This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957047.

H2020-LC-SC3-EE-2020-1/LC-SC3-B4E-6-2020

Big data for buildings

BIGG

Building Information aGGregation, harmonization and analytics platform

Project Nº 957047

D4.1

Description of the preliminary harmonization layer

Responsible:	Stoyan Danov (CIMNE)
Document Reference:	D4.1
Dissemination Level:	Public
Version:	Final
Date:	07/03/2022

Executive Summary

The document provides an overview of the development of the harmonisation layer of BIGG in the initial project period and presents the obtained results. During this phase the work focused on the development of the BIGG Standard Data Model 4 Buildings, which represents the common base for harmonisation of data from different sources. The steps for development are described in different sections and the methodological approach for building the data model is outlined.

The report starts with an introduction, overviewing the context of implementation, the needs the data harmonisation should address, and the state of the art.

The development starts with **analysis of the required data** necessary to support the planned Use Cases. It aims at defining the full range of data elements that should find their place in the data model. The analysis is extended also to the existing datasets from the local systems at the pilots. As a result, a set of atomic conceptual data groups are defined, attributes are assigned to them, and relationships between them defined. The atomic concepts related to each Use Case are listed and checked for completeness.

In parallel, relevant **existing ontologies in the Buildings domain are analysed** to pre-select concepts and relations that can be reused in BIGG in order to align as much as possible the data model to existing definitions since the start. Once the atomic concepts have been defined, semantic matches from the existing ontologies have been identified from each of the analysed ontologies. Recommendations for adoption of the identified concepts or not are included based on considerations such as level of adoption of the existing ontology, appropriateness of the relations for the BIGG use cases, etc.

Next, the report describes the methodology of building the data model and provides details on the development of the BIGG classes, attributes, relationships, and taxonomies that form part of the model.

The initial BIGG Standard Data Model 4 Buildings is then concisely presented and detailed description of its components is included in a document annex, as well as the procedures for mapping existing data sources to the BIGG data model.

Finally, conclusions and next steps are included, outlining the plans for evolving the BIGG Standard Data Model 4 Buildings into a BIGG ontology as a way towards full connection to existing knowledge description in domain ontologies and contribution to standardisation of buildings data.

Table of Contents

Table of Figures

List of Tables

Table of Acronyms and Definitions

Acronym	Definition
EC	European Commission
WP	Work Package
CWE	Collaborative Working Environment
PMB	Project Manager Board
PM	Project Manager
KoM	Kick-off meeting
RAF	Reference Architecture Framework

Contributors Table

DOCUMENT SECTION	AUTHOR(S)	CONTRIBUTOR(S) to results	REVIEWER(S)
All sections	Stoyan Danov (CIMNE)	Eloi Gabaldon (CIMNE), Jordi Carbonell (CIMNE), Gerard Mor (CIMNE), Nicolas Bus (CSTB), Guillaume Picinbono (CSTB)	Guillaume Picinbono (CSTB), Michel Hours (Inetum FR), Chris Develder (IMEC)
Section IV	Stoyan Danov (CIMNE), Nicolas Bus (CSTB), Guillaume Picinbono (CSTB)	Eloi Gabaldon (CIMNE), Jordi Carbonell (CIMNE),	Guillaume Picinbono (CSTB), Michel Hours (Inetum FR), Jarne Kerkaert (Inetum BE) Chris Develder (IMEC)
Section V	Stoyan Danov (CIMNE), Nicolas Bus (CSTB)	Nicolas Bus (CSTB), Eloi Gabaldon (CIMNE), Jordi Carbonell (CIMNE), Gerard Mor (CIMNE)	Guillaume Picinbono (CSTB), Michel Hours (Inetum FR), Chris Develder (IMEC)

I. INTRODUCTION

The interoperability with external data hubs is a core capability of BIGG. The Communication Layer developed in WP3 ensures, through connectors and/or definition of APIs, the interoperability with relevant data hubs at national and European level through referencing them through its standard data model and implementing the necessary transformations considering the BIGG Standard Data Model 4 Buildings developed in WP4. This data harmonization enables the semantic exchange of data between external data hubs and the BIGG internal storage – that stores the harmonized data ready to be used by the AI toolbox in WP5, ensuring at the same time the direct interoperability between external data collections that can use the BIGG infrastructure (standard data model, mappings, transformations, APIs) to exchange data between them and align their datasets.

The aim of this report is to present the BIGG Data Harmonization Layer developed within WP4 and its components during the first year of the project. This includes the development of documentation and supporting tools. These are shared with the BIGG consortium in the project repository.

The main objective of the work in the period covered by this report has been the development of the data model of BIGG, which forms the basis for each use case of the internal data storage and which serves as the recommended approach (to-be-proposed-as standard reference) for harmonisation of data from external sources. The report presents the methodological approach and outlines the different steps of the development process, which has been based on detailed analysis of BIGG use cases data requirements (see deliverable D2.1 of the project), conceptualisation of data structures and leveraging on relevant existing ontologies in the buildings domain.

As an outcome, the initial BIGG Standard Data Model 4 Buildings has been developed and is presented. The report covers also the mapping of the datasets involved in the pilot trials and presents the followed procedures and tools.

This work has been key for ensuring the alignment between the analytics, storage and communication components of the BIGG architecture (see deliverable D2.2 of the project) and for the implementation of the first version of demonstrators.

I.1. Organization of the document

This report is organised as follows:

- Section II describes the necessities for harmonisation of buildings data and how these are addressed by the project and provides the context in which BIGG Data Harmonization Layer is developed.
- Section III outlines the analysis of data requirements and the formulation of the preliminary definitions of the data concepts.
- Section IV describes the initial identification of relevant ontologies, outlines the analysed aspects, and presents recommendations for reuse of concepts in BIGG.
- Section V details the methodology for development of the data model.
- Section VI presents the overall BIGG Standard Data Model 4 Buildings in a condensed form.
- Section VII presents the procedures for the mapping of the available datasets.
- Section VIII provides conclusions from the development in the initial phase and outlines the next steps.
- The Annex contains the full description of classes and attributes of the BIGG Standard Model 4 Buildings and a detailed UML class diagram of the model.

I.2. Scope and audience

The results of this deliverable are important for all project participants, both in the Work Packages and in the pilot trials. The participants should take into account the provided definitions in the implementation of the work, as they constitute the base for alignment, exchange, and use of data in the project.

The deliverable is also of interest for a wide public, especially for implementors and providers of data-driven services related to buildings.

I.3. Definitions

The development of the preliminary harmonisation layer presented in this report involved a variety of steps and operations that are described using terminology from the field of data modelling and semantic technologies. There are numerous definitions of these terms that can be found in the literature and these are not always sufficiently clear and straightforward. For a better clarity of the explanation, the following definitions are meant to describe the adopted terms and how they are used in the context of BIGG.

Database Schema

Database schema is a physical implementation of data model in a specific database management system. It includes all implementation details such as data types, constraints, foreign or primary keys.

Data Model

Abstract model that organizes data elements and their relationships. Data model may be represented in many forms, such as Entity Relationship Diagram or UML Class Diagram complemented with description of the data elements. It is not related to any implementation but describes the structure of the data and details the information to be stored in an Information System.

In BIGG the term “data model” is referred to as a common detailed description of the data structure consisting of UML diagram, description of classes, attributes and relations that can be used for mapping to external database schemas and provide the necessary information for implementation of databases and physical systems that intend to use the BIGG Reference Architecture Framework components.

Ontology

Ontologies are part of the web semantic approach defined by the Wide Web Consortium. It is formal and explicit specification of the knowledge relatively in a domain expressed both in human-understandable and machine-readable form that enables its sharing across heterogeneous environments. An Ontology defines classes, properties and relations between classes as well as taxonomies. Ontologies can be combined to cover a complex domain. A wide range of ontologies are published as Open Data resources. Most of them are based on international standards and aligned with each other.

The BIGG ontology to be created in the next stage will be used for a full conceptualisation of the knowledge in the buildings domain that is necessary to support the BIGG Use Cases, as well as for the alignment of this knowledge with concepts from existing ontologies and data standards to enable its sharing, independently on the specific implementation.

Enumeration

A hierarchical structure (also called taxonomy) used in the data modelling to produce constrained lists of values for the data entities of enumerated data type. Sometimes referred to in the text as taxonomy. In an ontology, a enumeration can be implemented as a hierarchy of classes (general OWL approach) or as a graph of terms instances (SKOS approach).

Semantic technologies

Semantic technologies (also known as Semantic Web) are a combination of standards, specifications and software that help machine in understanding data. Semantic Technologies bring openness, structure, interoperability, universal identification and logical links between data. Ontologies are part of the semantic technologies. Semantic models exist in different levels of complexity, among which are ontologies or knowledge graphs, data models that represent sets of concepts belonging to a specific domain and the relationships between them.

II. NECESSITY TO BE ADDRESSED

II.1. Context

Buildings are responsible for the largest share of European final energy consumption (40%), according to the Energy Efficiency – first fuel for the EU economy report [1]. Buildings also present the greatest potential to save energy, as 75% of those standing in the EU were built during periods with minimal energy-related regulations and 75-90% of those standing today are expected to remain in use in 2050. Even in new construction and energy retrofits, significant gap between the expected and the actually measured consumption in operation is observed [2]. Human behavioural factors such as occupant behaviour, and quality of provided indoor environmental conditions are pointed as factors affecting the energy consumption in an extent at least as great as those of climate, building envelope, and energy systems characteristics [3].

Big data technologies together with the development of the Internet of Things (IoT) opened new possibilities. The continuously increasing amount of data generated by buildings through the adoption of digital technologies, integrating sensors, controllers, and their connectivity, offer enormous potential for increasing the efficiency in both existing and new buildings. However, this potential is hindered by ambiguity of the data definitions and lack of standardisation across applications and databases, which makes it difficult to exchange, compare and combine the data both at building and inter-building levels. As a result, (1) only a small fraction of the available data is analysed and effectively used for providing innovative building-related services; (2) available data is often used in silos and cannot be combined with other data or employed in other use cases.

The harmonisation of data by BIGG aims at contributing to overcoming these obstacles. In this context, the BIGG project aims to facilitate the implementation of big data analytics for buildings, reducing the effort to create applications. The project focuses on development of an open-source Big Data Reference Architecture and AI Analytics Toolbox with demonstrated capabilities to support a variety of use cases and applications covering the whole building life cycle. Combining of data from different sources for joint analytics and standardised input and output from the Analytics Toolbox components are indispensable features for the solution. Therefore, data harmonisation and interoperability are core in the concept of BIGG.

Whereas technical interoperability, i.e., the successful exchange of data, can be more easily achieved through the adoption of standardised communication protocols, the lack of harmonisation at the information level, which deals with the semantic (meaning) of data, constitutes a more complicated issue [4]. Semantic technologies, a combination of software and specifications that allow encoding the meaning of data and their interrelations in a machine-processable form, offer powerful tools to overcome the interoperability challenge [5]. Semantic models exist in different levels of complexity, including ontologies or knowledge graphs, data models that represent sets of concepts belonging to a specific domain and the relationships between them [6]. Previous research has shown that the association of raw data to terms belonging to a common ontology, and thus a common meaning, facilitates the uniform representation of data collected by different sources and therefore their informational interoperability [7], [5].

Most of the currently developed ontologies follow the World Wide Web Consortium (W3C) standards for the Semantic Web and are aimed at making internet data machine-readable. W3C standards-compliant ontologies are based on the Resource Description Framework (RDF) specifications, according to which information can be modelled combining triples in the form of “subject – predicate – object”, expressions called “triples” in RDF terminology [6]. Built on top of RDF, the Web Ontology Language (OWL) is designed for defining and instantiating Web ontologies.

Good practices for ontology development are based on two fundamental principles: reuse and alignment [8]. The rationale behind ontology reuse is to take advantage of the effort spent by others in encoding semantic structures. On the other hand, new applications require modification or definition of new concepts and relations between them for the same domain, introducing heterogeneity issues. To overcome the inconvenience of this, the process of ontology alignment aims at establishing relations between different ontologies while preserving the originals. This objective is achieved through ontology matching, the process of finding correspondences, called mappings when these are directional, between entities belonging to two distinct ontologies [9].

The broad scope of applications covered by the BIGG project requires combining of concepts from different ontologies, as well as definition of new concepts and relations that would enable the operational efficiency of big data analytics in the buildings' domain. An approach employing both ontology reuse and ontology alignment is needed to accomplish the objective of data harmonisation and standardisation set by the project.

II.2. How does BIGG address the necessities?

Data harmonisation is a process of bringing together data of varying file formats, naming conventions, and columns, and transforming them into one cohesive data set. In the context of the BIGG Reference Architecture Framework (see deliverable D2.2 of the project), the harmonisation layer ensures the capability to align, harmonise, and make comparable data from different sources over an internal standard data model. The harmonised data format will be adopted for the AI Analytics Toolbox components' input, making possible the integration of the developed tools with legacy systems and technologies as well as with new developments.

The harmonisation layer aims to make the process of data harmonisation systematic, reproducible, and operational. It is materialised through the development of the BIGG Standard Data Model 4 Buildings, open documentation, and specifications for implementation of open-source components for data mapping and harmonisation.

The harmonisation layer will be developed gradually throughout the project, initially creating the necessary artifacts, testing them in the first phase, and further evolving them as integrated components at a second stage.

In the first project stage the BIGG Data Model will be used as a base for implementing customised transformations of the raw data sources to the harmonised data structure. These transformations will be implemented in the local systems from where the data sources originate and will support the initial testing of the planned Business Cases.

In a second stage, the harmonisation layer process will evolve, developing specifications for generic components that will be implemented and embedded into the communication interfaces (WP3) and integrated within the BIGG architecture (WP2) to support operationally the Business Cases.

III. Analysis of data requirements

The broad scope of applications covered by the BIGG project requires combining static and dynamic data. The data comprise description of buildings' and building systems' characteristics, time series data from meters and sensors, maintenance actions, user actions, economic, utility tariffs, location weather data, etc. These data feed the AI Analytics Toolbox components, which in turn produce sets of data output that support the envisaged Business Case services. All these heterogeneous data need to find their place in the BIGG Standard Data Model 4 Buildings, appropriately allocated in conceptually related groups, in order to enable their operational use in the BIGG reference architecture. The followed approach of data requirements identification is described next.

III.1. Identification of data concepts

The preliminary analysis of data requirements was done over the Business Cases and their corresponding Use Cases and data sets outlined in D6.1-Detailed description of pilot's technical assets.

The analysis required significant effort and multiple iterations of collaborative work. with the first step comprised initial identification of atomic concepts based on the necessary and expectedly retrievable information from external sources and users involved in the Use Cases. In a next step, the atomic concepts were associated to the Use Cases and their relations required to support the envisaged functionalities were identified. Gradually, more granular data attributes with their data types were defined and added to each concept group to complete the data requirements. Table 1 presents a list of the identified data concepts.

Table 1: Identified atomic concepts

Atomic concept	Data represented by the concept
Area	The surface of a building, e.g., net, gross, heated
Building	The building characteristics, e.g., use type, year of construction
Building element	Construction element of the building, e.g., wall, slab
Building space	Enclosed part of the building dedicated to certain activity
Building system	Active building system, e.g., heating, ventilating, electrical
Cadastral data	Data about the building and the building parcel from cadastre
Contract	Contract for service provision, e.g., maintenance or energy performance contract
Cost saving	Monetary saving from improvement of some aspect of the building, e.g., from energy efficiency measure or project, maintenance, etc.
Device	Any meter, sensor, or actuator
Emissions	CO_2 emissions of the building or system
Energy cost	Cost of the different energy sources used in the building
Energy certificate	Energy performance certificate data

Energy Efficiency Measure (EEM)	EEM applied to or planned for a building
Energy saving	Saving of energy from EEM
Energy	Energy sources used
Financial indicators	Different indicators such as net present value and savings-to-investment ratio
Group	Group of building elements
Indoor quality	Data about the building indoor parameters
Investment	Investments in project or EEMs implemented in a building
Location	Location data, e.g., address, coordinates
Maintenance	Information about maintenance operations
Measurement	Measurements from sensors or meters
Occupancy	Data on building occupancy profile
Organization	Data on the organisation/owner of the building
Project	Includes a package of EEMs or renovation measures
Subsidy/incentive	Subsidies received for the implementation of the projects
User	Data on the user of the services provided by BIGG
Weather properties	Properties that influence the energy consumption and use, e.g. temperature, wind speed
Zone	Group of building spaces

The conceptual data groups of the required data input identified for each Use Case is presented in Table 2.

Table 2: Identified conceptual data groups in the use case data input

BIGG Use Cases	Identified data concepts
Business Case 1: Benchmarking and Energy Efficiency tracking in Public Buildings	
Use Case 1: Benchmarking and Monitoring Energy Consumption	Building, Area, Location, Organization, User, Device, Measurement, Energy cost, Emissions, Financial indicators, Weather
Use Case 2: Energy Efficiency Measures: Registration and Evaluation	Project, Energy Efficiency Measure, Energy saving, Cost saving, Investment, Subsidy, Financial indicators, Emissions, Device, Measurement, Weather

Business Case 2: Energy Certification in Residential and Tertiary Buildings	
Use Case 3: Integration of INSPIRE spatial data with Energy Performance Certification	Location, Cadastral data, Energy certificate
Use Case 4: Adoption of the sustainability indicators of EU framework Level(s) in Building Certification	Energy certificate, Building, Device, Measurement, Emissions, Weather
Business Case 3: Building life-cycle: from planning to renovation	
Use Case 5: Interoperability between BIM, BMS, CMMS and simulation engines	Building, Location, Building space, Building element, Building system, Zone, Group, Maintenance, Device, Measurement, User
Use Case 6: Interoperability of BIGG with EEFIG-DEEP	Project, Energy Efficiency Measure, Measurement
Use Case 7: Interoperability between EU Building Stock Observatory (EUBSO) and national/regional Energy Performance Certification hubs	Energy certificate, Building, Location, Measurement
Business Case 4: Energy Performance Contract based savings in Commercial Buildings	
Use Case 8: Building assets management	Organization, Building, Building element, Location, Contract, Project, Device, Measurement
Use Case 9: Actual savings tracking realized by the Energy Conservation Measures (ECMs) undertaken by the ESCO and monitoring on daily/weekly/monthly basis	Project, Contract, Energy Efficiency Measure, Energy cost, Energy saving, Financial indicators, Device, Measurement, Weather
Use Case 10: Energy Performance Contract management	Organization, User, Contract, Device, Measurement
Business Case 5: Buildings for occupants: comfort case	
Use Case 11: Optimization using weather forecasts	Weather, Device, Measurement, Energy cost
Use Case 12: Optimization using occupancy forecasts	Occupancy, Weather, Device, Measurement, Energy cost
Use Case 13: Optimization using price forecasts	Weather, Device, Measurement, Energy cost
Business Case 6: Flexibility potential of residential consumers on electricity and natural gas	
Use Case 14: Electricity Demand-Response	Occupancy, Weather, Device, Measurement, Energy cost

Use Case 15: Natural Gas Demand-Response	Occupancy, Weather, Device, Measurement, Energy cost

III.2. Data types

The analysis of the available data sets for supporting the BIGG Use Cases at the pilots identified a range of data types for the attributes in the conceptual data groups. These are listed below.

Table 3: Identified data types

Data type	Abbreviation	Definition
Unique Identifier	UID	An identifier that is unique among all the identifiers related to that attribute.
Integer	int	An integer number.
Float	float	A float point number.
String	string	A string of characters
Percentage	%	A fraction of a total represented with a percentual value.
Boolean	bool	Boolean value that can either take a “true” or a “false” value.
Date	date	Date in the format "YYYY-MM-DD", where YYYY indicates the year, MM the month, and DD the day.
Time	time	Point in time expressed in the format “hh:mm:ss.s” where hh indicates the hour, mm the minute, and ss.s the second.
Datetime	datetime	Point in time expressed in the format “CCYY-MM-DD hh:mm:ss.s” where YYYY indicates the year, MM the month, DD the day, hh the hour, mm the minute, and ss.s the second.
Point	point	Representation of a location on the Earth through its latitude and longitude WSG 84 coordinates.
Enumeration	enum	Datatype for those attributes which value can be chosen from a predefined set.
List	list	A sequence of ordered values of some datatype.

III.3. Summary of the outcomes

The preliminary analysis of data requirements represented the starting point for the development of the BIGG data model. The analysis of data requirements led to the definition of a preliminary set of concepts comprising the data attributes necessary to support the envisaged functionalities of the Use Cases. The process of iterative revision, listing, and addition of attributes to the concepts was accompanied by assigning of a data type for each attribute, resulting in definition of a full set of data types for the model.

A specific data type, designated as enumeration, representing a constrained list of admissible values, was identified as one of the keys for the harmonisation of data from different sources. Examples of enumerations are the list of possible energy efficiency measures that could be applied to a building, or the list of device types that can be installed. The analysis of the data sources evidenced the generalised use of ad hoc taxonomies for the enumerations which is an impediment for the data alignment. For that reason, particular effort was dedicated to the development of taxonomies.

The use of the preliminary analysis of the data requirements in the process of creation of the BIGG data model and the development of taxonomies for the enumerated data types is described in section V.

IV. IDENTIFICATION OF RELEVANT ONTOLOGIES

The analysis of relevant ontologies and data standards in the Buildings domain intended to identify data concepts and relations that can be reused in the BIGG data model. The objective was to discover semantic coincidences with the atomic concepts defined in the analysis of the data requirements of the BIGG Use Cases and to adopt already existing definitions from accepted standards rather than creating parallel definitions.

Reusing concepts from existing ontologies aims to provide a nexus of BIGG to the applications and databases that follow these ontologies, establishing the base for the creation of an ecosystem of semantically interoperable tools. At the same time, the analysis aimed to spot missing concepts in the existing ontologies. Defining such missing concepts in BIGG is hoped to contribute to further evolve the related existing standards.

The analysis of ontologies focused on some specific aspects relevant to be considered when deciding their reuse in BIGG. These are summarised in a compact tabular form in Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10 and Table 11 below.

Important aspect considered in the analysis was the level of adoption of the ontologies, as this is important factor for the decision to reuse them or not, due to the impact on the data interoperability in the broad Buildings domain their adoption could bring in practice. For simplifying the representation, we define a scale from 1 to 5 for the ontology adoption, with 1meaning only theoretic definition of the ontology and 5 full adoption in the respective domain.

Table 4: IFCOWL Ontology

Ontology name	IFCOWL
Full name	Industry Foundation Classes
Version	4.0.2.1 (Version 4.0 - Addendum 2 - Technical Corrigendum 1)
Status	ISO 16739-1:2018
Documentation	https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2_TC1/HTML
Ontology	https://standards.buildingsmart.org/IFC/DEV/IFC4/ADD2/OWL/index.html
Formats	Express, XSD, RDF, TTL
Description	The Industry Foundation Classes, IFC, are an open international standard for Building Information Model (BIM) data that are exchanged and shared among software applications used by the various participants in the construction or facility management industry sector. The standard includes definitions that cover data required for buildings over their life cycle. This release, and upcoming releases, extend the scope to include data definitions for infrastructure assets over their life cycle as well.
Targeted application / Use cases	• Widely used for building design and construction phases • Starting to be used in building exploitation phase • IFC model will be used as input for the BIGG data model. Part of IFC ontology is used to describe the building structure in the BIGG data model.

Level of adoption [1 - 5]	• 5 – full adoption as a standard
Relevant concepts / classes / relations for BIGG	• Area: IfcBuilding --> Qto_BuildingBaseQuantities --> GrossFloorArea • Building: IfcBuilding • Building element: IfcBuildingElement • Building space: IfcSpace • Building system: IfcSystem • Device: IfcDistributionElement • Group: IfcGroup • Occupancy: IfcBuilding --> Pset_BuildingCommon / Pset_BuildingUse • Zone: IfcZone or IfcSpatialZone (with geometry)
Recommendation for use in BIGG	• IfcOWL is a widely used data model, produced by most CAD softwares. • IfcOWL should be used for all architectural / spatial structure concepts

Table 5: SSN/SOSA Ontology

Ontology name	SSN/SOSA
Full name	Semantic Sensor Network Ontology / Sensor, Observation, Sample and Actuator (SSN/SOSA)
Version	OGC 16-079 / W3C Recommendation 19 October 2017
Status	W3C recommendation
Documentation	https://www.w3.org/TR/vocab-ssn/
Ontology	https://github.com/w3c/sdw/blob/gh-pages/ssn/rdf/sosa.ttl
Formats	RDF, TTL
Description	The Semantic Sensor Network (SSN) ontology is an ontology for describing sensors and their observations, the involved procedures, the studied features of interest, the samples used to do so, and the observed properties, as well as actuators. SSN follows a horizontal and vertical modularization architecture by including a lightweight but self-contained core ontology called SOSA (Sensor, Observation, Sample, and Actuator) for its elementary classes and properties. With their different scope and different degrees of axiomatization, SSN and SOSA are able to support a wide range of applications and use cases, including satellite imagery, large-scale scientific monitoring, industrial and household infrastructures, social sensing, citizen science, observation-driven ontology engineering, and the Web of Things. Both ontologies are described below, and examples of their usage are given.

Targeted application / Use cases	• SOSA ontology describes sensors and their observations and has application in all use cases these are involved. • A part of SOSA ontology is compatible with BIGG data model to describe the sensor network and measurements context.
Level of adoption [1 - 5]	• [TO BE EVALUATED]
Relevant concepts / classes / relations for BIGG	• sosa:Sensor • sosa:observes • sosa:Sample • sosa:Result • ssn:System
Recommendation for use in BIGG	• SOSA model will be used as input for the BIGG data model. • BIGG has its own measurement model but is it aligned with SOSA

Table 6: BOT Ontology

Ontology name	BOT
Full name	Building Topology Ontology
Version	0.3.2
Status	Public draft
Documentation	https://w3c-lbd-cg.github.io/bot/
Ontology	http://www.w3id.org/bot/bot.ttl
Formats	TTL
Description	The Building Topology Ontology (BOT) is a minimal OWL DL [owl2-primer] ontology for defining relationships between the sub-components of a building. It was suggested as an extensible baseline for use along with more domain specific ontologies following general W3C principles of encouraging reuse and keeping the schema no more complex than necessary. BOT is from design scoped to describe topology specific to the buildings domain. It does not provide a generic description of topological relationships as, for instance, the Regional Connection Calculus (RCC) [Randell]. See also in this regard the discussion in the accompanying publication [[Rasmussen2020], Sec. 3.3] .
Targeted application / Use cases	• Definition of the topological relations of building elements and spaces inside a building. • Often used for energetic simulation (thermal, acoustic).
Level of adoption [1 - 5]	• 3 – still limited adoption

Relevant concepts / classes / relations for BIGG	• Building: bot:Building • Building element: bot:Element • Building space: bot:Space • Device: bot:Element • Zone: bot:Zone
Recommendation for use in BIGG	• Could be used as a simplified version IfcOWL (converters IFC2BOT exist). • But it does not provide all the concepts included in IFC

Table 7: geoSPAQL Ontology

Ontology name	geoSPARQL
Full name	Geographic Vocabulary and Query Language for RDF Data
Version	1.0 Approved 1.1 Draft
Status	Approved OGC Implementation Standard
Documentation	https://www.ogc.org/standards/geosparql (v1.0) https://opengeospatial.github.io/ogc-geosparql/geosparql11/spec.html#_normative_references (v 1.1)
Ontology	geo: http://www.opengis.net/#geosparql geof: http://www.opengis.net/def/function/geosparql/ w3cGeo: http://www.w3.org/2003/01/geo/wgs84_pos# geor: http://www.opengis.net/def/rule/geosparql/ sf: http://www.opengis.net/ont/sf#
Formats	RDF, geoJSON-LD. Compatible with geoJSON, KML, GML, WKT
Description	This ontology is related to other complementary OGC standards and ontologies: WKT, GML, WGS84
Targeted application / Use cases	• GeoSpatial data • geospatial context, • geolocation (Buildings), • geometry description (Buildings footprints parcels).
Level of adoption [1 - 5]	• 4 – advanced adoption
Relevant concepts / classes / relations for BIGG	• CasdatralInfo.hasGeometry (WKT polygon)

Recommendation for use in BIGG	• WKT serialization for polygons is part of geoSPARQL https://en.wikipedia.org/wiki/Well-known_text_representation_of_geometry , https://event.cwi.nl/eswc2015-geo/03-stsparql-geosparql.pdf

Table 8: QUDT Ontology

Ontology name	**QUDT**
Full name	Units and Systems of Units
Version	2.1 Approved
Status	Approved W3C and NASA Standard
Documentation	http://qudt.org/
Ontology	unit : http://qudt.org/2.1/vocab/unit
Formats	RDF
Description	A set of vocabularies representing the various quantity and unit standards.
Targeted application / Use cases	Measurement data context and description
Level of adoption [1 - 5]	5 – full adoption
Relevant concepts / classes / relations for BIGG	• Measurement
Recommendation for use in BIGG	Should definitely be used to define units

Table 9: WGS84 Ontology

Ontology name	**WGS84**
Full name	Basic Geo (WGS84 lat/long) Vocabulary
Version	1.22 Approved
Status	Approved W3C Standard
Documentation	https://www.w3.org/2003/01/geo/
Ontology	geo: http://www.w3.org/2003/01/geo/wgs84_pos#
Formats	RDF

Description	This is a basic RDF vocabulary that provides the Semantic Web community with a namespace for representing lat(itude), long(itude) and other information about spatially-located things, using WGS84 as a reference datum.
Targeted application / Use cases	Location for buildings and other geolocated resources
Level of adoption [1 - 5]	5 – full adoption
Relevant concepts / classes / relations for BIGG	• LocationInfo
Recommendation for use in BIGG	This is the best standard way of defining location for a building

Table 10: FOAF Ontology

Ontology name	**FOAF**
Full name	Basic Geo (WGS84 lat/long) Vocabulary
Version	0.99 Approved
Status	Approved W3C Standard
Documentation	http://xmlns.com/foaf/spec/
Ontology	foaf: http://xmlns.com/foaf/0.1/
Formats	RDF
Description	Devoted to linking people and information using the Web. Regardless of whether information is in people's heads, in physical or digital documents, or in the form of factual data, it can be linked. FOAF integrates three kinds of network: social networks of human collaboration, friendship and association; representational networks that describe a simplified view of a cartoon universe in factual terms, and information networks that use Web-based linking to share independently published descriptions of this inter-connected world. FOAF does not compete with socially-oriented Web sites; rather it provides an approach in which different sites can tell different parts of the larger story, and by which users can retain some control over their information in a non-proprietary format. The Friend of a Friend (FOAF) RDF vocabulary, described using W3C RDF Schema and the Web Ontology Language.
Targeted application / Use cases	Describe human resources involved in retrofitting/renovation projects
Level of adoption [1 - 5]	5 – full adoption
Relevant concepts /	• Person

classes / relations for BIGG	• Organization • Group (of persons) • Company
Recommendation for use in BIGG	FOAF ontology can be extended by VCARD properties http://www.ibiblio.org/hhalpin/homepage/notes/vcardtable.html

Table 11: SAREF Ontology

Ontology name	**SAREF**
Full name	The Smart Applications REFerence
Version	V3.1.1
Status	Approved ETSI Standard
Documentation	https://saref.etsi.org/
Ontology	saref: https://saref.etsi.org/core/
Formats	RDF
Description	Smart Appliances REFerence Ontology (SAREF) is an ontology capturing high level aspects of smart and connected appliances. While SAREF does not capture the full spectrum of equipment and sensors that exist in buildings, SAREF models can be easily integrated into other ontologies. SAREF provides building blocks that allow separation and recombination of different parts of the ontology depending on specific needs.
Targeted application / Use cases	• Smart and connected appliances • IoT applications
Level of adoption [1 - 5]	3 – still limited adoption
Relevant concepts / classes / relations for BIGG	• saref:Device • saref:Measurement • saref:Property • saref:UnitOfMeasure
Recommendation for use in BIGG	• Relevant for description of devices and measurements in BIGG • SAREF and its extensions are based on other ontologies and BIGG could reuse them as a source rather than SAREF

V. METHODOLOGY

The focus in the first project period has been the development of the semantic BIGG Standard Data Model 4 Buildings in the core of the harmonisation layer. This section describes the methodological approach to that task.

V.1. Background and Overall approach

The data harmonisation in BIGG will enable semantic interoperability with both standardised data following existing ontologies and with non-standardised data from existing databases, allowing them to be used with the BIGG Reference Architecture Framework (RAF) and AI Analytics Toolbox in replications of the Use Cases, or other new applications.

The goal of achieving the harmonisation layer of BIGG is intended to be reached through the realization of a series of intermediate objectives. The first step is the development of the BIGG Standard Data Model 4 Buildings reflecting the data requirements of the Use Cases. In consecutive steps a Mapping Tool and a Harmonisation Component will be developed to respectively enable the alignment of the external sources to a common semantics and structure, and to ensure the operational transformation of the data coming from the systems into a harmonised data flow within the BIGG architecture.

The methodology for development aims to optimize the alignment of the BIGG data model to existing ontologies through semantic matching and reuse of data structures. This requires the use of ontological engineering. In the same time, in the practical application of the BIGG RAF technologies for implementing big data analytics for buildings, the BIGG Data Model aims to organise the data in a way ensuring optimal data storage, processing, and retrieving, maximising the systems' operational efficiency.

The work on harmonisation started with analysis of the Use Cases' data requirements and preliminary review of existing ontologies in the building and IoT domains that can potentially address the needs of BIGG. The identified data concepts in the data requirements analysis in section III were contrasted with the outcomes from the preliminary analysis of existing ontologies in section IV. This analysis made evident that the wide scope of concepts covered by the Use Cases of BIGG are not fully described by the semantic concepts of any of the existing ontologies, nor by a combination of them. In most cases the semantic coincidences were represented by synonyms of the identified concepts in the Use Cases rather than the same terms and finding the exact correspondence was a tedious process. Given the large number of attributes grouped in the BIGG concepts, it proved unfeasible to start the development of the BIGG data model from reusing existing ontologies, as this would require much more time than the available in the project.

An alternative and more direct path was adopted, similar to the suggested by [10]. According to it, an adequate approach to ontology engineering would start from the analysis of the data for which encoding of meaning is demanded. This analysis should aim at identifying the characteristics of every single piece of information, such as their value range, the domain concepts described in the dataset and their interrelations, and finally, define the data schema for the set, which is given by the structural organization and the logical relations among all the involved entities. Once a clear view of the data involved is obtained, the required ontology is built through a series of consecutive steps. First, the concepts needed by the application and the relations among them are encapsulated in a model. Only then, following the reusability principle of ontology engineering, a survey is conducted to identify the existing ontologies that match the chosen concepts, and a selection process is realized to choose which external ontology terms to integrate. Finally, the application ontology is completed by adding those classes and properties for which an existing match could not be found, according to existing naming conventions.

The work on the BIGG Harmonisation Layer presented in this report was focused on the development of theBIGG Standard Data Model 4 Buildings and a definition of a data schema during the first year of the project. The model was needed as a common reference for alignment of the data sources involved in the pilot Use Case trials, and for providing harmonised input for the AI Analytics Toolbox, which was essential for enabling the overall project progress in the first period. In addition to the provided semantic uniformity, the model had to organise the data appropriately to address the operational requirements of the IT systems and technologies used in the project.

The data harmonisation in version 1 (V1) of the BIGG solution is aimed through custom transformations to the BIGG Standard Data Model implemented in the local IT systems, as outlined in “D2.2-Initial technical specifications and preliminary design of BIGG Architecture building blocks”. In version 2 (V2) the BIGG data standardisation will be achieved by creation of a BIGG ontology, applying the principle of reuse and matching to existing ontologies and widely accepted data standards. The BIGG ontology will be used for alignment to existing ontologies, contribution to standardisation, and for mapping of external datasets that will be directly converted to RDF harmonised structure by the harmonisation component.

A schematic figure of the overall approach for development of the harmonisation layer of BIGG is presented in Figure 1. The figure outlines the process of development of the components and their intended use. The blue boxes in the figure correspond to the work done accomplished in the first period. The green boxes correspond to ongoing work and planned to be accomplished in the next period.

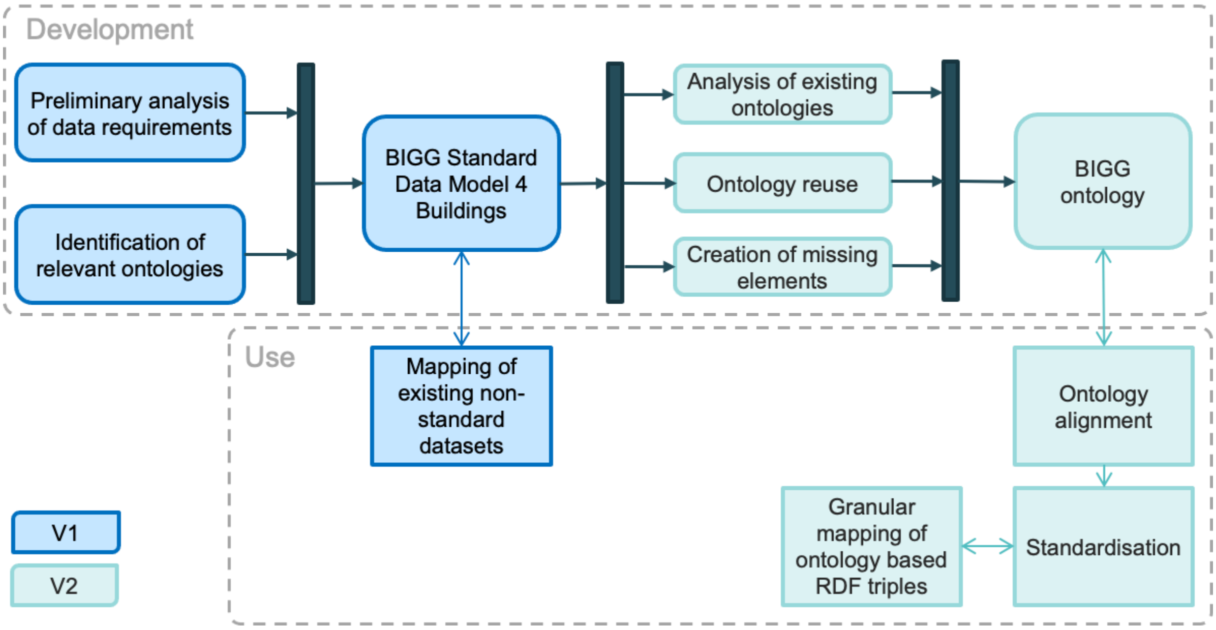

Figure 1 - BIGG data model development and use

V.2. BIGG data model design

The objective of the design was the creation of a conceptual model capable of incorporating the previously identified concepts in section III and IV, organizing them in an appropriate structure, and establishing logical relations among each bit of information.

The development of the BIGG data model followed an iterative process that required multiple stages of modification, re-arrangement, and joint revision with the consortium’s members. Thus, the steps listed further on in the text were not necessarily performed in order, nor the same number of times, and are described here with the intent of explaining the logic behind the work rather than providing an exact timeline of what was done. The design phase was conducted in parallel with the creation of a Unified Modelling Language (UML) diagram, used to represent the structure of the model in its totality, and a spreadsheet, containing detailed definitions and information for each of the elements included in the diagram.

Class definitions

The concepts recognized during the data requirements identification phase were revised to identify overarching conceptual groups, or on the other hand, to define more specific subgroups when a higher level of detail was expected to be required. New concepts were introduced to fill in eventual gaps, to accommodate proposals of the project's partners, or to emulate solutions adopted by other ontologies in the building and energy domains, always avoiding the occurrence of redundancies. Each fundamental concept was modelled as a class and assigned a name following the adopted conventions. According to the UML notation, these classes were represented as boxes in the diagram. When deemed necessary, some conceptual groups were modelled as subclasses of others, thus introducing inheritance relations. Meanwhile, a tab of the spreadsheet file was created for each of the defined classes.

Attribute definitions

The individual classes were studied to choose the necessary attributes. These were selected through an incremental process that took into account the existing databases supporting the BIGG Use Cases, the best practices identified from the review of existing ontologies, the compatibility requirements with third-party databases such as DEEP, and lastly, the expertise of the members of the project's consortium in developing analogous applications, who suggested additional data properties to facilitate the practical implementation of the analytics. The attributes' names were recorded in the UML diagram, within the class they belonged to, along with the expected value data type. At the same time, the attributes were inserted in the spreadsheet tabs of their corresponding classes, to allow the inclusion of descriptions and more information if necessary. A specific procedure was followed for those attributes of enumerated data type, as described in the corresponding step below.

Relationship definitions

Once a wider picture of the classes involved in the model was clear, the next step consisted in the definition of the object properties, called relationships in UML terminology, between them and their cardinalities. When the classes involved were found to match concepts modelled in one of the external tools presented in the Domain ontologies for energy and buildings section, an attempt was made to adopt the same or a similar relational structure, according to the reusability principle. According to the UML nomenclature, only relationships of the association and inheritance were established. The association relationships represent simple structural links between classes and are represented in the diagram as solid lines connecting the classes. The inheritance relations are represented with arrows with white heads starting from the subclass and pointing towards the parent class from which the attributes are inherited. The cardinalities of the associations, which, given a class instance, express the maximum number of other instances that can be related to it, were also decided and reported on the diagram following the convention shown in Table 12.

Table 12: Associations cardinalities

Notation	Cardinality
1	One and exactly one instance (mandatory relation)
0..1	One or zero instances (optional relation)
*	Zero or more instances (optional relation)
1..*	One or more instances (mandatory relation)

Defining a naming convention

The representation of meaning and knowledge through ontologies requires the adoption of an appropriate language. Among the existing standards, are the ones established by the World Wide Web Consortium (W3C) for the creation of the Semantic Web, an extension of the World Wide Web aimed at making internet data machine-readable. W3C standards- compliant ontologies are based on the Resource Description Framework (RDF). The RDF Schema (RDFS) built on top of RDF allows defining the fundamental constituents of ontologies: classes, or collections of resources, instances, which are particular members of a class, and properties.

Moreover, the RDFS introduces hierarchies of classes and properties and supplies some basic data types for property values (e.g., string, date-time, Boolean, etc.). To facilitate the next step of creation of BIGG ontology, the W3C naming convention was adopted for the definition of the classes and attributes in the BIGG data model. According to this model, the classes' names start with a capital letter and the data properties (attributes) and object properties (relationships) start with a lower case letter. To facilitate the understanding, the names are constructed from concatenated words, with those in the middle of the term starting with capital letter to facilitate the readability.

Creation of taxonomies for the enumeration data types

The production of the constrained lists of values admissible for attributes of enumerated data type required special attention during the design of the data model. Whereas some enumerations could be easily compiled as they referred to standard lists, such as the one of European countries or the one of currencies, others demanded a thorough classification procedure. Depending on the content of the list, different external sources were consulted to identify the best possible taxonomy from the point of view of its use in the BIGG Use Cases and the associated databases. These sources included, among others, the identified relevant ontologies previously discussed in Section IV for energy and buildings section, DEEP (De-Risking Energy Efficiency Platform), and analogous classifications adopted by third-party applications the model was meant to exchange data with. The preliminary analysis of existing taxonomies revealed the generalised lack of standardisation in the existing classifications. Therefore, the solution required the creation of specific taxonomies for the enumerated data types in BIGG, searching at the same time maximal alignment with the more widely adopted definitions. The process of creation of taxonomies involved iterations over the following steps, as appropriately for each case.

- Analysis of the classification adopted by the pertinent data sources
- Reuse and adaptation of the terminology from the analysed sources
- Creation of hierarchies
- Consecutive revisions and re-arrangements
- Definition of new items

VI. BIGG STANDARD DATA MODEL 4 BUILDINGS

The BIGG data model has been developed following the methodological steps presented in the previous section. Its documentation consists of a detailed description of the classes, attributes, taxonomies for the enumerated data type fields, as well the relations between the classes and the cardinalities. The full data model is described in a set of excel files and diagrams stored in the BIGG SharePoint repository. Different versions have been produced to reflect the evolution of the development and to track the exact relation between the files in the process of development. The full documentation of the model, close to the end of the project, will be available in an open source repository.

This section presents the last version of the model (v17) in a summarised form through a simplified UML class diagram. The description of the classes, attributes, and detailed UML class diagram are presented in Annex X.1.

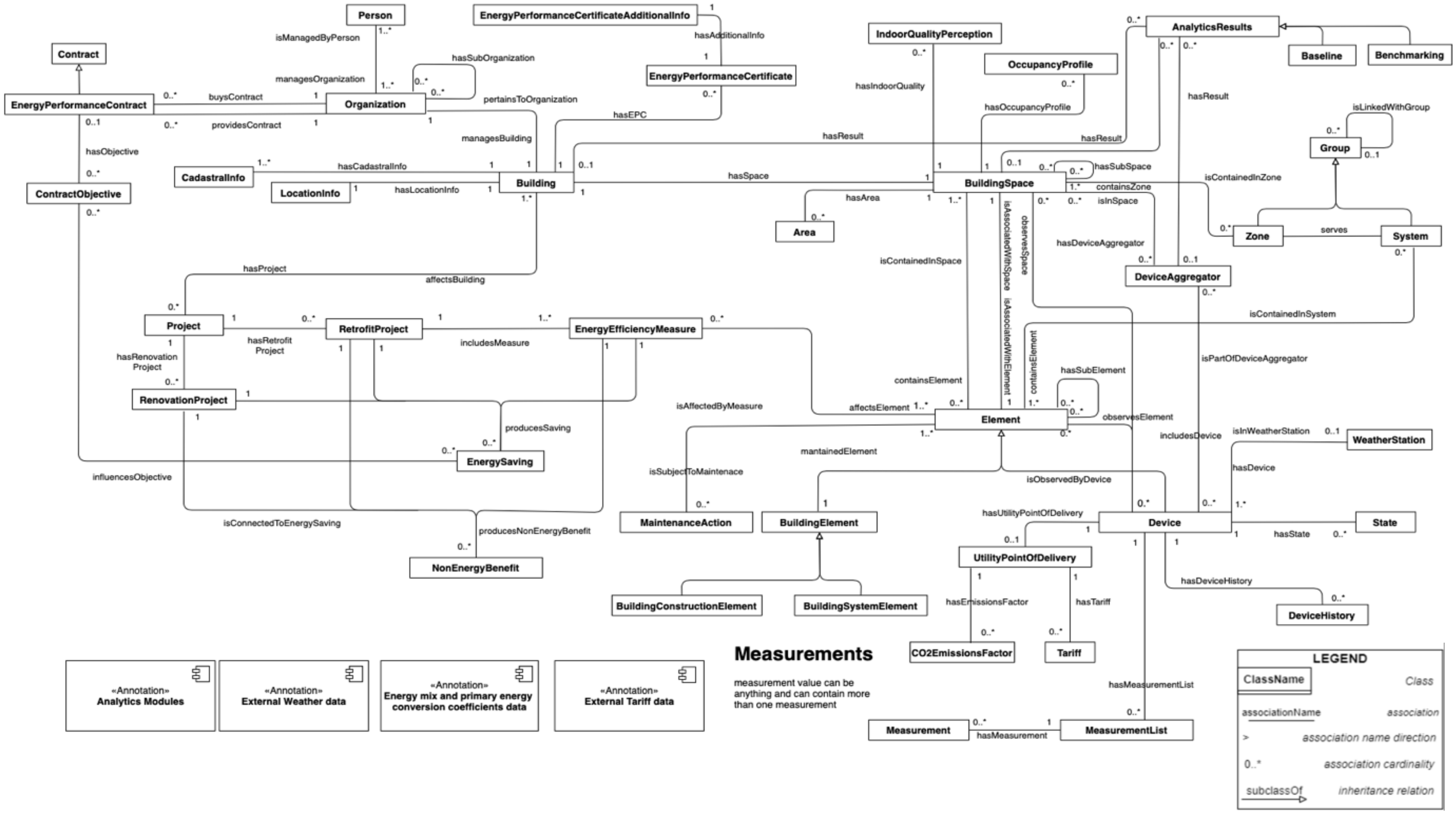

Figure 2 - UML class diagram of the BIGG data model

VII. Mapping to datasets

The mapping of the data sources providing input for the use cases tested in the BIGG pilots has two objectives:

1) to associate the data source attributes with those of the standard data model

2) to validate the BIGG data model and to detect missing attributes that should be added to the model.

This section presents the mapping template developed, outlines the mapping procedure, and provides some examples. The files with the actual mapping of the data sources are stored in the SharePoint repository and are not included in the deliverable, as they store proprietary data.

The mapping has been done using an Excel template developed for this purpose. The template incorporates the description of the BIGG data model's classes, attributes, and taxonomies, and provides a set of rules for mapping the external data sources. The rules for filling the template have been created to capture the structural relations between the external data sources and the BIGG data model and to allow their interpretation, as in some cases there is no one-to-one matching between them. Additional information about the format of the source data (e.g., JSON, XML, CSV, XLSX), the way they are available (e.g., via an API, from a database), and the version and the date of the mapping are collected.

The mapping distinguishes between "static data" and "time series data" and these are mapped in separate sheets of the template in order to capture the differences of the relationships between the classes in BIGG and the external datasets. Data is considered static if doesn't change after recording or is not frequently updated. In difference, timeseries data is coming from sensors and systems that are registering some properties of their state and are registered as pairs of time stamp and values.

Static data and timeseries data are mapped on separate sheets on the template.

VII.1. Static data mapping

On the left part of the static data sheet all classes and attributes of the BIGG data model are listed, each in a separate row. On the right, an empty labelled "My data model" is available, where the mapping should be done. Attributes of enumeration data type (taxonomies) are mapped on separate sheets – one for each taxonomy.

	A	B	C	D	E	F
1			BIGG data model			
2	Class name	Class definition	Attribute name	Attribute description	Attribute data type	(empty)
3	AnalyticsResults	(empty)	(empty)	(empty)	(empty)	
4	Area	An area measurement of a B	areaType	Type of measure area	enum	
5			areaUnitOfMeasurement	Unit of measurement of the area value	enum	
6			areaValue	Numerical value of the area	float	
7	Baseline	Energy performance of the c	baselineAdjustedValue	The value of the baseline after adjustment for	float	
8			baselineDefinition	Short textual description of the baseline and t	string	
9			baselineInfluenceFactor	Field to identify of the factors that impact the b	enum	
10			baselineNonAdjustedValue	The value of the baseline before adjustment fo	float	
11			baselinePerimeter	Definition of the baseline perimeter (eg. whole	string	
12			baselineUnit	The unit in which the baseline is defined	string	
13	Benchmarking	(empty)	(empty)	(empty)	(empty)	
14	Building	A building for which data is p	buildingClosingHour	Closing hour of the building in normal working	time	
15			buildingConstructionType	Building construction type	enum	
16			buildingConstructionYear	Year of building construction or major renovat	integer	
17			buildingID	Unique identifier for the building	UID	
18			buildingIDFromOrganization	Identifier given to the building from the organiz	string	
19			buildingName	Name of the building	string	
20			buildingOpeningHour	Opening hour of the building in normal working	time	
21			buildingOwnership	Indication of whether the building is owned or	enum	
22			buildingUseType	Building purpose of use / activity type	enum	
23	BuildingConstructionElement	Any static element of the bui	buildingConstructionElementType	Type of the building construction element	enum	
24	BuildingElement	Any element of the Building	buildingElementBrand	Brand of the building element	string	
25			buildingElementID	Unique identifier for the building	UID	
26			buildingElementInstallationDate	Date of installation of the building element	date	
27			buildingElementManufactureDate	Manufacture date of the building system elem	date	
28			buildingElementManufacturer	Manufacturer of the building system element	string	
29			buildingElementModel	Model of the building element	string	
30			buildingElementPurchaseDate	Date of purchase of the building element	date	
31			buildingElementSerialNumber	Serial number of the building element	string	
32			buildingElementState	(empty)	boolean	
33	BuildingElementKPIs	(empty)	buildingElementAvailability	(empty)	int	
34			buildingElementMandatoryMaintenance	(empty)	int	
35			buildingElementPresenceCompliance	(empty)	int	
36			buildingElementQualityCompliance	(empty)	int	
37	BuildingSpace	A space that can represent o	buildingSpaceName	Name of the building space	string	
38			buildingSpaceUseType	Purpose of use or activity conducted in the bu	enum	
39			builidingSpaceID	Unique identifier for the building space	UID	
40	BuildingSystemElement	Any system providing a serv	buildingSystemElementEfficiency	Percentual effieciency of the building system	%	
41			buildingSystemElementMaxOutput	Maximum output of the building system eleme	float	
42			buildingSystemElementMinOutput	Minimum output of the building system eleme	float	
43			buildingSystemElementType	Type of the building system element	enum	
44	CadastralInfo	(empty)	landArea	Area of the land	float	
45			landCadastralReference	Cadastral reference of the building	string	
46			landGeometry	Land geometry in in WGS84 crs format	polygon	
47			landGraphicalArea	(empty)	int	
48			landLocation	(empty)	(empty)	
49			landType	(empty)	enum	
50			propertyClass	(empty)	string	
51	Contract	(empty)	contractEndDate	Final date of validity of the contract	date	
52			contractID	Unique identifier for the energy performance c	UID	
53			contractName	Name of the energy performance contract	string	
54			contractPerimeter	Definition of the boundaries of the contract in t	string	
55			contractStartDate	Initial date of validity of the contract	date	
56	ContractObjective	(empty)	objectiveDeadline	Deadline date for the achievement of the obje	date	
57			objectiveDescription	Textual desciption of the objective	string	
58			objectiveID	Unique identifier for the energy performance c	UID	
59			objectiveName	Name of the energy performance contract obje	string	
60			objectiveTargetType	Field used o indicate whether the target of the	enum	
61			objectiveTargetValue	Value of the target of the contract. It has to be	float	
62			objectiveTargetValueUnit	Indication of the unit of measurement of the ta	enum	
63	DeviceAggregator	A virtual unit capable of perf	deviceAggregatorType	Physical property calculated	enum	
64			formula	Formula of the calculation performed by the de	string	
65	Element	Any generic element of the b	(empty)	(empty)	(empty)	
66	EnergyEfficiencyMeasure	Any measure for the improve	energyEfficiencyMeasureCurrencyExchangeRate	Exchange rate between the original investmen	float	
67			energyEfficiencyMeasureDescription	Description of the energy efficiency measure	string	
68			energyEfficiencyMeasureID	Unique identifier of the energy efficiency mea	UUID	
69			energyEfficiencyMeasureInvestment	Investment for the energy efficiency measure	float	
70			energyEfficiencyMeasureInvestmentCurrency	Original currency of the energy efficiency mea	enum	
71			energyEfficiencyMeasureOperationalDate	Date on which the energy efficiency measure	date	
72			energyEfficiencyMeasureSavingsToInvestmentRati	Estimated Savings To Investment Ratio (SIR)	float	
73			energyEfficiencyMeasureType	The type of energy efficiency measure	enum	
74			energySourcePriceEscalationRate	Escalation rate of the price of the energy sour	%	
75			shareOfAffectedElement	Percentage of the element that is affected by t	%	

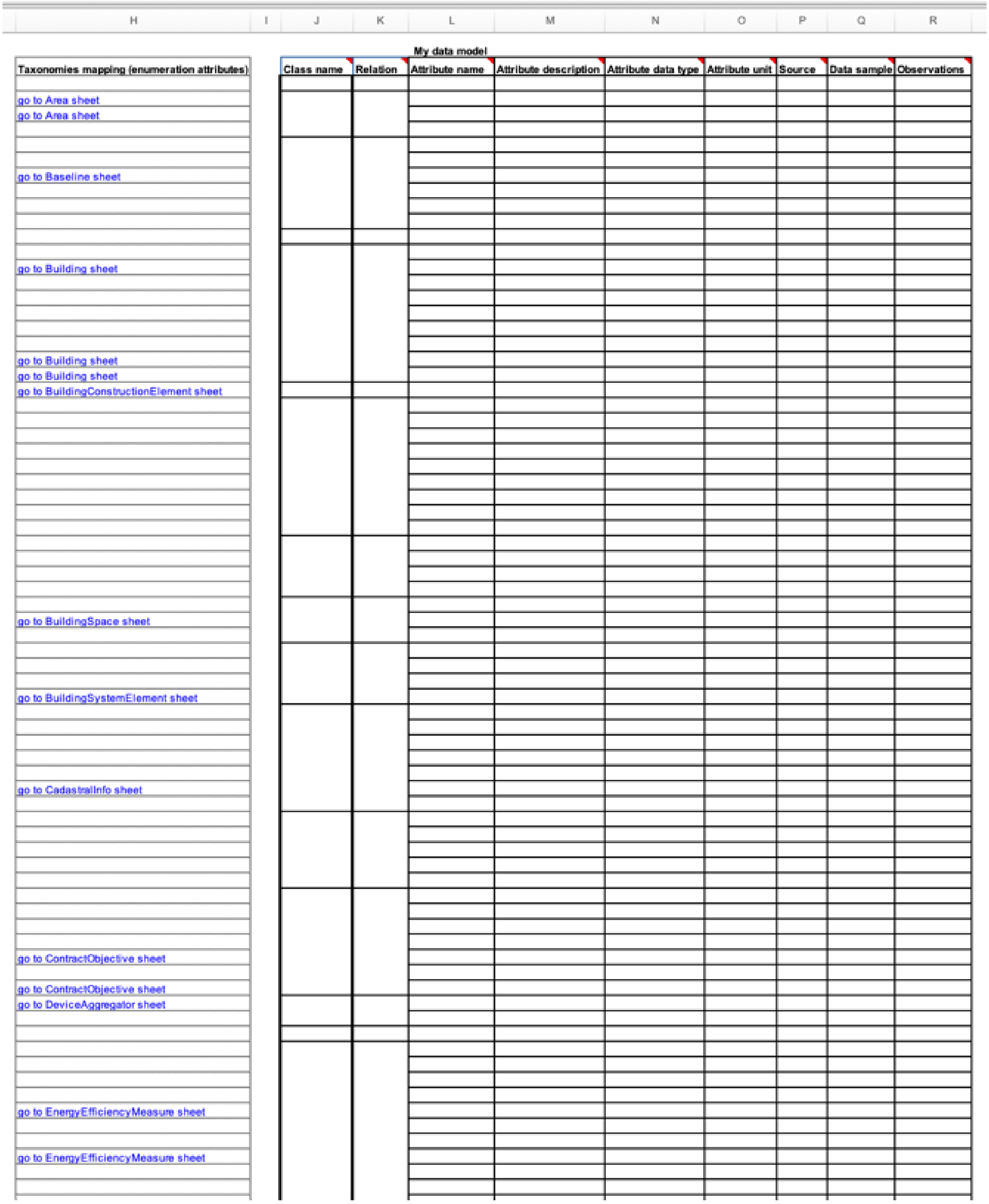

Figure 3 - Static data mapping sheet examples

The mapping table contains the following columns and instructions for filling.

Class name

In this column, put each attribute's class name.

For one-to-many matches separate the classes by commas.

Relation

Some data source schemas contain attributes that pertain to different resources, sometimes used as foreign keys for linking, contain duplicated values, or store different value.

When more attributes match the description of the data BIGG data model attribute, for describing the relation between them should be specified respectively: linking, duplicated or distinct. Only one relation per pair should be specified if fixed relations are present.

Attribute name

For one-to-one matches, put the name of the attribute that fits the best with the description.

For one-to-many matches, separate the attributes using commas.

For the attributes that at the same time contain numerical data and whose names are taxonomies, map them also at the corresponding taxonomy sheet adding as observation the word “name”.

Attribute description

Add a short description of each attribute. If more than one attribute was declared in the attribute name’s column, use commas to separate the descriptions.

Attribute data type

Information is stored using different data types. In this column specify the data type of each attribute.

Attribute unit

The attributes with numerical data type may be linked to a unit in an implicit or explicit way. Fill this column with the corresponding unit.

Source

Data could have distinct origins within a system.

Fill this column with each attribute’s name of the data source.

Data sample

Add an input example for each attribute.

Observations

Some attributes would require transformations to match with the objective data model fields. Write a short description of what should be done to match the data model attributes if required.

Missing fields table

If an incoming attribute, that you considered relevant, is missing in the BIGG data model, there is a table at the bottom of the mapping table (in yellow) where you can describe the field to be put into consideration for the next BIGG data model version.

VII.2. Time series data mapping

Within this sheet the BIGG data model classes and attributes related to time series data are represented as columns. A colour code indicates obligatory, desired and optional mapping fields. For a timeseries data input to be successfully mapped red coloured attributes (columns) should be filled, yellow columns are desired and white ones are optional.

Each time series input, filled by rows, should be implicitly or explicitly mapped following the previously described colour keys. If enumeration type input attributes have their counterpart in the BIGG data model they are considered as explicit and should be mapped using its attribute source name. On the other hand, implicit mapping occurs when there are input attributes that includes information regarding other attributes implicitly. For instance, units are common to be implicit within time series data given that the property measured has a commonly known/used unit of measurement, or when data providers know it without putting it explicitly. If the field in discussion is implicit and of a enumeration type, it should be filled according to its taxonomy sheet using quotation marks.

Time series mapping

		Class name						
		Device		Measurement			MeasurementList	
		deviceID	deviceType	measurementEnd	measurementStart	measurementValue	measuredProperty	measurementUnit
		Unique identifier for the device	Type of the device	Final timestamp of the reference period associated with the the measurement	Initial timestamp of the reference period associated with the measurement	Value of the property measured by the device	Physical property measured by the device	Unit of measurement of the property values
	Source class	UID	enum	datetime	datetime	any	enum	enum
1	hourly_consumption	cups	"SmartMeter"	datetime		consumptionKWh	"EnergyConsumptionGridElectricity"	"kWh"
2	quarter_hourly_consumption	cups	"SmartMeter"	datetime		consumptionKWh	"EnergyConsumptionGridElectricity"	"kWh"
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								

Figure 4 - Timeseries data mapping sheet example (obligatory and desired fields)

VII.3. Taxonomy mappings

After identifying a static attribute that maps the best to an enumeration data type attribute, the data provider should map its taxonomy with that of BIGG data model. For each enumeration type attributes a separate sheet has been created containing different taxonomy levels. Defined taxonomies were disaggregated up to four levels, but the provider can choose which level fits the best with his preferred taxonomies descriptions.

BIGG data model

Attri	Class name	Attribute name	Attribute description	Class definition	Taxonomy 1st level	Taxonomy 1st level descriptio
enum	Area	areaType	Type of measure area	An area measurement of a BuildingSpace	CadastralParcelArea	The area of the cadastral parcel over which the building is constricuted
					CooledFloorArea	Part of the net floor area that is cooled or air conditioned
					GrossFloorArea	Sum of the floor areas of all building spaces measured from the external faces of the exterior
					GrossFloorAreaAboveGround	Part of the gross floor area above ground
					GrossFloorAreaUnderGround	Part of the gross floor area below ground
					HeatedFloorArea	Part of the net floor area that is heated
					NetFloorArea	Gross floor area, excluding the area occupied by walls and partitions, the circulation area
		areaUnitOfMeasur	Unit of measurement of t	An area measurement of a B	ft2	Square feet
					m2	Square meters
					Other	Other unit of measurement for areas
Total Result						

My data model

Attribute name	Taxonomy	Description	Source	Observations

Missing fields

Attribute name	Taxonomy	Description	Source	Observations

Figure 5 - Taxonomies' mapping sheet example

VIII. CONCLUSIONS AND NEXT STEPS

The report covers the development of the harmonisation layer in the initial project period and presents the preliminary results. In this period the work was focused on the development of the BIGG Standard Data Model 4 Buildings in the core of the harmonisation layer that provides the semantics and structure of the data and enables their adequate allocation in databases and use in analytics services. The work started with analysis of the data requirements over the BIGG Use Cases and the available datasets from the pilots leading to identification of data concepts and relations between them. In parallel, a preliminary analysis of existing ontologies identified semantic coincidences with the BIGG data concepts that could be reused. On the base of these analyses, the initial BIGG data model was created in iterative steps of revisions, addition, and reorganisation of the data. The BIGG Standard Data Model 4 Buildings is comprised by detailed definition of classes, attributes, data types, relations, and a UML class diagram. In this initial phase, the data model was used as a common reference for mapping of the available data sources in order to enable the elaboration and testing of the first version (V1) of the pilot solutions and the AI Analytics Toolbox.

In a next step, during the second project period, the data model will further evolve into a BIGG ontology. By reusing and aligning to existing data standards, the BIGG Standard Data Model 4 Buildings implemented as ontology will set the base for high level of data interoperability for the targeted use cases in the Buildings domain. The analysis of existing ontologies and the recommendations following from it suggest that V2 of the BIGG data model should be more coherent with the IFC and SOSA data models, as these are the base for other domain ontologies, such as SAREF.

IX. References

[1]. Energy Efficiency – the first fuel for the EU Economy: How to drive new finance for energy efficiency investments, Final Report covering Buildings, Industry and SMEs, EEFIG, February 2015, ISBN: 978-84-606-6087-3.

[2]. Building Energy Performance Gap Issues: An International Review. IPEEC Building Energy Efficiency Taskgroup, 2019 https://www.energy.gov.au/sites/default/files/the_building_energy_performance_gap-an_international_review-december_2019.pdf

[3]. IEA EBC Annex 53: Total Energy Use in Buildings – Analysis and Evaluation Methods. https://www.iea-ebc.org/Data/publications/EBC_PSR_Annex53.pdf

[4]. Pritoni, M., et al., 2021. Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis. Energies, 14(7).

[5]. Cuenca, J., Larrinaga, F. & Curry, E., 2017. A Unified Semantic Ontology for Energy Management Applications. Vienna, s.n.

[6]. Bergman, M. K., 2018. A Knowledge Representation Practionary. Coralville: Springer. https://doi.org/10.1007/978-3-319-98092-8

[7]. Esnaola-Gonzalez, I., Bermúdez, J., Fernandez, I. & Arnaiz, A., 2020. Ontologies for Observations and Actuations in Buildings: A Survey. Semantic Web, 11(4), pp. 593-621

[8]. Corcho, O., Fernández-López, M. & Gómez-Pérez, A., 2007. Ontological Engineering: What are Ontologies and How Can We Build Them?. En: J. Cardoso, ed. Semantic Web Services: Theory, Tools and Applications. New York: IGI Global, pp. 44-70.

[9]. Euzenat, J., Mocan, A. and Scharffe, F., 2008. Ontology Alignments. In: M. Hepp, P. De Leenheer, A. De Moor and Y. Sure, ed., Ontology Management, 1st ed. New York City: Springer, pp.177-206.

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

X. ANNEXES

X.1. BIGG Standard Data Model 4 Buildings

X.1.1. Classes' definition

Class name	Class definition	Attribute name
AnalyticsResults	Contains all of the analytics results (to be further defined)	(empty)
Area	An area measurement of a BuildingSpace	areaType
		areaUnitOfMeasurement
		areaValue
Baseline	Energy performance of the company or a process before and after implementation of new actions for improving the energy efficiency	baselineAdjustedValue
		baselineDefinition
		baselineInfluenceFactor
		baselineNonAdjustedValue
		baselinePerimeter
		baselineUnit
Benchmarking	Benchmarking output (to be further defined).	(empty)
Building	A building for which data is provided in EN-TRACK's platform.	buildingClosingHour
		buildingConstructionType
		buildingConstructionYear
		buildingIDFromOrganization
		buildingName
		buildingOpeningHour
		buildingOwnership
BuildingConstruction Element	Any static element of the building construction (eg. walls, windows, roofs).	buildingConstructionElementType
BuildingElement	Any element of the Building which does not fall in the Device class. The type of the BuildingElement can be further specified through its subclasses BuildingConstructionElement and BuildingSystemElement	buildingElementBrand
		buildingElementInstallationDate
		buildingElementManufactureDate
		buildingElementManufacturer
		buildingElementModel
		buildingElementPurchaseDate
		buildingElementSerialNumber
		buildingElementState
		buildingElementIDFromOrganization
BuildingSpace	A space that can represent one or more rooms, floors, or zones of a Building, defined according to their use, or the necessity to separate monitoring and accounting of their energy use or performance. One BuildingSpace will be generated by default for each building, corresponding to the entire construction.	buildingSpaceName
		buildingSpaceUseType
		builidingSpaceIDFromOrganization
BuildingSystemElem ent	Any system providing a service to the Building (e.g. HVAC system, lighting system, electric power system, …) or any of their sub-	buildingSystemElementEfficiency
		buildingSystemElementMaxOutput
		buildingSystemElementMinOutput

	components (e.g. boilers, luminaries, solar photovoltaic panels, …).	buildingSystemElementType
CadastralInfo	Information from Cadastre	landArea
		landCadastralReference
		landGeometry
		landGraphicalArea
		landType
		propertyClass
CO2EmissionsFactor	Kilogram of CO2 equilvalent per kWh and its validity timeband	gridCO2EmissionsFactor
		gridCO2EmissionsFactorEnd
		gridCO2EmissionsFactorStart
Contract	Information about contract	(empty)
ContractObjective	Information about contract objective	objectiveDeadline
		objectiveDescription
		objectiveID
		objectiveName
		objectiveTargetType
		objectiveTargetValue
		objectiveTargetValueUnit
Device	Any meter, sensor, or actuator that can capture a signal, emit a signal, or assume a state that can be recorded in the form of time series data.	deviceInputProtocol
		deviceInputSignalType
		deviceLicenceVersionNumber
		deviceManufacturer
		deviceModel
		deviceName
		deviceNumberOfOutputs
		deviceOperatingSystem
		deviceType
		deviceIDFromOrganization
DeviceAggregator	A virtual unit capable of performing calculations, defined through a mathematical formula, and elaborating data from Buildings, BuildingSpaces, or Devices.	deviceAggregatorType
		formula
DeviceHistory	A set of information collected to contemplate the replacement of a Device (e.g. a smart meter) for maintenance reasons, in order to keep track of the device serial number and the period of installation.	deviceInstallationDate
		deviceManufactureDate
		deviceRemovalDate
		deviceSerialNumber
		deviceThresholdValue
Element	Any generic element of the building. The type of Element can be further specified through its subclasses BuildingElement and Device.	(empty)
EnergyEfficiencyMeasure	Any measure for the improvement of the efficiency of a Building or its Elements.	energyEfficiencyMeasureCurrencyExchangeRate
		energyEfficiencyMeasureDescription
		energyEfficiencyMeasureID
		energyEfficiencyMeasureInvestment

		energyEfficiencyMeasureInvestmentCurrency
		energyEfficiencyMeasureOperationalDate
		energyEfficiencyMeasureSavingsToInvestmentRatio
		energyEfficiencyMeasureType
		energySourcePriceEscalationRate
		shareOfAffectedElement
		energyEfficiencyMeasureCO2Reduction
		energyEfficiencyMeasureFinancialSavings
		energyEfficiencyMeasureLifetime
EnergyPerformanceCertificate	Information from the Energy Performance Certificate	annualCO2Emissions
		annualCoolingCO2Emissions
		annualCoolingEnergyDemand
		annualCoolingPrimaryEnergyConsumption
		annualEnergyCost
		annualFinalEnergyConsumption
		annualHeatingCO2Emissions
		annualHeatingEnergyDemand
		annualHeatingPrimaryEnergyConsumption
		annualHotWaterCO2Emissions
		annualHotWaterPrimaryEnergyConsumption
		annualLightingCO2Emissions
		annualLightingPrimaryEnergyConsumption
		annualPrimaryEnergyConsumption
		coolingCO2EmissionsClass
		coolingEnergyDemandClass
		coolingPrimaryEnergyClass
		energyPerformanceCertificateDateOfAssessment
		energyPerformanceCertificateDateOfCertification
		energyPerformanceCertificateReferenceNumber
		energyPerformanceCertificationMotivation
		energyPerformanceCertificationTool
		energyPerformanceClass
		energyPerformanceProcedureType
		heatingCO2EmissionsClass
		heatingEnergyDemandClass
		heatingPrimaryEnergyClass
		hotWaterCO2EmissionsClass
		hotWaterPrimaryEnergyClass
		lightingCO2EmissionsClass
		lightingPrimaryEnergyClass
		CO2EmissionsClass

EnergyPerformanceCertificateAdditionalInfo	Additional information from the Energy Performance Certificate	averageFacadeTransmittance
		averageWindowsTransmittance
		biomassSystemPresence
		buildingTecnicalInspectionCode
		constructionRegulation
		districtHeatinOrCoolingConnection
		electricVehicleChargerPresence
		geothermalSystemPresence
		regulationValueForFacadeTransmittance
		regulationValueForWindowsTransmittance
		solarPVSystemPresence
		solarThermalSystemPresence
EnergyPerformanceContract	Information about the Energy Performance Contract	contractEndDate
		contractID
		contractName
		contractPerimeter
		contractStartDate
EnergySaving	Any estimate or measure of energy savings triggered by a RenovationProject or EnergyEfficiencyMeasure	energySavingEndDate
		energySavingIndependentlyVerified
		energySavingStartDate
		energySavingType
		energySavingValue
		energySavingVerificationSource
Group	A collection of entities. It can be further differenciated into Zone (collection of building spaces) or System (collection of elements)	groupID
		groupName
IndoorQualityPerception	Information from the building occupant about his perception on the indoor comfort quality	indoorQualityEvaluationValidityEndDate
		indoorQualityEvaluationValidityStartDate
		indoorQualityUserPerception
LocationInfo	The collection of information related to the geographical location of a Building.	addressAltitude
		addressCity
		addressClimateZone
		addressCountry
		addressLatitude
		addressLongitude
		addressPostalCode
		addressProvince
		addressStreetName
		addressStreetNumber
MaintenanceAction	Information about maintenance actions	maintenanceActionDate
		maintenanceActionDescription
		maintenanceActionFrequency
		maintenanceActionID
		maintenanceActionIsPeriodic
		maintenanceActionName
		maintenanceActionType
Measurement		measurementEnd

	Any timeseries record registered by a Device.	measurementStart
		measurementValue
MeasurementList	A collection of Measurements from the same Device that measure the same property in the same measurement units.	measuredProperty
		measurementDescription
		measurementReadingType
		measurementSourceForEnergy
		measurementTypeForEnergy
		measurementUnit
		outputProtocol
		outputSignalType
		measurementListEnd
		measurementListFrequency
		measurementListStart
NonEnergyBenefit	Any additional benefit produced by Renovation Projects and Energy Efficiecy Measures other than Energy Savings.	nonEnergyBenefitImpactEvaluation
		nonEnergyBenefitImpactValue
		nonEnergyBenefitImpactValueDescription
		nonEnergyBenefitImpactValueVerifiedAndMeasured
		nonEnergyBenefitImpactVerificationMethod
		nonEnergyBenefitServiceWithNegativeImpact
		nonEnergyBenefitServiceWithPositiveImpact
		nonEnergyBenefitType
		nonEnergyBenefitImpactValueUnit
OccupancyProfile	The information related to the occupancy of a Building.	occupancyNumberOfOccupants
		occupancyProfileValidityEndDate
		occupancyProfileValidityStartDate
		occupancyVacationDates
Organization	A company or institution that provides data to EN-TRACK's platform and/or benefits from its services	organizationContactPersonName
		organizationEmail
		organizationID
		organizationName
		organizationTelephoneNumber
		organizationType
		organizationDivisionType
Person	A user of service or participant in a BIGG use case.	userEmail
		userID
		userName
		userRole
Project	General information about a project.	geometrySRID
		projectDescription
		projectID
		projectTitle
RenovationProject	A project that affects a whole Building or part of it, and can indirectly have an effect on the energy efficiency.	projectCurrencyExchangeRate
		projectDiscountRate
		projectGrantsShareOfCosts
		projectIncludedConfortmeterSurvey
		projectIncludedNonEnergyBenefitsEstimate
		projectInterestRate
		projectInternalRateOfReturn
		projectInventivesShareOfRevenues

		projectInvestment
		projectInvestmentCurrency
		projectMotivation
		projectNetPresentValue
		projectOperationalDate
		projectReceivedGrantFunding
		projectSavingsToInvestmentRatio
		projectSimplePaybackTime
		projectStartDate
		projectUsesIncentives
RetrofitProject	Any retrofit project that affects a whole Building or part of it, and that consists of one or more EnergyEfficiencyMeasures.	projectCurrencyExchangeRate
		projectDiscountRate
		projectGrantsShareOfCosts
		projectIncludedConfortmeterSurvey
		projectIncludedNonEnergyBenefitsEstimate
		projectInterestRate
		projectInternalRateOfReturn
		projectInventivesShareOfRevenues
		projectInvestment
		projectInvestmentCurrency
		projectMotivation
		projectNetPresentValue
		projectOperationalDate
		projectReceivedGrantFunding
		projectSavingsToInvestmentRatio
		projectSimplePaybackTime
		projectStartDate
		projectUsesIncentives
		projectCO2Reduction
State	A record of the particular condition that a Device is found in at a specific time.	state
		stateEnd
		stateStart
		stateType
System	A group of elements	systemType
Tariff	The specifications of a tariff associated with one of the metered commodities.	tariffCompany
		tariffEndDate
		tariffName
		tariffStartDate
		tariffAveragePrice
UtilityPointOfDelivery	A point on the utility distribution system where the deliverer makes the utility available to a receiver or to serve load.	utilityType
		pointOfDeliveryIDFromOrganization
WeatherStation	A weather station that provides weather data of interest for one or more Buildings.	weatherStationCoordinates
		weatherStationEndDate
		weatherStationStartDate
		weatherStationType
		weatherStationTimeStep
Zone	A group of building spaces	zoneType

X.1.2. Attributes' definition

Attribute name	Attribute description	Attribute data type
(empty)	(empty)	(empty)
areaType	Type of measure area	enum
areaUnitOfMeasurement	Unit of measurement of the area value	enum
areaValue	Numerical value of the area	float
baselineAdjustedValue	The value of the baseline after adjustment for the influence factors	float
baselineDefinition	Short textual description of the baseline and the reference situation it describes	string
baselineInfluenceFactor	Field to identify of the factors that impact the baseline and should be used to adjust it (eg. HDD; CDD; Occupancy; ProductionLevel...)	enum
baselineNonAdjustedValue	The value of the baseline before adjustment for the influence factors	float
baselinePerimeter	Definition of the baseline perimeter (eg. whole building, a given equipment, a group of equipment, a specific zone, a specific utility...)	string
baselineUnit	The unit in which the baseline is defined	string
(empty)	(empty)	(empty)
buildingClosingHour	Closing hour of the building in normal working days	time
buildingConstructionType	Building construction type	enum
buildingConstructionYear	Year of building construction or major renovation	integer
buildingIDFromOrganization	Identifier given to the building from the organization that owns it	string
buildingName	Name of the building	string
buildingOpeningHour	Opening hour of the building in normal working days	time
buildingOwnership	Indication of whether the building is owned or rented	enum
buildingConstructionElementType	Type of the building construction element	enum
buildingElementBrand	Brand of the building element	string
buildingElementInstallationDate	Date of installation of the building element	date
buildingElementManufactureDate	Manufacture date of the building system element	date
buildingElementManufacturer	Manufacturer of the building system element	string
buildingElementModel	Model of the building element	string
buildingElementPurchaseDate	Date of purchase of the building element	date
buildingElementSerialNumber	Serial number of the building element	string
buildingElementState	(empty)	boolean
buildingElementIDFromOrganization	Unique identifier for the building element reported from the organization	string
buildingSpaceName	Name of the building space	string
buildingSpaceUseType	Purpose of use or activity conducted in the building space	enum
builidingSpaceIDFromOrganization	Unique identifier for the building space as reported from the organization	string

Attribute name	Attribute description	Attribute data type
buildingSystemElementEfficiency	Percentual effieciency of the building system element	%
buildingSystemElementMaxOutput	Maximum output of the building system element in kW	float
buildingSystemElementMinOutput	Minimum output of the building system element in kW	float
buildingSystemElementType	Type of the building system element	enum
landArea	Area of the land	float
landCadastralReference	Cadastral reference of the building	string
landGeometry	Land geometry in in WGS84 crs format	polygon
landGraphicalArea	(empty)	int
landType	(empty)	enum
propertyClass	(empty)	string
gridCO2EmissionsFactor	Kilogram of CO2 equilvalent per kWh	float
gridCO2EmissionsFactorEnd	Emissions factor validity end datetime	datetime
gridCO2EmissionsFactorStart	Emissions factor validity start datetime	datetime
(empty)	(empty)	(empty)
objectiveDeadline	Deadline date for the achievement of the objective	date
objectiveDescription	Textual desciption of the objective	string
objectiveID	Unique identifier for the energy performance contract objective	UID
objectiveName	Name of the energy performance contract objective	string
objectiveTargetType	Field used o indicate whether the target of the contract is expressed either by an absolute or a relative value	enum
objectiveTargetValue	Value of the target of the contract. It has to be considered in combination with the ObjectiveTargetType	float
objectiveTargetValueUnit	Indication of the unit of measurement of the target value	enum
deviceInputProtocol	(empty)	enum
deviceInputSignalType	(empty)	enum
deviceLicenceVersionNumber	Number of the licence version of the device	string
deviceManufacturer	Name of the device manufacturer	string
deviceModel	Model of the device	string
deviceName	Descriptive name of the device	string
deviceNumberOfOutputs	Number of outputs the device produces. Each output should be connected to a different MeasurementList	int
deviceOperatingSystem	Operating system of the device	string
deviceType	Type of the device	enum
deviceIDFromOrganization	Unique identifier for the device reported by the organization	string
deviceAggregatorType	Physical property calculated	enum
formula	Formula of the calculation performed by the device aggregator	string
deviceInstallationDate	Date of installation of the device	datetime
deviceManufactureDate	Date of manufacturing of the device	date
deviceRemovalDate	Date of removal of the device	datetime
deviceSerialNumber	Serial number of the device	string
deviceThresholdValue	Threshold of recorded value after which the device will have to be replaced	float

Attribute name	Attribute description	Attribute data type
(empty)	(empty)	(empty)
energyEfficiencyMeasureCurrencyExchangeRate	Exchange rate between the original investment currency and euros	float
energyEfficiencyMeasureDescription	Description of the energy efficiency measure applied	string
energyEfficiencyMeasureID	Unique identifier of the energy efficiency measure	UUID
energyEfficiencyMeasureInvestment	Investment for the energy efficiency measure inplementation	float
energyEfficiencyMeasureInvestmentCurrency	Original currency of the energy efficiency measure investment	enum
energyEfficiencyMeasureOperationalDate	Date on which the energy efficiency measure became operational	date
energyEfficiencyMeasureSavingsToInvestmentRatio	Estimated Savings To Investment Ratio (SIR) for the energy efficiency measure	float
energyEfficiencyMeasureType	The type of energy efficiency measure	enum
energySourcePriceEscalationRate	Escalation rate of the price of the energy source related to the described energy efficiency measure (if applicable)	%
shareOfAffectedElement	Percentage of the element that is affected by the energy efficiency measure (eg. 70% of the whole building fabric, 50% of the HVAC system, …)	%
energyEfficiencyMeasureCO2Reduction	Annual reduction of CO2 emmisions by the energy efficiency measure, tCO2/year	float
energyEfficiencyMeasureFinancialSavings	Annual financial savings of energy efficiency measure	float
energyEfficiencyMeasureLifetime	Applied energy efficiency measure lifetime.	float
annualCO2Emissions	Value of annual CO2 emissions [kg CO2/m2*year]	float
annualCoolingCO2Emissions	Value of annual CO2 emissions associated with the cooling service [kg CO2/m2*year]	float
annualCoolingEnergyDemand	Value of annual energy demand associated with the cooling service [kWh/m2*year]	float
annualCoolingPrimaryEnergyConsumption	Value of annual consumption of non-renewable primary energy associated with the cooling service [kWh/m2*year]	float
annualEnergyCost	Annual energy cost for the reference building	float
annualFinalEnergyConsumption	(empty)	float
annualHeatingCO2Emissions	Value of annual CO2 emissions associated with the heating service [kg CO2/m2*year]	float
annualHeatingEnergyDemand	Value of annual energy demand associated with the lighting service [kWh/m2*year]	float
annualHeatingPrimaryEnergyConsumption	Value of annual consumption of non-renewable primary energy associated with the heating service [kWh/m2*year]	float
annualHotWaterCO2Emissions	Value of annual CO2 emissions associated with the domestic hot water service [kg CO2/m2*year]	float
annualHotWaterPrimaryEnergyConsumption	Value of annual consumption of non-renewable primary energy associated with the hot water service [kWh/m2*year]	float
annualLightingCO2Emissions	Value of annual CO2 emissions associated with the lighting service [kg CO2/m2*year]	float
annualLightingPrimaryEnergyConsumption	Value of annual consumption of non-renewable primary energy associated with the lighting service [kWh/m2*year]	float

Attribute name	Attribute description	Attribute data type
annualPrimaryEnergyConsumption	Value of annual consumption of non-renewable primary energy [kWh/m2*year]	float
coolingCO2EmissionsClass	Class letter assigned for CO2 emissions associated with the cooling service	string
coolingEnergyDemandClass	Class letter assigned for the energy demand associated with the cooling service	string
coolingPrimaryEnergyClass	Class letter assigned for the consumption of non-renewable primary energy associated with the cooling service	string
energyPerformanceCertificateDateOfAssessment	Date of assesment of the building for the production of the energy performance certificate	date
energyPerformanceCertificateDateOfCertification	Date of release of the energy performance certificate	date
energyPerformanceCertificateReferenceNumber	Reference number as reported on the energy performance certificate	string
energyPerformanceCertificationMotivation	Description of the reason behind the realization of the energy performance certification	string
energyPerformanceCertificationTool	Tool utilized to realized the energy performance certification	string
energyPerformanceClass	Class letter assigned for the consumption of non-renewable primary energy	string
energyPerformanceProcedureType	(empty)	string
heatingCO2EmissionsClass	Class letter assigned for CO2 emissions associated with the heating service	string
heatingEnergyDemandClass	Class letter assigned for the energy demand associated with the heating service	string
heatingPrimaryEnergyClass	Class letter assigned for the consumption of non-renewable primary energy associated with the heating service	string
hotWaterCO2EmissionsClass	Class letter assigned for CO2 emissions associated with the domestic hot water service	string
hotWaterPrimaryEnergyClass	Class letter assigned for the consumption of non-renewable primary energy associated with the hot water service	string
lightingCO2EmissionsClass	Class letter assigned for CO2 emissions associated with the lighting service	string
lightingPrimaryEnergyClass	Class letter assigned for the consumption of non-renewable primary energy associated with the lighting service	string
CO2EmissionsClass	Class letter assigned for CO2 emissions	string
averageFacadeTransmittance	(empty)	float
averageWindowsTransmittance	(empty)	float
biomassSystemPresence	(empty)	boolean
buildingTecnicalInspectionCode	(empty)	string
constructionRegulation	(empty)	string
districtHeatinOrCoolingConnection	(empty)	boolean
electricVehicleChargerPresence	(empty)	boolean
geothermalSystemPresence	(empty)	boolean
regulationValueForFacadeTransmittance	(empty)	float
regulationValueForWindowsTransmittance	(empty)	float

Attribute name	Attribute description	Attribute data type
solarPVSystemPresence	(empty)	boolean
solarThermalSystemPresence	(empty)	boolean
contractEndDate	Final date of validity of the contract	date
contractID	Unique identifier for the energy performance contract	UID
contractName	Name of the energy performance contract	string
contractPerimeter	Definition of the boundaries of the contract in terms of buildings, zones, equipment that are included in its scope	string
contractStartDate	Initial date of validity of the contract	date
energySavingEndDate	Final date of the reference period for the energy savings	date
energySavingIndependentlyVerified	Indication of whether the presented energy savings were independently verified or not	bool
energySavingStartDate	Initial date of the reference period for the energy savings	date
energySavingType	Energy savings type to be selected from a predefined list	enum
energySavingValue	Energy savings value	any
energySavingVerificationSource	Source of the verification, if the presented energy savings were independently verified	enum
groupID	Unique identifier for the group	UID
groupName	Name of the group	string
indoorQualityEvaluationValidityEndDate	Final date of the reference period for the indoor quality evaluation	date
indoorQualityEvaluationValidityStartDate	Initial date of the reference period for the indoor quality evaluation	date
indoorQualityUserPerception	(empty)	enum
addressAltitude	Altitude of the building location	float
addressCity	City where the building is located	enum
addressClimateZone	Identification of the climate zone that corresponds to the address coordinates	enum
addressCountry	Country where the building is located	enum
addressLatitude	Latitude coordinate of the building location	float
addressLongitude	Longitude coordinate of the building location	float
addressPostalCode	Postal code of the building location	string
addressProvince	Province/administrative district of the building location	enum
addressStreetName	Street name of the building address	string
addressStreetNumber	Street number of the building address	string
maintenanceActionDate	Date in which the maintenance action occurred	date
maintenanceActionDescription	Textual description of the maintenance action	string
maintenanceActionFrequency	Indication of the frequency of required maintenance in days, if the maintenace action is periodic	int
maintenanceActionID	Unique identifier for the maintenance action	UID
maintenanceActionIsPeriodic	Boolean field to indicate wether the maintenance action is periodically required or not	boolean
maintenanceActionName	Name of the maintenance action	string
maintenanceActionType	Type of the maintenance action	enum

Attribute name	Attribute description	Attribute data type
measurementEnd	Final timestamp of the reference period associated with the the measurement	datetime
measurementStart	Initial timestamp of the reference period associated with the measurement	datetime
measurementValue	Value of the property measured by the device	any
measuredProperty	Physical property measured by the device	enum
measurementDescription	Textual description of the measurement	string
measurementReadingType	Indication of the type of reading recorded and shown by the device (eg. average value, counter, etc.)	enum
measurementSourceForEnergy	Source from which the measurement list data was obtained (valid only for energy measurement lists)	enum
measurementTypeForEnergy	Indication of whether the measurement is actual or predicted (valid only for energy measurement lists)	enum
measurementUnit	Unit of measurement of the property values	enum
outputProtocol	(empty)	enum
outputSignalType	(empty)	enum
measurementListEnd	(empty)	datetime
measurementListFrequency	(empty)	string
measurementListStart	(empty)	datetime
nonEnergyBenefitImpactEvaluation	Evaluation of the project impact over the selected non energy benefit	enum
nonEnergyBenefitImpactValue	Approximate net value of the impact of the project over the non-energy benefit	any
nonEnergyBenefitImpactValueDescription	Description of the value provided for the impact of the project over the non-energy benefit	string
nonEnergyBenefitImpactValueVerifiedAndMeasured	Indication of whether the impact over the non-energy benefit has been measured and verifed	bool
nonEnergyBenefitImpactVerificationMethod	Description of the verification/measurment method used, in case the impact over the non-energy benefit has been verified or measured	string
nonEnergyBenefitServiceWithNegativeImpact	List of energy services that were determined to negatively contribute to the selected non-energy benefit	list (enum)
nonEnergyBenefitServiceWithPositiveImpact	List of energy services that were determined to positively contribute to the selected non-energy benefit	list (enum)
nonEnergyBenefitType	Type of non-energy benefit produced by the project	enum
nonEnergyBenefitImpactValueUnit	Units of the approximate net value of the impact of the project over the non-energy benefit	enum
occupancyNumberOfOccupants	Total number of occupants at full building occupation	integer
occupancyProfileValidityEndDate	Final date for the reference period of the given occupancy profile	date
occupancyProfileValidityStartDate	Initial date for the reference period of the given occupancy profile	date
occupancyVacationDates	List of vacation days when the building is closed	list (date)

Attribute name	Attribute description	Attribute data type
organizationContactPersonName	Name of the contact person for the organization	string
organizationEmail	Contact email of the organization	string
organizationID	Unique identifier for the organization	UID
organizationName	Name of the organization	string
organizationTelephoneNumber	Telephone number of the organization	string
organizationType	Nature of the organization	enum
organizationDivisionType	Nature of the organization division	string
userEmail	Email address of the EN-TRACK user	string
userID	Unique identifier for the organization	UID
userName	Complete name of the EN-TRACK user	string
userRole	Description of the role of the user within the organization	string
geometrySRID	International spatial Reference System for geometric types	int
projectDescription	Free text for the user to describe the project (e.g 300 characters)	string
projectID	Project unique identifier	UID
projectTitle	Project title	string
projectCurrencyExchangeRate	Exchange rate between the original investment currency and euros	float
projectDiscountRate	Discount rate used to calculate the financial metrics for the renovation project	%
projectGrantsShareOfCosts	Estimated share of the total project costs that were covered with grant funding, in case the project received it	%
projectIncludedConfortmeterSurvey	Indication of whether the project included a Comfortmeter survey	boolean
projectIncludedNonEnergyBenefitsEstimate	Indication of whether the non energy benefits produced by the project were estimated	boolean
projectInterestRate	Interest rate used to calculate the financial metrics for the renovation project	%
projectInternalRateOfReturn	Estimated Internal Rate of Return (IRR) of the renovation project	%
projectInventivesShareOfRevenues	Estimated share of the total project revenues that are represented by incentives schemes, in case the project benefitted/will benefit from them	%
projectInvestment	Investment for the project inplementation	float
projectInvestmentCurrency	Original currency of the project investment	enum
projectMotivation	Key reasons for the investment	enum
projectNetPresentValue	Estimated Net Present Value (NPV) of the renovation project	float
projectOperationalDate	Date on which the project became operational	date
projectReceivedGrantFunding	Yes or no data field to express whether the projects received grant funding	boolean
projectSavingsToInvestmentRatio	Estimated Savings To Investment Ratio (SIR) of the renovation project	float
projectSimplePaybackTime	Estimated Simple Payback Time (SPB) of the renovation project	float
projectStartDate	Date on which the project investment started	date

Attribute name	Attribute description	Attribute data type
projectUsesIncentives	Yes or no data field to express whether the projects benefitted or will benefit from incentive schemes	boolean
projectCurrencyExchangeRate	Exchange rate between the original investment currency and euros	float
projectDiscountRate	Discount rate used to calculate the financial metrics for the renovation project	%
projectGrantsShareOfCosts	Estimated share of the total project costs that were covered with grant funding, in case the project received it	%
projectIncludedConfortmeterSurvey	Indication of whether the project included a Comfortmeter survey	boolean
projectIncludedNonEnergyBenefitsEstimate	Indication of whether the non energy benefits produced by the project were estimated	boolean
projectInterestRate	Interest rate used to calculate the financial metrics for the renovation project	%
projectInternalRateOfReturn	Estimated Internal Rate of Return (IRR) of the renovation project	%
projectInventivesShareOfRevenues	Estimated share of the total project revenues that are represented by incentives schemes, in case the project benefitted/will benefit from them	%
projectInvestment	Investment for the project inplementation	float
projectInvestmentCurrency	Original currency of the project investment	enum
projectMotivation	Key reasons for the investment	enum
projectNetPresentValue	Estimated Net Present Value (NPV) of the renovation project	float
projectOperationalDate	Date on which the project became operational	date
projectReceivedGrantFunding	Yes or no data field to express whether the projects received grant funding	boolean
projectSavingsToInvestmentRatio	Estimated Savings To Investment Ratio (SIR) of the renovation project	float
projectSimplePaybackTime	Estimated Simple Payback Time (SPB) of the renovation project	float
projectStartDate	Date on which the project investment started	date
projectUsesIncentives	Yes or no data field to express whether the projects benefitted or will benefit from incentive schemes	boolean
projectCO2Reduction	Annual reduction of CO2 emissions by the project, tCO2/year	float
state	The state in which the device is found, within the selected state type category	any
stateEnd	Final timestamp of the reference period associated with the state	datetime
stateStart	Initial timestamp of the reference period associated with the state	datetime
stateType	The state type category that applies to the device (eg. On/off, start/stop, …)	enum
systemType	Type of the given system	enum
tariffCompany	Company that offers the tariff	string
tariffEndDate	Final date of the reference period for the energy savings	datetime

Attribute name	Attribute description	Attribute data type
tariffName	Name of the tariff	string
tariffStartDate	Initial date of the reference period for the energy savings	datetime
tariffAveragePrice	(empty)	float
utilityType	Indication of the type of utility delivered at the point of delivery of interest	enum
pointOfDeliveryIDFromOrganization	Unique identifier for the point of delivery as reported from the organization	string
weatherStationCoordinates	Latitude and longitude coordinates of the weather station	point
weatherStationEndDate	Final data of data retrieval from the weather station	datetime
weatherStationStartDate	Initial date of data retrieval from the weather station	datetime
weatherStationType	Type of the weather station	enum
weatherStationTimeStep	(empty)	string
zoneType	Type of the given zone	enum

X.1.3. UML Class Diagram