2026 European Conference on Computing in Construction
Corfu, Greece
July 12–15, 2026

TOPOLOGY-DRIVEN SEMANTIC INTEGRATION OF INDEPENDENT BIM MODELS FOR RENOVATION ANALYSIS

Jacopo Cassandro and Claudio Mirarchi
Department of Architecture, Built Environment, and Construction Engineering (DABC),
Polytechnic of Milan, Milan, Italy

Abstract

This paper proposes a topology-driven semantic integration approach to align as-is and to-be BIM models in building renovation scenarios. Starting from independent IFC files, a surface-based topological analysis infers physically meaningful contacts between envelope objects using coplanarity, opposite orientation, vertical overlap, and an overlap-coverage metric. The detected adjacencies are then represented as cross-model semantic relations by instantiating `bot:Interface` links between the corresponding `bot:Element` individuals in a dedicated knowledge graph, preserving model independence. The resulting topology enables cross-model SPARQL queries to extract and aggregate energy-related data at the envelope-system level without IFC merging or manual data preparation.

Introduction

Building renovation workflows increasingly rely on Building Information Modeling (BIM) based digital representations to support design, analysis, and decision-making processes (Han et al., 2025; Ismail et al., 2017). In practice, however, renovation projects are typically characterized by multiple BIM models representing different project phases, such as the existing (as-is) condition and the proposed retrofit (to-be) intervention. These models are often developed at different time stages, so they are simultaneously independent of each other, stored in separate IFC files, and lack explicit machine-interpretable semantic relationships that describe how the introduced objects physically relate to the existing building (Pauwels and Terkaj, 2016; Rasmussen et al., 2017).

This fragmentation represents a significant challenge for automatic data analysis by machines, such as the extraction of energy assessment for the envelope system. While the physical reality of a renovated building envelope is composed of coherent construction systems (e.g., walls upgraded with insulation layers), this coherence is not explicitly represented in conventional BIM models. As a result, properties relevant for performance evaluation are distributed across independent models, preventing automated system-level reasoning and requiring manual interpretation, model merging, or ad hoc data preparation.

Existing approaches to BIM-based analysis often assume that relevant relationships are already encoded within a single model or can be derived through explicit semantic links. However, such assumptions rarely hold in renovation scenarios, where as-is and to-be models remain semantically detached. Geometry-based methods can detect spatial proximity or collisions, but they typically do not provide a structured mechanism to transform geometric evidence into persistent, machine-interpretable relations suitable for higher-level analysis (Gan et al., 2022). Conversely, ontology-based approaches offer rich semantic representations, yet they rely on the availability of explicit relationships that are not present in independent BIM models (Karmakar and Delhi, 2023; Moreira, 2024).

This paper addresses this gap by proposing a geometry-driven and ontology-based approach for aligning independent as-is and to-be BIM models through surface-level topological analysis and semantic linking. Rather than merging models or imposing predefined semantic structures, the proposed method infers physically meaningful adjacency relations directly from geometric surface evidence and explicitly encodes them as semantic relations, reusing the Building Topology Ontology (BOT) and Building Element Ontology (BEO). In this way, objects that are fragmented across different models but belong to the same physical construction system can be coherently related in a machine-interpretable manner.

The main contribution of this work lies in enabling cross-model reasoning over building envelope systems without compromising model independence. By transforming geometric contact evidence into explicit semantic constructs, the proposed approach allows distributed BIM data to be aggregated, queried, and exploited for performance-oriented analyses, such as energy evaluation(Utkucu et al., 2025; Massafra et al., 2024). The approach is demonstrated through a case study focusing on envelope retrofit, showing how independent BIM models can be semantically linked to support automated data extraction and downstream simulation workflows.

The paper is structured as follows. Section *Methodology* presents the proposed methodology, including geometry-driven topological analysis and semantic enrichment. Section *Topological analysis and overlap metrics* formalizes the algorithm for the analysis of objects surface-based matching and the associated overlap metric. Section *Case study and results* reports the case study and discusses the results, while the final sections address semantic linking, energy-oriented data exploitation, and future work.

Background

Topology in BIM

Several studies have investigated the representation of topological relationships within BIM models to support advanced querying, navigation, and analysis tasks. Graph-based approaches encode building components and their relations as nodes and edges, allowing the computation of shortest paths, connectivity patterns, and higher-level structural properties (Han et al., 2025). Graph databases have been used to transform IFC models into object graphs, supporting advanced querying, algorithms, and topological analysis in IFC versions (Ismail et al., 2017). In parallel, IFC-based graph data models use weighted graph principles and automatic algorithms to extract and query topological relationships among 3D elements more efficiently than traditional 2D/3D representations (Khalili and Chua, 2015). More recent work has introduced the notion of Topological BIM (TBIM), a space-oriented BIM representation tailored for building performance management; TBIM relies on topological and conditional modeling principles to derive semantically enriched information models to build performance simulation (Massafra et al., 2024). Graph-based topological models have also been used to support semi-automated dataset generation for residential buildings, showing that explicit spatial graphs can serve as a robust backbone for downstream analytical tasks (Massafra, 2025).

Beyond purely topological representations, several approaches leverage geometry-based techniques—such as surface adjacency, coplanarity, and contact area analysis—to infer spatial relationships between building elements. These methods extend graph-based BIM representations by incorporating geometric evidence to identify physically compatible interfaces between components, for instance, in modular or surface-driven modeling strategies (Gan et al., 2022). However, geometry-based inferences are typically applied within a single BIM model and produce transient results that are not persisted as explicit semantic relations, limiting their applicability to cross-model renovation scenarios.

BIM-to-RDF

The translation of BIM data into semantic representations has been extensively explored as a means to improve interoperability, data integration, and querying capabilities. Ontologies such as IfcOWL enable a comprehensive mapping of the IFC schema into RDF triples, supporting formal reasoning and semantic querying. Pauwels and Terkaj (2016) formalized a standardized EXPRESS-to-OWL conversion procedure specifically targeting IFC, producing a recommendable ifcOWL ontology in OWL2 DL that balances fidelity to the original schema with practical usability. Due to the size and complexity of full IFC-to-RDF mappings, several lightweight ontologies have been proposed to represent core aspects of building topology and element classification in a more tractable way; notable examples include BOT and BEO ontologies, which focus respectively on spatial/elementary topology and on reusable element taxonomies (Rasmussen et al., 2017). More recently, domain-specific performance ontologies like the Holistic Building Performance Ontology (HBPO) provide unified schemas to connect BIM data with multi-domain building performance simulation tools (acoustics, lighting, energy) (Utkucu et al., 2025). In the Linked Building Data (LBD) context, these ontologies are combined with IFC and other domain schemas to support data-driven modeling of building topology, while frameworks such as Onto4Reuse address ontology reuse and alignment across BIM-related domains (Moreira, 2024). In parallel, semantic enrichment has been proposed as a strategy to reconstruct implicit or missing information in IFC-based models, particularly where exported data lack the relational structure needed for higher-level reasoning. Rule-based enrichment techniques infer missing topological relations, reorganise building hierarchies, and support automated code compliance checking by encoding domain knowledge as declarative rules (Karmakar and Delhi, 2023). Nevertheless, most of these semantic approaches assume that relevant relationships are explicitly present in the source models, or can be directly mapped from the IFC schema, and do not tackle the derivation of new topological relations from geometric evidence, nor the reconstruction of topology between independent as-is and to-be models. These approaches confirm that BIM models often lack explicit relational topology and that enrichment is necessary to make such information machine-interpretable. However, existing semantic enrichment methods are typically domain-specific, rely on rule sets tailored to particular use cases (e.g. regulatory checking or asset management), and operate on single BIM models. As a result, they offer limited support for renovation workflows in which as-is and to-be models are authored independently, stored in separate IFC files, and need to be semantically linked through derived topological relations.

Research gaps and research questions

The literature review highlights significant advances in graph-based BIM representations, geometry-driven spatial analysis, and semantic enrichment techniques. However, these approaches are typically developed in isolation and do not fully address renovation workflows characterized by the coexistence of independent as-is and to-be BIM models. In particular, there is a lack of methods capable of deriving persistent, machine- interpretable adjacency relations from geometric evidence and exploiting them for system-level performance analysis.

This research addresses the following questions:

- RQ1: How can adjacency relations between independent BIM models be inferred from surface-level geometric evidence?
- RQ2: How can such relations be semantically created

without model merging?

- RQ3: How can the resulting semantic model support performance-oriented data extraction?

Methodology

The proposed methodology enables the identification and semantic formalization of cross-model adjacency relations between building envelope objects belonging to independent as-is and to-be BIM models, which are typically disconnected in renovation workflows. The approach combines geometry-driven topological analysis with semantic knowledge graph representation, enabling aggregation of relevant properties (e.g., thermal transmittance) at the envelope-system level, despite their distribution across separate IFC files. This enables the use of domain-specific analysis tools (e.g., energy simulation software) to evaluate renovation scenarios without requiring the IFC model to be merged.

The proposed method transforms geometric surface-contact evidence into semantic envelope systems, thereby enabling the integration of distributed BIM data for energy performance evaluation. By decoupling geometric analysis from semantic representation, the approach preserves model independence while supporting cross-model reasoning and system-level assessment.

As illustrated in Figure 1, the workflow is structured into four main steps: independent IFC model, geometry-driven topological analysis and semantic translation, semantic enrichment through topological relation, and data extraction for domain-specific analysis.

Step 1 – Independent IFC model.

The IFC^{as-is} model is analysed and a corresponding IFC^{to-be} model is generated. Both models are treated as independent data sources, each containing specific objects and performance-related properties associated with different project phases. The objects of both models lack explicit relationships that a machine can interpret.

Step 2 – Geometry-driven topological analysis and semantic translation.

- *Step 2.a – Surface-based topological analysis.* A geometry-driven topological analysis is performed on the geometric representations of the as-is and to-be models to identify surface-level contact evidence between envelope objects. Topological adjacency is inferred from geometric conditions, including the detection of coplanar and oppositely oriented surfaces exhibiting a significant overlap area. This process produces a set of candidate adjacency relations expressed as pairs of element identifiers (GUID) and associated overlap metrics.
- *Step 2.b – Semantic representation of BIM data.* In parallel, the IFC^{as-is} and IFC^{to-be} models are independently translated into RDF reusing BOT and BEO ontology, generating two separate knowledge graphs (KG) that preserve the original model structure, geometry references, and performance-related properties.

Step 3 – Semantic enrichment based on geometry-driven topological analysis.

The geometric adjacency evidence obtained in Step 2.a is subsequently materialized as new semantic relations between the KG^{as-is} and KG^{to-be} generated in Step 2.b. Specifically, adjacency relations are instantiated as topological relations using the BOT ontology (e.g., `bot:Interface`), enabling the formal representation of cross-model envelope systems. This step represents the transition from geometry-driven evidence to explicit, machine-interpretable topological knowledge that can be queried, reasoned upon, and exploited in downstream analyses.

Step 4 – Data extraction for domain-specific analysis.

The enriched knowledge graph is queried using SPARQL to retrieve envelope-system-level data, including thermal properties originating from both the as-is and to-be models. These data are then provided as input to external energy analysis software to support detailed performance assessment of renovation strategies.

Topological analysis and overlap metric

This section builds upon the geometric processing and semantic preparation steps introduced in the previous methodology, focusing on the formal definition of topological analysis and its associated overlap metric. In the proposed approach, topological analysis is not assumed as an input condition but is derived from geometric surface evidence through a geometry-driven inference process. The objective is to identify physically meaningful adjacencies between building objects belonging to independent IFC_{as-is} and IFC_{to-be} models.

Geometry-driven topological analysis

Geometry-driven topological analysis is defined as the process of identifying adjacency relations between objects belonging to independent BIM models based on geometric surface evidence. Rather than relying on predefined semantic links or model merging, adjacency is inferred by analysing the spatial configuration and mutual compatibility of envelope surfaces.

This geometry-driven topological analysis enables the creation of cross-model relations that are not explicitly represented in the original IFC data, but are essential for system-level reasoning in renovation workflows.

Geometric preprocessing and surface abstraction

Each target building object (e.g., wall) is first converted from its native IFC geometric representation into a triangulated surface mesh. This tessellation step is required to homogenize heterogeneous IFC geometry definitions (e.g., swept solids, boundary representations) into a common computational format suitable for surface-based analysis.

Formally, each object Obj is represented as a set of trian-

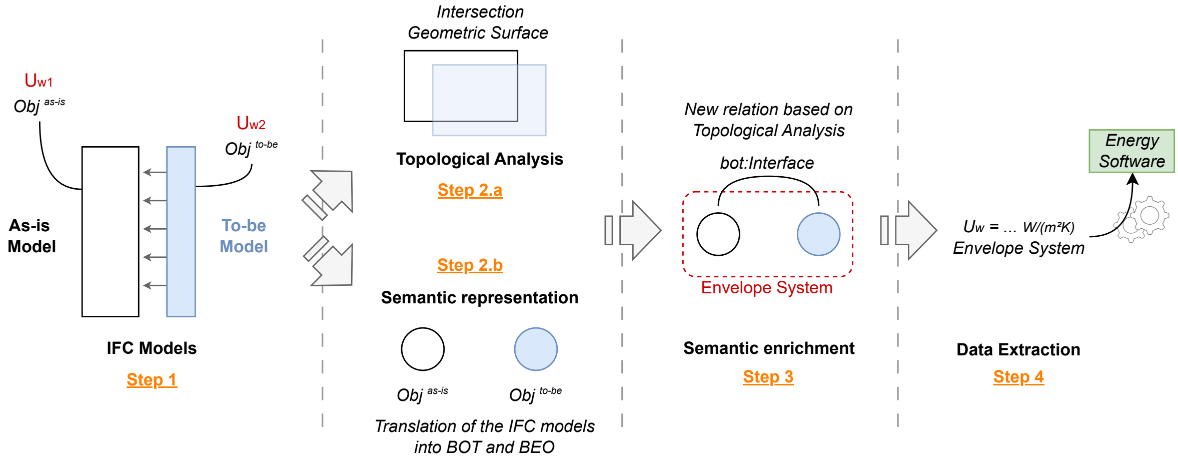

Figure 1: Methodology

gular faces:

$$\mathscr{T}(Obj) = \{t_k \mid t_k = (\mathbf{v}_{k1}, \mathbf{v}_{k2}, \mathbf{v}_{k3})\} \tag{1}$$

where $\mathbf{v}_{ki} \in \mathbb{R}^3$ are the triangle vertices expressed in a shared world coordinate system.

To reduce the computational complexity of subsequent operations, an axis-aligned bounding box (AABB) $B(Obj)$ is computed for each object and used for broad-phase spatial filtering. Only object pairs whose bounding boxes intersect are considered for further analysis.

For each triangle t_k, fundamental geometric descriptors are computed, including surface normal, area, and supporting plane. Individual triangles are then aggregated into *macro-faces* representing extended planar surface regions by grouping coplanar triangles with similar orientation. Each macro-face f is characterized by a unit normal $\mathbf{n}_f$, a supporting plane π_f, a total area $A(f)$, and a vertical extent $[z_{\min}, z_{\max}]$.

This surface abstraction step provides a compact and semantically meaningful representation of envelope geometry, enabling robust surface-level comparison across independent models.

Surface compatibility conditions

Let f_a and f_b be two macro-faces belonging respectively to an as-is object Obj_i^{as-is} and a to-be object Obj_j^{to-be}. The two surfaces are considered topologically compatible only if the following geometric conditions are simultaneously satisfied.

Opposite orientation

The two surfaces must face each other, ensuring the physical plausibility of a contact interface:

$$\angle(\mathbf{n}_{f_a}, -\mathbf{n}_{f_b}) < \varepsilon_\theta \tag{2}$$

where $\mathbf{n}_{f_a}$ and $\mathbf{n}_{f_b}$ denote the unit normal vectors of the two macro-faces.

Coplanarity

The supporting planes of the two surfaces must coincide within a geometric tolerance:

$$\text{dist}(\pi_{f_a}, f_b) < \varepsilon_d \tag{3}$$

where π_{f_a} denotes the plane associated with f_a and the distance is evaluated using a point of f_b.

Vertical overlap

To exclude geometrically coplanar but vertically misaligned surfaces, a vertical overlap constraint is introduced:

$$\min(z^a_{\max}, z^b_{\max}) - \max(z^a_{\min}, z^b_{\min}) > \tau_z \tag{4}$$

where $[z^a_{\min}, z^a_{\max}]$ and $[z^b_{\min}, z^b_{\max}]$ represent the vertical extents of the two macro-faces.

Together, these conditions ensure that only surface pairs capable of defining a physically meaningful interface are retained for further evaluation.

Overlap metric definition

For each pair of geometrically compatible macro-faces, a quantitative overlap metric is computed to assess the extent of the contact. Each surface is projected onto a local two-dimensional coordinate system defined on its supporting plane. The union of projected triangles forms a planar polygon representation P_f for each macro-face.

The effective contact surface area is computed as:

$$A_\cap(f_a, f_b) = \text{area}(P_{f_a} \cap P_{f_b}) \tag{5}$$

To provide a normalized and scale-independent measure, an overlap ratio is defined as:

$$\rho(f_a, f_b) = \frac{A_\cap(f_a, f_b)}{A(f_a)} \tag{6}$$

The overlap ratio expresses the proportion of the as-is surface that is covered by the to-be surface, enabling a quantitative comparison between different contact configurations.

Decision rule for adjacency inference

A topological adjacency relation between two objects Obj_i^{as-is} and Obj_j^{to-be} is inferred if at least one pair of their macro-faces satisfies the overlap condition:

$$\exists(f_a, f_b) \quad \text{s.t.} \quad \rho(f_a, f_b) \geq \tau \tag{7}$$

where τ is a user-defined threshold controlling the minimum required contact significance.
This rule allows multiple Obj_j^{to-be} to be associated with a single Obj_i^{as-is}, resulting in a one-to-many topological matching configuration. The inferred adjacency relations constitute the basis for subsequent semantic enrichment. Topological matching is thus formulated as a geometry-driven inference process, in which surface overlap metrics provide both qualitative adjacency relations and quantitative evidence supporting system-level aggregation.

Semantic linking and data exploitation

The set of inferred adjacency relations produced by the geometry-driven topological analysis is subsequently materialized in a semantic knowledge graph. To manage the complexity of IFC representations, the models are translated reusing the BEO ontology, which provides a lightweight taxonomy for objects such as `beo:Wall` and `beo:Insulation`, facilitating more efficient SPARQL querying compared to the direct translation of the EXPRESS schema into IfcOWL which could generate millions of RDF triples. Each objects pair (Obj_i^{as-is},Obj_j^{to-be}) is mapped to its corresponding RDF individuals using the *IfcGlobalUniqueId* as a linking key. To maintain ontological correctness—given that `bot:adjacentElement` is formally restricted to Zone-to-Element relations—the physical contact between the Obj_i^{as-is} and Obj_j^{to-be} is represented through the `bot:Interface` class. This approach allows for the creation of a topological "bridge" between the two models, where a single interface instance is declared as `bot:interfaceOf` both the existing objects and the new retrofit objects. The inferred relations are asserted in a dedicated KG that operates as a federated layer over the original KG^{as-is} and KG^{to-be}. This architecture preserves the integrity of the original BIM data while enabling the explicit representation of cross-model adjacencies derived from geometric evidence. This semantic enrichment enables advanced cross-model querying of envelope systems through SPARQL. By relating objects coming from independent IFC models into a KG, specific data (e.g., U-values) can be structured, extracted, and used as input data for domain-specific simulation tools.

Case study and results

As a case study, a typical building envelope retrofit scenario was considered, featuring an existing wall with an opening (a window), as illustrated in Figure 2. Despite its simplicity, the case study already includes relevant complexities, such as multiple elements, the presence of an opening, and partial geometric misalignments (e.g., vertically offset objects). These aspects allow demonstrating the method's ability to correctly identify meaningful adjacencies while effectively excluding false positives. The IFC^{as-is} model represents the existing state, characterized by a multi-layered composition (internal plaster, masonry, and external plaster). The IFC^{to-be} model includes only the retrofit objects, namely an external insulation system applied to the existing wall and a final plaster finish.

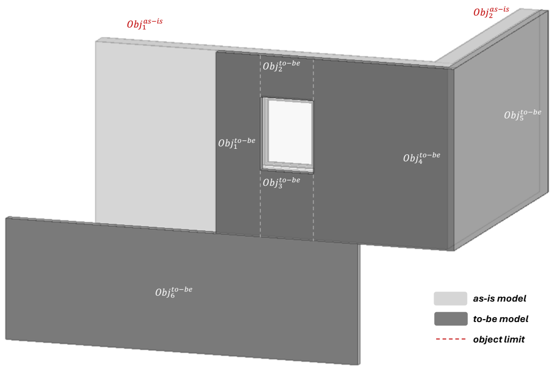

Figure 2: CaseStudy

To evaluate the proposed geometry-driven topological analysis, the retrofit objects were intentionally modeled as separate objects (Obj_1^{to-be}, Obj_2^{to-be}, Obj_3^{to-be}, Obj_4^{to-be}, and Obj_5^{to-be}) and some with a vertical offset relative to the existing objects (Obj_6^{to-be}). This modeling choice was adopted to test the algorithm's ability to distinguish between geometrically coplanar surfaces but with vertically misaligned surfaces and those with physically significant adjacencies, ensuring that only overlapping object surfaces are identified as valid contacts.
The case study was designed as a simplified yet representative scenario that reflects common building envelope renovation scenarios. Although limited in scale, it provides a controlled setting to evaluate the proposed method before its application to larger and more complex case studies within the ongoing research project.
Both models were exported using a shared coordinate system in IFC 4.3 ADD2 format and subsequently processed using a custom Python code (Python 3.10) to perform the geometry-driven topological analysis. In parallel, the IFC models were translated into lightweight semantic representations based on the BOT and BEO ontologies with the IFCtoLBD converter (Oraskari et al., 2024). The reuse of these ontologies provides a more compact and tractable structure compared to full ifcOWL representations, enabling efficient graph-based data management for the renovation process. The resulting RDF graphs were stored and queried using an Ontotext GraphDB.

Geometric matching results

The geometry-driven topological analysis was executed automatically on the two IFC models described in the pre-

vious section, following the methodology steps presented in the section *Topological Analysis and Overlap Metric*. The algorithm successfully identified the contact surfaces between existing objects (Obj_1^{as-is} and Obj_2^{as-is}) and new retrofit insulation objects (Obj_1^{to-be}, Obj_2^{to-be}, Obj_3^{to-be}, Obj_4^{to-be}, and Obj_5^{to-be}), while correctly excluding non-relevant surfaces (Obj_6^{to-be}).
No false-positive adjacencies were detected between geometrically coplanar but vertically misaligned surfaces (Figure 3).

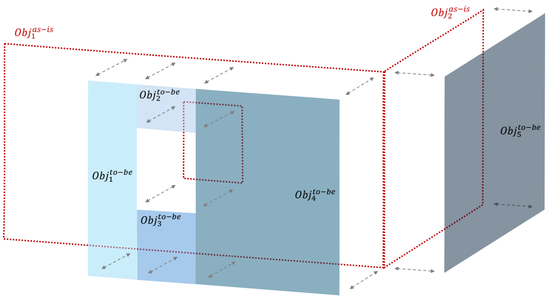

Figure 3: Geometry-driven topological analysis results.

The two existing walls of the IFC^{as-is} model were found to be adjacent to objects in the IFC^{to-be} model. In the scenario visible in Figure 2, a single existing object (Obj_1^{as-is}) was correctly associated with four different objects (Obj_1^{to-be}, Obj_2^{to-be}, Obj_3^{to-be}, and Obj_4^{to-be}), corresponding to the segmented area around the window opening. For all four objects, the coverage ratio computed with respect to the Obj_j^{to-be} surfaces exceeded 94%, indicating that each retrofit object is almost entirely supported by the existing wall, despite covering only a limited portion of its surface.
On the other side, a one-to-one correspondence between an existing object (Obj_2^{as-is}) and a retrofit object (Obj_5^{to-be}), with surface coverage ratios close to 100% on both sides, reflecting a full surface alignment.
The geometry-driven topological analysis results are summarized in Table 1, which respectively illustrate the contact surfaces identified between the existing objects (Obj_i^{as-is}) and the retrofit objects (Obj_j^{to-be}), including the corresponding surface coverage ratios for each detected adjacency. The results confirm that only geometrically and vertically aligned surfaces are selected, while areas associated with openings or without existing objects behind are correctly excluded. Furthermore, the results confirm the ability of the proposed algorithm to identify meaningful adjacency relations across independent IFC models and to quantify their geometric consistency through surface-based overlap metrics.

Semantic linking results

The adjacency relations inferred at the geometric level were automatically generated in a dedicated RDF case-study graph (KG^{link}). Starting from the set of matched object pairs produced by the geometry-driven topological analysis, the linking process was executed automatically.
Each pair of matched objects was identified across the independent KG_{as-is} and KG_{to-be} using the *IfcGlobalId* as a unique linking key. For each identified contact, a `bot:Interface` individual was automatically instantiated in KG^{link} and asserted as `bot:interfaceOf` both the existing `bot:Element` and the corresponding retrofit `bot:Element`. This approach enables the explicit and ontologically consistent representation of object-to-object adjacencies across independent BIM models.
All interface instances and their properties were generated directly from the output of the geometry-driven algorithm, ensuring full traceability between geometric evidence and semantic assertions (Figure 4). The resulting structure transforms a set of geometrically fragmented objects into a coherent, machine-interpretable system in which individual components are explicitly related through formal semantic constructs, thereby enabling consistent reasoning, querying, and exploitation via SPARQL.

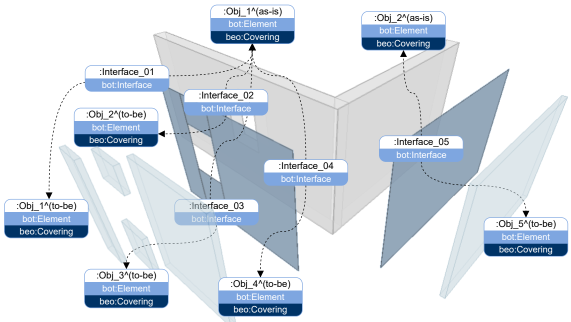

Figure 4: Ontology-based adjacencies as `bot:Interface` instances linking as-is and to-be objects.

Energy-oriented data extraction

The semantic linking introduced in the previous step enables the formulation of cross-model SPARQL queries that directly operate on physically meaningful object-to-object relations. By exploiting the `bot:Interface` individuals created in KG^{link}, existing and retrofit objects that are physically adjacent in the renovated envelope can be queried jointly, despite originating from independent KG^{as-is} and KG^{to-be} representations.
In this context, energy-related properties are retrieved by traversing the `bot:interfaceOf` relations and accessing the property states associated with the linked objects. Specifically, the thermal transmittance values of both the as-is and the to-be objects are extracted through their respective IFC property representations, using the *IfcGlobalId* as a stable linking key. This allows the selective retrieval of thermal properties only for object pairs that have been previously identified as physically adjacent through geometry-driven analysis.
As a result, thermal transmittance values that are distributed across separate BIM models can be aggregated at

Table 1: Geometry-driven topological analysis results with objects GUID and surface coverage ratios.

GUID Obj_i^{as-is}	GUID Obj_j^{to-be}	Surface Coverage Ratios
3CFf4AGlbEchvmFeRIHcCG	2FAjB5xJr21RdQcsKQE3TQ	Coverage Obj_1^{as-is}: 13.29% Coverage Obj_1^{to-be}: 100%
3CFf4AGlbEchvmFeRIHcCG	2FAjB5xJr21RdQcsKQE3UZ	Coverage Obj_1^{as-is}: 3.71% Coverage Obj_2^{to-be}: 100%
3CFf4AGlbEchvmFeRIHcCG	2FAjB5xJr21RdQcsKQE3Ut	Coverage Obj_1^{as-is}: 5.91% Coverage Obj_3^{to-be}: 100%
3CFf4AGlbEchvmFeRIHcCG	37NopSpLj37vnjyCBTBUoN	Coverage Obj_1^{as-is}: 40.14% Coverage Obj_4^{to-be}: 94.77%
3320QKjsPDs8e51pe4k$Fe	2ew7MBEPb01OZYaDjw2rmp	Coverage Obj_2^{as-is}: 100% Coverage Obj_5^{to-be}: 99.56%

the level of the renovated envelope assembly. The query returns, for each interface, the identifiers of the connected objects and their corresponding thermal transmittance values, providing a structured and machine-interpretable dataset that can be directly consumed for system-level energy performance evaluation. This demonstrates how geometry-derived semantic relations act as an enabling layer for performance-oriented data extraction without requiring IFC model merging or manual data preparation.
Importantly, the query operates at interface level rather than at individual object level, implicitly shifting the data extraction perspective from single BIM objects to physically reconstructed envelope systems. The query is independent of specific IFC schemas or modeling conventions and can be reused across different renovation scenarios, provided that geometry-derived interfaces have been materialized in the linking graph. In this context, a SPARQL query was defined to retrieve, for each detected interface, the identifiers of the connected as-is and to-be objects with their associated thermal transmittance values (Listing 1). The query selectively extracts thermal properties only for object pairs that have been previously identified as physically adjacent through geometry-driven analysis, ensuring that energy-related data are aggregated at the envelope-system level rather than at individual object level.

```
PREFIX rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX bot:    <https://w3id.org/bot#>
PREFIX props: <http://lbd.arch.rwth-aachen.de/props#>
PREFIX opm:    <https://w3id.org/opm#>
PREFIX schema: <http://schema.org/>
SELECT DISTINCT ?interface ?guidA ?U_A ?guidB ?U_B
WHERE {
  ?interface a bot:Interface .
  ?interface bot:interfaceOf ?objA, ?objB .
  FILTER(STR(?objA) < STR(?objB))
  ?objA props:globalIdIfcRoot ?gA .
  ?gA opm:hasPropertyState ?sA .
  ?sA schema:value ?guidA .
  ?objB props:globalIdIfcRoot ?gB .
  ?gB opm:hasPropertyState ?sB .
  ?sB schema:value ?guidB .
  ?objA ?pA ?propA .
  FILTER(CONTAINS(LCASE(STR(?pA)), "thermaltransmittance")) .
  ?propA opm:hasPropertyState ?psA .
  ?psA schema:value ?U_A .
  ?objB ?pB ?propB .
  FILTER(CONTAINS(LCASE(STR(?pB)), "thermaltransmittance")) .
  ?propB opm:hasPropertyState ?psB .
  ?psB schema:value ?U_B .
}
ORDER BY ?guidA
```

Listing 1: SPARQL Code for extracting thermal transmittance values

Discussion

The case study demonstrates the effectiveness of the proposed approach in identifying physically meaningful adjacency relations between objects belonging to independent BIM models and in enabling their subsequent semantic integration for performance-oriented analyses. Despite the limited scale of the test scenario, the results confirm the robustness of the geometry-driven topological analysis in accurately detecting real physical contacts while systematically excluding irrelevant adjacencies, such as geometrically coplanar but non-touching surfaces and regions associated with openings.
A key contribution of the proposed method lies in its ability to operate across independent as-is and to-be models without requiring prior alignment, model merging, or manual re-modelling. Through geometry-driven topological analysis and semantic linking, the method explicitly encodes which objects from different models are physically related in the built environment. This capability is particularly relevant in renovation scenarios, where the physical reality of the building envelope is distributed across multiple BIM representations that remain semantically disconnected in conventional workflows.
By creating object-to-object adjacencies as explicit semantic relations, the method enables the reconstruction of coherent envelope systems that exist in reality but are not directly represented in the virtual models. In practice, multiple retrofit components, such as insulation panels, can be correctly interpreted as belonging to the same functional

wall assembly, despite being modeled as separate objects in the to-be model. This transformation allows the machine to reason about which objects jointly contribute to the same physical construction package, overcoming a well-known limitation of geometry-only or model-centric BIM analyses.

From a data exploitation perspective, this semantic reconstruction directly supports the automated extraction and aggregation of performance-related properties. System-level data, such as thermal transmittance (U-values), can be computed by combining properties originating from distinct models but semantically linked through the identified interfaces. As a result, the output of the proposed workflow constitutes a ready-to-use, structured dataset that can be consumed by external energy simulation engines and emerging Digital Twin platforms, without additional manual interpretation or data preparation.

More broadly, the approach contributes to bridging the gap between raw geometric evidence and higher-level semantic representations that are required for advanced analysis and decision-making. This enables comparison of renovation strategies, supporting building-level decision-making. By making physical relationships between independently modeled objects explicit and machine-interpretable, the proposed method enhances interoperability, traceability, and automation in renovation-oriented BIM workflows.

Conclusion

This paper presented a geometry-driven and ontology-based approach to identify and semantically represent adjacency relations between independent BIM models in renovation scenarios, enabling cross-model reasoning and reconstruction of envelope systems without model merging. By combining surface-based topological matching with semantic linking through the BOT ontology, the proposed method enables the explicit representation of physical relations not encoded in BIM data, allowing fragmented objects to be integrated into coherent building systems.

The case study demonstrates the robustness of the approach in detecting adjacencies while excluding irrelevant configurations. It further shows that fragmented retrofit objects can be correctly interpreted as part of coherent envelope systems, reflecting real construction logic rather than the modeling constraints of individual BIM models.

Through the creation of object-to-object adjacencies as machine-interpretable semantic relations, the proposed workflow bridges the gap between raw geometric evidence and structured semantic data. This enables the automated aggregation and extraction of performance properties, supporting system- and building-level performance analysis and enabling comparison of alternative renovation scenarios, while providing a ready-to-use input for energy simulation, and emerging Digital Twin platforms, without requiring manual model alignment or re-authoring. In conclusion, the proposed approach enables a shift from object- to system-level reasoning, which is essential for supporting renovation decision-making processes.

Despite the promising results, some limitations can be identified. The accuracy of the geometry-driven topological analysis depends on the quality of the IFC geometric representations, which may vary across authoring tools and levels of detail. Furthermore, the present validation is limited to a small-scale case study and does not provide a comprehensive evaluation of computational performance on large and highly detailed models. Future research will focus on validating the approach on larger and more complex case studies, such as multi-wall assemblies, multiple openings, and non-orthogonal geometries, as well as on systematically evaluating its computational performance and scalability in real-world renovation scenarios.

References

Gan, V. J. L., Wong, J. K. W., and Tse, K. T. (2022). A graph-based data model for modular building design and analysis. Automation in Construction, 140:104337.

Han, Q., Xie, X., and Lu, W. (2025). Unified network-based representation of bim models for efficient data analysis. Automation in Construction, 165:105321.

Ismail, A., Nahar, A., and Scherer, R. (2017). Application of graph databases and graph theory concepts for advanced analysing of bim models based on ifc standard. In Proceedings of the 25th EG-ICE International Workshop.

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.

Khalili, A. and Chua, D. K. H. (2015). Ifc-based graph data model for topological queries on building elements. Journal of Computing in Civil Engineering, 29(3):04014046.

Massafra, A., Jabi, W., and Gulli, R. (2024). Topological bim for building performance management. Automation in Construction, 166:105628.

Massafra, A., A.-H. D. H. S. L. . J. W. (2025). Semi-automated dataset generation for residential buildings using graph-based topological modelling. Buildings, 15(8):1283.

Moreira, J., D.-A. P. P. B. E. v. E. T. (2024). Onto4reuse: Towards an ontology reuse framework for knowledge-intensive software engineering. In Proceedings of the Joint Ontology Workshops (JOWO).

Oraskari, J., Bonduel, M., McGlinn, K., Pauwels, P., Priyatna, F., Wagner, A., Kukkonen, V., Steyskaland, S., Lehtonen, J., and Lefrançois, M. (2024). Ifctolbd: Ifctolbd v 2.44.0.

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.

Rasmussen, M. H., Lefrançois, M., Bonduel, M., and Pauwels, P. (2017). The building topology ontology. Semantic Web, 9(2):1–11.

Utkucu, D., Sacks, R., et al. (2025). Ontology for holistic building performance modeling and analysis. Automation in Construction, 175:106197.