Persistence and linking of geometric-topological data models in BIM and GIS

Sophie Lipskaya[1], Robert Jürgensen[1] and Maximilian Sternal[1]
[1]Chair of Computing in Civil Engineering, Technische Universität Berlin, Germany
E-mail(s): s.lipskaya@campus.tu-berlin.de, r.juergensen@campus.tu-berlin.de

Abstract: The integration of Building Information Modeling (BIM) and Geographic Information Systems (GIS) requires robust geometric-topological consistency to avoid the costly extract-transform-load and repair processes that burden current workflows. This paper addresses three critical requirements for sustainable geometric-topological models: (a) consistent modeling tools that maintain topological integrity throughout the modeling process, (b) persistent data structures with stable identification for specific geometric and topological entities such as faces and solids, and (c) linking strategies that preserve geometric and semantic relationships without relying on error-prone data conversion pipelines. This approach is presented through the Partition Platform, a prototype modeling tool that begins modeling processes in partitioned rather than empty space, always ensuring a valid partition of the Euclidian space and maintaining full adjacency information throughout the workflow. The study demonstrates how modern JSON-based schemas, specifically IFC5 and CityJSON, represent progress toward geometric-topological persistence. Building on these foundations, it is shown how JSON-LD can extend beyond semantic linking to establish explicit connections between geometric elements, enabling direct logical and topological relationships between faces and spatial features across BIM and GIS domains. This approach eliminates the need for lossy geometric translations, instead enabling lightweight interoperability through persistent geometric identifiers. The resulting framework supports robust data exchange for digital twins, infrastructure modeling, and spatial analysis applications where geometric-topological integrity is essential for reliable computational results.

Keywords: Building Information Modeling (BIM), Geometric Modeling, Topological Relationships, Linked Data, Geographic Information Systems (GIS)

DOI: 10.18154/RWTH-CONV-254875. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,

1 Introduction

Geometry is the foundation of digital representations of the built environment, yet most of the common Building Information Modeling (BIM) tools show geometric and topological inconsistencies. Most geometric representations are not based on explicit topological entities, such as faces, but on procedural or mesh-based geometry, making boundaries ambiguous and preventing the storage of topological information such as inner faces bounding rooms and outer faces defining the building envelope. This limits both geometric clarity and the possibility of embedding topology directly in the model. Partially this comes from the object-based modeling paradigm used in current geometric modeling tools. Thus, adjacency and connectivity are not inherent, but must be reconstructed from raw geometry. Many models also show geometric inconsistencies such as overlaps or gaps, and there is typically no stable identification of geometry entities, making reliable linking and long-term data management difficult.

Transformation-based integration between BIM and Geographic Information Systems (GIS) further increases these issues. Geometric simplification, which represents the most common form of loss (see Figure 1), often occurs when parametric BIM geometry is converted into simpler primitives [1], while topological relationships embedded in the source format may be lost or require reconstruction. Level of Detail (LoD) mismatches create further problems, where reconstructing detailed BIM from GIS data often produces semantically poor models [2].

The GIS side, however, shows some advantages. Formats such as CityJSON already provide explicit boundary-based geometry, maintain consistent hierarchies and identification of every face, which allows elements to be linked easily, although they lack the semantic richness and parametric flexibility of BIM.

Improving the situation requires face-based persistence to capture both geometry and inner–outer face relationships, modeling techniques that ensure geometric and topological integrity, approaches that link rather than transform data, and formats that persistently store identifiers with geometry. Connecting BIM and GIS is essential because BIM offers building-scale semantics and object-oriented geometry in high detail, while GIS focuses on large-scale spatial organization and analysis. Together, they can support holistic, multi-scale digital twins that support applications such as city-wide energy analysis, infrastructure planning across multiple responsibilities, and climate resilience assessment, combining building details with their wider spatial context.

This paper addresses these challenges by analyzing geometric and topological persistence in BIM and GIS, focusing on modern JSON-based serialization formats such as CityJSON and IFC5. For the generation of consistent geometric-topological entities, it introduces the Partition Platform, a prototype for space partitioning–based modeling with robust topology via cellular decomposition and pointer-based navigation, providing persistent identifiers. JSON-LD is used to demonstrate cross-domain linking strategies in two examples that retain original data structures while enabling modular integration. The paper concludes with an outlook on the use of linked data with robust geometry persistence in the future.

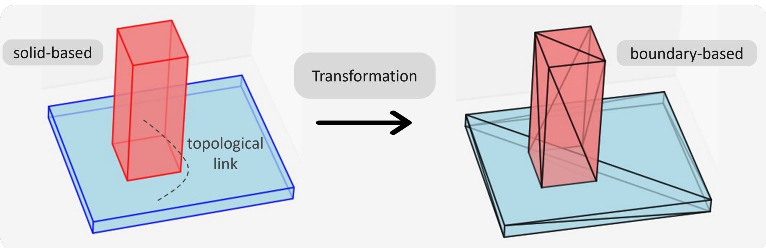

Figure 1: Information loss due to transformation from topological solid- to inconsistent mesh-based geometry.

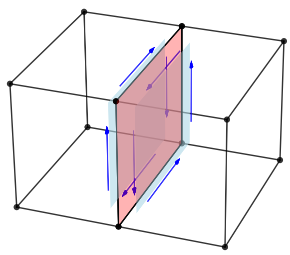

Figure 2: Geometric and topologigal model of common face (red) of two cubes in the Partition Platform.

2 Background

Understanding the geometric and topological persistence challenges in BIM-GIS integration requires examining the evolution of data models and serialization formats, the various approaches to format transformation, and emerging linked data alternatives. This background provides the foundation for understanding why current transformation approaches often result in information loss and how modern JSON-based formats, combined with robust topological modeling and linked data strategies, can address these limitations.

2.1 Introducing the Partition Platform

The Partition Platform (PP) is based on the partition of the entire unbounded Euclidean space from the beginning of the modeling process onwards. It defines clear polyhedral spatial boundaries to separate subspaces and maintain adjacency information and geometric robustness in every modeling step [3]. The topological pointer system successfully maintains edge-face and vertex-edge connectivity relationships, with cellular partitioning boundaries showing particular robustness in resolving modeling inconsistencies [4].

PP supports unbounded domains, as well as non-convex and multiply connected geometric configurations. It explicitly handles oriented and unoriented geometry for points, edges, faces, and volumes (cells). An example of a shared surface of two adjacent cubes in the Partition Platform is shown in Figure 2. The two cubes are a subset of the partition model. The face (red) shows the common geometry, and the blue facets are directed polygons with their normals pointing to the adjacent space (topology).

For persistent data storage, the topological model can be serialized into a nested JSON structure in which each topological element is uniquely identifiable by an ID. Coordinates are stored as a list of vertices, while topology and geometry are kept separate. Model reconstruction from this structure has a computational cost of order $O(n)$, with n as the total number of topological entities in the model.

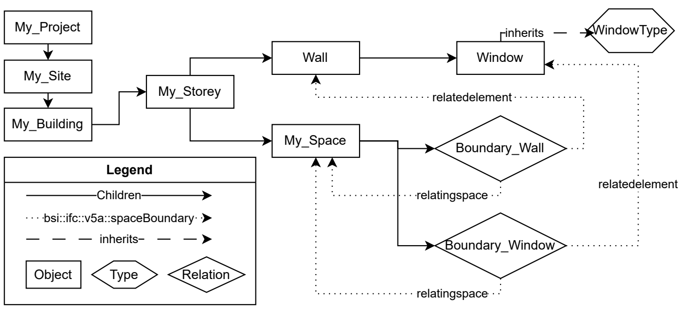

Figure 3: Hierarchical structure of IFC5 elements, illustrating parent-child and adjacency relationships

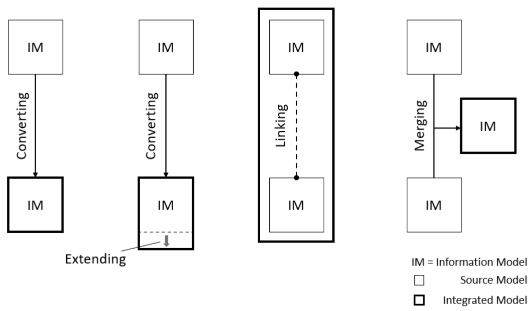

Figure 4: Different Information Model Strategies, with focus on linking [2]

2.2 Linked Data

Traditional approaches to BIM-GIS integration rely on format conversion to achieve interoperability, where data from one domain is transformed into data structures and geometric representations of another. Although conceptually straightforward, these transformation methodologies face, as mentioned above, significant challenges in preserving geometric accuracy and topological relationships throughout the conversion process.

The fundamental distinction between linking and conversion approaches lies in their treatment of source data and integration methodology (see Figure 4). Traditional conversion approaches create derivative models that represent information from one domain using the data structures of another, inevitably resulting in loss of information and structural adaptation [1]. In contrast, linking approaches maintain the original data models unchanged while establishing explicit cross-domain connections through definitions of external relationships [2], [5].

This way, the integrity of the source model data is preserved, which has several advantages for geometric and topological persistence. First, parametric geometric representations remain intact within their native formats, allowing continued model evolution and design iteration without degradation. Second, domain-specific topological relationships are preserved within their original structural contexts, maintaining the spatial hierarchies and connectivity information essential for domain-specific analysis. Third, the linking approach supports bidirectional integration without the asymmetric information loss characteristic of transformation workflows.

```
{  "path": "c8ecbf4c-e37a-4489-9133-15163
     b8a904e",
   "attributes":
   {
     "bsi::ifc::v5a::spaceBoundary": {
       "relatedelement": {
         "ref": "93791d5d-5beb-437b-b8ec-2
             f1f0ba4bf3b"
       },
       "relatingspace": {
         "ref": "e3035b71-bd9f-4cdc-86fd-
             b56e2f4605b6"
       }}}}
```

```
{  "@context": {
     "ifcx": "https://example.org/ifcx#",
     "cj": "https://example.org/cityjson#",
     "owl": "http://www.w3.org/2002/07/owl#"
   },
   "@graph": [
     { "@id": "ifcx:geometry424",
       "@type": "ifcx:WindowType",
       "owl:sameAs": "cj:window6"
     },
     { "@id": "cj:window6",
       "@type": "cj:Window",
       "owl:sameAs": "ifcx:geometry424" }]}
```

Figure 5: Side-by-side JSON representations: (left) topological relation, (right) linked window geometry in JSON-LD.

JSON-LD Implementation Requirements: JSON-LD employs triple-based structures specifying subject-predicate-object connections between elements across data sources, enabling definitions of spatial adjacency, geometric overlap, scale-based correspondence between BIM elements and GIS features (Figure 5, right), and explicit topological hierarchies that span domain boundaries, such as linking IFC spaces to CityJSON building rooms, as illustrated in Figure 7 with the green arrow. Its distributed nature allows data sources to remain in their original locations while maintaining logical integration through web-accessible relationship definitions [5]. Successful implementation depends on consistent topological models that preserve geometric precision and spatial relationship integrity across the linking framework, motivating specialized topological platforms for stable geometric foundations [3].

Implementation requires three fundamental components: unique identifiers for linkable elements, structured relationship definitions, and accessible data locations.

2.3 CityJSON and IFC5

BIM and GIS have traditionally relied on distinct data models and serialization approaches that reflect their different spatial scales and application domains. Although these formats have served their respective communities effectively, their structural differences and geometric representation methods present significant challenges for cross-domain integration.

Established Formats: CityGML employs an XML-based serialization format designed for representing 3D city models with semantic information across multiple levels of detail. The format structures geometric data using boundary representation (B-rep) approaches, where buildings and urban features are represented through their surface boundaries. However, the XML-based structure creates a hierarchical tree representation that requires deep navigation to access geometric primitives, making programmatic access and web-based applications less efficient [6].

IFC, with its latest version 4.3, represents an even more complex hierarchical approach to storing BIM data, typically serialized using the STEP format where each line represents a specific object with its attributes. The IFC schema supports numerous geometric primitives including Constructive Solid Geometry (CSG), boundary representations, extrusions, swept surfaces, and parametric shapes, providing extensive flexibility for detailed building component representation [7]. However, this geometric richness comes at the cost of accessibility. The hierarchical referencing system, while powerful, creates barriers to efficient geometric extraction and spatial relationship analysis.

Modern JSON-Based Approaches: The limitations of XML and STEP-based formats have driven the development of JSON-based alternatives that prioritize structural simplicity and web compatibility while maintaining semantic richness. CityJSON version 2.0, standardized in 2023 [6], provides a simplified JSON representation of the CityGML data model. The format's flat structure allows direct access to geometric objects without deep tree navigation, making it significantly more accessible for web technologies and modern development environments.

IFC5, currently under alpha version development by buildingSMART, represents a complete reboot and fundamental shift toward modular, JSON-based BIM data representation [8]. Unlike previous attempts to create JSON encodings of IFC4 that merely translated the hierarchical schema to JSON syntax, IFC5 redesigns the data model itself to leverage JSON's strengths [9]. The new format employs a flat, modularized structure where object order within files becomes arbitrary since geometries are identified through unique IDs rather than local references. This modularity enables selective loading of specific data components based on application requirements [8]. Parent-child relationships are maintained through explicit reference systems rather than structural nesting [8]. JSON approaches support enhanced topological modeling through three distinct relationship types maintained via identifier-based references (see Figure 3). Topological relationships are explicitly maintained through space boundary objects linking spaces to elements via relatingSpace and relatedElement attributes, preserving interface geometry (see Figure 5 left).

IFC5's flat structure with unique object IDs enables direct geometric and semantic linking, while CityJSON's identifier system supports feature-level and face-based connections, contrasting favorably with IFC4's reference-based geometry system, which lacks explicit geometry identifiers necessary for external linking [10].

3 Examples

The following examples demonstrate practical applications of the concepts discussed in the background section, illustrating how robust topological modeling and linked data approaches can address common challenges in BIM-GIS integration. These examples highlight both the potential and limitations of current approaches while demonstrating the effectiveness of combining topological consistency with modular JSON-based linking strategies.

3.1 Topological Relationships and Geometric Consistency

Analysis of the buildingSMART `Hello-Wall` IFC5 example model [8] reveals three significant geometric inconsistencies that demonstrate the need for robust topological modeling and persistence. First, multiple vertices and triangles appear in the geometry mesh where the space body contains each point three times, resulting in redundant representation (24 vertices instead of 8 for a simple box). Second, rounding inaccuracies and unnecessary transformations create potential geometric clashes or gaps, such as the boundary wall positioned at $-2.3841858e-8$ instead of the expected 0.0. Third, boundary surface mismatches occur as the `Boundary_Wall` consists of two triangles forming a rectangle, while the corresponding wall surface has several triangles due to window openings. The shape mismatch of the contact surfaces can be seen in the explosion graphic in Figure 6. The boundary surfaces are colored red and yellow for `Wall` and `Window` respectively.

By generating IFC5 objects from the Partition Platform or linking to corresponding topological model elements in the PP-JSON file via JSON-LD, consistency can be maintained while preserving the modular structure of the source data. Since not only solids (cells), but each topological entity (nodes, edges, faces) is identifiable in the Partition Platform, a correct adjacent surface with holes can be exported or linked to `Boundary_Wall` from the IFC5. This way, not only is the topological consistency guaranteed, but also the surfaces of neighboring elements would identically match.

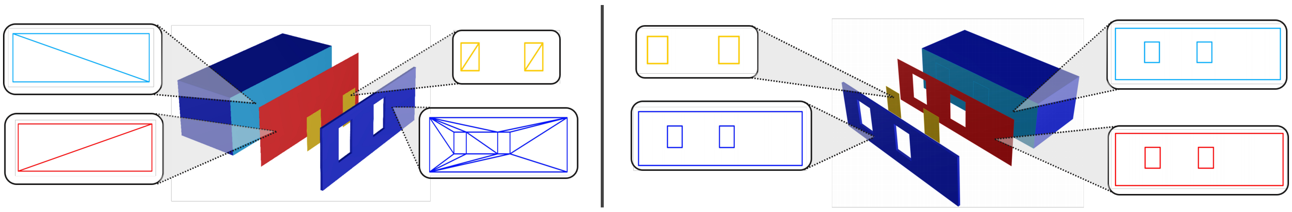

Figure 6: Explosion of selected objects of `Hello-Wall` IFC5 model [8] (left) and of respective objects in the Partition Platform (right)

3.2 BIM-GIS Linking with Modular JSON Structures

Using the IFC5 `Hello-Wall` model and a corresponding CityJSON representation created from [10], two linking approaches demonstrate the versatility of JSON-LD for BIM-GIS integration. The first approach links building elements at the object level, connecting the IFC space with the corresponding CityJSON `BuildingRoom` using a `owl:sameAs` relationship. This semantic linking maintains object identity across domains while preserving the original data structures in both files and is compatible with the current IFC standard, as object-level identifiers are already supported in IFC4. The second approach demonstrates geometry-level linking, connecting the detailed window geometry from the IFC5 file to the simplified window object in CityJSON, as shown in Figure 5, right. Unlike object-level linking, this geometry-level linking is only possible with IFC5, since the current IFC4 standard does not assign unique identifiers to geometry resources. Thus, IFC5 enables applications that require detailed BIM geometry combined with GIS spatial context without a complete format transformation. The linking visualization in Figure 7 illustrates how the same building in different formats can be simultaneously accessed through domain-specific viewers while maintaining logical connections through JSON-LD relationships. In this figure, the green arrow represents the object-level link between the room entities in both domains, while the red arrow indicates the geometry-level link between the IFC5 window geometry and the corresponding CityJSON window object, as seen exemplarily on the right in Figure 5.

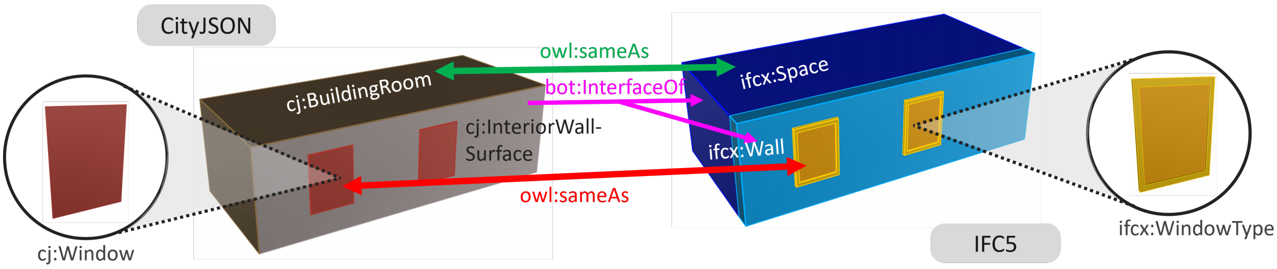

Figure 7: Visualization of linked elements: IFC viewer on the left and CityJSON viewer on the right

In the context of urban planning, another potential application of Linked Data is to establish direct connections between planned building models in BIM and the corresponding property features in GIS. This can be achieved by linking a property boundary and neighboring building connections, represented as a `WallSurface` in the CityJSON dataset, to the corresponding exterior wall geometry solid or boundary surface in an IFC5 model (blue and red meshes in Figure 6).

4 Conclusions and Outlook

This study demonstrates that modular JSON-based formats offer significant advantages for BIM-GIS integration over traditional hierarchical approaches. The transition to flat JSON structures with unique identifier systems, also for geometries, enables the possibility to add robust topological relationships and flexible cross-domain linking without information loss. The Partition Platform offers modeling geometry with topological consistency across serialization cycles while internally addressing geometric inconsistencies in current implementations that have also been shown in the minimal example presented in this paper. JSON-LD linking strategies preserve source model integrity while establishing semantic connections, representing a paradigm shift from transformation-based to preservation-based integration. This approach opens opportunities for integration with semantic web technologies and ISO 21597 (ICDD) standards [11]. The explicit linkable nature of geometries in IFC5 and CityJSON enables participation in broader knowledge graphs and standardized information containers, supporting digital twin and smart city applications.

Future research could focus on the integration of IndoorJSON for indoor navigation and connectivity modeling, extending the presented approach from city- and building-scale to interior environments. Another important aspect is the development and adoption of standardized linking vocabularies for spatial relationships, addressing data governance challenges in distributed linked data scenarios. The

Building Topology Ontology (BOT), developed by the W3C Building Data Group, already provides vocabulary for describing building topology [12]. The pink arrows in Figure (7) show such an application. The demonstrated approaches for persistent, identifiable geometric and topological data point toward building and urban models that maintain domain-specific strengths while participating in flexible, modular integration networks that are open to user-specific data models.

References

[1] C. Clemen, F. Gruner, and J. Pfeifer, “Infrastrukturdatenhaltung mit BIM und GIS”, Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt, 2023. DOI: 10.48755/DZSF.230015.01

[2] S. F. Beck, “Context-sensitive linking of heterogeneous information models from the building and the urban domain”, Ph.D. dissertation, TU München, 2023.

[3] M. Sternal, “Polyhedral partition of the euclidean space for consistent spatial modeling of the built environment”, Ph.D. dissertation, TU Berlin, 2025. DOI: 10.14279/depositonce-23869

[4] W. Huhnt, M. Sternal, and P. J. Pahl, “Modeling bounded and unbounded space with polyhedra: Topology and operators for manifold cell complexes”, *Advanced Engineering Informatics*, vol. 54, Oct. 2022. DOI: 10.1016/j.aei.2022.101790

[5] P. Pauwels, A. Costin, and M. H. Rasmussen, “Knowledge graphs and linked data for the built environment”, in *Industry 4.0 for the Built Environment*, Cham: Springer International Publishing, 2022. DOI: 10.1007/978-3-030-82430-3_7

[6] OGC. “CityJSON community standard”, Accessed: Jul. 1, 2025. [Online]. Available: http://www.opengis.net/doc/CS/covjson/2.0

[7] F. Noardo, “Multisource spatial data integration for use cases applications”, *Transactions in GIS*, vol. 26, no. 7, pp. 2874–2913, Nov. 2022. DOI: 10.1111/tgis.12987

[8] buildingSMART International. “Ifc5 development – github repository”, Accessed: Jul. 1, 2025. [Online]. Available: https://github.com/buildingSMART/IFC5-development

[9] L. van Berlo et al., “Future of the industry foundation classes: Towards IFC 5”, *Proc. of the 38th International Conference of CIB W78*, pp. 123–137, 2021.

[10] H. İ. Şenol and T. Gökgöz, “Integration of building information modeling (BIM) and geographic information system (GIS): A new approach for IFC to CityJSON conversion”, *Earth Sci Inform*, vol. 17, Aug. 1, 2024. DOI: 10.1007/s12145-024-01343-1

[11] J. Krischler, P.-C. Schuler, J. Taraben, and C. Koch, “Using icdd for bim and gis integration in infrastructure”, *LDAC 2024: 12th Linked Data in Architecture and Construction Workshop, June 13–14, 2024, Bochum, Germany*, 2024.

[12] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels, “BOT: The building topology ontology of the w3c linked building data group”, *Semantic Web*, vol. 12, pp. 143–161, 2020.