

Optimizing energy simulation processes through an ontology-based approach

Claudio Mirarchi¹, Jacopo Cassandro¹, Massimiliano Manfren¹, Claudio Del Pero¹

¹ Department of Architecture, Built Environment, and Construction Engineering (DABCE), Polytechnic of Milan, Milan, Italy

Corresponding author: claudio.mirarchi@polimi.it

Abstract— Building Energy Modelling (BEM) traditionally faces significant interoperability challenges due to manual model simplifications and the fragmented nature of data sources, needed in the modelling process. This paper presents SPHEERE, an innovative framework for ontology-based semantic integration that automates the translation of Industry Foundation Classes (IFC) into a temporary energy-oriented model. The methodology defines an Energy Information Model (EIM) to identify parameters across seven categories, only 40% of which are natively present in IFC. We utilize a layered ontological stack, including ifcOWL, BOT, SAREF4BLDG, and BRICK, to generate a federated Knowledge Graph (KG). This KG is further enriched through semantic derivation and external integration of climatic and operational data, creating a queryable "lean digital twin" without the need for a dedicated, isolated BEM. The framework is validated via a case study involving federated as-is and to-be retrofit models. Results demonstrate that SPARQL queries can successfully retrieve parameters that are relevant to drive the modelling process. This approach preserves the original BIM as the single source of truth while providing a transparent, auditable infrastructure for building performance modelling and optimization. The contribution lies in bridging the gap between geometric modeling and physics-grounded energy analysis through machine-interpretable semantic views.

Keywords—product and process modelling; BIM; digital construction; interoperability; digital twin; physics-informed ML, ontologies

I. INTRODUCTION

The design and retrofit of buildings require specialized analyses such as structural, energy, lighting and acoustic calculations, among others, each demanding specific levels of abstraction. While Building Information Modelling (BIM) and IFC have improved cross-disciplinary collaboration, practical interoperability between architectural models and calculation engines remains a bottleneck. Current workflows often rely on ad hoc model simplifications, leading to information silos, synchronization issues, and a lack of scalability.

Furthermore, there is an increasing demand for interpretable and physics-informed AI/ML to address the "design-operation performance gap". Modern energy management requires "digital twins" capable of providing traceable Key Performance Indicators (KPIs) that align with regulatory frameworks like the 2024 EPBD recast and related technical standards. To achieve this, the industry needs a semantic infrastructure that acts as the "connective tissue" between raw data and physically meaningful model parameters.

This study addresses these challenges by investigating:

- **RQ1:** The automated translation of IFC into energy simulation-ready ontological models.
- **RQ2:** The management of calculations within an ontology-based environment.
- **RQ3:** The practical application of semantic interoperability.

The primary objective is to evaluate a framework that transforms IFC into a temporary energy-oriented model. By leveraging a federated Knowledge Graph (KG), we integrate heterogeneous information, including geometric, physical, environmental, and operational parameters, while maintaining the BIM model as the single source of truth. This approach shifts the paradigm from generating isolated analysis models toward creating domain-specific semantic views that are queryable, reusable, and physically consistent.

II. RELATED WORK

Building performance modelling increasingly relies on the integration of standardized ontologies and interpretable data-driven approaches. In this context, the W3C Linked Building Data (LBD) ecosystem provides the semantic foundation for representing heterogeneous energy-related information, while energy-related information supports its organization and exploitation within BIM-to-BEM workflows.

Linked Building Data and Ontologies

The adoption of Semantic Web technologies within the AECO sector has led to the development of a rich ecosystem of ontologies supporting the representation, integration and querying of building information. The translation of BIM data into RDF through ifcOWL enables IFC models to be represented within a semantic environment while preserving their original structure and relationships [1]. The Building Topology Ontology (BOT) has become the standard for representing the building topology, including sites, buildings, storeys and spaces, while BEO provides a lightweight classification of building elements suitable for Linked Building Data applications [2]. More recent research has demonstrated the use of ontology-based approaches to support building performance modelling, semantic enrichment and multi-domain interoperability through the integration of BIM data with additional information sources [3,4]. Beyond BIM representation, several domain-specific ontologies have been proposed to model operational and performance-related information. BRICK provides a semantic schema for HVAC systems, sensors and operational control points, while SAREF4BLDG extends the SAREF framework to building-related concepts and energy properties. SOSA/SSN and GeoSPARQL enable the integration of observations, monitoring data and geospatial information, allowing heterogeneous information sources to be represented within a common semantic framework. Together, these ontologies form the basis of the Linked Building Data ecosystem and provide the semantic

infrastructure necessary to support interoperability across multiple building-related domains.

Interpretable and physics-informed data-driven modelling hierarchies

Regarding the simplification of detailed building energy simulation models while maintaining the physical meaning of the underlying quantities, various modeling approaches have been developed and reviewed, such as by Kavgić et al. [5], Foucquier et al. [6], and Chen et al. [7]. These reviews and other studies categorize models by their transparency: “white-box” models are detailed, physics-driven simulations; “grey-box” models combine physical and statistical methods; and “black-box” models rely purely on data-driven techniques like statistics and machine learning (ML). This distinction has become blurred today because statistical and machine learning techniques can now possess physical interpretation, address causality, and be generally interpretable—meaning understandable by humans (not a “black-box”), while still delivering satisfactory performance for the analytical task.

Recent research examples highlight how change-point (CP) regression models can be combined with ARX/ARMAX techniques to support at scale applications such as heat and electricity load profiles in buildings [8] and automated building performance certification [9]. Similar principles were leveraged in the EU H2020 project HEART for parametric energy performance analysis and monitoring of buildings [10]. Further, Time-of-Week-and-Temperature (TOWT) algorithm [11,12] can complement CP regression models in digital twin frameworks [13], to tackle dynamic building energy behaviour at scale [14].

Overall, these methods generate physically labeled coefficients that offer the diagnostic power necessary for ranking design retrofit options and optimizing operational performance. However, these coefficients often exist in isolation; combining them with a semantic framework and interpretability could significantly enhance the practical effectiveness of the modeling approach.

BIM-to-BEM semantic integration and further developments

Several studies have investigated the use of semantic technologies to support BIM-based performance analysis and BIM-to-BEM integration. Existing research demonstrates that graph-based and topology-oriented BIM representations can support advanced querying, spatial reasoning and performance-related analyses through the explicit representation of topological relationships among building elements [15,16,17]. More recently, semantic enrichment approaches have been proposed to reconstruct implicit information within BIM models and support higher-level reasoning, performance assessment and automated compliance checking [18].

Furthermore, while semantic technologies have been successfully applied to the integration of cost, time, resource and geometry information through federated knowledge graphs and advanced SPARQL querying [19,20], similar approaches remain relatively unexplored within the energy domain. This research gap motivates the development of the proposed Energy Requirements Model (EIM) and the generation of federated knowledge graphs capable of supporting energy-oriented semantic views through the integration of heterogeneous information sources.

Although significant progress has been achieved in semantic BIM representation, topology reconstruction and performance-oriented data integration, existing approaches typically focus on either the semantic representation of building information or the extraction of specific performance indicators.

III. METHODOLOGY

This paper proposes an ontology-based framework aimed at supporting the generation of a temporary energy-oriented model for building energy analysis starting from IFC-based information models and specific information stored in separate silos. Traditional approaches often require the creation of simplified or domain-specific Building Energy Models (BEM), leading to information duplication, synchronization issues and additional modelling efforts.

The proposed methodology adopts a different perspective. Instead of generating new analysis-specific models, the framework uses the IFC model as the primary information source and enriches it through semantic technologies. The objective is to create a semantic representation of the building capable of integrating heterogeneous information originating from multiple sources while preserving the original BIM model as the single source of truth.

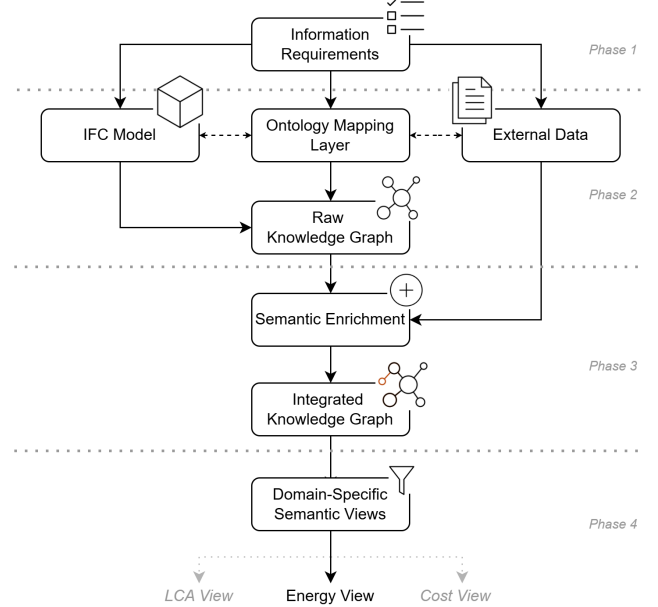


Fig. 1. Methodology workflow.

As illustrated in Figure 1, the proposed framework is structured into four main phases.

The methodology starts with the definition of EIM describing the parameters, relationships and data sources required to support energy-related assessments. A comprehensive set of information requirements was identified, including building geometry, envelope characteristics, thermal zones, building systems, operational parameters and environmental conditions. For each parameter, the potential source of information was identified, distinguishing between BIM-derived data, external datasets and user-defined inputs.

The second phase focuses on the semantic translation of the IFC model. Building information available in the BIM model is mapped into a KG using ontologies such as ifcOWL, BOT. This process generates a semantic representation of the building while preserving both the original BIM structure and the relationships among entities.

The third phase consists of semantic enrichment and data integration. Additional information required by the EIM is progressively connected to the KG through semantic relationships. Examples include the association of climatic information derived from building location, the integration of operational parameters, the classification of heated and unheated spaces and the aggregation of building components into energy-relevant systems. This phase enables the explicit representation of dependencies among geometric, physical, environmental and operational parameters.

The resulting enriched KG acts as a temporary energy-oriented model that provides a unified semantic representation of all the information required by energy assessment processes. Rather than generating a dedicated Building Energy Model, the proposed approach creates domain-specific semantic views that organize and contextualize the available information according to the requirements of a specific application domain. Such a representation improves interoperability between heterogeneous information sources and provides the basis for future energy assessment workflows.

IV. FRAMEWORK IMPLEMENTATION

Energy Information Model

The EIM constitutes the foundational phase of the proposed framework. Prior to any semantic translation, the EIM defines the complete set of parameters required for building energy assessments, their expected units, the origin of each datum, and the target ontological representation. This preliminary definition ensures that the subsequent KG is built purposefully rather than as a generic IFC conversion.

The EIM was structured around seven thematic categories: Identification, Location, Building Geometry, Construction Technology, Operational Settings and Control, Climate Data, and Pre-retrofit Building Performance. For each parameter, four attributes were recorded: (i) the parameter identifier and unit of measurement, (ii) the primary information source classified as IFC-native, Derived, External dataset, or Manual input (iii) the target ontology classes, and (iv) the specific IFC property set or ontology property to which the parameter is mapped. Table I presents a representative subset of the identified parameters.

Category	Parameter	Source	Ontology Classes
Identification	Project ID	IFC-native	ifc:IcfProject
Location	Location	IFC-native	ifc:refLatitude ifc:refLongitude
Climate	Climate data file (EPW)	External	sosa:Observation
Building geometry	Orientation	Derived	ifc:IcfDirection
Building geometry	U Exterior wall (Weighted average value)	IFC-native	ifc:IcfWall
Operational settings	Heating set-point temperature (°C)	Manual	brick:Zone_Air_Heating_Temperature_Setpoint

Table I. Representative subset of EIM parameters.

The classification of information sources reveals a significant finding: only approximately 40% of the required parameters are directly retrievable from the IFC model as-is. The remaining parameters require either geometric derivation (façade orientation, Window-to-Wall Ratio, shading factors), integration of external datasets (climate data via EPW files), or manual user input (setpoints, schedules, pre-retrofit consumption data). This distribution directly justifies the multi-phase semantic enrichment beyond simple IFC

conversion strategy described in subsequent subsections. Finally, the calculation engines used in dynamic building performance simulation are usually validated according to standards such as ASHRAE 140-2023 BESTEST and can leverage standardized procedures such as the ones described in ISO 52016-1:2017. The use of BESTEST cases, such as 600, 640, 900, and 940, out of a larger variety of cases in the standard, is generally considered the essential validation step for the dynamic simulation process and is considered as a reference in the implementation.

Ontology Mapping Strategy

The ontology mapping strategy establishes the formal correspondence between IFC entities and properties on one side, and semantic resources expressed in RDF/OWL on the other. The selection of ontologies was guided by three criteria: coverage of the target domain, maturity and community adoption, and alignment with existing IFC-based semantic resources.

The following ontologies were selected as the core vocabulary of the KG:

- ifcOWL, the OWL serialisation of the IFC schema (ISO 16739), used to preserve original BIM entity structure and property sets during translation;
- BOT (Building Topology Ontology), a W3C Linked Building Data community standard representing topological relationships among building components at the levels of site, building, storey, space, and element;
- BEO (Building Element Ontology), used to provide a lightweight semantic classification of building elements enabling a more tractable representation of construction components compared to full IFC-based representations;
- SAREF4BLDG (ETSI TS 103 410-6), the ETSI extension of SAREF for buildings, covering building objects, thermal and energy properties, and HVAC systems;
- BRICK Schema, used to represent operational control points, HVAC systems, sensors, and temporal schedules through classes such as `brick:HVAC_System`, `brick:AHU`, `brick:Schedule`, and `brick:hasPoint`;
- SOSA/SSN (W3C), adopted for representing climate observations and metered consumption data as `sosa:Observation` instances with `sosa:hasResult` linkage;
- GeoSPARQL, used to represent spatial geometries and geographic location;
- QUDT, used to express units of measurement ($\text{W/m}^2\text{K}$, m^2 , m^3 , $\text{kJ/m}^2\text{K}$, vol/h) in a machine-readable and dimensionally consistent form.

The ontology mapping strategy preserves IFC-native information, including thermal and operational properties, and mapped to the most appropriate ontology. Semantic enrichment is applied only to information unavailable in IFC or requiring geometric derivation or external integration.

Table II illustrates representative mappings between IFC and their corresponding ontology classes and properties.

IFC Entity / Property	Ontology	Mapped Class/Property
IfcBuilding	BOT	bot:Building
IfcBuildingStorey	BOT s	bot:Storey
IfcSpace	BOT	bot:Space
IfcWall	BOT / ifcOWL	bot:Element / ifc:IfcWall
IfcWindow	ifcOWL / SAREF4BLDG	ifc:IfcWindow / s4bldg:Window
Pset_WallCommon. ThermalTransmittance	SAREF4BLDG	s4bldg:thermalTransmittance
Thermal zone name	BOT / IFC	bot:Space / bot:Zone

Table II. Representative ontology mappings.

Where direct mappings between IFC properties and ontology classes were not available, intermediate RDF resources were introduced to represent derived or aggregated concepts. This was particularly relevant for energy-oriented groupings such as thermal zones and envelope systems, which do not have a direct structural counterpart in the IFC.

Two custom properties were introduced where no established ontology property covers a required concept: `ex:WindowWallRatio` for the per-orientation Window-to-Wall Ratio (WWR) KPI, and `ex:infiltrationRate` for the space-level air infiltration rate. These are defined within a project-specific namespace and documented in the EIM to ensure reproducibility.

A deliberate design choice was made regarding the handling of façade orientation. Although IFC provides the geometric information required to determine element orientation, no standard ontology property explicitly represents the cardinal orientation (N/S/E/W) of façade elements. Therefore, the azimuth of each external element is derived from its geometric placement, materialised as the custom property `ex:azimuth`, and subsequently classified into cardinal orientation groups to support orientation-dependent analyses such as WWR computation.

Knowledge Graph Generation

The KG generation phase translates the IFC model into an RDF graph according to the mappings defined in the previous subsection. The process operates on the IFC file in STEP format and proceeds through three sequential steps.

The pipeline consists of three sequential steps. First, relevant IFC entities are extracted from the STEP file: `IfcProject`, `IfcSite`, `IfcBuilding`, `IfcBuildingStorey`, `IfcSpace`, and element classes (`IfcWall`, `IfcRoof`, `IfcSlab`, `IfcWindow`, etc.). Second, each entity is instantiated as an RDF resource using its corresponding ontology class, and object properties represent relationships between resources. Third, datatype properties are populated with literal values extracted from Property Sets and Quantity Sets.

The resulting raw KG preserves the hierarchical BIM topology through the BOT ontology (`bot:containsElement`, `bot:hasSpace`, `bot:adjacentElement`) and attaches all directly available thermal and geometric properties from the IFC source. The following excerpt (Listing I) illustrates the RDF representation of an external wall with its thermal property.

At this stage, the graph contains exclusively information present in the IFC source model. Parameters classified as Derived, External, or Manual in the EIM remain as unresolved placeholders addressed in the following phase.

```
@prefix bot:<https://w3id.org/bot#>.
@prefix
ifc:<https://standards.buildingsmart.org/IFC/DEV/IFC4/ADD2/OWL#>.
@prefix s4bldg:<https://saref.etsi.org/saref4bldg/>.
@prefix qudt:<http://qudt.org/schema/qudt/>.
@prefix ex:<https://example.org/building#>.
ex:Wall_ExtNorth_012
  a ifc:IfcWall, bot:Element ;
  rdfs:label "External Wall North-East" ;
  s4bldg:thermalTransmittance [
    qudt:numericValue "0.28"^^xsd:decimal ;
    qudt:unit qudt:W-PER-M2-K ] ;
  ex:azimuth "15.3"^^xsd:decimal .
ex:Space_Office_01
  a ifc:IfcSpace, bot:Space ;
  rdfs:label "Office 01 - Level 2" ;
  bot:containsElement ex:Wall_ExtNorth_012,
ex:Window_007 ;
  ifc:isHeated "true"^^xsd:boolean .
```

Listing I. RDF representation of an external wall with its thermal property

Semantic Enrichment

The semantic enrichment phase progressively integrates into the KG all parameters identified in the EIM that are not directly available from the IFC model. Enrichment operations are classified into three types: derivation, external integration and classification.

Derivation computes parameters through geometric or spatial reasoning applied to existing graph resources. The WWR is computed per cardinal orientation by aggregating `ex:WindowWallRatio` values from windows and walls sharing the same `ex:azimuth` classification. Façade orientation grouping (N/S/E/W) is assigned through a SPARQL CONSTRUCT rule that maps azimuth ranges to orientation labels. Thermal capacity per net surface area is derived according to the ISO 52016 classification procedure. Shading factors are assigned from ISO lookup tables based on window orientation and glazing type.

External integration links the KG to datasets originating outside the BIM model. Climatic data, including the EPW file reference, Heating Degree Days, and design outdoor temperatures, are associated with the building via the `ifc:IfcSite` geographic coordinates (`ifc:refLatitude`, `ifc:refLongitude`) mapped to `geo:hasGeometry`. A `sosa:Observation` instance is created for each climatic parameter and linked to the building's site node, preserving full traceability to the external source.

Manual annotation is applied only when the required information is unavailable in the IFC model or cannot be derived geometrically. IFC-native thermal and operational information is preserved and semantically mapped whenever available. When not available in the source IFC model, typical examples include HVAC control points (heating and cooling setpoints represented as `brick:Zone_Air_Heating_Temperature_Setpoint` via `brick:hasPoint`), operational schedules for heating, cooling, internal gains, and ventilation (represented as `brick:Schedule` instances).

The output of this phase is the integrated KG illustrated in Figure 3. The graph constitutes the temporary energy-oriented model proposed in this paper: a unified, queryable RDF representation of all EIM parameters, expressed through the adopted ontology vocabulary, and structured to support domain-specific semantic views through targeted SPARQL queries. The Energy View, one such view, extracts the subset of parameters required for an energy assessment workflow without modifying the underlying graph structure, preserving the IFC model as the single source of truth throughout.

V. RESULTS

Case Study

The proposed framework was validated on a practical case study consisting of two independent IFC 4 ADD2 TC1 models representing different phases of the same building (Figure 2). The first model describes the existing building configuration (as-is), while the second model represents the retrofit intervention (to-be) consisting of an external thermal insulation system applied to the building envelope.

The as-is model includes two storeys, 65 wall elements (40 classified as external walls), 40 windows, one roof element and ten floor slabs. The retrofit model contains 16 insulation elements represented as `IfcCovering` instances with predefined type `INSULATION`. Unlike the as-is model, which mainly contains geometric and topological information, the retrofit model provides thermal performance information associated with the insulation layer, including thermal transmittance values associated with the insulation system.

This configuration was selected to demonstrate the capability of the proposed framework to integrate information distributed across multiple BIM models without generating a dedicated BEM. Previous research demonstrated the feasibility of establishing semantic relationships between independent as-is and to-be IFC models through ontology-based linking mechanisms and topology-driven semantic integration, enabling the aggregation of properties distributed across separate BIM representations [21]. Building upon this principle, the proposed framework associates thermal properties contained in the retrofit model with the corresponding building elements represented in the as-is model, creating a unified semantic representation of the building envelope while preserving model. Both models were authored in Revit and exported to IFC. The U-value of the external envelope ($0.35 \text{ W/m}^2\text{K}$) is derived from to-be model and integrated into the KG.

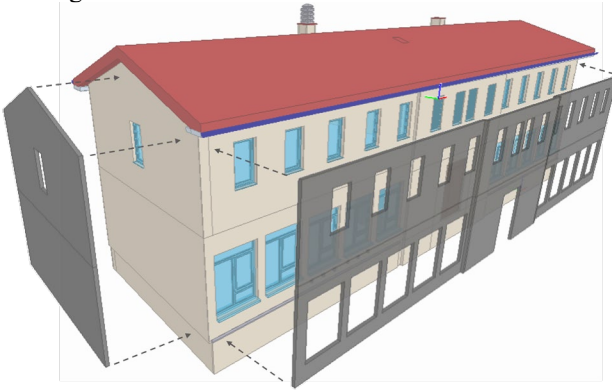


Fig. 2. Case Study - as-is and to-be IFC models.

Knowledge Graph

The two IFC models were translated into RDF KG following the ontology mapping strategy described in Section IV. Rather than producing a single monolithic graph, the framework generates four distinct graph layers that are then federated into a unified knowledge base: (i) the as-is BIM graph, (ii) the to-be BIM graph, (iii) the cross-model adjacency graph, and (iv) the energy enrichment graph. This layered architecture preserves the independence of the original BIM models while enabling semantic integration across heterogeneous information sources.

Figure 3 illustrates a representative subset of the federated knowledge graph. The figure highlights the interaction between the four graph layers and demonstrates how information originally distributed across independent IFC models and external sources becomes accessible through a unified semantic representation.

The as-is KG contains the building topology extracted from the original IFC model, represented through `ifcOWL` and `BOT`. The retrofit KG contains the external insulation system represented classified through `BEO`.

The cross-model adjacency graph acts as a semantic integration layer between the two BIM representations. A `bot:Interface` instance links the existing wall element to the corresponding insulation covering through two `bot:interfaceOf` assertions, enabling explicit machine-interpretable relationships between elements belonging to different IFC models without requiring model merging.

The energy enrichment graph integrates all parameters not available from the IFC models through three categories of enrichment:

- Semantic aggregation, which associates thermal properties stored in the retrofit model with the corresponding envelope elements represented in the as-is model. Thermal transmittance values are represented through `SAREF4BLDG` properties and `QUDT` quantity values while maintaining traceability to the originating insulation layer.
- Geometric derivation, which generates energy-oriented indicators from BIM geometry. This includes façade orientation derived from IFC placement information through azimuth calculation, orientation-based façade grouping, and WWR indicators represented as semantic resources linked to the corresponding façade entities.
- External and operational integration, which introduces climatic observations and operational parameters not available in the BIM models. Climatic data are represented through `SOSA` observations connected to the building site, while operational settings, such as heating and cooling setpoints, are represented through `BRICK` control points associated with `BOT` zones derived from the building spaces.

The resulting federated knowledge graph provides an explicit representation of the relationships among building topology, envelope systems, thermal properties, climatic information and operational parameters. Unlike traditional BIM-to-BEM workflows, no dedicated analysis model is generated. Instead, the knowledge graph acts as a semantic integration layer capable of supporting domain-specific semantic views while preserving the original IFC models as the single source of truth.

The distribution of triples across ontology namespaces reflects the multi-domain nature of the framework. `BOT` provides the topological backbone of the building representation, `BEO` supports element classification, `SAREF4BLDG` captures thermal properties, `QUDT` ensures dimensional consistency values, `BRICK` represents operational information, and `SOSA` enables the integration of external observations. Custom properties are used only where no established ontology concept was available, such as façade orientation, azimuth and WWR indicators.

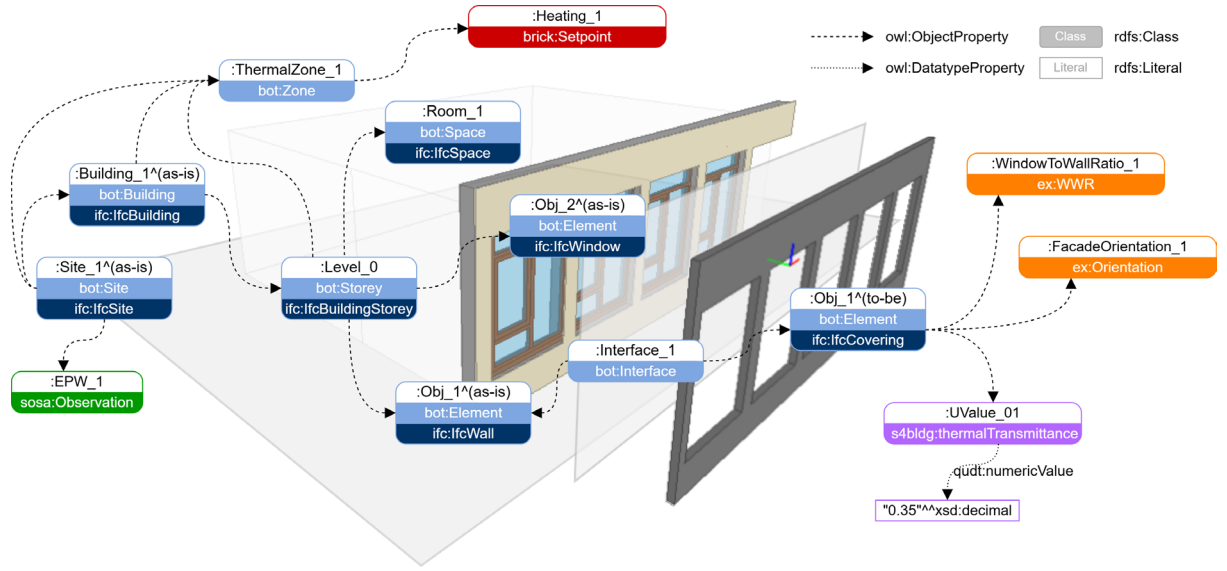


Fig. 3. Federated knowledge graph integrating BIM-derived, operational, climatic and derived information.

SPARQL Query

The federated KG can be queried through SPARQL queries to generate domain-specific semantic views. Two representative queries are presented to demonstrate the energy-oriented semantic views extractable from the graph without modifying the underlying data model.

The first query (Listing II) demonstrates the capability of the federated knowledge graph to retrieve thermal properties distributed across independent BIM models. Rather than performing a numerical aggregation, the query traverses the semantic relationships established during the enrichment process to reconstruct the envelope system composed of the existing wall and the corresponding retrofit insulation layer.

The query starts from the `bot:Interface` instances contained in the cross-model adjacency graph. Each interface represents a semantic relationship connecting an element belonging to the as-is BIM model with the corresponding insulation covering represented in the retrofit BIM model. Through the `ex:hasRetrofitLayer` relationship, the query retrieves the insulation element associated with the wall and accesses the thermal transmittance value materialized during the enrichment phase as a `s4bldg:thermalTransmittance` property represented through a `qudt:QuantityValue`.

In this way, the query combines information originating from three different graph layers: the wall geometry contained in the as-is BIM graph, the insulation component represented in the retrofit BIM graph, and the thermal property generated in the energy enrichment graph. The resulting semantic view reconstructs the envelope system without requiring IFC model merging or the generation of a dedicated BEM.

```

PREFIX bot:<https://w3id.org/bot#>
PREFIX ex:<https://caseStudy/ontology#>
PREFIX s4bldg:<https://saref.etsi.org/saref4bldg/>
PREFIX qudt:<http://qudt.org/schema/qudt/>
SELECT ?interface ?wall ?insulation ?uValue
WHERE {
  ?interface a bot:Interface ;
    bot:interfaceOf ?wall ;
    bot:interfaceOf ?insulation .
  ?wall ex:hasRetrofitLayer ?insulation ;
    s4bldg:thermalTransmittance ?uValueNode .
  ?uValueNode qudt:numericValue ?uValue .
}

```

Listing II. Cross-model thermal transmittance retrieval

Table IV reports the results obtained from the federated knowledge graph. The query successfully retrieves the wall

element belonging to the as-is model, the corresponding insulation layer contained in the retrofit model, and the associated thermal transmittance value. In the analysed case study, the retrieved U-value is equal to 0.35 W/m²K, corresponding to the thermal performance of the external insulation system represented in the retrofit BIM model.

The query demonstrates the practical application of semantic interoperability within the proposed framework. Thermal properties originally stored in an independent BIM model become accessible through a unified semantic representation, enabling the generation of energy-oriented views without modifying the source IFC models.

Interface	As-is Wall	Retrofit Insulation	U-value (W/m ² K)
inst:Interface_1	https://example.org/as-is#wall_59cb93ee-2221-4804-b9f1-df683cacfd2e	tobe:covering_55ab4281-8201-4e36-92f3-4fb7e1a42601	"0.35"^^xsd:decimal

Table IV. Example of U-value of a generated envelope system

The second query (Listing III) demonstrates the retrieval of derived energy indicators generated during the semantic enrichment phase. Unlike thermal transmittance values, which originate from BIM information, the WWR is not directly available in the source IFC models. Instead, it is computed through a geometric derivation process that combines window areas and external wall areas extracted from the as-is BIM model and groups them according to façade orientation.

The derived WWR values are represented in the enrichment graph as dedicated semantic resources of type `ex:WindowWallRatio`, associated with the corresponding façade orientation and stored as `qudt:QuantityValue` instances. By materializing the indicator within the knowledge graph, the framework enables its retrieval through standard SPARQL queries without requiring any additional geometric processing during the query stage.

```

PREFIX ex: <https://caseStudy/ontology#>
PREFIX qudt: <http://qudt.org/schema/qudt/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT ?orientation ?wwr
WHERE {
  ?node a ex:WindowWallRatio ;
    ex:facadeOrientation ?orientation ;
    qudt:numericValue ?wwr .
}
ORDER BY ?orientation

```

Listing III. Window-to-Wall Ratio retrieval

The query extracts the Window-to-Wall Ratio values associated with each façade orientation directly from the energy enrichment graph. The results demonstrate that the framework is capable of generating and storing energy-oriented indicators as first-class semantic entities that can subsequently be queried and reused by downstream applications.

Table V reports the WWR values generated for the analysed case study. The North façade presents the highest glazing ratio (17.6%), reflecting the predominance of north-oriented external walls and windows in the building geometry. The West façade exhibits a significantly lower glazing ratio (1.3%), while the South and East façades return null values due to the absence of windows or external wall geometry associated with those orientations in the source BIM model.

Orientation	WWR (%)
“East”	"0.000"^^xsd:decimal
“West”	"0.013"^^xsd:decimal
“North”	"0.176"^^xsd:decimal
“South”	"0.000"^^xsd:decimal

Table V. Window-to-Wall Ratio values derived through semantic enrichment and retrieved through SPARQL.

Unlike the first query, which demonstrates interoperability between independent BIM models, this second query illustrates the capability of the ontology-based environment to generate, store and retrieve derived energy indicators.

Energy-oriented exploitation of the Knowledge Graph

The EIM not only defines the parameters required for energy-related analyses but also establishes the semantic structure necessary to support downstream analytical workflows. Once represented within the federated knowledge graph, the resulting Energy Semantic View provides a unified and queryable representation of geometric, physical, environmental and operational information extracted from BIM models and external data sources.

As illustrated in Figure 4, the Energy Semantic View integrates BIM-derived information, climatic observations, operational parameters and energy indicators derived within a common semantic environment. Parameters such as thermal transmittance, WWR, climatic indicators, occupancy schedules and operational setpoints can be retrieved from the knowledge graph and supplied as inputs to interpretable and physics-informed modelling approaches.

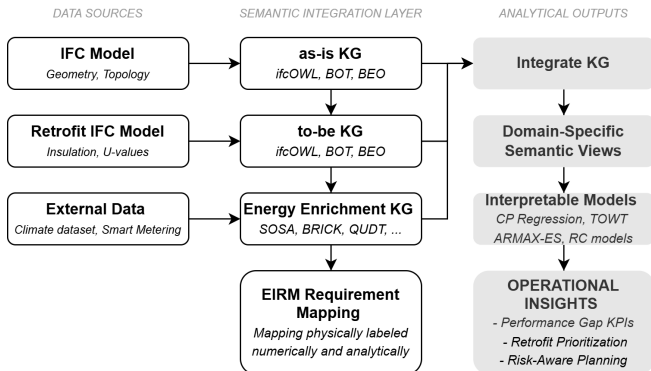


Fig. 4. Energy-oriented exploitation of the federated knowledge graph.

The modelling parameters can then be manipulated with analytical models, built on the approximated definition of building energy balance components [8], that could, in turn, be compared with the actual models' parameters (e.g., CP regression and TOWT models) derived from model fitting on measured data. On the one hand, this is meant to ease the

calibration of energy models. While standards such as ASHRAE 140-2023 provide testing criteria for dynamic energy simulation models, and ISO 52016-1:2017 provides a standardized dynamic calculation methodology, they cannot cope with the variability of real-world operational conditions. This renders calibration process necessary to address the "design-operation" performance gap. Approximated building energy balance components definition is meant to be used for this scope.

On the other hand, it essentially gives the possibility to achieve a better understanding of actual building behaviour by physically grounded estimates of model coefficients to be compared to the actual ones. These estimates can be compared to assumptions in the calibration process. This comparison can involve, for example, the overall heat transfer coefficient, effective solar gain factor, base heat load, and internal gains estimates [22]. Furthermore, this comparison is not strictly limited to "static" parameters and can be extended to dynamic operational regimes (e.g., occupied vs. unoccupied) and actual typical recurrent operating schedules identified by model fitting. In this perspective, the federated knowledge graph acts as a semantic bridge between BIM-derived information, operational data and interpretable building performance models, enabling the generation of traceable and physically meaningful Key Performance Indicators that can support calibration, performance-gap analysis and retrofit decision-making.

VI. DISCUSSION

The results demonstrate the feasibility of the proposed ontology-based framework for integrating BIM-derived, external and derived information within a unified semantic environment. Together, these ontology layers provide a unified semantic environment capable of representing, integrating and querying the information required for energy-oriented analyses without requiring the generation of dedicated BEM.

The generation of the federated knowledge graph directly addresses RQ1, which investigated whether an IFC model can be translated into a temporary representation suitable for energy-related applications. The proposed workflow successfully transforms the original IFC models into interoperable RDF knowledge graphs and enriches them with additional information required for energy-oriented analyses. The resulting federated knowledge graph therefore acts as a temporary semantic meta-model capable of supporting domain-specific applications while preserving the original IFC models as the single source of truth.

The second research question (RQ2) explored whether energy-related calculations and indicators can be managed within an ontology-based environment. The results obtained through Listing III demonstrate that derived energy indicators such as the Window-to-Wall Ratio can be generated during the semantic enrichment phase, represented as semantic resources within the knowledge graph and subsequently queried through standard SPARQL mechanisms. This confirms the capability of the ontology-based framework to support the management and retrieval of derived energy-related information.

Finally, RQ3 focused on the practical application of interoperability within the proposed framework. Listing II demonstrates that semantic relationships can be established between independent BIM models through the cross-model

adjacency graph. The retrieval of wall elements, retrofit insulation layers and associated thermal properties through a single SPARQL query illustrates how heterogeneous information sources can be integrated without model merging or the generation of dedicated BEMs.

The proposed framework does not aim to replace existing energy simulation environments. Instead, it provides a semantic integration layer capable of organizing and contextualizing the information required by energy-related analyses. Future work will focus on extending the enrichment process to additional parameters required by energy performance assessment standards and on the automatic generation of simulation-ready semantic views.

VII. CONCLUSION

This paper presented an ontology-based framework for the generation of energy-oriented semantic views starting from IFC-based BIM models and heterogeneous information sources. The proposed approach combines an EIM, semantic translation mechanisms and KG technologies to integrate IFC-native information, retrofit interventions, derived indicators and external datasets within a unified semantic environment.

The results demonstrate the feasibility of translating IFC models into interoperable KG and enriching them through semantic relationships and derived energy indicators. The federated KG enables the integration of independent BIM models through explicit cross-model relationships while preserving the original IFC files as the single source of truth. Furthermore, the case study showed that both retrofit information and derived indicators, such as WWR, can be represented and queried through standard semantic technologies without generating dedicated BEMs.

The current implementation has been validated on a single case study and a limited subset of energy-related parameters. Consequently, additional validation is required to assess the applicability of the approach to different building typologies, BIM authoring environments and performance assessment scenarios. Future work will focus on extending the enrichment process to additional parameters and domains, further expanding the EIM and investigating the generation of analysis-oriented semantic views capable of supporting a broader range of performance assessment workflows.

ACKNOWLEDGMENTS

This work was carried out in the framework of the European Horizon SPHEERE Project, Grant agreement ID: 101235544

REFERENCES

- [1] Pauwels, P. and Terkaj, W. (2016). Express to owl for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in Construction*, 63:100–133.
- [2] Rasmussen, M. H., Lefrançois, M., Bonduel, M., and Pauwels, P. (2017). The building topology ontology. *Semantic Web*, 9(2):1–11.
- [3] Utkucu, D., Sacks, R., et al. (2025). Ontology for holistic building performance modeling and analysis. *Automation in Construction*, 175:106197.
- [4] Moreira, J., Donkers, A., Pauwels, P., Bektas, E., Ee, T.v. (2024). Onto4reuse: Towards an ontology reuse framework for knowledgeintensive software engineering. In *Proceedings of the Joint Ontology Workshops (JOWO)*.
- [5] M. Kavcic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, M. Djurovic-Petrovic, A review of bottom-up building stock models for energy consumption in the residential sector, *Build. Environ.* 45 (2010) 1683–1697. <https://doi.org/http://dx.doi.org/10.1016/j.buildenv.2010.01.021>.
- [6] A. Fouquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: A review, *Renewable and Sustainable Energy Reviews* 23 (2013) 272–288. <https://doi.org/http://dx.doi.org/10.1016/j.rser.2013.03.004>.
- [7] Y. Chen, M. Guo, Z. Chen, Z. Chen, Y. Ji, Physical energy and data-driven models in building energy prediction: A review, *Energy Reports* 8 (2022) 2656–2671. <https://doi.org/https://doi.org/10.1016/j.egyr.2022.01.162>.
- [8] J. Smertinas, H.T. Walnum, B. Ludvigsen, Å.L. Sørensen, I. Sartori, H. Madsen, Building-PROFet: A scalable data-driven modelling approach for estimating heat and electricity load profiles of buildings, *Energy* 353 (2026) 140814. <https://doi.org/https://doi.org/10.1016/j.energy.2026.140814>.
- [9] J. Smertinas, M.Y.C. Van Hove, P. Bacher, H. Madsen, Digital energy performance certification of single-family houses in Denmark, *Sustain. Cities Soc.* 141 (2026) 107283. <https://doi.org/https://doi.org/10.1016/j.scs.2026.107283>.
- [10] M. Manfren, N. Aste, F. Leonforte, C. Del Pero, M. Buzzetti, R.S. Adhikari, L. Zhixing, Parametric energy performance analysis and monitoring of buildings—HEART project platform case study, *Sustain. Cities Soc.* 61 (2020) 102296. <https://doi.org/https://doi.org/10.1016/j.scs.2020.102296>.
- [11] P. Price, Methods for analyzing electric load shape and its variability, *Lawrence Berkeley National Laboratory Report LBNL-3713E*, 2010.
- [12] J.L. Mathieu, P.N. Price, S. Kiliccote, M.A. Piette, Quantifying changes in building electricity use, with application to demand response, *IEEE Trans. Smart Grid* 2 (2011) 507–518.
- [13] M. Manfren, P.A.B. James, V. Aragon, L. Tronchin, Lean and interpretable digital twins for building energy monitoring – A case study with smart thermostatic radiator valves and gas absorption heat pumps, *Energy and AI* 14 (2023) 100304. <https://doi.org/https://doi.org/10.1016/j.egyai.2023.100304>.
- [14] LFEnergy Open DSM (<https://lfeenergy.org/projects/opendsm/>), (n.d.).
- [15] Han, Q., Xie, X., and Lu, W. (2025). Unified network based representation of bim models for efficient data analysis. *Automation in Construction*, 165:105321.
- [16] Massafra, A., Jabi, W., and Gulli, R. (2024). Topological bim for building performance management. *Automation in Construction*, 166:105628.
- [17] Massafra, A., A.-H. D. H. S. L. . J. W. (2025). Semiautomated dataset generation for residential buildings using graph-based topological modelling. *Buildings*, 15(8):1283.
- [18] Karmakar, A. and Delhi, V. S. K. (2023). Restructuring elements in ifc topology through semantic enrichment for automated compliance checking. In *Proceedings of the 40th International Symposium on Automation and Robotics in Construction (ISARC)*, pages 1–8.
- [19] Cassandro, J., Hagedorn, P., Sigalov, K., Gatto, C., Mirarchi, C., Pavan, A. & König, M. (2025) Semantic Integration of Cost, Time and Geometry for a Resource-centred Perspective of Managing Construction Projects, in *Proceedings of the 2025 EC³ and the 42nd CIB W78*. Porto, Portugal, 14–17 July
- [20] Hagedorn, P., Cassandro, J., Sigalov, K., Mirarchi, C., Pavan, A. and König, M. (2025) A Shared Construction Resource Ontology for Semantically Aligning Cost and Time Domains in Construction Projects, in *Proceedings of the 13th LDAC 2025*, Porto, Portugal, 9–11 July. *CEUR Workshop Proceedings*, 4113, pp. 10–24.
- [21] Cassandro, J., & Mirarchi, C. (2026) Topology-Driven Semantic Integration of Independent BIM Models for Renovation Analysis, in *Proceedings of the 2026 EC³*. Corfu, Greece, 12–15 July.
- [22] L. Tronchin, M. Manfren, B. Nastasi, Energy analytics for supporting built environment decarbonisation, *Energy Procedia* 157 (2019) 1486–1493. <https://doi.org/https://doi.org/10.1016/j.egypro.2018.11.313>.