Qualifying spatial information for underground volumes

Kamel Adouane[a], Fabian Boujon[a], Bernd Domer[a]
[a]HES-SO, University of Applied Sciences and Arts Western Switzerland, Geneva, Switzerland
kamel.adouane@hesge.ch

Abstract. Urbanization and condensation of habitants per m^2 have led to an intense use of subsurface volumes as construction space. Planning and constructing in such spaces is a very challenging task, since knowledge of existing objects is fragmentary and imprecise. An intelligent identification of present objects and thereby detecting available volumes would increase the design quality of projects, since incidents reported during field excavations (Tanoli et al., 2019) are numerous and costly. Combining existing official territorial data with intelligent methods for information completion, compliance checking and data management, is a promising approach as it has been partially demonstrated by the use of ontologies (Caselli et al., 2020; Métral et al., 2020). The minimum level of necessary information for a model-checking framework is identified and formalized by an ontology. The ontology then serves as a basis schema for a triple store database, storing data, completion and compliance rules. The process of data completion allows to qualify the confidence in spatial information delivered.

1. Introduction

Worldwide, the field of construction is influenced by the development of urban underground spaces (Bobylev and Sterling, 2016). However, this increasing utilization should be in phase with functionalities deployed in cities (Admiraal and Cornaro, 2017). The UUS density metric is proposed as an indicator to improve the management of energy demand in urban areas (Bobylev, 2016a). Overall, the Underground Sustainable Project Appraisal Routine (Uspear) provides guidelines structure subsurface projects (Zargarian et al., 2018).

Geneva based experts (SIG, 2020), analyzed ground-penetrating radar (GPR) as a solution to measure missing data of existing utility networks. It has been found, that this solution is costly and not precise enough. Additionally, time consuming post-processes are required to handle measured data-sets. The UK-based project Mapping the Underworld (MTU) proposes ray tracing (Shan et al., 2006) as an alternative.

Shortly after the emergence of the BIM methodology in the AEC industry, it became clear that combining BIM with GIS-type data would increase the quality of territorial and urban planning. Use cases, as checking the occupation of subsurface volumes, explicitly need GIS data to discover free installation space and BIM data to represent, *e.g.*, utility networks.

3D geoinformation embedded in city models can serve as a basis for several use cases, as presented by Biljecki (Biljecki, Stoter, et al., 2015). Examples cited are the energy demand estimation on small scale to assess the return of average building energy retrofits, the visibility analysis using 3D city models in order to determine the sky view factor metric required for thermal comfort analyses or the automatic identification of suitable roof surfaces for the installation of photovoltaic panels. Solar installation potential is especially sensitive to the positioning of the city model (Biljecki, Heuvelink, et al., 2015), as an uncertainty of 50 cm could lead to a variation of 10 % in the estimation of produced energy. It would be advantageous, if such use cases could integrate BIM data. The challenge is that BIM and GIS systems apply different concepts for interoperability, which are difficult to match.

Several proposals have been made to provide convergence for BIM and GIS data structures. Stouffs (Stouffs et al., 2018) uses a triple grammar approach: a solution is developed to map

BIM-IFC type graph towards GIS-CityGML graphs. Adouane (Adouane et al., 2019) presented a specific use case to map a complete building from a BIM-IFC format towards GIS-CityGML. His methodology has been validated on a general architectural model containing complex geometries. Biljecki has developed an ADE (Application Development Extension) to automate the BIM-IFC conversion towards GIS-CityGML format. The conversion strategy has been tested in collaboration with the Building and Construction Authority (BCA) of Singapore (Biljecki et al., 2021).

Interoperability for BIM systems is supported worldwide by buildingSMART international, in particular by their IFC standard. Pauwels (Pauwels et al., 2017) proposes to use Web ontology language (OWL) to specify IFC. An OWL representation facilitates the mapping between data models, like IFC and CityGML.

Xu and Cai (Xu and Cai, 2020) are using ontologies to describe and to manage heterogeneous data sets of underground utilities: they integrate digital conversion tools for spatial relations among objects. Building code compliance is checked through SPARQL queries on triple store databases.

As shown by the examples above, BIM and GIS convergence can be achieved. It should be mentioned, that due to the different objectives of such systems, not all BIM information is transferable into a GIS system and vice versa.

Official GIS databases, like the Geneva "SITG" (SITG, 2020) system, contain precise data for surface objects. When subsurface volumes are considered, existing data is less precise and complete and in most cases not sufficient for a 3D representation. Data completion through in-situ measurements is complicated and costly. This hinders the correct representation of position and geometry of existing subsurface objects.

2. The InnoSubsurface project

The overall objective of the "InnoSubsurface" project is to support subsurface project planning by proposing solutions for a better management of such volumes. As a first step, a taxonomy of subsurface objects and their necessary attributes for 3D representation and planning has been created. This structure is called "minimal data model". The geometric model of the subsurface objects uses only primitives like extruded polygons, cylinders or truncated cones. It accommodates natural elements, like trees, manmade objects, like utility lines and public law restrictions, like contaminated sites. The data model has been transferred into an ontology, which integrates IfcOwl as well as CityGML elements (Caselli et al., 2020; Métral et al., 2020).

The ontology serves as a data schema for a triple store, populated by data from the "SITG" database. As expected, provided information is not sufficient for a 3D representation as defined by the minimal model. Therefore, a completion strategy for positioning and geometrical attributes had to be developed.

Object attributes of the minimal model related to position and geometry are associated to a confidence level. The confidence level might represent measurement precision or the confidence associated to an attribute derived by a completion strategy. Completed objects are stored in the triple store database. Hypotheses used for completion are called "Completion rules". They are derived by construction codes or interviews with practitioners and formulated according to a generic rule model, described in (Caselli et al., 2020). Rules are stored together with subsurface objects.

Although completion rules can be defined for the majority of attributes on a theoretical level, the completed data might give unrealistic results on a practical level. A first proposal for a metric to qualify subsurface volumes is made by Bobylev (Bobylev, 2016b). He relates subsurface volumes to ground surface. His metric does neither provide an indicator on the quality of data used to calculate the volume of subsurface objects nor on the quality of their position.

The paper highlights the following aspects of the “InnoSubsurface” project:

- The integration of confidence levels for positioning and geometric attributes of subsurface objects
- The combination of object confidence on a class level
- The visual representation of the degree of confidence
- A first approach to qualify spatial information for underground volumes with respect to constructability, integrating data completeness, precision as well as the metric proposed by Bobylev.

3. Methodology

3.1 Using probability functions to represent confidence in subsurface object position and geometry attributes.

Positioning and geometrical representation of objects are based on measured and empirical elements (completion rules). In order to represent the precision of such attributes, we propose to use simple probabilistic functions. Their integration vary according to the nature of the imprecision: has the value been measured or derived by a completion rule?

Each object possesses two visual representations:

- The core representation is called “primary object” and specifies the completed object derived by SITG and completion rules.
- The second representation is called “secondary object“. The secondary object envelops the primary one and indicates the confidence, that a given object can be found inside its boundaries. The developed stochastic model allows the user to choose the desired confidence level, which affects the size of the secondary object.

3.1.1 Triangle probability function in multiple dimensions

Measured attributes can reach, according to expert interviews, a maximum of 95%. A triangle distribution, with an overall degree of confidence of 95%, is assigned to model the precision of positioning measurements. Based on SITG description, the precision of x and y coordinates of the cantonal database is +/-10 cm. Figure 1 illustrates the model for the x coordinate of a tree root.

The concept of “primary” and “secondary” objects is shown Figure 2. The blue cylinder represents the primary object, the red cylinder the secondary object.

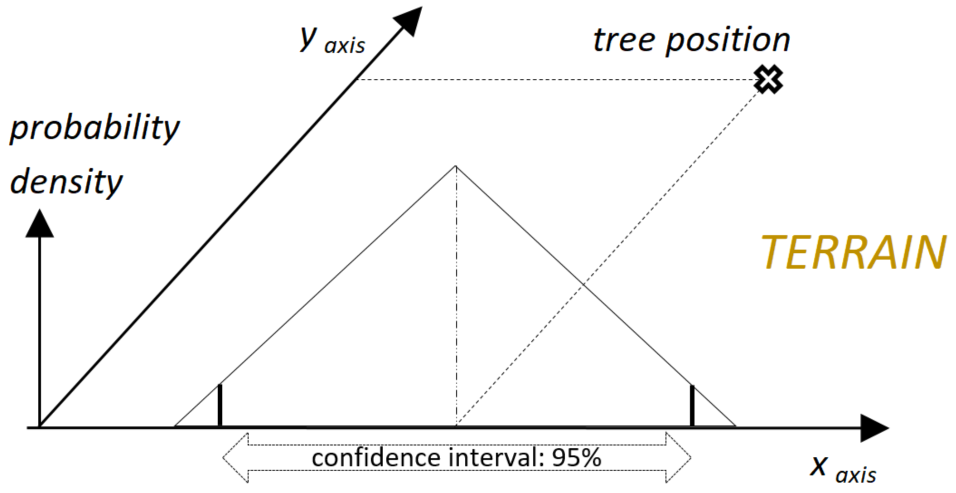

Figure 1: Triangle probability density distribution function for the x coordinate of the tree position

Figure 2 indicates how multiple probabilities for single attributes are modeled. Since horizontal positioning needs x and y coordinates, the confidence interval of the two has to be combined. Depth information is related to a "step" probability function. The uncertainty for the tree root model in Figure 2 is estimated by Equation 1.

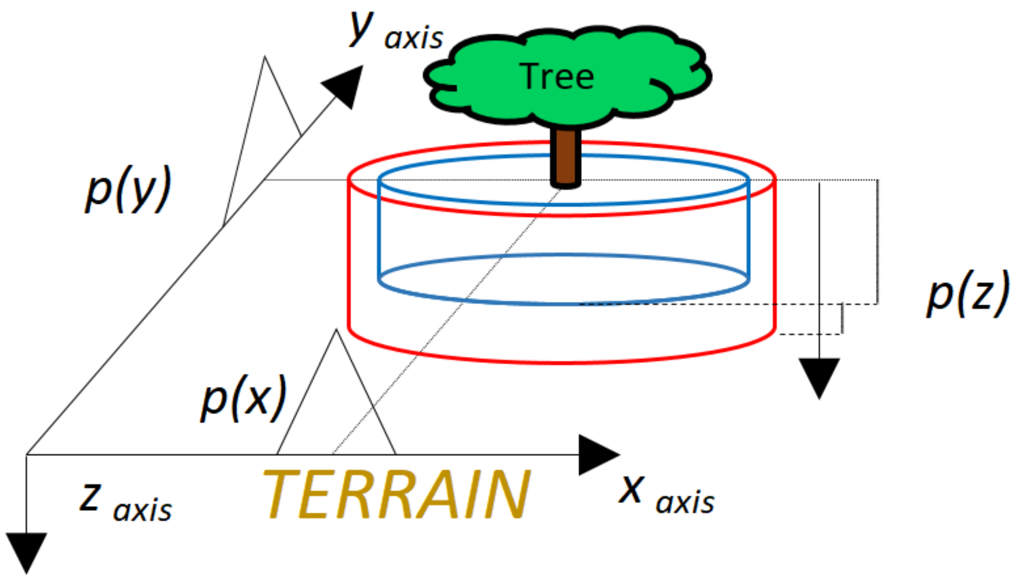

Figure 2: Example for primary (blue cylinder) and secondary objects (red cylinder) of tree roots with primitive probabilistic density functions associated to positioning and depth attributes

Equation 1

$$Object_{uncertainty} = \prod_{dimension} p$$

3.1.2 The Dirac probability density function

The Dirac probability density function represents the confidence used in completion rules for empirical single values. These are, for example, a standard height and quantity for basement floors, a standard diameter for utility pipelines, etc. Figure 3 presents the model of the Dirac primitive for a gas network node.

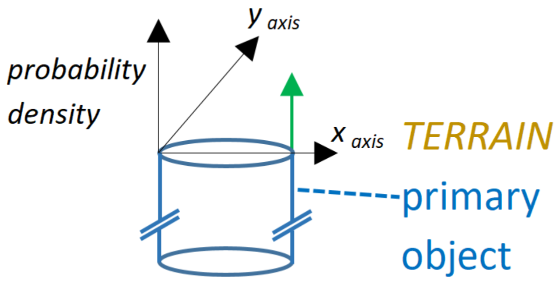

Figure 3: Dirac probability density distribution function (green arrow) applied to the diameter of a gas node

The maximum level of confidence for such a completion is set to 80 %, based on expert interviews.

3.1.3 The Pert probability density function

The confidence in the depth of utility networks is modeled by a Pert probability density function. When depth information for a particular network is unknown, a Pert function based on neighboring networks of the same type, containing the desired depth information, is established. The function is characterized by the triplet {a,b,c}, fitted by the least square method to the depth distribution histogram.

Figure 4 shows how the Pert function is used to place the primary and secondary object of a gas network. The top of the primary object is placed at a depth of PERT coeff b, (90 cm in this example). The secondary object is modeled by a bounding rectangle around the pipeline diameter. The maximal lateral limits of the secondary object are obtained using the triangle probability function of Figure 1, since x and y coordinates are known. The upper and lower bounds of the secondary object are calculated by adding a second component, obtained by subtracting the measurement uncertainty from PERT coeff a (for the upper limit) and by adding the measurement uncertainty to PERT coeff c (for the lower limit).

As Figure 4 demonstrates, the size of the secondary object varies with the confidence interval chosen by the user.

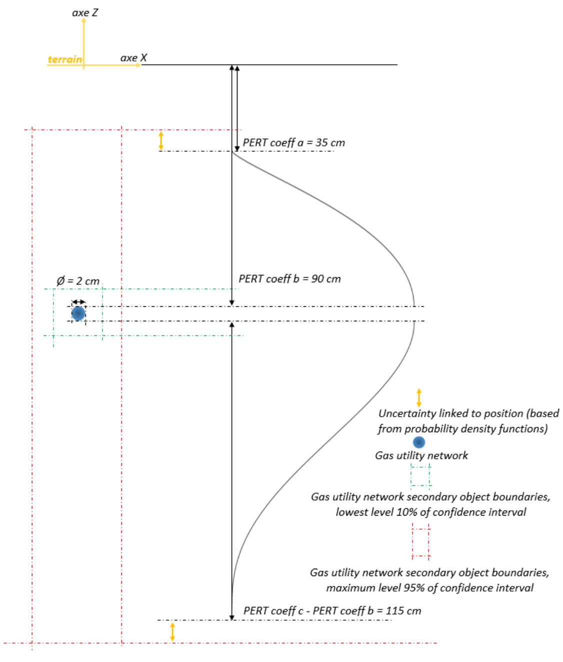

Figure 4: Pert probability density distribution function applied to the positioning of a gas utility network

3.2 Combining attribute confidence for a class of objects

For a given class of objects, like all tree roots, the confidence level can be consolidated according to Equation 2. $Volume_{secondary}$ represents the volume of the secondary volume, $Object_{uncertainty}$ represents the object uncertainty introduced in section 3.1.1.

Equation 2

$$confidence = \frac{\sum(Volume_{secondary} * Object_{uncertainty})}{\sum Volume_{secondary}}$$

3.3 Visual representation of confidence

Each object is visualized by a twofold 3D representation, a primary and a secondary object, as introduced in section 3.1. In general, the secondary object possesses the same geometry as the primary. Figure 5 shows the only exception: for practical reasons, conducts are associated with a cuboid. As the user can choose the confidence level, the right side of the same figure shows the effect on the size of the secondary object.

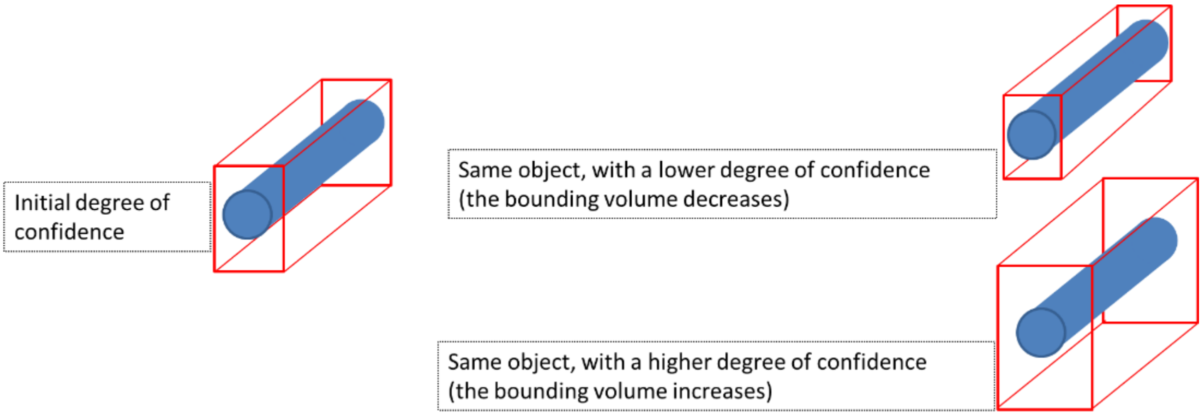

Figure 5: Primary and secondary objects for utility networks. The effect of varying confidence levels is shown on the right.

3.4 A first approach to qualify spatial information for underground volumes

An underground volume contains a finite number of objects with a finite number of geometrical and positioning attributes, which are required to visualize the primary object. Available data is analyzed to identify the number of missing attributes. This number is related to the total number of attributes required.

As the volume of objects is not taken into account, the Completeness Ratio (Equation 3) only describes the information maturity level within the database. Small objects are given the same weight as larger ones.

Equation 3

$$CompletenessRatio = \frac{\sum_{Object_{Class} \epsilon Taxonomy} \sum_{Object \epsilon Object_{Class}} PrimitiveParametersCount_{Existing}(Object)}{\sum_{Object_{Class} \epsilon Taxonomy} \sum_{Object \epsilon Object_{Class}} PrimitiveParametersCount_{Required}(Object)}$$

The Completeness Ratio can be refined (Equation 4) when calculated separately for each object class. An average can then be obtained for all present object classes. This leverages the parasite effects created by objects with a bigger number of attributes or are present in a greater number than others in the evaluated volume.

Equation 4

$$RefinedCompletenessRatio = \frac{\sum_{Object_{Class} \epsilon Taxonomy} \frac{\sum_{Object \epsilon Object_{Class}} PrimitiveParametersCount_{Existing}(Object)}{\sum_{Object \epsilon Object_{Class}} PrimitiveParametersCount_{Required}(Object)}}{Cardinal(Object_{Class} \epsilon Taxonomy)}$$

4. Results/Validation

The methodology has been applied to two subsurface volumes in the center of Geneva: a first one being nearby the main train station (Cornavin, 0.32 km^2) and a second one located around the Arve river (PAV, 0.31 km^2).

4.1 Visual representation of results

Information related to subsurface objects has been extracted from the SITG database for the two zones. The content has been stored in the triple store and missing positioning and geometrical attributes have been found by applying completion rules. Secondary objects are created based on the desired confidence level. In this sector, the confidence for all utility networks is evaluated to 92% by Equation 2. Finally, a GIS-Frontend is used to visualize the results (Figure 6).

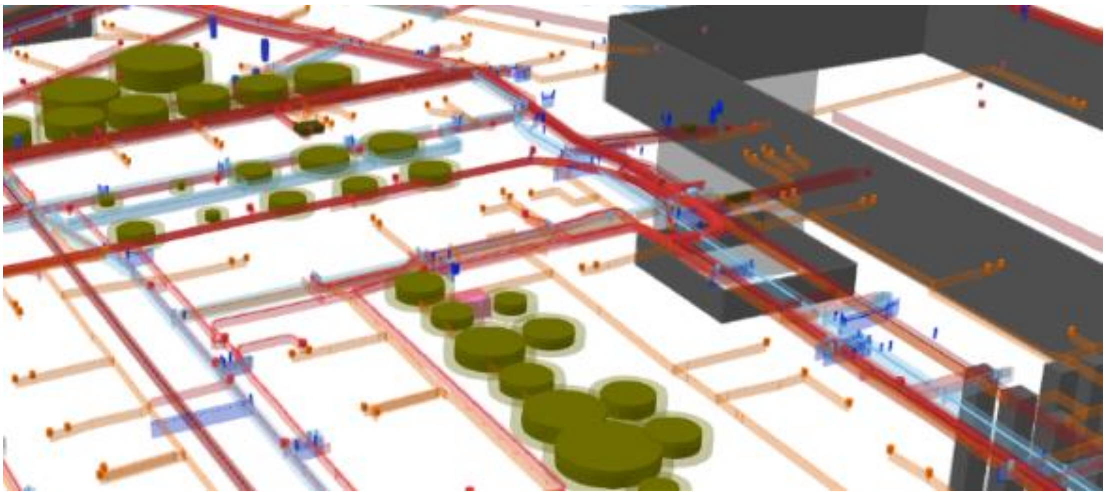

Figure 6: 3D viewing of underground volumes, developed (Topomat, 2021)

Table 1 indicates the colors applied to the different objects, based on Swiss construction codes.

Table 1 : Color codes used in Figure 6

Colour	Object
green	tree root
blue	natural water network
brown	lighting network
red	electricity network
pink	recycling site
grey	geotechnical

4.2 Qualifying existing spatial information for two volumes

The Urban Underground Space metric (UUS) (Bobylev, 2016b) is determined in order to validate the results of our project. For PAV and Cornavin areas, we obtain a density that is comparable to the results of underground volumes in Berlin (Table 2).

Table 2: UUS for PAV and Cornavin zones

metric	**PAV sector**	**Cornavin sector**	**Berlin** (Bobylev, 2016b)
surface, km^2	0.31	0.32	
primary volumes, m^3 (000 000)	0.49	0.42	
UUS m^3/m^2 (in cm)	156.8	133.7	128.0

Table 3 exposes the results of Equation 3 and Equation 4 applied to the two zones (Cornavin and PAV). The completeness ratio is calculated to approximately 80%, the refined metric results in approximately 73%. In addition, the total number of geometric parameters required to represent the volume, is indicated.

Table 3: description of data completeness ratios for PAV and Cornavin zones

Area	**Completeness ratio**	**Refined completeness ratio**	**Gross geometric parameters**
Cornavin	80.22%	72.33%	387’282
PAV	79.36%	73.76%	147’468

5. Conclusion and future work

The AEC industry needs to capture possibilities offered by the digital transition in order to speed up to industry 4.0. Data driven civil and underground engineering are two domains affected by this transition. The wideness and variety of the data available is advantageous but subject to errors. In addition, the data is heterogeneous in precision, completeness, accuracy, level of details and format. Intelligence based processes to automatically correct datasets are therefore required to make those data useful for analysis and design purpose.

Curation and processing of uncertain and incomplete subsurface data prompts research on models to represent uncertainty and to process data with different confidence levels. This paper shows that even imprecise and incomplete data can be applied to provide a coherent representation of subsurface volumes. The proposed concept to associate objects to a confidence level and to inform the user about data quality is unusual but helpful.

Only simple geometric representation and probability functions have been used to facilitate the understanding and control of the workflow. The developed methodology is independent from structure and quality of available subsurface data. Of course, completion strategies and confidence model will have to be checked before being applied to other locations with different database concepts. The overall architecture of the system, based on an ontology and a generic rule model, will nevertheless ease such an adaptation.

A threshold, indicating when data completion strategies will become senseless, is needed. The qualifying metrics employed (UUS, completeness ratio, refined completeness ratio) have to be tested on this question and improved.

The InnoSubsurface project investigated into the application of "Compliance rules", defining spatial constraints on objects, as well. Besides the detection of geometric conflicts, these rules are good candidates to be employed, *e.g.*, in order to automatically disentangle the crossing of multiple utility lines.

6. Acknowledgements

This research is supported by Innosuisse in the framework of the Innovation project 35265.1 IP-ICT, "Impulse-Subsurface: Efficient data exploitation in subsurface planning". The authors would like to thank, Mr Loïc Neuenschwander and Mr Yohann Schatz (HEPIA)

Ontology creation, conceptualization of the generic rule model and set up of the triple store database: UNIGE (Pr Giovanna Di Marzo Serugendo, Dr Claudine Metral, Pr Gilles Falquet, Dr Vincenzo Daponte, Mr Ashley Caselli)

Industrial partner, GIS-expert, representation and analysis of subsurface volumes in a GIS-Frontend: TOPOMAT (Mr Stéphane Couderq, Mrs Marie-Christine Nicolle, Mr Christophe Suter, Mr Alexandre Gauch)

Provision of data and professional knowledge: State of Geneva (State Direction for Territory Information, Department of Urbanism, Department of Transports, Department of civil engineering, Department of Energy, Department of Water, Industrial Services of Geneva and Geneva Airport)

References

Admiraal, H. and Cornaro, A. (2017) Underground space., Tonbridge, Ice Publishing.

Adouane, K., Stouffs, R., Janssen, P. and Domer, B. (2019) A model-based approach to convert a building BIM-IFC data set model into CityGML, Journal of Spatial Science, pp. 1–24 [Online]. DOI: 10.1080/14498596.2019.1658650.

Biljecki, F., Heuvelink, G. B. M., Ledoux, H. and Stoter, J. (2015) Propagation of positional error in 3D GIS: estimation of the solar irradiation of building roofs, International Journal of Geographical Information Science, vol. 29, no. 12, pp. 2269–2294 [Online]. DOI: 10.1080/13658816.2015.1073292.

Biljecki, F., Lim, J., Crawford, J., Moraru, D., Tauscher, H., Konde, A., Adouane, K., Lawrence, S., Janssen, P. and Stouffs, R. (2021) Extending CityGML for IFC-sourced 3D city models, Automation in Construction, vol. 121, p. 103440 [Online]. DOI: 10.1016/j.autcon.2020.103440.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S. and Çöltekin, A. (2015) Applications of 3D City Models: State of the Art Review, ISPRS International Journal of Geo-Information, vol. 4, no. 4, pp. 2842–2889 [Online]. DOI: 10.3390/ijgi4042842.

Bobylev, N. (2016a) Transitions to a High Density Urban Underground Space, Procedia Engineering, vol. 165, pp. 184–192 [Online]. DOI: 10.1016/j.proeng.2016.11.750.

Bobylev, N. (2016b) Underground space as an urban indicator: Measuring use of subsurface, Tunnelling and Underground Space Technology, vol. 55, pp. 40–51 [Online]. DOI: 10.1016/j.tust.2015.10.024.

Bobylev, N. and Sterling, R. (2016) Urban underground space: A growing imperative, Tunnelling and Underground Space Technology, vol. 55, pp. 1–4 [Online]. DOI: 10.1016/j.tust.2016.02.022.

Caselli, A., Daponte, V., Falquet, G. and Métral, C. (2020) A Rule Language Model for Subsurface Data Refinement,.

Métral, C., Daponte, V., Caselli, A., Di Marzo, G. and Falquet, G. (2020) ONTOLOGY-BASED RULE COMPLIANCE CHECKING FOR SUBSURFACE OBJECTS, London, UK, vol. XLIV-4/W1-2020, 2020 [Online]. DOI: https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-91-2020.

Pauwels, P., Zhang, S. and Lee, Y.-C. (2017) Semantic web technologies in AEC industry: A literature overview, Automation in Construction, vol. 73, pp. 145–165 [Online]. DOI: 10.1016/j.autcon.2016.10.003.

Shan, Q., Pennock, S. R. and Redfern, M. A. (2006) Investigation of GPR Configurations by Ray-Tracing Methods, 2006 IEEE Conference on Radar, Syracuse, NY, USA, IEEE, pp. 335–341 [Online]. DOI: 10.1109/RADAR.2006.1631821 (Accessed 31 March 2021).

SIG, S. I. de G. (2020) [Online]. DOI: https://ww2.sig-ge.ch/.

SITG, L. territoire genevois à la carte, Accessed 26 Feb (2020) [Online]. DOI: https://ge.ch/sitg/.

Stouffs, R., Tauscher, H. and Biljecki, F. (2018) Achieving Complete and Near-Lossless Conversion from IFC to CityGML, ISPRS International Journal of Geo-Information, vol. 7, no. 9, p. 355 [Online]. DOI: 10.3390/ijgi7090355.

Tanoli, W. A., Sharafat, A., Park, J. and Seo, J. W. (2019) Damage Prevention for underground utilities using machine guidance, Automation in Construction, vol. 107, p. 102893 [Online]. DOI: 10.1016/j.autcon.2019.102893.

Topomat (2021) Switzerland partner for Geographical Information Systems,.

Xu, X. and Cai, H. (2020) Semantic approach to compliance checking of underground utilities, Automation in Construction, vol. 109, p. 103006 [Online]. DOI: 10.1016/j.autcon.2019.103006.

Zargarian, R., Hunt, D. V. L., Braithwaite, P., Bobylev, N. and Rogers, C. D. F. (2018) A new sustainability framework for urban underground space, Proceedings of the Institution of Civil Engineers - Engineering Sustainability, vol. 171, no. 5, pp. 238–253 [Online]. DOI: 10.1680/jensu.15.00013.