Grant Agreement N° 872592

PLATOON

Digital platform and analytic tools for energy

Deliverable D2.3
PLATOON Common Data Models for Energy

Contractual delivery date:
M27

Actual delivery date:
31st of March 2022

Responsible partner:
P01: ENGIE, France

Project Title	PLATOON – Digital platform and analytic tools for energy
Deliverable number	D2.3
Deliverable title	Report on Common Data Models for Energy
Author(s):	Sarra BEN ABBES and Lynda TEMAL
Responsible Partner:	P01 - ENGIE
Date:	31.03.2022
Nature	R
Distribution level (CO, PU):	PU
Work package number	WP2 – Reference Architecture, Interoperability and Stand-ardization
Work package leader	ENG

Abstract:	The purpose of this document is to provide a common semantic data models of all needs described in the business cases of the WP1-T1.1. This includes semantic knowledge about all relevant concepts and its relationships in each pilot. A common semantic data model allows heterogenous systems to share the same meaning of entities that enables them to interoperate.
Keyword List:	Use cases requirements, Semantic data models, Ontology, Methodology, Templates, Formalization

The research leading to these results has received funding from the European Community's Horizon 2020 Work Programme (H2020) under grant agreement no 872592.

Editor(s):	Sarra BEN ABBES (ENGIE) Lynda TEMAL (ENGIE)
Contributor(s):	
Reviewer(s):	Philippe CALVEZ (ENGIE) Martino Maggio (ENG)
Approved by:	Philippe Calvez (ENGIE) – PLATOON Coordinator Erik Maqueda (TECN) – Technical Coordinator Eduardo Jimenez (IND) – Exploitation Coordinator
Recommended/mandatory readers:	WP4, WP5 and WP6 partners

Document Description

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
V0	28/09/2020	First draft version	Sarra Ben Abbès
V1	19/10/2020	First draft version	Sarra Ben Abbès and Lynda Temal
V2	30/10/2020	Subsections related to Pilot #2a	Valentina Janev
V3	30/10/2020	Subsections related to Pilot #3a	Romain Petinot
V4	01/11/2020	Subsections related to Pilot #3b	Valeria Di Pasquale
V5	02/11/2020	Subsections related to Pilot #2b and #3c	Erik Maqueda Pau Joan Cortés Forteza
V6	02/11/2020	Add some parts in the inventory of existing semantic data models	Sarra Ben Abbès and Lynda Temal
V7	03/11/2020	Add sections related to step 1 and 2 of the methodology	Lynda Temal
V8	03/11/2020	Add sections related to step 3 and 4 of the methodology	Sarra Ben Abbès
V9	10/11/2020	Add seas ontology in the inventory of existing semantic data models	Oumy Seye
V10	12/11/2020	Add Saref ontology in the inventory of existing semantic data models	Oumy Seye
V11	17/11/2020	Add Background section. Update the step 1 step 2	Sarra Ben Abbès Lynda Temal
V12	18/11/2020	Update Background section and existing ontology construction methodologies	Sarra Ben Abbès
V13	19/11/2020	Update existing ontology methodologies	Sarra Ben Abbès
V14	20/11/2020	Update references of existing ontology methodologies Add the overview of the proposed methodology Add the existing semantic data models Add overview of main topics of pilots Add Figures in the consolidation part	Sarra Ben Abbès
V15	23/11/2020	Add overviews section, Step 1 and step2.	Lynda Temal, Sarra Ben Abbès

V16	24/11/2020	Add the application sections of the proposed methodology	Sarra Ben Abbès, Lynda Temal
V17	25/11/2020	Part II: semantic data model	Lynda Temal, Sarra Ben Abbès
V18	26/11/2020	Part II /Formalization Add ontologies modules	Oumy Seye
V19	26/11/2020	Part II: semantic data model	Lynda Temal, Sarra Ben Abbès
V20	27/11/2020	Part II: semantic data model	Lynda Temal, Sarra Ben Abbès
V21	30/11/2020	Add description of Pilot #4a	Marco Musetta
V23	30/11/2020	Update Ontology formalization	Oumy Seye
V22	30/11/2020	Add Introduction Conclusion Executive Summary Update state-of-the-art Update Part II	Sarra Ben Abbès Lynda Temal
V23	01/12/2020	Review Section of Ontology Formalization Add Annexes Review all parts of the stable version of D2.3.	Sarra Ben Abbès Lynda Temal
V24	02/12/2020	Review Formalization Review all sections	Sarra Ben Abbès Lynda Temal
V25	04/12/2020	Add Rome Capitale section Review all sections of D2.3	Sarra Ben Abbès Lynda Temal
V26	17/12/2020	Update document related to Erik’s comments	Sarra Ben Abbès Lynda Temal
V27	15/10/2021	Update document according to the EU commission comments	Sarra Ben Abbès Lynda Temal
V28	21/03/2022	Update the semantic data models deliverable	Sarra Ben Abbès Lynda TEMAL

Acknowledgements

The following people are hereby duly acknowledged for their considerable contributions, which have served as a basis for this deliverable:

Name	Partner
Valentina Janev	INSTITUT MIHAJLO PUPIN
Erik Mqueda	FUNDACION TECNALIA RESEARCH & INOVATION
Oumy Seye	ENGIE
Romain Petinot	ENGIE
Jan Helsen	VRIJE UNIVERSITEIT BRUSSEL
Valeria Di Pasquale	POSTE ITALIANE - SOCIETA PER AZIONI
Armando Calvanez	POSTE ITALIANE - SOCIETA PER AZIONI
Pau J. Cortés Forteza	SAMPOL INGENIERIA Y OBRAS S.A.
Patrick Maurelli	ROMA CAPITALE
Marco Mussetta	POLITECNICO DI MILANO
Hantong Liu	FRAUNHOFER IAIS
Franco La Torre	ROMA CAPITALE
Martino Maggio	ENG
Seo-Ra Kwon Kim	INDRA
Fathoni A. Musyaffa	UNIVERSITY OF BONN
Andrej Campa	COMSENSUS

Table of Contents

List of Figures

List of Tables

Terms and abbreviations

AFNOR	French Standardization Association
ANSI	American National Standard Institute
BFO	Basic Formal Ontology
BOT	Building Ontology Topology
CA	Consortium Agreement
CIM	Common Information Model
CO	Confidential
DM	Dissemination Manager
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
DSO	Distribution System Operators
EC	European Commission
EC/EP	Energy Consumption / Energy Production
ENTSO-E	European Network of Transmission System Operators
EV	Electric Vehicle
HVAC	Heating Ventilation Air Conditioning
IEC	International Electrotechnical Commission's
ISO	International Organization for Standardization
LLUC	Low Level Use Case
LLUC	Low Level Use Case
NTL	Non-Technical Losses
OEMA	Ontology for Energy Management Applications
OWL	Ontology Web Language
PV	Photovoltaic Panel
RDF	Resource Description Framework
RUL	Remaining Useful Life
SAREF	Smart Appliances REFerence ontology
SCADA	Supervisory Control And Data Acquisition
SDM	Semantic Data Model
SEAS	Semantic Energy Aware System
SSN	Semantic Sensor Network Ontology
SUMO	Suggested Upper Merged Ontology
TOVE	TOronto Virtual Enterprise
TSO	Transmission System Operators
UC	Use Case
WP	Work package
PDM	Politecnico di Milano

Executive Summary

This document aims to contribute to the harmonization and representation of semantic data models used in future energy management systems. For a complete process that supports interoperability (from heterogeneous data to semantic data and eventually knowledge-based services), it is imperative to interpret the information exchanged in a coherent and accurate manner to maintain the coherence of data. The task 2.3-Common data models play an important role to obtain semantic knowledge and to tackle the issue of semantic interoperability. These semantic data models allow heterogeneous systems to effectively be able to interoperate with each other (data and services).

This document presents a background on data models, explains an overview of existing semantic data models, describes the proposed methodology for this task and details the application of all steps of this methodology. The modeling process was accomplished in two phases. The first phase is based on the scenarios descriptions and the second phase is based on dataset provided in different pilots.

1 Introduction

Ensuring the semantic interoperability issues between heterogeneous systems is very challenging task. Considering the use of a common semantic data model is a central element to ensure this semantic interoperability. In fact, a common semantic data model allows heterogenous systems to share the same meaning of entities that enables them to interoperate, when they come to use them in their different processes and services. This work concerns the task 2.3 that aims to propose a common semantic data model to meet the needs of the pilots described in the WP1 task 1.1. Sematic Data Models (SDM) or more known as Ontologies are recognized as the corner stone element to build a common semantic data model. Indeed, ontology allows to describe domains in unambiguous manner allowing experts to reach a consensus in a specific domain. Furthermore, ontologies could be shared and reused by different actors. SAREF and CIM are examples of ontologies in the energy domain that is considered as reference frameworks. Reusing standard(ized) ontologies foster the interoperability between systems. However, existing ontologies (like SAREF, CIM, …) in the energy domain do not cover yet all needs assessed from the different pilots of the PLATOON project. Consequently, there is need to design a new ontological module to cover the PLATOON project needs.

The work presented in this deliverable has several contribution aspects: (i) the first aspect concerns the inventory of existing ontologies in the energy domain and analyse them according to their domain coverage, the overlapping between them, and the community that work on them etc., (ii) the second aspect is to show how to reuse and combine some existing ontology to represent some needs described in pilots, (iii) the third aspect is to propose new ontological modules that will cover the missed part in existing ontologies, (iv) the fourth aspect, concerns detecting modelling errors in certain ontology, (v) the fifth aspect is the strong interaction between participants and stakeholders that allows to understand more the needs and to propose a relevant modelling, and (vi) the last aspect concerns the methodology proposed, the teamwork organization and the meetings performed in order to reach the goal.

The result of the task 2.3 will be used directly in the task 2.4 that aims to transform the non-semantic data into a semantic data according to the common semantic data models. In PLATOON framework, these semantic data models will be applied in several work packages such as WP3-Data Governance, Security and Privacy and WP4 – Analytical Toolboxes, etc.

This deliverable D2.3 summarizes the work done by project partners during two phases of the project. The first phase started from month 6 to month 12 according to the definition of the use cases. The second phase started from month 25 to 27 according to real datasets of pilots. This version 2 of the deliverable D2.3 is divided into two main chapters; the first chapter covers a proposed methodology of this task, that includes the main steps to create common semantic data models for energy domain based on the description of the low-level use cases for each pilot. The second chapter focus on the application of the data models methodology in each LLUC of each Pilot described in the WP1.

2 From Classical Data Models to Semantic Data Models: PLATOON Methodology

In this section, the general concepts regarding data models will be briefly explained and we will describe the proposed methodology to create semantic data models for each LLUC of Pilots.

2.1 Introduction

The semantic interoperability is the ability of different agents, services, and applications to communicate data, information, and knowledge in a coherent purpose [33]. The integration of heterogeneous data makes it difficult to answer questions or queries (for standard requests integrated into existing processes), and to pursue data sharing (e.g., several domains in energy sector: smart grids, building, wind turbine, city).

This part is organized as follows. Section 2 clarifies some important notions concerning existing types of data models and methodologies of ontology construction. Section 3 details the proposed methodology to build the Semantic Common Data Models. Section 4 presents an overview of all use cases of the different pilots in PLATOON and finally the section 5 concludes this part with the key topics of these use cases.

2.2 Background

Before starting any presentation of the methodology used to design semantic data models. It is important to recall some keys notions that help to avoid meaning confusing between them. This section presents some definitions and distinction between different kind of data models (relational, geographic, and semantic). Section 2.2.1 briefly presents the classic data models such us relational, geographic data model. Section 2.2.2 explains the semantic data models such us terminology, thesaurus, taxonomy, and ontology. Because, in our work we focus on constructing data models based on ontologies, section 2.2.3 outlines a brief state-of-the-art of existing ontology construction methodologies.

2.2.1 Classic Data Models

Data model is an abstract model that describes the structure of a database, the operations for manipulating these structures, and certain constraints that the database should obey. There are different types of classic data models: relational, geographic, etc.

a) **Relational data model** [1]: describes the data and its relationships in the relational database. The relational model expresses the data and relationships between the data in the form of tables.
b) **Geographic data model** [2]: represents geographic objects or surfaces as data. It can be represented as a vector data model (set of geography points lines and polygons), raster data model (geography cell matrices) and TIN data model (a set of geography of triangles).

2.2.2 Semantic Data Models

Semantic data model: "represents data in terms of named sets of objects, named sets of values, named sets of relationships, and constraints over these object, value, and relationship sets" [8].
A semantic data model is generally defined as a set of semantic entities (e.g., concept, term, descriptor, instances, properties) linked together by a set of relationships (e.g., hierarchical relationship, associative relationship). Each semantic data model has a nature and a specificity that is described in the following subsection [20].

a) **Taxonomy:** A taxonomy is defined by [14] as a set of concepts linked by “is-a” relations. More precisely, a taxonomy is a hierarchy of concepts organized in the sense of specialization (from the general to the specific). Figure 1 shows an extract of the building taxonomy.

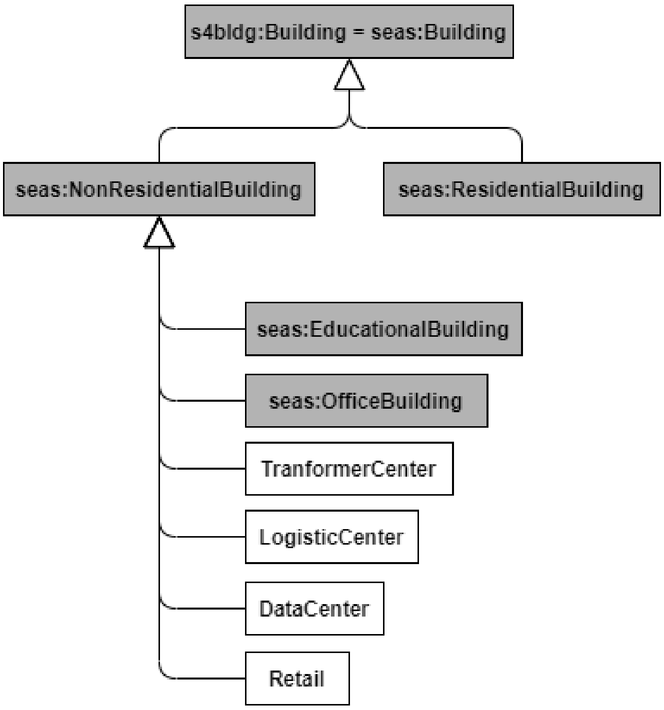


Figure 1: Taxonomy of the concept "Building"

b) **Ontology:** the most widely cited definition of an ontology is that of Gruber (1993) [11] who said that an ontology is “a formal, explicit specification of a shared conceptualization, used to help programs and humans share knowledge.” This definition highlights the following characteristics of an ontology:
   - **Conceptualization**: defined the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them [34]. A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose [11].
   - **Explicit:** corresponds to the precise definition of the concepts and the constraints of their use.
   - **Formal:** refers to the fact that the expressions must be machine readable.
   - **Shared:** refers to a common understanding of domain knowledge among people or agents.

Different types of ontologies [9] are proposed in the literature: (i) Upper ontologies (e.g., SUMO, Dolce, BFO), (ii) Core ontologies (e.g., Agent Ontology, Time Ontology), (iii) Domain ontologies for a specific domain (e.g., SEAS, SAREF, SSN) and (iv) Application ontologies that can reuse a former type of ontologies and extends them to meet a specific need of the application. Figure 2 shows an extract of building ontology.

c) Other semantic data models:

Different other semantic data models are defined such as:

- **Thesaurus**: To index documents in a specific domain, a controlled and structured terminology is used, called the thesaurus. The definition of the ISO2788-1986 [35] and ANSIZ39 [36] standards, seems to us to be the most complete: "a thesaurus is an organized authority list of descriptors and non-descriptors obeying their own terminological rules and linked together by semantic relations (hierarchical, associative or equivalence)."
- **Dictionary**: a kind of reference book that provides a list of words with their meaning, pronunciation, phonetic symbols, etc. [13].
- **Terminology**: [19] has defined a terminology as "a list of terms in a specific domain or subject, referring to notions that are frequently used. This list is unstructured." To structure the terms of a domain, it is important to link them together. This is called a terminological network. [5] have defined the terminological network as a set of terms (of words) linked together by lexical relationships, namely [6], synonymy, hypernymy/hyponymy, antonymy.
- **Glossary**: it covers only the words and its definition relating to a particular subject [15]

In the Task 2.3: Common Data Models, we will focus on two types of semantic data models which are **Ontology** and **Taxonomy**.

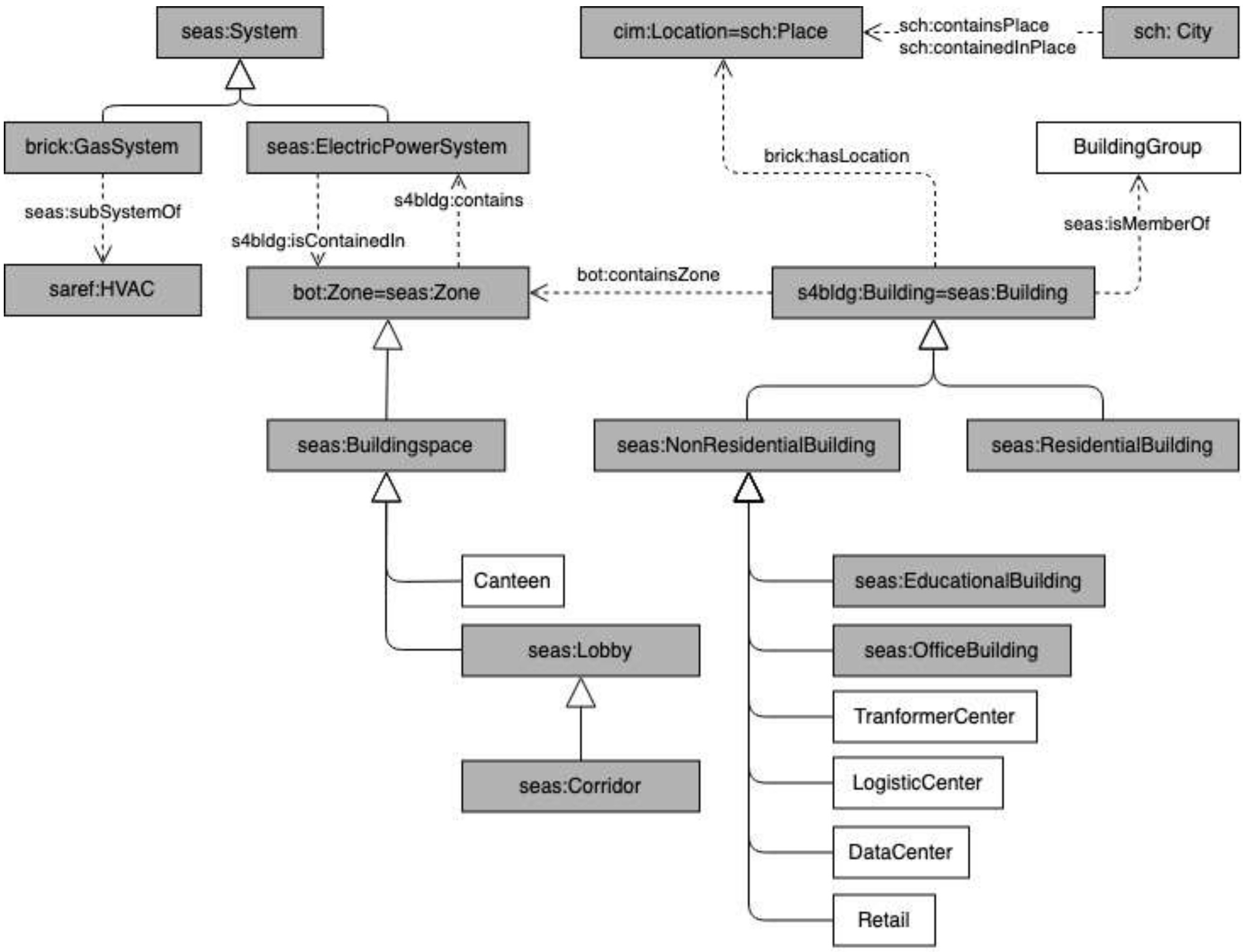


Figure 2: Ontology related to the concept "Building"

2.2.3 Ontology Design methodologies

Developing ontologies is a fastidious and time-consuming task for which there exists only some very general principles to guide the ontology designers [17;22] who still must face with many modelling choices. Indeed, there is not a single, consensual ontology-design methodology. Several existing methodologies have been proposed such as:

a) **Cyc methodology [7**]: presents to process a large amount of common-sense knowledge which is being built upon a core of over a million hand-entered assertions designed to capture a large portion of what people normally consider consensus knowledge about the world. Three phases are defined: (i) the first phase proposes manually coding the explicit and implicit knowledge appearing in the knowledge sources, (ii) the second phase proposes knowledge codification, and (iii) the third phase delegates most of the work to the tools.

b) **Hitzler *et al.* [18]**: presents a tutorial that is based on use case studies and design pattern application and follows different steps from existing ontology design methodologies.

c) **KACTUS methodology [21]**: this methodology is proposed as a part of the Esprit KACTUS project. The main objective of this project is to investigate the feasibility of knowledge reuse in complex technical systems and the role of ontologies to support it.

d) **Methodology of Uschold and King [24]**: this methodology aims to develop and evaluate ontologies. It includes different stages such as identify the purpose, build the ontology, ontology coding, integration of existing ontologies, evaluation, etc.

e) **METHONTOLOGY [10]**: this methodology is developed within the Laboratory of Artificial Intelligence at the Polytechnic University of Madrid. It aims to construct ontologies at the knowledge level. This methodology considers relationships between the life cycle of different ontologies. It is also included detailed recommendations for re-engineering ontologies.

f) **Methodology of Bravo *et al.* [16]**: this methodology promotes the reutilization of ontologies by implementing ontology modules from the beginning of the ontology design, ontologies are seen as reusable modules not as general design patterns. It is based on 4 main steps: (i) requirements specification, (ii) formal design, (iii) construction, and (iv) evaluation.

g) **SENSUS methodology [23]**: it based on the use of a huge ontology for building specific ontologies and knowledge bases to be used in applications. SENSUS is an ontology to be used in natural language processing.

h) **TOVE [12]**: this methodology is defined to build an ontology for the TOronto Virtual Enterprise (TOVE) modelling project. The TOVE ontology was constructed with the objective of representing a common-sense enterprise model.

In the state-of-the-art, there is no consensus in which methodology is better than other. In fact, different input parameters can determine the choice of the methodology adopted to design ontologies. Each group uses its own methodology, and it is odd to find someone who uses a methodology elaborated by a different group.

In the case of T2.3, we inspired from the methodology described in [16;18] and adapt some steps to fit our need for describing of 19 LLUCs issued from the WP1-T1.1.

2.3 Methodology for semantic data model design and construction

In this section, we propose an overview of the methodology which incorporated the most relevant principals that could be applied to construct and to reuse ontological models. The main stages of this methodology, after examining the existing semantic data models (ontologies) of the energy sector, are (see Fig. 3):

1) **Ontology Requirements Specification:** it aims to analyse each LLUC to: (i) assess and define the scope of the ontology according to the application domain, (ii) extract the relevant terms that need to model the LLUC (concepts and relationships), and (iii) specify a list of natural language questions that the ontology should answer to.

2) **Ontology Analysis:** the goal of these steps is to reuse well-known ontologies or to design new ontological modules by respecting the practices for ontology design patterns.

3) **Overview of ontological modules:** it targets a consolidation of all conceptual modules together and providing an example for each pilot.

4) **Interaction with stakeholders and Ontology formalization:** the purpose of this step is to: (i) discuss with the stakeholders to check if their needs are designed, and (ii) formalize all ontological modules by using an ontology editor and a standard language and integrate all modules into an ontology system

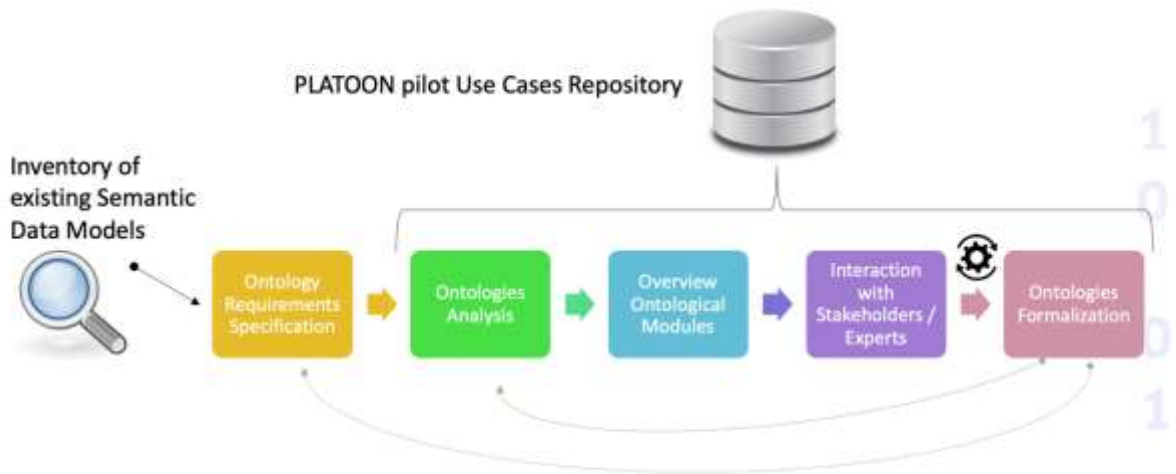


Figure 3: Overview of Semantic Data Models methodology

2.4 Inventory of existing semantic data models

A lot of information on the Web remains difficult to exploit. The Semantic Web has emerged as an interesting solution for structuring and sharing information on the Web, which is stored and represented in semantic resources.

This section describes some existing semantic data models and specially energy domain ontologies.

a) Smart Appliances REFerence ontology (SAREF)

SAREF is a modular ontology for the internet of things domain. It provides a family of ontologies for the semantic interoperability. The SAREF family is composed of the main SAREF ontology and some extensions:

- SAREF main module contains core concepts on IOT Domain

- SAREF extension for smart cities module, SAREF4CITYontology [25] provides a semantic representation of topology, administrative area, event, city object key performance indicator and public service.
- SAREF extension for Building module, SAREF4BLDG ontology [26] describes building devices their location in buildings.
- SAREF extension for energy domain SAREF4ENER [27]
- SAREF extension for agriculture and food domain SAREF4AGRI [28]
- SAREF extension for environment SAREF4ENVI [29]

b) Smart Energy Aware Systems (SEAS):

SEAS ontology [30] describes energy systems and their interactions. It is designed to cover de use cases of the SEAS (Smart Energy Aware Systems) Project. SEAS ontology is versioned modular to favourite the separation and recombination of different parts of the ontology. Seas ontology is composed of different modules according to specific needs:

- SEAS mains modules:
  - seas:FeatureOfInterestOntology for description of features of interest and their properties.
  - seas:EvaluationOntology for the evaluation of this properties.
  - seas:SystemOntology for virtual systems and their connections description.
- SEAS Smart Grids and Micro-Grids use cases modules: seas:PlayerOntology, seas:EnergyFormOntology, seas:ElectricPowerSystemOntology, seas:PhotovoltaicOntology...
- SEAS Smart Homes uses cases modules: seas:ZoneOntology, seas:ZoneLightingOntology seas:BuildingOntology, seas:ThermodynamicSystemOntology
- SEAS Electric cars use cases modules: seas:ElectricVehiculeOntology, seas:BatteryOntology.
- SEAS Electricity market use cases modules: seas:PricingOntology, seas:OfferingOntology, seas:Tradingontology
- SEAS Weather forecasts use cases modules: seas:ForcastingOntology

c) Common Information Model (CIM)

In 2009, ENTSO-E (European Network of Transmission System Operators) was established by the EU's Third Package for the Internal energy market which aims at further liberalizing the gas and electricity markets in the EU. As part of its development, cooperation and coordination, ENTSO-E has accepted International Electrotechnical Commission's (IEC) CIM (Common Information Model) standard as the preferred model for describing power grids. To ensure that the International Electrotechnical Commission's (IEC) CIM standards are developed in line with TSO requirements, ENTSO-E established liaisons with IEC TC 57/WG13 [31] (the working group dealing with CIM for transmission) and IEC TC57/WG16 [32] (the working group responsible for CIM for energy markets). The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic, and related technologies. Nowadays, Common Information Model (CIM) is a widely accepted electricity information model being part of the IEC 61970 standards. Its main objective is to develop a platform independent data model for enabling better grid interoperability. This model includes the exchange between market participants and market operators as well as communication between market operators. Based on the commonly accepted conceptual models (NIST conceptual model, EU M/490 Smart Grids conceptual model), IEC TS 62913 has organized the Smart Grids domains in five clusters (CIM allows extensions):

- Grid-related domains transmission, distribution, micro-grids
- Market-related domain
- Resource-connected-to-the-Grid domains (the Bulk Generation, Distributed Energy Resources, Storage, Smart Industry, Buildings and Home)
- Electric Transportation
- Support functions domains (Smart substation automation, Smart metering, Asset Management)

As CIM is a continuously evolving model, in PLATOON framework we consider CIM as a potential meta-model, e.g., for the Pilot 2a interoperability layer. CIM defines the main concepts needed to describe for power grids, while CIM instances are models of concrete power grids (e.g., actors in Serbia, infrastructure, power systems, etc.).

d) Other energy domain ontologies

Several other domain ontologies of energy sector have been proposed such as SSN, OEMA, BOT, ThinkHome, Semanco, OntoWind, etc. These ontologies have some correspondences with the standard ones (see Tab. 1).

As we notice in Table 2, there is a strong overlapping between the existing semantic data models. This overlapping can be explained by the fact that semantic data models shared common concepts or relations and they covered the same domains.

Domain	Format	Link	Semantic Data Models	Renewable Generation	Smart grids		End Use of Energy			Generation, Distribution and End Use of Energy
				Predictive Maintenance of Wind Farms	Electricity Balance and Predictive Maintenance	Electricity grid stability, connectivity and Life Extension	Office building: Operation performance thanks to physical models and IA algorithms	Advanced Energy Management System and Spatial (multi-scale) Predictive Models	Advanced Energy Management System and Energy Efficiency and Predictive Maintenance in the Smart Tertiary	Energy Management in microgrids
Home appliances	TTL	https://forge.etsi.org/rep/SA	Saref*	✓	✓	✓	✓	✓	✓	✓
Energy	TTL	https://ontology.tno.nl/sa	Saref4Ener*	✓	✓	✓	✓	✓	✓	✓
Environment	Ntriple, RDF/XML, TTL	https://w3id.org/def/saref4	Saref4Envi*	✓						
Building	Ntriple, RDF/XML, TTL	https://w3id.org/def/saref4	saref4Bldg*				✓	✓	✓	
Smart city	Ntriple, RDF/XML, TTL, JSON-LD	https://w3id.org/def/saref4	Saref4City*			✓		✓		
Industry	Ntriple, RDF/XML, TTL, JSON-LD	https://w3id.org/def/saref4	Saref4INMA*							
Agriculture	Ntriple, RDF/XML, TTL, JSON-LD	https://w3id.org/def/saref4	Saref4Agri*							
Building, device,	JSON-LD	https://fiware-datamodels.r	FIREWARE*	✓			✓	✓	✓	
Building, device	TTL, OWL	http://ifcowl.openbimstan	IFC				✓	✓	✓	
IOT	JSON-LD	https://www.gsma.com/iot	GSMA*	✓	✓	✓	✓	✓		
Electricity	UML, RDF	https://ontology.tno.nl/IE	CIM*		✓	✓				✓
Sensor device iot	TTL	http://www.w3.org/ns/ss	SSN	✓	✓	✓	✓	✓	✓	✓
SSN and spatial data	TTL	http://www.w3.org/ns/sosa	SOSA	✓	✓	✓	✓	✓	✓	✓
Energy system	TTL	https://w3id.org/seas	SEAS	✓	✓	✓	✓	✓	✓	✓
Energy heating	TTL, TAG	https://brickschema.org/c	Brick schema				✓	✓	✓	
Building topology	TTL	http://www.w3id.org/bot/b	BOT				✓	✓	✓	
Energy efficiency in Future Smart Homes	OWL	https://www.auto.tuwien	ThinkHome				✓	✓	✓	
Prosumer-Oriented Smart Grid	OWL, TTL	http://data-satin.telecom-st	ProSGV3							✓
OEMA (Ontology for Energy Management Applications): Infrastructure, Energy, Equipment, Geographical data, Smart Grid Stakeholders, Units of Measure	RDF/XML, Ntriples, TTL	https://innoweb.mondrag	OEMA (reutilise ThinkHome, Saref, Energyuse, ProSGV3)	✓	✓	✓	✓	✓	✓	✓
Energy flexibility for a specific device	TTL	https://sites.google.com/site	Mirabel	✓	✓	✓	✓	✓	✓	✓
Smart city	RDF	http://wlode.disit.org:8080	KM4city					✓		
Energy domain: Building energy consumption	OWL	http://semanco-tools.eu/s	SEMANCO				✓	✓	✓	
LCC (Leeds City Council): Energy Consumption	TTL	http://smartcity.linkeddata.	LCC	✓	✓	✓	✓	✓	✓	✓
Smart Building	OWL	http://lpis.csd.auth.gr/onto	BOnSAI				✓	✓	✓	
Intelligent Domotic Environments	JSON LD, RDF/XML, Ntriples, TTL	http://iot-ontologies.github	DogOnt	✓	✓	✓	✓	✓	✓	✓
Wind farms/turbines	OWL	https://raw.githubuserconte	OntoWind	✓						

Table 1: Overview of existing semantic data models

Domain	Semantic Data Models	#Concepts	#ObjectProperties	#DataProperties	Update Date
Home appliances	**Saref***	94	30	6	mai-18
Energy	**Saref4Ener***	66	17	40	déc-16
Environment	**Saref4Envi***	21	19	8	juin-19
Building	**saref4Bldg***	67	176	80	avr-20
Smart city	**Saref4City***	16	15	3	juin-19
Industry	**Saref4INMA***	24	21	11	avr-19
Agriculture	**Saref4Agri***	28	13	5	mai-19
Building, device,	**FIREWARE***	0	0	0	
IOT	**GSMA***	0	0	0	
Electricity	**CIM***	38	49	64	juil-15
Sensor device Iot	**SSN**	6	15	0	avr-17
SSN and spatial data	**SOSA**	13	21	2	avr-17
Energy system	**SEAS**	678	1947	56	January-2017
Energy system	**Brick schema**	2039	30	`0	2020
Building topology	**BOT**	9	14	2	2020
Energy efficiency in Future Smart Homes	**ThinkHome**	1206	513	378	2014
Prosumer-Oriented Smart Grid	**ProSGV3**	614	58	52	May-2014
OEMA (Ontology for Energy Management Applications); Infrastructure, Energy, Equipment, Geographical data, Smart Grid Stakeholders, Units of Measure	**OEMA (reutilise Thinkhome, Saref, Energyuse, ProSGV3)**	108	69	39	March-2018
Energy flexibility for a specific device	**Mirabel**	26	16	0	April-2015
Smart city	**KM4city**	667	104	262	July-2017
Energy domain: Building energy consumption	**SEMANCO**	981	420	358	July-2014
LCC (Leeds City Council): Energy Consumption	**LCC**	62	12	9	sept-14
Smart Building	**BOnSAI**	99	76	41	April-2019
Intelligent Domotic Environments	**DogOnt**	1065	37	72	Aug-2017
Smart Building	**RealEstateCore**	163	69	64	April-2020
Building	**FIEMSER**	48	5	34	April-2015
Building	**SBIM**	84	14	1	May-2017
Wind Turbine Blade Materials	**WTBA (Wind Turbine Blade Material Analysis)**	1600	295	13	July-2018
Wind farms/turbines	**OntoWind**	94	13	28	March-2018

Table 2: Additional information about available semantic data models

2.5 Proposed methodology definition

To design the semantic data model for the PLATOON project, we define a methodology that allow us to reach the requirement of each LLUCs defined its different pilots. The Methodology is inspired from several works in state-of-the-art of ontology design domain [16;18]. The methodology is divided in four steps detailed below. Each step is applied to each LLUC to cover all the specified needs (see Fig. 4).

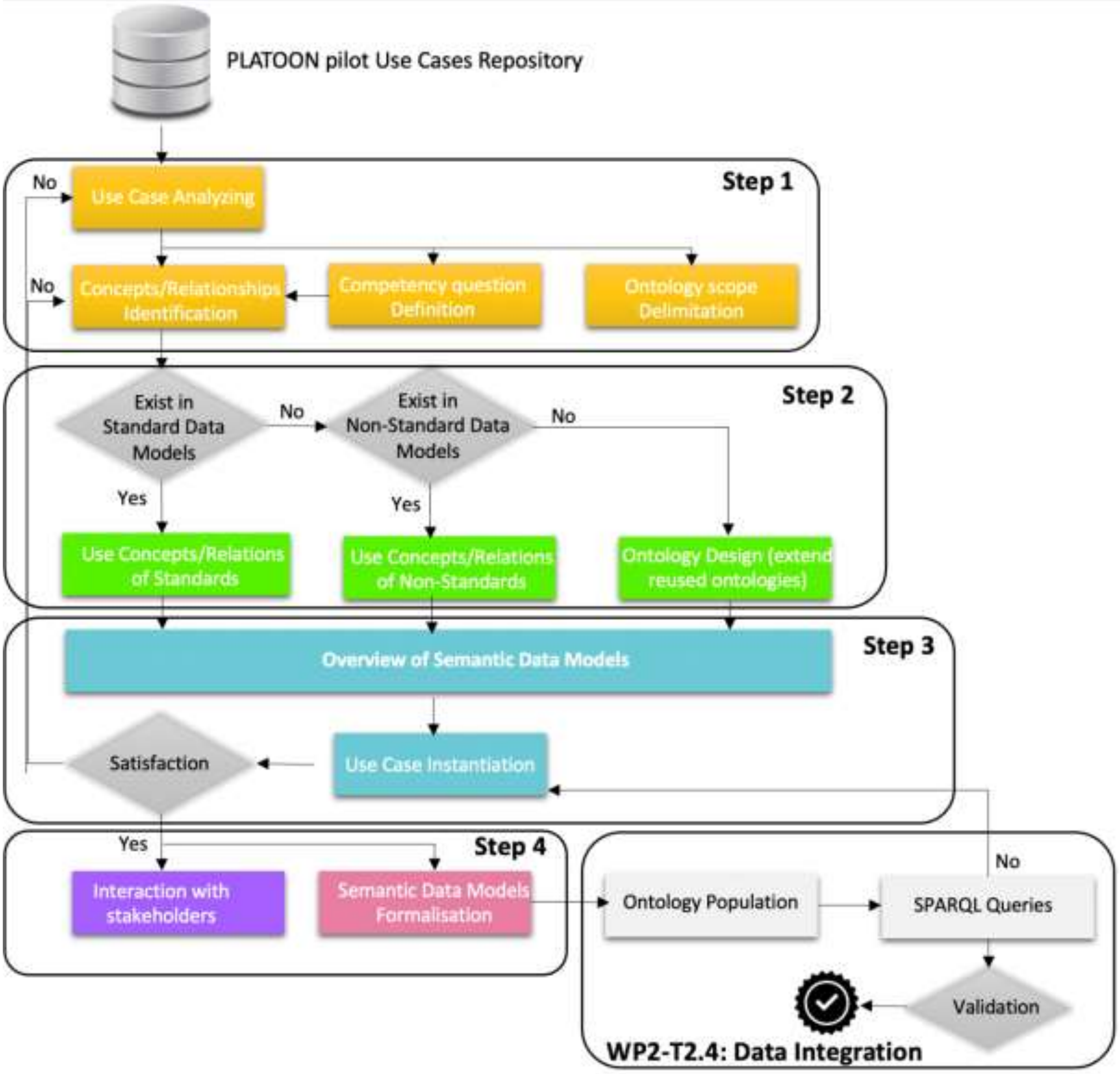


Figure 4: **Detailed Steps of the proposed methodology**

2.5.1 Step 1: Ontology Requirements Specification

Before starting to design ontology, it is important to start to analyse the requirement and the specification expected from the ontology or semantic data model. Then in this step1, is composed of different important tasks to meet the ontology requirement. These tasks are detailed below (see Fig. 4):

a) **Use case analysing:** First, a deep analyse of the requirement (Use Case) is mandatory. In the PLATOON project, each business use case was described in two documents:
   a. the PM^2 template where a High-level description, was given for a specific pilot
   b. the IEC-62559 based document that describes in detail each Low-Level Use Case (LLUC) for each pilot. The analysing phase leads to have a big picture and help to assess the motivations, objectives for each pilot.

b) **Ontology scope delimitation:** The delimitation of the scope of ontology is very important to design a modular ontology. Indeed, often use cases are cross domains and because we cannot define everything in the world, an ontology should have a limited scope which can facilitate its definition and sharing. Then, in this task the scope of the ontology is defined according to the application domain. An example of scope limitation could be a Building Energy efficiency and Heating and Cooling Systems.

c) **Competency questions definition:** are questions in natural language that domain experts want the ontology help to answer. The knowledge engineer together with the group of domain experts should produce a list of competency questions. Such competency questions are generated by asking the group of domain experts to enunciate direct questions that they expect the ontology system will be able to answer once it is implemented and in production. The list of competency questions will also be useful to evaluate the final ontology. An example of competency questions can be:

   - What are the sensors used for energy consumption in the building x, and what are the type of these sensors?
   - What is the energy consumption of each system contained in a building x?
   - What is the occupancy of a zone and what is the forecast of the occupancy?

d) **Term elicitation:** to do this task, the knowledge experts analyse the IEC-62559 documents that describe the low-level use cases and identify and extract all terms or notions that are relevant for a particular domain. Furthermore, they also analyse the list of competency questions to extract key terms relevant to be included in the semantic model. Example of key terms are sensors, building, energy consumption, what (type of sensors), where (sensors are located).

In summary the output of this step 1 is a template document with the scope of the ontology, a list of competency questions and a tab with a list of extracted relevant terms. These terms were analysed with the domain business experts to validate if the terms extracted are effectively relevant to be designed in the semantic model. Furthermore, the domain business experts can add terms in the tab of extracted term if needed to cover notions missed in the IEC-62559 document.

2.5.2 Step 2: Ontology Analysis

The step 2 aims to create a socle to build the semantic data model. Following the principle of ontology domain, the ontology reusing is recommended before starting any new designing. According to this principle, we define several tasks detailed below:

a) **Identification of concepts and relationships,** from the list of extracted terms during the elicitation task of the step1, we associate for each term or a key notion a concept name or a relation name that will be used in semantic datal model. For example, we associate a concept Sensor to the term *sensor*, and the concept `HumiditySensor` to the term *humidity sensor*.

b) **Reusing Ontology** as previously noticed, reusing ontology should be privileged before creating new concepts. For each identified concept in the previous task, we search in existing ontology if this concept is already defined. When the concepts exist in several ontologies, our strategy is to propose a mapping between these equivalent concepts to ensure the interoperability. However, before effectively reuse a concept we analyse the hierarchy of the concept to be sure that the subsumption relation (`is-a`) is not confused with the part-whole relationship (`part-of`), composition or location relationships.

c) **Extending Ontology** in the second case, the ontological module in the domain exists and covers a part of the use case but the concept doesn't exist in these modules. In this case we extend the existing ontology with the new concept.

d) **Ontology Construction** in the third case, the existing ontologies don’t cover the use case domain. In this case we create new ontological modules that cover the use case. However, we extend or add equivalent concepts when it is relevant to increase the interoperability.

In summary the output of this step 2 is: (i) a list of identified existing ontological modules, (ii) a list of reused concepts and relations and (iii) a list of concepts that need to be designed in a new ontological module to cover the scope of the LLUC.

2.5.3 Step 3: Overview of ontological modules

The third step of the proposed methodology aims at integrating all modules together into a harmonized semantic data model and producing an example for each pilot. This phase consists of the following procedures: diagrams integration, ontology evaluation (scoping of Use Case, consistency, competency) and pilot instantiation with an illustrative example.
This step takes as input: (i) a list of identified ontological modules, (ii) a list of concepts and relations coming from the list of ontological modules, and (iii) a new designed ontological module.

a) **Diagrams integration** the knowledge engineer puts different modules together in a schema diagram to check all classes, properties and possibly to improve them. This task is important to join the different schema diagram of modules with semantic relations (e.g., subsumption, equivalence, etc.). A global overview schema of all used modules should be provided (see Fig. 4).

b) **Ontology evaluation (scoping of Use Case, consistency, competency)** The knowledge engineer will evaluate the ontological modules to ensure that its definitions correctly implement the use case requirements and competency questions. The goal of ontology evaluation is to prove compliance of the world model with the world modeled formally.

Two important aspects are used for evaluation:

- Competency of the ontology: verify that a representational model is complete with respect to a given set of competency questions.
- Quality requirements: can be measured as the degree of compliance it has with respect to established design criteria (Clarity, Coherence, Modularity).

c) Use Case instantiation with an illustrative example the knowledge engineer instantiates scenarios of the use case with an illustrative diagram (see example in Fig. 5).

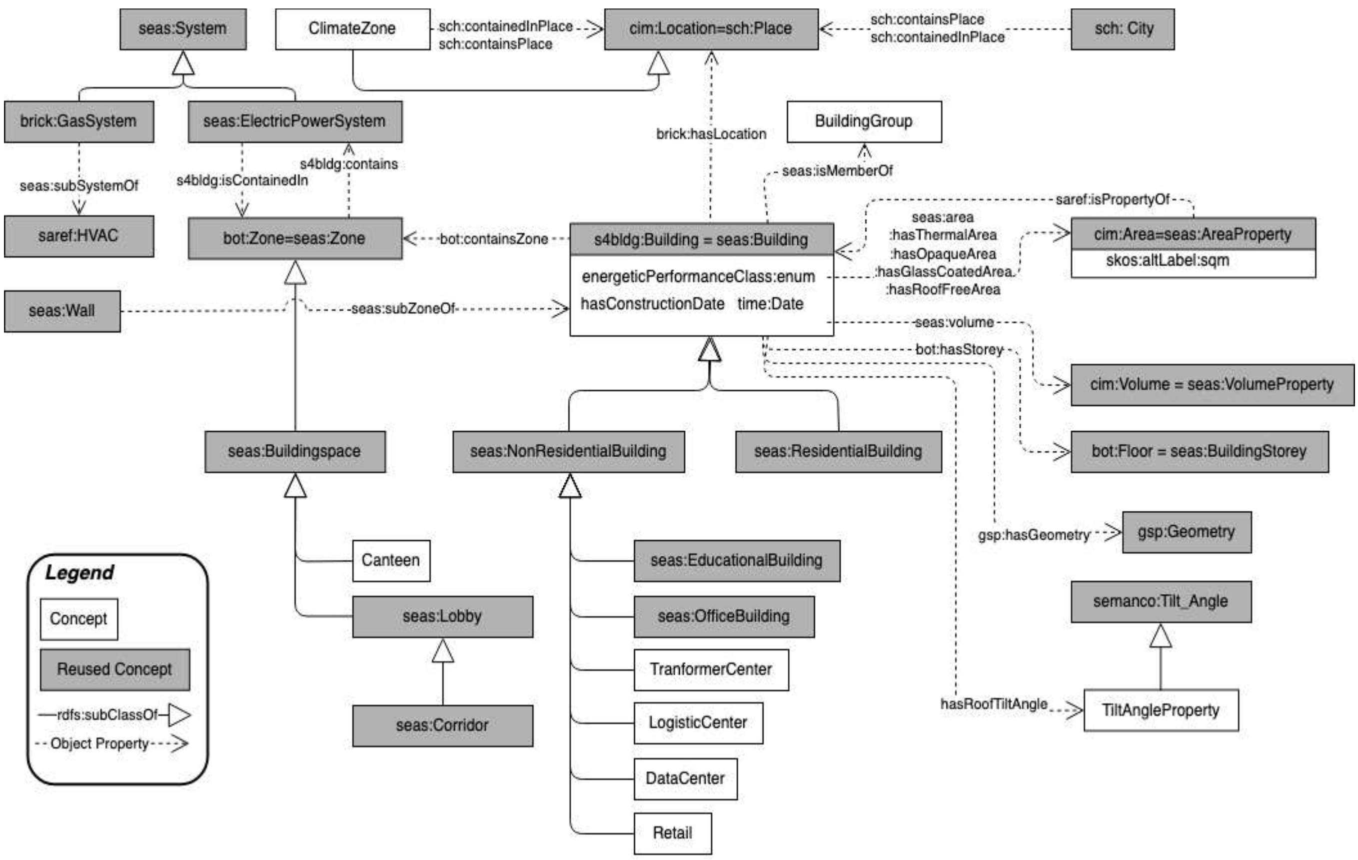


Figure 5: **Example of the concept "Building" and its properties**

2.5.4 Step 4: Interaction with stakeholders and ontology formalization

The objective of the fourth step of the methodology is to interact with stakeholders and to code all ontology modules by using an ontology editor and a standard language and integrate all modules into an ontology system. This step consists of the following tasks: (i) discussion with stakeholders and (ii) ontology formalization.

a) Discussion with stakeholders If the stakeholder is unsatisfied, the knowledge engineer lists a set of issues and returns to Step 1.

b) Ontology formalization Process If the stakeholder is satisfied, the knowledge engineer proceeds to formalization process:

- provide an owl file for each new ontological module if created in the Step 2
- provide an RDF file for each use case, describing how the overview model is instantiated. This file should be taken as an example to feed the knowledge base (in WP2-T2.4).

2.6 Pilots' description

2.6.1 Pilot #1a – Predictive Maintenance of Wind Farms

Pilot #1 focusses on predictive maintenance in wind energy. Today, the O&M cost of wind energy is substantial. Having better failure prediction models can help in substantially reducing these costs by allowing to perform maintenance during clustered and well-planned repair campaigns. The industry is therefore looking for accurate monitoring and diagnostics tools that can continuously process data to many wind turbines and pinpoint to root-causes of failures.
The main goal of this pilot use case is to achieve early detection of faults in the electrical components of the powertrain: generator and converter. Different modelling approaches are used to predict behaviour of the electrical components. On the one hand a physics based digital twin model is used; on the other hand, data-driven normal behaviour models are considered. These models will serve anomaly detection schemes. Through reasoning based on semantics the fault will be diagnosed based on these anomalies and potential root cause pinpointed out. Particularly for the latter a good semantic representation is required, since analytics-driven reasoning should be done on the data.
In addition to batch processing at cloud level, edge computing will be used to perform processing of the high frequency current data of the generator for signature extraction. In this pilot focus is on offshore and onshore wind turbines equipped with a doubly fed induction generator.

2.6.2 Pilot #2a – Electricity Balance and Predictive Maintenance

PLATOON Pilot #2a will take place in Serbia, while the overall goal of the pilot is to integrate and deploy different PLATOON analytical services with the Institute Mihajlo Pupin (IMP) proprietary VIEW4 Supervisory control and data acquisition (system). The VIEW4 SCADA is deployed at many parts in the energy value chain in Serbia, starting from control on the production side (in the large hydro and thermal power systems, as well as RES), via transmission management to distribution and electricity dispatching, see Figure 6. To better understand the needs, several different scenarios have been defined in WP1. Scenarios elaborated to different Smart Grid domains and were prioritized in the requirements specification phase (Task 1.1 and Task 1.4):

- **Market-related domain (low priority):**
  The first two scenarios have low priority because services are currently under development in H2020 TRINITY project. These two scenarios have not been modelled in detail in the PLATOON semantic model.
  - LLUC P-2a-01 Balancing on regional level:
    Electricity balancing is a set of actions and processes performed by a TSO to ensure that total electricity withdrawals (including losses) equal total injections in a control area at any given moment. At each point in time, total production, combined with interchange, i.e., export or import of energy from/to control area, must be equal to total consumption to stabilize system frequency and to maintain exchange at scheduled levels; it is therefore also called load-frequency control.

  - LLUC P-2a-02 Balancing on country level, reserve/energy exchange process:
    Electricity production from solar and wind plants is subject to considerable forecast errors that drive demand for balancing i.e., for operational reserves. To enable a regional exchange of operational reserves, it would be necessary to ensure that the corresponding volumes could be made physically available when required. The unplanned (intended and unintended) deviations from day-ahead schedules are required to be balanced in intra-day and balancing markets. Hence, the TSO and balance providers need different types of forecasting services.
- **Grid-related domains transmission, distribution, micro-grids (high priority):**
  - LLUC P-2a- 03 Load/Demand forecasting:
    Electricity demand forecasting is a central and integral process for planning periodical operations and facility expansion in the electricity sector. The aims of this use case are load forecasting and prediction of the load pattern. Load demand forecasting involves accurate prediction of both magnitudes and geographical locations of electric load over the different periods of the planning horizon. Load forecasting can be divided into three categories: short-term forecasts (from one hour to one week), medium-term forecasts (from a few weeks to a few months and even up to a few years) and long-term forecasts which is a crucial part in the electric power system planning, tariff regulation and energy trading. A long-term forecast is generally known as an annual peak load.
    There are several factors that will be taken into consideration for load forecasting, which can be classified as time factor, economic factor, weather condition and customer factor.
- **Resource-connected-to-the-Grid domains (Distributed Energy Resources):**
  - LLUC P-2a-04 RES forecasting (high priority):
    The goal of this scenario is to develop and test a PLATOON service for more accurate prediction of renewable energy generation. Different energy supplier strategies for RES should be considered that will support and improve strategic optimization of the resources on the side of the balancing reserve provider. The wind power forecasting yields an estimate of the variable power injected in the distribution grid. This allows prediction of when the transformer connecting the distribution grid to the transmission grid will be overloaded, i.e., when local wind turbine generator production will be very high. The various forecasting approaches can be classified according to the type of input (weather prediction, wind turbine generators data, historical production data). Statistically based approaches allow very short-term predictions (2 hours). One of the key challenges for day-ahead forecasting of wind energy remains unscheduled outages that can have large effects on the forecasts for small systems, while the effect is small on the overall grid.
  - LLUC P-2a-05 RES effects on the Power System (medium priority):
    Constantly increasing number of renewable energy resources such as photovoltaic and wind power plant has a significant impact on the stability and power quality of electricity transmission. The goal of this scenario is to test the PLATOON analytical services for analysis of unexpected variations (voltage profile of the power system) before and after RES integration to the power system. The service will calculate the effects of RES on the existing power system and presented graphically. To this aim Phasor Measurement Units (PMU) for collecting and edge computing device for analysis of the real-time power flows will be used. Additionally, phasor-based

control could be exploited to perform power flow optimization and improve the power quality in the real-time domain.

- **Support functions domains (Asset Management) - (high priority)**
  - LLUC P-2a-07 Predictive maintenance in RES power plants

  The continuous monitoring of asset performance generates input that can be used for predictive analytics and to provide early warnings of component/object failures (e.g., RES plant/component). Identifying problems before they occur helps to reduce unscheduled downtime, improve plant maintenance, and optimize asset performance. In this Use case, a PLATOON service will be developed that identify rare events that could occur in power plant infrastructure due to infrastructure health problems, progressive degradation, or failure. By monitoring the output from the RES power plant using the PMU unit and doing advanced power quality (PQ) analytics close to the source, events can be detected and labeled. By gradually creating a database of events by learning from historical data, one could use this classification to find abnormal functioning of the system before it leads to failure.

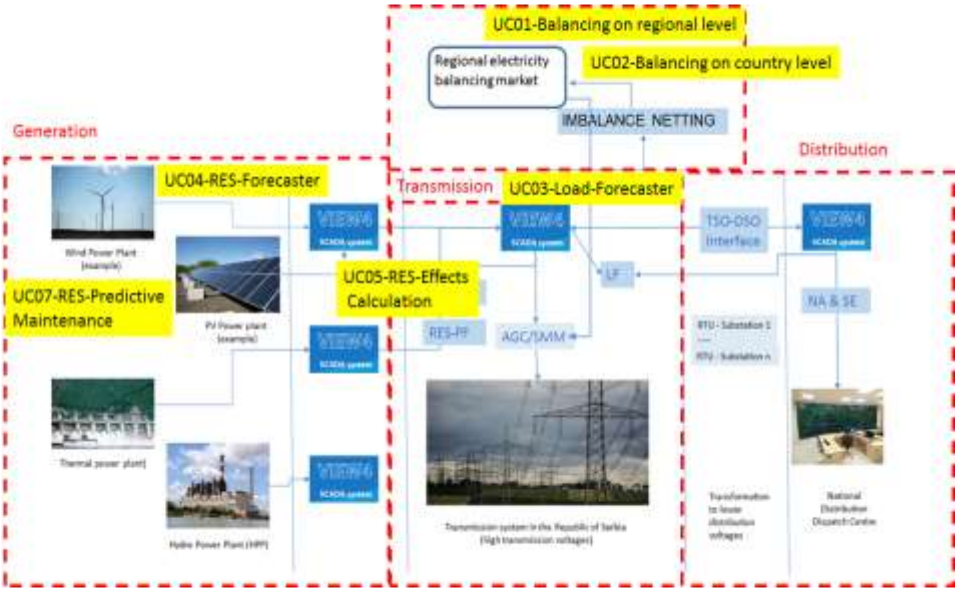


Figure 6: **Serbian energy supply chain and relation to PLATOON services**

2.6.3 Pilot #2b – Electricity Grid Stability, Connectivity, And Life Cycle

Parc Bit is a technological park located in Mallorca where its electrical grid will be under study in pilot 2b. Two different use cases have been defined described in the following list:

- **LLUC P-2b- 01 - Predictive Maintenance for MV/LV Transformers**

  This UC will focus on LV/MV transformer predictive maintenance, estimating transformer components health and its maintenance costs, planning maintenance actions, monitoring transformer alarms and studying different grid scenarios in case of replacing old transformers or adding complementary transformers. To predict transformer maintenance is crucial to estimate transformer components health and its maintenance costs, planning maintenance actions, monitoring transformer alarms and studying different grid scenarios in case of replacing old transformers or adding complementary transformers This tool will use available data in LV/MV transformers, which usually have a small budget for monitoring and maintenance. Maintenance actions are based in the Remaining Useful Life (RUL), in this project, different failures modes of the transformer critical components will be estimated and reflected to the health index of the transformer. Once it is obtained the health index and considering maintenance and failure costs, the transformer maintenance plan will be defined. Finally, a prescriptive analytics tool will be developed. This tool will allow to evaluate the effect of different operational actions in the grid O&M cost sheet.

- **LLUC P-2b- 02- Non-technical loss detection in Smart Grids**
  The main objective of this use case is to develop a tool for the quantification of losses in the distribution grid of a DSO and the detection of non-technical losses (NTL), using the available smart meter data from Sampol´s smart grid in ParcBit, Majorca (Spain). NTL in the electric system is one of the biggest fraud factors, generally due to smart meter and/or connection to the grid (bypass of certain consumer loads) manipulation. NTL can be attributed to i) administrative losses due to accounting errors and record keeping, ii) customer theft, iii) customer non-payment and iv) theft by non-customers. For instance, in the last years, losses in the Spanish electric system have been about the 8% and almost half of that amount could be due to NTL. NTL detection is a complex task as DSOs do not have the appropriate tools and resources to avoid fraud techniques, and the only effective way to fight against NTL is to detect it once it has been committed, and to establish legal actions in that case. Many theorical studies in detecting NTL using different techniques have been conducted in the last years. One of these techniques is data science, very popular in the academia. The main output of this UC is to develop a solution for NTL detection using data analytics, which just requires measurement data available from the Automatic Metering Infrastructure-AMI, and optionally information on the grid topology.

2.6.4 Pilot #3a – Office Building: Operation Performance Thanks To Physical Models And IA Algorithms

Pilot #3a concerns an office building with a focus on developing Use Cases to optimize the HVAC system performance or provide new kind of services (to help with the grid management). Two main use cases have been identified to be implemented on the pilot (see Fig. 7):

- **LLUC P-3a-01- Optimizing HVAC control regarding occupancy**
  The use case aims at providing a smart module for an office building that optimize HVAC operation in function of real occupancy. Occupancy data are available via dedicated sensors, and the comfort and HVAC controls are available via the Building Management System (BMS) of the building. Using historical data, some learning algorithms are implemented to predict occupancy and anticipate heating and cooling period in the building and its different zones. A first optimization loop can be implemented to control the overall building occupancy planning and HVAC operation. A second optimization loop is used to adapt HVAC controls in the different zones of the building. The building manager can supervise and update some parameters in the system and access some regular assessment of the system controls. It also collects data from occupancy sensor to map the occupancy in the different zones of the building.
- **LLUC P-3a- 02 - Providing Demand Response Service through HVAC control**
  The use case aims at providing a smart module to supervise the implementation of Demand Response services in an office Building using HVAC control and building inertia. Through the supervision of the building parameters and weather forecast, the module developed is providing predictions of the HVAC load and the potential flexibility available in the building, considering that a certain thermal comfort level must be maintained. These predictions are regularly transmitted to an aggregator that is then able to

engage reliable flexibility services with the grid operator. The aggregator can then send or plan orders to stop the HVAC system of the building for a given time. If the orders are validated (within conditions of the contract and minimum comfort parameter are respected), they are implemented in the BMS. Feedback and KPI are shared with the aggregator concerning the load shifting operations.

Figure 7: Data exchange between the different actors in Pilot #3a

2.6.5 Pilot #3b – Advanced Energy Management System and Spatial (Multi-scale) Predictive Models In The Smart City

PLATOON Pilot #3b will take place in Rome (Italy) and the **overall goal** of the pilot is to acquire, aggregate and process energy consumptions and related data of different buildings (often in various form not coherent between them) to make energy domain specific data analysis as consumption forecasting, predictive maintenance, benchmarking and so on.
PLATOON methodology was applied to model in detail the scenarios that will be tested in WP6 framework and are related to analytical services that will be developed in WP4 framework, as follows:
Poste Italiane buildings to be considered in the use cases are all located in Rome Municipality Area. They span from 180 to 28.000 sqm.
Four different destinations for the building spaces are considered: Datacentre, Logistics, Retail and Offices (Directional), for a total of 16 buildings.

- ***LLUC 3b-01 - Building Heating & Cooling consumption Analysis and Forecast***

The correlation with external weather conditions, building characteristics and past performances together with benchmark with similar building, represent an area of optimization for both cooling and heating systems. Sensors, meters, and other hardware produce information that, through processing with forecasting algorithms and machine learning techniques, could be used to predict plants consumption and for the energy efficiency benchmarking. Different objectives are fixed for this LLUC such as: (i) energy efficiency plans (heating, cooling), (ii) daily and hourly energy consumption forecast, (iii) building energy usage benchmark, and (iv) reduction of emissions (CO_2 / TOE correlation).

- ***LLUC 3b-02 - Predictive maintenance of cooling & heating plants***

Predictive maintenance allows equipment users and manufacturers to assess the working condition of machinery, diagnose faults, or estimate when the next equipment failure is likely to occur. If we can diagnose or predict failures, we can plan maintenance in advance, reduce downtime, and increase operational efficiency. Using systems energy consumption data and historical information about fault and maintenance it will be possible to identify anomalies and predict failures in the systems. Several objectives are defined: (i) improve plants efficiency, (ii) technical plants fine tuning, and (iv) increase the availability of heating/cooling plants.

- ***LLUC 3b-03 - Lighting Consumption Estimation & Benchmarking***

Understanding lighting consumptions as accurately as possible it is crucial; on the other hand, they are often aggregated with other energy usage, so the specific consumption is often estimated using algorithms and benchmark tools. Knowing other consumption usage data (such as heating and cooling...), total consumption of the building, lighting installations number and type and other building characteristics (such as category, square meters, generic occupancy profiles...) we want to estimate the specific building lighting consumption, to benchmark, plan optimization actions and detect anomalies and outliers. Estimating the lighting consumption will also be possible to better compare the new performance with the previous lighting technology where new installations are made. Several objectives are defined: (i) estimation and benchmarking between different lighting solutions, (ii) optimization and reduction of lighting consumption, (iii) correlation between the number of building user and the lighting consumption, and (iv) GHG emissions reduction.

- **LLUC 3b-04 - Monitoring and Analysis of energy meters data of ROM large asset**

This LLUC is related to Roma Capital large asset of municipal buildings assuming the energy consumptions data coming from the power and gas meters and the structural data of the buildings to be processed with the general scopes such as: (i) improving energy efficiency, (ii) increasing the responsiveness of the municipal offices, (iii) improving their awareness in terms of energy consumptions and efficiency dynamics, and (iv) reducing losses of money and time. The Public Works and Infrastructures Department of Roma Capital (SIMU Department) includes Plants Division with at least 3 offices managing energy issues: the Energy Manager Office of Roma Capital (EMO), the Utilities Meters Office (UMO) and the Thermal Plants Office (TPO). This Unit manages around 8950 energy meters (6500 electric meters and 2450 gas meters) related to almost 2000 buildings and complex of buildings owned by the municipality. To help the offices in this activity, considering the huge amount of data coming from the meters each month, an integrated monitor and analysis system shall be implemented. In the initial context the management of these data is fragmented and far from being fully integrated in coherence with a set of general objectives, while the energy consumptions datasets, for electricity and gas, are quite heterogenous. The data should be analysed automatically to produce a benchmarking focusing on Energy Performances (EP), to highlight anomalies, to generate reports for different purposes, to produce forecasts in terms of energy consumptions (EC) and other reports and assessments to tackle energy efficiency activities more effectively. The mentioned municipal offices, beneficiaries of the PLATOON 3b-04 pilot project, expressed a series of needs allowing to define a set of 4 Services that PLATOON should offer, based on the energy meters Big Data, to be submitted to a test phase with the users and deployed fully working on different services: spatial reporting (e.g., spatial visualization of the buildings, spatial and attributes queries to produce Flash reports and aggregated results), benchmarking analysis (e.g. reporting the benchmarking of EC & EP in terms of ranking, peaks, anomalies, averages for typologies, etc.), forecasting (e.g. predicting energy usage of one or a set of buildings by analysing multiple factors), and RES Potentialities (e.g. selecting a sub-set of buildings, describing actions or recommendations to optimize/enlarge/improve the installed plants (PV) based on different simulation criteria such as max. self-consumptions, max. Plants dimensions on roofs, min exceeding energy, and economical optimum).

Resulting in terms of PV plants power, PV energy production, Storage capacity if needed to update Energy AUDITS and EE scenarios for the buildings and to assess the existing RES plants efficiency.

2.6.6 Pilot #3c – Energy Efficiency and Predictive Maintenance In The Smart Tertiary Building Hubgrade

Pilot #3c takes place in Nanogune, a tertiary sector smart building dedicated to nanotechnology research, based in San Sebastian (Spain). This building is divided into different areas such as offices and laboratories and has both thermal and electrical meters to differentiate the areas (see Fig. 8).

Figure 8: Nanogune Building

Two main use cases have been identified to be implemented on the pilot:

- **LLUC P-3c 01 Advanced EMS for Tertiary Buildings:** The Advanced EMS will optimize the local renewable energy resources (RES) and HVAC operation as function of building and RES characteristics, building comfort constraints, ambient conditions and energy market price following a multi-objective pattern which targets to reduce the overall energy bill and maximize the usage of RES. For this the study area will be the second floor of offices in the Nano area (see Fig. 9) and we will use both comfort sensors, such as electrical and thermal constants in this area, as well as the consumption of the AHU that heat this area and finally the PV panels that are installed in Nanogune.

Figure 9: Zone of study

- **LLUC P-3c 02 Predictive Maintenance in Smart Tertiary Building Assets:** Development and implementation of predictive maintenance tools for the thermal control assets of smart tertiary buildings (specifically chillers, pumps, and distribution rings). The objective is to improve the maintenance policy increasing the availability and useful life of these assets and reducing the general maintenance costs. For this use case we will use the signals that are currently integrated into the SCADA of this equipment, as well as the history of the maintenance tasks carried out on all the chillers that GIROA manages and the maintenance ranges that it applies.

Figure 10 depicted the ICT architecture showing the data exchange between different pilot partners:

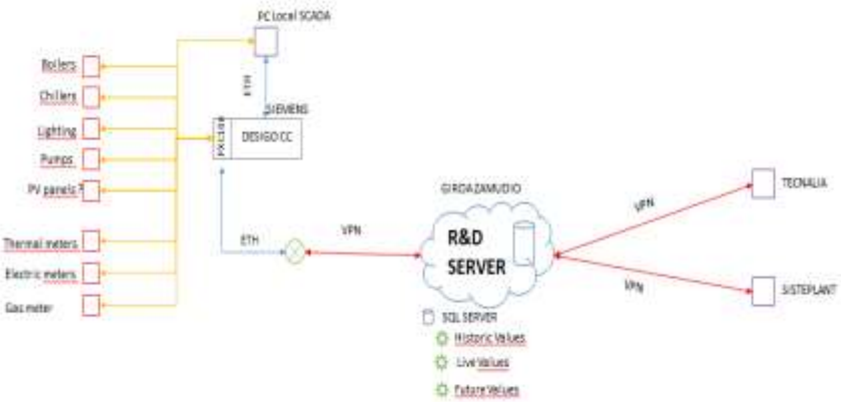


Figure 10: data exchanges between the different actors in the pilot #3c

2.6.7 Pilot #4a – Energy Management Of Micrograms

PLATOON Pilot #4a will take place in Milano, Italy. Figure 11 shows the Politecnico di Milano’s Multi-Good Micro-Grid Laboratory (MG2lab) which is an experimental facility for real-life scale research, simulation, and test purposes, thus, allowing to study new data-driven paradigms for energy management able to deal with increased complexity of the energy systems and to assess the advantages of innovative strategies.

Figure 11: The Multi-Good Micro-Grid Laboratory (MG2lab) at PDM

The main use case to be implemented on this pilot is the following:

LLUC P-4a- 01 - Energy Management of Microgrids

The goal of the functionality described in the current use case is to study data-driven energy management able to deal with increased complexity of the energy systems and to assess the advantages of innovative strategies, by means of EMS with real-time processing and optimization for small-scale/renewable electricity generation, based on power generation and load forecasts (see Fig. 12).

Indeed, EMS plays a crucial role in the management, real-time processing and optimization of assets connected to the grid (small-scale/renewable electricity generation and those used for demand response). The development of new microgrids control and management strategies involves the integration of data analytics toolboxes and optimization platforms on the EMS that manages the microgrid operation.
The EMS development includes on one side the incorporation of predictive algorithms able to forecast renewables production and load profile and on the other the exploitation of an Optimal Power Flow ('OPF') algorithm, able to consider the fluctuation of Renewable Energy Resources (RES) and to optimize the economic unit dispatch, the reliability of the operation of the electricity network (e.g. by predictive maintenance) and energy efficiency (e.g. by reliably predicting and monitoring energy savings).
Optimization Algorithms and Control Predictive functions, for both off-grid and on grid applications, require accurate load and energy production profiles to exploit the potentiality: AI-based models are employed to improve the accuracy of the forecast. Models including on-site measurement and sky-images are currently developed to further improve the PV and load forecast and will be tested in the micro-grid.

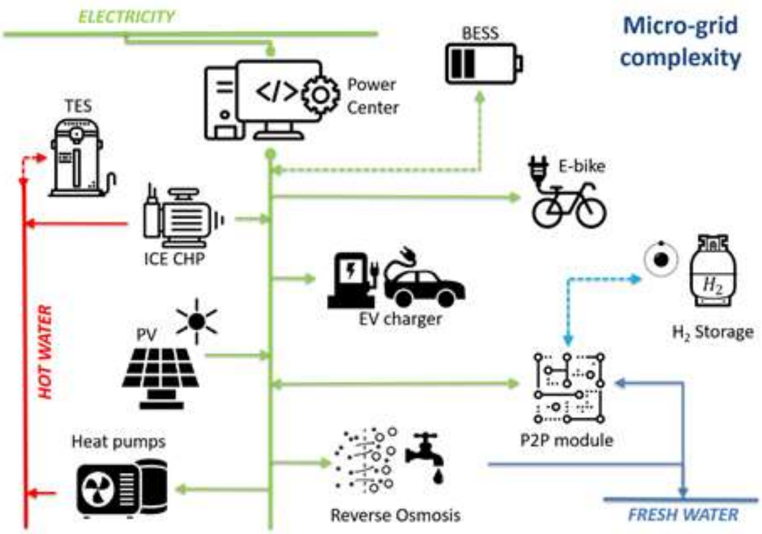


Figure 12: Assets in the considered experimental micro-grid.

2.7 Overview of main Pilots' topics

In PLATOON project, as presented in the previous sections, seven pilots are involved. Each pilot can have its specific needs and use cases covering overlapping topics.
Indeed, these pilots share several common notions in different domains. For these uses cases design to be harmonized, we need to determine the set of overlapping concepts that are similar in meaning and are unique to each of the use cases. This step is important in our process to create harmonized semantic data models that include the information of all the use cases.
From the analysis of these pilots, we have roughly identified five topics (see Fig. 13).

- Topic 1 (orange circle) concerns the building, its zones, and the associated other types of building such as retail, logistic centre, data centre, etc. This topic is also related to the properties of a building for instance, energy load, gas, or electric consumption.
- Topic 2 (green circle) concerns the HVAC equipment, its subsystems such as heating, cooling and ventilation system. It is also related to the Air Handler Unit.
- Topic 3 (blue circle) concerns smart/microgrid, electricity generation, balance, storages, and its properties.

- Topic 4 (grey circle), concerns wind turbine, its components and PV plant.
- Topic 5 (white circle), concerns common notions such as sensors and meters, weather, schedule, failures, etc.

Pilots #3a, #3b and #3c share 2 topics (topic 1 and topic 2) related to Building and HVAC domain. Pilots #2a, #2b and #4a share topic 3 related to the Grid domain. The pilot #1a is in an independent topic (topic 4) related to the energy renewable but could share some notions on power production and electricity generation. All pilots share common concepts (topic 5) which are related to the weather, measurement, forecasting, damages, and failures, etc.

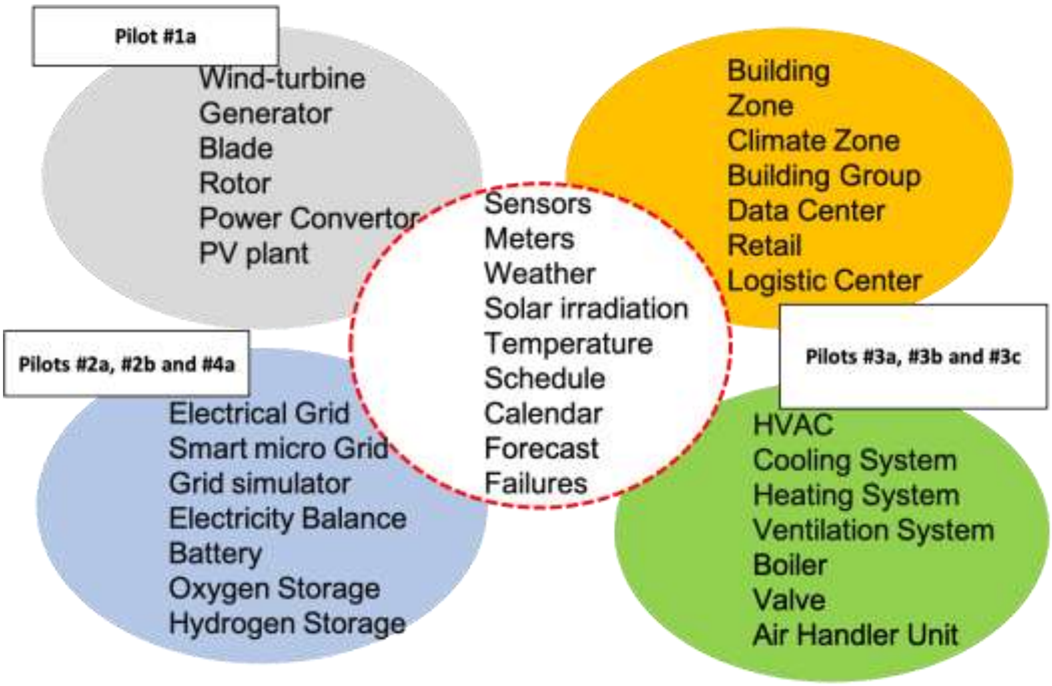


Figure 13: Overview of main topics

In the section 2, we started by presenting some definitions in the domain of semantic data modelling. Then, we presented the methodology proposed to create the data model. Next, we give a short description of the LLUCs. We finished with an overview of the main topics that cover LLUCS and shared by the different pilots, to give us an idea on different domains of pilots and it also allows us to define the right strategy to harmonize the semantic data models. The next section shows the application of the proposed data models methodology.

3 Proposed methodology application

This section aims to describe case studies of the application of the proposed methodology for all the pilots that are presented in PLATOON. We defined a common template that is provided to the different partners to correctly define the scope of the domain ontology and to easily exchange with the stakeholders.

3.1 Application of Step 1 – Ontology Requirements Specification

The aim of step 1 of the proposed methodology is, from use cases repository defined in WP1, to (i) analyse each use case of pilots, (ii) delimit the scope of the ontology, (iii) define the competency questions and, (iv) list the relevant terms.
In the common template, each responsible of pilot briefly described the goal of the use case that will be designed, the main mission of the ontology design and proposed a set of questions that the ontology should be able to answer.

Table 3 points to an example of a use case of pilot #3a named “Optimization of HVAC operation regarding building occupancy”. This use case aims to design semantically a building, HVAC systems used in this building and its properties. As a result of the meetings with the business experts, different natural language questions are defined to validate the semantic data model such as:

- *What are the zones of a building?*
- *What are sensors located in a zone?*
- *In which zone an HVAC valve is located?*

STEP 1: Ontology Requirements Specification	
Tasks	**Description**
Use Case Analysis	The ontology model aims at designing a building where a BMS is installed in order to review and monitor HVAC systems used in this building
Ontology scope	HVAC systems, Building, Building occupancy and consumption, heating/cooling.
Competency questions	Some competency questions: 1. What are the zones of a building? *Ok* 2. What are zones adjacent to a zone x? *Could be interesting, not key* 3. What is the area of a zone x? *Could be interesting, not key* 4. What is the volume of a zone x? (geometry) *Could be interesting, not key* 5. What are kinds of zones (corridor, stairs, open space zone, meeting room)? No (*Not this level of detail, one zone could aggregate many type of these zones)* 6. Where zone x is located in building y? *Could be interesting, not key* 7. What kind of opening (windows / doors) zone / building? *Too precise* 8. What are sensors located in a zone? *Ok* 9. What property a sensor measure? *Ok* 10. What frequency the sensor takes the measurement. *What frequency the data is collected?* 11. In which zone an HVAC valve is located? Ok 12. What are the % of opening of the valve ? At a given time Ok 13. Ask if a zone is occupied/unoccupied? Ok 14. What is the air temperature of a zone? Ok 15. What is the air temperature of the building? Ok

Table 3: Example of use case specification in T2.3

In this step, a list of terms that are relevant for the domain of knowledge, are also extracted. The extraction of terms is done by identifying the list of nouns and verbs. Nouns represent the list of candidate terms that will be used as the preliminary concepts in ontology modules. This list of terms should be non-redundant and is used to validate the concept coverage after designing the ontology.

Figure 14 shows an example of the important extracted terms from the description of LLUC-P-3a-01 such as building, zone, sensors, valve, hvac, etc.

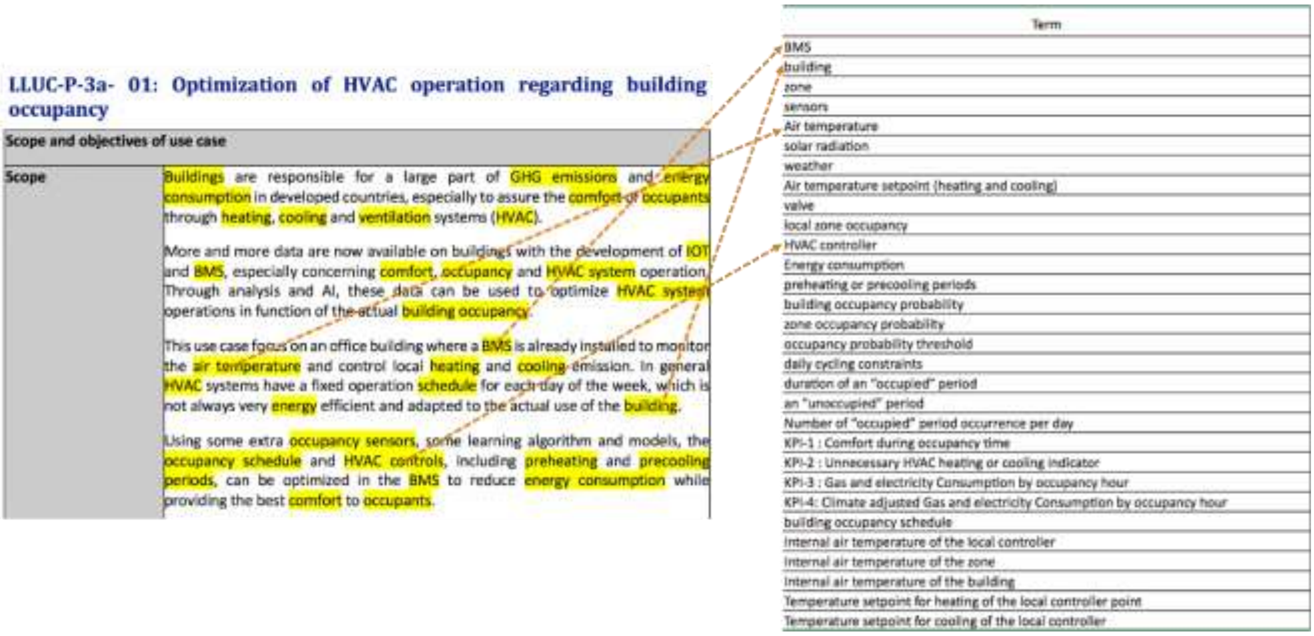


Figure 14: Example of term elicitation process

3.2 Application of Step 2 – Ontology Analysis

The goal of this step is to identify the relevant concepts needed to be modelled in the semantic data model and the reused or extended ontologies. Concretely, for each term or a group of terms extracted in the previous step presents in the table below, we associate a concept name (see Concept Column) or a relation name (see the Relation Column) that will be used in semantic data model. Then, if the concept already exists in an ontology, we put the existing concept in the column "existing ontology". In the case that the concept exists in several ontologies we put all the concepts in the same cell to prepare the mapping between these concepts. Otherwise, if the concept doesn't exist yet in ontologies, we seek for the concept that could be extended with our new concept and put it in the "Extending ontology" column. To illustrate this, we can take the example of the term `Valve` present in the tab below. The concept associated to this term is `Valve`. Because the concept *`Valve`* exist in saref for building ontology and in the brick ontology we reuse *`s4bldg:Valve`* concept from saref for building ontology. However, for cooling and heating valve, saref for building doesn't include this specification but brick ontology does. However, *`CoolingValve`* and *`HeatingValve`* are sub concept of *`Valve`* which is defined as a kind of *`HVAC`* (*`brick:Valve is a brick:HVAC`*). Obviously, a valve is not a kind of HVAC but rather a component of it. Then in this case, we faced the problem of confusion of the relation *`is_a`* with the relation *`part_of`* or *`composition_of`* which is a frequent problem in ontology modeling. In our side because we do not agree this modeling, we choose to not reuse directly the cooling and heating valve and extends instead the saref for building ontology with new concept *`CoolingValve`* and *`HeatingValve`* with the relation *`rdfs:seeAlso brick:HeatingValve.`*

Reusing ontologies is not straightforward, as we can unfortunately be confronted with modelling errors such as the confusing between *`is_a`* and *`part_of`* relationship. This error is obviously serious because it corrupts the global consistency that leads to inference errors when querying the knowledge graph using these ontologies.

STEP 1		STEP 2					
Term	**is-it considered in the**	**Concept**	**Concept definition**	**Relation**	**Reusing Ontology**	**Extending Ontology**	**Comment**
HVAC							
HVAC System	Yes	HVAC	heating, cooling and air conditioning System		saref:HVAC; dogont:HVACSystem,brick:HVAC;		
HVAC operations; heating, cooling, preheating,	Yes	HeatingExecution/CoolingExec				pep:ProcedureExecutio	
HVAC controller	Yes	HVACValveController				s4bldg:Controller	HVAC controller is a subsystemOf HVAC. what does it controls exactly ?
valve	Yes	Valve			s4blg:Valve	subSystemOf saref:HVAC,	
Valve cooling control for the local controller point	Yes	CoolingValve				s4blg:Valve	
Valve heating control for the local controller point	Yes	HeatingValve				s4blg:Valve	
% Opening of valve	Yes	OpeningPercentageProperty	the opening percentage of a valve	hasOpeningPercentage		saref:Property; seas:PercentageProper...	
Heating period	Yes	HeatingExecution	Heating execution during a time	hasTempralContext		pep:ProcedureExecution	
Cooling period	Yes	CoolingExecution	Cooling execution during a time interval	hasTempralContext		pep:ProcedureExecution	
Schedule HVAC operation	Yes	HVACOperationSchedule				sch:Schedule	
Schedule HVAC operation includes period (preheating/precooling)	Yes	HVACOperationSchedule				sch:Schedule	
preheating periods	Yes	PreheatingExecution		hasTempralContext		pep:ProcedureExecutio	
precooling Period	Yes	PrecoolingExecution		hasTempralContext		pep:ProcedureExecutio	
Electricity consumption for cooling of the building	Yes	CoolingElectricityEnergyConsumptionProperty		hasCoolingElectricityEnergyConsumpti		seas:ConsumptionProperty	same than above
HeatingSystem	Yes	HeatingSystem				saref:HVAC; dogont:HVACSystem,bri	
CoolingSystem	Yes	CoolingSystem				saref:HVAC; dogont:HVACSystem,brick:HVAC; seas:HVAC;	
Building							
BMS (Building management System)	Yes	Building Management System			-		https://en.wikipedia.org/wiki/Building_management_system
building	Yes	Building			s4bldg:Building		
Office Building	Yes				semanco:Office; seas:Office; lcc:OfficeSite		
Cooling electricity consumption by hour of building occupancy	Yes	CoolingElectricityEnergyConsumptionProperty		hasElectricEnergyConsumption		seas:ConsumptionProperty	
Electricity consumption for cooling of the building	Yes	CoolingElectricityEnergyConsumptionProperty		hasCoolingElectricEnergyConsumptio		seas:ConsumptionProperty	same than above
Cs_heat_gas_occ: Heating gas consumption by hour of building occupancy	Yes	HeatingGasEnergyConsumptionProperty		hasGasEnergyConsumption		seas:ConsumptionProperty	
Gas consumption for heating of the building	Yes	HeatingGasEnergyConsumptionProperty		hasHeatingGasEnergyConsumption		seas:ConsumptionProperty	
Building zone	Yes	Zone			brick:Zone; fiemster:BuildingZone;		

Table 4: Example of ontology analysis

3.3 Application of Step 3 – Overview of ontological modules

The purpose of this step is to harmonize all diagrams proposed for each pilot and provide an illustrative example for some semantic data models. The harmonization plays an important role in pilots that present different diagrams for each use case. Regarding the similarity or difference existing between these models, some effort is being made to address this issue; an example is the efforts focusing on the analysis of same concepts between models of pilots. The fact is that each responsible of pilot has defined his model, but he/she doesn't have a sufficient overview of the existing ontologies, so he/she created a new concept or relation instead of reusing an existing one. We hoped to have a tool that allow us to generate these similitudes, but all this step is done manually. We also deal with all aspects related to the unified use of other concepts involved in a harmonization process, e.g., granularity, harmonization strategy, model quality, etc.

This section outlines the harmonization of multiple semantic data models that will be bring some benefits such as a clear definition of concepts and relations used in each pilot, a reference semantic data models under the same structure, with uniform and formal vocabulary, etc.

3.3.1 Common semantic data models

As we mentioned in the previous section, there are domains that are pilot independent and common to all pilots such as weather, sensor, meter, forecast, etc…

This section presents some extracts of these common modules

- **Weather:** The diagram of Figure 15, represents properties concerning the weather that are used in some pilot and the sensors dedicated to measure these properties. To represent these properties and their values we use the pattern used in seas EvaluationsOntology[1] (https:/W3id.org/seas/EvaluationOntology) that allow us exhaustively these aspects. This pattern is compatible with the measures pattern of saref ontology and ssn/soa ontologies. Furthermore, a mapping between these ontologies will be provided in second period of this task. As an example of modelling properties and its evaluation according to the seas pattern we can take, wind speed (`seas:WindSpeedProperty`) wind direction (`seas:WindDirectionProperty`), humidity (`saref:Humidity`) and air temperature (`plt:AirTemperatureProperty`) are properties (`owl:subClassOf saref:Property`) that can has multiple evaluation (`seas:evaluation`) for each property (`seas:WindSpeedEvaluation seas:WindDirectionEvaluation, seas:HumidityEvaluation and plt:AirTemperatureEvaluation`) respectively. These evaluations were observed or measured by a corresponding sensor represented by the relation `seas:observesProperty`. For example, the `plt:Anenometer` is `ow:subClassOf saref:Sensor` which `seas:observesProperty seas:WindSpeedProperty` and produces a `seas:WindSpeedEvaluation` related with the relation `seas:hasTemporalContext` to a specific time such as `time:Instant` or `time:Interval`. Another example is `plt:SolarRadiationProperty` that is extended into three concepts `plt:DiffuseSolarRadiationProperty`, `plt:DirectSolarRadiationProperty` and `plt:ReflectedSolarRadiationProperty` like previously each one of these properties has it corresponding evaluation which has its associated temporal context . All these concepts are defined in the PLATOON_weather_ontology (version 1). The semantic data model of weather is extended with new concepts and relationships. For example `plt:CloudCapacityProperty`, `plt:CloudCoverProperty`, and `plt:PrecipitationProperty` are new properties. Two other properties are defined as an `plt:AngleProperty; plt:ZenithProperty` and `plt:AzimuthProperty`. All and other new concepts and relationships are mentioned in Figure 16.

- **Sensor:** the concept Sensor is generic and common to different domains. However, each property has it dedicated type of sensor to observes it. For example, we have sensors used to observe weather properties like `seas:HumiditySensor` that `seas:observesProperty` the `saref:Humidity`, `dogonto:TemperatureSensor` that `seas:observesProperty` the `saref:Temperature` and `plt:Pyranometer` that `seas:observesProperty` the `plt:SolarRadiationProperty`. We have also some sensors used in the wind turbine domain, such as `plt:Anenometer` that `seas:observesProperty` the `seas:WindSpeedProperty` and `plt:WindVanSensor` that `seas:observesProperty` the `seas:WindDirectionProperty`. Furthermore, according to the design pattern used in seas, each sensor implements a `seas:Sensing` procedure and do some `seas:Observation` that generate the evaluation value of the property. In this manner we can link between the evaluation of one property to the observation made by a sensor via the relation `prov:wasGeneratedBy` from provenance ontology (https://www.w3.org/TR/prov-o).

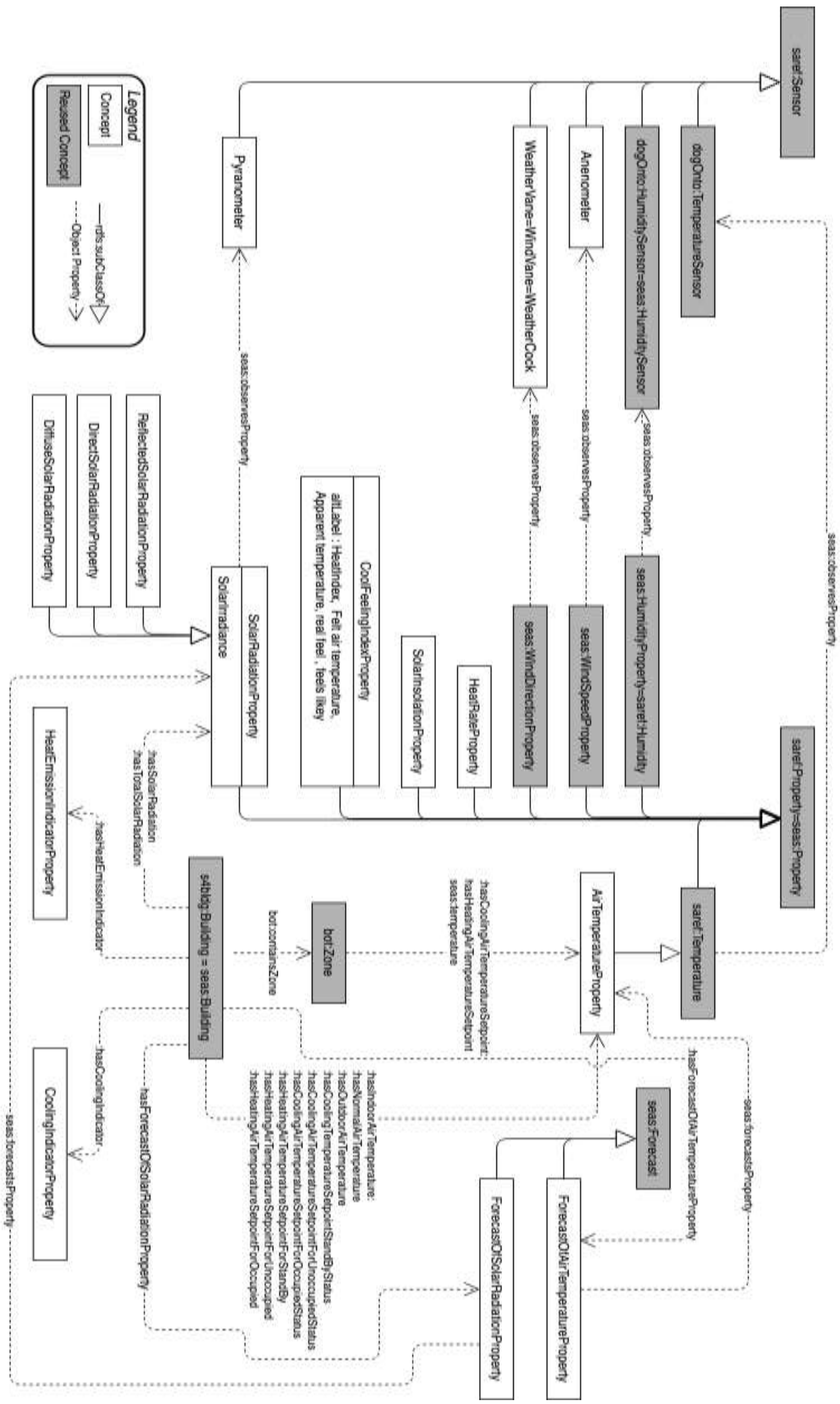


Figure 15: Semantic Data Model of Weather – version 1

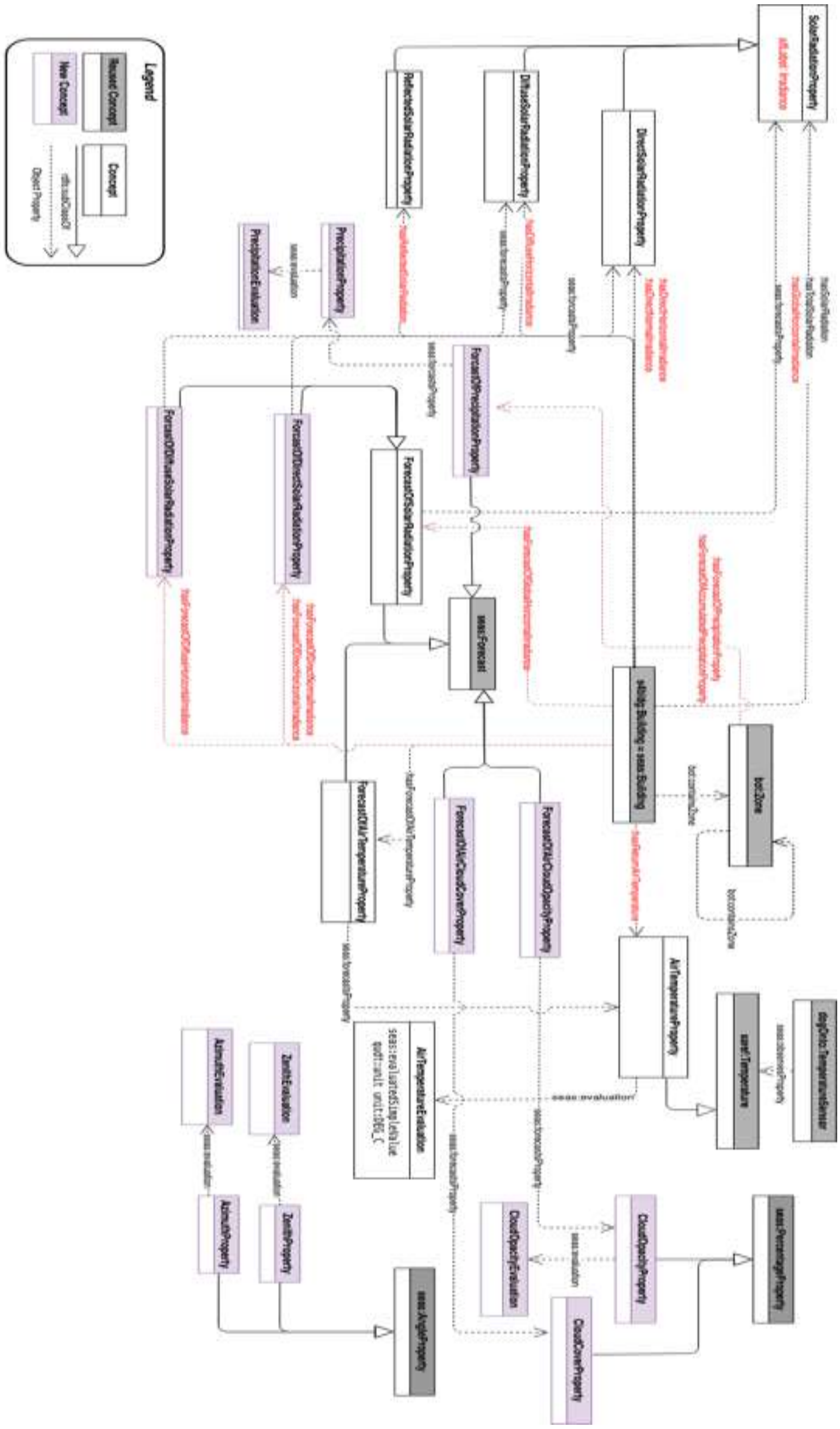


Figure 16-Semantic Data Model of Weather (extension)- version 2

- **Meter**: as Sensor, Meter is also a generic concept that is comment to different model. However, in the energy domain we have a specific meter to measure either consumption or production of energy of a building. Figure 17 shows an example; the `seas:ElectricEnergyMeter` that `seas:measuresProperty` the `seas:ElectricEnergyProperty` which could be either seas:ElectricEnergyConsumptionProperty or seas:ElectricEnergyProductionProperty. Another example is seas:GasMeter which is a meter that seas:measuresProperty the plt:GasFlowProperty. Like sensors, each meter implements seas:Metering procedure and do some seas:MeteringExecution.

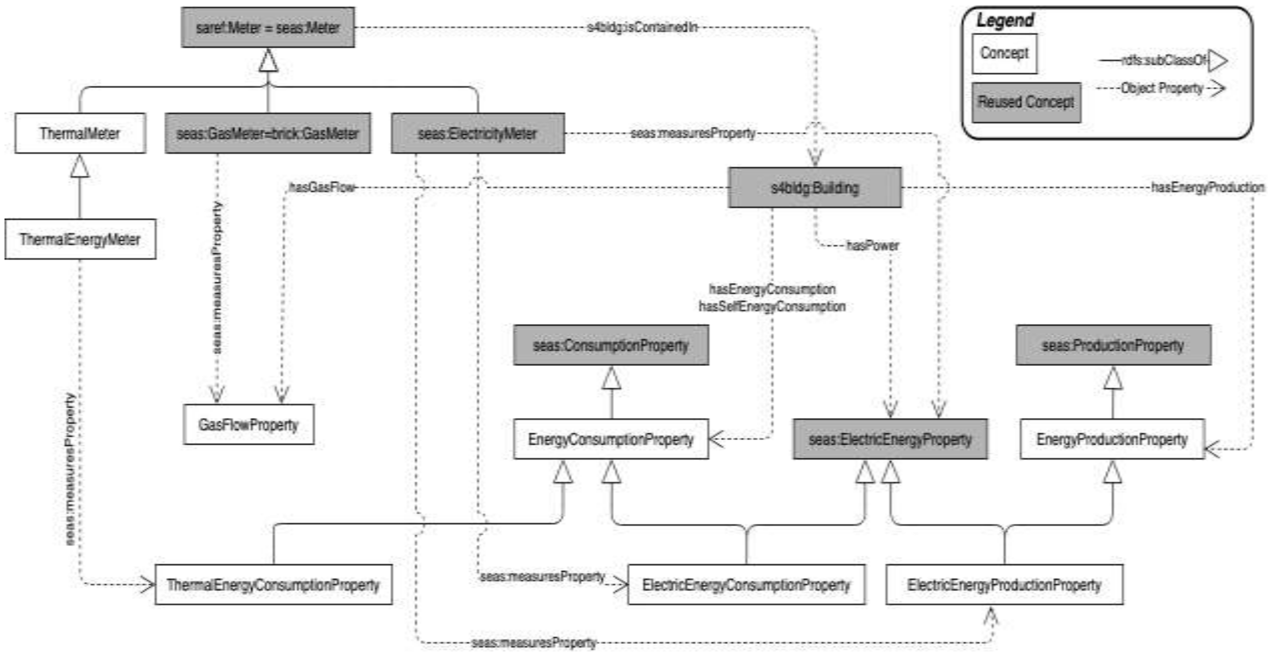


Figure 17: Semantic data model of Meter

- **Failures and Damages** are notions that concern several pilots such as #1a, #2a and #2b (see Fig. 18). Roughly speaking damages and failures can concern each entity that perform a function with a certain efficiency. In this ontology we try to describe failures in general way that should be extended to describe failures and damages in specific domain like wind turbine domain. We consider plt:Damage as more generic concept to represent all damages. It is a dependent concept that is inherent to a specific entity. Then, each plt:Damage is linked to an entity (seas:FeatureOfInterest) with the relation plt:isDamageOf. The inverse relation is plt:hasDamage that links an entity to its damage. The plt:Damage can leads to (plt:leadsTo) to failure event (cim:FailureEvent) which is an plt:unscheduledEvent. The cim:FailureEvent is related to the seas:FeatureOfInteres with the relationship plt:affects. A damage can have a level of criticality that is represented by the plt:hasSeverityLevel relationship that relates the plt:Damage to plt:SeverityLevelProperty. This later has an seas:evaluation plt:SeverityLevelEvaluation. Sometimes these failures are preceded by a plt:FailureAlert to alert about (saref:isAbout) the plt:FailureEvent. Furthermore, for damaged entity, in the case where a maintenance event is scheduled, we can relate this entity to plt:ScheduledMaintenance with plt:hasMaintenance relationship. Then from this general representation of failures and damages, we can represent a specific failure case of some pilot. For example, in the pilot #2a we want to represent the plt:DegradationConstantDamage which is a plt:Damage related to seas:SolarPanel with the relation plt:isDamageOf. This plt:DegradationConstantDamage may cause (`plt:cause`) an efficiency degradation (`plt:EfficiencyDegradation`) that `plt:leadsTo` a `cim:FailureEvent` and `plt:affects` the `seas:SolarPanel`.

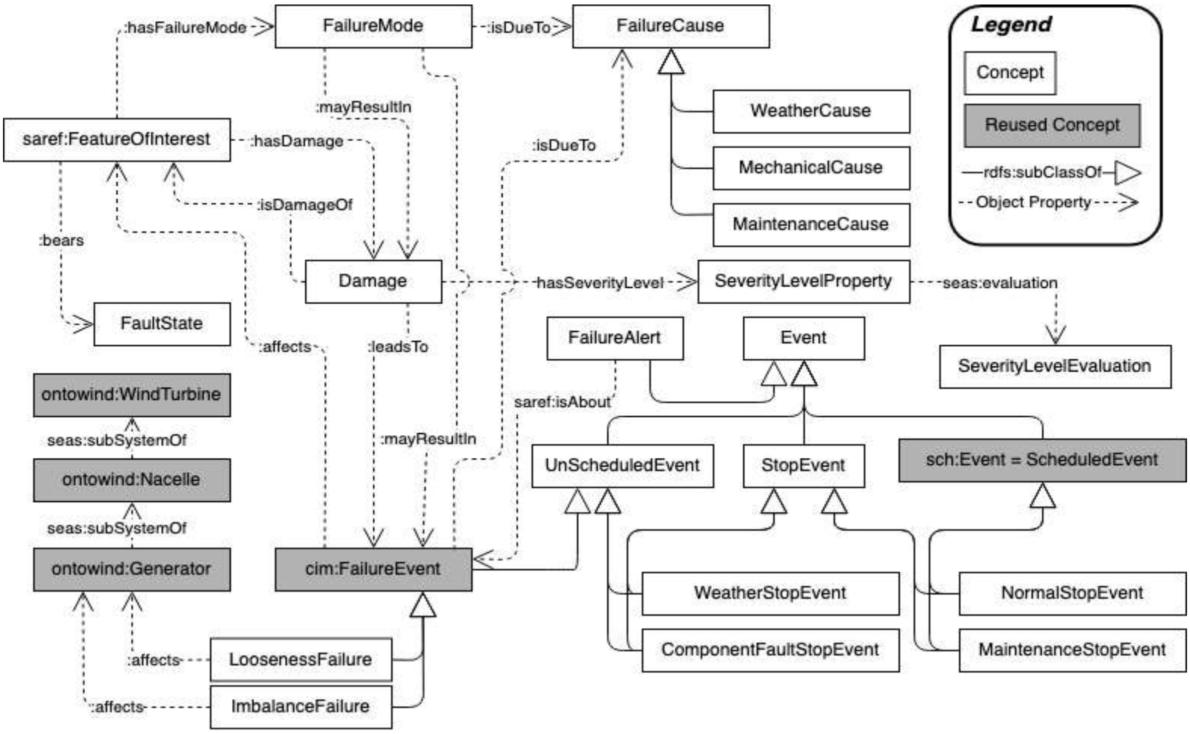


Figure 18: Common Semantic data model of Damage and Failure – version 1

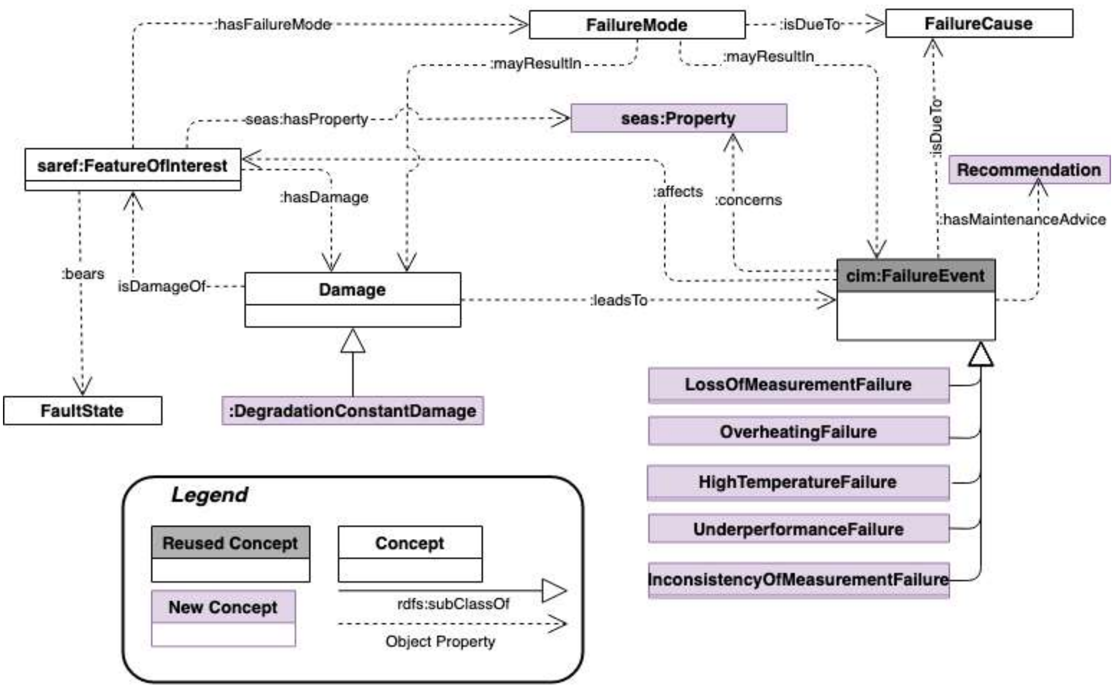


Figure 19 – Common Semantic data model of Damage and Failure (extension) – version 2

Figure 19 presents the extension of the module Failures and Damages, mainly according to the needs of the pilot #1a. These needs were identified during the transformation phase of the raw data into semantic data to create the knowledge graph. The added concepts are represented in violet color. The main added notion concerns the enrichment of the concept failure events adding some new specific failure events to meet the kind of failures represented in the data. Among this specific failures events we can cite: `plt:LossOfMeasurementFailure`, `plt:OverheatingFailure`, `plt:HighTemperatureFailure`, etc..

Another aspect of this extension is that failure event may concern (`plt:concerns`) a specific property (`seas:Property`) of an entity, for example `plt:HighTemperatureFailure plt:concerns saref:Temperature`. Recommendation (`plt:Recommendation`) is also a new aspect added to represent the knowledge about the maintenance advice (`plt:hasMaintenanceAdvice`) associated to a failure event.

- **Status code alarm** is the module of the ontology that represents the vocabulary of the status code and alarms related particularly to wind turbine systems (see Fig. 20). The goal of this module is to clearly represent the status code related to wind turbine to the notions of alarms. This modeling is closely cooled with status code information given by wind turbine control system. In this purpose, for each wind turbine, one status code property (`plt:StatusCodeProperty`) is associated through the relation `plt:hasStatusCode`. This property triggers (`plt:triggers`) alarms (`saref:Alarm`), and has evaluations (`plt:StatusCodeEvaluation`). The the status code evaluation which occurs intime has a specific status (`plt:hasStatus plt:Status`) and has renewable energy production status (`plt:hasRenewableEnergyProductionStatus`) that are status specific to renewable energy production. Four renewable energy production status are defined to distinguish between: (i) the information status (`plt:RenewableEnergyProductionInformationStatus`) that has the alarm type `pltr`[1]`:Information`, (ii) the warning status (`plt:RenewableEnergyProductionWarningStatus`) that has the alarm type `pltr:Warning`, (iii) the manual stop status (`plt:RenewableEnergyProductionManualStopStatus`) that has the alarm type `pltr:Breakdown` and (iv) the automatic stop status (`plt:RenewableEnergyProductionAutomaticStopStatus`) that has the alarm type `pltr:Breakdown`. The three alarm types (`pltr:Information, pltr:Warning and pltr:Breakdown`) are instances of `saref:Alarm`.

[1] Prefix pltr: <https://w3id.org/platoon/resource/> is the prefix for the URI associated to the instance defined in the platoon ontology

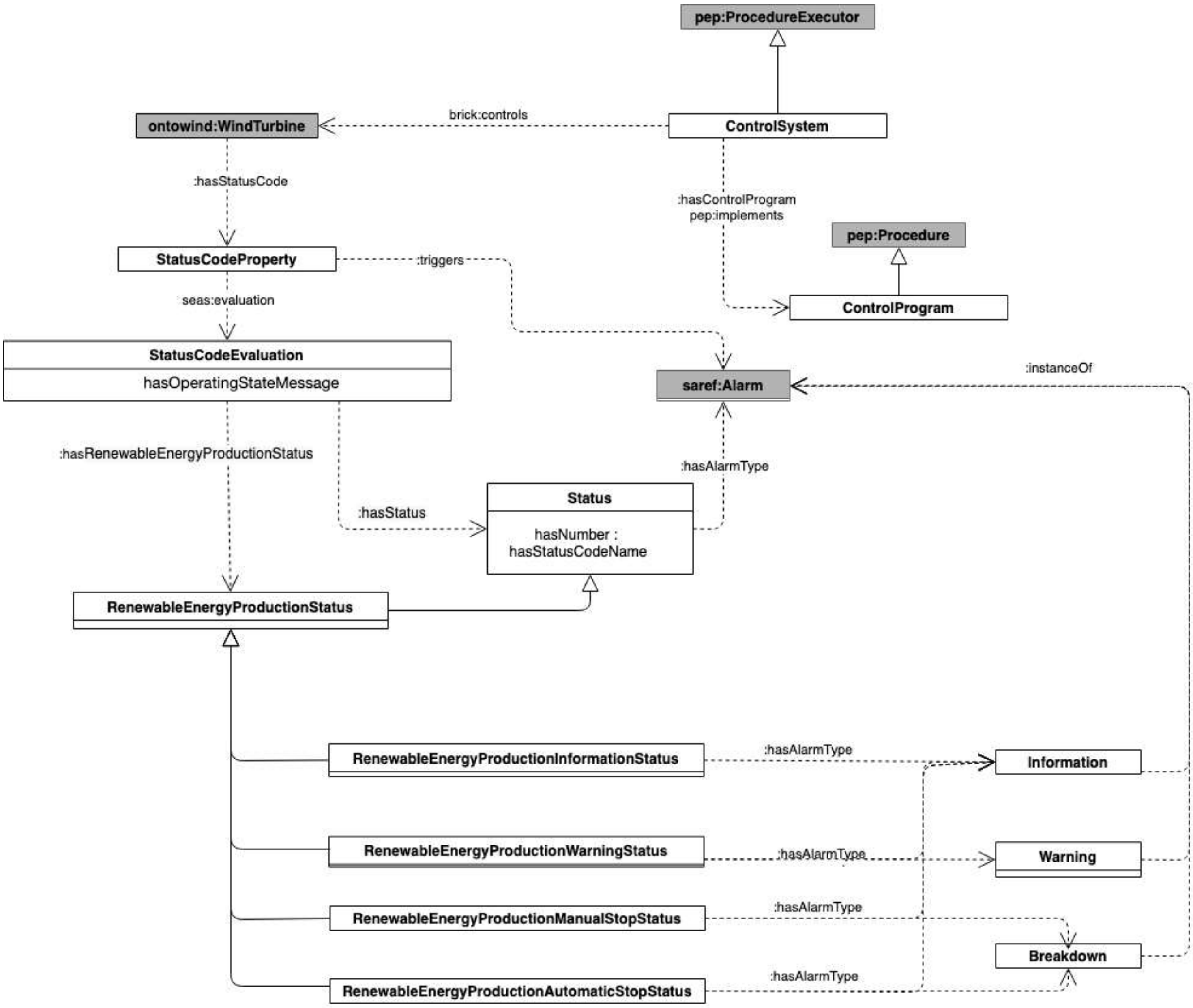


Figure 20 - Semantic data model of Status Code related to a Wind Turbine

3.3.2 Semantic data models of renewable energy – Pilot #1a

This Section presents the work performed within task T2.3 by main partners involved in development of applications for Pilot #1a (VUB, ENGIE, TECN, IAIS). The goal is to develop a harmonized PLATOON Semantic model for renewable energy and specially for wind turbine. Pilot #1a focuses on predictive maintenance of wind turbines.

Figure 21 detailed the concept of wind turbine and its components. The concept of wind turbine (`ontowind:WindTurbine`) is designed in the ontology OntoWind that described some concepts and relationships related to wind energy. There are two different types of wind turbines: onshore wind turbine (`plt:OnshoreWindTurbine`) which is located on land, and offshore wind turbine (`plt:OffshoreWindTurbine`) which is located in bodies of water. Three key components of wind turbine are defined in this pilot:

- Blade (`ontowind:Blade`): is the most critical piece of a wind turbine. It is designed to harness wind energy and drive the rotor of a wind turbine.
- Nacelle (`ontowind:Nacelle`): contains the mains subcomponents of the wind turbine. It includes the gearbox, the controller (`plt:Controller`) and the electrical generator (`ontowind:Generator`). The generator can be an induction generator (`ontowind:inductionGenrator`) or double fed induction Generator

(ontowind:DoubleFedInductionGenerator. Different systems are linked to the electrical generator such as the rotor (ontowind:Rotor), the stator (plt:Stator) and the plt:GeneratorCooling and plt:GeneratorBearing. The plt:GeneratorBearing can be plt:NonDriveEndBearing and plt:DriveEndBearing. The rotor is the system that transforms the kinetic energy of wind to mechanic energy. It is linked to the rotor winding (plt:RotorWinding) and a slip ring (plt:SlipRing). A rotor is connected to (seas:connectedTo) a stator (plt:Stator). The stator is an important electrical part of the wind turbine. It contains all the coils of wire which comprised of fiber glass and resin. A stator winding (plt:StatorWinding) and stator frame (plt:StatorFrame) are sub systems of the stator.

- Convertor (plt:Convertor): is linked to the system of IGBT (plt:InsulatedGateBipolarTranslator. It can adjust the generator frequency and voltage to the grid.

An electric power transformer (seas:ElectricPowerTransformer) is an electric power system that can transform electrical voltage within a power network, between a primary connection and a secondary connection point. This system is connected to (seas:connectedTo) ontowind:WindTurbine.

Different other properties are designed in this pilot that are related to the current, voltage, electric power, vibration, temperature which are observed by specific sensors. The rotational speed of the rotor is also measured to indicate the original or the nominal value (see details in github[2])

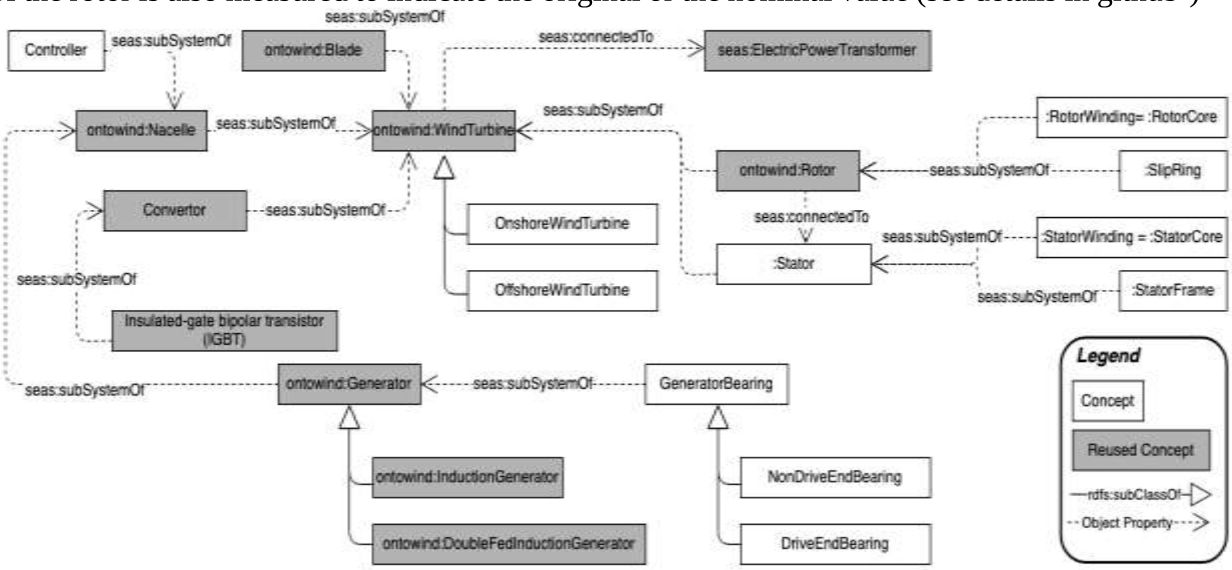


Figure 21: Semantic data model of Wind Turbine components – version 1

2 https://github.com/PLATOONProject/semantic-data-model/tree/main/Semantic-Data-Models-Diagrams/Pilot%231a-Renewable-Energy

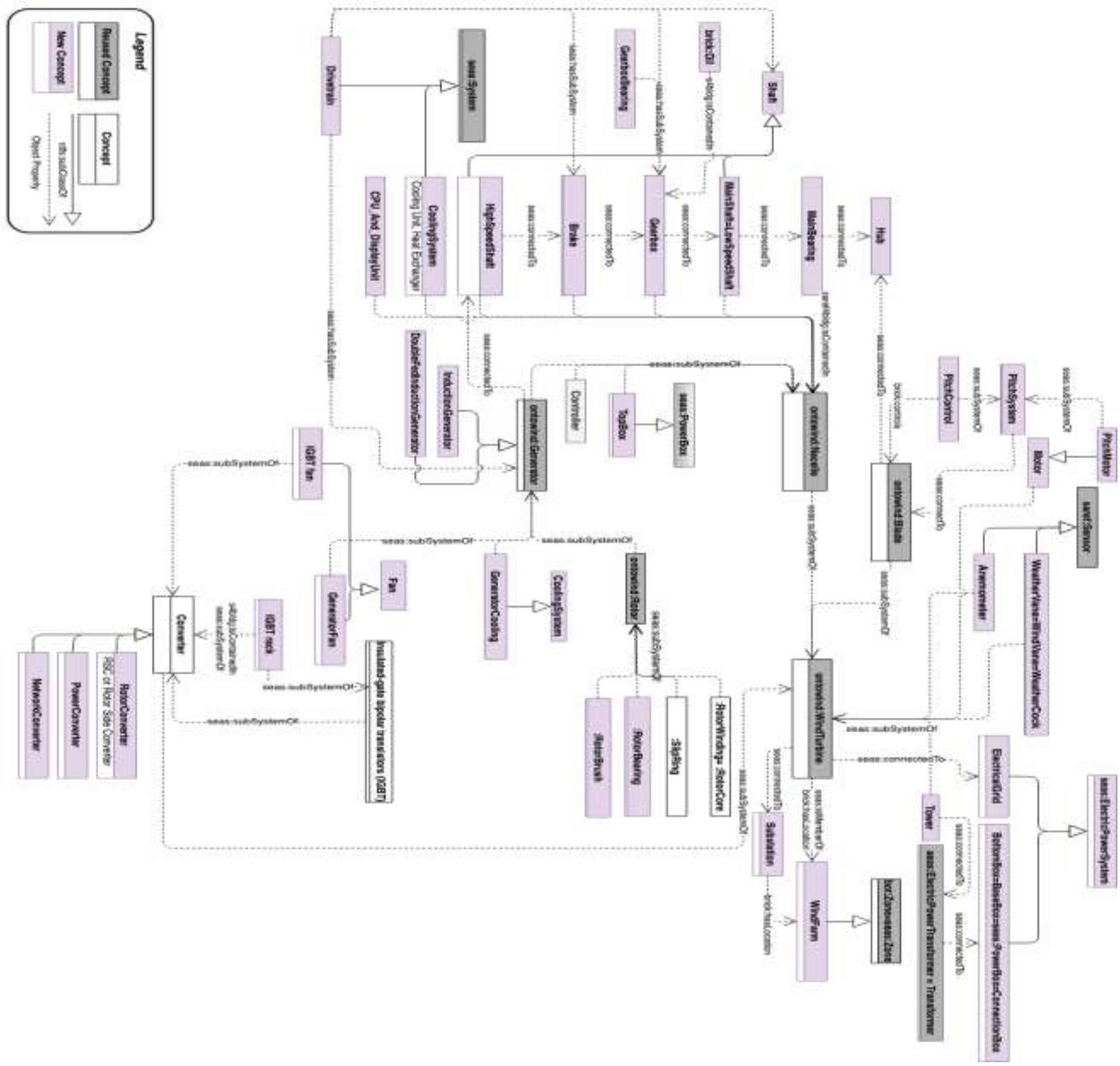


Figure 22- Semantic data model of Wind Turbine components – version 2

During the creation step of the knowledge graph associated to the pilot #1a based in data coming from heterogenous systems like SCADA data, maintenance data, prediction system, etc., several new concepts and relationships are added to the wind turbine ontology module. These new elements are represented in purple in the Figure 22. The following section, details some of these concepts and relations:

- Wind farm (`plt:WindFarm)` is composed of a group of wind turbine in the same location used to produce electricity. The wind turbine is related to the wind farm by the (`seas:isMemberOf`) relation.
- In the wind turbine there are two kind of shafts: (i) Main shaft also called Low speed shaft (`plt:MainShaft owl:equivalentClass plt:LowSpeedShaft`) is a subclass of shaft (`plt:Shaft`). The Main shaft is connected to (`seas:connectedTo`) main bearing (`plt:MainBearing`) and gearbox (`seas:GearBox`). and contained in (`s4bldg:isContainedIn`) the nacelle (`plt:Nacelle`); (ii) High speed shaft (`plt:HighSpeedShaft`) is a subclass of shaft (`plt:Shaft)` .

The high speed shaft is connected to the brake (plt:Brake) and contained in (s4bldg:isContainedIn) the nacelle (plt:Nacelle).

- In the nacelle, there are components that are contained in it:
  - CPU and display Unit (plt:CPU_And_DisplayUnit) is an unit that allows to display locally the information concerning the wind turbine.
  - Cooling system (plt:CoolingSystem) is a system which has the role to cool the components in the Nacelle
  - In the nacelle there are two kind of shafts: (i) Main shaft also called Low speed shaft (plt:MainShaft owl:equivalentClass plt:LowSpeedShaft) is a subclass of shaft (plt:Shaft). The Main shaft is connected to (seas:connectedTo) main bearing (plt:MainBearing) and gearbox (seas:GearBox). and contained in (s4bldg:isContainedIn) the nacelle (plt:Nacelle); (ii) High speed shaft (plt:HighSpeedShaft) is a subclass of shaft (plt:Shaft). The high speed shaft is connected to the brake (plt:Brake) and contained in (s4bldg:isContainedIn) the nacelle (plt:Nacelle)Brake (plt:Brake) is a system that allows to slow down the blade, keeping the turbine rotation at a safe speed even in fast wind. The brake is connected to the high speed shaft in one side and to the gearbox (seas:GearBox) in the other side.
  - Gearbox (plt:Gearbox): a gearbox is typically used in a wind turbine to increase rotational speed from a low-speed rotor to a higher speed electrical generator. Gearbox is connected to the main shaft (plt:MainShaft) in one side and the brake (plt:Brake) in the other side. The gearbox has subsystem GearboxBearing (plt:GearboxBearing), which is a kind of rolling-element bearing similar to an epicyclic gear.
  - Main bearing (plt:MainBearing) realized to bear a shaft of a wind turbine, which shaft is caused to rotate by a number of blades connected to the shaft. The Main bearing is contained in (s4bldg:isContainedIn) the nacelle (plt:Nacelle) and connected to the hub.
- Hub (plt:Hub) : is a component that connects the blades to the main shaft. It is connected to (seas:connectedTo) the main bearing (plt:MainBearing) .

  - TopBox (plt:TopBox) is a kind of power box (seas:PowerBox) that is a subsystem of (seas:subSystemOf) the Nacelle (plt:Nacelle).
  - WeatherVane (plt:WeatherVane) also called WindVane or WeatherCock is a sensor (saref:Sensor) which measures (seas:observesProperty) the wind direction (seas:WindDirectionProperty) and communicates with the yaw drive to orient the turbine properly with respect to the wind.
  - Anemometer (plt:Anemometer) is a sensor that measures (seas:observesProperty) the wind speed property (seas:WindSpeedProperty).
  - GeneratorFan (plt:GeneratorFan) is a fan and a sub-system of a generator that ensures the needed load, safe life time and efficiency of the generator.
  - Fan (plt:Fan) use wind to make electricity.

- InductionGenerator (`plt:InductionGenerator`) : is a type of generator that produces an alternating current (AC) that uses the principles of AC to produce electric power.
- DoubleFedInductionGenerator (`plt:DoubleFedInductionGenerator`) : is a generator that feeds the rotor with currents of varying frequency, in order to reach the desired rotor speeds. It is connected to (`seas:connectedTo`) to the high speed shaft (`plt:HighSpeedShaft`).
- RotorConverter (`plt:RotorConverter`) is a convert that is connected to the Rotor
- GeneratorCooling (`plt:GeneratorCooling`) is a cooling system (plt :CoolingSystem) that has for function to cool (`seas:subSystemOf`) the generator (`plt:Generator`)
- IGBT rack (`plt:IGBTRack`) is a container of the plt:Insulated-GateBipolarTransistor(IGBT) that are (`seas:isMemeberOf`) It.the IGBT rack is contained is (`s4bldg:isContainedIn`) and sub system of (`seas:subSystemOf`) the converter (`plt:Converter`).
- IGBT fan (`plt:IGBTDFan`) is a fan (`plt:Fan`) that is used to cool the IGBT it is also sub system of the converter.
- PitchMotor (`plt:PitchMotor`) is a motor (`plt:Motor`) which is a subsystem of pitch system (`plt:PitchSystem`).
- Pitch system (`plt:PitchSystem`) is a system which is connected to the blade.
- Tower (`plt:Tower`) is a vertical structure that carries the nacelle and the rotor. It is connected to the transformer (`seas:ElectricPowerTransformer`)
- RotorBearing (`plt:RotorBearing`) is a bearing that is subsystem of the rotor (`plt:Rotor`)
- BottomBox (`plt:BottomBox`) is a power box that is connected to the transformer

3.3.3 Semantic data models of Grids – Pilots #2a, #2b and #4a

This Section presents the work performed within task T2.3 by main partners involved in development of applications for Pilot 2a (PUPIN, CS, UBO, ENGIE), Pilot 2b (TECN, SAM, IND, ENGIE) and Pilot 4a (PDM, ENGIE).

The goal is to develop a consolidated PLATOON Semantic model for Smart Grids. The modelling activities performed during this task have concentrated on gathering information about the common semantic concepts and properties applicable for targeted scenarios. Different existing data models have been consulted and considered for reuse in these pilots such as CIM, SEAS, SAREF, OEMA, etc.

The goal of the pilot #2a is to analyse the balancing of the energy and the predictive maintenance.

The pilot #2a semantic model is mainly influenced by two existing ontologies:

- the IEC Common Information Model (CIM) standards, see CIM V2.53.0 Schema (MOF, PDF, and UML, https://www.dmtf.org/standards/cim/cim_schema_v2530). In our analysis, we have used the semantic CIM model available at https://ontology.tno.nl/IEC_CIM/. It is a canonical taxonomy in the form of packages of UML class diagrams referring to the components of power utility networks with functional definitions and measurement types to a high degree of granularity (packages: Core, Topology, Wires, Generation, LoadModel, Outage, SCADA, ControlArea and others), and

- SEAS - Smart Energy Aware Systems, https://w3id.org/seas/.

In this step 3 of the proposed methodology, we redefined the relations between concepts of CIM used in the PLATOON semantic data models.
The concepts that will be reused in Pilot #2a come from different packages (see Fig. 23). For instance, `cim:PowerSystemResource` (Core package) can be an item of equipment such as a Switch, an `cim:EquipmentContainer` containing many individual items of equipment such as a Substation. Each `cim:PowerSystemResource is registered on the grid (cim:RegisteredResource)` and belong to a control area (`cim:HostControlArea`) that is operated by a `cim:ControlAreaOperator`.

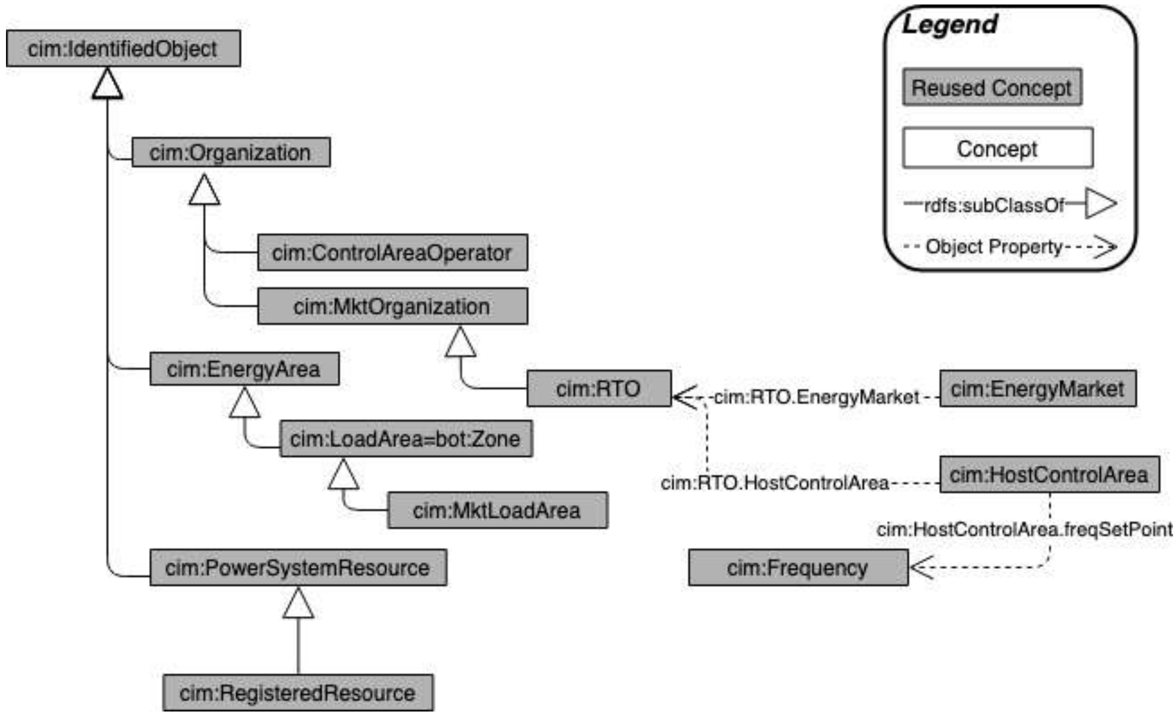


Figure 23: Semantic data model of Organization

The `cim:ControlAreaOperator` is responsible for stabilizing the system frequency (`cim:Frequency`); it is therefore also called frequency control. The system is balanced by utilizing both supply and demand resources. However, the existing electric power systems were not initially designed to incorporate different kinds of generation technology (`cim:Plant`) in the scale that is required today.
Historically, balancing the system has been maintained mostly by directing thermal power plants to increase or reduce output (`cim:ActivePower`) in line with changes in demand. With significant penetration of distributed generation, the distribution network has become an active system with power flows () and voltages () determined by the generation as well as by the loads (see `cim:AreaLoadCurve`). In order to forecast the load for the next period `cim:ControlAreaOperator` needs a service for load prediction (LLUC P-2a- 03) for

different `cim:LoadForecastType.` The service is modelled with common semantic data models (see details in GitHub[3]).

The use case #2b-01 focuses on transformer predictive maintenance, estimating transformer components health and its maintenance costs, planning maintenance actions, monitoring transformer alarms, and studying different grid scenarios in case of replacing old transformers or adding complementary transformers. The key concept of this use case is the transformer *(s4bldg:Transformer)*. The electric power transformer is an electric power system that is capable of transforming electricity within a power network, between a primary connection point and a secondary connection point. Different systems are linked to the transformer such as *plt:SecondaryWinding*, *plt:PrimaryWinding*, *plt:Insultation* and *plt:Casing* (see Fig. 24). The transformer is connected to *(seas:connectedTo) plt:ElectricalSubstation* and it is contained in a transformer center *(:TranformerCenter)*. Different properties are specified in this pilot to estimate the Transformer Oil Temperature the and the health index *(plt:HealthIndexProperty)* to estimate the remaining useful life of the transformer, taking into account some of the most important characteristics of its insulation system.

The subsystems of the transformer possess additional properties such as active power *(cim:ActivePower)*, reactive power (*cim:ReactivePower*), voltage (*seas:VoltageProperty)* and current (*seas:CurrentProperty*).

The transformer is connected to a prosumer (*oema:Prosumer*) that blurs the distinction between consumer and producer. Furthermore, use case 2b-02 focuses on the Non-Technical Loss (NTL) detection based on the energy balance of the consumed electric power *(saref:Power)* of the transformer and Smart Meter readings from the prosumers.

[3] https://github.com/PLATOONProject/semantic-data-model/tree/main/Semantic-Data-Models-Diagrams/Pilots%232a_%232b_%234a-Grids

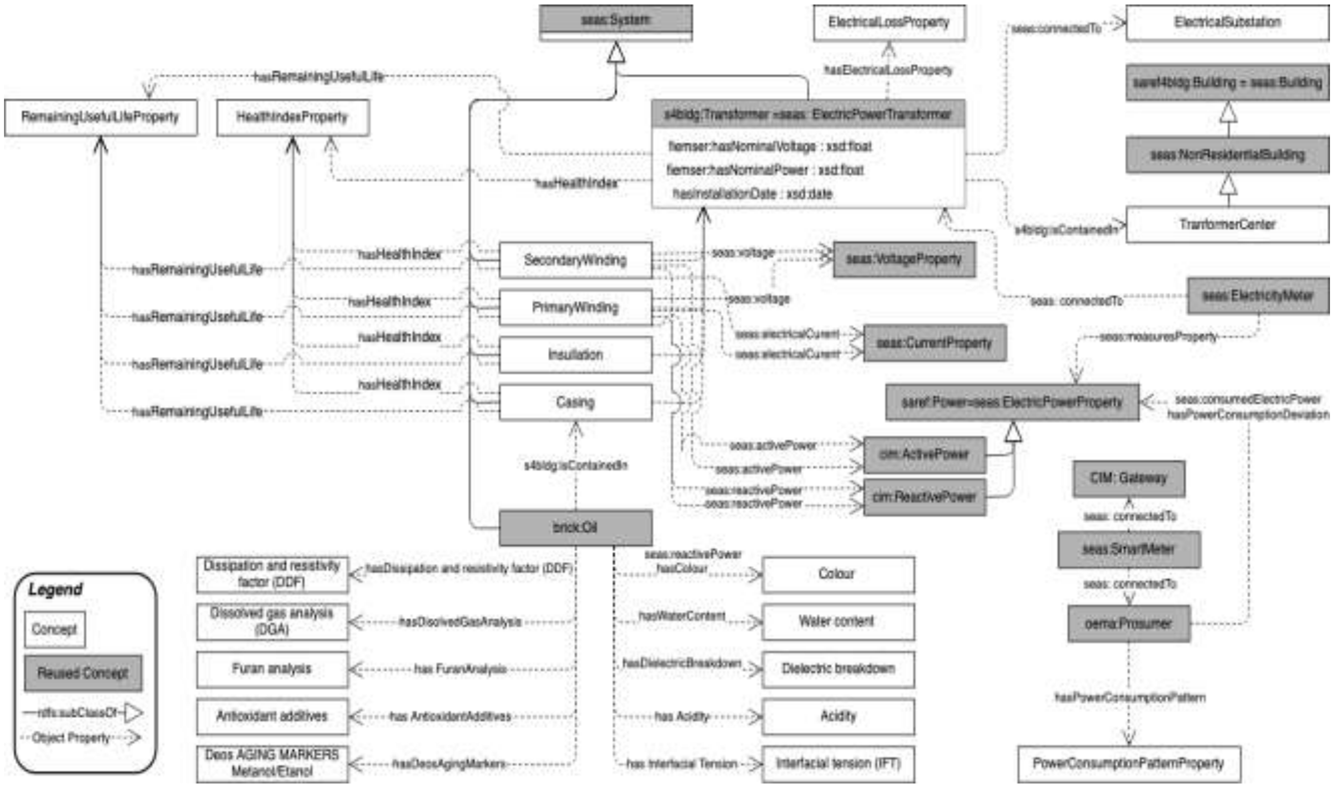


Figure 24: Semantic data model of Electric Power Transformer

The pilot #4a aims to study data-driven energy management able to deal with increased complexity of the energy system. The key concept of this pilot is the storage system (`plt:StorageSystem`). Figure 25 shows different types of storages such as `seas:ElectricPowerStorageSystem`, `plt:HydrogenPowerToPowerSystem`, `plt:OxygenStorageSystem`, `plt:HydrogenStorageSystem`, `plt:ThermalStorageSystem` and a `seas:Battery`.

A battery (`seas:Battery`) has the ratio (`seas:stateOfChargeRatio`) and the setpoint (`seas:stateOfChargeRatioSetpoint)` of the charging state, which is a percentage (`seas:PercentageProperty`). Figure 26 shows an extension of the semantic data model of the storage system. A new Storage system is defined `plt:SolarStorageSystem`. Different properties are also specified such as `plt:StorageGoodProperty`, `plt:StorageTimeReserveProperty`, `plt:StorageCapacityProperty`. In this extended semantic data model, several relationships are described `(e.g., plt:hasStorageChargingEfficiency, plt:hasStorageTimeReserve).`

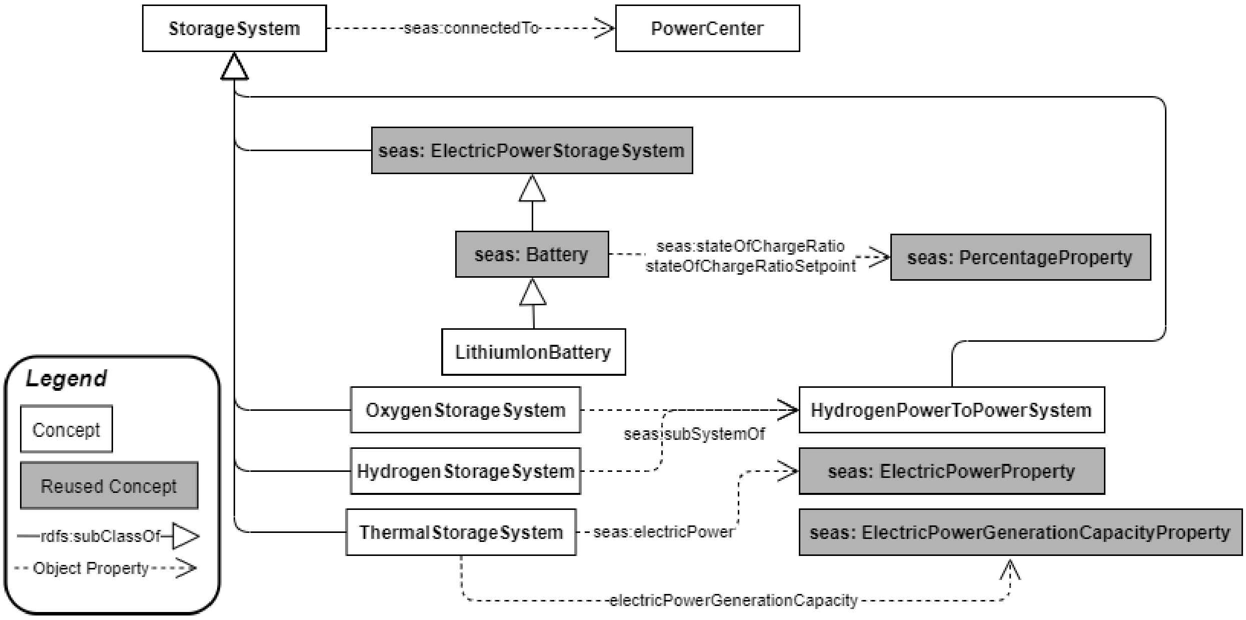


Figure 25: Semantic data model related to the storage systems – version 1

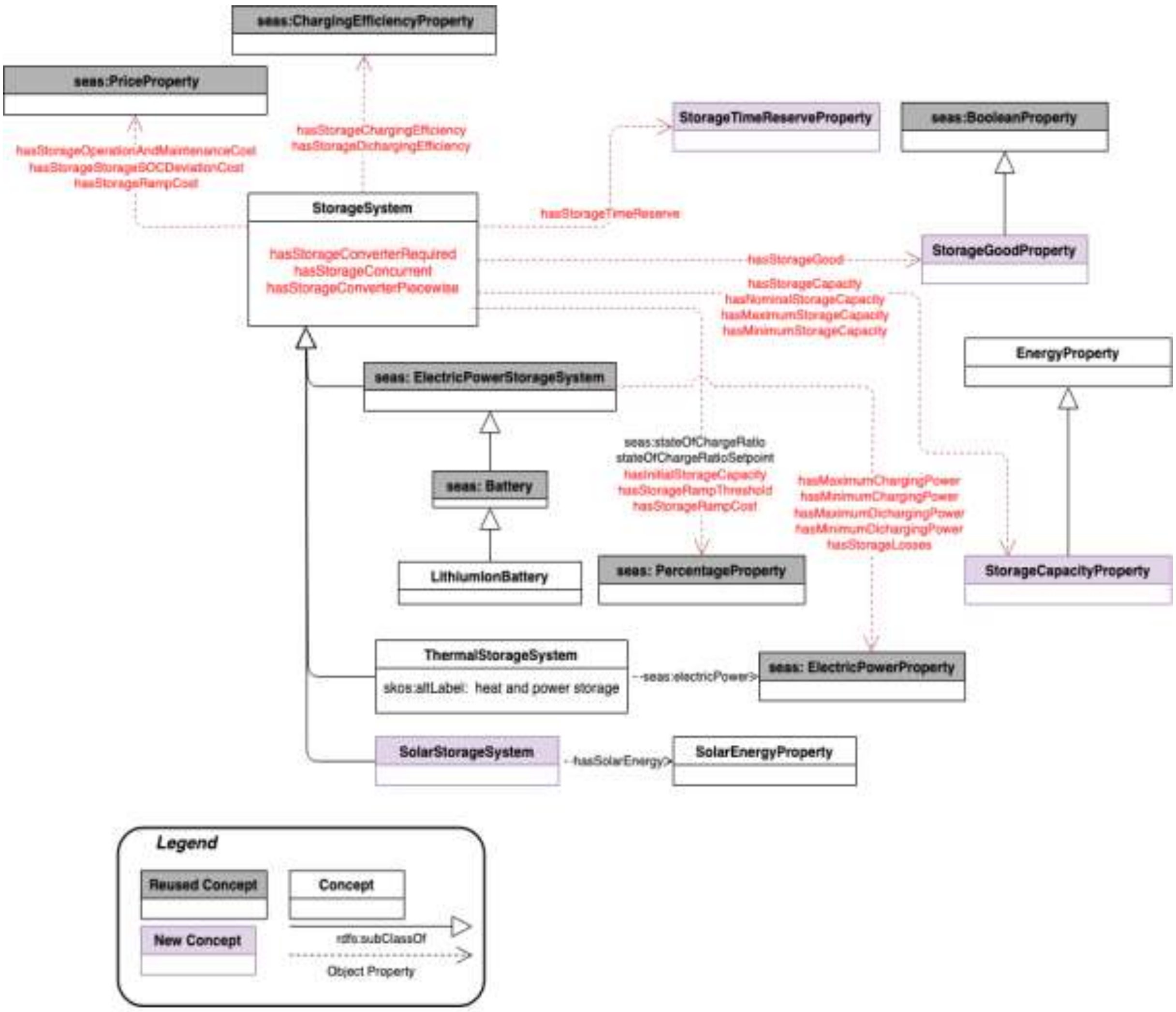


Figure 26: Semantic data model related to the storage systems (extension) - version 2

Another key concept in the pilot #4a is related to the grid (*plt:ElectricalGrid*) which is an *seas:ElectricPowerSystem* and defined as an interconnected network for distributing electricity from producers to consumers (see Fig. 27). Various kinds of electric power systems are designed such as *plt:ChargingStation* (two types are cited: *seas:ElectricVehiculeChargingStation, plt:BikeChargingStation), seas:ElectricVehicule*, *plt:ElectricBike, plt:ElectricHub, etc.* On type of the electrical grid is *plt:SmartMicroGrid* which is connected to *seas:ElectricPowerProducer.* From the following types of the electric power producer, we have a photovoltainc plant (*semanco:PVSystem*). This system is composed of solar strings (*plt:SolarArray*) and solar modules (*plt:SolarModule*). All *plt:SolarArray* are connected in parallel (*plt:isConnectedInParallel)* and all *plt:SolarModule* are connected in series *(plt:isConnectedInSeries).*

In the Figure 27, many other properties have been defined (e.g., *plt:ElectricityEnergyConsumptionProperty, plt:ElectricEnergyLoadProperty*, etc).

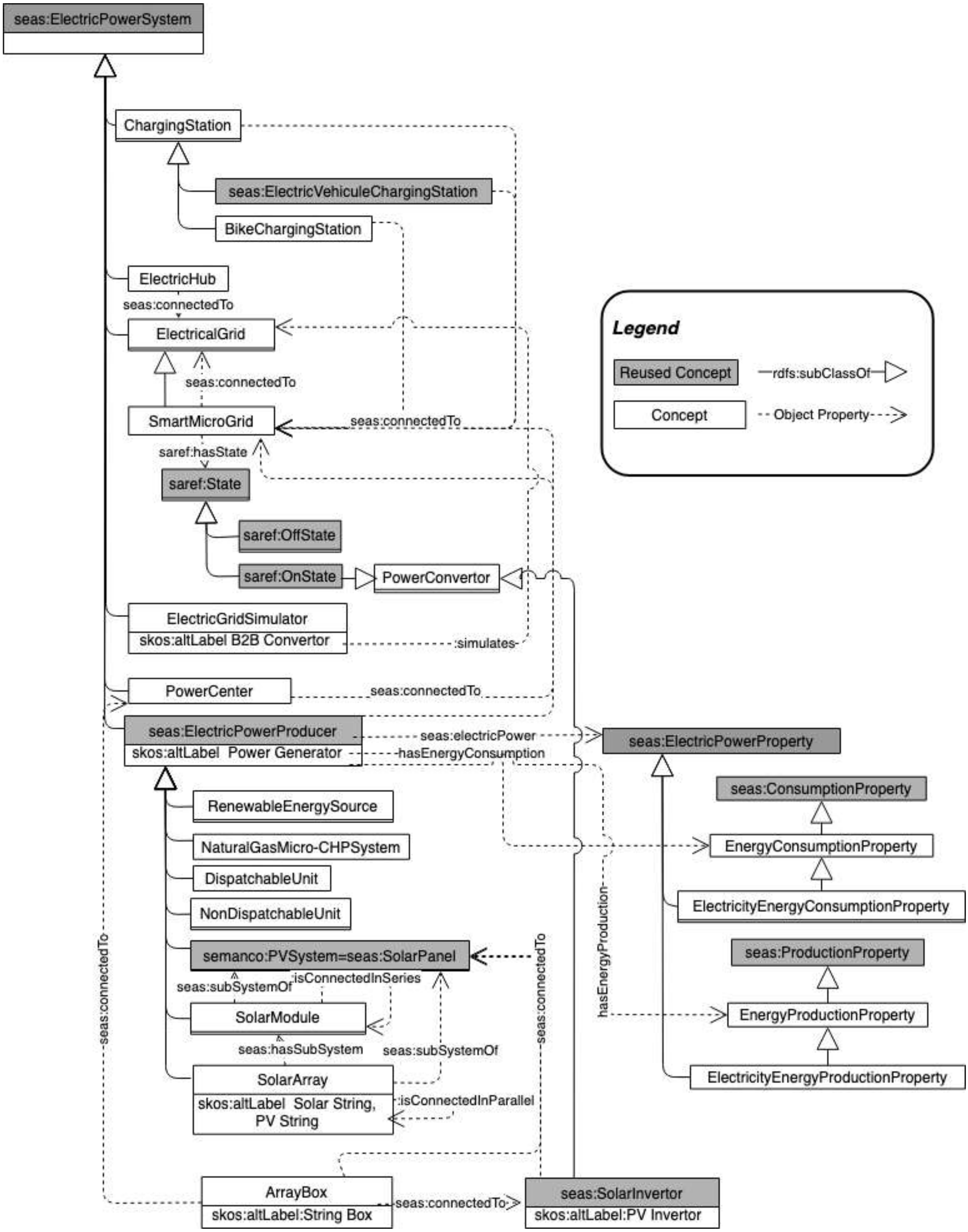


Figure 27: Semantic data model related to the Grid

3.3.4 Semantic data models of Building and HVAC – Pilots #3a, #3b and #3c

The pilot #3a aims to meet two goals in an office building. On the one hand, it aims to optimize the control of the HVAC regarding occupancy. On the other hand, it aims to providing a demand response service through HVAC control. To reach this purpose, two mains' elements are in this model: Building and HVAC. We present below an extract of our modelling of these two mains elements. Obviously, the modelling elements related to the building and HVAC are the same used in the pilot #3b and #3c.

Figure 28 represents an extract of Building and its occupancy and comfort properties. The building concept is represented by *s4bldg:Building* from SAREF for building ontology which is equivalent to the concept *seas:Building* from seas ontology. The building can be subdivided in different zones (*bot:Zone*) and is linked to these zones with the (*bot:containsZone*) relationship. Each zone (*bot:Zone*) has an occupancy property (*saref:Occupancy*). The zone and its occupancy property can be related with three types of occupancy relationships: the first one *plt:hasOccupancy* which is a generic property of occupancy that can be specialized on two sub properties *plt:hasClientsOccupancy* to indicate that the zone is occupied by customers and *plt:hasEmployeesOccupancy* to indicate that the zone is occupied by employees. The occupancy can be assessed either by some statistical algorithm or detected by occupancy sensor (*dogOnto:OccuapncySensor*) which observes the occupancy property (*seas:ObservesProperty*). We can associate two types of evaluation to the occupancy. The first is the occupancy status (*plt:OccupancyStatusEvaluation*) to know if the zone is occupied or unoccupied. The second is the number of persons that occupy the zone (*plt:OccupiedNumber-Evaluation*). As each Evaluation, these two evaluations have a temporal context of validity. For optimization of HVAC control, forecasting of the occupancy is important to schedule the cooling and the heating of the zones. This forecast is represented by *plt:ForecastOfOccupancy* concept that forecasts (*seas:forecastsProperty*) the *saref:Occupancy* Property. The optimization should respect for each *bot:Zone* the comfort level (*plt:ComfortLevelProperty*) which depends on (*plt:dependsOn*) *saref:Temperature* and *saref:Humidity*. This comfort level has four levels of evaluation (*seas:evaluation*). We can distinguish between four levels of comfort: *plt:ThresholdComfortLevelEvaluation*, *plt:MinimumThreshold ComfortLevelEvaluation*, *plt:BasicComfortLevelEvaluation* and *plt:BestComfortLevelEvaluation*.

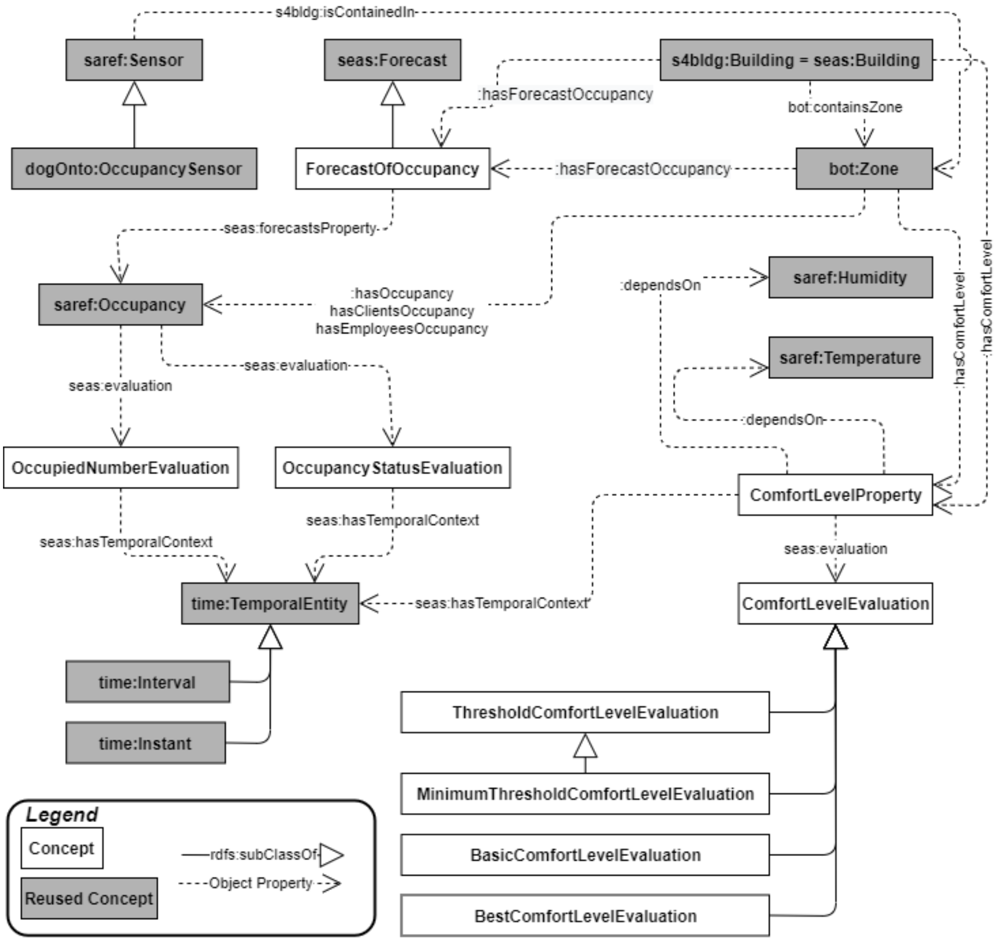


Figure 28: Semantic data model of Building properties

Figure 29 presents the HVAC data model. For the HVAC domain, brick ontology was identified as a good candidate that covers a big part of the HVAC domain. However after analyzing the taxonomy of this ontology, we detect modelling error concerning the use of subsumption relationship (`rdfs:subClassOf`). Indeed, the subsumption relationship was confused with the part of (`part_of`) relationship. For example, the boiler, pump, fan, filter, and valve are all considered as HVAC whereas that are part of the HVAC system. Then for brick concepts we prefer to use the `rdfs:seeAlso` relationship. The `HVAC` is a system has three subsystems: the `plt:HeatingSystem` that is related by the relation `plt:hasHeatingSystem`, the `plt:CoolingSystem` that is related by the relation `plt:hasCoolingSystem` and the `plt:VentilationSystem` that is related by the relation `plt:hasVentilation-System`. `Saref:HVAC` is located (`s4bldg:is-ContainedIn`) in `s4bldg:Build-ing`. This HVAC is controlled by a `plt:BuildingManagementSystem` that controls `(brick:controls)` the HVAC.

The cooling system and heating system has respectively `plt:CoolingValve` and `plt:HeatingValve` that are their subsystem (`seas:subSystemOf`). Each valve is located in (`s4bldg:isContainedIn`) a zone and has opening percentage property (`plt:OpeningPercentageProperty`) related by the relationship `plt:hasOpening Percentage`. The HVAC has three kind of operations (`plt:HVACOperationEvent`) that are a type of `pep:ProcedurExecution` and type of `plt:ScheduledEvent`. The operations are:(i) `plt:HeatingExecution` performed by (`ssn:madeByExecutor`) the `plt:HeatingSystem`. The heating execution can has two modes represented by `plt:hasMode` data property that can has the value "`preheating`" or "`heating`". (ii) `plt:CoolingExecution` performed by (`ssn:madeByExecutor`) the `plt:CoolingSystem`. The cooling execution can has two modes represented by `plt:hasMode` data property that can has the value "`precooling`" or "`cooling`". (iii) `plt:VentilationExecution` performed by (`ssn:madeByExecutor`) the `plt:VentilationSystem`. Each these kind of procedure executions has a temporal context (`seas:hasTemporalContext`) because that happened in a specific time.

Sometime, in HVAC system there is also an `plt:AirHandlerUnit` that is a `seas:subSystemOf` the `saref:HVAC`. This air handler is connected to the heating, the cooling and the ventilation systems. A boiler (`plt:Boiler`) or heating coil (`plt:HeatingCoil`) can be subSystem of the heating system and the chiller (`plt:Chiller`) or cooling coil (`plt:CoolingCoil`) is a subSystem of cooling system. The `plt:Fan-coil` that has a `plt:Fan` and `s4bldg:Coil` as a subsystem, is connected to the air handler unit.

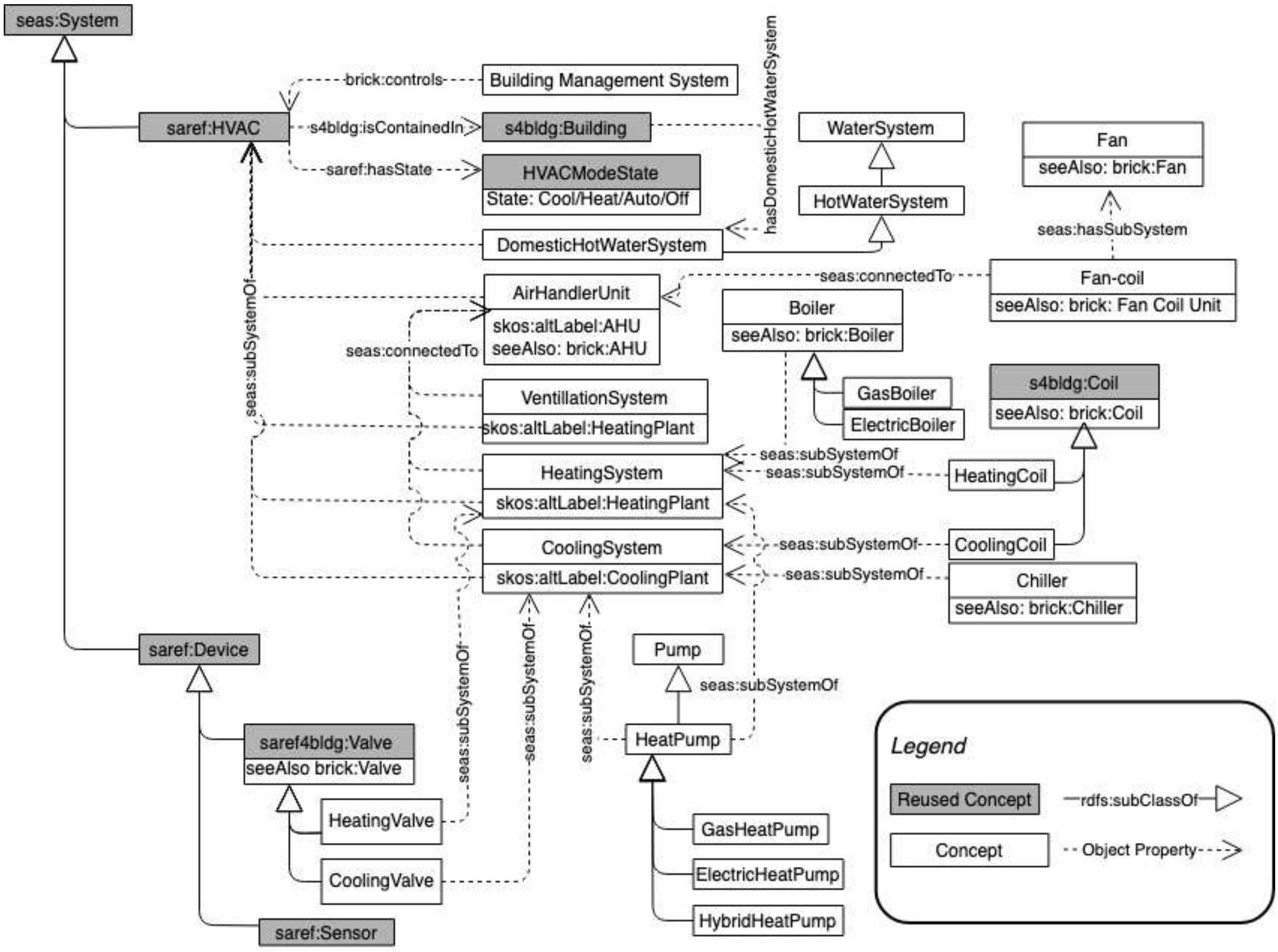


Figure 29: Extract of HVAC system – version 1

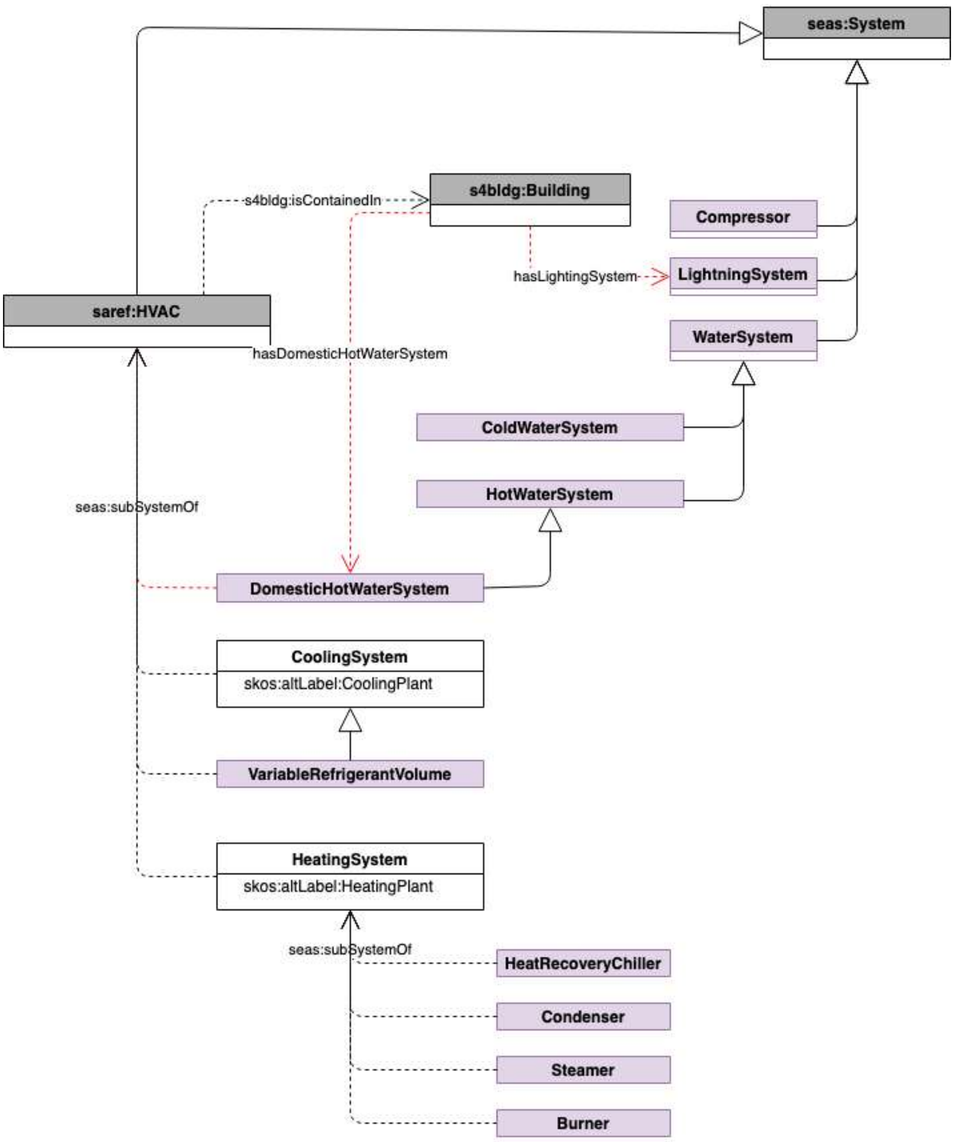


Figure 30 - HVAC System (Extension) - version 2

Figure 30 presents the extension of the representation of the HVAC System with concepts concerning the domestic hot water system (`plt:DomesticHotWaterSystem`) which is a type of `plt:HotWaterSystem` and subsystem of (`seas:subSystemOf`) the HVAC. The building is related to the domestic hot water system by the relation `plt:hasDomesticHot-WaterSystem`. Variable refrigerant volume (`plt:VariableRefrigerantVolume`) is also a subsystem of the HVAC.

The pilot #3c will optimize the local renewable energy resources (RES) and HVAC operation as function of building and RES characteristics, building comfort constraints, ambient conditions and energy market price following a multi-objective pattern which targets to reduce the overall energy bill and maximize the usage of RES. To achieve this, a key concept related to the contract (`seas:Contract`) is discussed in below.

A `seas:Contract` is a container for transactions (see Fig. 31). It is specifically defined in FIBO ontology as “voluntary, deliberate agreement between competent parties to which the parties agree to be legally bound, and for which the parties provide valuable consideration”. Two types of contracts are enumerated: (i) a gas contract (`plt:GasContract`) means the agreement under which gas is sold and delivered by a seller to a buyer, and (ii) an electricity contract (`plt:ElectricityContract`) means the agreement for the purchase of electricity. Each contract is about the following properties `plt:GasEnergyProperty` and `seas:ElectricEnergyProperty`. It is also characterized by a price (`saref:Price`) that can be the selling price (`plt:hasSellingPrice`) and the buying price (`plt:hasBuyingPrice`).

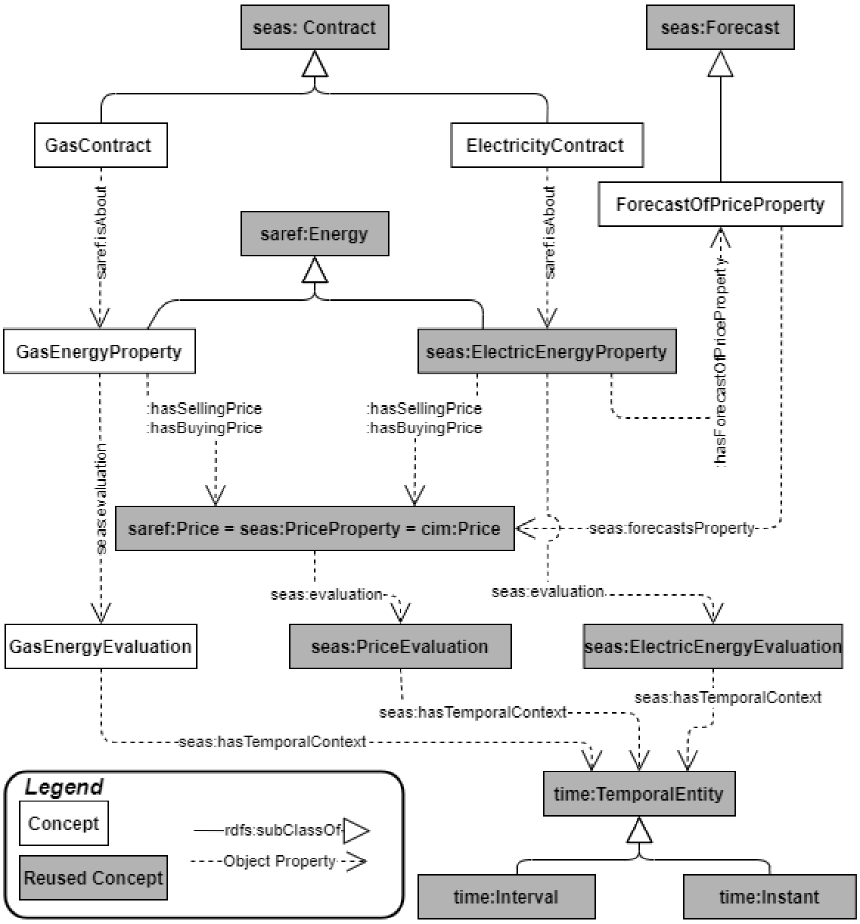
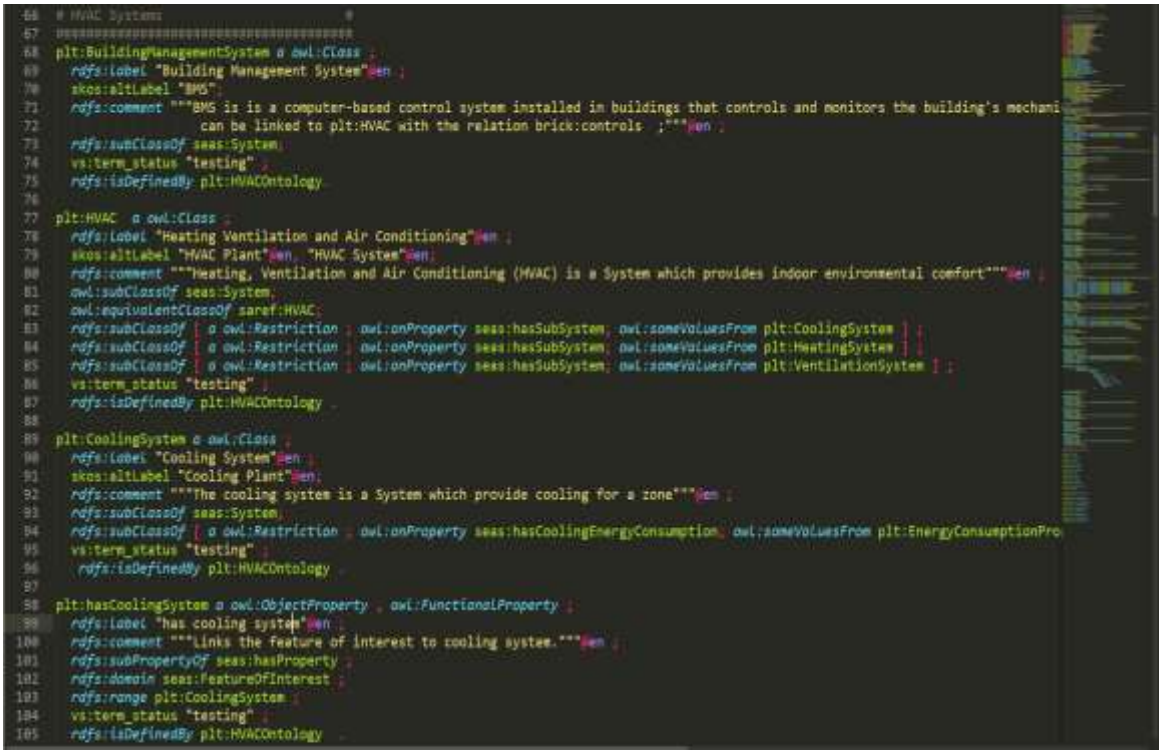


Figure 31: Semantic data model of energy market

3.4 Application of Step 4 – Formalization of semantic data models

Formalization is the last step to do in the proposed methodology. It consists of to formalize, thanks to an ontology web language, the result of Step 3 which is illustrated in some diagrams. Concepts related to a specific domain, are defined in the same module. All resulted modules will constitute the PLATOON ontology. We suggest using the name space https://w3id.org/PLATOON/ as a preliminary URI and `plt` as a prefix. Note that this URI should be validated by all the project partners before to put the ontology online and it does not currently work.

We use the OWL 2 DL language to describe each module, and the TURTLE serialization which is human and machine readable. Figure 32 shows an extract of the formalization of the HVAC ontology module. Each new concept is defined as `owl:Class` (see the line 77) . Its label is defined with `rdfs:label` annotation in the line 78. If the concept can have other labels, `skos:altLabel` is used to define all the alternative labels (see the line 79). The definition associated to the concept is defined with the `rdfs:Comment` annotation in the line 80. The subsumption relation is defined by the `rdfs:subClassOf` relationship. If there is an overlapping with other ontologies we use the relation `owl:equivalentClassOf` to associate to other existing concept like in line 82 where we put `plt:HVAC` as equivalent to `saref:HVAC`. From the line 83 to 85, there are restrictions that are necessary properties to be an HVAC system. That means, every HVAC system has as subsystem a cooling system, a heating system and a ventilation system. The line 86, `vs:term_status` is an annotation to indicate if the concept is on testing phase or already validated. The line 87 is annotation to indicate which ontology module define this concept.



```
66  # HVAC Systems                         #
67  ########################################
68  plt:BuildingManagementSystem a owl:Class ;
69    rdfs:label "Building Management System"@en ;
70    skos:altLabel "BMS";
71    rdfs:comment """BMS is is a computer-based control system installed in buildings that controls and monitors the building's mechani
72                can be linked to plt:HVAC with the relation brick:controls  ;"""@en ;
73    rdfs:subClassOf seas:System;
74    vs:term_status "testing" ;
75    rdfs:isDefinedBy plt:HVACOntology .
76
77  plt:HVAC  a owl:Class ;
78    rdfs:label "Heating Ventilation and Air Conditioning"@en ;
79    skos:altLabel "HVAC Plant"@en, "HVAC System"@en;
80    rdfs:comment """Heating, Ventilation and Air Conditioning (HVAC) is a System which provides indoor environmental comfort"""@en ;
81    owl:subClassOf seas:System;
82    owl:equivalentClassOf saref:HVAC;
83    rdfs:subClassOf [ a owl:Restriction ; owl:onProperty seas:hasSubSystem; owl:someValuesFrom plt:CoolingSystem ] ;
84    rdfs:subClassOf [ a owl:Restriction ; owl:onProperty seas:hasSubSystem; owl:someValuesFrom plt:HeatingSystem ] ;
85    rdfs:subClassOf [ a owl:Restriction ; owl:onProperty seas:hasSubSystem; owl:someValuesFrom plt:VentilationSystem ] ;
86    vs:term_status "testing" ;
87    rdfs:isDefinedBy plt:HVACOntology .
88
89  plt:CoolingSystem a owl:Class ;
90    rdfs:label "Cooling System"@en ;
91    skos:altLabel "Cooling Plant"@en;
92    rdfs:comment """The cooling system is a System which provide cooling for a zone"""@en ;
93    rdfs:subClassOf seas:System;
94    rdfs:subClassOf [ a owl:Restriction ; owl:onProperty seas:hasCoolingEnergyConsumption; owl:someValuesFrom plt:EnergyConsumptionPro
95    vs:term_status "testing" ;
96     rdfs:isDefinedBy plt:HVACOntology .
97
98  plt:hasCoolingSystem a owl:ObjectProperty , owl:FunctionalProperty ;
99    rdfs:label "has cooling system"@en ;
100   rdfs:comment """Links the feature of interest to cooling system."""@en ;
101   rdfs:subPropertyOf seas:hasProperty ;
102   rdfs:domain seas:FeatureOfInterest ;
103   rdfs:range plt:CoolingSystem ;
104   vs:term_status "testing" ;
105   rdfs:isDefinedBy plt:HVACOntology .
```

Figure 32: An extract of the formalization of the HVAC module

During the formalization step, as we notice previously, concept related to the same domain, are defined in the same module. Then we defined different modules that complete the domain covering of existing ontology. Among these modules, we can cite (see details in GitHub[4]):

- **Common ontological modules:**
  - **Meter ontology module** that includes new type of meter such as thermal energy meter
  - **Sensor ontology module** that includes new type of sensor such as presence sensor, wind speed sensor, and voltage sensor. Each sensor is associated to the property it's observes
  - **Weather ontology module** includes all properties related to the atmospheric condition such as radiation, wind direction, and air temperature
  - **Failure and Damage ontology module** includes the minimal concept describe damages and failures. These concepts are extended to represent more specific failure such as Looseness failure and imbalance failure.

[4] https://github.com/PLATOONProject/semantic-data-model/tree/main/Semantic-Data-Models-Formalization

- **Ontological modules of renewable energy:**
  - Wind turbine ontology module that includes wind turbine component such as the generator and power converter and subcomponents such as rotor and stator.

- **Ontological modules of Grids:**
  - Smart Grid ontology module includes the smart grid and the components of this grid such as power generator, charging station and power centre.
- **Ontological modules of Building and HVAC:**
  - Building ontology module that includes the new concepts of building type (e.g., Data Centre and Retail), the occupancy property and the comfort associated to this building.
  - HVAC ontology module that includes all subsystem of the HVAC such as heating system, heating valve and the properties inherent to these systems.

3.5 Use case instantiation with an illustrative example

This section aims to describe how the semantic data model (ontology) can be used for a given dataset.
Table 5 shows an example of dataset (raw data) that related to the temperature in a building.

Five columns are defined:

1. BuildingID: includes all the identifiers of buildings.
2. ZoneID: includes all the identifiers of zones.
3. TempSensor: includes all the names of sensors.
4. Value C°: includes all temperature values in degrees Celsius.
5. Date: includes all dates of temperature measurement.

BuildingID	ZoneID	TempSensor	Value C°	Date
1	1	S1	22	20201008:11h40
1	1	S1	21	20201008:11h50
1	2	S2	20	20201008:11h40
1	2	S2	19	20201008:11h50
2	1	S1	17	20201008:07h00
2	1	S1	20	20201008:08h00

Table 5: Datasets of temperature in a building

Figure 33 shows the result obtained by transforming raw data into semantic data. This example gives ideas for D2.4 (Data Integration) and other tasks that will use semantic data models.

The building of line 1 is identified by this URI `<engie/building/1>` and we type it as `s4bldg:Building` (see orange box). This building contains (`bot:containsZone`) different zones identified as follows `<engie/building/1/zone/1>` and `<engie/building/1/zone/2>`.
Each zone is characterised by (see green box):

- Type (`bot:Zone`)
- Temperature, by this relation (`seas:temperature`) that links with the URI of temperature `<engie/building/1/zone/1/airtemperature/property>`
- Provenance of this data (`prov:atLocation`)

A sensor `<engie/sensor/s1>` is a `saref:TemperatureSensor` (see yellow box). It is contained in (`s4bldg:isContainedIn`) in a zone `<engie/building/1/zone/1>` and measures (`ssn:measures`) the air temperature `<engie/building/1/zone/1/airtemperature/property>`.

The temperature `<engie/building/1/zone/1/airtemperature/property>` is a `plt:AirTemperatureProperty` (see pink box). It is a property of (`saref:isPropertyOf`) this zone `<engie/building/1/zone/1>`.

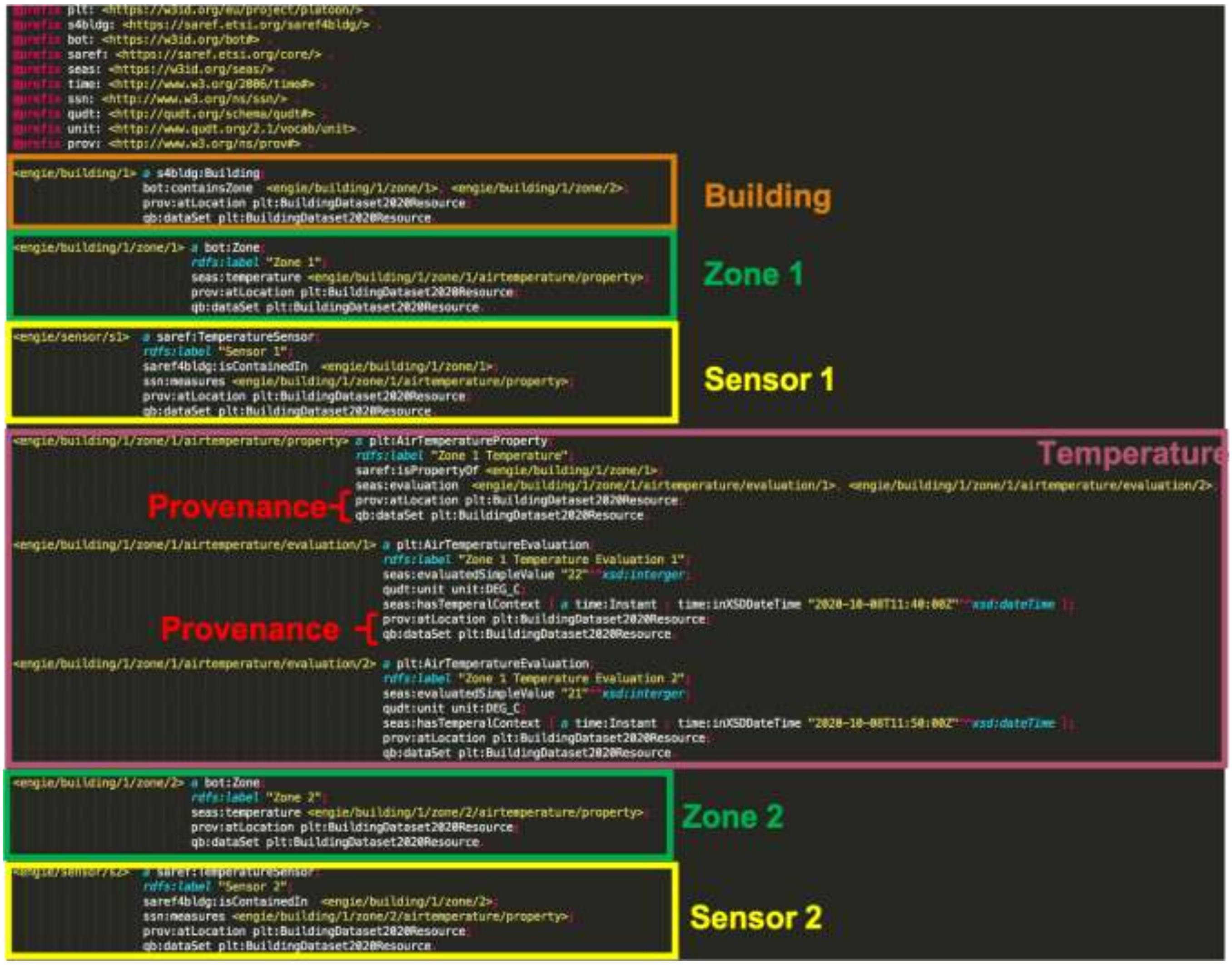


Figure 33: Extract of Temperature in building: RDF description (TTL file)

4 Conclusion

The main objective of this task 2.3 is to provide common semantic data models for 19 LLUCs of 7 Pilots, by applying a specific ontology construction methodology for the PLATOON project. These LLUCs fully address PLATOON's objectives and challenges. Each pilot identifies a set of datasets and data exchange requirements which served as the basis for the construction of the common semantic data models, as described in the section 2. From the descriptions of LLUCs, we identified five topics that covers the pilot's needs. These pilots shared a set of common-sense concepts related to different domains in particular Building and HVAC domain (pilots #3a, #3b and #3c), Renewable energy domain (pilot #1a) and Grids domain (pilots #2a, #2b and #4a). All pilots used also similar notions about the weather, the measurement, forecasting, failures, etc. To achieve a richer interoperability between pilots through the handling of data heterogeneity and to plan to bridge the gap between this task 2.3 and all dependent PLATOON's work packages and tasks such WP2-T2.4, WP4, WP5, etc., we propose a specific methodology to create harmonized semantic data models that include all the needs of the PLATOON project's use cases.

This document puts forward the construction of common semantic data models by reusing existing domain ontologies (SAREF, CIM, SEAS, etc.), as an essential part, or/and constructing ontology modules from the beginning of the ontology design process. We have elicited four major steps required for performing semantic interoperability, after examining the existing ontologies of the energy sector. The step 1 called "ontology requirements specification", allows use cases analysis, ontology scope definition, terms elicitation and competency questions determination. Step 2 named "ontology analysis", aims at analysing the competency questions, and identifying ontology modules by reusing, extending and/or creating an ontology. Step 3 "overview of ontological modules", is about putting all the modules together, evaluating ontologies and instantiating some use cases with an illustrative example. The last step concerns the discussion with stakeholders and the formalization process of the common semantic data models. All four steps are intertwined, and each step provides the necessary input to perform the next step.

We applied the steps of the proposed methodology in different LLUCs of pilots. We presented some fragments of semantic data models (see Section 3) such as Building, HVAC, Wind Turbine, Grid, etc. Finally, we presented an example of formalization of the semantic data model thanks to the ontology web language; and illustrated an example of semantic representation that transform the data in knowledge by using the semantic data model created in the PLATOON project during the two phases of the WP2 Task2.3

The output of this task 2.3 is to provide an OWL file for each new ontological module and an RDF file for each use case. To publish the new modules, we suggest using the URI https://w3id.org/PLATOON to have a secure, permanent URL re-direction service. However, this proposition should be discussed among partners before publishing the semantic data models.

As a perspective on the results of this work in WP 2.3 we can of course cite the use of these ontological SDM modules in the PLATOON project, but the results can also be used as base modules in the ongoing initiatives within GAIA-X and European Data Space and in particular the Energy and BDVA Energy Task Force which will be launched in 2022.

5 References

[1]. https://binaryterms.com/relational-data-model.html

[2]. https://en.wikipedia.org/wiki/Data_model_(GIS)

[3]. European Commission, "Third energy package," 2009. Available: https://ec.europa.eu/energy/topics/markets-and-consumers/market-legislation/third-energy-package_en

[4]. ENTSO-E. Common Grid Model Exchange Standard (CGMES) – Based on IEC Common Information Model, version 2.4, August 2014.

[5]. D. Bourigault et J. Charlet (2005). Construction d'un index thématique de l'ingénrierie des connaissances. In Conférence d'Ingénierie de Connaissances, pages 29-47, De Régine Teulier.

[6]. D. Cruse (1986). Lexical Semantics. Cambridge University Press, UK.

[7]. Douglas B. Lenat and R. V. Guha. (1989). Building Large Knowledge-Based Systems; Representation and Inference in the Cyc Project (1st. ed.). Addison-Wesley Longman Publishing Co., Inc., USA.

[8]. Embley D.W. (2009) Semantic Data Model. In: LIU L., ÖZSU M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA.

[9]. G. van Heijst, A.Th. Schreiber, B.J. Wielinga, Using explicit ontologies in KBS development, In International Journal of Human-Computer Studies, Volume 46, Issues 2–3, 1997, p183-292.

[10] Gómez-Pérez, A, Fernández-López, M and de Vicente, A, 1996, "Towards a method to conceptualize domain ontologies" Workshop on Ontological Engineering. European Conference of Artificial Intelligence (ECAI), pages: 41–52.

[11] Gruber, T., 1993, A translation approach to portable ontology specifications, Knowledge acquisition. Elsevier, Vol. 5, n°2, pp. 199-220

[12] Enterprise Engineering. In Workshop on Implemented Ontologies, European Conference on Artificial Intelligence (ECAI).

[13] Jacquemin, Bernard & Ploux, Sabine. (2008). Corpus spécialisé et ressource de spécialité : l'information forme le sens. In Actes des Journées Scientifiques du CRTT: Corpus et dictionnaires de langues de spécialité, Lyon.

[14] Kefi, Hassen, Reynaud, Chantal and Safar, Brigitte. (2006). Techniques structurelles pour l'alignement de taxonomies sur le Web. In Atelier Fouille du Web – EGC 2006.

[15] Lassila, O. & Mcguinness, Deborah. (2001). The role of frame-based representation on the Semantic Web. Electronic transactions on Artificial Intelligence (ETAI) Journal. 6.

[16] Maricela Bravo, Luis F. H. Reyes, José A. R. Ortiz. (2019). Methodology for ontology design and construction, Contaduría y Adm., vol. 0, no. 0, pp.1–24, 2019.

[17] McMorran, A.W.: An introduction to IEC 61970–301 & 61968–11: The common information model. University of Strathclyde, 93, 124 (2007)

[18] Pascal Hitzler, Adila Krisnadhi. (2018). A Tutorial on Modular Ontology Modeling with Ontology Design Patterns: The Cooking Recipes Ontology. CoRR abs/1808.08433 (2018).

[19] Philippe Lefèvre (2000). La recherche d'informations. Du texte intégral au thésaurus. Hermes Science.

[20] Sarra Ben Abbès, Building a domain cartography from heterogeneous semantic resources, Paris 13 University, France, 2013.

[21] Schreiber, Guus & Wielinga, Bob & Jansweijer, Wouter. (2003). The KACTUS View on the 'O' Word. IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing: Proceedings of the 7th Dutch National Conference on Artificial Intelligence NAIC'95 159–168.

[22] Sure, Y., Staab, S., Studer, R.: Methodology for development and employment of ontology-based knowledge management applications. SIGMOD Record 34 (2002) 8–24.

[23] Swartout, William & Patil, Ramesh & Knight, Kevin & Russ, Tom. (1997). Toward Distributed Use of Large-Scale Ontologies. Symposium on Ontological Engineering of AAAI 138–148.

[24] Uschold, Michael & Grüninger, Michael. (1996). Ontologies: Principles, methods, and applications. The Knowledge Engineering Review. 11(2), 93-136. https://doi.org/10.1017/S0269888900007797

[25] https://w3id.org/def/saref4city

[26] https://w3id.org/def/saref4bldg

[27] https://w3id.org/def/saref4ener/

[28] https://w3id.org/def/saref4agri/

[29] https://w3id.org/def/saref4envi/

[30] https://w3id.org/seas

[31] http://www.iec.ch/dyn/www/f?p=103:14:0::::FSP_ORG_ID:2392

[32] http://www.iec.ch/dyn/www/f?p=103:14:0::::FSP_ORG_ID:2388

[33] https://www.isko.org/cyclo/interoperability.htm#2

[34] Michael R. Genesereth and Nils J. Nilsson. Logical Foundation of Artificial Intelligence, 1987.

[35] https://www.iso.org/fr/standard/7776.html

[36] https://www.niso.org/publications/ansiniso-z3919-2005-r2010