

# Processing BIM and GIS models in practice: experiences and recommendations from a GeoBIM project in the Netherlands

Ken Arroyo Ogori      Abdoulaye Diakité  
Thomas Krijnen      Hugo Ledoux      Jantien Stoter

This is an author's version of the paper. The authoritative version is:

**Processing BIM and GIS models in practice: experiences and recommendations from a GeoBIM project in the Netherlands.** Ken Arroyo Ogori, Abdoulaye Diakité, Thomas Krijnen, Hugo Ledoux and Jantien Stoter. *ISPRS International Journal of Geo-Information* 7(8), August 2018. ISSN: 2220-9964. DOI: 10.3390/ijgi7080311

Related source code is available at:

[https://github.com/aothms/IfcOpenShell\\_CGAL](https://github.com/aothms/IfcOpenShell_CGAL)

It is widely acknowledged that the integration of BIM and GIS data is a crucial step forward for future 3D city modelling, but most of the research conducted so far has covered only the high-level and semantic aspects of GIS-BIM integration. This paper presents the results of the GeoBIM project, which tackled three integration problems focussing instead on aspects involving geometry processing: (i) the automated processing of complex architectural IFC models, (ii) the integration of existing GIS subsoil data in BIM, and (iii) the georeferencing of BIM models for their use in GIS software. All the problems have been studied using real world models and existing datasets made and used by practitioners in the Netherlands. For each problem, this paper exposes in detail the issues faced, proposed solutions, and recommendations for a more successful integration.

# 1 Introduction

Geographic information systems (GIS) have long been used to model the environment and to perform 2D spatial analyses of large areas. However, with the increasing availability of computing power, inexpensive data acquisition methods, and automated workflows that generate detailed 3D data, GIS models have become increasingly detailed and have started to contain detailed models of individual buildings—the traditional domain of building information modelling (BIM).

At the same time, the same increase in computing power and the availability of better software have enabled once rare BIM methodologies to be applied on a large scale, disrupting more traditional building design platforms based on 2D drawings. As users of BIM software want to incorporate the surrounding features of a building or another structure into their workflow, it is only logical that the BIM domain is currently enhancing its standards and software to support environmental information such as infrastructure<sup>1</sup>, and that BIM users increasingly turn to existing GIS datasets containing this environmental information. Both domains are thus now overlapping, increasingly modelling the same objects, even if the data is represented and stored in rather different ways.

While the GIS and the BIM domains clearly overlap when the modelling of cities is concerned, each domain retains its own focus and has its own characteristics. The BIM domain focuses on information about the design and construction of building sites, and thus has very detailed and semantically rich information about all the physical elements that comprise an individual building as it is designed or built. Meanwhile, GIS represent information about the environment ‘as built’ (or as captured) at different points in time, thus yielding less detailed but regularly updated datasets covering wide regions.

<sup>1</sup><http://www.buildingsmart-tech.org/infrastructure>

Due to the overlap in the features that are modelled in both domains, as well as their differing strengths and weaknesses, it is widely acknowledged that the integration of data from both domains is beneficial and a crucial step forward for future 3D city modelling [Kalantari, 2017]. This integration can avoid unnecessary efforts in redundant modelling and allow for new data flows in both directions. In this way, more detailed BIM data can feed more general GIS data, and GIS data can provide the context that is usually missing in BIM data. By pursuing the integration of GIS and BIM data, many ideas become possible:

- with contextual GIS information, BIM methodologies can be better applied to infrastructural works;
- more detailed 3D city models can be built by reusing BIM data;
- smart city concepts can perform integrated reasoning on terrain, buildings and city infrastructure;
- and spatial analyses can support multiple levels of detail and the complete life cycles of objects.

Despite this desire for integration, the disciplines of GIS and BIM are currently disconnected by their modelling paradigms, software tools and open standards, such as CityGML for GIS and IFC for BIM (Section 2.1). Consequently, GIS and BIM datasets differ fundamentally with respect to their semantics, geometry and level of detail, and because of the different modelling approaches of both, there is not one optimal nor uniform conversion between the information models. Even as researchers and practitioners have studied how to best share information between BIM and GIS and how to address all the differences from different perspectives (Section 2.2), it is still very hard (if not impossible) to share 3D information among different users throughout the life cycle of urban and environmental processes, i.e. from plan, design and construction to maintenance—especially when trying to do so using open standards. Moreover, most of the research conducted so far

has focused on the high-level and semantic aspects of GIS-BIM integration (e.g. conceptually mapping equivalent types) [El-Mekawy et al., 2011, 2012; Amirebrahimi et al., 2016; de Laat and van Berlo, 2011], mostly avoiding the more complex tasks involving geometric processing (e.g. conversions, geometric validation and automated tests using said geometries in a manner that is consistent with how geometries are represented in both domains). Without a domain-aware geometry conversion/integration, it is not possible to use the resulting datasets as input in the existing software packages used in the BIM and the GIS communities.

In view of the lack of research about the GIS-BIM integration, in the beginning of 2017, the GeoBIM project was started in the Netherlands with the aim of developing methodologies to process complex BIM and GIS models concurrently, in an automated fashion, and in a manner that is usable in practice. In pursuit of this goal, three different smaller research subprojects were defined:

1. the automated processing of complex architectural IFC models into more generalised GIS models (Section 3);
2. the integration of information about the geological constitution of the subsurface at the design stage of infrastructure projects (Section 4);
3. the georeferencing of IFC models (Section 5).

The GeoBIM project is a collaboration of two research groups on BIM (in TU Eindhoven) and 3D GIS (in TU Delft), the two respective national standardisation bodies (BIM Loket and Geonovum) and several users who have a high interest in closer BIM/GIS integration, i.e. the local waterway and roads authority (*Rijkswaterstaat*), the cadastral land registration authority (*Kadaster*), and the cities of the Hague and Rotterdam.

As the first stage of the GeoBIM project is now finished, there is a series of results, which are summarised in this paper and are presented in full in the reports available in the project website (<https://3d>.

[bk.tudelft.nl/projects/geobim/](https://3d.bk.tudelft.nl/projects/geobim/)). A series of errors have been found that seem to be pervasive in IFC models and which make automated processing of complex architectural models very difficult, and since fully dealing with such errors would require automatic repair algorithms, a complete working interface for GIS and BIM integration could unfortunately not be developed within the timeframe of the GeoBIM project. However, the project has had significant advances on several fronts. First, as the result of further looking into these errors and how to avoid them, a set of recommended guidelines has been written, which should facilitate the automated processing of IFC models (Section 3.3). Second, it has been found that it is technically possible to provide BIM practitioners with subsurface data, and a set of recommendations to make this process more straightforward have also been formulated (Section 4.3). Third, a simple tool to add georeferencing information to IFC files based on a web map has been developed. The paper finishes with some conclusions from this first stage of the project and an outlook for future research in Section 6.

## 2 Background

### 2.1 Open standards in GIS and BIM

Due to a wider availability of information, ease of analysis and for pragmatic reasons, the studies on the exchange of BIM and GIS data often focus on the two most prominent open standards in the two domains: the OGC standard *CityGML* [Open Geospatial Consortium, 2012] for the 3D GIS domain (Section 2.1.1), and the *Industry Foundation Classes* (IFC) [ISO, 2013; Building SMART International, 2013] for the BIM domain (Section 2.1.2). IFC models represent the physical elements of single constructions in great detail, while *CityGML* models represent entire cities in a simpler format that is usable for exchange, dissemination and spatial analyses, such as solar potential and energy consumption estimations. The two modelling paradigms embodied by IFC and

CityGML are representative of BIM and 3D GIS data in general, and they are both widely used in their respective domains.

### 2.1.1 CityGML

CityGML [Open Geospatial Consortium, 2012] is a prominent standard to store and exchange 3D city models with semantics in the GIS domain. It presents a structured way to describe the geometry and semantics of topographic features such as buildings and roads. CityGML as a data format is implemented as an application schema for the Geography Markup Language (GML)<sup>2</sup> [OGC, 2004].

CityGML contains a small number of classes structured into 12 modules, most of which are meant to model different types of objects (e.g. Building, Bridge, WaterBody). These classes differ in the way objects are structured into smaller parts and the attributes that are expected for each. However, CityGML geometries are essentially the same for all classes: objects are represented as surfaces embedded in 3D and consist of triangular and polygonal faces.

CityGML defines five levels of detail (LODs). Figure 1 illustrates the five LODs (LOD<sub>0</sub> to LOD<sub>4</sub>) for the building object:

**LOD<sub>0</sub>** is non-volumetric and is an horizontal footprint and/or roof surface representation for buildings;

**LOD<sub>1</sub>** is a block-shaped model of a building (with an horizontal roof);

**LOD<sub>2</sub>** adds a generalised roof and installations such as balconies;

**LOD<sub>3</sub>** adds, among others, windows, doors, and a full architectural exterior;

**LOD<sub>4</sub>** models the interior of the building, potentially with pieces of furniture (CityGML does not mandate which indoor features need to be modelled, in practice resulting in models with a different granularity [Goetz, 2013; Boeters et al., 2015])<sup>3</sup>.

<sup>2</sup>CityGML uses version 3.1.1 of GML

<sup>3</sup>LOD<sub>4</sub> will be removed in CityGML 3.0. Instead, in-



Figure 1: A building represented in LOD<sub>0</sub> to LOD<sub>4</sub> (image from Biljecki et al. [2016b]).

### 2.1.2 IFC

The Industry Foundation Classes (IFC)<sup>4</sup> standard is an open data model used in the Building Information Modelling (BIM) domain for the exchange of construction models, mainly including 3D models of buildings. It has also been adapted as the ISO 16739 international standard [ISO, 2013]. The geometry definitions are, however, mostly defined or derived from a different standard, ISO 10303 [ISO, 2014], which also specifies the STEP Physical File (SPF) encoding that is most commonly used in IFC files (.ifc).

IFC files can contain many types of classes (130 defined types, 207 enumeration types, 60 select types, 776 entities in IFC 4 Addendum 2). A considerable subset of these pertain to several different representation paradigms which can be combined freely. In practice, most IFC objects are built using sweep volumes, explicit faceted surface models and Constructive Solid Geometry [CSG, El-Mekawy and Östman, 2010]. Elements are modelled in local coordinate systems defined by a hierarchical set of transformations that correspond to the levels in a decomposition structure (typically a site, project, building and individual floors). The representation paradigms include:

- **Primitive instancing:** an object is represented based on a set number of predefined parameters. IFC uses this paradigm to define various forms of 2D profiles (Figure 2), as well as volumetric

door and outdoor features will each be modelled at LOD<sub>0-3</sub>.

<sup>4</sup><http://www.buildingsmart-tech.org/specifications/ifc-releases>

objects such as spheres, cones and pyramids.

- **CSG and Boolean operations:** an object is represented as a tree of Boolean set operations (union, intersection and difference) of volumetric objects (see Requicha [1982] for more details). Half-spaces are often used to cut out the undesired parts of surfaces or volumes.
- **Sweep volumes:** a solid can also be defined by a 2D profile (a circle, a rectangle or an arbitrary polygon with or without holes) and a curve [Wang and Wang, 1986] along which the surface is extruded.
- **B-rep:** an object is represented by its bounding surfaces, either triangulated meshes, polygonal meshes or topological arrangements of free-form surfaces.

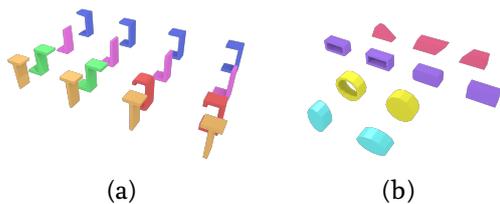


Figure 2: IFC defines various types of parametric curved profiles such as (a) those based on the characters U, L, Z, C and T and (b) those based on trapezia, (rounded) rectangles, circles with/without holes and ellipses. Note the various types of tapered and curved parts of the profiles. These are most commonly used in extrusions such as those shown here.

## 2.2 Previous GIS-BIM integration efforts

The integration of BIM and GIS data is a complex topic that has been tackled in different ways. As Liu et al. [2017] state, the integration can involve semantics and/or geometry, and it can involve the conversion of BIM and GIS data to either a unified model or to the standards of each other (either

bidirectional or only one-way). The Geo-BIM project instead focused on realising the integration in practice and at the *data* level. By contrast, Amirebrahimi et al. [2016] discusses how the integration can also be performed at the *process* and *application* levels.

Most of the related work on this topic is concerned with converting IFC models to a GIS model like CityGML because that implies simplifying and removing details and extraneous information in the data. The inverse operation, from GIS data to BIM, is rarely discussed as it is deemed less useful and it might involve adding more details to the data. Nonetheless, there are methods to create BIM from existing models, which can be considered as GIS to BIM [Song et al., 2017; Volk et al., 2014]. One use case for the conversion from GIS to BIM is to be able to import GIS data about the environment into BIM software so that designers can consider the environment in their design. An example is converting data about the geological subsurface (defined in GIS formats) into IFC [Diakit  and Stoter, 2017].

El-Mekawy et al. [2011] and El-Mekawy et al. [2012] propose to combine information from both domains and create a unified model in which all the semantic properties of IFC and CityGML are present, and they propose using bidirectional mappings for all the semantic classes relevant to IFC and CityGML. Amirebrahimi et al. [2016] extends the data model of GML (and not CityGML) to create a unified model supporting two specific applications (visualisation and flood damage assessment to specific buildings). Kang [2018] introduces a conceptual mapping standard B2GM, which links BIM to GIS based on the ISO 19166 standard.

When the geometry is considered, most existing conversion algorithms from BIM to GIS convert *all* the geometries (from one of the 4 paradigms listed in Section 2.1.2), which yields GIS models having poor usability in practice. Using these geometries in software for simulations or spatial analyses requires a huge amount of manual work [McKenney, 1998]. Little attention

has been paid to a more meaningful conversion which requires not only selecting the appropriate classes, but also performing spatial operations (such as Boolean operations) to select only the features, or part of these, that are appropriate in the other model. One concrete example is how in CityGML the buildings in LOD<sub>2</sub> and LOD<sub>3</sub> should be modelled using only their surfaces that are visible from the outside, while a simple conversion of all the BIM classes (in most cases) will yield volumes for individual wall segments with thickness and thus interior and exterior surfaces. As Benner et al. [2005] remarks, the surfaces forming the exterior of a building are not explicitly marked as such in an IFC file and cannot be deducted directly from the semantics of the objects. Entire building elements can be marked with an *IsExternal* property, but this does not offer the granularity to tag (parts of) individual surfaces, which are often implicitly defined.

A few programs offer the possibility to convert IFC models into CityGML models. Three examples are: the Building Information Modelserver<sup>5</sup>, IfcExplorer<sup>6</sup> and Safe FME<sup>7</sup>. All of them allow the users to convert IFC models to CityGML at different LODs. The users can in some cases choose which IFC objects should be used. However, all the features are converted, without any selection or post-processing to keep only the relevant ones. Different projects use such approach to obtain integrated datasets useful for visualisation-based analysis, e.g. Rafiee et al. [2014]. Observe here that for visualisation and other applications, such as daylight simulation, converting all the geometries is not a major hindrance since internal elements will not be displayed or will not affect the end result. The size of the dataset will however increase and slow down the visualisation process.

Hijazi et al. [2010] built an open-source web-GIS in which IFC objects can be imported after having been converted to b-rep models. Unfortunately, only sweep-volume

objects can be converted and all the geometries are kept (thus resulting in non-manifold building models). De Laat and van Berlo [2011] developed an CityGML ADE (application domain extension) called GeoBIM<sup>8</sup>, so that new semantic classes defined in IFC are added in a CityGML model. However, no geometric manipulation is performed.

There have been different attempts at converting IFC models to CityGML LOD<sub>2/3</sub> models by processing the geometries. Benner et al. [2005] describe the general steps needed to convert an IFC file to an alternative data model closely related to CityGML (the QUASY model). They first map the semantics from IFC to QUASY and select the relevant boundary objects, and then the outer visible surfaces are extracted by selecting a subset of the input objects. For the (equivalent of) the LOD<sub>3</sub> model, they discard geometries inside the building by projecting each floor of a building to horizontal and vertical ‘footprints’ and keeping only those touching the envelope. This technique may yield building models having holes/gaps in the exterior envelope. Moreover, while the output models appear to be LOD<sub>3</sub> models, the walls and the roof are volumetric. Deng et al. [2016] converts IFC to the different LODs of CityGML. To obtain the exterior envelope of each building, they use a ray-tracing algorithm: they define a few points-of-view and determine whether a given surface is visible from them. If so, it is assumed to form part of the exterior boundary. This method will however yield buildings with several holes as several surfaces (e.g. those for a roof overhang, or small ones near a window sill) will not be visible from the finite set of points of view. Donkers et al. [2016] convert IFC models to LOD<sub>3</sub> models by selecting a subset of the objects and then extracting the exterior envelope by using a series of Boolean set operations in 3D. Their algorithm does not yield holes/gaps if the input does not contain any and they can close small gaps by buffering all primitives; this however introduces artefacts in the entire model. In addition, in their implementation the semantics of the

<sup>5</sup><http://bimserver.org>

<sup>6</sup><http://iai-typo3.iai.fzk.de/www-extern/index.php?id=1566&L=1>

<sup>7</sup><http://www.safe.com>

<sup>8</sup>[http://www.citygmlwiki.org/index.php/CityGML\\_GeoBIM\\_ADE](http://www.citygmlwiki.org/index.php/CityGML_GeoBIM_ADE)

objects cannot be stored and thus different tricks are used to recover it for each surface in the final model. Kang and Hong [2018] automates the generation of LOD geometries using Screen-Buffer scanning-based Multiprocessing (SB-MP) using semantic mapping rules.

Recently, the Open Geospatial Consortium carried out research on the use of IFC and CityGML in Urban Planning [Kalanitari, 2017]. This project was done in the context of the Future City Pilot Phase 1. Their main conclusions align to the findings presented in this paper. They firstly conclude that the architecture design processes often do not require georeferencing, and unreferenced architectural models seriously hinder their reuse in GIS environments. Instead, for IFC data to be of use in GIS applications, the geographical coordinate system of IFC file should be set in advance. Secondly, the OGC project identified several inconsistencies in coding IFC elements in practice that made transformation to CityGML complicated. Therefore they conclude that for adopting IFC in urban planning a clear set of specifications needs to be set for the preparation of IFC files. This is in line with the findings presented here. However, this project goes one step further and actually proposes such guidelines in Section 3.3.

### 3 Automatically processing complex architectural models in IFC

The first part of the GeoBIM project was an experiment- and use case-driven scoping study with two aims: (i) to develop a CityGML/IFC interface for reusing GIS data in the BIM domain and vice versa, and (ii) to formulate recommendations for further integration, such as modelling guidelines for bidirectional integration based on the main issues identified and the preferred solutions to these. This part of the GeoBIM project focused its experiments on three IFC files provided by the city of the Hague (Figure 3). These are complex models with

several thousand objects each and which use all the main representation paradigms that are possible in IFC (see Section 2.1.2).

#### 3.1 Initial methodology

Initially, as shown in Figure 4, the initial workflow was based on the use of two well-known and used open-source libraries: CGAL<sup>9</sup> and IfcOpenShell<sup>10</sup>. An attempt was made to parse and process every object in an IFC file independently using a modified version of IfcOpenShell, where its Open CASCADE<sup>11</sup>-based kernel has been substituted for a new one based on CGAL. The aim of this approach was to use spatial analysis algorithms from the GIS domain and the robust Boolean set operations on Nef polyhedra available in CGAL in order to solve various use cases [Bieri and Nef, 1988; Nef, 1978; Hachenberger, 2006]. This approach was chosen since previous experience showed that the Boolean set operations in Open CASCADE are not as robust as those available in CGAL. For instance, compliance with height regulations in 3D zoning maps can be checked with CGAL using Boolean intersections of the model and the polyhedra generated from zoning maps.

The main geometric IFC entity types were thus converted into appropriate CGAL-based representations for points, polygonal curves, polygons with holes, planes, etc. In this process, implicit and parametric curves, surfaces and volumes were interpreted into explicit boundary representations and discretised into polygonal curves and polygonal meshes (Figure 2). It is worth noting that this discretisation necessarily creates a mesh that deviates from the input curve or surface. All placements and transformations in IFC are converted into 3D affine transformations defined by a matrix, which can then be recursively applied to each object as necessary. In this manner, a polyhedral representation of every volumetric object has been obtained and stored as a CGAL Polyhedron\_3.

<sup>9</sup>The Computational Geometry Algorithms Library: <https://www.cgal.org/>

<sup>10</sup><http://ifcopenshell.org>

<sup>11</sup><https://www.opencascade.com>

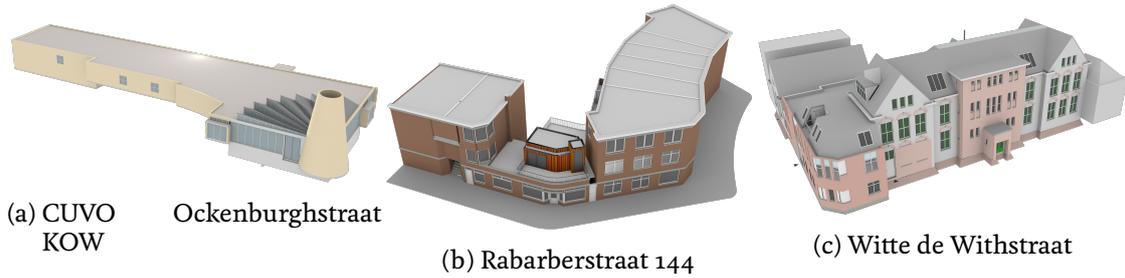


Figure 3: Test IFC files from the City of The Hague

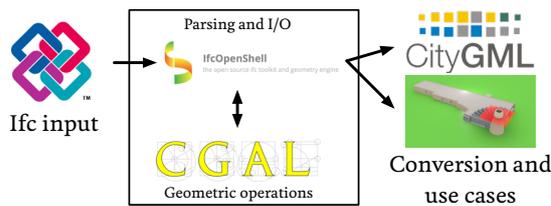


Figure 4: The initial methodology

Some geometric IFC entity types however required additional processing, such as the CSG solid, half-space and Boolean result representations which are obtained by first converting a polyhedron to a CGAL `Nef_polyhedron_3` and then performing Boolean set union, intersection or difference operations.

This methodology yielded good initial results (Figure 5), being able to parse and generate CGAL geometries for a rapidly increasing number of the objects present in the test IFC models.

However, invalid objects are widespread in the IFC models that were received as part of this study—something that is consistent with previous experience with BIM and GIS data and with the experience of other practitioners; see for instance Biljecki et al. [2016a]. Self-intersections and intersections between objects were the most common errors (Figures 6a and 6b), but there were also uneven surfaces that were supposed to be planar and disconnected objects that were modelled as one. One of the most interesting errors found was the presence of several objects that are seemingly valid and form topological 2-manifolds (as checked by making sure that surfaces join in pairs along common edges), but in reality

contain self-intersections (Figure 6c). Note that this is something explicitly disallowed by the IFC standard<sup>12</sup>, but not enforced by most current implementations.

After asking the architects producing these models, it was found that such errors are generated by the software they work with. They are mostly unaware of these errors, or at least these do not significantly hinder their design, building and constructions processes.

Since invalid objects would often cause the implemented processing methods to crash, catching and correcting as many errors as possible was attempted, and so a series of validation tests on every object and to its openings (if any) were added. These are done at appropriate places during the construction of every object, before the conversion of a CGAL `Polyhedron_3` to a `Nef_polyhedron_3`, and before it is triangulated for the generation of a file used for visualisation (the simple OBJ format). Among others, tests were made for:

- combinatorial validity (2-manifoldness),
- surfaces that enclose a space,
- crashes/failures of CGAL's triangulation algorithm (e.g. when a surface self-intersects),
- self-intersections,
- CGAL crashes/failures when converting to a `Nef_polyhedron_3`.

<sup>12</sup><http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/ifcgeometricmodelresource/lexical/ifcfacebasedsurfacemodel.htm>

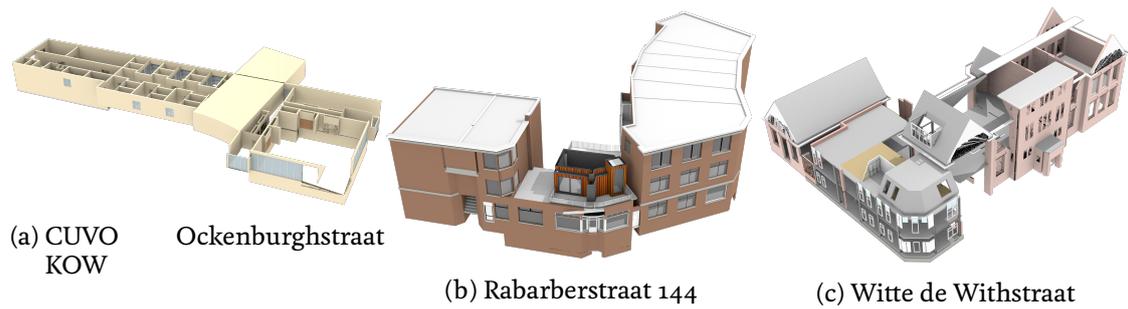
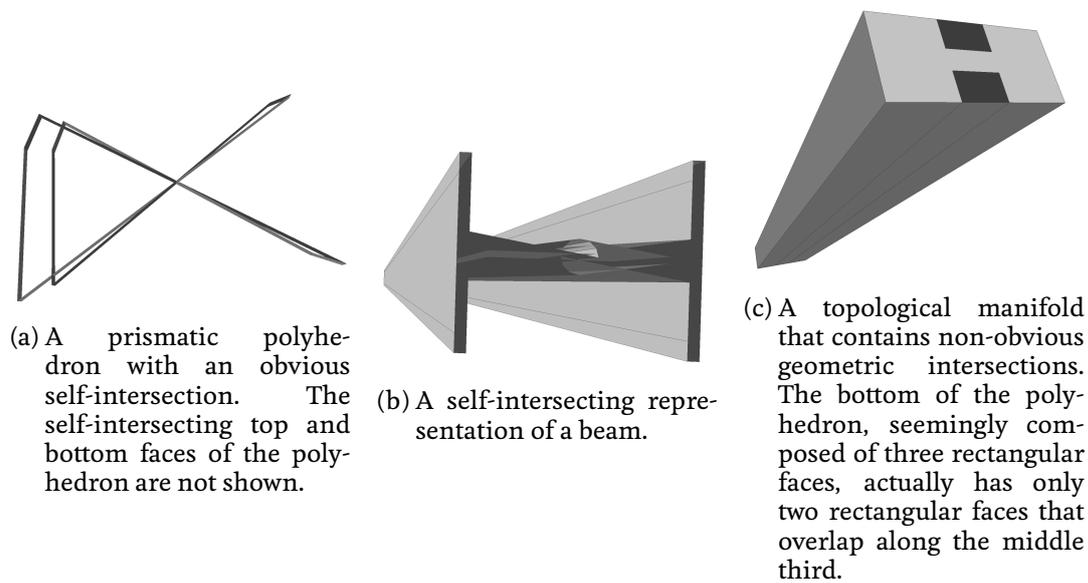


Figure 5: Test IFC files from the City of The Hague after processing with initial methodology



(a) A prismatic polyhedron with an obvious self-intersection. The self-intersecting top and bottom faces of the polyhedron are not shown.

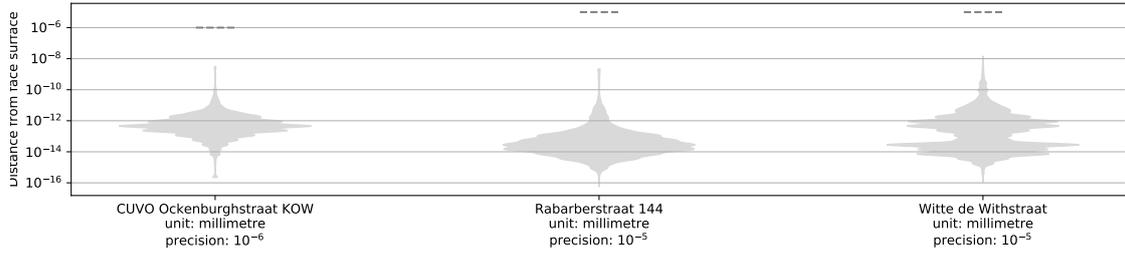
(b) A self-intersecting representation of a beam.

(c) A topological manifold that contains non-obvious geometric intersections. The bottom of the polyhedron, seemingly composed of three rectangular faces, actually has only two rectangular faces that overlap along the middle third.

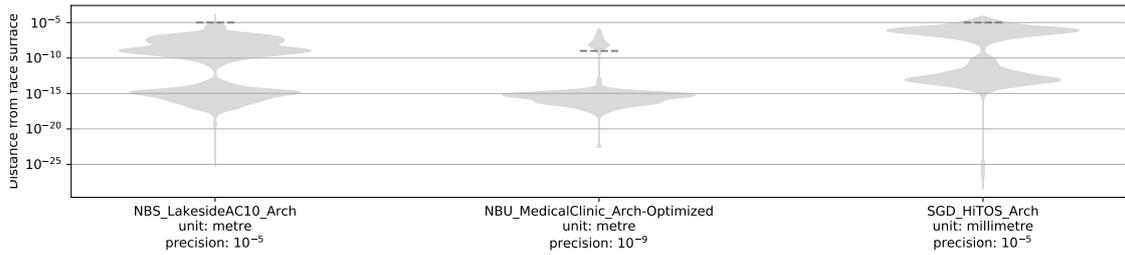
Figure 6: Some of the errors present in the test IFC files

There are two classes of faces in IFC: `IfcFace` and `IfcFaceSurface`. The latter has an associated elementary surface that describes the face geometry, which can be non-planar, while the former is always planar and its surface is given implicitly as the plane through the points of the associated polygonal loops. The polygonal boundary loops are given in three-dimensional coordinates and not in parametric coordinates of the corresponding surface. Hence, there can be issues with vertices not in agreement with the underlying surface. Building geometries are often polyhedral and support for curved geometries is limited in importing and exporting applications. For that reason only the planar class is encountered in the set of models assessed in this paper.

Figure 7 shows an analysis of the deviation of face vertices from the corresponding surface. Due to the imprecision of floating point arithmetic used in BIM authoring tools, some inaccuracy is always expected. For an IFC geometry to be considered valid, this imprecision should be smaller than the geometric modelling precision supplied as part of the `IfcGeometricRepresentationContext`. However, CGAL Nef polyhedra pose stricter requirements on the planarity of faces. Therefore faces of every object are triangulated whenever a Nef polyhedron is created. This ensured that the conversion from a CGAL `Polyhedron_3` to a `Nef_polyhedron_3` is able to compute a plane passing through every face. Another possibility would have been to compute the



(a) IFC files from the City of The Hague



(b) A set of three arbitrary public test models from `data.dunaark.eu`

Figure 7: Distributions of maximal vertex distance from underlying surface. To assess the distance from implicitly defined planes, a surface normal is found using Newell’s algorithm [see e.g. Sutherland et al., 1974, p. 15]. Arbitrary precision floating point arithmetic is used during computation. (a) shows that all distances are within the geometric precision defined in the file. As shown in (b) this is not always the case.

best fitting plane per face and then snap a vertex to the intersection of the planes of its incident faces [Arroyo Ogori, 2016].

As CGAL and the typical prevalent geospatial databases only support polyhedra, an issue identified is the potential for different results in the generation of discrete and explicit b-rep linear geometries (e.g. polygons and polygonal meshes) from the implicit and curved geometries in IFC. Explicit geometries will invariably yield different b-rep representations according to the chosen discretisation method and its parameters. For instance, a choice was made to discretise ellipses into closed polygonal curves using a customisable number of equal-angle intervals, and to discretise spheres into icosahedral approximations with a customisable number of refinements. However, alternative methods could involve discretising ellipses using equal-length line segments and spheres using equal angle rectangular patches.

From a more practical standpoint, it was found that the available features of CGAL are not sufficient to comfortably model all the complex features of IFC—something that is also true for many other geometric processing libraries. For instance, CGAL Nef polyhedra do have support for half-space representations as long as an extended kernel is used, which incorporates polynomial representations of various classes such as planes. However, extended CGAL kernels appear to be incompatible with various parts of the Polygon mesh processing package, which is used for triangle-triangle intersection tests, 2-manifoldness tests, stitching the faces of a polyhedron together, and reversing normals, among other functions. Another example is that Nef polyhedra do not offer an efficient way to construct meshes that do not enclose a volume or to create a `Polyhedron_3` from such a mesh stored in a Nef polyhedron. While workarounds around these CGAL problems were found (and partly implemented), these

involve very complex code, are too slow for practical use or do not cover all the many cases commonly present in IFC files.

Unfortunately, while trying to fix all the errors present in the models, it thus became clear that getting all IFC entities to work correctly using this conversion methodology was going to be a significantly more time-consuming undertaking than the project's allotted time allowed. In addition, the many errors that were found in the IFC models were unexpected, and therefore solving these was outside the scope of the project.

### 3.2 Final methodology

After running against the time constraints for the project, and due to the unsolvable issues identified with the initial methodology, a simplified process that minimised the complex interaction between CGAL and IfcOpenShell was applied (Figure 8). The idea was to sidestep the limitations of CGAL in representing certain geometries and avoid many of the CGAL crashes caused by the interplay between invalid geometries in the IFC files and the stricter geometry requirements of CGAL compared to Open CASCADE.

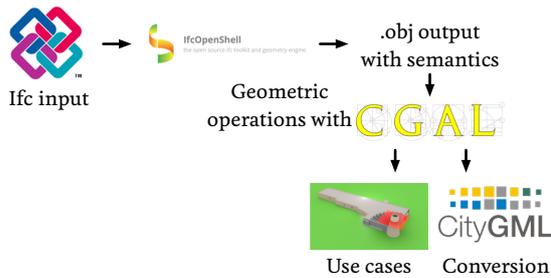


Figure 8: The final methodology

In this final methodology, a script was written that used the standard version of IfcOpenShell (with its Open CASCADE kernel) to export a series of Wavefront .obj files containing all relevant objects present in the IFC files. For this, the subtypes of IfcBuildingElement were analysed, which according to the IFC standard define a ‘major functional part of a building’, those

being: IfcBeam, IfcBuildingElementComponent, IfcBuildingElementProxy, IfcChimney, IfcColumn, IfcCovering, IfcCurtainWall, IfcDoor, IfcFooting, IfcMember, IfcPile, IfcPlate, IfcRailing, IfcRamp, IfcRampFlight, IfcRoof, IfcShadingDevice, IfcSlab, IfcStair, IfcStairFlight, IfcWall, IfcWindow. Thus, all geometries in every IFC file were extracted to an easily-parsable format while preserving the most relevant semantic information present in the original input file. This allowed us to overcome a shortcoming of the methodology of Donkers et al. [2016].

Afterwards, a different program that parsed the individual .obj files and created a CGAL Nef polyhedron for every object was created. The parsed objects were processed using the same defensive programming methods and validation tests described for the initial methodology. The obtained Nef polyhedra were then selectively processed together using Boolean set operations so as to obtain a single Nef polyhedron containing all of the relevant parts of a building for a given spatial analysis.

The simplified processing pipeline in the final methodology resulted in a more workable methodology for the time frame of the project. All objects in the three test IFC datasets were able to be exported correctly to .obj and appeared to be visually correct. Moreover, many of the problems present in the IFC files could indeed be sidestepped by using the Open CASCADE kernel of IfcOpenShell, resulting in a significantly smaller number of problematic objects (e.g. such as a very complex staircase with many self-intersections), which could in general be disregarded.

By using Boolean set operations, it was also possible to correctly handle overlapping objects, and by combining these with Minkowski sums, reasonable solutions to many of the remaining problems in the IFC files were found. For instance, non-manifolds could be converted into manifolds by finding the non-manifold edges and vertices in an object, applying a Minkowski sum of these with a small volumetric kernel, and subtracting the result from the original objects using a Boolean

set difference. Similarly, small gaps between objects that would cause the interiors of a space (e.g. a room) not to be enclosed could be fixed by applying a Minkowski sum with a small kernel of a width equivalent to a snapping threshold between adjoining volumes. As another example, some CGAL crashes could be avoided by applying an iterative process where program crashes are caught, and an object that otherwise causes a crash is processed by applying a Minkowski sum with a small kernel until the object can be handled properly. These kinds of automated shape healing is what constitutes the advantage of using CGAL as an arbitrary-precision and robust geometry engine over the floating-point accuracy precision of Open CASCADE.

Unfortunately, even as significantly more objects could be processed using the final methodology, it was still not possible to fix all outstanding issues in the time allotted for the project, and thus some objects in the IFC files were necessarily discarded. This is a problem, since the automated tests that were considered (e.g. zoning and shadow casting) cannot yield authoritative results if the models have missing objects, i.e. a lack of errors does not guarantee that a model passes a test with absolute certainty. Similarly, meaningful conversions from IFC to CityGML cannot be performed when spaces that should be well enclosed are not, due to missing objects at their boundaries.

### 3.3 A proposed set of IFC modelling guidelines

The development process of the two methodologies pointed to a series of issues that are common in IFC files, including in the files that were received for this project. After a further look into these issues, a set of guidelines to avoid them were produced. If properly followed, they would greatly simplify the automated processing of IFC files.

However, it is first important to note that many rules and recommendations regarding the proper use of IFC are already

given in the IFC standard<sup>13</sup>, implementation guidance<sup>14</sup>, implementers guide for IFC2x3<sup>15</sup>, implementer agreements<sup>16</sup>, and external guidelines such as BIM Basic IDM<sup>17</sup>. These range from fundamental aspects, such as how each IFC entity is defined and the possible values for each attribute, to common-sense practical rules, such as schemes for the consistent naming of objects.

While the above mentioned sources already discuss plenty of guidelines on the proper use of IFC, during the course of this project and based on previous experiences, it was found that these sources of rules and recommendations are not always followed in practice, either because the authoring tool does not enforce or enable such rules, or because its users are unaware of certain issue. Moreover, they do not cover certain desirable aspects, resulting in a series of issues that make processing IFC models less than ideal. Therefore, a series of recommendations that aim to improve their reusability in GIS applications were proposed:

**Georeferencing** IFC files should contain their precise real-world location using the latitude, longitude and altitude values in the `IfcSite` taking into account the offset given by the `WorldCoordinateSystem` of the `IfcGeometricRepresentationContext` for the 3D model and the 2D plan (if used). Due to practical difficulties, it cannot be expected that these values match the reality with the `Precision` given in the `IfcGeometricRepresentationContext`, but the values should be easy to set to within a few meters of the real location. In addition, if the `y-axis` of the `WorldCoordinateSystem` in the `IfcGeometricRepresentationContext` does not match the

<sup>13</sup><http://www.buildingsmart-tech.org/ifc/IFC4/final/html/>

<sup>14</sup><http://www.buildingsmart-tech.org/implementation/ifc4-implementation>

<sup>15</sup><http://www.buildingsmart-tech.org/downloads/accompanying-documents/guidelines/IFC2xModelImplementationGuideV2-0b.pdf>

<sup>16</sup><http://www.buildingsmart-tech.org/implementation/ifc-implementation/ifc-impl-agreements>

<sup>17</sup><http://bimloket.nl/BIMbasisILS>

true North direction, the *TrueNorth* attribute should be set as well. See Section 5 for more details on how these values are stored.

**Valid volumetric objects** Use volumetric objects as much as possible (e.g. sweeps and geometric primitives), test that they conform to their entity definition and take care to make sure these are watertight when using boundary representations.

**No intersections** There should be no self-intersections or intersections between objects, much like the BIM Basic IDM disallows intersections between objects by stating: ‘There are no duplicates or intersections permitted. Make sure this is checked in IFC.’ Among others, the Solibri Model Checker<sup>18</sup> and the simplebim Space Boundary add-on<sup>19</sup> are able to detect overlapping objects.

#### **Forming enclosed spaces that are also modelled as IfcSpaces**

The necessary objects so as to form enclosed spaces should be modelled, and these objects should fit properly with each other and without gaps between them. In addition, enclosed spaces should also be modelled explicitly and with their precise geometry as IfcSpaces. Simpler applications that make straightforward use of these IfcSpaces can then avoid more complex geometry processing.

**Using specific entities** Always use the most specific entity type possible, avoiding generic types such as *IfcBuildingElementProxy*, use these specific entity types consistently across all objects of a model, and ensure that all related geometries of an object are marked as such. Additionally, when conversions to CityGML are concerned, it is worth focusing on entities that have a direct mapping to CityGML classes, such as *IfcSlab*.

<sup>18</sup><https://www.solibri.com/faq/using-the-space-validation-rule-to-ensure-model-accuracy/>

<sup>19</sup><http://datacubist.com/support/addon-spaceboundary.html>

## 4 Integrating subsurface information with IFC

Data about the subsurface is critical for any project involving shallow or deep digging of the ground, such as the construction of buildings and infrastructure. By taking into account such information at the design stage, the risks of accidents can be better handled and costs can be reduced significantly because safety and risk analysis can be based on data about the subsurface constitution.

The Netherlands possesses a wide range of datasets about the subsurface that could be used for this purpose, and most of them are open data. In particular, the Netherlands Organisation for Applied Scientific Research (TNO), through the Geological Survey of the Netherlands (GDN), has made available the largest databank of the ground beneath the Netherlands. This databank, known by the acronym DINO<sup>20</sup>, is the single clearinghouse for data on the shallow and deep geology of the Netherlands. It comprises borehole data, groundwater data, cone penetration test data, vertical electrical soundings, the results of geological, chemical and mechanical sample analyses, borehole logs and seismic data.

However, despite the availability of subsurface data in the Netherlands as open data for years, the data is rarely used in the design and planning phase of major infrastructure projects. This might be due to the fact that the teams involved in those projects are not aware of the availability of such data, or could also be explained by incompatibilities between the formats in which the data is encoded and delivered by TNO and the tools that are used by designers of infrastructure projects.

### 4.1 Aim and scope of this subproject

A Dutch registry of subsurface information known as the BRO (*Basis Registratie Onder-*

<sup>20</sup><https://www.dinoloket.nl>

grond)<sup>21</sup> is currently being developed. The aim of the BRO is that subsurface information can easily be used in applications that can benefit from this data. To provide input for optimal development of the BRO, the project presented in this report examined how information about the geological subsurface can be made more accessible to designers and constructors, such that the BRO can be developed accordingly.

The subsurface data that is investigated here is GeoTOP (Figure 9), a countrywide dataset on the geological composition of the shallow subsurface (maximum 50 meters below the ground surface), which is disseminated under the NetCDF (Network Common Data Form) format and available online. GeoTOP divides the subsurface in millions of voxels of  $100 \times 100 \times 0.5$  m. Every voxel has properties assigned to it which characterize the physical and chemical properties of the soil, such as the lithographical unity and class.

On the design side, the focus was on BIM projects, as well as the tools (software) used to design these projects. The aim in particular was to develop an interface between GeoTOP and construction/design applications, so that the latter can easily check and eventually integrate such data to their workflow. The focus is put as much as possible on solutions based on open standards to maximize the interoperability possibilities. However, tailored commercial solutions are also acceptable if they improve the usability of the data, but these should be developed upon open standards.

In addition to a prototype of the interface, this subproject aimed to also understand how the surface data can be best accessed in the design phase and what next steps are needed to achieve this. Some recommendations were formulated in the form of action points that can be addressed in a follow up project.

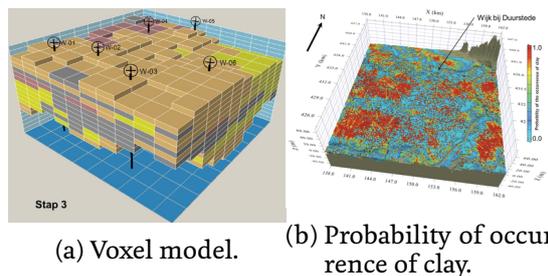


Figure 9: GeoTOP data is encoded based on a voxel model with a resolution of  $100 \times 100 \times 0.5$  m. Each voxel in the model contains information on the lithostratigraphy and lithological classes, including the probability of occurrence for each lithological class (the values are based on interpolations of data from boreholes).

## 4.2 The solution

In order to make open subsurface data such as GeoTOP readily available in designing tools, the project opted for the development of an interface able to translate the GeoTOP data into IFC data in order to make it understandable for the BIM software products. Figure 10 illustrates the workflow of this approach.

Since subsurface information is not the focus of IFC, there is no class directly suitable for the representation of such data. It is however possible to bypass such a limitation by relying on a class that is meant to represent generic spaces known as the *IfcSpace* class and which can easily represent the voxel structure of GeoTOP.

Open source tools were relied upon to build a solution, which is implemented in C++. For parsing NetCDF data, the NetCDF C++ API<sup>22</sup> was used. To deal with IFC reading and writing, the *IfcOpenShell* library was relied upon. Finally, a simple viewer to visualize the extracted voxels was developed, which relies on the *libQGLViewer* library<sup>23</sup>.

<sup>21</sup><https://www.digitaleoverheid.nl/voorzieningen/gegevens/inhoud-basisregistraties/bro/>

<sup>22</sup><https://www.unidata.ucar.edu/software/netcdf/netcdf-4/newdocs/netcdf-cxx.html>

<sup>23</sup><http://libqglviewer.com>

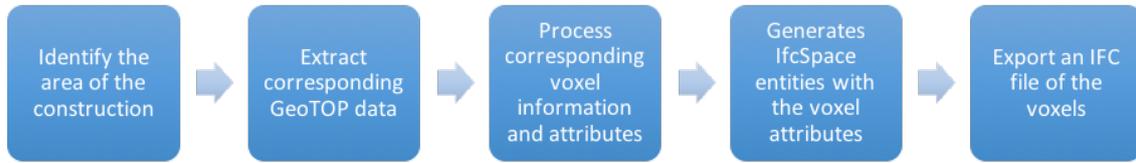


Figure 10: Workflow of this solution

Based on discussion with the GeoTOP expert from TNO, it was decided that the solution would store only the most important values from GeoTOP into the Description attribute of IfcSpace. These are the GeoTOP *strat* and *lithok* classes (Figure 11).

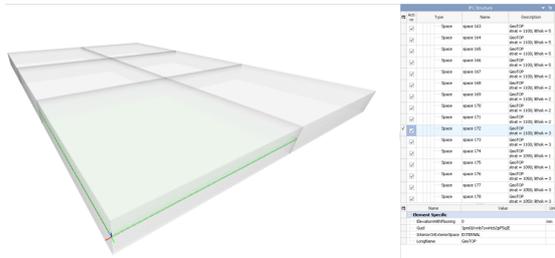


Figure 11: Produced voxels from the GeoTOP underneath the Faculty of Architecture of the TU Delft (visualised on BIM Vision software).

Although the implementation of the proposed solution could be achieved properly, several issues tend to make the integration between BIM and GIS data rather difficult. Among these, the most relevant was the problem of georeferencing the data from BIM or localising the GeoTOP inside a BIM platform (Figure 12), which needed to be done manually.

### 4.3 Recommendations

From the experiments, it is possible to conclude that technically it is feasible to provide BIM practitioners with subsurface data. However, since using 3D GIS data in BIM software and vice versa is currently difficult, this is a topic that requires further attention. In order to guide this in the right

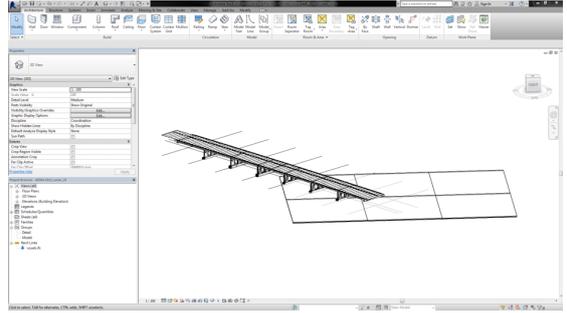


Figure 12: Attempting to put a BIM model of a bridge (provided by Volker-Infra for testing) and the generated GeoTOP voxel model.

direction, the following recommendations were formulated:

- Further investigation of practices of the BIM practitioners is needed regarding subsurface issues. To obtain initial understanding of how they use (or would like to use) subsurface data, a small questionnaire was set up, which was distributed as part of the BIM-loket newsletter. The responses show that BIM practitioners are interested in subsurface data, but currently have difficulties using it. Also, it must be said that it was very hard to get the BIM community involved. In this project, an attempt was made to organise a session for BIM vendors, but there was little to no interest.
- It is necessary to develop more tools that are able to handle other types of subsurface data with more/different coverage.
- Discussions should be started between BIM and GIS practitioners for better handling of georeferencing/alignment

practices and the establishment of guidelines to support this.

- It is desirable to investigate possible approaches to allow the automatic integration/alignments of the models in either BIM or GIS systems.
- Finally, the issues of integrating rough and fuzzy subsurface data with detailed objects represented in BIM that have a much higher accuracy should also be investigated. Most BIM sites fall within a complete voxel or a few voxels only, so it is worth considering what is the value of conclusions drawn from this data integration which may have severe impact on a risk analysis? This requires the close involvement of geological data experts.

## 5 Georeferencing IFC models

Properly georeferencing an IFC file makes it possible to link the (local) coordinates inside an IFC model with their corresponding real-world coordinates, and thus to place the model of a single building or construction within the virtual environment. From discussions with other researchers and from the results of the previously presented two subprojects, it is clear that this is a major issue in practice. Problems with the georeferencing of a file might be due to the accuracy of GIS base map, georeferencing methods, georeferencing accuracy, accuracy of the IFC file, and the error propagation process.

In order to georeference an IFC file, it is possible to use the IFC entity *IfcSite*, which defines an area where construction works are undertaken, and optionally allows for storage of the real-world location of a project using the *RefLatitude*, *RefLongitude* and *RefElevation* attributes. The latitude and longitude are defined as angles with degrees, minutes, seconds, and optionally millionths of seconds with respect to the world geodetic system WGS84 (EPSG:4326). Positive values represent locations north of the equator, west of the geodetic zero meridian (nominally the Greenwich prime meridian)

in IFC2x3, or east of the zero meridian IFC4. Note that this difference in the two IFC versions is risks confusing users and implementers in practice. Negative values represent locations south of the equator, east of the zero meridian in IFC2x3, or west of the zero meridian in IFC4. These angles are expressed according to the type *IfcCompoundPlaneAngleMeasure* and all components (i.e. degrees, minutes, seconds and millionth-seconds of arc) should have the same sign. According to the IFC standard, the geographic reference given might be the exact location of the origin of the local placement of the *IfcSite* or it might be an approximate position for informational purposes only. The elevation is defined according to the datum elevation relative to sea level.

In addition, *IfcSite* contains a few other attributes that allow for an approximation of the real-world location to be given. The *LandTitleNumber* can store the designation of the site within a regional system (e.g. a cadastral record ID), and the *SiteAddress* can store the postal address of the site.

The IFC entity *IfcGeometricRepresentationContext* is used to define the coordinate space of an IFC model in 3D and optionally the 2D plan of such a model. This entity can be used to offset the project coordinate system from the global point of origin using the *WorldCoordinateSystem* attribute, it defines the *Precision* under which two given points are still assumed to be identical, and it defines the direction of the *TrueNorth* relative to the underlying coordinate system. The latter attribute defaults to the positive direction of the y-axis of the *WorldCoordinateSystem*.

In theory, the use of the latitude, longitude and altitude values in the *IfcSite* with an optional offset and true North direction given by the *IfcGeometricRepresentationContext*, should make it possible to precisely georeference an IFC model. In fact, it seems that most IFC files do fill in the requisite values in the *IfcSite*. However, these values are almost always set to zero, to a default or wrong location, or to a very rough approximation of the real location (e.g. a point in the same city). This

is unfortunately compounded by the mismatched definitions of the positive direction for the longitude in IFC2x3 and IFC4.

## 5.1 Adding georeferencing information in Revit 2018

The georeferencing information that is found in an IFC file depends on the BIM software from which it was exported. Since for many users this means Autodesk Revit, this section briefly describes how it can be set in this software, where there is a set of relevant functions made available for designers to incorporate georeferencing information in their models within the *Manage* tab (Figure 13a).

While the *Coordinates* and *Position* buttons are related to the local coordinate reference system of Revit, which only indirectly affect the georeferencing of the model, the *Location* button allows the user to directly link it to real-world coordinates. The location can be defined in three different ways: (i) using the *Internet Mapping Service* option and clicking in a map window or using the address of the project (Figure 13b), (ii) by choosing a city for a rough location (Figure 13c), or (iii) by manually specifying a latitude and longitude.

By default, the project's location is set to a place in Boston, USA, with a latitude value of  $42.3584^\circ$  and a longitude of  $-71.0597^\circ$ . For this reason, any IFC model exported from Revit and in which no project location has been specified will be georeferenced at that location. Among the test files, this is the case for the Witte de Withstraat model, which is supposed to be located in the Hague, but which contains the default location coordinates of Revit (Figure 14).

Next, Revit offers the possibility to set the *TrueNorth* properly in the *Floor Plan* view using 2D coordinates (even if the IFC standard supports 3D coordinates as well). For this, it is necessary to set the *Project Base Point* and the *Survey Point*, which are toggled by enabling their visibility under the *Properties* panel (usually on the left side of the Revit window) with the *Visibility/Graphics Overrides* option (Figure 15a). Here, the project

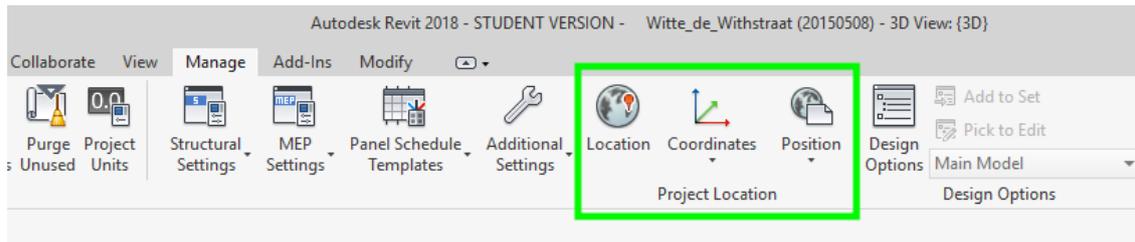
base point (Figure 15b) represents the reference point of all the geometry of the model, i.e. the origin of the project at  $(0, 0, 0)$ , and it is therefore the reference point of the *IfcSite* as well. Meanwhile, the survey point (Figure 15c) stores the geodetic data related to the project, which is assumed to be imported from other sources provided by the surveyors of the project (e.g. some shared CAD files of the construction site). When such source files are missing, all the values are set to 0 by default, including the angle to true North, thus assuming that it coincides with the y-axis of the project.

## 5.2 Example: correction of the georeferencing of the Witte de Withstraat model using Revit

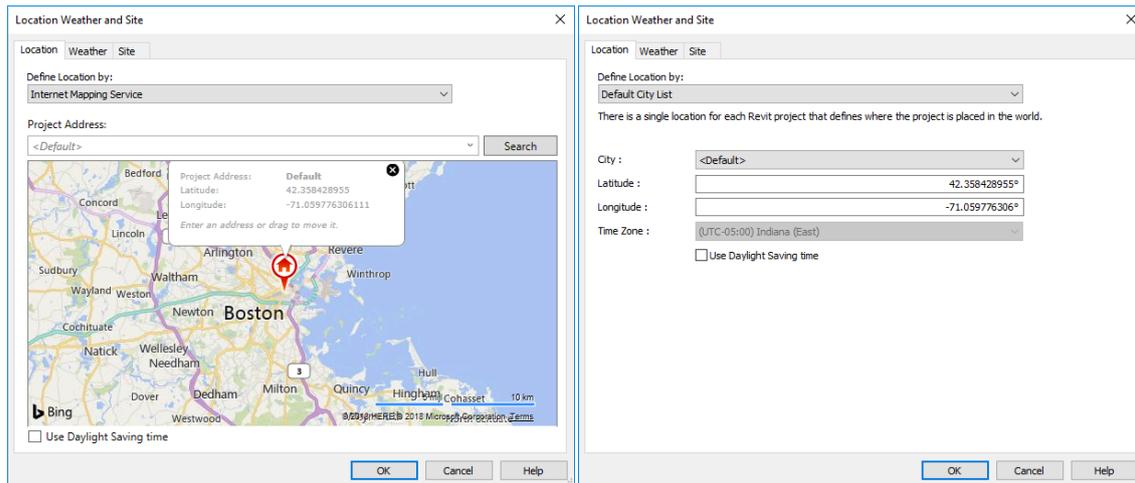
As discussed previously, the Witte de Withstraat model was received with the default georeferencing from Revit, i.e. located in Boston and with its true North set to the y-axis. Firstly, the latitude and longitude can be specified quite easily with the *Location* tool using the *Internet Mapping Service* option (Figure 16a). Note that the new latitude and longitude are inserted in the *IfcSite* attributes (Figure 16b), but the elevation is kept at 0.0 (which is not really a problem for this model in the Netherlands).

Afterwards, correcting the model's true North requires the angle difference between the project North and the true North. As it is possible to import images in Revit, a simple screenshot of the satellite image of the real location was used, with the building properly oriented with respect to the true North (Figure 17).

By roughly finding on the image the point that corresponds to the project base point as well as the survey point in the case of the test model, it is possible to visually recognize the corresponding line of the project North. Then it is necessary to determine the angle between that line and the geographic North one. This can be done by drawing two lines corresponding to those directions (Figure 18a). The measuring tools of Revit then allow the user to compute the angle, which is  $127^\circ$  and corresponds to the angle



(a) Available tools



(b) Map-based

(c) City-based

Figure 13: Tools to select a project's location in Revit 2018.

```
#123593= IFLOCALPLACEMENT ($, #123592);
#123594= IFCSITE ('1nS$S07rASAMoJLctMJUt', #41, 'Surface:243539', $, ', ', #123593, #123590, $, .ELEMENT, ((42,21,30,344238), (-71,-3,-35,-194702)), 0., $, $);
#123599= IFCPROPERTYSINGLEVALUE ('Comments', $, IFCTEXT ('90_terrain'), $);
```

Figure 14: Wrong georeferencing of the Witte de Withstraat model.

from the green line to the orange line in the counter-clockwise direction. The correct value of the angle to the true North from the project North is thus equal to  $360 - 127 = 233^\circ$ , which is then entered in the *Angle to True North* attribute of the *Project Base Point* (Figure 18b).

Once this parameter is set, one can switch from the project North to the true North, and vice versa in the Revit interface, as shown in Figures 19a and 19b. Furthermore, all the changes are reflected in a new IFC file when exported. The newly generated IFC classes integrate the correct information as shown in Figure 19c and the model can be directly oriented after being loaded. In Revit, it is then possible to see that the new Survey Point gives the right direction (Figure 19d). However, it is worth noting that this process is not trivial for

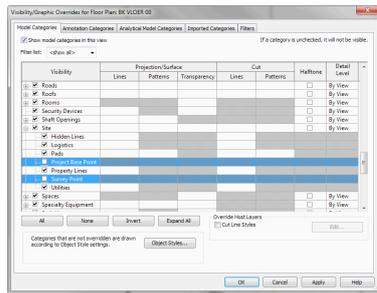
many designers.

### 5.3 IfcLocator: an open-source web service for georeferencing IFC models

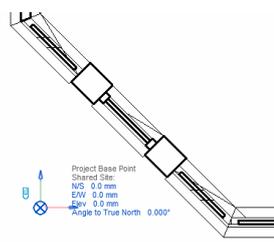
The above mentioned process requires the use of Revit, and thus limits who can correct the georeferencing in an IFC file. For the project, **IfcLocator**<sup>24</sup> was developed, a supporting tool that allows the user to view and check the location of the project stored in an IFC file and to correct that information when necessary. The tool is based on Cesium<sup>25</sup>, which is an open source Javascript

<sup>24</sup><https://github.com/tudelft3d/IfcLocator>

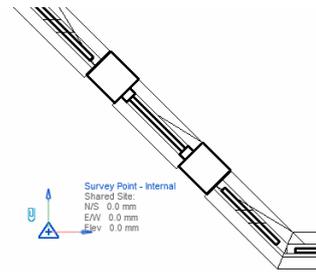
<sup>25</sup><https://cesium.com/>



(a) Making the base and survey points visible in Revit.

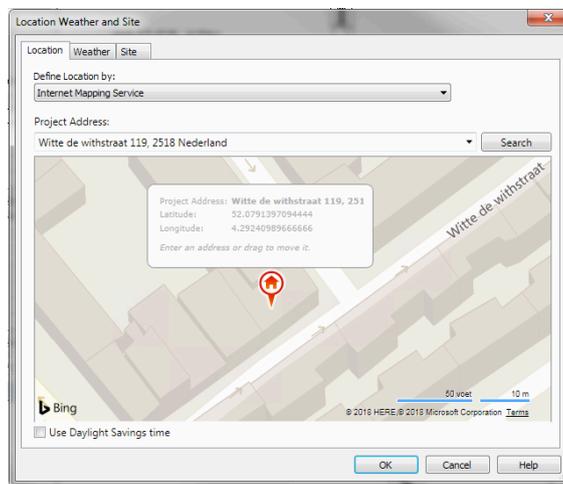


(b) Project Base Point



(c) Survey Point

Figure 15: Setting the Project Base Point and Survey Point of the *Witte\_de\_Withstraat* model.



(a) Setting in correct location in the Internet Mapping Service tool.

```
#1432827= IFCLOCALPLACEMENT($,#1432826);
#1432828= IFCSITE('In$$$07rASAMoJLcMJUt',#41,'Surface:317636',$,'',#1432827,#1432822,$,.ELEMENT.,(52,4,44,902954),(4,17,32,675628),0.,,$);
#1432833= IFCPROPERTYSSINGLEVALUE('Material',$,IFCLABEL('90_zand'),$);
```

(b) The new coordinates in the IFC file.

Figure 16: Correcting the location of the *Witte de Withstraat* model.

library for 3D mapping, including an intuitive 3D viewer of the globe. The choice of a web-based Javascript open source code was motivated by the high flexibility offered by such approach, which is cross-platform. Furthermore, the users can perform all the operation locally, which means they do not have to send their model to a remote server.

## 6 Conclusions

This paper presents a research project that aims at improving the reuse of BIM data in geo applications and vice versa. Based on the results of the project, it can be concluded that a full integration of GIS and BIM data is far from straightforward given current practice. This is partly because existing BIM models contain many geometrical and topological errors which need to be properly handled and often fixed before a conversion or automated processing can be successful. It should be observed that these may not be problematic when used in a BIM

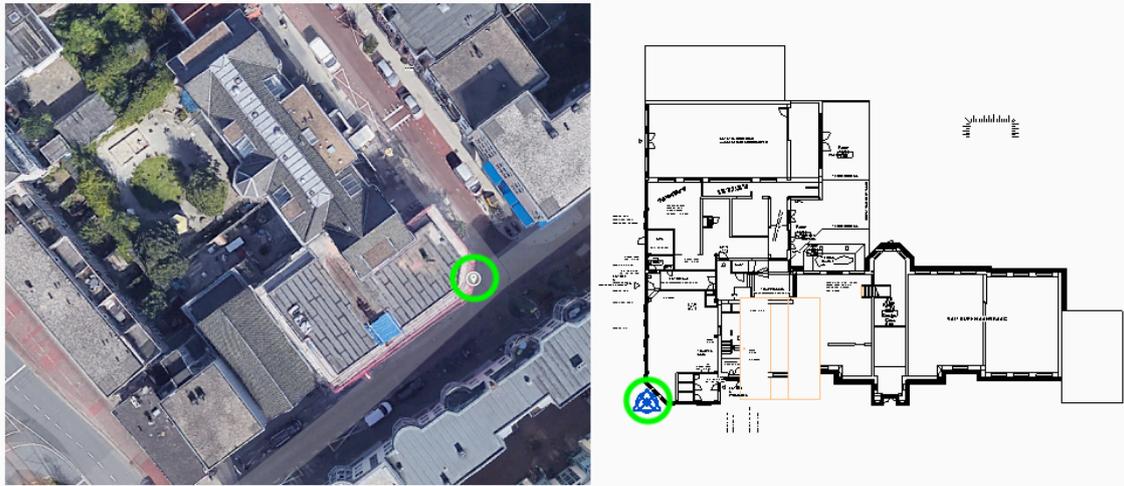


Figure 17: A screenshot from Google Maps is used to identify the orientation of the building with respect to the true North. The green circles indicate the project base and survey points.

environment because of a few reasons:

1. many more geometry types are usually natively supported in BIM software than in GIS software;
2. geometric errors (e.g. self-intersections) are often found in places that are inconspicuous or invisible from the outside;
3. data is often only used for visualisation purposes, which does not require geometric and topological correctness;
4. or the errors only show after the implicit and parametrised types have been converted into explicit geometries (which is during the conversion to IFC).

However, these errors are very problematic in many applications that involve the automated processing of IFC data, such as in GIS operations that perform spatial analyses, which often involve complex operations such as Boolean set operations. Therefore, functionalities to validate the geometry of a BIM model are required; see Ledoux [2013] for an equivalent for 3D GIS datasets.

Another conclusion that can be drawn is that it is unrealistic to develop a robust process for all IFC geometry entity types, particularly so for independent developers and

smaller vendors. In the experiments, a solution for only a subset of the IFC standard was developed. It would take years to extend these efforts to cover all entities, while in practice many of them are rarely used. With IFC, there are many different ways to model an object, which hinders the adoption of the standard and standardisation in the community. Ideally, standards that are intended for exchange should define only one way to model something—this both fosters data exchange and makes it easy to create compliant implementations of the standard. Therefore, it is recommended to formulate and agree on guidelines to standardise the way a specific situation should be modelled in IFC.

Data exchange and its accompanying problems has a much longer history in the GIS domain than in the BIM domain. As also concluded by Liu et al. [2017] from Cerovsek [2011], the relatively younger concept of BIM and its standard IFC have not satisfied the requirements of standards yet: competitiveness, conformity, and connectivity. BIM has to further develop in this respect, including paying attention to formal definitions for correct geometries, adhering to these and developing functionality to validate models.

While many academic research has demon-

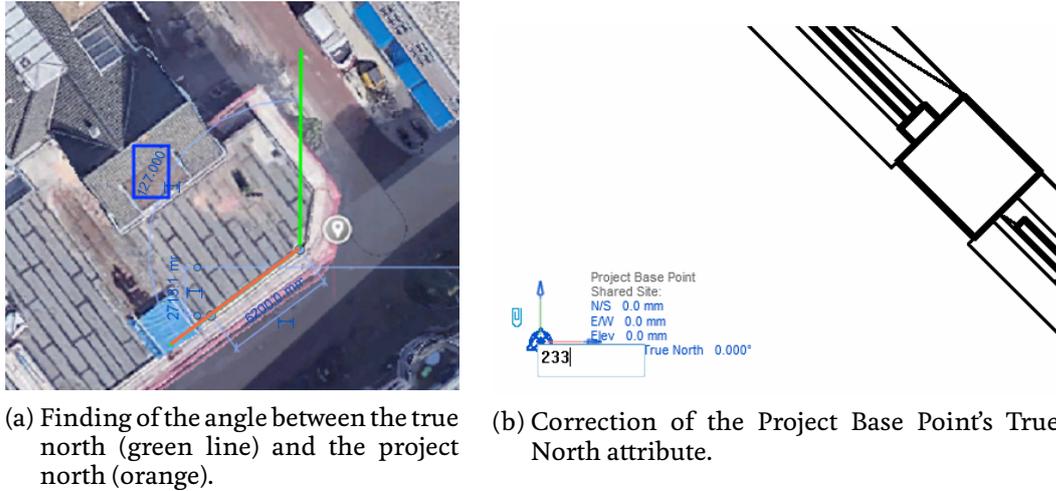


Figure 18: Detection and correction of an IFC file's *TrueNorth*

strated successful IFC-CityGML transformations, they have mostly developed their own solution, based on the use case or data at hand. A transformation from IFC to CityGML that works in practice requires a standardised transformation as there are currently many different interpretations of how to best transform an IFC file into CityGML. From a geometric perspective, this diversity in transformation is mainly due to the many classes of IFC that need to be converted into the relatively few classes of CityGML. The processing needed to make explicit geometries from implicit geometries and to change geometry classes (like conversion of volumes in IFC into surfaces in CityGML for walls) can result in different outcomes in different implementations. For a sustainable information chain that supports the life cycle of objects, this is unwanted, and therefore there is a need to define one uniform and standardised transformation.

From the experiments, it is also possible to conclude that additional information for building modellers is needed in order to support a better automated processing pipeline for IFC geometries in a GIS environment. The guidelines should be a good first step in this direction, which should help to create CityGML/GIS-ready IFC datasets. However, additional efforts are needed since the export process from other BIM formats to IFC is rather opaque

and the implementation of the guidelines is not straightforward.

More understanding is also needed in order to find out how BIM data are used in GIS applications and vice versa. For example, how are IFC designs checked against the existing physical world and against a 3D zoning plan (both GIS data). Is this done in BIM software?, and thus GIS data needs to be imported and localised in BIM software; or is this done in GIS software?, and thus BIM design data needs to be imported in GIS software and properly georeferenced.

BIM data is usually much more detailed than what is expected from GIS data. It is thus very unlikely that all the details of a BIM dataset will be integrated into GIS data. Instead, a generalised version of the BIM model (with relevant attributes for the GIS world) will be converted into a GIS model. A 3D city model may serve as a connection between the two, with unique identifiers and update mechanisms in order to keep the separated BIM models consistent with their generalised counterparts.

## Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant



(a) Superposed floor plan and image oriented with respect to the true North



(b) Superposed floor plan and image oriented with respect to the project North

```

Before correction
#65= IFCAXIS2PLACEMENT3D(#6,$,$);
#66= IFCDIRECTION((2.,6.12303176911189E-17,1.));
#68= IFCGEOMETRICREPRESENTATIONCONTEXT($,'Model',3,1.00000000000000E-5,#65,#66);

After correction
#80= IFCAXIS2PLACEMENT3D(#6,$,$);
#81= IFCDIRECTION((0.79863551004729,-0.601815023152052));
#83= IFCGEOMETRICREPRESENTATIONCONTEXT($,'Model',3,0.01,#80,#81);

```

(c) New **IfcDirection** of the true north.



(d) New *Survey Point* of the project

Figure 19

agreement No 677312 UMnD). The subproject on the automated processing of complex architectural IFC models received funding from a collaboration of Rijkswaterstaat, the cities of The Hague and Rotterdam, Kadaster, Geonovum, and BIM-loket. The subproject on the integration of information about the geological constitution of the subsurface at the design stage was partly funded by the former Ministry of Infrastructure and Environment of the Netherlands.

## References

Sam Amirebrahimi, Abbas Rajabifard, Priyan Mendis, and Tuan Ngo. A BIM-GIS integration method in support of the assessment and 3D visualisation of flood

damage to a building. 61(2):317-350, April 2016.

Ken Arroyo Ohori. *Higher-dimensional modelling of geographic information*. PhD thesis, Delft University of Technology, apr 2016.

J Benner, A. Geiger, and K. Leinemann. Flexible generation of semantic 3D building models. In G. Gröger and T. H. Kolbe, editors, *Proceedings 1st International Workshop on Next Generation 3D City Models*, pages 17-22, Bonn, Germany, 2005.

H. Bieri and W. Nef. Elementary set operations with  $d$ -dimensional polyhedra. In Hartmut Noltemeier, editor, *Computational Geometry and its Applications*, volume 333 of *Lecture Notes in Computer Science*, pages 97-112. Springer Berlin Heidelberg, 1988.

- Filip Biljecki, Hugo Ledoux, Xin Du, Jantien Stoter, K. H. Soon, and V. H. S. Khoo. The most common geometric and semantic errors in CityGML datasets. volume IV-2/W1 of *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 13–22, Athens, Greece, 2016a. doi: <http://dx.doi.org/10.5194/isprs-annals-IV-2-W1-13-2016>.
- Filip Biljecki, Hugo Ledoux, and Jantien Stoter. An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59:25–37, 2016b.
- Roeland Boeters, Ken Arroyo Ogori, Filip Biljecki, and Sisi Zlatanova. Automatically enhancing CityGML LOD2 models with a corresponding indoor geometry. *International Journal of Geographical Information Science*, 29(12):2248–2268, December 2015.
- Building Industry SMART International Foundation Classes (IFC), IFC4, 2013. Available from: <http://www.buildingsmart-tech.org/specifications/ifc-releases/summary>.
- Tomo Cerovsek. A review and outlook for a ‘building information model’ (BIM): A multi-standpoint framework for technological development. *Advanced Engineering Informatics*, 25:224–244, 2011.
- Ruben de Laat and Léon van Berlo. Integration of BIM and GIS: The development of the CityGML GeoBIM extension. In T. H. Kolbe, G. Köning, and C. Nagel, editors, *Advances in 3D Geo-Information Sciences*, Lecture Notes in Geoinformation and Cartography, pages 211–225. Springer-Verlag, Berlin Heidelberg, 2011.
- Yichuan Deng, Jack C P Cheng, and Chima Anumba. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, 67:1–21, July 2016.
- Abdoulaye Diakité and Jantien Stoter. Eindrapport scoping studie voor integratie geotop en bim: Als input voor de ontwikkeling van basis registratie ondergrond. Technical report, Delft University of Technology, jun 2017.
- Sjors Donkers, Hugo Ledoux, Junqiao Zhao, and Jantien Stoter. Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Transactions in GIS*, 20(4):547–569, 2016.
- Mohamed El-Mekawy and Anders Östman. Semantic mapping: an ontology engineering method for integrating building models in IFC and CityGML. *Proceedings of the 3rd ISDE Digital Earth Summit*, pages 12–14, 2010.
- Mohamed El-Mekawy, Anders Östman, and Khurram Shahzad. Towards interoperating CityGML and IFC building models: A unified model based approach. *Advances in 3D Geo-Information Sciences*, pages 73–93, 2011.
- Mohamed El-Mekawy, Anders Ostman, and Ihab Hijazi. A unified building model for 3D urban GIS. *ISPRS International Journal of Geo-Information*, 1:120–145, July 2012.
- Marcus Goetz. Towards generating highly detailed 3D CityGML models from OpenStreetMap. *International Journal of Geographical Information Science*, 27(5):845–865, May 2013.
- Peter Hachenberger. *Boolean Operations on 3D Selective Nef Complexes Data Structure, Algorithms, Optimized Implementation, Experiments and Applications*. PhD thesis, Saarland University, 2006.
- Ihab Hijazi, Manfred Ehlers, and Sisi Zlatanova. BIM for geo-analysis (BIM4GEOA): set up of 3D information system with open source software and open specifications (OS). In *Proceedings 5th International 3D Geoinfo conference*, pages 45–49, Berlin, Germany, 2010.
- ISO. *Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries*. International Organization for Standardization, March 2013.

- ISO. *Industrial automation systems and integration - Product data representation and exchange*. International Organization for Standardization, August 2014.
- Mohsen Kalantari. Future city pilot-1: Using ifc/citygml in urban planning engineering report. Ogc engineering report, Open Geospatial Consortium, jun 2017.
- Tae Wook Kang and Chang Hee Hong. IFC-CityGML LOD mapping automation using multiprocessing-based screen-buffer scanning including mapping rule. *KSCE Journal of Civil Engineering*, 22(2):373-383, 2018.
- Taewook Kang. Development of a conceptual mapping standard to link building and geospatial information. *ISPRS International Journal of Geo-Information*, 7(162), 2018.
- Hugo Ledoux. On the validation of solids represented with the international standards for geographic information. *Computer-Aided Civil and Infrastructure Engineering*, 28(9):693-706, 2013. ISSN 1467-8667. doi: <http://dx.doi.org/10.1111/mice.12043>.
- Xin Liu, Xiangyu Wang, Graeme Wright, Jack Cheng, Xiao Li, and Rui Liu. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 6(2):53, February 2017.
- Dan McKenney. Model quality: The key to CAD/CAM/CAE interoperability. Technical report, International TechneGroup Incorporated, 1998.
- Walter Nef. *Beiträge zur Theorie der Polyeder*. Herbert Lang, Bern, 1978.
- OGC. Geography markup language (GML) encoding specification. Open Geospatial Consortium inc., 2004. Document 03-105r1, version 3.1.1.
- Open Geospatial Consortium. OGC City Geography Markup Language (CityGML) Encoding Standard 2.0.0. Technical report, April 2012.
- Azarakhsh Rafiee, Eduardo Dias, Steven Fruijtier, and Henk Scholten. From BIM to geo-analysis: View coverage and shadow analysis by bim/gis integration. *Procedia Environmental Sciences*, 22:397-402, 2014. ISSN 1878-0296. doi: <http://dx.doi.org/10.1016/j.proenv.2014.11.037>. URL <http://www.sciencedirect.com/science/article/pii/S1878029614001844>.
- A. A. G. Requicha. Representation of rigid solids—theory, methods and systems. *ACM Computing Surveys*, 12(4):437-464, 1982.
- Yongze Song, Xiangyu Wang, Yi Tan, Peng Wu, Monty Sutrisna, Jack C. P. Cheng, and Keith Hampson. Trends and opportunities of BIM-GIS integration in the architecture, engineering and construction industry: A review from a spatio-temporal statistical perspective. *ISPRS International Journal of Geo-Information*, 6(397), 2017.
- Evan E. Sutherland, Robert F. Sproull, and Robert A. Schumacker. A characterization of ten hidden-surface algorithms. *ACM Computing Surveys*, 6(1):1-55, jan 1974. doi: [10.1145/356625.356626](https://doi.org/10.1145/356625.356626). URL <https://doi.org/10.1145/356625.356626>.
- Rebekka Volk, Julian Stengel, and Frank Schultmann. Building information modeling (BIM) for existing buildings—literature review and future needs. *Automation in Construction*, 38:109-127, 2014.
- WP Wang and KK Wang. Geometric modeling for swept volume of moving solids. *Computer Graphics and Applications, IEEE*, 6(12):8-17, 1986.