generic ontology for integrating sensory data, infrastructure types, electrical appliances, electrical generation systems, weather reports, and so on. Gomes et al. [40] also proposed an ontology to represent a time-series of multiple observations in microgrids. These studies provide information for different aspects of power grids; nevertheless, the detailed information related to energy end-use buildings is still unconsidered.

There are precedents to the generation of linked data regarding either buildings or microgrids; however, few have focused on using ontologies to integrate the related data sources in both buildings and microgrids for enhancing multi-level EM. To date, the unique study found by the authors that integrates buildings and utilities includes a 3D city modelling method using CityGML [41].

2.3. Ontology for Performance Assessment

To harness the knowledge and insights in the linked data for energy performance improvement, it is essential to define a way in which to evaluate such performance. There are several ontologies regarding performance assessment in buildings, such as the Performance Information Model (PIM) ontology [42] and the Performance Framework (PF) ontology [43]. These two ontologies provide a framework to evaluate building performance, but without detailed descriptions of the required data sources. Díaz et al. [44] proposed an Energy Efficiency Ontology (EEOnt) for a unified representation of energy efficiency in buildings, which contains information related to building structure, systems and devices, and the EEI_B (Energy Efficiency Index) and EEL_B (Energy Efficiency Landscape). EEOnt aims to supply useful information for the diagnosis and correction of inefficiencies in buildings. However, all of these performance ontologies above are designed for the representation of performance information in buildings. The ontologies that describe energy performance in a microgrid or in a district have not yet been identified, except in the case of a Global City Indicator Environment Ontology [45] which assesses the environment in a city but does not describe the aspects of energy management.

A key finding from the existing ontology review is the absence of an ontology to integrate the key performance information and the related data sources on both the district and building scales. A noticeable opportunity emerges when both the multi-level evaluation and performance problem identification are considered together. It is quite important that a new ontology can represent the key performance information of stakeholders' concerns and build upon the variety of information available from district-scale EM. In the following section, a new EM-KPI ontology is developed to represent the multi-level key performance information and integrate the master data domains.

3. EM-KPI Ontology Development

The development of the EM-KPI ontology follows the NeOn ontology engineering methodology [46]. This method builds ontology networks through reusing and re-engineering knowledge resources as opposed to building new ontologies from scratch. Since EM at district and building scales is complicated, the targeted EM-KPI ontology may be relatively complex. If each term needs to be defined anew, there will be a huge amount of work involved. Fortunately, a range of reusable ontology resources already exists. Linking the existing knowledge to generate a new ontology network which has a specific aim of serving the new application can save time and work; furthermore, it could facilitate the implementation of the ontology. Reusable ontological resources include ontology patterns and vocabularies, which can be reused as a whole or partially, while it is essential to justify the reasons why the resources are chosen. To complete the ontology, new patterns and concepts are also defined in order to represent newly created content.

Using the NeOn methodology, the first step adopted to develop the ontology is the definition of ontology requirements, such as identifying the ontology's scope and intended end-users. The second step involves extracting the terms for building the ontology. Following this, the third step defines the conceptual model of the ontology, whose purpose is to offer a global view of the main relationships between different domains. Afterwards, the ontology's detailed model is developed in parallel with the fourth step, ontology search and selection. The fifth step is to implement the ontology by integrating the reused ontologies and the newly developed ones. If the ontology can represent all of the data pointing to the applications or use-case experiments, it is completed. Otherwise, additional work to complete the ontology should be performed.

3.1. Ontology Requirements Definition

Ontology requirements definition includes the identification of 1) the ontology's purpose, scope and implementation language; 2) the intended end-users; 3) the intended uses and 4) the non-functional and functional requirements of the ontology.

3.1.1. Identifying Purpose, Scope and Implementation Language

As mentioned above, the targeted EM-KPI ontology is aimed at exchanging both the multi-level key performance information and the cross-domain master data between various stakeholders. Therefore, the proposed ontology is to be represented by two components, namely EM master data and KPIs, as shown in Figure 1.

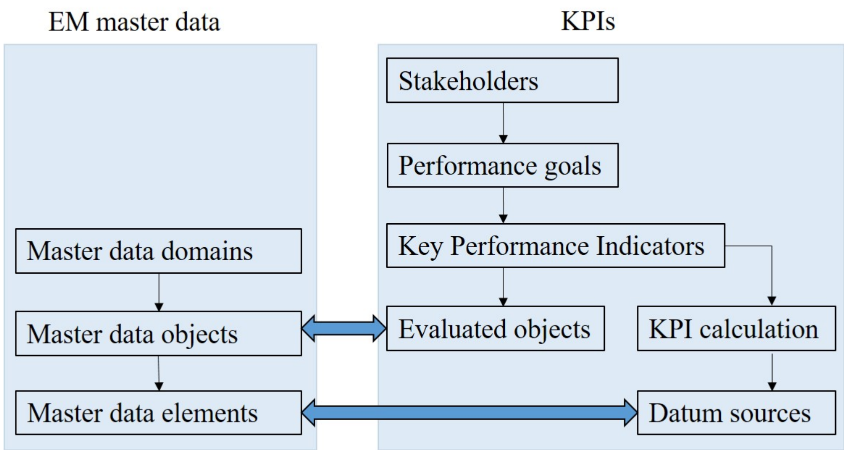

Figure 1: The proposed EM master data and KP components of the EM-KPI ontology.

The purpose of the KPI component is to represent the multi-level key performance information. This includes KPI definitions, associated performance goals and stakeholders, evaluated objects, calculations and required datum sources. This part provides the basis for energy performance tracking and assessment. The purpose of the EM master data component is to integrate key cross-domain data from districts and buildings that should be shared among stakeholders. This component intends to describe the relationships between the different master data domains, objects and elements. The EM master data sources are associated with KPIs, which provide the basic data for KPI calculation and analysis.

The combination of these two components helps to track and assess performance, to exploit the knowledge and insights within the master data sources for the identification of performance problems and key areas for improvement, and to support energy managers in making informed decisions with regard to energy efficiency measures. The link between the EM master data and the KPI components occurs through the data objects and elements, as the master data objects are associated with the KPI-evaluated objects and the master data elements provide the datum sources for KPI calculations.

The KPIs that underpin the stakeholders' performance goals should represent the energy performance information at different levels. Generally, KPIs in EM are classified into strategic, tactical and operational [47,48]; respectively, these represent the energy performance at the district level, building and system levels, and the zone and equipment levels [11].

The targeted ontology was implemented in OWL (Web Ontology Language) [49].

3.1.2. Identifying the Intended End-users

Since the ontology aims to enhance energy management, its intended end-users should be those stakeholders who have interest in, who have influence in and who are impacted by the actions of energy management. In this case, energy managers are the main actors, as they are responsible for the management of energy operations. District energy managers and building energy managers interact through the ontology and use the shared data from any other stakeholders to analyse and improve energy performance, so as to achieve defined performance goals. In addition, other stakeholders should become informed with regard to relevant energy performance by gaining access to related information, and thereby engage in decision-making. Therefore, the end-users of the ontology include the following:

- User 1. District energy managers who perform district energy operations;
- User 2. Building energy managers who fulfill building energy optimisation; and
- User 3. Other stakeholders involved in energy management.

3.1.3. Identifying the Intended Uses

The main intended uses of the ontology include the following:

- Use 1. To exchange the energy-related master data among different stakeholders;
- Use 2. To provide key performance information for various stakeholders; and
- Use 3. To support performance analysis through linked data for identifying key areas for improvement and achieving stakeholders' goals.

3.1.4. Identifying the Non-functional and Functional Requirements

Regarding the non-functional requirements, the ontology should strive to adopt the concepts and patterns in existing ontologies where possible, combining them with the newly developed terms and patterns.

The competency question (CQ) technique [50] is used for the identification of functional requirements that specify the knowledge which should be encapsulated within the ontology model. Considering the intented application and use-case experiments, the CQs are defined from the KPIs to the related master data domains using a top-down approach. Some examples of the CQs are as follows:

- Who are the stakeholders involved in energy management at district and building levels?
- What KPIs can be used to measure the performance goals of stakeholders?

- How can KPIs be calculated?
- What observation provides the datum sources for KPI calculation?
- Where is the district located?
- What type of buildings does the district contain?
- What kind of energy-generating units are installed in the district?
- When is the energy production monitored?
- What is the unit of measurement of the energy production?
- What energy aspects are influenced by the weather?

Table 1: Range of the identified domains for the EM-KPI ontology.

Domain	Range
KPI	KPI definition and its hierarchy, KPI calculation and its value, the evaluated object, the associated stakeholders and performance goals
Observation	Observation of any parameter involved in energy management, including the observation value, the observed property, the observation time and feature of interest
Utility	Only electrical system considered, i.e. power system and its equipment, and the related parameters
Building	Building basic information, building envelope and its thermal properties, building schedule and event, building energy facilities and its equipment, and the related parameters
Occupancy	Occupant number and energy-related behaviors, and the acpects that impact the behaviors
Energy	The parameters of energy generation, storage, supply and consumption, energy price, energy cost, and energy forecast
Weather	Weather conditions, weather parameters, and weather forecast
Location	Location of each building and power system resource
Date time	The evaluated time of each KPI, and the obaservation time of each measurement
Unit	The unit of each value

Subsequently, the CQs are categorised according to the domains to which their knowledge belongs. For example, the CQ (what KPIs can be used to measure the performance goals of stakeholders) is classified into the KPI domain, since the answer to this question refers to the related KPIs; and the CQ (what observation provides the datum sources for the calculation of KPIs) is grouped into the observation domain, because its answer involves the observation data. Each CQ is assigned to the respective domain. As a result, 10 groups are sorted, including KPI, observation, utility, building, occupancy, energy, weather, location, date time, and unit. The range of each domain is listed as Table 1. For example, the buildings domain contains basic building information, information pertaining to the building envelope and its thermal properties, energy systems and components, etc. Envelope type and the associated thermal properties as represented by U-value are typically obtained within design documents. As mentioned above, the listed domains and ranges are identified based on the intended application and use-case experiments. In this case, only electricity is considered as opposed to other energy vectors such as gas or heat. For further work, the related domains and energy types could be extended.

3.2. Terms Extraction

Since the CQs contain the knowledge that should be covered by the ontology, most of the ontology terms can be extracted from the CQs. Table 2 lists the extracted terms, with their synonyms contained in brackets. Other terms can also be directly extracted from the data sources of use cases and/or the existing ontologies.

Table 2: Terminology from competency questions, with their synonyms in brackets.

Top terms
key performance indicator (KPI), district, building, observation (measurement), location, KPI calculation, weather, interval and date time (temporal entity), unit, power system (utility), occupant
Other terms

stakeholder, energy performance goal, KPI-evaluated object, interval, KPI value, datum source, geographic coordinate point, weather condition, weather phenomenon, weather forecast, energy type, power equipment, generating unit, storage unit, power delivery unit, energy consumer, energy facility, building type, building dimension (area, volume, etc.), building element (wall, window, floor, roof, etc.), thermal property (U-value), building equipment, schedule, event, occupant behaviour, indoor comfort, energy parameter, energy production, energy storage, energy delivery (energy supply), energy use (energy demand), energy cost, energy tariff, energy forecast, equipment parameter

3.3. Ontology Conceptualisation

An initial conceptualisation of the EM-KPI ontology was drafted in order to gain a global view of the main classes and relationships within the different domains. Figure 2 illustrates the conceptual model of the proposed ontology using the extracted terms. This model has been built taking into account both the ontology's purpose and scope as stated in Section 3.1.1, and the functional requirements identified in Section 3.1.4. It describes both the KPI and the master data components, which are shown respectively in the left and right parts. As utilities and buildings in the district are both infrastructure, they are combined into one module. The time and unit domains are included in the observation module, as they are essential for unambiguous data descriptions. The energy parameters are represented in a domain parameter module. As a result, the target ontology is divided into seven ontological modules, namely KPI, infrastructure, weather, location, occupancy, observation and domain parameter.

The KPI module represents the main classes and relationships related to the stakeholders, strategic performance goals, KPIs and the calculation of the KPIs. Since the calculation requires the datum sources provided by observations, it is linked to the observation module. The observation module illustrates the various concepts for the description of the observation results, the observed property and its feature of interest. The time and unit domains are used for unambiguous descriptions for observation data; and they are of equal importance to represent the evaluation time and the unit of the KPIs. In any case, each KPI and observation has an associated object.

The infrastructure and the occupancy modules have been developed to represent the KPI-evaluated objects and the features of interest related to observations. The infrastructure module describes the power system resources and the buildings in the district, including their subclasses and components. Meanwhile, the occupancy module represents the occupants in the buildings and occupants' behaviour related to energy usage. Additionally, the objects in these two modules offer different parameters for observations, such as the building, occupancy, equipment and energy parameters. Such parameters are included in the domain parameter module; they can be treated as subclasses of the observed property from the observation module. Furthermore, in order to identify the external environment of the objects, the location module and the weather module are indispensable, among which the weather module provides the outdoor environmental parameters for observation.

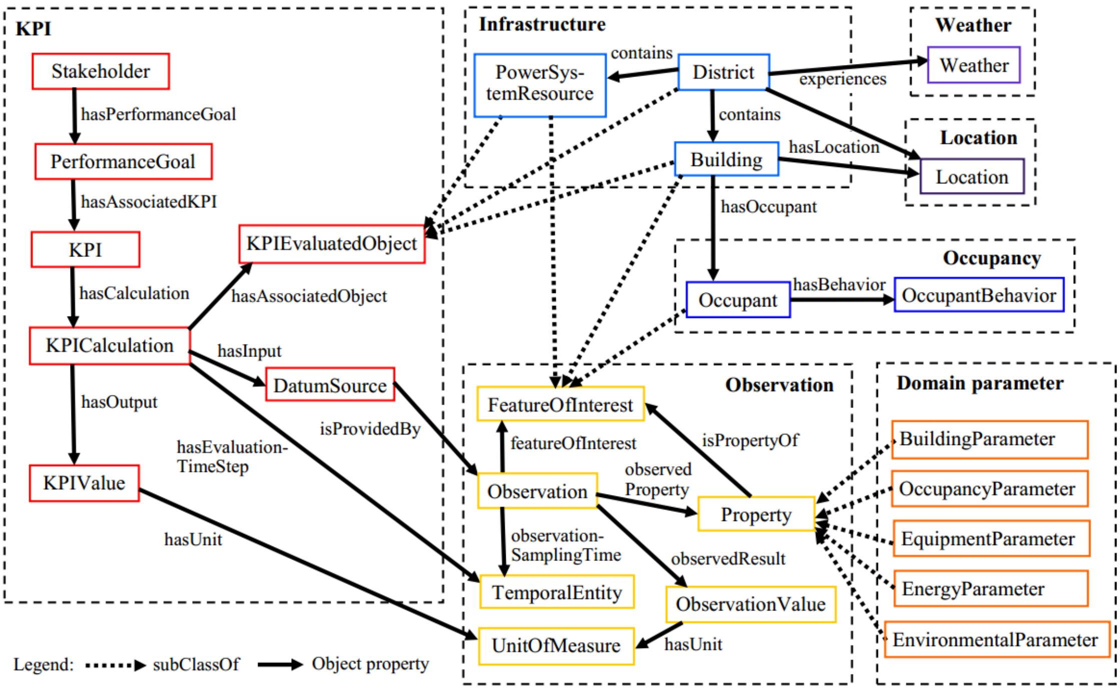

Figure 2: The initial conceptual model of the EM-KPI ontology.

The following ontology search and selection for implementing the ontology uses this conceptual model as a guide. Following this, a detailed model of each module has been developed.

3.4. Ontology Search and Selection

Some widely-known ontologies in related domains contain classes and/or properties that can be reused in the EM-KPI ontology. A literature review and online search of existing ontologies in regard to the related domains was carried out in order to find those that best fit the previously extracted terms. The online search was performed using the Google search engine, based on domain-specific keywords. In addition, the smart cities ontology catalogue developed in the READY4SmartCities project[1] was also used as a tool for searching reusable ontologies; this catalogue contains ontologies that describe the different domains in smart cities, including the building, facility and/or energy, among others. Furthermore, some widely-recognised ontologies were selected beforehand, since they are standard ones or are already well-known for describing certain classes and/or properties.

Table 3 lists the ontologies whose patterns and/or vocabulary have been reused for building the EM-KPI ontology, including their namespaces, prefixes and example terms. These ontologies are all available on the Web and the reason for their selction is offered below.

The DUL (DOLCE+DnS Ultralite) ontology includes descriptions and situations ontology; the purpose of its reuse is to provide the related upper-level concepts [51], including the class *PhysicalObject*, which represents any object that has a space region, and the object property *hasLocation*, which describes the spatial location of any entities. The data properties *identifier* and *title* for distinguishing each entity are reused from the widely-used Dublin Core Ontology [52], which also provides another concept: *Location*. To describe the location, a geographic coordinate point and/or a postal address can be used, whose vocabularies are, respectively, provided by the well-known WGS84 Geo Positioning Ontology and the website schema.org. The latter also contains a wide range of vocabularies for event description. To represent the weather related to the location, the specific Weather Ontology (WO) provides the reusable patterns and terms [53].

[1] http://smartcity.linkeddata.es/

Several ontologies have been found for building representation, such as the IFC2X3 – University of Ghent Ontology [54], the gbBuilding Information Ontology (BIO) [55], the Architecture and Building Physics Information Ontology [56] and the SimModel ontology [57]. Considering that the focus of the target ontology is the energy-related information, rather than the highly detailed building physics, the BIO ontology is a good choice. BIO provides a range of defined classes, axioms and datatypes for reuse, such as the *Building*, *BuildingElement* and *BuildingParameter*. In addition, the Energy and Resource Ontology (ERO) [58] is used to complete the energy information for buildings, since it provides various concepts for energy description, such as *EnergyParameter*, *EnergyFacility*, *EnergyType*, *EnergySupply* and *EnergyTariff*. With regard to the occupants in buildings, the concept *OccupancyParameter* is selected from the User Behaviour and Building Process Information Ontology (PO) [59], which is an ontology used to represent the behaviours and processes involved in smart home systems.

Table 3: Ontologies selected for the development of the EM-KPI ontology.

Ontology	**Namespace**	**Prefix**	**Example of term**
DUL ontology	http://www.ontologydesignpatterns.org/ont/dul/DUL.owl	dul	*PhysicalObject, hasLocation, isLocationOf*
Dublin Core ontology	http://purl.org/dc/terms/	dct	*identifier, title, description, type, Location*
WGS84 Geo Positioning Ontology	http://www.w3.org/2003/01/geo/wgs84_pos#	geo	*Point, lat, long, alt*
schema.org	http://schema.org/	schema	*Event, Postal Address*
gbBuilding Information Ontology	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/building/1_10/gbBuildingOntology.owl	bio	*Building, Building Element, Zone, containsArea, Area, containsVolume, Volume, BuildingStorey, Weather*
Energy Resource Ontology	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/EnergyResourceOntology.owl	ero	*EnergyFacility, Equipment, Appliance, consumesEnergy, producesEnergy, EnergySupply, EnergyDemand, EnergyType*
Weather Ontology	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/WeatherOntology.owl	wo	*WeatherCondition, WeatherPhenomenon, Humidity, SolarIrradiance*
User Behavior and Building Process Information	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/ProcessOntology.owl	po	*OccupancyParameter, hasInfluenceOn*
Semantic Sensor Network Ontology (SSN)	http://purl.oclc.org/NET/ssnx/ssn	ssn	*Observation, ObservationValue, observedProperty, Property, observationSamplingTime, observationResult*
Ontology of units of Measure (OM)	http://www.wurvoc.org/vocabularies/om-1.8/	om	*Unit_of_measure, Compound_unit, Singular_Unit, Unit_multiplication*
OWL-Time Ontology	http://www.w3.org/2006/time#	time	*Interval, hasEnd, hasBeginning, Instant*
Mathematical Modelling Ontology	http://identifiers.org/mamo/	mamo	*Mathematical_model, Variable, Independent_variable, Dependent_variable*

To describe the observation of various parameters, the widely recognised W3C Semantic Sensor Network (SSN) ontology [60] is selected. It provides a complete representation for observations, including terms such as *Observation*, *ObservationValue*, *FeatureOfInterest* and *Property*. However, the SSN ontology has not included the time and unit domains, which are intended to be imported from separate ontologies. The well-known OWL-Time Ontology [61] and the Ontology of Units of Measure (OM) [62] are used, respectively, to describe the observation time and the unit of observation value.

Most importantly, to represent the KPI calculation, the ontologies related to mathematical modelling have been searched. The Mathematical Modelling Ontology (MAMO) [63] provides concepts such as *Mathematical_model*, *Variable, Independent_variable* and *Dependent_variable*. However, MAMO is still unable to completely describe the KPI calculation. Therefore, the patterns of the Model ontology in the OntoMODEL (Ontological Mathematical Modeling Knowledge Management) ontology [64] are also reused. This ontology represents the different components of the mathematical model, including the equation, assumption, variables and constants.

Lastly, there are also some other ontologies which are not available on the Web but part of whose patterns have been used, such as the ontology to represent energy-related occupant behaviour in buildings, proposed by Tianzhen Hong et al. [65,66], the CIM ontology proposed by Neumann et al. [28] and the CIM extension of the microgrid energy management system proposed by Ming Ding et al. [67].

3.5. Ontology Implementation

The detailed EM-KPI ontology model is developed through restructuring the selected ontology resources using the ontology network proposed in Figure 2, and integrating it with the newly defined patterns and vocabularies (with the prefix *eko* in the following description). The subsections below illustrate each of the ontology modules in detail. The domain parameters will be described together with the related objects in order to make their relationship clearer. All of the elements are represented with the prefix before their names.

3.5.1.The KPI Module

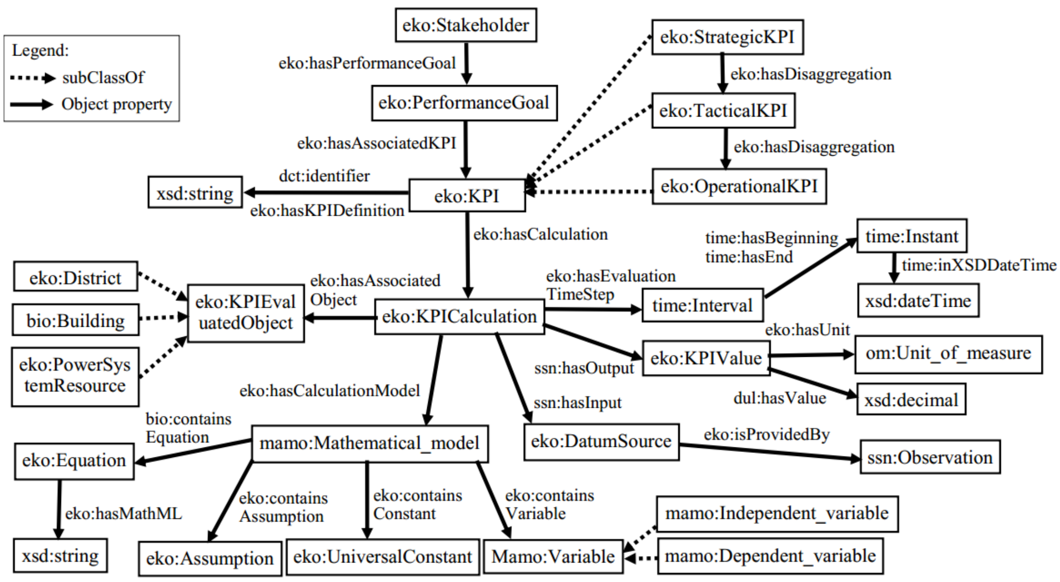

Figure 3: Detailed model of the KPI module.

Most of the classes and properties in the detailed model of the KPI module are presented with the prefix *eko*, as they are new contributions from this paper (Figure 3). A noticeable refinement of the initial model is that each KPI has an identifier and a definition for its distinction and interpretation. The class *KPI* is further divided into three subclasses, namely *StrategicKPI*, *TacticalKPI* and *OperationalKPI*, for representing multi-level key performance information. The *StrategicKPI* is the hyper-aggregated KPI at the district level, which can be disaggregated to the *TacticalKPI* at the building and system level, and then the *OperationalKPI* at the zone and equipment level. The model represents, in detail, the *KPICalculation* in terms of its calculation model, input, output, associated object and evaluation interval. The *Equation* of the

Mathematical_model is linked to a string of MathML, which is an XML language for describing mathematical expressions, and can be converted to Content MathML for calculation [68]. The time step for a KPI calculation is represented as an *Interval*, which has a beginning instant and an end instant in accordance with the Time Ontology [61]. The output of the KPI calculation is a *KPIValue*; and the input is the *DatumSource*, which is provided by the *Observation*.

3.5.2. The Observation Module

The SSN Ontology [60], Time Ontology [61] and OM [62] are reused to construct the detailed model of the observation module (Figure 4). Since the purpose of the EM-KPI ontology is to describe the master data rather than the sensors, only the observation aspects of the SSN Ontology are represented in the model. The reused terms include *Property, FeatureOfInterest* and *ObservationValue*. The feature of interest for the observation could be either districts, buildings, power system resources, occupants or their subclasses. Each observation corresponds to an observed *Interval* or *Instant*, which are subclasses of *TemporalEntity*. The units of observation values are classified based on OM [62]. Most importantly, the observed properties include the five types of domain parameters as mentioned previously, which are needed for the KPI calculation and performance analysis.

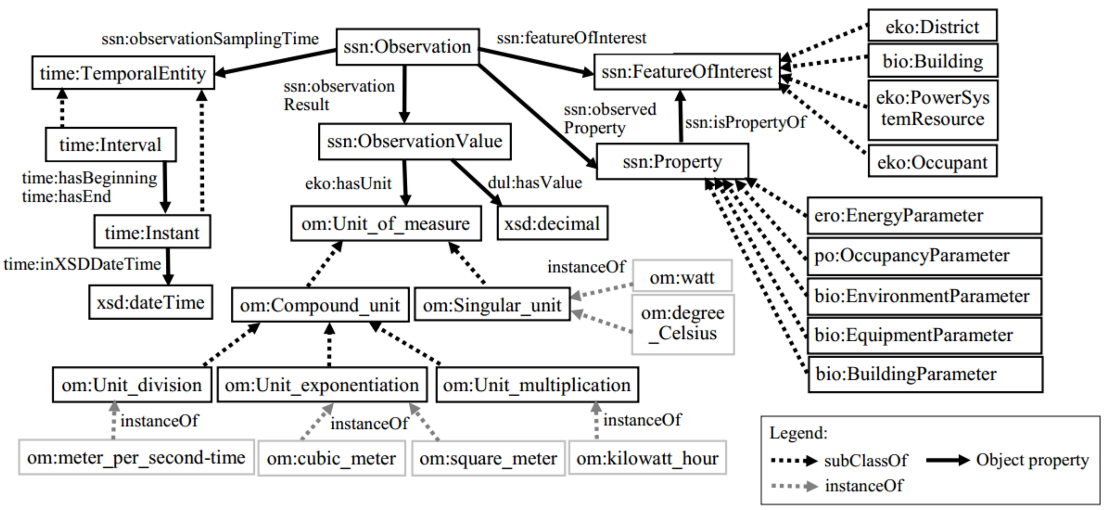

Figure 4: Detailed model of the observation ontology module.

3.5.3. The Location Module

The detailed model of the location module (Figure 5) describes the districts, buildings and power system resources, with identifiers and locations. The location can be represented by a geographic coordinate which details the latitude, longitude and altitude, or a postal address which contains the country, region, postal code and street address.

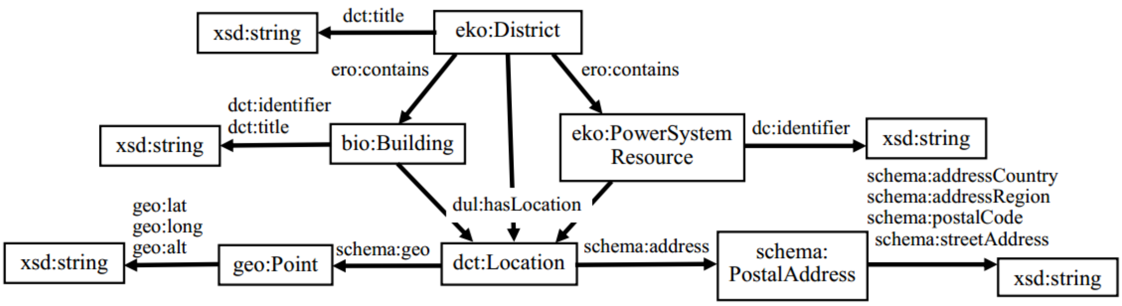

Figure 5: Detailed model of the location ontology module.

3.5.4. The Infrastructure Module

The power system resource in the infrastructure module can be classified into subclasses, as shown in Figure 6. The terms *PowerSystemResource* and *EquipmentContainer* are extracted from the Common Information Model (CIM) [69]. CIM is a series of standards developed by EPRI

(Electric Power Research Institute) for the exchange of power system networks and data between different organisations. It describes the components of the power system at the distribution level [70]. *EquipmentContainer*, *PowerEquipment* and *EnergyFacility* are defined as three main subclasses of *PowerSystemResource*. An equipment container is a group of equipment, such as a substation or distributed energy resources. Meanwhile, the *EnergyFacility* describes the energy systems that produce, store or consume energy in buildings; and it can be classified into *EnergyConsumerFacility* and *EnergyProducerFacility*. An energy facility may also contain different equipment. The class *PowerEquipment* has subclasses including *GeneratingUnit*, *StoringUnit, PowerDeliveryUnit* and *EnergyConsumer* [67], which describe equipment ranging from energy production, storage and supply to consumption. The most commonly used generating unit is the *PhotovoltaicGeneratingUnit*. The class *PhotovoltaicType* is defined, because the generating efficiency depends on the type of PV unit. In any case, each type of power equipment has its own equipment parameters.

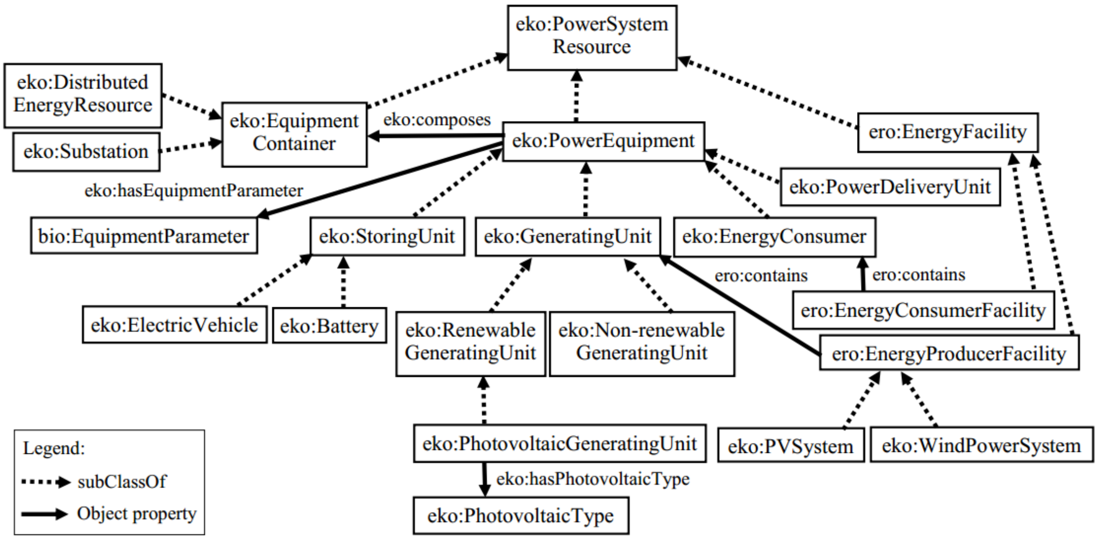

Figure 6: Detailed model of the power system resource in the infrastructure module.

Buildings are connected to the power system; the detailed model of the building is shown in Figure 7. Each building has a description of the building type, the year of construction and building parameters such as *Area* and *Volume*, since they are all related to the energy use in buildings. Energy facilities in a building include the *HVACSystem*, *LightingSystem*, *Appliance* and others; they may contain various building equipment. Building equipment could be, but does not necessarily have to be, power equipment. The description of *EquipmentParameter* is provided by BIO [55], including *Capacity*, *Efficiency*, *InputWatts* and *Power*. Building elements such as *BuildingStorey*, *Zone* and *Room* are also reused from BIO. The zone, here, refers to a building thermal zone, whose heating/cooling load is related to the thermal properties of the building envelope. Therefore, the classes *ExtWall*, *ExtWindow*, *Floor* and *Roof* are defined; and the thermal property of the envelope is represented as a *U-value*, which is a subclass of *BuildingParameter*. Finally, the occupants are those who use the zone; they are the attendees of the event taking place in the zone.

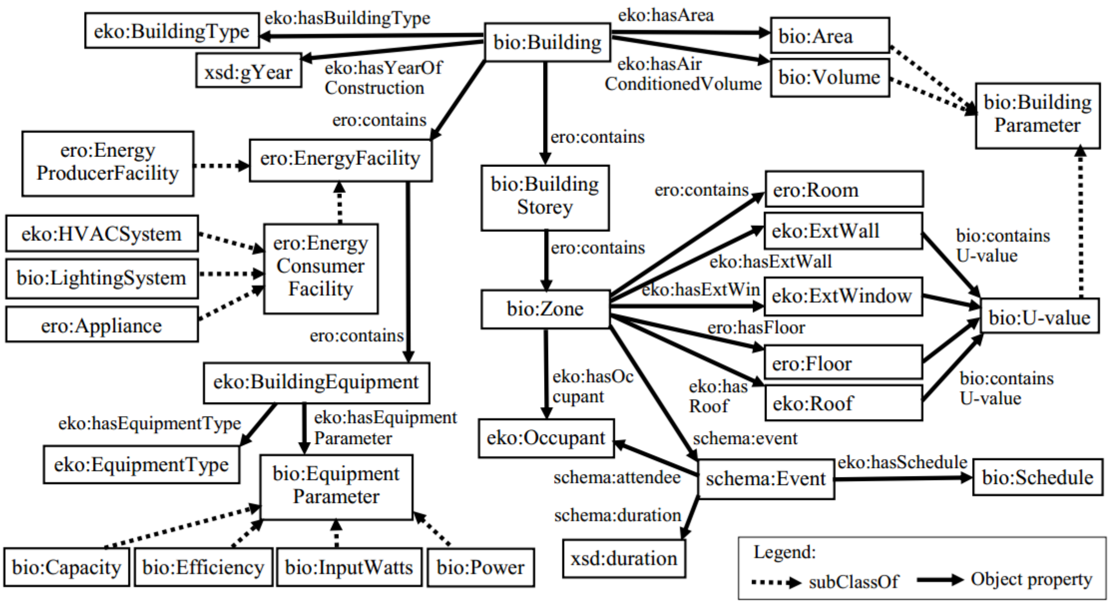

Figure 7: Detailed model of the building in the infrastructure module.

3.5.5. The Occupancy Module

In order to describe the occupancy parameters and the factors which influence occupants' behaviour, Figure 8 offers a detailed model of the occupancy module. The occupancy parameter concerned in this case only includes the occupant number, regardless of his/her gender, age, etc. Some patterns of this model are extracted from the ontology proposed by Tianzhen Hong et al. [66]. The main drivers of occupants' behaviour include the event and the level of indoor comfort existing in the zone. Furthermore, the outdoor weather influences the level of indoor comfort, so it indirectly impacts the occupants' behaviour. The occupants' behaviour has consequences for the energy consumer facilities, thus influencing the energy demand. The *IndoorComfort* and *Weather* are the subclasses of *EnvironmentalParameter*, among which *IndoorComfort* is further divided into *ThermalComfort*, *VisualComfort* and *IndoorAirQuality*.

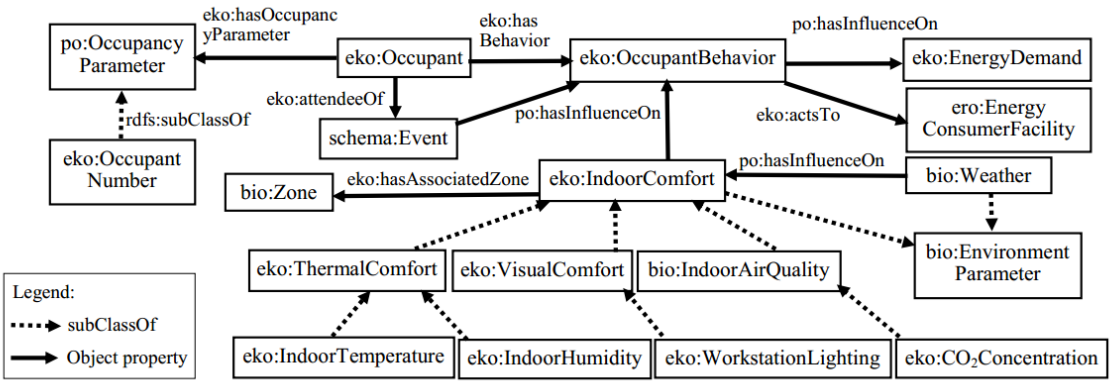

Figure 8: Detailed model of the occupant ontology module.

3.5.6. The Weather Module

A detailed model of the weather module is depicted in Figure 9. It mainly reuses the concepts and patterns of the Weather Ontology [53]. The *Weather* class is defined here to describe different types of outdoor environment, and is divided into two subclasses, *WeatherCondition* and *WeatherPhenomenon*. The weather condition describes conditions such as rain, snow and sun, and the weather phenomenon includes outdoor temperature, humidity, wind and solar irradiance. *SolarIrradiance* is classified into *GlobalHorizontalSolarIrradiance* and *GlobalInclinedSolarIrradiance*, because one of these represents the horizontal solar radiation, while the other one represents the strongest solar radiation at the location of application. The weather has a direct influence on energy production, and the weather forecast influences the energy forecast in districts and buildings, thus influencing energy supply.

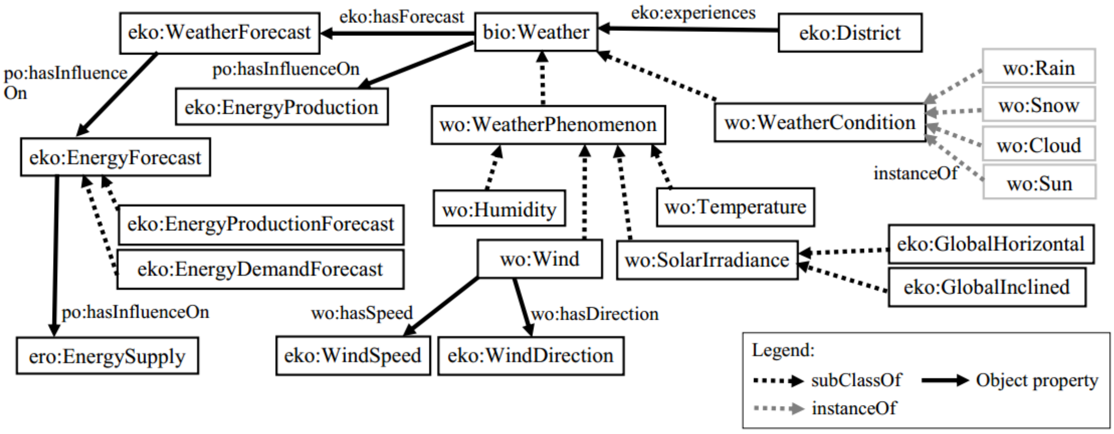

Figure 9: Detailed model of the weather module.

3.5.7. The Energy Parameter Module

Finally, Figure 10 illustrates the detailed model related to the energy parameters. The model describes the energy type produced, stored, delivered and consumed. To describe the various aspects of energy, the class *EnergyParameter* is divided into subclasses, including *EnergyProduction*, *EnergyStorage*, *EnergySupply*, *EnergyDemand*, *EnergyGain*, *EnergyCost*, *EnergyTariff* and *EnergyForecast*. The class *EnergySupply* is further classified into the subclasses *EnergySupplyFromStorage*, *EnergySupplyFromSiteProduction* and *EnergySupplyFromExternal Grid*, which respectively represent the energy suppliers of off-peak storage, the site-renewable resources and the external grid. The *EnergyGain* class describes the surplus energy that could be sold to the external grid. Therefore, the *EnergyTariff* class is divided into *EnergyCostTariff* and *FeedInTariff*, which respectively refer to the tariff for purchasing energy from, and for selling energy to, the external grid. Lastly, the energy cost depends on the energy tariff, energy gain and energy supply from the external grid.

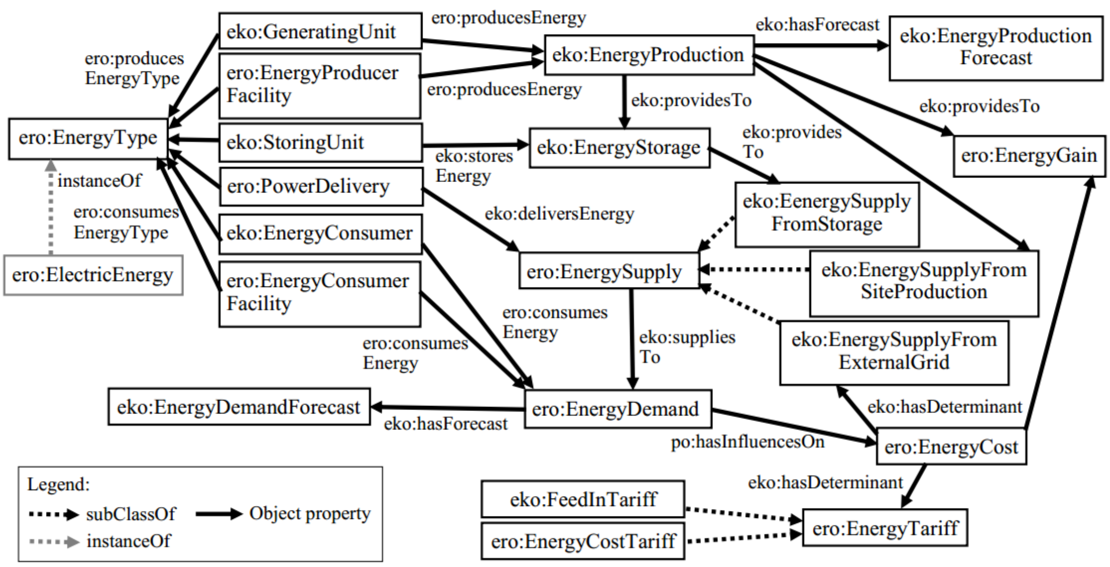

Figure 10: Detailed model of the energy parameter in the domain parameter module.

Once completing the detailed model of the ontology modules, the EM-KPI ontology is implemented in OWL. Its final version has been evaluated using the OOPS! ontology pitfall scanner[1] to ensure that no modelling or reasoning problem exists in the ontology. In addition, the

[1] OOPS! OntOlogy Pitfall Scanner!, http://oops.linkeddata.es/response.jsp#.

ontology has been published for open access, which is available at http://energy.linkeddata.es/em-kpi/ontology.

4. Demonstration

An example using the ontology has been constructed to demonstrate the feasibility of the ontology. The example dataset represents a real scenario from the Solar Decathlon Europe (SDE) 2012[1], which took place in Madrid, Spain. SDE 2012 is an international competition of solar houses, which is an initiative committed to sustainable buildings. From September 14 to 30, 19 teams from 12 countries participated in the event and built their houses in Villa Solar, Madrid. The established district called Villa Solar contained 19 solar houses (though one of the houses quit the competition at the end) and five service buildings. Each solar house was equipped with photovoltaic panels, high energy-efficient measures and energy management systems in order to achieve zero-energy buildings. All of the buildings were connected to a microgrid; both the buildings and microgrid were monitored from September 17 to 28, 2012 but initially without information exchange among the different domains and scales.

In this section, the EM-KPI ontology is used to facilitate the interchange of multi-level key performance information and cross-domain master data. The key stakeholders identified in this case include a microgrid system company, district energy engineers, building owners, building energy managers, occupants, and an organising committee [11]. The original data sources gathered from the different stakeholders are all stored in Excel, including the list of KPIs, building static data, energy demand data, energy production data, energy supply data and indoor comfort data, as well as the outdoor weather data. The district energy supply data were measured and collected separately, not through the aggregation of building-level data. Generally, the monitored data have good quality and consistency, due to its purpose for the competition; for this reason, the definition of methods to deal with potential data inconsistency issues is beyond the scope of this work.

The process adopted to map the data sources to the EM-KPI ontology and to generate the linked dataset is shown in Figure 11. The distributed data sources should be firstly converted into RDF (Resource Description Framework) [71] in order to instantiate the ontology. The tool OpenRefine is used to clean and transform the Excel data, to create the instances for the classes and to assign values to the properties for each instance. Afterwards, the converted RDF data and the ontology are gathered together in a triple store. The tool RDF4j is used to integrate the various RDF files and to query the linked dataset. Support services for the calculation of KPIs and the data visualisation are needed in order to enable the energy performance analysis.

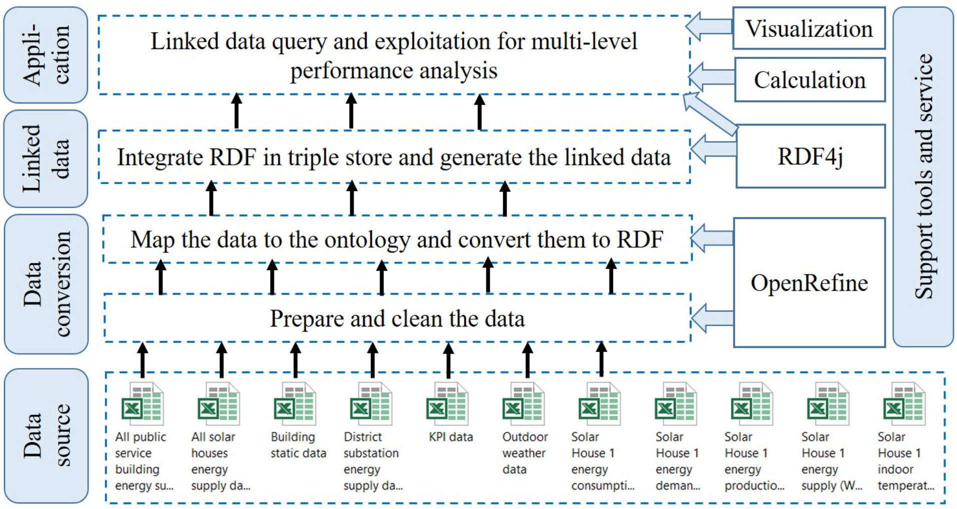

Figure 11: The process used to map the data sources to the ontology and generate linked data.

[1] Solar Decathlon Europe 2012, http://www.sdeurope.org/.

Three representative KPIs related to the energy balance have been studied, namely I03 (district energy balance), I05 (individual building energy balance) and I07 (time correlation between generation and consumption). This involves two main groups of stakeholders, namely district energy engineers and building energy managers. The related information about such KPIs and their associated master data sources were mapped to the ontology. In total, 27 Excel data files were converted.

Listing 1, below, shows an RDF snippet of the generated linked data that describes I03, I05 and I07, and which is written in Turtle. The snippet shows that I03 is a strategic KPI and that it is defined as the district energy balance between energy generation and consumption during a given time step. In addition, I05 and I07 are tactical KPIs which are disaggregated from I03. The stakeholders, including the district energy engineers and the building energy managers, have a performance goal – i.e., energy self-sufficiency – and this performance goal has the three associated KPIs.

```
@prefix dct: <http://purl.org/dc/terms/>.
@prefix eko: <http://energy.linkeddata.es/em-kpi/ontology#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
eko:I03 a eko:StrategicKPI;
        dct:identifier "District energy balance "^^xsd:string;
        eko:hasKPIDefinition "District energy balance between generation and consumption
during given time step (TS)"^^xsd:string.
eko:I05 a eko:TacticalKPI.
eko:I07 a eko:TacticalKPI.
eko:I03 eko:hasDisaggregation eko:I05 , eko:I07.
eko:EnergySelfsufficiency a eko:PerformanceGoal;
        eko:hasAssociatedKPI eko:I03. eko:I05, eko:I07.
eko:DistrictEnergyEngineers a eko:Stakeholder;
        eko:hasPerformanceGoal eko:EnergySelfsufficiency.
eko:BuildingEnergyManagers a eko:Stakeholder;
        eko:hasPerformanceGoal eko:EnergySelfsufficiency.
```

Listing 1: An RDF snippet in Turtle relating to the KPIs I03, I05 and I07.

To enable a deeper analysis of the energy balance performance, the calculation of the representative KPIs has been requested using the SPARQL query language [72], including their mathematical model, the related datum sources, the calculated value, the evaluated objects and the evaluation time step. The mathematical model and datum sources for the calculation of I03 are retrieved as Listing 2.

```
PREFIX eko: <http://energy.linkeddata.es/em-kpi/ontology#>
PREFIX ssn: <http://purl.oclc.org/NET/ssnx/ssn#>
PREFIX bio:
<https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/gbBuildingOntology.owl#>
SELECT DISTINCT ?mathematicalModel ?equation ?mathML ?datumSource
WHERE
{ eko:I03 eko:hasCalculation ?i03Calculation.
  ?i03Calculation ssn:hasInput ?datumSource.
  ?i03Calculation eko:hasCalculationModel ?mathematicalModel.
  ?mathematicalModel bio:containsEquation ?equation.
  ?equation eko:hasMathML ?mathML. }
```

Listing 2: The SPARQL query concerning the mathematical model and datum sources for I03 calculation.

I03 is calculated using *eko:Equation_I03*, whose MathML, after being translated, is presented as Equation 1, where $P_{District\ generation}$ and $P_{District\ consumption}$, respectively, represent the power generated and consumed in the district. Since I03 reflects the difference between the energy generation and consumption during the time step (TS), if the value is 0, district energy balance is achieved. If the value is positive, surplus energy is exported to the external grid. Otherwise, the energy balance performance target is not achieved.

Equation 1: $I03 \quad = \int_0^{TS} P_{District\ generation}\, dt - \int_0^{TS} P_{District\ consumption}\, dt\ (kWh)$

The collective of data sources, *eko:DatumSource_EnergySupplyExternalkW_VillaSolar*, is the input for the calculation of I03. The data represent the external energy supply in the Villa Solar; i.e. ($P_{District\ consumption}$ - $P_{District\ generation}$). Such data sources are provided by observations, which can be retrieved through a simple SPARQL query, as illustrated in Listing 3, and visualised as Figure 12. The observation time is represented by the *xsd:dateTime* format “CCYY-MM-DDThh:mm:ss”.

```
PREFIX eko: <http://energy.linkeddata.es/em-kpi/ontology#>
PREFIX ssn: <http://purl.oclc.org/NET/ssnx/ssn#>
PREFIX dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>
PREFIX time: <http://www.w3.org/2006/time#>
SELECT ?x ?instant ?dateTime ?value ?unit
WHERE
{ ?x eko:providesTo eko:DatumSource_EnergySupplyExternalkW_VillaSolar.
  ?x ssn:observationResult ?observationValue.
  ?observationValue dul:hasValue ?value.
  ?observationValue eko:hasUnit ?unit.
  ?x ssn:observationSamplingTime ?instant.
  ?instant time:inXSDDateTime ?dateTime }
ORDER BY ?instant
```

Listing 3: The SPARQL query for the observations that provide the input data sources.

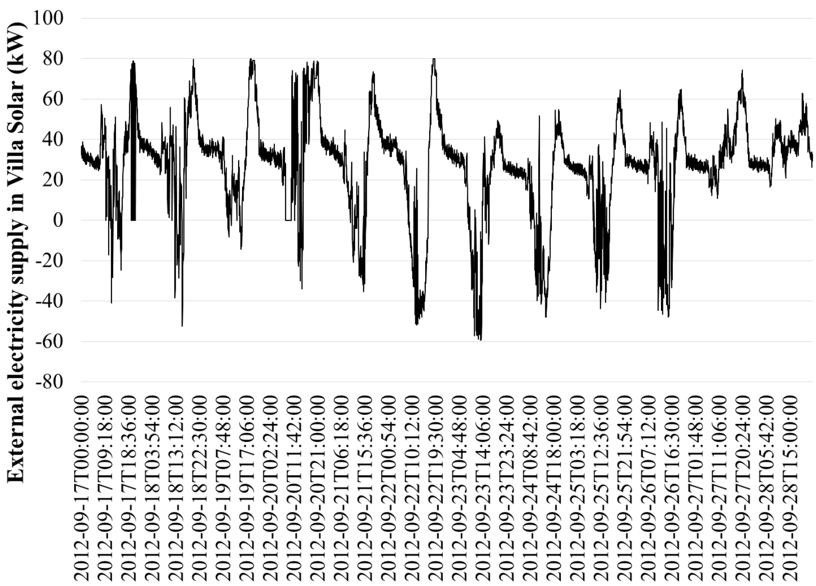

Figure 12: The visualisation of the observation data for external electricity supply in Villa Solar.

Table 4: The query results of the I03 calculation, including the evaluation interval, value and unit.

i03Calculation	**interval**	**value_I03**	**unit**
eko:Calcu01_I03_VillaSolar	eko:Interval20120917	-747.535	eko:kilowatt_hour
eko:Calcu02_I03_VillaSolar	eko:Interval20120918	-802.770	eko:kilowatt_hour
eko:Calcu03_I03_VillaSolar	eko:Interval20120919	-787.730	eko:kilowatt_hour
eko:Calcu04_I03_VillaSolar	eko:Interval20120920	-857.385	eko:kilowatt_hour
eko:Calcu05_I03_VillaSolar	eko:Interval20120921	-628.870	eko:kilowatt_hour
eko:Calcu06_I03_VillaSolar	eko:Interval20120922	-458.180	eko:kilowatt_hour
eko:Calcu07_I03_VillaSolar	eko:Interval20120923	-359.555	eko:kilowatt_hour
eko:Calcu08_I03_VillaSolar	eko:Interval20120924	-310.020	eko:kilowatt_hour
eko:Calcu09_I03_VillaSolar	eko:Interval20120925	-483.635	eko:kilowatt_hour
eko:Calcu10_I03_VillaSolar	eko:Interval20120926	-495.795	eko:kilowatt_hour
eko:Calcu11_I03_VillaSolar	eko:Interval20120927	-827.310	eko:kilowatt_hour
eko:Calcu12_I03_VillaSolar	eko:Interval20120928	-845.055	eko:kilowatt_hour

In order to evaluate I03 during the monitored day, the outputs of the I03 calculation have also been requested. The query result is listed in Table 4 and visualised in Figure 13. *Interval20120917* refers to the date 17th September. The negative value of *value_I03* implies that the energy balance performance target was not achieved in Villa Solar. The reason underlying the undesirable performance can be analysed with the input data sources, as shown in Figure 12. It can be understood that the power consumed was generally higher than the power generated, due to the positive value of the energy supply, especially on the 17^{th}, 18^{th}, 19^{th} and 20^{th}, which had low surplus power but high energy supply. Moreover, on the 27^{th} and 28^{th}, there was no generated power exported to the external grid.

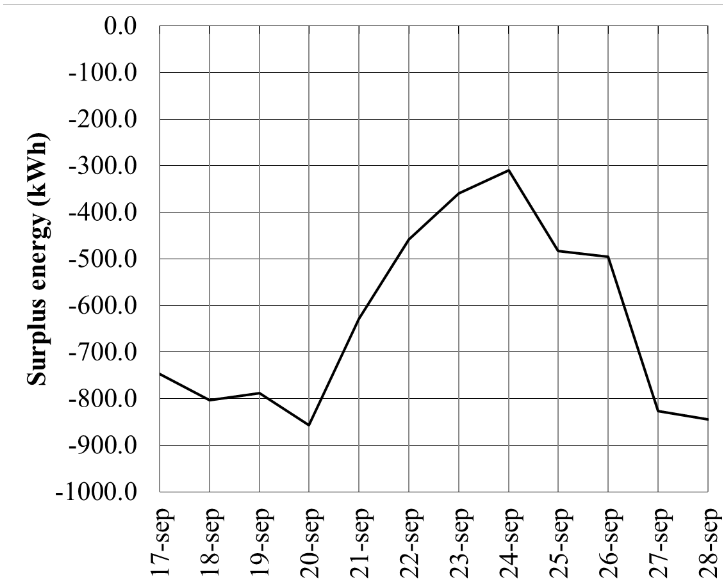

Figure 13: I03 district energy balance showing a constant need for energy to be supplied from the grid

A further analysis of the factors which influence energy production has been conducted to ascertain the reason underlying the low level of energy generation on the 27^{th} and 28^{th}. The SPARQL query concerning the factors is detailed in Listing 4. It has been found that global solar irradiance has a direct influence on energy production. Therefore, the observation data of the global solar irradiance were retrieved, and visualised as Figure 14. It shows that the solar irradiance in *Interval20120927* and *Interval20120928* was very low, which was unfavourable for energy production, and thus led to poor energy balance performance on these two days.

```
PREFIX eko: <http://energy.linkeddata.es/em-kpi/ontology#>
PREFIX po:
<https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/ProcessOntology.owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT DISTINCT ?influenceFactor
WHERE
{ ?influenceFactor po:hasInfluenceOn ?energyProduction.
  ?energyProduction rdf:type eko:EnergyProduction  }
```

Listing 4: The SPARQL query concerning the factors which influence energy production.

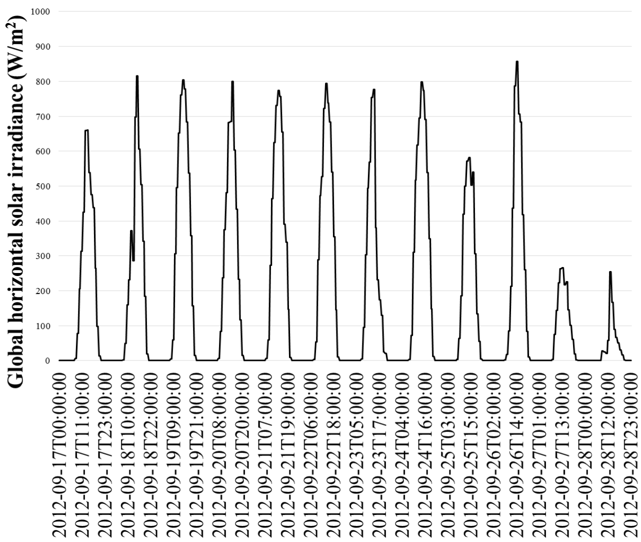

Figure 14: Global horizontal solar irradiance at Villa Solar.

Since solar houses are the energy producers in the district, the disaggregated indicator I05 (individual building energy balance) has also been explored to examine the energy balance performance in each house. Through a SPARQL query, the calculation results of I05 associated with each solar house are retrieved and compared, as shown in Figure 15. The comparison shows that Solar House 2 had the worst energy balance performance. In order to determine the reason, the observation data of the energy production in Solar House 2 is queried as Listing 5, and visualised as Figure 16. It can be understood that the highest power generated in Solar House 2 was 324 W, which was insufficient to cover the energy demand. Excluding the weather factors, the low level of generated power may be due to the problems of the energy producer facilities in Solar House 2.

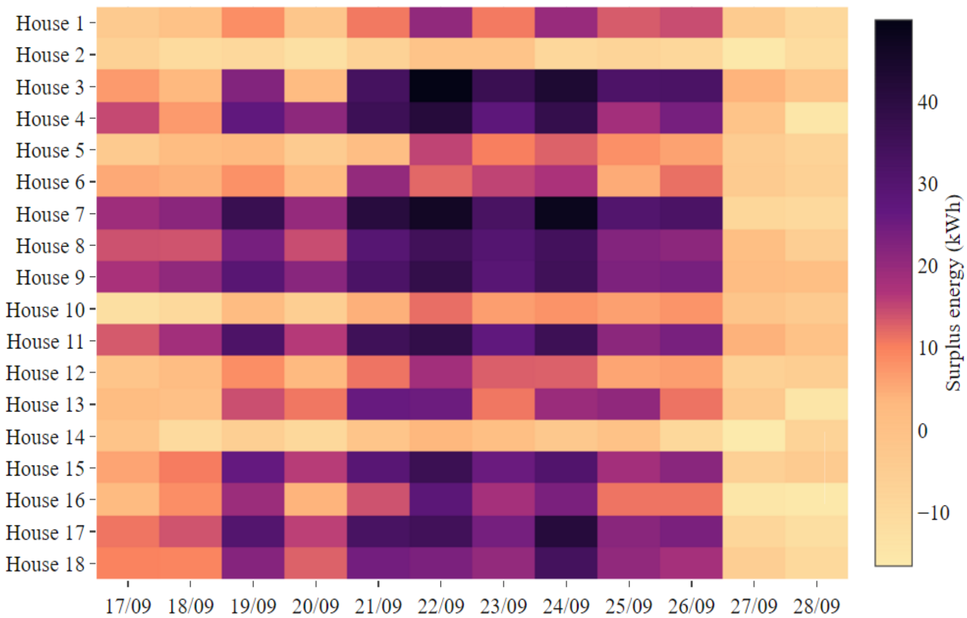

Figure 15: I05 building energy balance

```
PREFIX ssn: <http://purl.oclc.org/NET/ssnx/ssn#>
PREFIX eko: <http://energy.linkeddata.es/em-kpi/ontology#>
PREFIX dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>
PREFIX time: <http://www.w3.org/2006/time#>
SELECT DISTINCT ?dateTime ?energyProduction ?unit
WHERE
{ ?observation ssn:observedProperty eko:EnergyProductionW_Solarhouse_2.
  ?observation ssn:observationSamplingTime ?instant.
```

```
    ?instant time:inXSDDateTime ?dateTime.
    ?observation ssn:observationResult ?observationValue.
    ?observationValue dul:hasValue ?energyProduction.
    ?observationValue eko:hasUnit ?unit. }
ORDER BY ?dateTime
```

Listing 5: The SPARQL query concerning the observation data of energy production in Solar House 2.

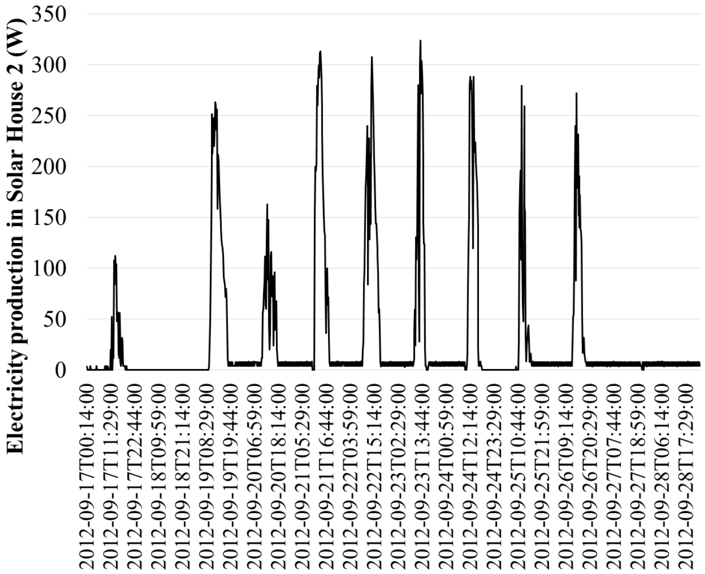

Figure16: The visualisation of the observation data of electricity production in Solar House 2.

The analysis through the linked data found the following: a PV system generated electrical energy in Solar House 2; the PV system contained a microcrystalline silicon PV unit; the equipment parameter of the PV unit includes the installed power, whose value is 11.35 kW. Comparing the installed power, 11.35 kW, with the actual highest generated power, 0.324 kW, it can be inferred that there were some problems or a system failure in the PV unit in Solar House 2, which resulted in such a low level of generating efficiency.

5. Discussions

The exploration and analysis of the generated linked data for Villa Solar shows the feasibility of the EM-KPI ontology, particularly for exchanging multi-level key performance information and cross-domain master data. The ability to access related information about the district energy balance (I03) combined with individual building energy balances (I05) proves that stakeholders can retrieve relevant performance information at different levels. Furthermore, the stakeholders can also access and exploit the master data from various domains. In this case study, the retrieved master data includes data relating to district level energy supply, weather, multi-level KPI calculations, building energy production and building facility data. This combined dataset spans different domains and scales, which are normally heterogeneous and stored in isolation. Thus, it is quite difficult to exchange this data amongst different stakeholders in the absence of a linked data approach. In addition, the query of observation data that provides the datum sources for the I03 calculation shows the effectiveness of the link between the KPIs and related master data. Such links facilitate performance tracking and analysis along with performance problem identification.

SPARQL queries for the linked dataset can be predefined and saved in advance, thus empowering stakeholders who need to access the specific information. Such stakeholders can simply enter linked dataspaces with their user authorisation and execute predefined queries; thus enabling the use of linked datasets by individuals without expertise in linked data.

The developed ontology facilitates energy performance tracking and improvement analysis at both district and building levels. However, electricity is the only energy type currently represented. Furthermore, the infrastructure module contains the power system resources, including the equipment and systems used for electricity generation, storage, distribution and consumption; while other possible energy utilities, such as the district heating network, gas delivery network and/or combined heat and power, are not included. Although electricity is one of the primary energy types used in buildings, other energy types especially gas and thermal energy are also present. It is therefore essential to optimise the use of all types of consumed or generated energy in order to enable thorough and robust multi-level energy management, it. An extension of the current ontology to describe gas and thermal energy is needed.

The current ontology only represents the required domain parameters for the described use case. For instance, building level parameters are limited to area, volume and U-value, and the occupancy parameter is restricted to occupant number. In real-world applications, many other parameters may be required to support integrated energy management. Therefore, another improvement would be to include a more comprehensive list of domain parameters.

The practice of reusing ontologies is encouraged within in the ontological engineering community but this practice also creates a dependency between the EM-KPI ontology and the ontologies reused within; a relationship that could change over time. These changes would have: (1) a minimal impact on the KPI, power system and occupancy modules, since their main design patterns and terms are defined anew; (2) a moderate impact on the building, energy parameters and weather modules, since they only reuse terms; and (3) a high impact on the observation module, which is predominantly built on existing ontologies. In any case, the EM-KPI ontology depends on specific versions of the reused ontologies; and in the case where a new version of some ontologies appears, a thorough analysis of the impact of such versions on current development would be performed prior to adoption.

6. Conclusions and Further Work

The exchange, sharing and exploitation of multi-level energy performance information and data from different stakeholders help to improve energy performance and to achieve stakeholders' performance goals. However, the main barriers are the interoperability problems associated with heterogeneous data and the vast amount of information involved. Therefore, an EM-KPI ontology is proposed in this paper to facilitate the interchange of key performance information and insightful data among different stakeholders.

The conventional mechanism used for the exchange of heterogeneous data is usually ineffective and time-consuming. The use of linked data in this paper provides an efficient means to facilitate the data interchange, using ontologies as the foundation. The majority of existing ontologies related to energy management focus on enabling system control rather than generating linked data. Additionally, the ontologies, including those for performance assessment, are usually designed for individual buildings. Thus, the EM-KPI ontology aims to integrate performance information and data for both the demand and supply sides in a district in order to enhance multi-level energy management.

The developed ontology only represents the key performance information and the key data that underpin stakeholders' performance goals which include a KPI (key performance indicator) component and an EM (energy management) master data component. The KPI component enables the interchange of multi-level key performance information, while the master data component facilitates cross-domain sharing of insightful data. The stakeholders who use the ontology can not only exchange and obtain access to their relevant performance information, but can also track and analyse energy performance related to their respective goals.

The linked data generated by the ontology provides a novel mechanism that engages different stakeholders in energy management. This is demonstrated by the querying and analysis of the linked dataset example. Additionally, the demonstration illustrates how to leverage the ontology for the generation of linked data and how to link multi-level KPIs and master data. Most importantly, the sharing of the linked dataset enables cross-domain analysis that identifies meaningful insights for energy performance improvement.

To generate a linked dataset, data preparation and curation is one of the most time-consuming processes, especially when seeking to exploit a considerable amount of data. In our use case, these tasks have mostly been performed manually. However, in order to apply the approach on a large scale, other techniques or tools to prepare and curate the data automatically or semi-automacially are more favorable, and we will research this in future work.

Another requirement in order to use the EM-KPI onology to generate linked data in real-world applications is that of scalability. Dealing with distributed district and building data on a large scale imposes new hardware and software requirements that may not be satisfied by the current approach used in our case study, and which is also part of future research.

Acknowledgements
We acknowledge China Scholarship Council's funding of Li Yehong to undertake doctoral studies in Universidad Politécnica de Madrid, Spain. The work in China was financially supported by Guangzhou Science and Technology Innovation Talent Project, Project No: 201710010156; Guangzhou energy saving special fund project, J-2016-14. The data access granted by Solar Decathlon Europe 2012 and Schneider Electric Madrid is highly appreciated. The authors would like to acknowledge the help offered by Shushan Hu in the UCD Energy Institute. Other parts (no duplications) of the preliminary research work in Ireland were supported by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme, project title SuPerB, project number 631617.

References

[1] U. Berardi, A cross-country comparison of the building energy consumptions and their trends, Resource, Conservation and Recycling, 123 (2017), pp 230–241. doi:10.1016/j.resconrec.2016.03.014.

[2] C.A. Balaras, A.G. Gaglia, E. Georgopoulou, S. Mirasgedis, Y. Sarafidis, D.P. Lalas, European residential buildings and empirical assessment of the Hellenic building stock, energy consumption, emissions and potential energy savings, Building and Environment, 42 (2007), pp 1298–1314. doi:10.1016/j.buildenv.2005.11.001.

[3] R. Madlener, Y. Sunak, Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management?, Sustainable Cities and Society, 1 (2011), pp 45–53. doi:10.1016/j.scs.2010.08.006.

[4] M. Manfren, P. Caputo, G. Costa, Paradigm shift in urban energy systems through distributed generation : Methods and models, Applied Energy, 88 (2011), pp 1032–1048. doi:10.1016/j.apenergy.2010.10.018.

[5] S. Lee, B. Kwon, S. Lee, S. Member, Joint Energy Management System of Electric Supply and Demand in Houses and Buildings, IEEE Transactions on Power Systems, 29 (2014), pp 2804–2812. doi: 10.1109/TPWRS.2014.2311827.

[6] SunGard solutions, Big Data - Challenges and Opportunities for the energy industry, (2013). https://www.sungard.com/~/media/fs/energy/resources/white-papers/Big-Data-Challenges-Opportunities-Energy-Industry.ashx (accessed January 28, 2017).

[7] Strategic R&D Opportunities for the Smart Grid: Advancing measurement science and standards for smart grid technologies, (2013). http://www.nist.gov/smartgrid/upload/Final-Version-22-Mar-2013-Strategic-R-D-Opportunities-for-the-Smart-Grid.pdf (accessed January 29, 2017).

[8] M. Donnelly, Building Energy Management: Using Data as a Tool, (2012). http://www.institutebe.com/InstituteBE/media/Library/Resources/Existing Building Retrofits/Using-Building-Data-as-a-Tool.pdf (accessed January 29, 2017).

[9] A. Dreibelbis, E. Hechler, I. Milman, M. Oberhofer, P. van Run, D. Wolfson, Enterprise Master Data Management: An SOA Approach to Managing Core Information, IBM Press. ISBN 978-0-13-236625-0.

[10] D. Loshin, D. Loshin, Chapter 1 – Master Data and Master Data Management, in: Master Data Management, (2009), pp 1–21. doi:10.1016/B978-0-12-374225-4.00001-1.

[11] Y. Li, J. O'Donnell, R. García-Castro, S. Vega-Sánchez, Identifying stakeholders and key

performance indicators for district and building energy performance analysis, Energy and Buildings, 155 (2017), pp 1–15. doi.org/10.1016/j.enbuild.2017.09.003.

[12] J.C. Van Gorp, Using Key Performance Indicators to Manage Energy Costs, Strategic Planning for Energy and the Environment, 25 (2005), pp 9–25. doi:10.1080/10485230509509683.

[13] E. Curry, J. O'Donnell, E. Corry, S. Hasan, M. Keane, S. O'Riain, Linking building data in the cloud: Integrating cross-domain building data using linked data, Advanced Engineering Informatics, 27 (2013), pp 206–219. doi:10.1016/j.aei.2012.10.003.

[14] T. Heath, C. Bizer, Linked data: Evolving the Web into a global data space (1st edition), Synthesis Lectures on the Semantic Web: Theory and Technology 11, (2011). doi:10.2200/S00334ED1V01Y201102WBE001.

[15] E. Corry, J. O'Donnell, E. Curry, D. Coakley, P. Pauwels, M. Keane, Using semantic web technologies to access soft AEC data, Advanced Engineering Informatics, 28 (2014), pp 370–380. doi:10.1016/j.aei.2014.05.002.

[16] 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), pp 223–235. doi:10.1016/j.scs.2014.09.006.

[17] The NewTrend project. http://newtrend-project.eu/ (accessed March 4, 2017).

[18] The OptEEmAL project. https://www.opteemal-project.eu/ (accessed March 4, 2017).

[19] F. Radulovic, R. García-Castro, M. Poveda-Vilalón, M. Weise, T. Tryferidis, READY4SmartCities - ICT Roadmap and Data Interoperability for Energy Systems in Smart Cities Deliverable D4.1 : Requirements and guidelines for energy data generation, (2014). http://xueshu.baidu.com/usercenter/paper/show?paperid=65e43d5886dd2bcb3e2ea6586cc13ac2&site=xueshu_se (accessed November 22, 2018).

[20] M. Weise, M. Poveda-Villalón, M.C. Suárez-Figueroa, R. García-Castro, J. Euzenat, L. Maria-Priego, B. Fies, A. Cavallaro, J. Peters-Andres, K. Zoi-Tsagkari, READY4SmartCities - ICT Roadmap and Data Interoperability for Energy Systems in Smart Cities Deliverable D2.2: Ontologies and datasets for energy management system interoperability v1, (2014). https://hal.archives-ouvertes.fr/hal-01180932 (accessed November 22, 2018).

[21] L. Ding, P. Kolari, Z. Ding, S. Avancha, Using ontologies in the semantic web: A survey, Ontologies, (2007), pp 79--113. doi:10.1007/978-0-387-37022-4_4.

[22] M. Grassi, M. Nucci, F. Piazza, Towards an ontology framework for intelligent smart home management and energy saving, in: 2011 IEEE International Symposium on Industrial Electronic (ISIE), Gdansk, Poland (2011), pp 1753–1758. doi:10.1109/ISIE.2011.5984327.

[23] M.J. Kofler, C. Reinisch, W. Kastner, A semantic representation of energy-related information in future smart homes, Energy and Building, 47 (2012), pp 169–179. doi:10.1016/j.enbuild.2011.11.044.

[24] H. Wicaksono, S. Rogalski, E. Kusnady, Knowledge-based Intelligent Energy Management Using Building Automation System, in: 2010 Conference Proceedings IPEC, Singapore (2010), pp 1140–1145. doi:10.1109/IPECON.2010.5696994.

[25] J. Han, Y. Jeong, I. Lee, Efficient Building Energy Management System Based on Ontology , Inference Rules , and Simulation, in: 2011 International Conference on Intelligent Building and Management, Singapore, (2011), pp 295–299. http://www.ipcsit.com/vol5/53-ICCCM2011-B0023.pdf (accessed November 22, 2018).

[26] J. Caffarel, S. Jie, J. Olloqui, R. Martínez, Bat-MP: An Ontology-Based Energy Management Platform, in: J. Bravo, D. López-de-Ipiña, F. Moya (Eds.), Ubiquitous Computing and Ambient Intellifence: 6th Internaltional Conference, Vitoria-Gasteiz, Spain, (2012), pp 9-16. doi:10.1007/978-3-642-35377-2_2.

[27] S. Rohjans, M. Uslar, H. Appelrath, OPC UA and CIM : Semantics for the Smart Grid, in: 2010 IEEE PES Transmission and Distribution Conference and Exposition, New Orleans, LA, USA (2010), pp 1–8. doi:10.1109/TDC.2010.5484299.

[28] S. Neumann, J.P. Britton, A. DeVos, S. Widergren, Use of the CIM Ontology,

Distribution, (2006), pp 26–32. http://uisol.com/uisol/papers/Use_of_the_CIM_Ontology_DistribuTech_2006_Paper.pdf (accessed July 10, 2017).

[29] K. Salameh, R. Chbeir, H. Camblong, A Generic Ontology-Based Information Model for Better Management of Microgrids, in: R. Chbeir, Y. Manolopoulos, I. Maglogiannis, R. Alhajj (Eds.), The 11th International Conference on Artificial Intelligence Applications and Innovations, Bayonne, France (2015), pp. 451–466. doi:10.1007/978-3-319-23868-5_33.

[30] K. Macek, K. Mařík, P. Stluka, Ontology-Driven Design of an Energy Management System, Computer Aided Chemical Engineering, 29 (2011), pp 2009–2013. doi:10.1016/B978-0-444-54298-4.50180-X.

[31] A. Anvari-moghaddam, J.M. Guerrero, A. Rahimi-Kian, M.S. Mirian, Optimal Real-time Dispatch for Integrated Energy Systems : An Ontology-Based Multi-Agent Approach, in: 7th International Symposium on Power Electroics for Distributed Generation Systems, Vancouver, Canada, (2016), pp 1–7. doi:10.1109/PEDG.2016.7526997.

[32] V. Custodio, J.I. Moreno, M. Sikora, P. Moura, N. Ferna, G. Lo, Modeling Smart Grid neighborhoods with the ENERsip ontology, Computers in Industry, 70 (2015), pp 168–182. doi:10.1016/j.compind.2015.01.008.

[33] P. Brizzi, D. Bonino, A. Musetti, A. Krylovskiy, E. Patti, M. Axling, Towards an ontology driven approach for systems interoperability and energy management in the smart city, in: International Multidisciplinary Conference on Computer and Energy Science 2016, Split, Croatia, (2016), pp 1–7. doi:10.1109/SpliTech.201607555948.

[34] B. Jayan, An ontological approach to energy management for buildings and their districts using artificial intelligence. http://www.bre.co.uk/filelibrary/BRE Trust docs/Cardiff/Bejay-Jayan-poster.pdf (accessed January 26, 2017).

[35] E. Corry, D. Coakley, J. O'Donnell, M.M. Keane, The role of Linked Data and the Semantic Web in Building Operation, in: Proceedings of the 13th International Conference for Enhanced Building Operations, Montreal, Quebec, (2013). http://hdl.handle.net/1969.1/151454 (accessed November 22, 2018).

[36] N. Shah, K.-M. Chao, T. Zlamaniec, A. Matei, Ontology for Home Energy Management Domain, in: H. Cherifi, J.M. Zain, E. El-Qawasmeh (Eds.), DICTAP 2011: Digital Information and Communication Technology and Its Application, Dijon, France, (2011), pp 337–347. doi:10.1007/978-3-642-22027-2_28.

[37] Q. Zhou, S. Natarajan, Y. Simmhan, V. Prasanna, Semantic Information Modeling for Emerging Applications in Smart Grid, in: 2012 Ninth International Conference on Information Technology - New Generations, Las Vegas, NV, USA (2012), pp 775–782. doi:10.1109/ITNG.2012.150.

[38] Y. Simmhan, Q. Zhou, V. Prasanna, Semantic Information Integration for Smart Grid Applications, Green IT: Technologies and Applications, (2011), pp 361–380. doi:10.1007/978-3-642-22179-8_19.

[39] S. Gillani, F. Laforest, G. Picard, A generic ontology for prosumer-oriented smart grid, in: CEUR Workshop Proceedings. 3rd Workshop on Energy Data Management at 17th International Conference on Extending Database Technology, Athènes, Greece, (2014), pp 134–139. https://www.hal.inserm.fr/OPENAIRE/emse-00948316 (accessed November 22, 2018).

[40] L. Gomes, M. Lefrancois, P. Faria, Z. Vale, Publishing Real-time Microgrid Consumption Data on the web of Linked Data, in: 2016 Clemson University Power Systems Conference, Clemson, SC, USA, (2016), pp 1–8. doi:10.1109/PSC.2016.7462861.

[41] J.-M. Bahu, A. Koch, E. Kremers, S.M. Murshed, Towards a 3D spatial urban energy modelling approach, International Journal of 3-D Information Modeling, 3 (2014), pp 1-16. doi: 10.4018/ij3dim.2014070101.

[42] Performance Information Model (PIM) ontology. http://smartcity.linkeddata.es/ontologies/www.bimtoolset.orgontologiesIntUBE-EnergyPIM.owl.html (accessed January 23, 2017).

[43] 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), pp 249–259. doi:10.1016/j.autcon.2015.05.002.

[44] J.J.V. Díaz, M.R. Wilby, A.B.R. González, J.G. Muñoz, EEOnt : An ontological model for a unified representation of energy efficiency in buildings, Energy and Buildings, 60 (2013), pp 20–27. doi:10.1016/j.enbuild.2013.01.012.

[45] D. Dahleh, M.S. Fox, An Environment Ontology for Global City Indicators (ISO 37120), Enterprise Integration Lab, University of Toronto, 2016. http://ontology.eil.utoronto.ca/GCI/ISO37120/Environment.owl (accessed November 22, 2018).

[46] M.C. Suárez-Figueroa, Neon methodology for building ontology networks: Specification, Scheduling and Reuse, Universidad Politécnica de Madrid, 2010. ISBN: 1614991154.

[47] BS EN 15221-3:2011 Facility Management. Guidance on quality in Facility Management, BSI, 2012. https://www.document-center.com/standards/show/BS-EN-15221-3 (accessed November 22, 2018).

[48] G. May, M. Taisch, V. Prabhu, I. Barletta, Energy Related Key Performance Indicators– State of the Art, Gaps and Industrial Needs, in: APMS 2013: Advances in Production Management Systems. Sustainable Production and Service Supply Chains, State College, PA, USA, (2013), pp 257–267. doi:10.1007/978-3-642-41266-0_32.

[49] OWL Web Ontology Language Overview. https://www.w3.org/TR/owl-features/ (accessed March 12, 2017).

[50] M. Gruninger, M.S. Fox, Methodology for the design and evaluation of ontologies, in: Workshop Notes of IJCAI-95, Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI-95), Montreal, Canada, (1995), pp 1–10. http://www.eil.utoronto.ca/wp-content/uploads/enterprise-modelling/papers/gruninger-ijcai95.pdf (accessed November 22, 2018).

[51] Ontology:DOLCE+DnS Ultralite, (2010). http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+DnS_Ultralite (accessed January 19, 2017).

[52] The Dublin Core Ontology, DCMI Usage Board. (2012). http://dublincore.org/documents/dcmi-terms/ (accessed January 20, 2017).

[53] Weather Ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/WeatherOntology.owl (accessed February 12, 2017).

[54] D. Van Deursen, IFC2X3 ontology, (2010). http://multimedialab.elis.ugent.be/organon/ontologies/IFC2X3# (accessed January 20, 2017).

[55] gbBuilding Information Ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/building/1_10/gbBuildingOntology.owl (accessed February 12, 2017).

[56] Architecture and Building Physics Information. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/BuildingOntologySharedVocabulary.owl (accessed January 20, 2017).

[57] P. Pauwels, E. Corry, J. O'Donnell, Representing SimModel in the Web Ontology Language, in: International Conference on Computing in Civil and Building Engineering, Orlando, Florida, USA, (2014). doi:10.1061/9780784413616.282.

[58] Energy and Resource Ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/EnergyResourceOntology.owl (accessed March 14, 2017).

[59] User Behavior and Building Process Information Ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/ProcessOntology.owl (accessed February 12, 2017).

[60] M. Compton, P. Barnaghi, L. Bermudez, R. García-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog, V. Huang, The SSN ontology of the W3C semantic sensor networks incubator group, Journal of Web Semantics, 17 (2012), pp 25–32. doi:10.1016/j.websem.2012.05.003.

[61] OWL-Time Ontology. http://www.w3.org/2006/time# (accessed February 12, 2017).
[62] Ontology of units of Measure. http://www.wurvoc.org/vocabularies/om-1.8/ (accessed February 12, 2017).
[63] Mathematical Modelling Ontology (MAMO). https://bioportal.bioontology.org/ontologies/MAMO/?p=summary (accessed February 12, 2017).
[64] P. Suresh, G. Joglekar, S. Hsu, P. Akkisetty, L. Hailemariam, A. Jain, G. Reklaitis, V. Venkatasubramanian, Onto MODEL: Ontological mathematical modeling knowledge management, Computer Aided Chemical Enginnering, 25 (2008), pp 985–990. doi:10.1016/S1570-7946(08)80170-8.
[65] T. Hong, S. D'Oca, S.C. Taylor-Lange, W.J.N. Turner, Y. Chen, S.P. Corgnati, An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework, Building and Environment, 92 (2015), pp 764–777. doi:10.1016/j.buildenv.2015.08.006.
[66] T. Hong, S. D'Oca, S.C. Taylor-Lange, W.J.N. 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), pp 196–205. doi:10.1016/j.buildenv.2015.08.006.
[67] M. Ding, Z. Zhang, X. Guo, CIM extension of Microgrid Energy Management System, in: 2009 Asia-Pacific Power and Energy Engineering Conference, Wuhan, China, (2009), pp 1–6. doi:10.1109/APPEEC.2009.4918216.
[68] What is MathML. https://www.w3.org/Math/ (accessed November 22, 2018).
[69] A.W. Mcmorran, An Introduction to IEC 61970-301 & 61968-11: The Common Information Model, (2007). https://cimphony.com/cim-intro.pdf (accessed November 22, 2018).
[70] IntelliGrid Common Information Model Primer: Second Edition, (2013). https://www.epri.com/#/pages/product/000000003002001040/?lang=en-US (accessed November 22, 2018).
[71] J. Z.Pan, Resource Description Framework, in: S. Staab, R. Studer (Eds.), Handbook on Ontologies, (2009), pp 71–90. doi:10.1007/978-3-540-92673-3_3.
[72] SPARQL Query Language for RDF. https://www.w3.org/TR/rdf-sparql-query/ (accessed November 22, 2018).