PEDvolution

Interoperable solutions to streamline PED evolution and cross-sectoral integration

Deliverable 7.1

Data Exchange Platform Design

Co-funded by the European Union

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement Nº 101138472.

The Swiss project participants received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

Document Summary Information

Grant Agreement No	**101138472**	Acronym	**PEDvolution**
Full Title	**Interoperable solutions to streamline PED evolution and cross-sectoral integration**		
Start Date	**01/01/2024**	Duration	**36 months**
Project URL	**https://www.pedvolution.eu/**		
Deliverable	**Data exchange platform design**		
Work Package	**WP7**		
Contractual due date	**31/12/2024**	Actual submission date	**26/12/2024**
Nature	**R — Document, report**	Dissemination Level	**PU**
Lead Beneficiary	**ICOM**		
Responsible Author	**Ilia Petri (ICOM)**		
Contributions from	**Nikolaos Charitos (ICOM)** **Marios Charalampidis (ICOM)** **Maria Tsaroucha (ICOM)** **Igyso Zafeiratou (ICOM)** **Isidoros Kokos (ICOM)**		
Reviewed by	**Martin Vodnik (OFFSET)** **Sašo Brus (OFFSET)** **Markos Psimitis (INLE)** **Makis Kouloumbis (INLE)** **Reda El Makroum (TUW)**		

Copyright message

Disclaimer

Revision history (including peer reviewing & quality control)

VERSION	ISSUE DATE	COMPLETE[1]	CHANGES	CONTRIBUTOR(S)
v0.1	22/10/24	0%	Initial Deliverable Structure (TOC)	ICOM
V0.2	29/11/24	30%	Chapters 1,2,5	ICOM
V0.3	02/12/24	80%	Chapters 3,4	ICOM
V0.4	05/12/24	95%	Chapters 6, Executive Summary, Glossary Deliverable Ready for internal Review.	ICOM
V0.5	16/12/24	96%	Review comments.	OFFSET, INLE
V0.6	22/12/24	97%	Quality control check.	TUW
V0.7	23/12/24	99%	Address review and quality control comments.	ICOM
V1.0	26/12/24	100%	Final Version for submission.	ICOM

1 According to Project's Quality Assurance Process

Table of Contents

List of Figures

List of Tables

Glossary of terms and abbreviations used

ABBREVIATION / TERM	DESCRIPTION
CHP	Combined Heat and Power
CIM	Common Information Model
CSDM	Common Service Data Model
CSV	comma-separated values
DHC	District Heating and Cooling
DS	Data Space
DSP	Data Space Protocol
EDC	Eclipse Data Space Connector
EnergyADE	Energy Application Domain Extension
EPC	Energy Performance Certificate
EV	Electric Vehicle
FQDN	Fully Qualified Domain Name
HLUC	High-Level Use Case
HVAC	Heating, Ventilation, and Air Conditioning
IDS	International Data Space
IDSA	International Data Spaces Association
JSON	JavaScript Object Notation
LCC	Life-cycle costs
OWL	Web Ontology Language
PED	Positive Energy District
PV	Photovoltaic
R&I	Research and Innovation

RAM	Reference Architecture Model
RDF	Resource Description Framework
SAREF	Smart Applications REFerence
SAREF4BLDG	SAREF extension for building
SAREF4ENER	SAREF extension for Energy Flexibility
SAREF4CITY	SAREF extension for Smart City
VCs	Verifiable Credentials
XML	Extensible Markup Language

EXECUTIVE SUMMARY

PEDvolution aims to develop a cross-sectorial data exchange and interoperability platform that enables the integration of the different platforms, systems and assets in the PED ecosystems. This deliverable reports on the work produced during the first phase of T7.1 "Integration of Smart Homes and Buildings and Grid Infrastructure" which mainly includes two activities: (i) the design, development and deployment of the Interoperability Platform, (ii) the analysis of existing ontologies and state-of-the-art information models towards the building of a common information model for the abstraction of information with the PED ecosystem.

The Interoperability Platform of PEDvolution follows the EU Data Spaces paradigm, aiming towards increased interoperability, security and self-sovereignty data sharing among various stakeholders and solutions. The Data Space solution incorporates all the required security and privacy mechanisms while allowing the implementation of data usage and market policies to support authorised access to data and controlled exposure to third parties.

A bottom-up approach was followed for building a semantic data model for PEDvolution project, initiating from the data needs of the different solutions of the project. To define a Common Information Model (CIM) for the PEDvolution project, a comprehensive analysis of state-of-the-art ontologies and standardised information models was conducted. This included examining frameworks such as SAREF and its extensions (SAREF4BLDG, SAREF4CITY, SAREF4ENER), OMEGA-X CSDM, Energy ADE for CityGML, BOT, and FLEXOFFER. The analysis revealed that while these ontologies collectively address a significant portion of PEDvolution's data requirements, they exhibit partial overlaps and require extensions to fully meet the specific needs of all PEDvolution pilot projects.

To address these gaps, online workshops were held with PED partners and solution providers to identify and document all data exchanges facilitated through the Interoperability Platform. These workshops informed the development of PEDvolution's CIM, which encompasses diverse categories of infrastructure and assets such as batteries, HVAC systems, district heating and cooling (DHC) networks, and renewable energy systems like solar panels. The CIM also integrates characteristics at both the building and district levels and accounts for the multifaceted nature of Positive Energy Districts (PEDs), considering their social, technical, and economic dimensions.

The work presented in this report comprises the basis for the integration of the different systems, platforms and assets in PEDvolution in the subsequent WP activities. The Interoperability Platform will be enriched with mechanisms for the interfacing with the energy market and federation services for the interconnection to the wider EU Data Space ecosystem in T7.2 and T7.4 respectively, whereas the CIM will be finalised addressing the final needs of the solution providers (T7.3) and the pilots (T7.1). The final solution will support the pilot activities in the context of WP8 and WP9 by providing a cloud-based platform that will enable data exchanges between all PEDvolution pilots and solutions, according to the project's requirements.

1 INTRODUCTION

This deliverable outlines the integration and interoperability approach to be followed in PEDvolution project towards enabling smooth integration of the various systems, platforms and devices in PED ecosystems. The report presents, the methodology followed for the definition of PEDvolution's Common Information Model with which all different solutions and platforms need to conform, as well as the design of the Interoperability Platform to be used for secure data sharing, following the requirements analysis and system architecture defined in D1.3.

1.1 Mapping Project's Outputs

This section outlines the PEDvolution project's commitments as specified in the Grant Agreement, describing how the formal deliverable and task descriptions align with the outputs of the "Data exchange platform design" deliverable and the performed work. By systematically cross-referencing each commitment with the corresponding activities and results, we provide an overview of how the deliverable adheres to the project's objectives. This includes the definitions of the common information model and the design of the PEDvolution Interoperability Platform.

Table 1: Adherence to Project's GA Deliverable & Tasks Descriptions.

PROJECT GA COMPONENT TITLE	PROJECT GA COMPONENT OUTLINE	RESPECTIVE DOCUMENT CHAPTER(S)	JUSTIFICATION
DELIVERABLE			
D7.1	Integration of the different platforms/systems/devices of the PED ecosystems via a cross-sectorial data exchange platform. Provision of abstraction layer of assets in the PED, enabling transparent data exchanges and enabling system's extensibility. Analysis of standard information models (e.g., SAREF, IEC CIM) for the creation of a common information.	Chapters 2-5	Presents the integration and interoperability approach to be followed in PEDvolution. It provides documentation on the architectural design of the Interoperability Platform and the mechanisms employed for enabling secure data exchange as well as the methodology for the provisioning of the abstraction layer of assets in the PED to achieve semantic interoperability.
TASKS			
T7.1	Task 7.1 Integration of Smart Homes and Buildings and Grid Infrastructure will provide an abstraction layer of assets in the PED, enabling transparent	Chapters 2-5	Initiating from the context view of D1.3, data needs were analysed to build a Common Information Model within PEDvolution which fully covers the project's requirements and relevant PED

	data exchanges and enabling system's extensibility (D7.1).		aspects. Also, the technical specifications of the system architecture in D1.3 are further analysed to design the Interoperability Platform and specify the integration approach to be followed.

1.2 Deliverable Overview and Report Structure

The structure of the deliverable is organised in five main chapters, each focusing on different interoperability and integration aspects. From the interoperability approach with the design of the data exchange platform to the semantic analysis and the building of a common semantic information model. More specifically, the deliverable is structured as follows:

- Chapter 2 outlines the interoperability methodology used for the seamless integration of the different platforms, systems, and assets in PEDs through the design of the data exchange platform and the use of common models for semantic and syntactic interoperability.
- Chapter 3 focuses on the description of the interoperability platform designed based on a Data Space framework, describing its architecture aspects and outlining its core services and their functionalities.
- Chapter 4 presents the semantic interoperability approach with the analysis of PEDvolution's solutions data needs, the investigation of the relevant standardised ontologies and existing data models and the formation of PEDvolution's information model.
- Finally, Chapter 5 concludes the report outlining the main outcomes and identifying the next steps for the successful integration of services and platforms in PED ecosystems.

The integration and interoperability approach and specifications reported in this deliverable aim to facilitate the alignment of all the different solutions developed in the implementation WPs (WP3, 4 and 5) as well as the interactions between the different stakeholders and data access from PED assets. The design of interoperable interfaces for the integration of PEDvolution solutions will continue at T7.3 and the final data exchange platform prototype with the integrated PEDvolution solutions, the incorporated market mechanisms (T7.2) and the federation services for the exposure of PEDvolution data and services outside the project (T7.4) will be reported in the final WP7 deliverable D7.3 "Data Exchange Platform".

2 METHODOLOGY

One of the objectives of PEDvolution project (O4) [1] is to design, develop and deploy an Interoperability Platform incorporating all necessary mechanisms for secure data exchange between different solutions, platforms and devices within the PED ecosystem as well as with the external environment. The design of the Interoperability Platform of PEDvolution follows the EU Data Spaces paradigm [2] aimed at increased interoperability and secure/seamless/self-sovereign data sharing, hence a brief introduction of the key initiatives in the domain is presented in the chapter along with the methodology followed for the design, building upon existing developed tools, methodologies and standards.

Conversely, another critical dimension of interoperability lies at the syntactic and semantic levels of data. By leveraging open data standards (e.g., SAREF) and protocols, the project seeks to promote interoperable data exchanges and ensure compatibility with solutions developed in ongoing EU initiatives and R&I projects. This chapter further outlines the methodology employed to address data interoperability challenges, aiming to facilitate unified data access for PED platforms in alignment with established methodologies and standards.

2.1 EU Data Spaces Paradigm

The European Union's Data Spaces initiative [2] is a strategic framework designed to promote the secure, seamless, and interoperable exchange of data across sectors and borders within the EU, providing a unified digital ecosystem where businesses, governments, and individuals can access and share data under clear rules and trust mechanisms, in respect to the EU Data Act. Data Spaces are sector-specific, targeting areas such as health, energy, agriculture, and finance, and are foundational to the EU's digital and green transitions.

2.1.1 Key Initiatives

International Data Spaces

The International Data Spaces (IDS) [3] is an association that has data sovereignty as its core goal. International Data Spaces Association (IDSA) provides a Data Space framework to enable secure, standardised data exchange and integration within a trusted ecosystem, the Reference Architecture Model (IDSA-RAM) [4]. The IDS-RAM serves as a comprehensive framework designed to facilitate secure and trusted data exchange across organisational boundaries. It is built with an open and technology-agnostic approach, ensuring compatibility and flexibility with diverse technologies and systems. Furthermore, its universal applicability makes it a valuable tool across various industries and domains, enabling the seamless integration of data-sharing practices regardless of specific technological constraints or sector-specific requirements. IDS-RAM enables a comprehensive description of data assets and interoperability required for seamless data exchange. According to IDS-RAM, the distributed network that enables the connection of different participants through IDS Connectors is separated into the following layers:

- <u>Business Layer</u>: Specifies and categorises the different roles of the participants, their main activities and interactions.
- <u>Functional Layer</u>: Specifies the functional requirements of the DS & concrete features to be derived from these.

- Process Layer: Describes the interactions taking place among the different components.
- Information Layer: Defines a conceptual model describing both static and dynamic aspects of IDS's constituents.
- System Layer: Decomposes the software components of the DS, specifying integration, deployment, and extensibility aspects (amongst others).

IDSA has fostered the development of the **Data Space Protocol** (DSP) [5], a universal standard that defines the schemas, protocols, and procedures for publishing, negotiating, and accessing data across diverse platforms and systems within a Data Space to facilitate seamless data sharing and ensure technical interoperability between the different solutions.

Gaia-X

Gaia-X is a European initiative aimed at creating a federated, secure, and transparent data infrastructure to promote digital **sovereignty** and innovation. Gaia-X framework [6] establishes an interconnected ecosystem comprising an infrastructure layer and a data ecosystem using **federation services**. Key enablers of Gaia-X architecture are Verifiable Credentials (VCs). VCs are used to establish trust, verify compliance, and enable secure interactions within the federated ecosystem. VCs are cryptographically secure proofs issued by trusted authorities, validating participants' identities, certifications, and adherence to Gaia-X standards without relying on centralised systems. Gaia-X provides a set of open-source tools enabling the creation and operation of a DS, the Digital Clearing House [5].

EDC

Eclipse Data Space Components (EDC) [6] are open-source tools designed to enable the secure and interoperable exchange of data in Data Spaces. They provide essential building blocks, such as **data connectors** for secure data transfer, **policy enforcement** for managing data usage rights, and **federation services** to support interoperability between participants. EDC aligns with standards like Gaia-X to ensure trust, compliance, and scalability in decentralised data ecosystems.

EDC provides a modular architecture allowing its core runtime to be extended through specific modules. These modules provide the flexibility to support various protocols, enforce diverse policy rules, and interact with different catalogue systems, making the platform highly adaptable to a wide range of use cases. Leveraging the IDS protocol, EDC ensures trusted data exchanges by embedding mechanisms for identity management, data sovereignty, and policy enforcement. The Control Plane oversees decisions regarding data management, routing, and processing. This includes tasks such as identity management and enforcement of access and usage policies. On the other hand, the Data Plane is responsible for the actual transfer of data, ensuring efficient and reliable exchange across systems while maintaining compliance with data usage agreements. The Control Plane and Data Plane are separated to enhance extensibility and support the modular integration of various protocols.

2.1.2 Design Methodology

The approach of PEDvolution is to combine different initiatives, following the paradigm of other EU projects (i.e., OMEGA-X [7]), enabling the creation of a DS for the PED ecosystem. Leveraging existing open-source solutions, such as the Digital Clearing House of Gaia-X and the Data Space Protocol of EDC, the design of the Interoperability Platform will provide a robust and scalable mediation layer for secure, seamless, and standardised data exchange. By integrating these solutions, the platform will facilitate trusted collaboration between stakeholders, streamline data access, and support the efficient exchange of information within and across various Data Spaces.

In terms of documenting the design of the solution, the paradigm of viewpoint architecture, presented in D1.3 [8], is used complementing the high-level architecture of PEDvolution project. Towards enabling a granular view of the design of the Interoperability Platform, the functional and process viewpoints will be presented. The former provides a decomposition to individual building blocks and their main interrelations, whereas the latter will outline workflows, interactions, and data flows.

2.1.3 Alignment with PEDvolution Requirements

The selection of the Data Spaces paradigm for the system architecture of the Interoperability Platform is justified by considering the advantages of a Data Space compared to Middleware Platforms, as well as the individual technical requirements that should be met by the various PEDvolution solutions, which will be interconnected via the Interoperability Platform. The key categories of non-functional technical requirements are interoperability, performance, security, and privacy [9]. The key advantages of a Data Space architecture are summarised below, while references to related requirements categories are made:

1. Decentralisation: The Data Space architecture's decentralised approach is ideal for PEDvolution's system architecture as it allows data to remain at its source while enabling secure and controlled sharing among participants. This decentralisation aligns with the performance requirement of ensuring timely and efficient data exchanges. By avoiding central bottlenecks, the architecture facilitates real-time or periodic updates based on business needs, minimising latency. Furthermore, it supports security and privacy by removing the reliance on a centralised repository, reducing the risk of data breaches, and enabling direct implementation of sharing policies at the source.
2. Data Sovereignty: A core advantage of a Data Space is its ability to ensure data sovereignty, where data providers retain full control over their data. This is critical for meeting PEDvolution's security and privacy requirements, as providers can implement granular sharing policies, ensure compliance with GDPR [10] for EU and Federal Act on Data Protection [11] for Switzerland, and adopt techniques like anonymisation and pseudonymisation to protect sensitive user data. It also fosters trust among participants, which is vital when integrating diverse stakeholders and pilots in the project.
3. Interoperability: Data Spaces are designed to facilitate interoperability through standardised data exchange protocols and semantic models. This advantage directly addresses PEDvolution's requirement for compliance with common standards, information models, and protocols, such as SAREF, CSDM, and FlexOffer, which will be adopted by the Interoperability Platform to secure interoperability. By adhering to open standards, Data Spaces provide a flexible and modular architecture that simplifies integration with diverse pilot systems while reducing the risk of vendor lock-in.
4. Scalability and Flexibility: The federated nature of a Data Space makes it inherently scalable and adaptable to evolving project needs. This scalability ensures that new participants or data sources can be integrated seamlessly, which is crucial for PEDvolution as its pilots and solutions expand. Moreover, its flexible mechanisms for periodic and ad-hoc data exchanges align perfectly with the performance requirements, ensuring operations are completed within required time thresholds.

5. Trust and Security: A Data Space employs robust mechanisms for trust and security, such as federated identity management, encryption, and secure communication protocols. These features directly support PEDvolution's security and privacy requirements, safeguarding both stored and transmitted data while ensuring compliance with GDPR. By enabling fine-grained access control and secure data exchanges, a Data Space architecture mitigates risks and provides a solid foundation for collaboration among stakeholders.
6. Vendor Independence: The Data Space architecture promotes vendor independence by adopting the use of open-source code, such as EDC (cf. Section 2.1.1). This approach reduces reliance on proprietary technologies and ensures that the Interoperability Platform remains adaptable to evolving technological requirements even beyond the project's lifetime and can integrate with a wide range of tools and systems without encountering compatibility or licensing barriers.

By aligning the advantages of a Data Space with PEDvolution's requirements for interoperability, performance, security, and privacy, this system architecture proves to be the most suitable and future-proof choice for the project.

2.2 Data Interoperability

2.2.1 Common Models for Semantic and Syntactic Interoperability

For systems to interoperate and have a common understanding, different levels of interoperability are required. Syntactic interoperability ensures a common language is used i.e., same structures and symbols among the different parties. On the other hand, semantic interoperability ensures that the parties exchanging the data have a common agreement on the precise meaning of the syntax and symbols, enabling unambiguous communication.

To ensure a shared understanding of the PED ecosystem, a Common Information Model (CIM) - model that provides a shared structure and vocabulary for describing information within a specific domain – is built, which directly supports syntactic and semantic interoperability by providing a unified framework for data exchange and interpretation. In terms of semantic interoperability, the CIM defines the meaning of data elements and their relationships and establishes a shared vocabulary and ontology for understanding. Towards this, PEDvolution provides a Semantic Data Model (presented in Chapter 4), a conceptual model which encompasses semantic description of the data exchanged, defining the meaning and relationships between them. In terms of syntactic interoperability, the CIM defines the supported structures, providing a shared schema and format for data exchange and ensures that all data exchanges adhere to a common syntax for data representation. Syntactic interoperability is addressed from the adopted Data Space framework leveraging common formats and widely accepted standards (such as JSON, XML and CSV) for data structuring, while tools and connectors within the Data Space framework validate compliance with syntactic rules during data exchange. Adherence to standard syntactic conventions remove integration complexity and enhances the scalability in the Data Space ecosystem.

2.2.2 Principles

Following the example of OMEGA-X Energy Data Space Horizon Europe project [7], as set of principles was adopted aiming at increased interoperability and reusability of PEDvolution models:

- **Standards alignment**: Ensure alignment with domain standard models and/or reference models in the domain.
- **Reusability**: Reuse existing concepts (classes, attributes, etc.) published in selected semantic data models.
- **Modularity**: Adopt a loosely coupled approach with smaller ontologies, each on focusing on a certain aspect of the domain, enabling the effective combination of different ontologies towards the creation of application specific and domain specific profiles.
- **FAIR**: Support Findable, Accessible, Interoperable and Reusable (FAIR) principles for data exchanged in PEDvolution, as specified also in PEDvolution's data management plan [12].

2.2.3 Approach

A bottom-up approach was followed in order to identify the different datasets, their properties and the semantic meaning for the different stakeholders in the project as elaborated in the next steps:

- Identifying the data needs: The data needs of the different solution providers have been investigated through workshop sessions and following the data requirements analysis of WP1 core data categories and the properties describing them have been identified. The PEDvolution data needs are documented in Section 4.2.
- Analysis of state-of-the-art: Existing ontologies, data models and standards have been analysed to select relevant ontologies, and gap analysis has been performed to discover areas of extension. An analysis of the examined standards, ontologies, and data models is provided in Section 4.3.
- Convergence to common semantics: The common semantic model is built extending the selected ontologies with missing information and cover PED conceptualisation, implementation and development according to PEDvolution needs. More information on PEDvolution's common semantic model can be found in Section 4.4.

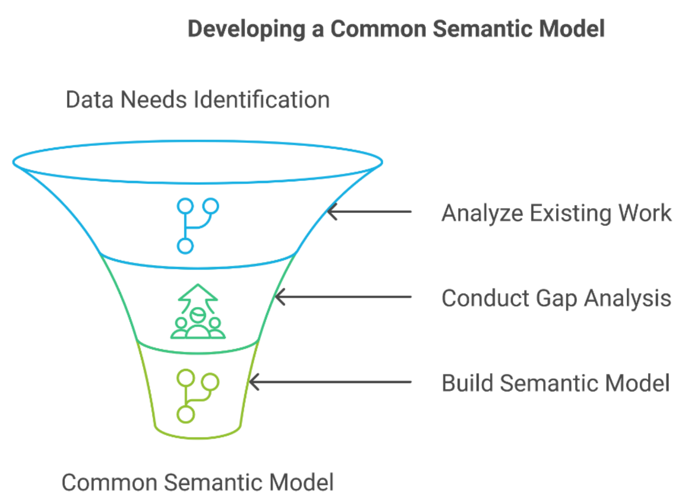

Figure 1: Semantic harmonisation methodology.

The results of the application of the methodology in PEDvolution context, is presented in Chapter 4.

3 PEDVOLUTION'S INTEROPERABILITY PLATFORM

This chapter documents the design of a data exchange platform for PEDvolution (i.e., the Interoperability Platform) that will be used for the communication between PEDvolution's solutions, platforms, and assets in the PED ecosystems. The positioning of the Interoperability Platform in the context of PEDvolution is documented in D1.3 [8]. The purpose of this chapter is to provide the reader with more details on reference architectures and technologies for Data Spaces, as well as dive deeper into the system architecture of the Interoperability Platform by elaborating on its functional and process layers.

The Interoperability Platform incorporates mechanisms that enable:

(i) Seamless integration of the different parties to securely access and transfer data.
(ii) Connection to the wider EU Data Space ecosystem to expose any data and services externally.
(iii) Interfacing with the energy market.
(iv) Compliance with security standards and appropriate privacy policies to safeguard access to/exchange of sensitive information.

To achieve these goals, the design of the Interoperability Platform, follows the EU Data Spaces (DS) paradigm [13], providing a platform that allows different parties to access and collaborate on data within the context of PEDs, in conformance to the EU Data Act [14]. The DS solution will provide all necessary mechanisms for semantic, syntactic and technical interoperability to ensure secure data exchanges between the various systems, platforms, and assets in the context of the PED demonstrators and the PEDvolution solutions assisting stakeholders in PED planning and operation. The DS is built upon the Eclipse Data Space Components (EDC) framework [15] which implements the IDS Data Space Protocol (DSP) [16] following the framework of Gaia-X [6], to support decentralised data exchange in a secure, sovereign, and standards-based way.

3.1 Data Spaces Reference Architectures and Technologies

The development of the Interoperability Platform will be based on established architectures and technologies on Data Spaces, such as the IDSA-RAM and Gaia-X Framework. The IDSA-RAM, developed by the IDSA, is an open, technology-neutral framework for secure data exchange. Its versatile design ensures compatibility across industries and supports seamless integration regardless of specific technologies or domains. The Gaia-X Framework is founded on decentralised data services enabling trust, interoperable data sets and services across sectors, and anchored contract rules governing data access and usage.

3.1.1 IDSA RAM

The reference architecture of IDS architecture [4], is built around several key components to ensure seamless, secure, and efficient operations, as described below.

- **Identity Provider**: Ensures secure and trusted identification of participants within the International Data Spaces (IDS) ecosystem, enabling authentication and authorisation.
- **IDS Connector**: Acts as a gateway for secure data exchange between participants, ensuring data sovereignty and compliance with IDS standards.
- **App Store and Data Apps**: Hosts and distributes certified applications (Data Apps) that provide specialised data processing, analytics, and transformation capabilities, enhancing the functionality of the IDS Connector.
- **Metadata Broker**: Maintains a registry of available data and services within the IDS, enabling discovery and matchmaking between data providers and consumers.
- **Clearing House:** Provides an auditing mechanism for transactions within the IDS, ensuring transparency and accountability by logging data exchanges and interactions.
- **Vocabulary Hub:** Centralises and standardises vocabularies, ontologies, and taxonomies to ensure semantic interoperability across participants in the ecosystem.

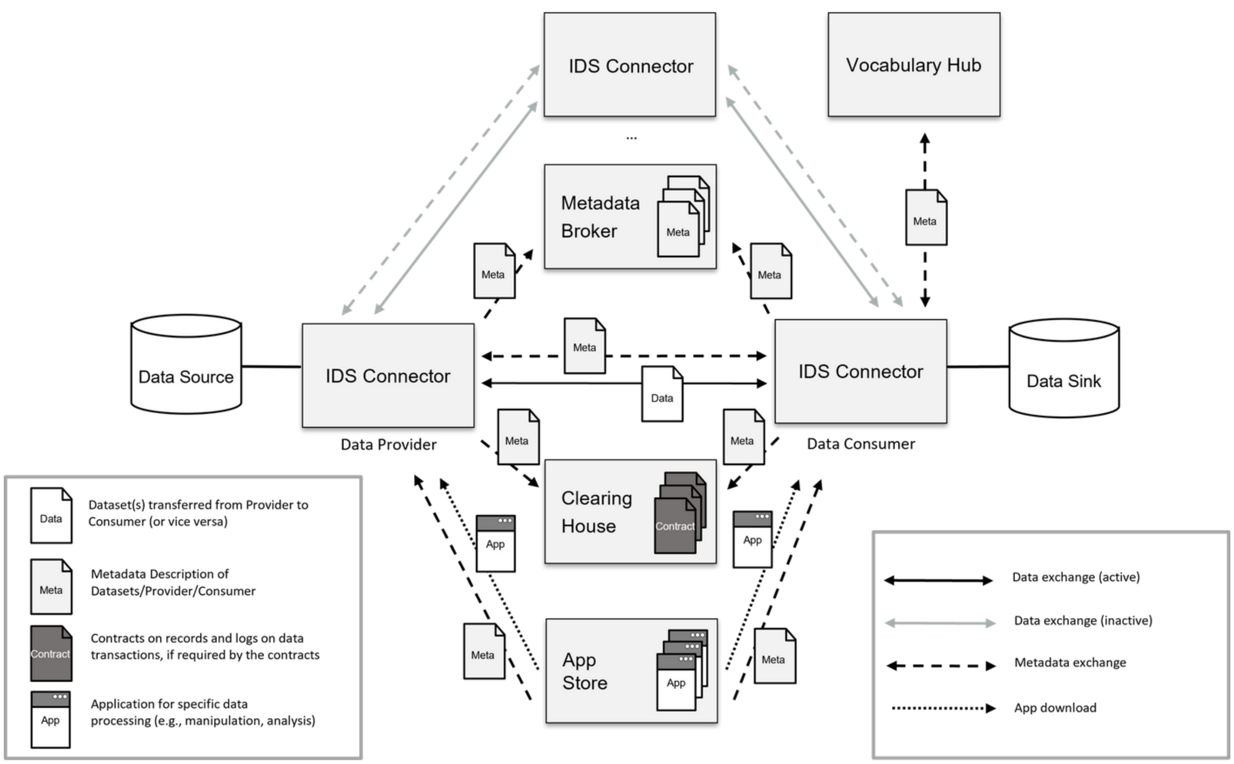

Figure 2: Data Space System Layer by IDSA [4].

3.1.2 Gaia-X Framework

The Gaia-X framework [6] covers three pillars:

- **Compliance**: for a common digital governance based on European values.
- **Federation**: enables interoperable & portable (Cross-) Sector data sets and services.
- **Data exchange**: A mean to perform data exchange and anchor data contract negotiation results into the infrastructure.

Key components of this framework concern:

- **Identity and Trust Services**: Ensure participant authentication, authorisation, and trust through a unified identity framework (notary, registry).
- **Compliance Services**: Verify adherence to Gaia-X principles, including security, interoperability, and European regulatory standards.
- **Data Exchange Services**: Enabling the authentication, policy negotiation and contraction as well as cataloguing mechanisms for data and services. It may also provide additional functionalities such as for data offerings as well as logging and auditing functionalities.
- **Federated Catalogue**: A searchable directory of participants, services, and datasets, enabling discovery and collaboration.
- **Support Tools**: Facilitate seamless integration across various technologies, providers, and sectors (e.g., Verifiable Credential Wizard, Credentials Manager).

3.2 Interoperability Platform Design

The design of the Interoperability Platform is elaborated in this section by documenting its functional and process layers. The functional layer presents the key components of the DS, while the process layer describes the workflows enabling data exchanges between two parties in the context of PEDvolution.

3.2.1 Functional Layer

The functional layer of the Interoperability Platform, as depicted in the figure below, encompasses the key components that enable seamless, secure, and efficient data exchanges within the PEDvolution ecosystem. These components form the backbone of the DS solution, ensuring interoperability between PEDvolution solutions, pilot platforms, and external stakeholders.

- Data Exchange Services include Data & Services connectors required for data access and sharing from PEDvolution solutions or the data platforms. Data exchanges take place in a decentralised way directly between the participating Data & Services Connectors. Any PEDvolution solution may act as a data consumer for accessing and processing data coming from other PEDvolution solutions or PED platforms and assets or may have the role of a data provider exposing their outputs to other PEDvolution solutions or externally. Integrated policy management services in the data exchange services will enable selective access to datasets, allowing data to be shared only with authorised participants. Also, privacy policies will be applied to protect sensitive information, ensuring compliance with relevant regulations and standards.
- The Catalogue will serve as the central repository of the Data Space, offering an overview of available data and services. The DS solution will integrate federation services allowing the data

to be exposed beyond the PEDvolution solutions to the broader Data Space community, as part of efforts detailed in Task T7.4.

- The Identity Management Service will ensure secure and efficient management of participants' identities within the Data Space through a centralised token provider which validates the participants' certificates to ensure that only trusted entities are granted access and issues secure tokens for managing access control.
- A telemetry service will be utilised for tracking and monitoring transactions within the Data Space and log traces from Data & Services Connectors.
- Finally, a Vocabulary Hub can serve semantic interoperability within PEDvolution by hosting its common information model built to ensure that DS participants adhere to the unified vocabulary for data description and exchange.

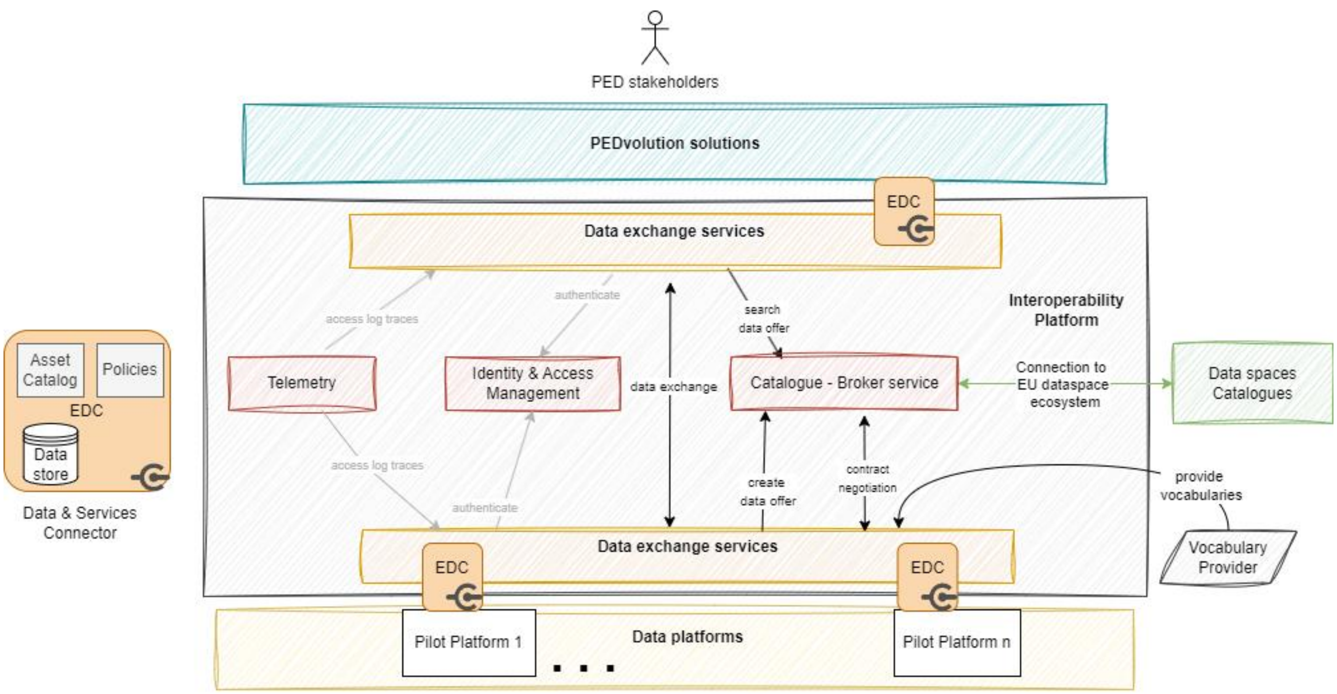

Figure 3: Functional layer of Interoperability platform.

3.2.2 Process Layer

In this section the workflows that facilitate data exchange between two participating parties in PEDvolution through the DS solution are outlined. The data exchange process, along with the associated actions required from both parties, is depicted in Figure 4 below.

The data exchange process begins with the Data Provider, who prepares the dataset to be shared by uploading the data to a data source (different types of data source(s) can be supported) and then creating its metadata to their Data connector. These elements define the policies under which this dataset can be accessed by potential consumers, and they are essential for describing the dataset's attributes. The dataset is then made available by creating the corresponding data offer (also referred to as a data contract) to their Data connector. The data offer is then published to the Catalog Service (a Broker) for making it discoverable by potential data consumers.

On the other hand, a Data consumer may access a provider's shared datasets following the procedure described next. In a typical scenario, a data consumer can search for an appropriate data provider - a Connector registered to the Catalogue Service (the Broker) - within the Data Space to obtain their information, including the provider's self-description. In a simple scenario, the Data connector of the data consumer may already know the data provider's IP address or Fully Qualified Domain Name (FQDN) and contact the provider's data connector directly. In either case, he can find the most suitable data provider and get metadata information responding to their query.

Before the actual data transfer, a contract negotiation triggered by the Consumer takes place, as the two parties need to come to an agreement on the terms of the data offer (contract). The supported policies can be extracted from the self-description of the Contract Definition. By default, the data offer gets automatically accepted by the Data Consumer that eventually obtains a data agreement, based on the predefined policy. The negotiation process utilises the DSP, while the agreement process typically uses HTTPS or MQTT protocol. Having a data agreement, the Data Consumer can now retrieve the data utilising DSP triggering the data transfer of the Asset (pulling the data) or waiting for the Data Provider to push the dataset to them (the Data Consumer), depending on the supported mechanism.

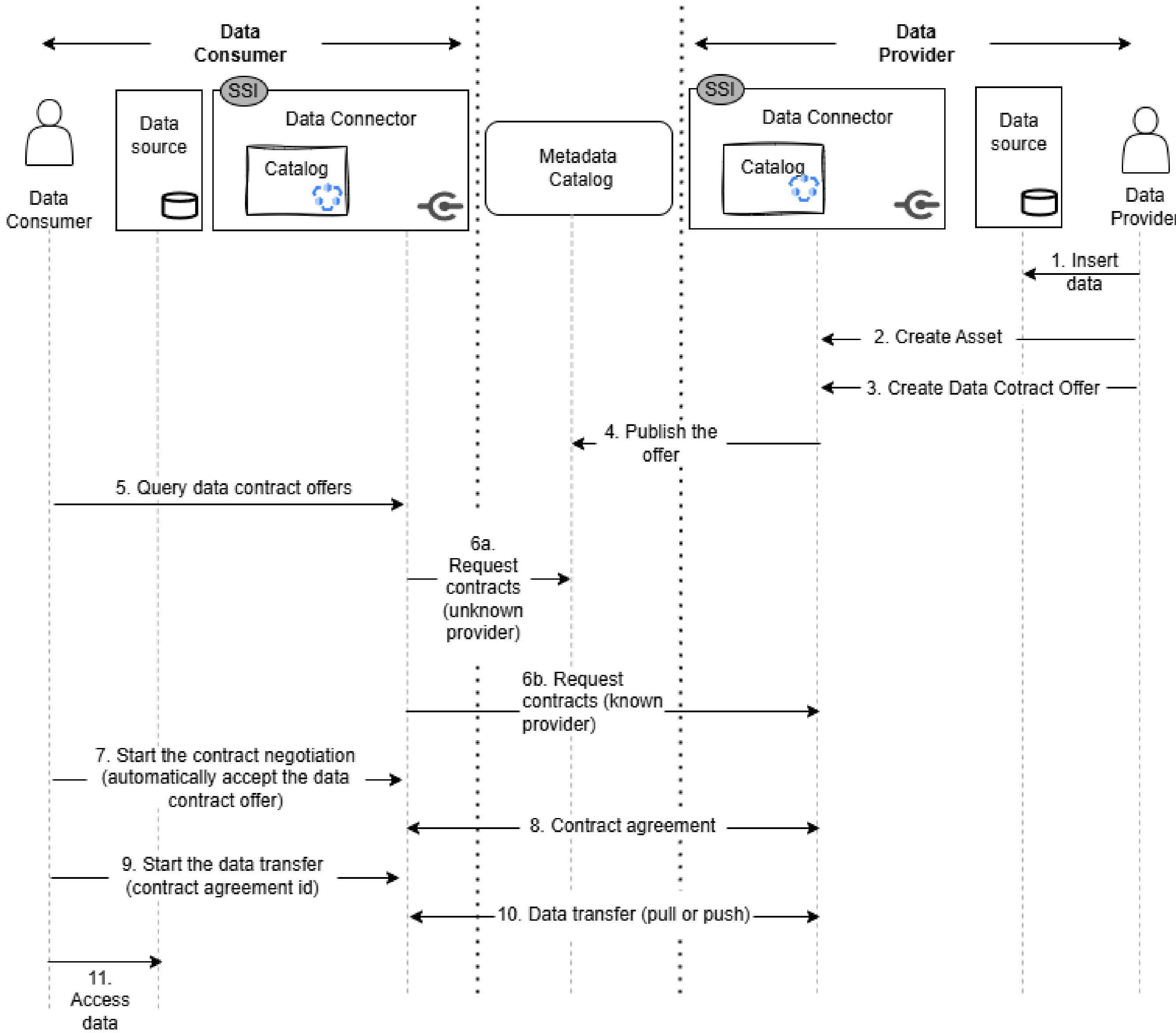

Figure 4: Data exchange workflow.

An example in the context of a PED pilot is described as follows. An Energy Supplier's platform (Data Provider) shares energy measurements of its customer portfolio via its Data Connector, defining access policies and publishing a data offer to a Catalog Service. The Demand Response Management System (DRMS) of an Aggregator (Data Consumer) discovers this offer, negotiates a data agreement, and retrieves the energy measurements using DSP. Subsequently, the Aggregator uses this data to manage and monitor Demand Response (DR) event harvesting energy flexibility.

4 SEMANTIC INFORMATION MODELS FOR PEDVOLUTION

This chapter focuses on semantic interoperability and the building of a semantic data model that covers the different aspects of PED conceptualisation, implementation and development. Initially, the data needs of PEDvolution solutions and stakeholders are analysed listing the identified data categories and their characteristics. Next, relevant state-of-the-art ontologies, standards and data models are presented followed by a gap analysis. The chapter concludes with the definition of PEDvolution's Common Information Model.

4.1 Background in Semantic Interoperability

Semantic interoperability ensures that precise meaning and interpretation of exchanged information is commonly and accurately understood by the parties or systems exchanging data. The key technologies used to construct the required ontologies are Resource Description Framework (RDF) [17] and Web Ontology Language (OWL) [18] described below.

Resource Description Framework (RDF) comprises the foundational model for representing interconnected data and standards in the semantic web. An RDF statement is a triple format "subject – predicate – object" consisting of resources where every resource is identified by a URI. In RDF, data or information is often depicted as labelled, directed graphs. In such representations, ovals and rectangles signify resources and literals, respectively, while labelled arrows denote relations or predicates. RDF supports multiple serialisation formats, such as RDF/XML, Turtle and JSON-LD.

The structure of the data is defined using RDF Schema (RDFS) [19], while the OWL describes semantic relationships. The RDF Schema (RDFS) serves as the vocabulary for RDF data, providing the semantic framework for its interpretation. The class, property, domain, and range are key concepts defined as described in the Tables below:

Table 2: RDF classes.

Class name	Description
rdfs:Resource	The class resource, everything.
rdfs:Literal	The class of literal values, e.g., textual strings and integers.
rdf:langString	The class of language-tagged string literal values.
rdf:HTML	The class of HTML literal values.
rdf:XMLLiteral	The class of XML literal values.
rdfs:Class	The class of classes.
rdf:Property	The class of RDF properties.
rdfs:Datatype	The class of RDF datatypes.

rdf:Statement	The class of RDF statements.
rdf:Bag	The class of unordered containers.
rdf:Seq	The class of ordered containers.
rdf:Alt	The class of containers of alternatives.
rdfs:Container	The class of RDF containers.
rdfs:ContainerMembershipProperty	The class of container membership properties, rdf:_1, rdf:_2, ..., all of which are sub-properties of 'member'.
rdf:List	The class of RDF Lists.

Table 3: RDF properties.

Property name	Description	Domain	Range
rdf:type	The subject is an instance of a class.	rdfs:Resource	rdfs:Class
rdfs:subClassOf	The subject is a subclass of a class.	rdfs:Class	rdfs:Class
rdfs:subPropertyOf	The subject is a subproperty of a property.	rdf:Property	rdf:Property
rdfs:domain	A domain of the subject property.	rdf:Property	rdfs:Class
rdfs:range	A range of the subject property.	rdf:Property	rdfs:Class
rdfs:label	A human-readable name for the subject.	rdfs:Resource	rdfs:Literal
rdfs:comment	A description of the subject resource.	rdfs:Resource	rdfs:Literal
rdfs:member	A member of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:first	The first item in the subject RDF list.	rdf:List	rdfs:Resource
rdf:rest	The rest of the subject RDF list after the first item.	rdf:List	rdf:List
rdfs:seeAlso	Further information about the subject resource.	rdfs:Resource	rdfs:Resource
rdfs:isDefinedBy	The definition of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:value	Idiomatic property used for structured values.	rdfs:Resource	rdfs:Resource

rdf:subject	The subject of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:predicate	The predicate of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:object	The object of the subject RDF statement.	rdf:Statement	rdfs:Resource

The Web Ontology Language (OWL) is a semantic language designed for the web, which aims at representing complex knowledge and relationships between entities (referred to as "things"). Documents of OWL models are ontologies which are constructed using RDF statements. OWL incorporates all statements from RDFS and extends them with additional ontology-related declarations, such as:

- owl:equivalentClass and owl:equivalentProperty for defining equivalences,
- owl:imports for importing existing information models, and
- owl:versionInfo for specifying the version of an information model.

Initially, at the beginning of an OWL declaration, the necessary prefixes for entities defined in RDF are specified to ensure clarity and consistency in referencing resources, as shown below:

@prefix : <http://www.example.com/ontology/exampleontology#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

Any subsequent statements provide additional information about the newly declared entity, enriching its semantic definition and contextual relevance.

4.2 PEDvolution Data Needs

A comprehensive understanding of diverse and interrelated data domains is required to secure interoperability between PEDvolution solutions, considering all the relevant aspects in the context of a PED, such as energy, economic, social, and environmental[2]. To achieve this, the Interoperability Platform of PEDvolution should address a broad spectrum of data needs, spanning energy-related aspects such as electricity and heat generation, consumption, and storage; economic factors like energy markets' data, retail energy prices, and financial incentives; social dimensions, including community engagement and demographic data; and environmental considerations such as weather patterns and climate impacts. This multidisciplinary data collection is needed by the PEDvolution solutions, so that they can be holistic, resilient, and adaptable to diverse urban contexts, highlighting the importance of robust, accurate, and interoperable data for informed decision-making. An overview of the basic data categories needed for the design and operation of PEDvolution solutions can be seen in Figure 5.

[2] An initial documentation of the needed information per High-Level Use Case (HLUC) can be found in D1.2 [9], which includes the complete documentation of HLUCs (Annex III).

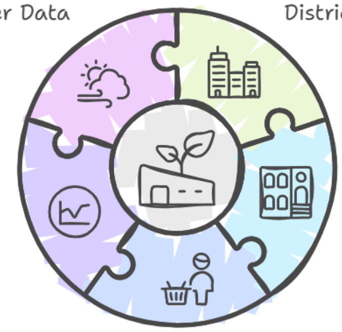

Figure 5: Basic data needs for PEDvolution solutions.

To effectively design and implement PEDs, data must be collected and aggregated across multiple levels, reflecting the layered complexity of urban energy systems. Hence, data will be collected at the prosumer level. The prosumer level, in this case, is a reference to individual connections to the grid, either these concern apartments or detached buildings. Building-level data concern entire buildings, which may consist of multiple prosumers (e.g., apartments in residential buildings, offices in commercial buildings, etc.). The district level encompasses a broader geographical and organisational scale, integrating multiple buildings and infrastructure within the PED boundaries. In addition to the above levels, energy market and weather data are considered, so that all the data needs of the PEDs are accounted. More details on each category are provided in the following sections.

4.2.1 Prosumer Level

At the prosumer level, data collection focuses on both the specifications of prosumer's energy assets and detailed energy measurements. This includes data on the specifications of energy assets such as solar panels, battery systems, and HVAC systems, alongside granular measurements of energy production, consumption, and storage. These datasets are critical for enabling contributions of individual prosumers to the PED's energy balance, as well as to offering flexibility services to the energy market through flexibility aggregation. Additionally, prosumer energy prices form a vital component, encompassing district heating (DH) prices, retail electricity prices, network electricity prices, and feed-in tariffs for PV generation. Together, these data categories ensure that technical, operational, and economic aspects relevant to prosumers are fully represented, enabling effective integration into the broader PED framework.

4.2.2 Building Level

At the building level, data needs encompass multiple critical categories that provide a detailed understanding of building characteristics, operations, and energy dynamics. Construction characteristics include information such as year of construction, building type (e.g., residential, commercial), total floor area, number of floors, window area, and renovation-related information. These factors influence the building's energy performance and retrofitting needs. Building occupancy patterns focus on the number of occupants and their typical activity schedules, such as working hours for businesses or peak usage times for residential buildings, which are crucial for modelling energy demand. Building energy assets characteristics refer to shared systems such as a centralised HVAC system or an EV charging station, providing details on the assets' specifications. Building energy measurements include energy measurements for shared building energy assets and aggregated measurements that account for the energy usage of individual prosumers within the building, whether apartments, offices, or other types. Additionally, the building's Energy Performance Certificate (EPC) provides a standardised assessment of energy efficiency. These datasets provide insights into building energy efficiency and informed retrofitting or energy optimisation strategies, as well as facilitate the provision of energy and flexibility services from buildings of various types.

4.2.3 District Level

At the district level, data collection encompasses diverse categories that reflect the physical, economic, social, and energy-related characteristics of the area. District geospatial data include location coordinates such as district boundaries and the locations of buildings and other infrastructures, which are crucial for urban planning. District energy assets characteristics focus on specifications of energy assets within the district, such as combined heat and power (CHP) plants, hydropower plants, photovoltaic (PV) parks, stationary battery systems, Electric Vehicle (EV) charging stations, and District Heating and Cooling (DHC) infrastructure. District energy measurements include energy generation data from various district-level units, covering diverse energy sources, as well as aggregated measurements of electricity and heat production and consumption collected at the building level. District economic characteristics capture key indicators such as life-cycle costs (LCC), land use efficiency, value stability, and exposure to environmental risks. District social characteristics pertain to aspects such as population demographics, while district social assessment involves metrics like social engagement indicators, providing insights into community involvement and support for PED initiatives. Lastly, district assessment and certification data document the evaluation and certification of a district towards becoming a PED. District-level data are crucial for understanding the holistic performance of a PED, enabling system-wide energy and flexibility optimisation, strategic planning, and alignment with broader economic, social, and environmental goals.

4.2.4 Energy Market Level

Apart from the above data categories, energy market data are essential in the design and operation of PEDs, facilitating their effective interaction with energy systems. Electricity market data cover key components (e.g., day-ahead market, intraday market, balancing market, forward market, and ancillary services market), thus providing the framework for trading electricity, balancing supply and demand, and accessing flexibility services. Additionally, other relevant markets, such as those for natural gas, may also be considered, depending on the exact context of PEDs and the specific needs of PEDvolution solutions. This data is essential for optimising energy costs, enabling market participation, and aligning PED operations with broader energy system dynamics.

4.2.5 Weather Data

Weather data are also fundamental for the planning and operation of PEDs, given their significant impact on energy production and consumption patterns. Weather measurements provide information on parameters such as temperature, solar radiation, wind speed, and humidity, enabling precise adjustments to energy systems. Weather forecasts support PED operation by predicting these conditions, allowing for optimised energy scheduling. Additionally, climate data offer long-term insights into historical and projected weather patterns, which should be considered by PEDvolution planning tools to account for changing environmental conditions. Together, these datasets ensure that PEDs remain efficient, adaptable, and aligned with their sustainability goals.

4.2.6 Summary

Table 4 provides an overview of various data categories that are needed by the PEDvolution solutions, exposing the versatile nature of these data – in line with the broad scope of PEDvolution – as well as the different levels at which data will be collected or aggregated. To conclude with, it is noted that timeseries data, such as electricity consumption and generation, can be converted to different frequencies according to the data needs of various PEDvolution solutions.

Table 4: Collection of needed data at different levels.

LEVEL	PROSUMER	BUILDING	DISTRICT	REGION
GEOSPATIAL DATA			●	
ECONOMIC CHARACTERISTICS			●	
SOCIAL CHARACTERISTICS			●	
SOCIAL ASSESSMENT			●	
ASSESSMENT AND CERTIFICATION DATA		●	●	
ENERGY ASSETS CHARACTERISTICS	●	●	●	
ENERGY MEASUREMENTS	●	●	●	
CONSTRUCTION CHARACTERISTICS		●		
OCCUPANCY PATTERNS		●		
ENERGY PRICES	●			
ELECTRICITY MARKET DATA				●
WEATHER DATA			●	●

4.3 Analysis of Existing Data Models, Ontologies and Standards

In this section, state-of-the-art ontologies and data models which cover aspects relevant to the PEDvolution's data needs are analysed focusing on ontology parts which are of interest to PEDvolution and identifying gaps for extension to cover the PED context. These include:

- The **Smart Applications REFerence (SAREF)** ontology is a standardised semantic model developed by ETSI (European Telecommunications Standards Institute) to enable interoperability between Internet of Things (IoT) devices, services, and platforms across various domains structures. It contains a core ontology and extensions covering specific vertical domains. In PEDvolution, we will focus on SAREF4ENER (SAREF for Energy Flexibility), SAREF4BLDG (SAREF for Buildings) and SAREF4CITY (SAREF for Cities).

- **OMEGA-X CSDM** is a modular ontology, developed by EDF in the context of the Horizon Europe project [20] which ensures seamless data integration and supports semantic interoperability within the energy Data Space, as well as cross-project collaboration in the wider Data Spaces ecosystem. It facilitates standardised communication across various stakeholders in the energy sector, promoting efficient data exchange and innovation.
- The **Brick** schema is an open data model for buildings that defines a uniform framework for representing and organising metadata about building systems, devices, and their relationships.
- **BOT** (Building Topology Ontology) is a lightweight ontology for describing the spatial and structural aspects of buildings, providing a foundation for integrating building data with other ontologies and facilitating interoperability in building information systems.
- **EnergyADE** (Application Domain Extension) for CityGML is an extension specifically designed to represent energy-related information for urban environments, enriching the CityGML model with concepts and attributes for energy simulations, analysis and planning at the city scale.
- **FlexOffer** is a well-established data model focusing on energy flexibility and adopted by many EU Research and Innovation projects. On the energy market level the FlexOffer data model is investigated for modelling flexibility transactions.

4.3.1 SAREF and its Extensions

The SAREF ontologies family [19] structures the core semantic layer for devices to communicate within smart buildings, achieving energy management and optimisation both at building level and district level. SAREF Core provides a shared vocabulary representing smart appliances and their functionalities. Extensions of SAREF ontology, such as SAREF4ENER, SAREF4BLDG and SAREF4CITY, expand this framework towards more complicated energy-related UCs. Below the core ontology and its extensions are briefly introduced.

SAREF Core

The **SAREF Core** (Smart Appliances REFerence ontology) [21] is a standardised ontology developed to enable seamless interoperability between devices (*saref:Device*) and systems (*s4syst:System*) in the IoT ecosystem. Initially designed for smart appliances, SAREF Core provides a structured framework for describing devices, their capabilities, and interactions. By standardising the way devices communicate and exchange information, SAREF ensures that products from different vendors and manufacturers can work together efficiently, fostering a more interconnected and smart environment. It improves device interoperability by focusing on fundamental IoT concepts, including device properties (*saref:Observation*) and functions (*saref:Function*) and measurement units (*saref:isMeasuredIn*).

As part of the broader SAREF family of extensions, SAREF Core is often used as a base ontology that integrates with domain-specific extensions tailored for energy, building automation, and more. Its modular design allows for adaptability and extension, making it appropriate for IoT applications that demand scalability and flexibility. In compliance with international standards, SAREF Core supports the development of innovative smart solutions while addressing critical challenges such as compatibility, efficiency and user-centric schemes following the rapidly evolving IoT technology and advancements.

Within the SAREF Core ontologies, various extensions have been developed to address specific domains and enhance interoperability across diverse systems. These extensions, such as SAREF4ENER for energy, SAREF4BLDG for buildings, and others provide energy domain-specific vocabularies and frameworks.

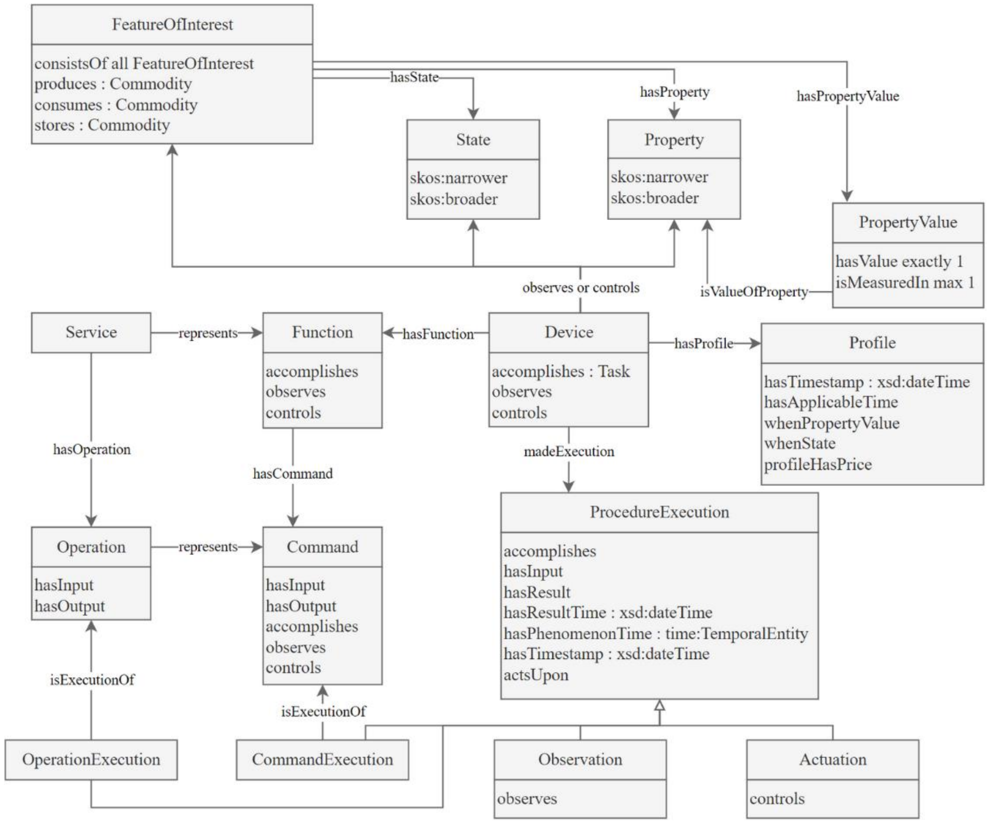

Figure 6: SAREF Core ontology [21].

Within SAREF Core ontology, Building Energy Assets datasets are included, containing details and specifications of the equipment installed in residential or other type of buildings, such as appliances (*saref:Appliance*), meters (*saref:Meter*), sensors (*saref:Sensor*), actuators (*saref:Actuator*) and other utility loads, that enable monitoring and control to enhance energy efficiency and management. More specifically, this ontology provides the principal information including the types of devices, their properties, and operation mode profiles (*saref:Command*), to enable an initial comprehensive energy management within the system. Other assets-, energy-, price-related description details, that were previously included in the SAREF Core framework, will be transferred and deployed through SAREF4ENER, SAREF4BLDG and SAREF4CITY extension ontologies introduced below.

SAREF4ENER

The **SAREF4ENER** (Smart Appliances REFerence ontology for the energy domain) [22] extension builds upon the SAREF Core ontology to facilitate interoperability and efficient communication among energy-related devices, systems, and services. It is designed to model energy-related data and processes, enabling smarter energy management in residential, business and industrial sectors. By standardising how energy consumption, production (*saref4ener:hasPowerSequenceContainer, s4ener:hasEnergy*), and storage (*s4ener:Storage*) are described, SAREF4ENER fosters seamless integration between energy systems, including smart grids, renewable energy sources and energy management platforms.

This ontology supports the optimisation of energy usage by modeling key concepts such as demand-response (*s4ener:DemandDrivenProfile*) scenarios, in which customers can manage their smart devices and offer flexibility to the grid, energy generation and storage, and load management, including scenarios of flexibility aggregation and activation (*s4ener:FlexibilityProfile*). It provides a framework for representing and exchanging data on energy flows, consumption patterns and system states (e.g., *s4ener:hasOperationModeElement*), enabling informed decision-making (e.g., *s4ener:EventStateConsume, s4ener:EventStateProduce*) for energy efficiency and sustainability. More specifically, at building and asset level, the power consumption/production measurements data (*s4ener:Power*) describes detailed metrics on power usage and generation and enables the monitoring (*s4ener:LoadControlEventState*) and optimisation (*s4ener:PPBCInstruction*) of energy flows to improve efficiency and support grid stability.

Alongside occupancy patterns are recorded enabling better energy management and optimisation designs in case of demand-driven solutions (*s4ener:DemandDrivenProfile*). This data category contains attributes, such as number of occupants (residential, business, etc.), occupancy rates (patterns), and the like.

Economic Data are also introduced, such as energy costs (*ResumeCostEstimated, s4ener:hasTransitionCosts*), and any energy-related expenses e.g., *s4ener:RelativeCost, s4ener:AbsoluteCost* for incentive types activities), assisting with cost-benefit analyses and decision-making for sustainable energy management.

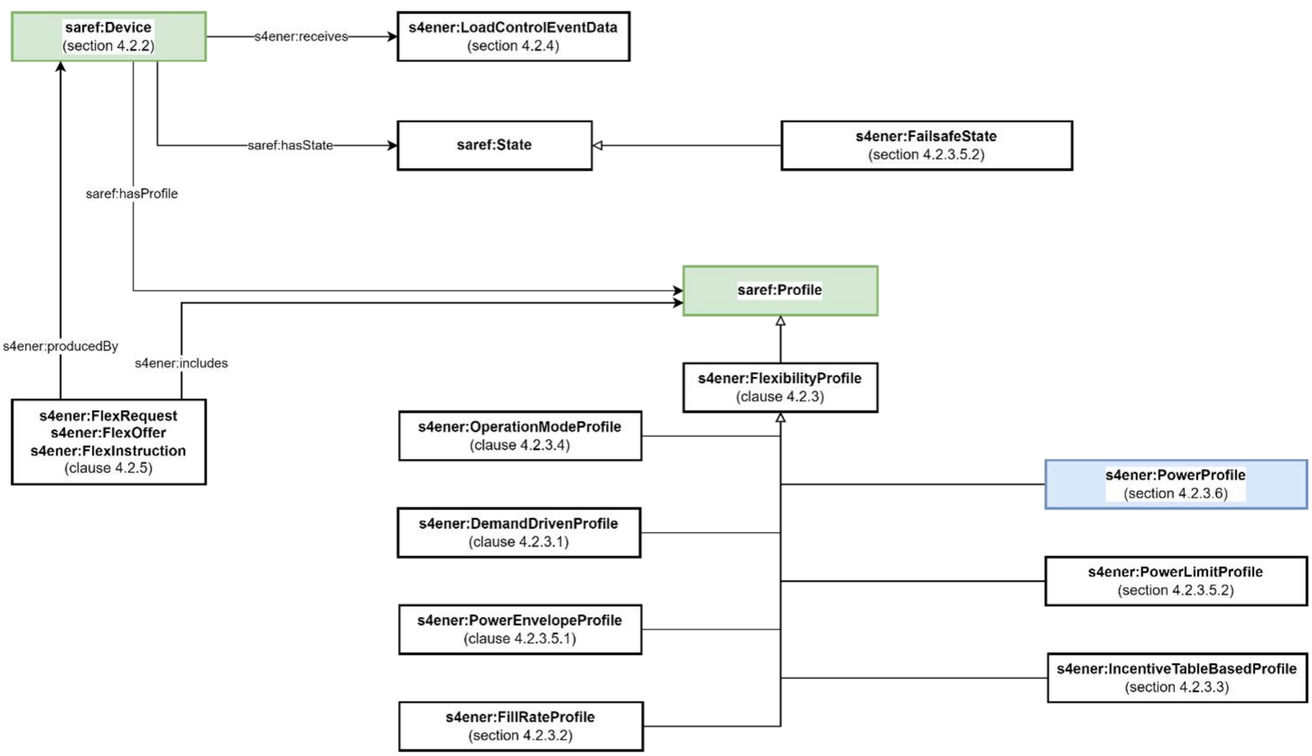

Figure 7: SAREF4ENER ontology [22].

Finally, *CO_2 Emissions Data* are also included which concerns carbon emissions data related with energy consumption and production, enabling analysis on the environmental impact and facilitating compliance with sustainability and regulatory aspects (*s4ener:FlexibilityProfile*).

SAREF4BLDG

SAREF4BLDG [23] is an extension of the SAREF Core ontology designed for the building sector, focusing on improving the interoperability of smart building systems. It provides a standardised framework to represent devices (*s4bldg:BuildingObject*), systems and processes within buildings (e.g., *s4bldg:DistributionDevice*), enabling seamless communication between heterogeneous technologies. SAREF4BLDG supports applications such as energy management, building automation and occupant's well-being and comfort by facilitating data exchange between systems such as lighting (*s4bldg:integratedLightingType*), boiler (*s4bldg:isWaterStorageHeater*) and the like. This information model plays a vital role in improving the efficiency and sustainability of buildings by enabling the integration of IoT devices and smart services, being ideal for managing interactions among smart appliances and building energy management systems (BEMS). With SAREF4BLDG, developers can model building-specific entities, their functions and interconnections, ensuring that different components work in cohesion. This not only optimises resource usage and reduces energy consumption but also enhances user experience and operational supervision. Different datasets are described by the SAREF4BLDG ontology. First, Building Space data describes the physical attributes of buildings, such as size (*internalSurfaceArea*), and spaces (*s4bldg:hasSpace*). Regarding Building Energy Assets (Building Objects and Devices) datasets describing building equipment are included, such as distribution flow/control devices, transportation devices, energy conversion devices (electric generators, evaporator, health devices, solar devices, shading devices, heat exchangers). This extension covers a wide range of devices that serve various occupants needs offering opportunities for comfort enhancement.

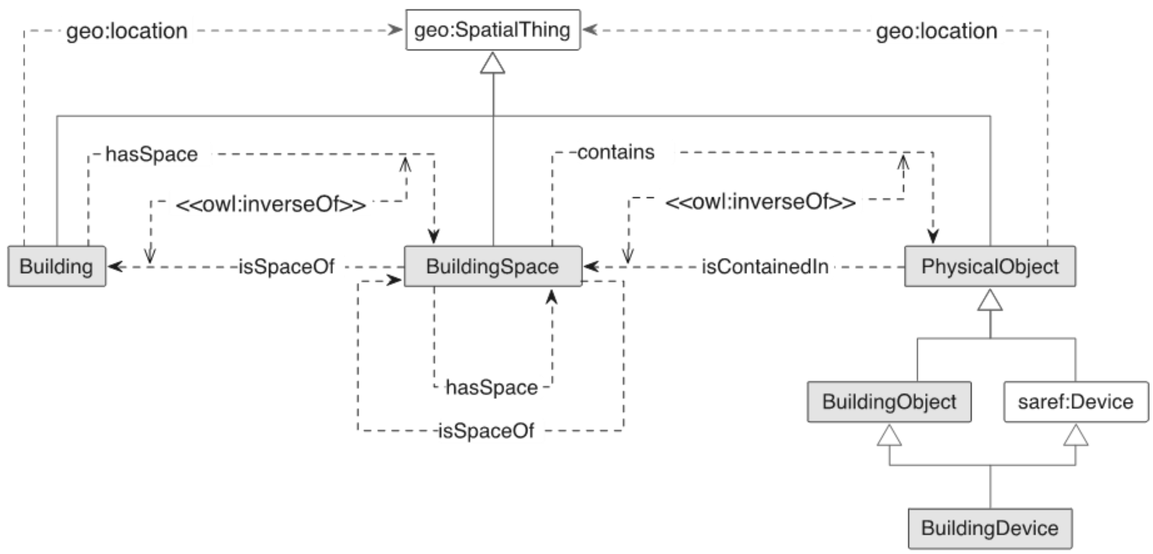

Figure 8: SAREF4BLDG ontology [23].

SAREF4CITY

SAREF4CITY [24] is an extension of the SAREF ontology designed specifically to address the needs of smart cities. It facilitates interoperability between various IoT devices, systems, and platforms used in urban environments. By providing a standardised framework, SAREF4CITY enables efficient data exchange and integration across domains such as energy. The ontology supports the modeling of urban entities, services and their relationships, ensuring a unified representation of city operations and enabling the transition toward a more connected and intelligent urban living. It is deployed for broader city-level scenarios such as traffic management, public lighting and energy flows among buildings and public infrastructure.

At district level, datasets record key details about the district area (*s4city:AdministrativeArea*), including its geographical boundaries, district coordinates, demographical characteristics of the population, infrastructure layout, building locations, existing utilities and infrastructure, energy demand patterns. This type of data encourages efficient urban planning, energy distribution and initiatives of sustainable development.

Furthermore, data about the building assets, as already mentioned in SAREF4ENER ontology, are also involved with SARE4CITY for the Smart Cities domain. This dataset provides detailed information on housing equipment, such as appliances and energy systems, alongside district-level technical characteristics like heating and cooling networks, energy distribution infrastructure and connectivity.

Then, District Heating and Cooling Infrastructure data includes the technical characteristics and connectivity of district-level heating and cooling systems, enabling coordinated energy distribution and management across multiple buildings. Some of the data attributes are: pipes network, pipe length, insulation, pressure zones, heat generator, heating storage, heating/cooling peak power, substation capacity, heat loss in pipes, installation cost, operational cost, substation efficiency, etc.

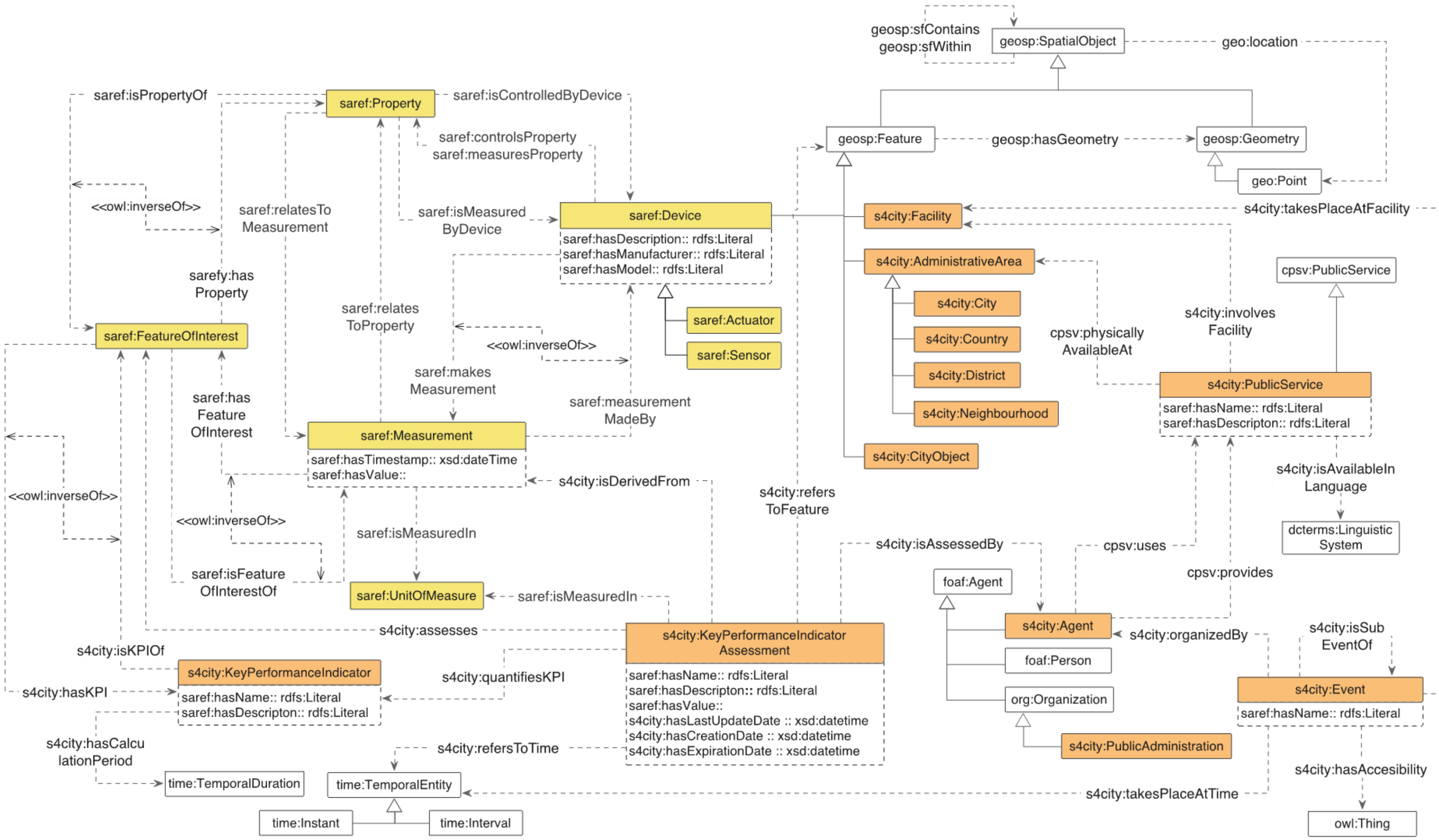

Figure 9: SAREF4CITY Ontology [24].

4.3.2 Energy ADE for cityGML

The **Energy Application Domain Extension (Energy ADE) for CityGML** [25] is a data model extension of CityGML 2.0 standard [26], designed to provide a conceptual model and exchange format for the representation, storage and exchange of virtual 3D city models, with energy-related information for urban and building energy performance analysis and simulations. Energy ADE enriches cityGML datasets by incorporating key energy-related features and characteristics of urban cities and buildings, such as building energy demand, energy systems, occupants, and energy-related building properties. These extensions address requirements for urban energy planning, performance analysis, and the integration of renewable energy systems enabling various applications like energy demand diagnostics, solar potential study and simulation of low-carbon energy strategies for developing sustainable urban environments.

The Energy ADE is developed based on the INSPIRE Directive [27] of the European Parliament, along with the US Building Energy Data Exchange Specification (BEDES) [28] towards representing energy-related data collected at urban scale. It is also aligned with international standards such as ISO 13790 for energy balance methods and enabling sub-hourly dynamic simulations using tools like CitySim or EnergyPlus to provide information data required by different urban energy models and simulation.

The Energy ADE is organised into six main modules described next.

The **Energy ADE Core** module extends the relevant CityGML building features (*AbstractBuilding*) with energy-related parameters useful for the energy assessment. It classifies the building usage (*buildingType*) and its construction structure (*constructionWeight*) and describes important geometrical (*volume*, *floorAarea*), locational (*referencePoint*, *heightAboveGround*) and average material parameters of the building's exterior shell (*aggregatedBuildingConstruction*). It also describes the energy performance certificate information if the building has any as well as information regarding its refurbishment (refurbishmentMeasure) like the date and level of refurbishment. CityGML accounts for different building feature types being either a complete building or a building part.

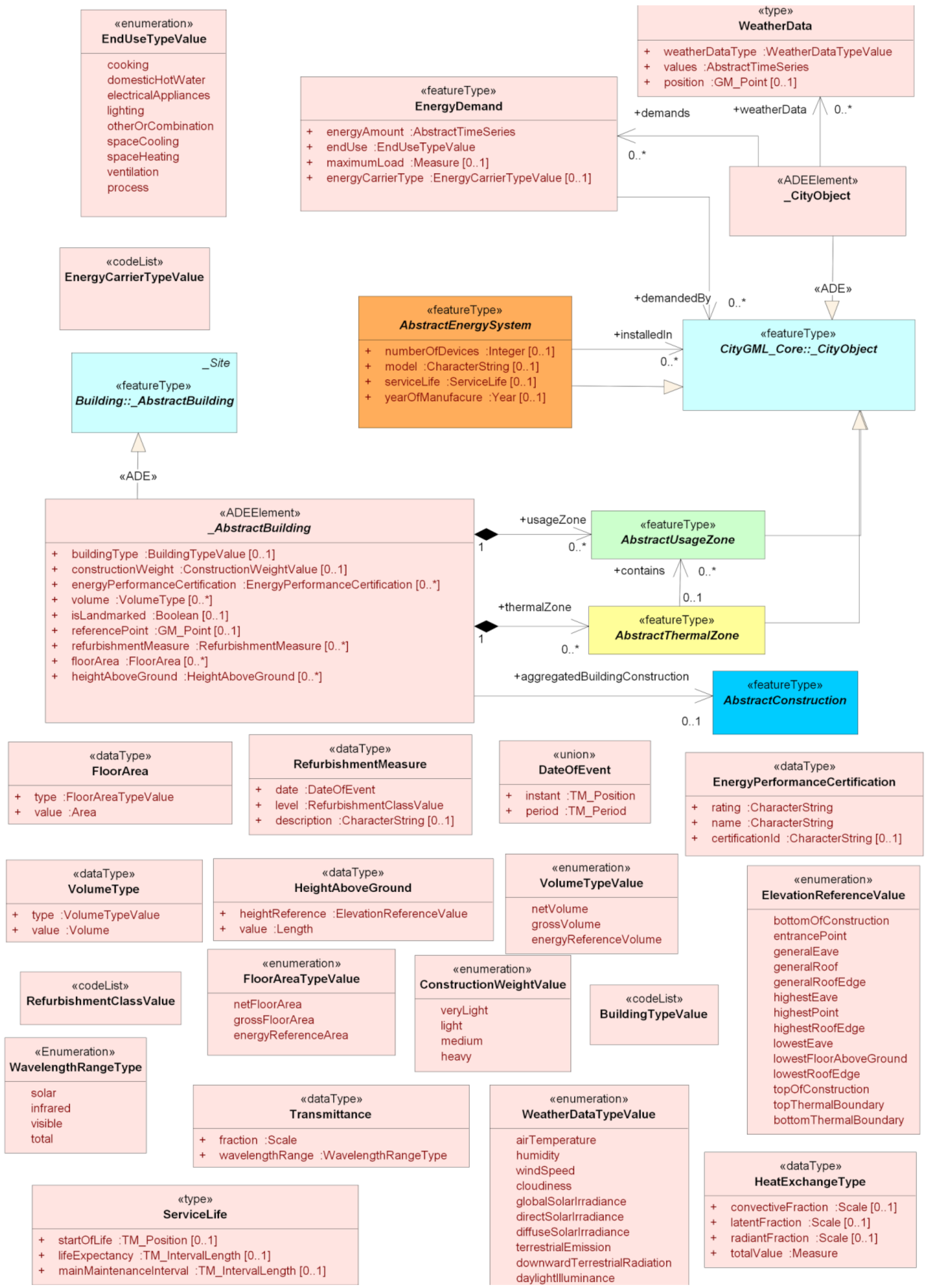

Figure 10: Energy ADE Core [25].

The **Building Physics** module describes parameters for single- or multi-zone building energy simulations such as thermal zones which can be useful for building heating and cooling demand calculation. For example, thermal zones can be used to distinguish rooms with different orientations (solar gains) and/or thermal behaviour.

The **Occupant Behaviour** module models the occupants' energy-related behaviour. A building may contain different usage zones, i.e., zones with homogeneous usage conditions and indoor climate control settings. Usage zones can be useful to calculate energy demand related to occupants' end-uses, such as domestic hot water, specific electrical appliances and the like. Also, occupants are categorised by their *occupancyType* (residents, etc.), the number of persons (*numberOfOccupants*) in the usage zone and their time schedule (*occupancyRate*). For the thermal modelling of the zone, the heat dissipation (*heatDissipation*) is modelled. The occupant behaviour model also contains the households of occupants describing parameters like the household type (a family, adult, etc.) or the type of residence (main, secondary or vacant). Finally, the facilities in the building are also described specifying parameters regarding their usage (*operationSchedule*) and emitted heat (*heatDissipation*), while they are further distinguished into domestic hot water (*DHWFacilities*), electrical appliances (*ElectricalAppliances*) and lighting facilities (*LightingFacilities*). HVAC systems are not included as they are described in the Energy System module.

The **Material and Construction** module describes physical parameters of building materials like walls, roofs and windows, their structure, and their thermal and optical properties.

The **Energy Systems** module represents the energy conversion, distribution, storage and emission devices of a building or city object as well as the energy flow between them. Elements of interest include the *Emitter which* describes the end unit emitting useful energy (e.g., heat) in the zone where it is installed. The emitter is characterised by its type (e.g., radiator), the installed power and the thermal exchange type providing the proportions of different thermal exchange mechanisms (e.g., convective or radiative). Flow of energy and relevant environmental factors like CO_2 emissions are also modelled. The Energy Systems module also classifies energy conversion systems, such as boilers, chillers, heat pumps, etc. describing parameters relevant to their physical properties, such as their efficiency (e.g., thermal, electrical etc.) and technical specifications (e.g., condensationType etc. for a chiller). Among others, solar energy systems are further modelled covering both photovoltaic systems and solar thermal systems.

The **Supporting Classes** module represents time series of physical values, schedules of different kinds (like operations for electrical devices, occupancy rates, temperatures for heating and cooling systems) and weather or climate data supporting time series of meteorological parameters.

Energy ADE elements from the Energy System module such as for example solar systems can be used to describe energy facilities at the district level, however centralised energy infrastructures, like district heating system or gas networks are only covered by Utility Network ADE (Kutzner and Kolbe, 2016). Also, Energy ADE does not cover other district characteristics like social (e.g., demographic factors, energy awareness) and economic aspects (e.g., energy pricing, operational costs, construction costs). Energy ADE focusing on the building construction characteristics and occupancy patterns will be used to enrich PEDvolution's information model with additional building parameters. Weather data will also complement any missing properties in the formed information model.

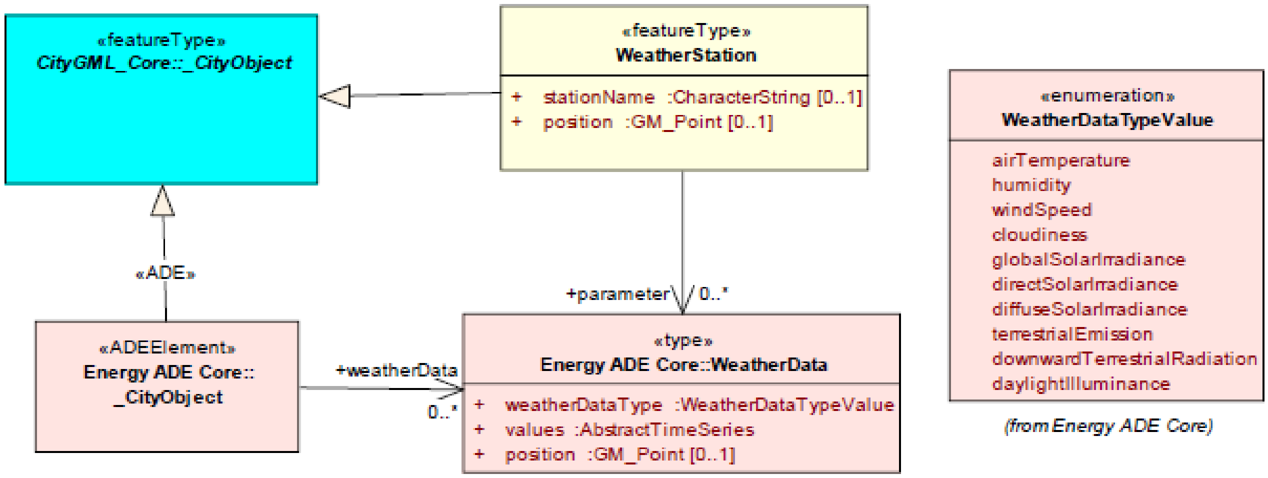

Figure 11: Weather data in Energy ADE.

4.3.3 Building Topology Ontology (Bot)

The **Building Topology Ontology** (BOT) [29] is a minimal OWL ontology that describes the building's sub-components and their relationships. It has been designed as a minimal, extendable framework tailored to building-specific topology which can be used as baseline for more specialised domain specific ontologies (e.g., in smart city) following W3C principles by promoting reuse and maintain a schema that is as simple as possible. Key entities to model building topology are the following:

- **Site:** It represents a spatial area containing one or more buildings.
- **Building:** It represents a physical structure, i.e., an independent unit, on a site.
- **Storey:** Represents a level or floor within a building.
- **Space:** Represents a bounded area within a building, such as rooms or zones.
- **Element:** Represents physical components within a building, like walls or doors, chairs and devices, e.g., temperature sensor.
- **Zone:** Represents a conceptual or functional grouping of spaces or elements, often defined by specific criteria (e.g., HVAC zones, a site, a building, a space, an apartment etc.).

Alignments with SAREF4Bldg and BRICK are available. Bot elements have been adopted in CSDM ontology.

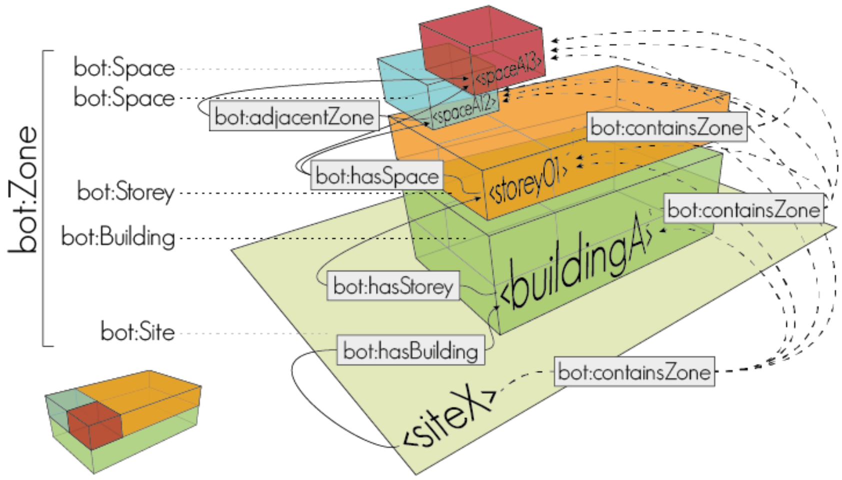

Figure 12: Bot ontology [29].

4.3.4 Brick

Brick is an open-source ontology-based data scheme [30] capturing the entities and relationships within a building to represent buildings and their subsystems covering all the different building aspects (operational, structural and functional aspects). In Brick entities represent physical, logical and virtual assets in buildings covering equipment (e.g., air handling units etc.), points (e.g., sensors readings, setpoints, alarms, commands/actuations, etc.), locations (e.g., rooms, floors etc.), and logical collections (e.g., HVAC zones, lighting zones etc.), supporting the use of tags (e.g., sensor, setpoint, air, water,) for entities, borrowed from Project Haystack, to provide flexibility for annotation. Brick models describe the context of data sources, i.e., what and where they are, and how they relate to building systems. Brick schema is integrated with RealEstateCore (REC) such as for location classes and relations where possible. For example, spaces (Space) such as Building, Site, Zone, Room, Level etc., as well as building elements (BuildingElement) like Roof, Wall, Façade etc. are adopted by REC. On the other hand, building facilities can be described using Brick collections such as for Electric Vehicle Charging Stations (or Hubs) and PV systems (photovoltaic arrays with their panels), as well as Brick systems like Domestic Hot Water, Electrical, Gas, HVAC, Lighting, Safety and Shading System. All the above collections and systems are linked with their Meter measuring consumption/production of the relevant brick: substance (e.g., gas, oil, water, CO_2, air, light etc.). All equipment classes like PV panel, gas distribution, solar thermal connectors, water distribution, water heater and dedicated weather stations are described in detail by their characteristics. Finally, events model temporally indexed entities such as an observation which can be instantaneous (timestamp property assigned) or have temporal extent (start and end properties assigned).

Building equipment classes like HVAC and EV charging systems and their characteristics will be utilised within PEDvolution to cover relevant classes in PEDvolution's CIM while building information integrated by REC in Brick will enrich missing building information e.g., from Energy ADE for CityGML.

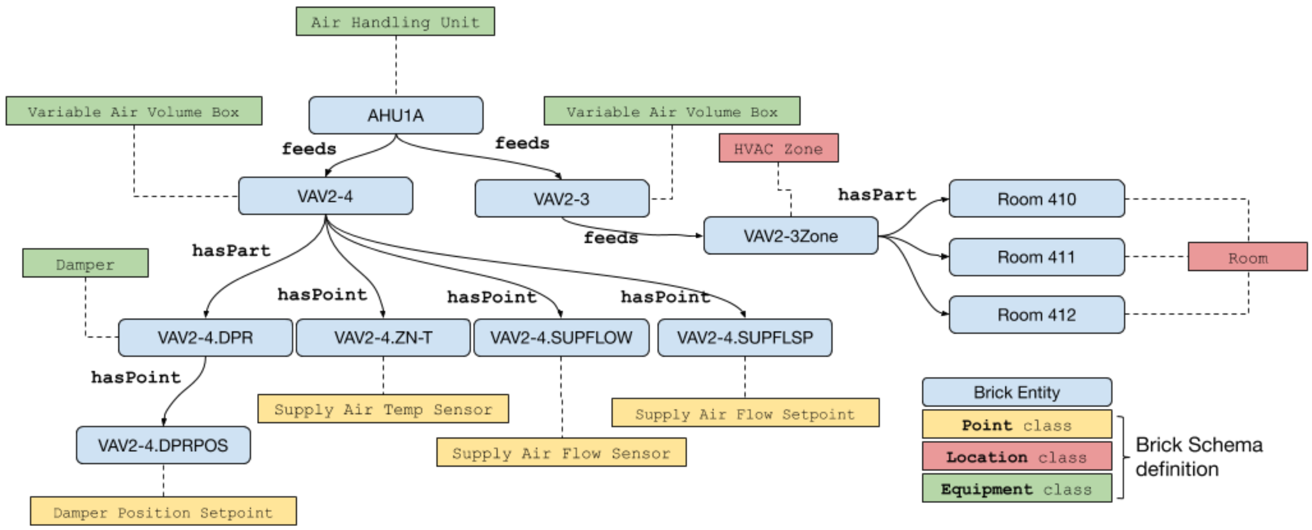

Figure 13: Brick model example [30].

4.3.5 OMEGA-X CSDM

OMEGA-X Common Semantic Data Model (CSDM) [20] is a modular ontology developed as part of the OMEGA-X project [7] to facilitate data integration and semantic interoperability in its energy Data Space among different energy stakeholders. OMEGA-X CSDM was designed to meet the needs of data exchanges across four distinct energy use cases: flexibility, renewable energy, local energy communities, and electromobility. OMEGA-X CSDM primarily focused on modeling timeseries data, units, and measurement types. It consists of the following modules described below.

Events and Time Series

The **Events and Time Series (ETS)** ontology defines a general structure of datasets and their underlying elements to be shared. The *ValueSet* is a generic concept describing a shared dataset and its information (the dataset's unique identifier, its name and time created) covering different structures such as a *DataPoint*, a *DataCollection*, an *Event*, or *TimeSeries* each having their own specific attributes. The building of the ontology is based on timeseries and datapoints from Interconnect [31] and SEDMOON [32] events and energy properties evaluation.

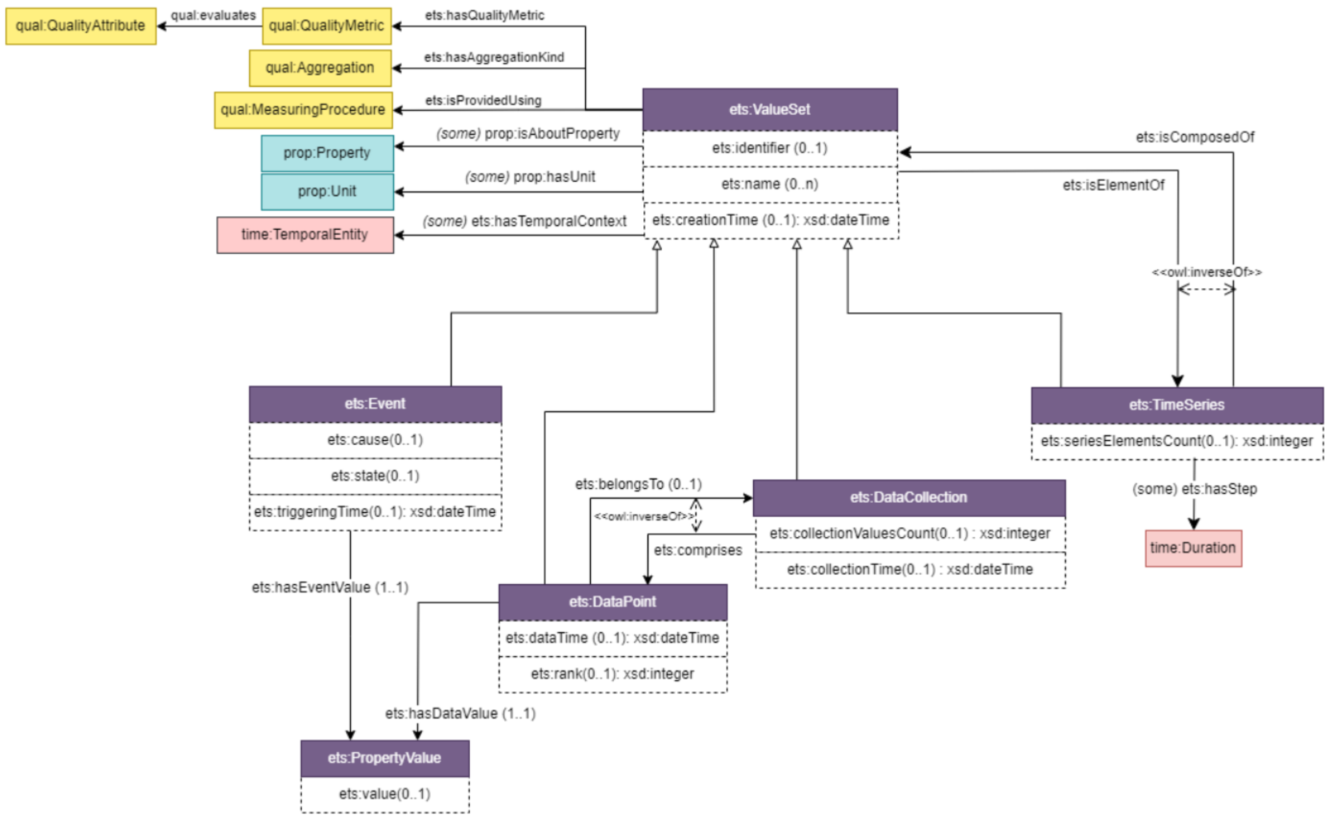

Figure 14: OMEGA-X CSDM ETS ontology [20].

ETS ontology can be used as the baseline to build PEDvolution's datasets structures to be transferred through the platform. These include profiling data, assets data, as well as energy performance data for the case of building, district characteristics at its four aspects (technical, economic, social and geospatial data) and its consumption/production data and finally weather and climate data from external services.

Quality

Quality (qual) ontology describes the quality dimensions of the data. It utilises the *Aggregation* level to indicate the granularity of the data e.g., average, sum, median etc. *QualityMetric* can be used to define any quality metrics determining the quality of as data value, while the MeasuringProcedure describes the procedure used for generating the specific data e.g., measurement, simulation, prediction etc.

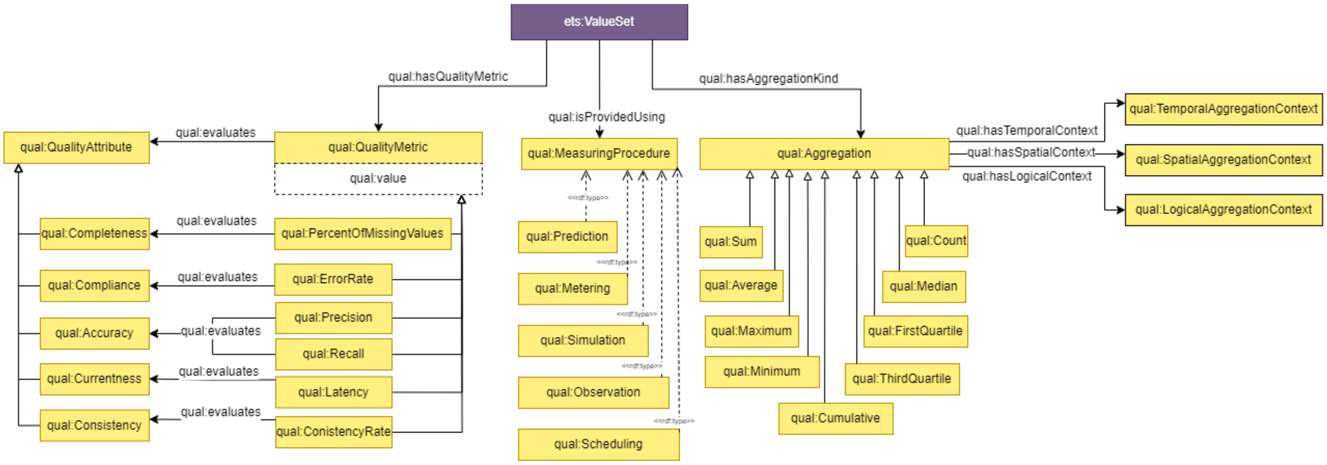

Figure 15: OMEGA-X CSDM qual ontology [20].

Aggregation and Measuring procedure are of interest to PEDvolution to accompany the characterisation of the shared datasets by ensuring that shared datasets are accurately characterised and analysed. Concerning aggregation, it involves combining data from various sources or time intervals and provides a comprehensive view of energy performance and consumption patterns within the PED framework. Meanwhile, the measuring procedure ensures that data is consistently collected, validated and aligned with predefined metrics or standards. These processes enhance data reliability and facilitate informed decision-making towards energy usage optimisation and achievement of PED project goals.

Energy Data Set

The **Energy Dataset (EDS)** ontology is used to add contextual information in the datasets to be exchanged through the energy Data Space that facilitate data interpretation. Its main feature is the EnergyDataSet which corresponds to the eds: ValueSet to be exchanged. TechnicalContext can be used to describe technical characteristics for the dataset while ExchangeContext can be used to describe the participants in the exchange. The ontology is generic and can be used for the Data Space approach of PEDvolution.

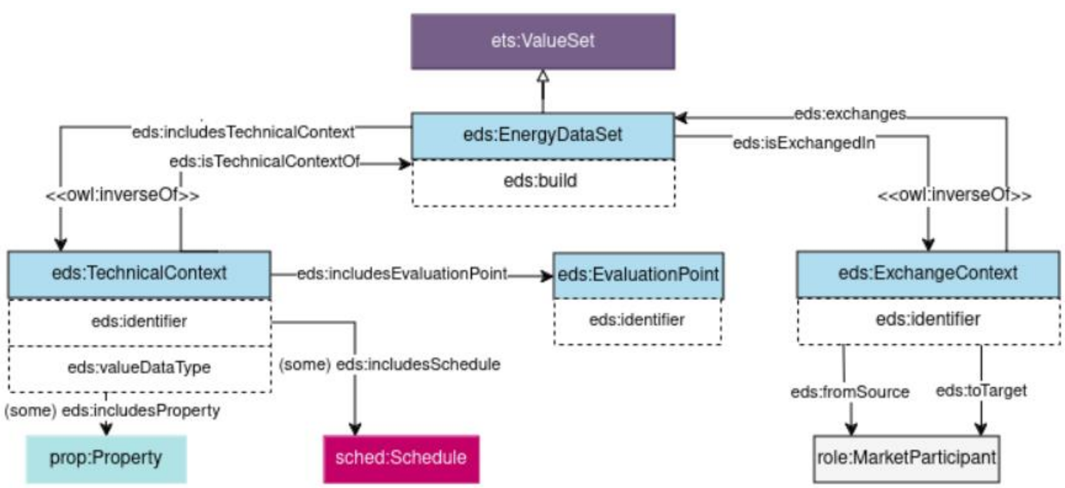

Figure 16: OMEGA-X CSDM EDS ontology [20].

Energy role ontology

The **Energy Role (role)** ontology describe the stakeholders (*MarketParticipant*) in the energy market, their role/behaviour (*MarketRole*) and relationships (*Association*) between them.

The roles, participants and associations between them to be used in PEDvolution have been defined in D1.3 [8].

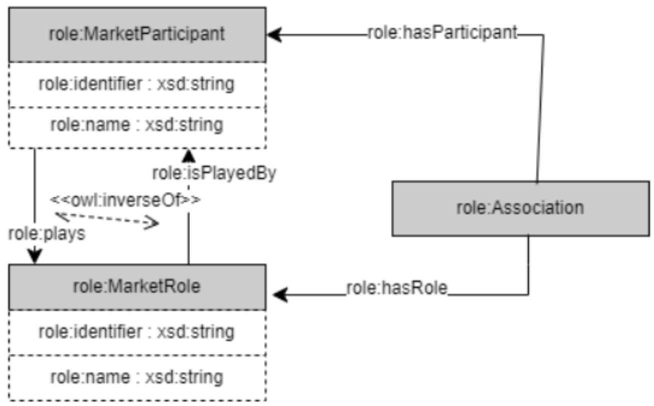

Figure 17: OMEGA-X CSDM role ontology [20].

Property

The **Property (prop)** ontology provides a taxonomy for energy domain properties including their units and relevant categories.

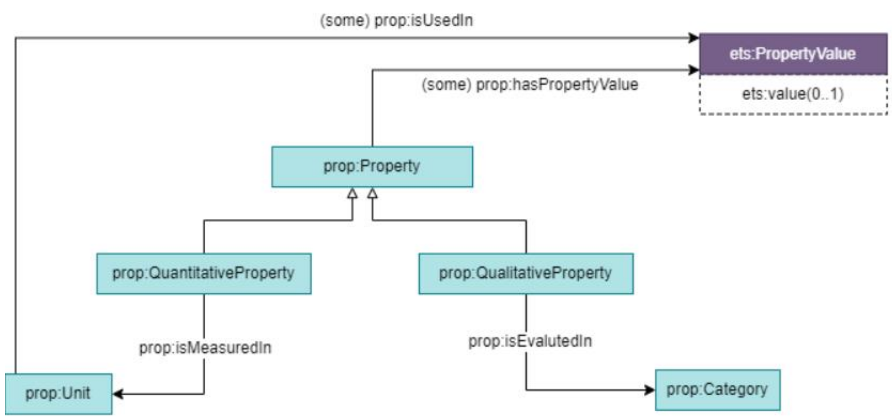

Figure 18: OMEGA-X CSDM prop ontology [20].

EU Metering

OMEGA-X CSDM adopts the **EUMED (European My Energy Data) Metering** ontology representing EUMED Metering profile.

The EME ontology can be used for energy metering data in PEDvolution.

Infrastructure ontology

Infrastructure (Infra) ontology describes infrastructure of an energy system. The *System* (equipment or site) is characterised by a system configuration with a set of properties. The building of the ontology is based on system definition in SEAS and adopts the Zone definition from BOT.

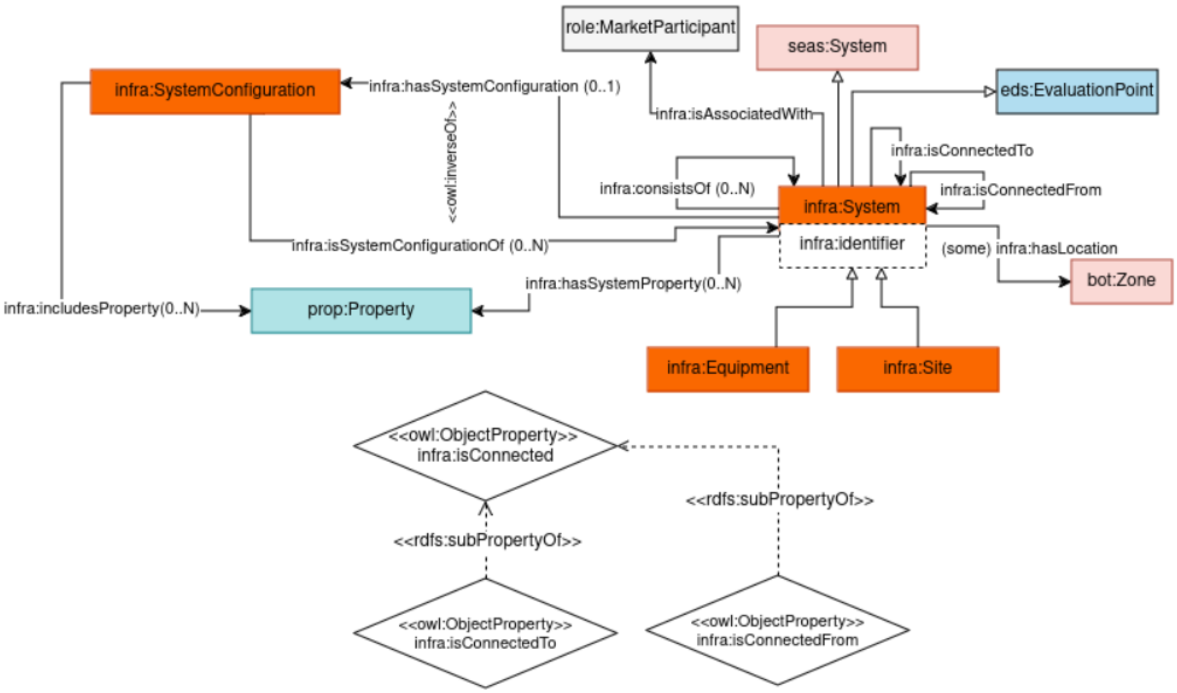

Figure 19: OMEGA-X CSDM Infra ontology [20].

Flexibility ontology

OMEGA-X CSDM's Flexibility (flex) ontology describes Flexibility operations (offers, request and orders) and their data. Infrastructures from its use case specific ontologies for electromobility, renewables and local energy communities like charging stations and renewable energy production systems can be considered to describe FlexibilityInfrastructure and cover different flexibility use cases. The ontology is inspired by Interconnect (flexibility operations and profiles), Saref4Bldg (for the building features like spaces and devices) and IEC62746 for grids.

Within PEDvolution FlexOffer data model will be used for flexibility use cases enabling efficient coordination and optimisation of energy demand and supply in sector-coupling scenarios, e.g., electricity, heating, cooling, and mobility sectors.

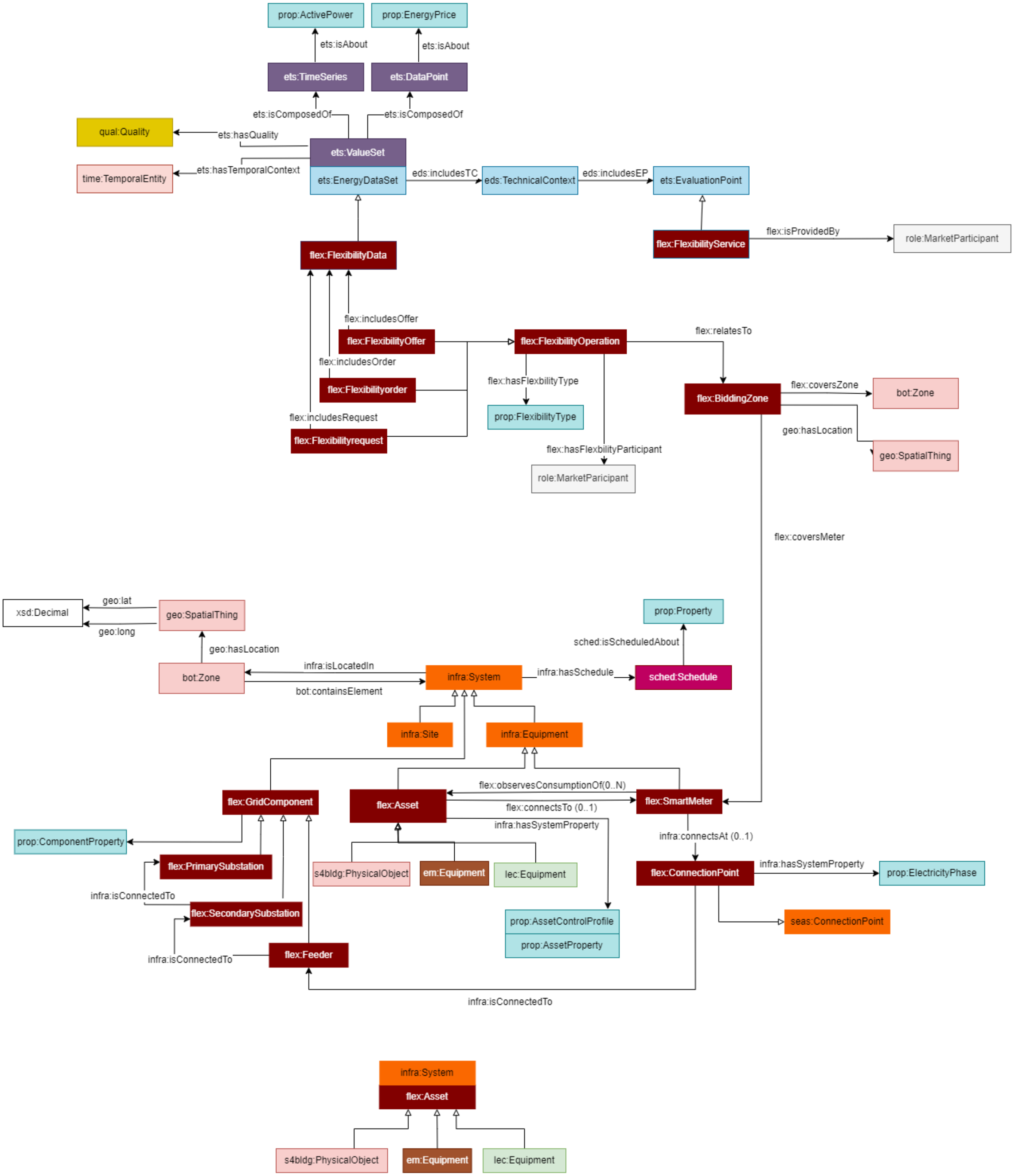

Figure 20: OMEGA-X CSDM Flex ontology [20].

Scheduling

The Scheduling (sche) ontology defines the tasks, events and time-based schedules, created by and optionally assigned to specific actors.

Similarly, FLEXOFFER will be utilised within PEDvolution for the case of schedules.

Renewables solar energy

The Renewables solar (solar) ontology describes solar plant information required towards executing a set of services. The ontology is influenced by Platoon, SEAS and SAREF4SYSTEM and includes the following key entities:

- SolarTracker: Controls the orientation of photovoltaic (PV) panels.
- SolarArray: Comprises multiple SolarString components, which are parallel connections of SolarPanel units.
- SolarModel: Consists of several SolarCell devices that convert light energy into electrical energy.
- InverterStation: Includes PowerInverter units that convert DC to AC power, potentially subdivided into PowerModule components.
- Transformer: Represents an electric power transformer.
- PowerCabinet: Houses electrical equipment.
- CombinerBox: A specialised cabinet for combining energy from multiple solar arrays.
- Substation: A collection of equipment used for switching or modifying electrical energy characteristics, excluding generation or utilisation.
- Feeder: A collection of equipment for organisational purposes, used for grouping distribution resources. The organisation of a feeder does not necessarily reflect connectivity or current operation state.
- Meter: It is the meter measuring the production of PV Plant.
- Finally, the ontology defines the WeatherStation measuring atmospheric conditions for making weather forecasts and climate studies.

Within PEDvolution solar ontology will be used to describe PV systems, while enriching weather stations information from Energy ADE for CityGML.

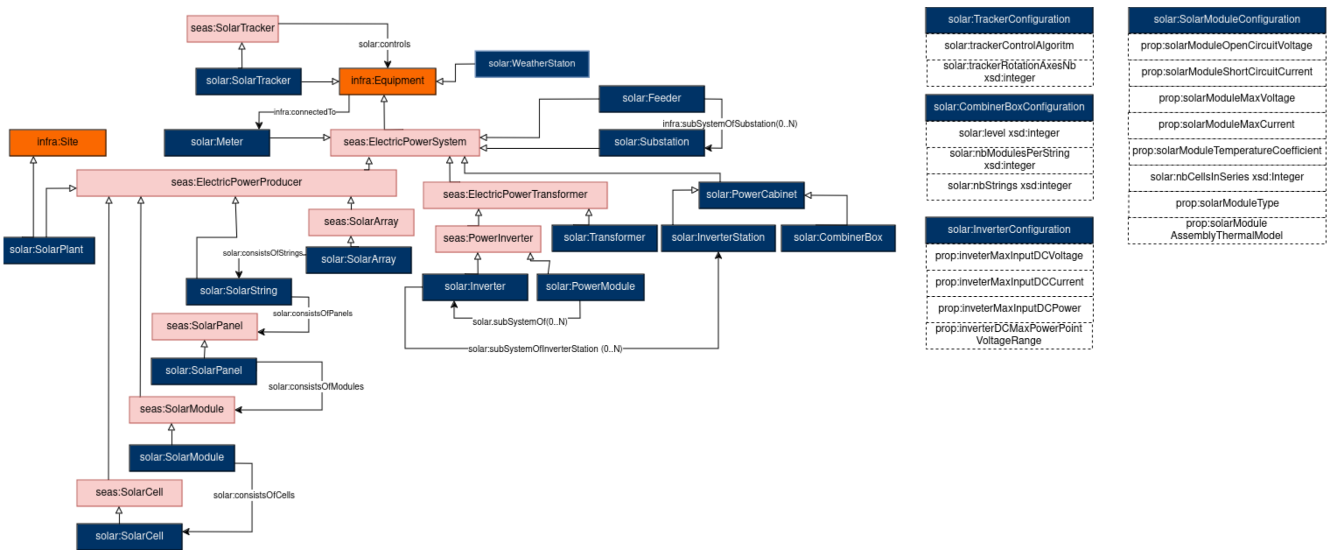

Figure 21: OMEGA-X CSDM solar ontology [20].

Local energy communities ontology

The Local energy communities (LEC) module describes local energy equipments and their static parameters such as batteries (Battery) and transformers (Transformer) and their configuration (batteryconfiguration, transformerconfiguration). The ontology is inspired by SEAS and PLATOON. The batteries definition from CSDM which aims to cover local energy communities may not fit well to the battery concept needed in PEDvolution to describe PEDs storage needs.

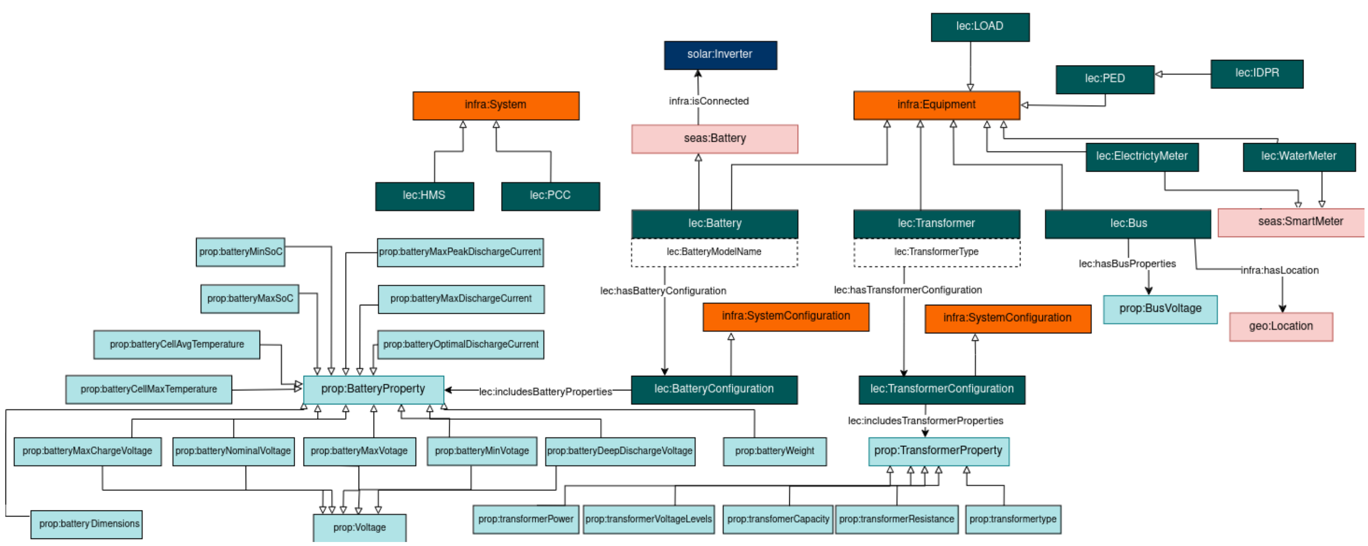

Figure 22: OMEGA-X CSDM LEC ontology [20].

Electromobility

The Electromobility (em) module is inspired by SEAS ontology and describes the charging infrastructure. Its key elements include EV charging pool, station, charging point and connector, while the property module describes their relevant characteristics.

EM module will be adopted by PEDvolution to describe EV charging infrastructure and complement where needed by BRICK data schema.

4.3.6 FlexOffer

The FlexOffer [33] Information Model is designed to represent and manage flexibility in energy consumption and production. It enables energy consumers and producers to specify their flexible energy usage or generation capabilities in a standardised format. By capturing the temporal and quantitative aspects of flexibility, such as energy amounts over specific time intervals, the model facilitates efficient communication between energy providers, consumers and market operators who aim at optimising energy use, balancing supply and demand in smart grids. This standardised representation supports advanced demand response programs, enabling optimised energy scheduling and improving grid stability. PEDvolution will utilise a FlexOffer extension, which accounts also for sector-coupling, namely considering flexibility provision in different energy vectors, such as electricity, heat, etc. The specifications of FlexOffer which will be utilised in the context of PED Energy Manager (WP5), are presented thoroughly in D2.3 [34].

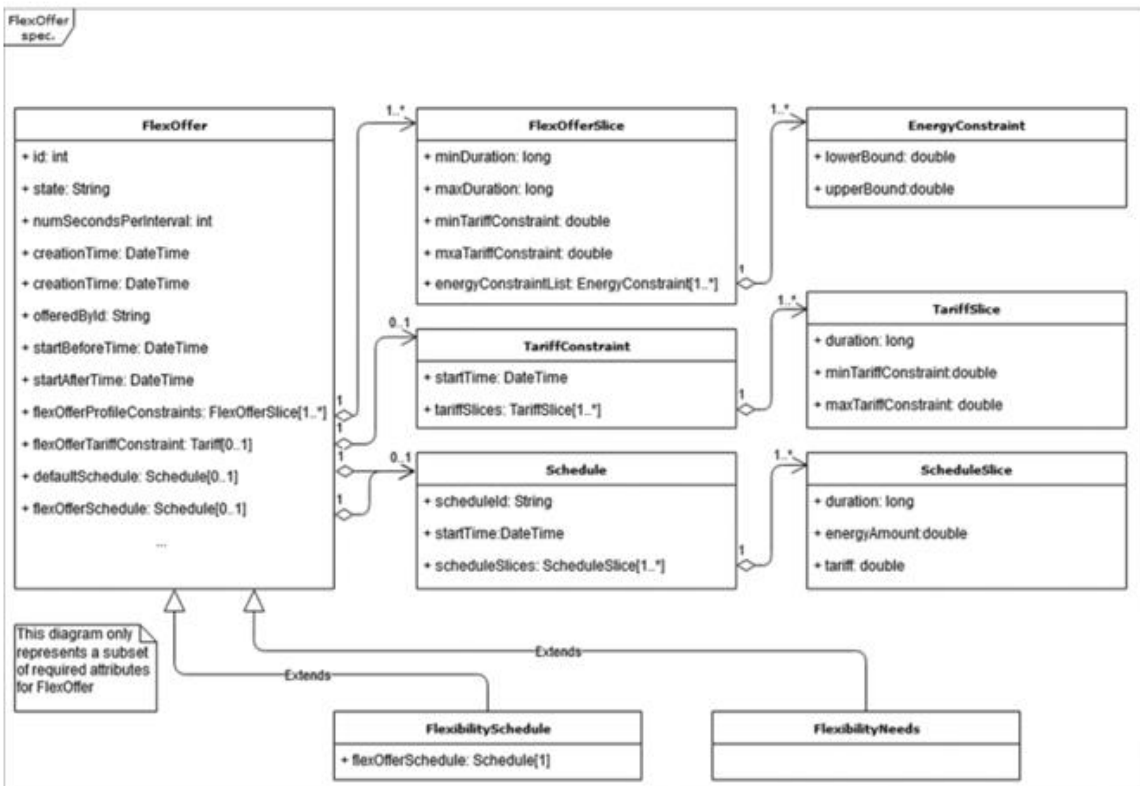

Figure 23: FlexOffer data model [33].

4.4 PEDvolution's Common Information Model

This section describes the formation of the semantics data model which is based on the analysis of the selected data models described in the previous section adopting the necessary classes, their properties, combining and extending them to cover PEDvolution's data needs. PEDvolution's semantic data model will be published and maintained in Github[3]. The common information model of PEDvolution will be continuously updated as the task progresses to cover identified requirements and its finalised version will be presented in D7.3.

The ontologies analysed in the previous section are listed in the table below (Table 5) grouping them based on the aspects and semantics areas they focus and according to the identified core data categories which are relevant to PEDvolution data needs for the PED conceptualisation, implementation, and development.

Table 5: Mapping of selected ontologies and covered data categories in the context of PEDs.

DATA CATEGORIES/ DATA MODEL	SAREF	CSDM	BRICK	ENERGYADE FOR CITYGML	BOT	FLEXOFFER
BUILDING CONSTRUCTION CHARACTERISTICS	●	●	●	●	●	
BUILDING OCCUPANCY PATTERNS			●	●		
BUILDING ENERGY ASSET CHARACTERISTICS	● (SAREF4BLDG)		●	●		
SOLAR PANEL CHARACTERISTICS		●	●	●		
BATTERY SYSTEM CHARACTERISTICS	● (SAREF4ENER)			●		
EV CHARGING CHARACTERISTICS		●	●			
HVAC SYSTEM CHARACTERISTICS			●	●		
DISTRICT HEATING & COOLING INFRASTRUCTURE				●		
DISTRICT GEOSPATIAL DATA	● (SAREF4CITY)		●	●	●	
DISTRICT TECHNICAL CHARACTERISTICS	● (SAREF4CITY)		●	●		
DISTRICT ECONOMIC CHARACTERISTICS	● (SAREF4CITY)			●		
DISTRICT SOCIAL CHARACTERISTICS	● (SAREF4CITY)	●				
CERTIFICATION AND ASSESSMENT DATA	●	●				
ENERGY MEASUREMENTS	●	●	●	●		
PROSUMER ENERGY PRICES	● (SAREF4ENER)					●
ASSET SCHEDULE AND FLEXIBILITY OPERATIONS	● (SAREF4ENER)			●		●
WEATHER AND CLIMATE DATA	●	●	●	●		

Building construction characteristics: Construction-related data like building geometry and materials, are useful for energy simulations. Bot, adopted by CSDM, provides minimal information for the building structure. Saref4BLDG describing the logical structure of the building with limited construction details can be utilised by PEDvolution towards a general building representation. Brick through its integration with REC describes building's elements and spaces and general building characteristics like the year built and its type, while some additional properties can be described utilising SAREF. However, Energy ADE for CityGML integrates detailed building construction data relevant to energy simulations that provide insights into how buildings consume energy based on their physical structural and material properties and by extending CityGML. As described in Section 4.3.2, it defines general characteristics for the building usage (e.g., the building type) and its different areas, number and size of rooms, floors etc., thermal properties, geometrical and locational parameters, as well as information regarding the

[3] https://github.com/PEDvolution/cim

building's renovations and energy performance certificate. Thus, Energy ADE will be utilised as the primary source of detailed construction data required within PEDvolution and enriched with any missing properties.

Building occupancy patterns: Building occupancy patterns and data are useful for optimising energy consumption and operational efficiency in buildings. As described in Section 4.3.2, the *Occupant Behaviour* module in EnergyADE can describe occupancy data which are of interest to PEDvolution e.g., through the *Household* and *Occupants* objects, as well as the *household* and *residence types*. It also describes occupancy schedules regarding when spaces are occupied and in what capacity supporting temporal profiles for characterising occupancy patterns over time e.g., hourly, daily, seasonal, and finally usage profiles describing parameters like number of occupants and expected usage scenarios e.g., peak hours based on different types of buildings. Within PEDvolution, Energy ADE will be the primary ontology of building occupancy data and patterns use. Additionally, Brick can complement the modelling with real time occupancy data such as the occupancy status and percentage of a zone.

Building energy assets characteristics: SAREF4BLDG and SAREF4SYSTEM can be combined to describe the energy assets providing a conceptual link between the building and its systems and equipment installed within it in a lightweight way. Brick and Energy ADE could describe more granular subsystem connectivity and flows which are not required within PEDvolution.

The energy assets which are of interest to PEDvolution project either at building or at district level are divided in the following categories covered by different ontologies:

Solar panels: OMEGA-x CSDM provides a detailed description of solar system through the solar (Renewables) ontology covering information about solar trackers, inverters, the combination of multiple solar arrays, as well as the collection of equipment utilised for management purposes, as mentioned in Section 4.3.5. CSDM can adequately cover PEDvolution needs. However, if further information is needed such as conversion efficiency, Brick properties could complement solar description Finally, the Energy ADE ontology with the *EnergyAde.solarthermalsystem* as previously mentioned in Section 4.3.2, contributes with additional properties such as the panel's inclination apart from the panel's azimuth.

Battery characteristics: Regarding the Battery system, Energy ADE will be firstly deployed to import storage system attributes with *EnergyADE.storagesystem* and *EnergyADE.powerstoragesystem* classes at building and district level. In addition, SAREF4ENER will cover the incorporation of missing data with Power, Voltage, Storage, Energy classes (see also Section 4.3.1). OMEGA-x CSDM prop ontology covers batteries characteristics however focusing on local energy communities (LEC) aspects that are not relevant to PEDs.

EV charging characteristics: Electromobility and Electrical vehicles are well-described in OMEGA-x CSDM ontology, as cited in Section 4.3.5, including the EV charging equipment (pool, station, charging points and connectors) and relevant charging properties in the prop ontology. Brick is an alternative ontology for EV charging.

HVAC characteristics: Concerning HVAC system, Brick ontology will primarily be used taking into account entities that represent physical, logical, and virtual assets within buildings. As mentioned in Section 4.3.4, the assets include equipment (e.g., air handling units), points (e.g., sensor readings, setpoints, alarms, commands/actuations), locations (e.g., rooms, floors), and logical collections (e.g., HVAC zones) appropriate to support and integrate HVAC characteristics. Then, Energy ADE will be additionally implemented to cover the gaps for heating or ventilation type devices data description with Energy Systems (Section 4.3.2) and its subclass *AbstractEnergyConversionSystem* providing detailed information regarding energy conversion systems, such as heat exchangers, chillers, heat pumps, boilers, etc.

Other aspects relevant to the district level are divided into the following categories, while the ontologies that can be employed to cover them are documented:

District Heating and Cooling infrastructure: Similarly for Heating and Cooling infrastructure as for HVAC systems, Energy ADE will be used. In addition, if there is need to report on relevant devices at building level, SAREF4BLDG can be deployed for device monitoring and power flow control with *s4bldg:DistributionControlDevice* and *s4bldg:DistributionFlowDevice* (see also Section 4.3.1).

District geospatial data: Geospatial data at district level within PEDvolution project will primarily rely on the SAREF4CITY ontology with *s4city:District/city/AdministrativeAreacontrolsProperty* class, which offers comprehensive support for representing and managing urban and spatial information. As already introduced in Section 4.3, at district level, datasets capture essential information about the district area (*s4city:AdministrativeArea*), such as its geographical boundaries, coordinates, population demographics, infrastructure layout, building locations, existing utilities and energy demand mechanisms. In cases where additional details or specific building-related data are required, the BOT ontology will be used complementarily to enhance the data model and ensure a more complete representation of the district-level geospatial information (Zone entity as described in Section 4.3.3).

District technical characteristics: For the district's technical characteristics, the Brick ontology will be utilised first, incorporating its aforementioned entities to represent physical, logical, and virtual devices that describe the data sources framework, including their role and location in relation with the district and the buildings' inner environment. Following this, the Energy ADE will be applied to provide detailed information on energy-related aspects, such as energy demand, systems, and performance (see Energy Systems in Section 4.3.2). Finally, the SAREF4CITY ontology can be additionally used to address any remaining data gaps, particularly those related to external facilities and infrastructure, ensuring a comprehensive representation of the district's technical characteristics (Section 4.3.1).

District economic characteristics: Although District Economic Characteristics are not directly accounted by the examined standards, ontologies, and data models, a combination of Energy ADE for CityGML and SAREF (especially SAREF4CITY) could be utilised as a basis to build upon them tailored extensions on district economic characteristics (e.g., LCC, land use efficiency, etc.), as these provide the most relevant foundations for economic and environmental risk indicators.

District social characteristics: Key ontologies from SAREF and CSDM will be deployed to measure and assess critical performance metrics, evaluating the success of a particular activity in which the society engages, such as community engagement, energy poverty reduction, social awareness and education, well-being, economic empowerment. Regarding *saref4city:keyperformanceindicator* and *keyperformanceindicatorassessment*, they provide robust tools to assess systems' efficiency and effectiveness of different activities, focusing on indicators like energy behaviour, mobility or environmental impact. In addition, *csdm:qual:qualitymetric, measurementprocedure*, and *aggregation* of CSDM ontology (see also Section 4.3.5), the project ensures a standardised framework to aggregate, measure and evaluate the quality of the various imported data. These ontologies collectively allow informed decision-making, fostering community engagement and citizen-centric district transformation.

Certification and assessment data: Similarly as in District Social Characteristics case, SAREF and CSDM ontologies can be used to evaluate performance, compliance and alignment with sustainability and energy goals, such as certification compliance, data quality and availability, lifecycle assessment and maintenance compliance.

Weather and climate data: SAREF includes basic support for environmental data, sensed by IoT devices (e.g., weather sensors covered by SAREF and SAREF4ENER or weather stations within the context of *SAREF4AGRI*). However, Energy ADE explicitly links to climatic zones, weather datasets, and temporal weather variations as described in Figure 11 by the *Supporting classes* module, which can be used as the primary source within PEDvolution by extending *Weather Stations* with *WeatherData* objects that describe time series of meteorological parameters. Additionally, weather stations modelled in CSDM refer to weather data (e.g., temperature) to further characterise the solar panels equipment. Finally, Brick covers dedicated weather stations to describe weather data useful for building systems, especially in managing HVAC and energy efficiency.

As regards energy measurements, they will be collected at different levels, namely prosumer, building, and district level, as explained in Section 4.2.

Energy measurements: OMEGA-x CSDM can be utilised for energy measurements, as it employs the EUMED Metering Ontology leveraging *eme:MeterReadingSet* to describe energy measurements. It is highlighted that SAREF can also be exploited to cover the aspect of energy measurements, e.g., through *saref:measurement, saref:meter, saref:UnitOfMeasure*, etc. Finally, it is noted that although EnergyADE and Brick are also capable of describing energy measurements by supporting timeseries data, they won't be utilised for this purpose, as CSDM is deemed to be more detailed and consistent with Data Space modeling.

Asset schedule and flexibility operations: For the asset scheduling and flexibility operations information model, FlexOffer will be the primary ontology utilised due to its robust capabilities in representing flexibility exchange, demand response and asset scheduling (Section 4.3.6). It effectively supports modeling flexibility operations, such as energy consumption patterns, scheduling and pricing structures. As a secondary option, SAREF4ENER (Section 4.3.1) can be employed to complement FlexOffer where necessary, providing additional support for energy-related data, including pricing, energy profiles, and operational details.

Prosumer energy prices: SAREF4ENER (see also Section 4.3.1 reference to Economic Data) can be used as an option for energy pricing providing further support in pricing models. The goal will be to enhance cost optimisation and maintain transparency in energy transactions. As regards flexibility exchange and demand response cases, FlexOffer will be prioritised, focusing on flexibility pricing and tariff declaration.

To conclude with, it is noted that the Energy Data Set ontology of CSDM (*eds*) will be employed to provide information on the technical and exchange contexts of PEDvolution data sets. Furthermore, the upper-level module of the CSDM on Event and Time Series (*ets*) can be utilised to capture the structure of the various PEDvolution datasets and their underlying elements. Based on ets, the data categories of Table 5 that are structured as *ets:TimeSeries* are building occupancy patterns, energy measurements, prosumer energy prices, weather forecasts and measurements, asset schedule and flexibility operations. The rest of the data categories (i.e., building construction characteristics, building energy asset characteristics, solar panel characteristics, battery system characteristics, EV charging characteristics, HVAC system characteristics, District Heating and Cooling infrastructure, district geospatial data, district technical characteristics, district economic characteristics, district social characteristics, certification and assessment data, climate data) are structured as *ets:DataCollection*. More information on both *eds* and *ets* can be found in the respective paragraphs of Section 4.3.5.

5 CONCLUSIONS

This deliverable presented the outcomes of the work performed in the first set of activities of T7.1 regarding (i) the design of the Interoperability Platform, (ii) the formation of a common semantics model based on analysis of state-of-the-art ontologies and standards and (iii) the integration of the different systems, devices and platforms in the PED ecosystems through the Interoperability Platform.

The Data Space paradigm embraced by other EU projects like OMEGA-X approach has been adopted for the design of PEDvolution's Interoperability Platform to facilitate secure communication and sovereign data exchange between the different systems and stakeholders in the PED ecosystems in a standards-based and decentralised way. The design methodology used followed the viewpoint architecture from D1.3 complementing the architecture modelling with the design of functional and process viewpoints to present the underlying building blocks and their interactions in more detail. The DS implementation will be built by leveraging well established protocols and standardised frameworks, using EDC as the foundational framework which implements technologies derived from IDS standards and Gaia-X principles.

Regarding data interoperability, a bottom-up approach has been used for building a semantic data model for PEDvolution based on analysis of the data needs and state-of-the-art ontologies and standards to describe a common vocabulary for PEDvolution's data exchanges and ensure precise meaning and interpretation of exchanged information by different solutions and stakeholders within the project. Semantics cover all different building and district characteristics from building construction data, occupancy patterns and their assets to the district's geospatial, social, technical and economic characteristics, while distinguishing between different categories of energy infrastructures and assets and their characteristics such as batteries, HVAC systems and DHC infrastructures, and renewables like solar panels. SAREF family and CSDM from OMEGA-X project comprise the core of the data model, enriched with parts from Energy ADE for CityGML and BRICK to complement missing entities or attributes, while the part of the energy flexibility is covered by FLEXOFFER used in other EU projects and can be aligned with existing standards like SAREF4ENER. Syntactic interoperability on the other hand is handled by the adopted Data Space approach leveraging common standards like JSON for the data schema to be used.

The selected interoperability approach including the design of the Interoperability Platform and the common information model will comprise the basis for the integration of the different systems, platforms, and assets in PEDvolution (T7.1/T7.3) in the subsequent WP activities. Development of PEDvolution solutions in the context of WP3-6 should also be aligned with the CIM, so that interoperability amongst all systems and platforms can be achieved. The solution will be enriched with mechanisms for interfacing with the energy market and federation services for the interconnection to the wider EU Data Space ecosystem in T7.2 and T7.4 respectively. Finally, the semantics data model will be finalised based on other pilot needs. The final prototype will support the pilot activities in WP8 and WP9. As regards the remaining WP deliverables, D7.2 will report on the developed mechanisms for interfacing with energy markets, while D7.3 will document the final specifications of PEDvolution's Interoperability Platform, thus capturing essentially all the work to be conducted in WP7.

6 REFERENCES

[1] PEDvolution project, "Description of the Action - part B," 2023.

[2] "European Data Space Initiative," [Online]. Available: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12491-Data-sharing-in-the-EU-common-European-data-spaces-new-rules-_en.

[3] "International Data Space Association," [Online]. Available: https://internationaldataspaces.org/.

[4] International Data Spaces Association, "IDS Reference Architecture Model, version 4.0," https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/, 2022.

[5] International Data Spaces Association, "Advancing interoperability: The Dataspace Protocol," [Online]. Available: https://internationaldataspaces.org/offers/dataspace-protocol/. [Accessed 10 October 2024].

[6] GAIA-x, "https://docs.gaia-x.eu/framework/," [Online].

[7] "OMEGA-X European project," [Online]. Available: https://omega-x.eu/.

[8] PEDvolution project, "D1.3 - System architecture and technical specifications," 2024.

[9] PEDvolution project, D1.2 - Functional and Operational Requirements of the Demo Sites and Reference Use cases, 2024.

[10] European Union, *General Data Protection Regulation,* https://eur-lex.europa.eu/eli/reg/2016/679/oj, 2016.

[11] Swiss Federal Assembly, *Federal Act on Data Protection,* https://www.fedlex.admin.ch/eli/cc/2022/491/en, 2020.

[12] PEDvolution project, "D13.2 - Data Management Plan," 2024.

[13] European Commission, "Common European Data Spaces," [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/data-spaces. [Accessed 10 October 2024].

[14] European Parliament, European Council, "Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance)," 13 December 2023. [Online]. Available: https://eur-lex.europa.eu/legal-

content/EN/TXT/?uri=CELEX%3A32023R2854&qid=1704709568425. [Accessed 17 December 2024].

[15] EDC, "Eclipse Dataspace Components. Retrieved from https://github.com/eclipse-edc," [Online].

[16] IDSA, "Internation Dataspace Protocol," [Online]. Available: https://docs.internationaldataspaces.org/ids-knowledgebase/dataspace-protocol.

[17] [Online]. Available: https://www.w3.org/RDF/.

[18] [Online]. Available: https://www.w3.org/OWL/.

[19] [Online]. Available: https://www.w3.org/TR/rdf-schema/.

[20] "Omega-X Common Semantic Data Model," [Online]. Available: https://fz-hannou.github.io/Omega-X/.

[21] "SAREF Core ontology," [Online]. Available: https://saref.etsi.org/core/v3.2.1/.

[22] "SAREF4ENER Ontology," [Online]. Available: https://saref.etsi.org/saref4ener/v1.2.1/.

[23] "SAREF4BLDG Ontology," [Online]. Available: https://saref.etsi.org/saref4bldg/v1.1.2/.

[24] "SAREF4CITY Ontology," [Online]. Available: https://saref.etsi.org/saref4city/v1.1.2/.

[25] "CityGML Energy ADE," [Online]. Available: https://github.com/citygml4j/energy-ade-citygml4j/releases.

[26] "CityGML standard," [Online]. Available: https://github.com/citygml4j/citygml4j/tree/citygml4j-v2.

[27] "INSPIRE Directive," [Online]. Available: https://knowledge-base.inspire.ec.europa.eu/index_en.

[28] "Building Energy Data Exchange Specification (BEDES)," [Online]. Available: https://www.energy.gov/eere/buildings/building-energy-data-exchange-specification-bedes.

[29] "Building Topology Ontology," [Online]. Available: Building Topology Ontology .

[30] "Brick," [Online]. Available: https://brickschema.org/.

[31] "INTERCONNECT EU project," [Online]. Available: https://interconnectproject.eu/.

[32] "Semantic Data Models of Energy," [Online]. Available: https://github.com/PLATOONProject/SEDMOON.

[33] FlexOffer, "FlexOffer principles," [Online]. Available: https://flex-community.eu/data/groups/documents/D09FE8EF-9F2F-4813-943F-82AFA9F3ABDE_FlexOffer_principles_v1.0.pdf. [Accessed 10 October 2024].

[34] PEDvolution project, "PEDvolution D2.3 - Solution specification and ideation for the PED energy manager," 2024.