



**HAL**  
open science

## Semantic topological querying for compliance checking

N. Bus, F. Muhammad, B Fies, A. Roxin

► **To cite this version:**

N. Bus, F. Muhammad, B Fies, A. Roxin. Semantic topological querying for compliance checking. 12th European Conference on Product and Process Modelling (ECPPM 2018), Sep 2018, Copenhagen, Denmark. 10.1201/9780429506215-57 . hal-02270841

**HAL Id: hal-02270841**

**<https://cstb.hal.science/hal-02270841>**

Submitted on 26 Aug 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Semantic topological querying for compliance checking

N. Bus, F. Muhammad & B.Fies

*Centre Scientifique et Technique de Bâtiment, Sophia-Antipolis, France*

A. Roxin

*Université de Bourgogne, Dijon, France*

**ABSTRACT:** Architects and construction engineers need services to check their designs against specific standards, regulations and policies. Many works have been done during the last years to develop checking software. French National project “Digitizing building regulations” aims at formalizing regulation for automatic compliance checking purposes. Our approach is cloud ready, opened, extensible and based on international standard. In this paper we focus on semantic topological aspects to show gains, limits and perspectives of this approach.

## 1 INTRODUCTION

In a world where industrial operations are increasingly complex and globalized, there is a growing need for delivering services informing the users about the levels of “safety” and compliance of a given building. While technical experts and audit specialists are moving towards BIM (Building Information Modeling), model-checking software allow to control building operations and processes, starting from the first lifecycle phase – building design. This improves project standard compliance while ensuring related costs remain affordable.

In this study, we will focus on how geometrical and semantic constraints can simultaneously be computed in a compliance checking context. This work is part of the French national project “Digitizing building regulations” (« Numérisation des règles ») which aims at converting existing AEC laws and regulations into logic rules.

## 2 RELATED WORK

### 2.1 Model checking solutions

The traditional approach of verification using MVDXML [1] [14] is very limited. Major limitations have been identified such as restricted scope of applying conditions and constraints on several branches of an IFC model; poor geometric analysis of an IFC model; lack of mathematical calculations;

support of only static verification of a model, etc. Several works introduced a semantic rule engine oriented [2] approach as a viable alternative.

### 2.2 Semantic approach

In the context of BIM (Building Information Modeling), modelling building elements as resources has been identified as an interesting approach for achieving information interoperability (Pauwels, et al., 2011 and Farias, et al., 2014). Existing IFC-related ontologies were conceived as direct syntax mappings between EXPRESS and OWL languages (Beetz, et al., 2009 and Pauwels, et al., 2011). One of the latest and the most solid implementations of an IFC ontology is IfcOWL proposed in (W3C Linked Building Data Community Group, 2014). This version of IfcOWL is also a Candidate Standard (buildingSMART, 2017) for buildingSMART (meaning it is considered as an activity in the process of acquiring international consensus before being submitted to the Standards Committee for a final vote).

In the last years, several initiatives were proposed by the research community, most of them relying on Semantic Web technologies for addressing the issues that appear when relying solely on IFC models. Based on these, buildingSMART International accepted ifcOWL (OWL serialization of the IFC standard) as a candidate standard. Indeed, ifcOWL and its sub-related graph structure allow solving some issues regarding data partitioning, querying and reasoning. By doing so,

buildingSMART International (notably its Linked Data Working Group - LDWG) have opened the road to expert model-checking solutions. Still, at this stage, several questions remain unanswered. Most requirements address simultaneously geometrical and semantics aspects.

Taking IFCOWL and IFC to RDF converter as a starting point, we have built transformation services and a methodology that is described in detail in the following chapters.

### 2.3 Collaborative BIM platform

In March 2018, the KROQI platform has been officially launched. It is a collaborative cloud platform for construction SMEs available free of charge to all construction stakeholders including SMEs and VSEs.

It provides document management functionalities, collaborative services, and innovative business services all along the building life cycle.

The MVXML technology is deployed on the KROQI platform to check the quality of the IFC files that are uploaded by the users. It checks if the structure of the IFC files is compliant with project policies.

The semantic checker to be developed in the frame of the « Numérisation des règles » project will complement this first checker. It will be integrated to the KROQI platform and provides extended rule checking capabilities such as checking topology.

## 3 METHOD

### 3.1 Validating modeling options

The Building Executive Plan (BEP) describes how construction project stakeholders agree to model and share information. This document describes data exchange processes, modelling best practices, exports parameters, nomenclatures and classifications to be used. Most of the time the document references international classification frameworks like Unifomat or Omniclass. Having all actors of the project share the same modeling rules drastically decrease the number of models' variants. The BEP is also used to extend IFC properties with ad-hoc property sets dedicated to specific topics, such as certification level, equipment performances or environmental impact of product. CSTB aims at providing a

national modeling charter core to be used as a common starting point when engineering the modeling plan of a project. We are working together with architects, building owners, engineers, audit experts, building quality control agencies and government to reach this goal.

### 3.2 Enriching model with semantic inferences

The semantic approach allows to describe a building as well as a requirement by using the same atomic fact formalism called triplet. The IFC and so ifcOWL standard vocabulary includes hundreds of classes and properties representing the building physical layer. This vocabulary is not natively designed to deal with national considerations, building performances, requirements and building functionalities.

During the first phase of the « Numérisation des règles » project, a first regulation vocabulary has been extracted from a set of regulatory documents. While trying to align IFC vocabulary with regulatory vocabulary we showed that this alignment cannot be based on a simple bijective relation. Indeed, some regulatory concepts can only be defined with a logical statement aggregating various IFC classes. For instance, in our regulation ontology, a simple regulatory concept like “highest storey” is inferred by comparing the “elevation” (ifcOWL property) values of the all “IfcBuildingStorey” (ifcOWL class) for a given building.

The Regulatory ontology is composed by complex regulatory concepts defined on top of ifcOWL. Regulatory concepts are organized in layers. The ifcOWL vocabulary is the ground layer of a vocabulary pyramid. At the very top of the pyramid (see figure 1), we find the very regulatory-specific vocabulary. A term of a specific layer is defined by using terms belonging to the lower layer. This layered approach is flexible as it keeps high-level definitions simple by increasing the number of terms.

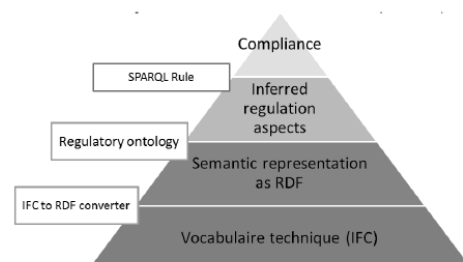


Figure 1. Regulation vocabulary layers

The whole process that provides a semantic regulatory view from the IFC source model works as follows:

firstly, we convert IFC model into an ifcOWL triplet database by using the IFC-to-RDF-converter [3] - provided by BuildingSmart-;

A custom class and relations filter is used to remove non-significant information (concepts not involved in any rule). This operation can be compared to a MVD -Model View Definition- dedicated to regulation. As some IFC terms are almost never addressed by regulatory texts -such as high-level geometry, element history or sensor states, they can be filtered from the model to reduce the amount of data.

Then a “geometrical preprocessor” renders and computes additional geometry aspects. This preprocessor is detailed in the following chapters. Eventually a “semantic preprocessor” infers the model according to the regulatory ontology.

As a major advantage, this last transformation step offers a simplified graph from ifcOWL model [11].

For instance, referencing to classification (e.g. Unifomat) takes at least ten triplets with ifcOWL vocabulary whereas one triplet is sufficient with the regulatory ontology.

### 3.3 Geometrical preprocessor

The regulation describes constraints that involve topological relations between objects such as: is below, is inside, is adjacent to. IFC allows to describe elements with various representation models: extrusion, BREP, boolean operations. To compute topological relations a low-level geometry such as bounding boxes is sufficient.

The building boxes are described by using only 6 semantic relations from the regulation ontology: Xmin, Xmax, Ymin, Ymax, Zmin and Zmax (see figure 2).

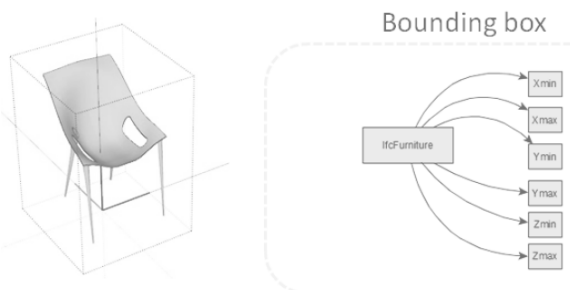


Figure 2. Bounding box corresponding to geometry of a chair

The geometric preprocessor infers on geometry to compute triplets materializing bounding box representation of building objects (e.g. furniture20 xMin 50.12). According to our tests reasoning on the bounding box representation instead of BREP or geometrical operations representation are reliable and low CPU consuming.

### 3.4 Formalizing regulatory requirements

This section provides details on the methodology used to transform the regulatory constraints, expressed in natural language, into processable queries.

Firstly, regulatory texts are prepared and interpreted by working groups. Experts use a text editor with autocompletion and syntax highlighting capabilities to re-write regulatory texts into semi-formal rules [18]. Each semi-formal rule is supposed to detect non-compliant building elements concerning a very specific aspect. Depending on its complexity, a regulatory rule, in texts, can be divided into several atomic semi-formal rules dealing with complementary aspects. The idea is to keep each semi-formal rule as simple as possible. As a guideline, we suggested that each semi-formal rule begin with “IF” followed by a condition on specific elements and ends with “THEN NON-COMPLIANT”.

The second step of this transformation brings the regulatory constraints from the semi-formal to the formal stage. This step is performed by computer scientists familiar with the BIM and with semantic technologies. The semi-formal stage makes it convenient to build formal rules. Each semi-formal rule is translated to a SPARQL query using the same level of vocabulary coming from the regulatory ontology. The direct use of the regulatory controlled vocabulary within the constraint-queries keeps them understandable for construction experts. Geometrical constraints are described by using the geoSPARQL (OGC standard) relations covering all the possible topological relations between two geometries as introduced by the Trinity College of Dublin [5]. Geometry simplification makes it easier to deal with geometry coordinates. This approach could be extended if needed by using functional extensions [6].

### 3.5 Checking regulatory requirements as a KROQI service

Automating code-compliance checking consists of chaining conversion algorithm with geometrical and semantic preprocessors introduced in §3.3. Each time new IFC model is submitted to the semantic checker this chain is triggered so that we obtain a triplet database with the right level of details, aligned with the regulatory ontology. The geometrical preprocessor consists of a java code that returns geometrical triplets from IFC. The code executes a sequence of semantic forward chaining operations that filters useless information or enhances model with high level vocabulary. Ontology alignment and extensions to IFCOWL are declared in the regulation ontology. Constraint-queries are organized by regulatory topics so that the KROQI user can select a set of constraints corresponding to his specific needs (fire safety, accessibility, ventilation, acoustics...).

## 4 SYSTEM ARCHITECTURE

Technically, the checking service is hosted by a microservice architecture among other BIM service like account manager, document sharing, annotations, instant messaging, BIM viewer. Each service is pro-grammatically independent but can communicate with others through synchronous HTTP/REST queries or through an asynchronous messages queue.

The checking service is developed according to the SaaS pattern. This service can be executed locally or integrated to any collaborative BIM platform. The service (see Figure 3) can be executed through its own user interface that provides a list of rule set to check or silently.

The REST API published by the service is simple. It consists of a unique method “check” and two parameters. The first parameter defines the path to the IFC or IFCZIP model. Constraint-queries are packed as SPARQL files in a ZIP archive. The archive also contains the metadata for end-user's information purpose. The second parameter defines the path to the ZIP file corresponding to the Constraint-queries set.

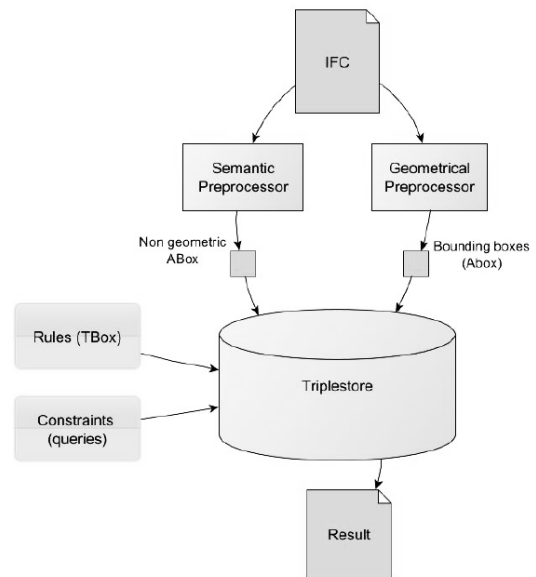


Figure 3. Checking Service Architecture

Next step of the automation algorithm consists of executing, one by one, each formal constraint-query corresponding to the chosen topic.

The result can be returned as XML, BCF, PDF or JSON format according to the accept request-header defined by the REST specification. PDF is used for user display, JSON allow to chain services and BCF can interoperate with BIM-native tools. Whatever the format asked, the result contains for each constraint-query a list of non-compliant building elements identified by their GUID.

The BCF format [15] is simple and easy to implement [16]. The basic content is an issue with comments and references to affected objects (using IFC mechanisms for Global Unique ID's). The format also supports status-information, since issues may be discussed by different users during the workflow. The BCF file makes it convenient to display within a standard IFC viewer, on top of a graphical representation of a building element the results of the evaluation. The BCF format is independent of the IFC schema version, so it can be used with IFC2x3 and IFC4.

## 5 RESULTS

We applied this approach on several constraints of the French regulation on both fire safety and accessibility domains. A first version of a French regulation ontology based on previous CSTB research works has been adapted [7]. A group of domain experts had analyzed, interpreted and finally converted a dozen of regulation texts into about one hundred semi-formal constraints

implementing concepts of the regulatory ontology vocabulary. The regulation ontology was then extended with experts' new concepts suggestions. The following paragraphs illustrate in detail two specific use cases.

### 5.1 Checking fire safety requirements

The fire safety domain regulation put some constraint on structure element performances. Indeed, from the fire safety perspective, the structure elements of the building must bear load during at least one hour in case of fire. This constraint differs according to the building height. On a fire safety point of view, the building height is measured from the ground up to the floor of the latest storey. Translated as a SPARQL query, this constraint implies various high-level concepts such as: highest level (storey with highest elevation), storey floor (lower slab of a storey), structure element (building element that bear load), loadbearing duration in case of fire. The following sequence is triggered by the constraint-query:

1. The building height according to the fire safety standard is calculated on the fly
2. The material performance threshold is determined according to the building height
3. All structure elements are retrieved and their performances are checked
4. Finally, the query returns all structure elements with fire resistance below than the threshold.

Fire safety regulation needs a circulation graph to be checked (see figure 4).

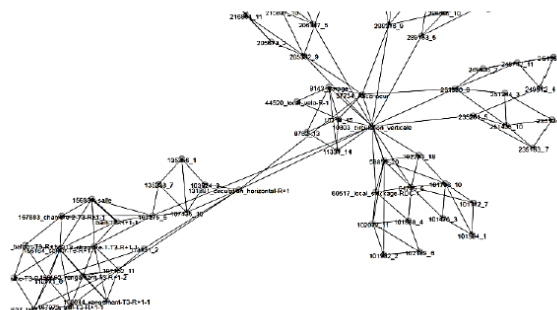


Figure 4. Circulation graph inferred for IFC model

This circulation graph can be inferred from the IFC model by using the *IfcRelSpaceBoundary* relation between spaces and openings. Sometime some expected relations are missing in the IFC model. To

complete information we also infer spaces adjacencies with a fuzzy intersection rule. The fuzzy intersection is a clash detection rule with tolerance on boundary boxes.

This architecture has been implemented on the KROQI platform providing a full automated checking process on various fire safety and accessibility topics.

### 5.2 Checking accessibility requirements

According to French building regulation, an accessible WC seat (water closet seat) must have a free space of 0.8 x 1 meter on its left OR on its right side. The WC is a high level regulatory concept corresponding to an *IfcFlowTerminal* [8] with a predefined type equal to *WCSEAT* [8]. Each free space is classified by using the high level regulatory concept *FreeSpace* (see figure 5). It represents a virtual object (not physical) with three dimensions (bounding box). According to the regulation, this object must not intersect with any other physical element of the building.

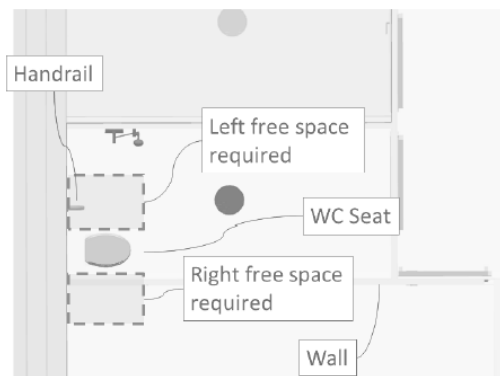


Figure 5. Materialized free space required around WC seat

The checking process of this constraint is detailed by the following sequence:

1. For each WC (water closet seat), two *FreeSpace* elements -and their bounding box- are created on-the-fly. One on the left and one on the right.
2. For each *FreeSpace*, the query retrieves all building elements that intersect.
3. Finally, the query returns all WC with at least one *FreeSpace* intersecting at least one physical building element.

In this use case, a bounding box representation of a building element is accurate enough to detect clashes with the WC. Yet, the bounding box must be oriented to be able to locate left and right sides.

This orientation can either be explicit or inferred by comparing dimensions.

In addition to the list of elements that break rules, an explanation query displays the list of elements that intersects with a *FreeSpace*. During our test we manage to detect a handrail that was settled in a wrong location.

### 5.3 Test set

This approach has been validated on various IFC models from the simplest to the heaviest by using a widely used triplestore implemented in Java with forward and backward chaining capabilities.

Table 1. Models tested

Building name	IFC size		RDF size
	Mb	Spaces	M triplets
Chanteloup	17	166	0.8
AC9 Haus	4	26	1.2
CSTB Building	15	59	2.7
EM Building	2	17	0.4
Model 23	29	17	1.9
Hitos	62	243	13.6
Liberty	12	210	2.1

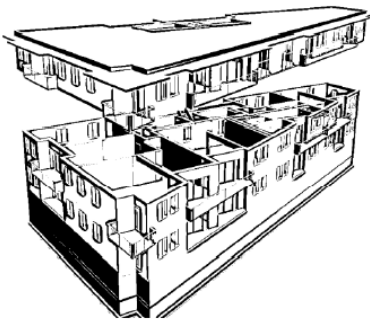


Figure 6. The “Liberty” model displayed in a IFC viewer

While the whole process takes 7 minutes for the Liberty model (2.1 M triplets) on a quad core i5 CPU at 2,5Ghz, it takes 10 minutes for the HITOS model (13.6 M triplets). IFC conversion to RDF is the most consuming phase and should be optimized.

## 6 CONCLUSION

We showed that our semantic topological approach allows addressing significant issues raised by the computation of building regulation requirements. We’ve built, explained and executed queries validating the approach by using portable SPARQL syntax and a triple store implementing W3C recommendations.

Obviously, this approach does not address all constraints defined in the regulations. For instance, when considering a complex corridor with multiples aisles - and their different related widths - the bounding box approximation will output erroneous results while computing intersection between spaces.

## 7 DISCUSSION AND FUTUR WORK

### 7.1 Enriching the regulatory ontology

Web semantic technologies bring a powerful formalism to express terminology and constraints. Yet many BIM platform practitioners do not have full understanding of web semantic technics which limit the number of contributors to a regulation ontology. We believe that it is possible to design tools that support visual query building and enable quick, easy and reusable presentation of terminologies and constraints.

### 7.2 Explaining results

Assigning an object to a rule to point out non-compliance can be insufficient to explain complex cases. The EXPLAIN capabilities of SPARQL could be used to give information on contexts and causes, other elements, properties and relation involved in the computation. Ideally the system should give advices on how to solve compliance issues. Topological requirement could be visualized within the model to point out obstructions, clashing or wrong dimensions. We could materialize those requirement as extra elements and then use the BIM-Snippet [15] brought by BCF 2.0 to return those small partial models in a standard way.

In addition, as each BCF-issue can carry a camera/viewport and even a screenshot of the situation those elements could be generated by the checker itself to illustrate the issue.

### 7.3 Expressing compliance on partial elements

For complex elements we note that compliance should be expressed on a single part of the element. For instance, a specific aisle of a complex corridor could be less large than required. As the corridor is generally described as a single space we need a new formalism to assign the issue to a specific part of the space.

#### 7.4 Improving topological treatments

Next, we should improve the geometrical preprocessor to compute and reason on topological quantities such as intersection volume, distance between elements or projection surfaces. A full integration of geoSPARQL using WKT [17] description of geometric could help.

#### 7.5 Standardizing regulation ontology

To reinforce domain interoperability and build solutions on international and acknowledged standards, we envision to improve the regulation ontology by considering the latest works done in the Linked Building Data Community group such as the Building Topology Ontology (BOT) [9], the ontology defining the core concepts of a building and OntoBREP [10] - ontology for CAD Data and Geometric Constraints.

## 8 REFERENCES

1. MVDXML specification and support implementation - <https://github.com/BuildingSMART/mvdxml>
2. Fahad, M., BUS, N., Andrieux, F.: Towards Validation of IFC Models with IfcDoc and SWRL A Comparative Study, In the Twelfth International Conference on Internet and Web Applications and Services, pp. 7-13 (2017)
3. Pauwels P., and Oraskari J.: IFC-to-RDF Converter, <https://github.com/IDLabResearch/IFC-to-RDF-converter>
4. Preidel, C., Bus, N., Borrmann, A., Fies, B.: Pre-Processing IFC Building Models for Code Compliance Checking based on Visual Programming, to appear in the proceedings of ICCBE (2018).
5. McGlenn, K., Debruyne, C., McNerney, L., O'Sullivan, D., Integrating Ireland's Geospatial Information to Provide Authoritative Building Information Models, In proc. of 13th conf. on semantic systems, pp. 57-64 (2017)
6. Chi Zhang, Jakob Beetz: Querying Linked Building Data Using SPARQL with Functional Extensions
7. Bouzidi, K. R., Zucker, C. F., Fies, B., Corby, O., Thanh, N-Le.: Towards a Semantic-based Approach for Modeling Regulatory Documents in Building Industry, 9th European Conf. on Product & Process Modelling, ECPPM (2012).
8. IFC 2x3, Specification, <http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/ifcsharedbldgs-erviceelements/lexical/ifcflowterminal.htm>
9. Rasmussen, M.H., Schneider, G.F., Pauwels, P.: Building Topology Ontology (BOT), <https://github.com/w3c-lbd-cg/bot>
10. Perzylo, A.C., Somani, N., Rickert, M., Knoll, A.: An ontology for CAD data and geometric constraints as a link between product models and semantic robot task descriptions. In proceedings of IROS 2015: 4197-4203
11. Pauwels P., Roxin, A.: SimpleBIM: from full ifcOWL graphs to simplified building graphs, Ework and ebusiness in architecture, engineering and construction. WORK AND EBUSINESS IN ARCHITECTURE, ENGINEERING AND CONSTRUCTION. pp.11-18 (2016).
12. Daum, S., Borrmann, A.: Processing of Topological BIM Queries using Boundary Representation Based Methods, Advanced Engineering Informatics, vol. 28(4), pp. 272-286 (2014)
13. Maïssa, S., Frachet, JP., Lombardo, JC., Bourdeau, M., Soubra, S.: Regulation checking in a Virtual Building, 2002
14. Zhang, C., Beetz, J., Weise, M.: Model view checking: automated validation for IFC building models. In Mahdavi, ed. eWork and eBusiness in Architecture, Engineering and Construction: ECPPM (2014).
15. BuildingSMART, BCF specification, <http://www.buildingsmart-tech.org/specifications/bcf-releases>
16. Mahdavi, A., Martens, B., Scherer, R., eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2014
17. Open Geospatial Consortium (OGC), geoSPARQL specification, <http://www.opengeospatial.org/standards/geosparql>
18. Wieringa, R., Dubois, E., Huyts, S., Integrating semiformal and formal requirements, 2005