

# HBIM-BASED INFORMATION SHARING FOR THE EXCHANGE AND SAFEGUARD OF MODELLING EXPERTISE

O. Rosignoli<sup>1</sup>

<sup>1</sup> He.Su.Tech. group, MantovaLab, Dept. of Architecture Built environment and Construction engineering (ABC), Politecnico di Milano, 46100 Mantua, Italy, olga.rosignoli@polimi.it

**KEY WORDS:** datasets, HBIM, Semantics, 3D, expertise, RDF.

## ABSTRACT:

This paper demonstrates an approach for sharing and safeguarding modelling expertise based on the use of Historic Building Information Modelling (HBIM). The approach involves the exchange of information between subjects involved in the creation of HBIM models, an exchange enhanced by Semantic Web Technologies.

HBIM datasets usually contain rich and detailed information about historic buildings and their construction; this experiment, in a slightly unusual way, deals with another type of information regarding something that happens in the background during the definition of an HBIM model: the procedures that create the complex three-dimensional geometries typical of historical architecture. This kind of information does not usually live inside BIM models and is even less likely the subject of semantic enrichment. It is however beneficial to record and make available this kind of information for several reasons, discussed in the text.

## 1. INTRODUCTION

### 1.1 Research question

HBIM research has focused important efforts on how to reproduce the complex shapes typical of historical architecture elements, making the Scan-to-BIM/HBIM process a heavily investigated issue throughout all projects involved. As the digital modelling of these shapes is oftentimes based on very precise digital surveys, the three-dimensional copies of their geometry have the potential to achieve optimal degrees of accuracy, a result towards which these researches aim intently.

Although great improvements have been obtained on the subject of automation (Andriasyan et al., 2020),(Nieto-Julián et al., 2019) several actions are still necessary to translate the geometries coming from a detailed point-cloud into an HBIM environment with the same good accuracy; these passages can be numerous and complex, according to the requirements of the project, the time available and the skills of the professionals in charge, as well as more or less automated (Banfi, 2019), (Moyano et al., 2022).

When the built assets are large (as in for the majority of historically important complexes) several modelling strategies have to be implemented for the homogeneous representation of all the different shapes involved. Therefore, during the process, several choices are made by the expert (or the group of experts) creating the model.

No matter how automated or precise, a digital representation of a real object always involves upstream discretizations that affect the appearance of the replica obtained and therefore the performance of the digital model. Level of Detail (LOD) indications and the information agreements made beforehand by the figures involved in the projects are the framework guiding the discretization choices; afterwards the somewhat tacit expertise of the professional figures in charge of the decisions and the decision processes plays a major role in the representation of these complex shapes.

This work investigates how to translate and incorporate in the HBIM process the knowledge that unfolds behind the scenes of the three-dimensional modelling phase of the HBIM workflow.

Documenting this knowledge is crucial for many reasons: one in particular is the structural analysis, which could benefit from detailed explanation about how the geometries are formalised. The article also investigates how the information collected has the potential to become instructional for future projects where similar difficulties could arise.

In essence, the safeguard of instrumental knowledge that enables to shape these digital assets can serve many purposes.

### 1.2 Relevance

BIM technologies and processes allow for information of various type to be incorporated into digital copies of built environment and exchanged between several actors. The possibilities for interchange of said information increase when Semantic Web Technologies (Berners-Lee et al., 2002), through their interaction with BIM technologies are taken into consideration. By exploiting these technologies (and the semantic structures upon which they are built) new and unpredicted categories of information can be involved in the information exchange process. The objects of this work are, in particular, the pieces of information explaining the 3D modelling choices; the intention is to translate this information in a Semantic Web friendly way and exploit this environment and the new possibilities it opens.

### 1.3 Proposed solution

The connection between HBIM and Semantic Web Technologies (Cursi et al., 2022), allows for more freedom and more possibilities of dealing with information: the rigid structure they are built on (both in their native BIM environments and in their IFC translation) can in fact be expanded in a web environment where other semantic structures can be associated. This operation is called "semantic enrichment".

A dataset coming from the HBIM model of a heritage building is used for the experiment.

Ontologies are the structures upon which the Semantic Web is structured. Ontology reuse is foreseen in this project, according to on-going research. Therefore, some classes and relations of existing ontologies are chosen and applied on the HBIM dataset, so that the information about the modelling process would be

structured accordingly to existing knowledge structure and their meaning would be elicited. Given the peculiarity of the research objective for this work, specific existing ontologies (regarding the construction of an HBIM model geometry) were not found available. However, the most common knowledge structures approved by the World Wide Web Consortium (W3C, <https://www.w3.org/>) offer classes for broader concepts such as "note", "creator", "label", very close to the BIM model data. Somewhat surprisingly, their being more generic and less specific can possibly ensure better sharing possibility and a longer lasting life. Other concepts, inevitably and for the specificity of the project, had to be built from scratch. The article also discusses the dilemma between structuring the data in a IFC compliant way and converting the modelling with the information incorporated or bypassing this step, as other examples in literature have performed (Bassier et al., 2019). The end goal, in any case, is to have the information structured according to the Linked Open Data framework, which is the standard for information exchange for the Semantic Web.

## 2. MAIN BODY

### 2.1 Description of the study sample:

One particularly rich case study is used as an example here: the as-built HBIM model of the large church complex of San Giovanni Evangelista in Parma (Italy). The model is the translation of a very comprehensive laser-scanner and photogrammetric (both from drone and terrestrial) survey (Figure 1).

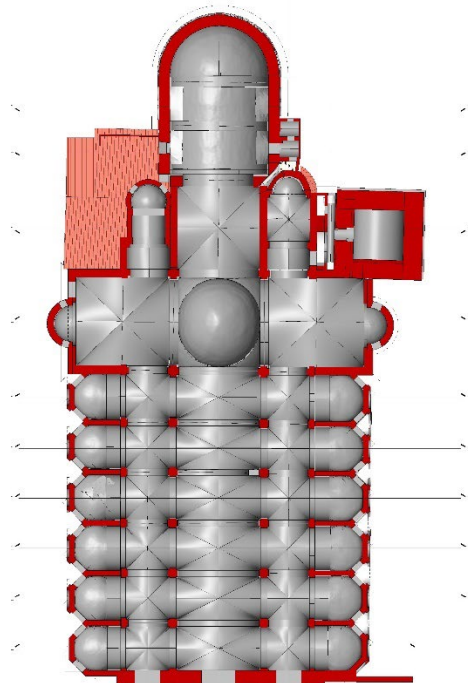


**Figure 1:** point-cloud of the complete digital survey for San Giovanni Evangelista church, Parma (courtesy of HeSuTech group).

The point-cloud was georeferenced according to the topographic survey network. Different datasets were the output of this phase: the point-cloud of the interior, of the exterior, and of the attic spaces. The clouds were exported in the .e57 format as well as the proprietary .rcp Autodesk Recap© format. The datasets were then inserted into the Autodesk Revit© environment and used to plot the object according to the surveyed conditions. Thanks to their combination, the thickness of all architectural elements was described.

The goal of the digitization process was to provide material for the structural analysis of the complex, performed by the engineers involved in the project contract. The main technology discussed in this work are the vaults of the church building: vaults are some of the most representative elements of historic architecture, and at the same time amongst the most difficult to represent. Oftentimes, other software (e.g. McNeel Rhinoceros©, Autodesk AutoCAD©) is used as an intermediate modelling tool (Banfi, 2017). This software is capable of conveying highly complex surfaces based on the point-clouds, elements which BIM then uses to draft its components on. The information exchanged involves and explain also these important external passages. The vaults of San Giovanni (Figure 2) are several and very diverse: their side length can span from a value of 1.5 m to a value of 10 m. Sometimes, tools native to Revit such as local masses and adaptive components were sufficient; other times the CAD approach was necessary.

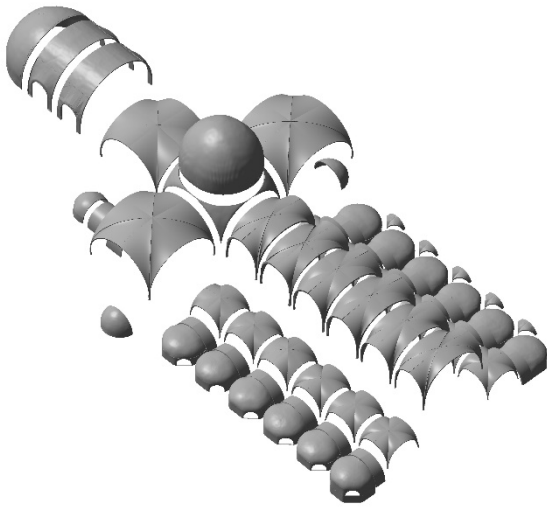
The point-cloud resulted from the survey was the starting point and the reference check for the geometries plotted, according to the requirements of the engineering team (less than 0.02 m deviation). The vaults modelled for the main body of the church are 63.



**Figure 2:** the very diverse vaulted ceiling of San Giovanni Evangelista; inverted plan view of the HBIM model.

## 2.2 Description of the strategy

One dataset is extracted from the whole model of the Church and constructed for this experiment: the "vaults" dataset. The dataset is only a partial description of the church building. More than often, BIM modelling operations satisfy specific requests from the commissioning party, and therefore focus on specific themes (in this case the theme was structural information). The complex vaulted ceiling of the Church, therefore, underwent specific 3D modelling processes, compared to other more simplified items (e.g. decorations and openings). This is the reason why annotation is considered necessary above all for this category of objects. Nothing prevents the same procedure from being implemented for other objects.



**Figure 3:** the vaults of San Giovanni in their complexity; HBIM model axonometric view.

A complex toponomy is created for the church spaces and is inherited by the BIM objects; however it might seem difficult to interpret, the coded names are both unambiguous and understandable intuitively if not immediately. The decomposition of church spaces is always a complicated matter. In the case of San Giovanni, the complexity is given by the different generative shapes (there are cross vaults, barrel vaults, domed vaults, barrel and lunetted barrel vaults) and the fact that more than one vault often covers a single space (meaning one has to differentiate or enumerate the different parts). For example: "VaultNC01" indicates the vault covering the first span of the central nave.

**2.2.1 Data exchange context:** As much as full interoperability is always the end goal of semantic enrichment projects, in the context of this publication, is limited to the engineering team, the HBIM drafting team, and the local Superintendency (the commissioning party). Therefore, some of the semantics remain decipherable only for the design team.

**2.2.2 The "vaults" dataset from the HBIM model:** The 3D modelling phases allowed to have the objects inside the BIM environment. To get to the level of definition wanted, sometimes the vaulted elements were deconstructed into smaller parts: for example, each panel of a cross vault constituted of two separate items, making a total of eight elements necessary to describe a singular vault.

"Assemblies" (objects native to Revit software) are chosen to group the elements. These objects allow scheduling and labelling, therefore are fit for data exchange. The assemblies created for the vaults are scheduled as "walls assemblies" and this is not manageable differently by the program. Each created vault is therefore a new "type" of walls assembly. The parameter "type" acts as an identifier and refers back to the custom made toponomy created for the project.

A total of 63 walls assemblies with their unique identifier correspond to the total number of vaults in the project.



**Figure 4:** the groin vault, the boundaries of the eight surfaces composing it are visible. Assemblies have a unique "Type Name".

Once the objects "assemblies type" and their identifiers were established, it was possible to begin annotating properties upon them. A list of significant parameters to describe these properties was set-up (in the perspective of telling how the vault was created), namely: the type (the identifier), the 3D modelling techniques, and the creator of the geometry. Later, two other parameters were added, one describing the path to the position of the geometry (3D files) used as intermediate tool (for the cases where a CAD intermediate surface was used) on a local server, and the other for the same purpose but on a common online storage system (Microsoft OneDrive). To resume:

1. Type: text parameter
2. Local\_RefGeom: URL parameter
3. Online\_RefGeom: URL parameter
4. ModTechnique: text parameter
5. Creator: text parameter

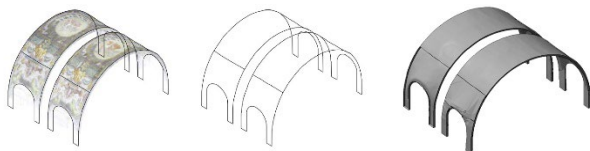
The textual parameter for the modelling technique ("ModTechnique") is the one where the record of expertise happens; the record consists in the detailed explanation of the manufacturing process of the digital vaulted element, from the point-cloud to the BIM object. Although the operations were monitored by the HBIM drafting team, the person responsible for the creation of the object ("Creator") is also the responsible for the entry of the information. The requirements given upstream were to give an easy to understand and very procedural explanation of the sequence of actions performed.

The quality of the information recorded is varied: explanations can be both short and lengthy, concise and wordy. Although some concepts repeat themselves (e.g. the use of point-clouds, the use of CAD sections, the use of some Revit families) it was counterproductive to translate it into a few clear concepts that are repeated. Once again, the many peculiarities of an historic building and their representation difficulties make it a struggle with being framed into precise categories. Consequently, almost all modelling actions presented many difficulties that made each modelling process a process on its own. Therefore, a single piece

of textual information was considered more adapt then information structured into many parameters/concepts. This approach allowed more flexibility. Even if this work does not want to become a catalogue for explaining all of the modelling techniques used, there follow some examples, to clarify the quality of the information. The premise is that people receiving the information are familiar with the jargon from the 3D modelling world and of the software used, and with the reference system of the project.

An example of the information, for the lunetted barrel vault:

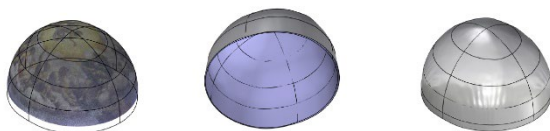
*"A wireframe CAD surface is necessary; the CAD surface is modelled in Rhino v.6, command Patch; the polylines making the wireframe structure are 6; one YZ section at the northern boundary, another YZ section at the southern boundary; one XZ section at the eastern boundary another XZ at the western boundary; one XZ section in correspondence of the joint between the vault and the eastern nail another XZ section in correspondence of the joint between the vault and the western eastern nail; the wireframe creates a three-parted .3dm geometry. Import as Cad inside a Revit local mass. Apply wall-by-surface command onto the three parts of the surface."*



**Figure 5:** the lunetted barrel vault, with the CAD wireframe on the left and the BIM object on the right

Another example for the main dome:

*"A wireframe CAD surface is necessary; the CAD surface is modelled in Rhino v.6 command Patch; the polylines making the wireframe structure are 6; a YZ section and a XZ orthogonal section at the centre. 4 XY sections starting from the bottom going up at a 0.9 m pace. Points from the point-clouds were used in the calculation of the Patch. The wireframe creates a three-parted .3dm. Import as CAD inside a Revit local mass. Apply wall-by-surface command onto the surface."*

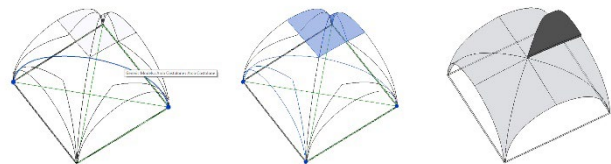
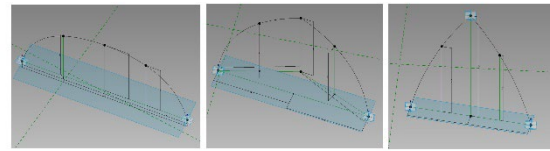


**Figure 6:** From left to right: the point-cloud points and CAD generative sections; the CAD surface in purple; the vault object in BIM.

Last example to bring to attention is the one of the cross-vault. Here the modelling solution explored (which gave the better results faster) is quite different from previous examples. The passages are also more numerous. The surface is generated inside a Revit "mass" object without the help of additional other native objects (Revit "adaptive" components). The information collected is the following:

*"The reference geometry is created inside a local mass. A wireframe of lines is created by 3 nested families of arched parametric lines (Arch01, Arch02, Arch03) inside this project. 4*

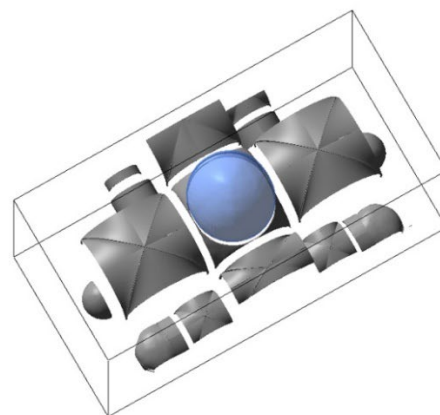
*instances of families Arch01 and Arch02 are positioned on each side of the vault according to the point-cloud. Inner height parameters are adjusted to adhere to the point-cloud. 2 instances of Arch03 are positioned diagonally. Inner height parameters are adjusted to adhere to the point-cloud. Surface is created with command Create Form selecting the arch families. 2 panels are then created for each panel of the cross vault. Apply wall-by-surface command onto all 8 the parts of the surface."*



**Figure 7:** a cross vault created with the the adaptive objects native of Revit.

The last example raises the problem of "nested" components, which are subcomponents loadable inside Revit objects. Their use is very practical and widespread in BIM modelling; those family on their own would benefit of additional explanation. An additional link to other more detailed tutorials could serve here.

Walls Assembly VaultCup	
Assemblies (1) Edit	
Comments	
Mark	
Online_RefGeom	https://polimi365-my.sharepoint.com/...
Local_RefGeom	Import:VaultCup
ModTechnique	A wireframe CAD surface is necessary; the...
Creator	Designer1
Phasing	
Phase Created	Stato di Fatto
Phase Demolished	None
IFC Parameters	
Export to IFC	By Type
Export to IFC As	
IFC Predefined Type	
IfcGUID	3WS3kHgRn6PBxVicTJR15E



**Figure 8:** the properties of the vault object, with highlighted the associated parameters in the model.

It goes without saying that the reference geometries should favour formats as open as possible for the members of the project. To end the creation of the vaults dataset, in a proper BIM schedule a list of all the parameters associated to each item was therefore created. This list was then exported in .csv format.

**2.2.3 Debate over the use of IFC:** Resuming the open format concept expressed just before, it is important to dedicate some attention to IFC (2013), the Industry Foundation Classes, which is an open, international, and standardized specification for data sharing in the construction and facility management industries. IFC is a very popular and very powerful, and its use is also being studied also in relation to semantic enrichment scopes (Bonduel et al., 2018). Since the experiment required limited interoperability, and the use of "assemblies" proved to be suitable temporarily, IFC use was discarded.

### 2.3 Dataset enrichment

The nature of the vault dataset experiment was ideal to experiment with the Semantic Web technological stack and have a simple database on which to start performing semantic enrichment.

RDF, Resource Description Framework (RDF, <https://www.w3.org/RDF/>) a standard defined by the World Wide Web Consortium was chosen to perform the enrichment on the dataset, as a good standard for data storage and exchange in the Semantic Web.

The extraction of the significant data to perform the enrichment on was described in section 2.2.2. This following passage narrates the RDF triples generation and serialisation based on this data. Triples are the foundational element of any RDF file.

A data conversion script was developed using Python 3.11 programming language (Python Software Foundation, 2022) and the RDFlib library (RDFlib, n.d.), a package created to work with RDF. By feeding the .csv dataset to the script, the output becomes a set of triples serialised in an RDF file, in its Turtle format (.ttl).

The nature of RDF allows to elicit the semantics of the information. Semantics are detected because stored and recorded in the form of digital knowledge structures called "ontologies". Some ontologies are recorded on the web in the form of "web ontologies". A large number of them are also approved by the W3C and are part of the so-called "Linked Open Data framework". Full interoperability is reached when data adheres to the latter type of knowledge structures. In the case of this experiment, this happens only partially, while some other concepts are built from scratch. The list of the HBIM properties created and their corresponding concept in other existing structures are described in **Table 1**.

HBIM property	knowledge structure	corresponding class
Type	RDFS	label
Online_RefGeom	SCHEMA.org	url
ModTechnique	DC	Description
Creator	DC	Creator

**Table 1:** concepts inherited from the HBIM dataset and their correspondence (the ontology "class")

The existing knowledge structures chosen are some of the most commonly used: RDF Schema (or RDFS, W3C, 2014) Dublin Core (Dublin Core Metadata Initiative, 2012) and Schema.org

(Schema.org, n.d.). A this point in the process, the indication for the path in the local server was discarded. The way local references work inside Revit (the relative path), in fact, loses its meaning here. The description for each of this concept can be easily assessed online at the location (or "namespace") where the data structure exists (e.g. "http://purl.org/dc/elements/1.1/" for the Dublin Core.)

A brand-new namespace was also created for the experiment, the fictional namespace dedicated to the project: "http://sangiov.org". The concept referring to this namespace is the concept of the HBIMVault (a concept referring to an HBIM object describing a vault).

When running the script the translation happens. As the unique identifier used so far (the assembly name derived from the building toponymy) becomes a label, a new counter appears, for each vault; the counter indicates each vaulted item and belongs to the <http://sangiov.org/> namespace as well. The online path to the reference geometry becomes a fully functioning URL directing to the online folder. "Creator" (DC.creator) and "ModTechnique" (DC.description) because of their textual nature, become literal strings.

```

1 @prefix dc: <http://purl.org/dc/elements/1.1/> .
2 @prefix sdo: <http://www.w3.org/2009/01/rdf-schema#> .
3 @prefix sangiov: <http://sangiov.org/> .
4 @prefix schema: <http://schema.org/> .
5
6
7
8
9
100 sangiov:} a sangiov:HBIMVault ;
101 rdfs:label: "VaultRefModWord" ;
102 dcterms: "examples" ;
103 dc:description "A wireframe CAD surface is necessary: the CAD surface is modelled in Rhino v.6,
command Patch; the polylines making the wireframe structure are 62 mm VS section at the northern
boundary, another 12 section at the southern boundary, one X2 section at the eastern boundary another
X2 at the western boundary; one X2 section in correspondence of the joint between the vault and the
eastern wall another X2 section in correspondence of the joint between the vault and the western
eastern wall; the wireframe creates a three-parted .3dm geometry, import as cad inside a revit local
mass. Apply wall-by-surface command onto the three parts of the surface."@en ;
104 nameOfLocal: "Report\\Vault\\VaultModWord" ;
105 schema:
106 <http://polina.95-mv.sharepoint.com/_layouts/personal/10385118_polina_it/Email_n2onP0N01330@rock-e-nsl-
8fc7D0e6e33d8b0e401?> .

```

**Figure 9:** part of the RDF dataset, with the namespaces on top and the information associated to a vault at the bottom.

### 3. CONCLUSIONS AND FUTURE WORKS

This experiment recounts the experience of the semantic enrichment of the dataset coming from an HBIM model. The dataset enriched through the use of Semantic Web Technologies is describing some of the actions happening in the background during the definition of an HBIM model: the procedures that create the complex 3D geometries typical of historical architecture. This kind of information does not usually live inside the HBIM model, and even less likely is subject of semantic enrichment.

Explanation and incorporation of the modelling choices coming from the HBIM process can have several positive outcomes. In the immediate context of HBIM environment, it can help to interpret and use the model when exchanged with the other professionals working on the model, and serve as an instructional tool for other case studies, where similar issues should arise. In the larger context of enrichment with Semantic Web Technologies Create a knowledge base on which to derive decision for the future on similar projects.

Evidently, this experience takes on full meaning within a framework where semantic enrichment becomes a standard (a possible future scenario) in a context larger than that of this experiment. But, meanwhile (as with any experience of digitization) the very act of creating a digital version of something precious or volatile is an act that has value in itself, for it preserves something that might otherwise get lost.

The RDF dataset obtained in the context of this research is stored and kept in the closed context of a working research group, with limited interoperability potential. To expand the potential of the information collected, the following step would be to experiment

the consolidated interoperability standard IFC for the export of the vaults datasets. The use of IFC (and then ifcOWL ontology) under similar conditions, however, raises many issues which are worth discussing in separate dedicated research. Although an alternative strategy is presented here, the author is very aware of the advantage to use an open and diffuse format for direct exchange with the experts as well as for internal storage. IFC.

Simplifying the information regarding modelling expertise in a single field served the purpose of the experiment. Nevertheless, future works should cover how to detail this information and give it more structure. Some final remarks on the accessibility and property of data, and the issue of proprietary format are necessary.

#### ACKNOWLEDGMENTS

I want to express gratitude for the assistance provided in modelling the church of San Giovanni Evangelist, by thesis student Giuliano Zanini and collaborator Barbara Fazio.

#### REFERENCES

- Andriasyan, M., Moyano, J., Nieto-Julián, J. E., Antón, D., 2020. From Point Cloud Data to Building Information Modelling: An Automatic Parametric Workflow for Heritage. *Remote Sensing* 2020, Vol. 12, Page 1094, 12(7), 1094. <https://doi.org/10.3390/RS12071094>
- Banfi, F., 2017. BIM orientation: Grades of generation and information for different type of analysis and management process. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2W5), 57–64. <https://doi.org/10.5194/isprs-archives-XLII-2-W5-57-2017>
- Banfi, F., 2019. HBIM GENERATION: EXTENDING GEOMETRIC PRIMITIVES and BIM MODELLING TOOLS for HERITAGE STRUCTURES and COMPLEX VAULTED SYSTEMS. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2/W15), 139–148. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-139-2019>
- Bassier, M., Bonduel, M., Derdaele, J., Vergauwen, M., 2019. Towards the Semantic Enrichment of Existing Online 3D Building Geometry to Publish Linked Building Data. In *Communications in Computer and Information Science* (Vol. 1029). Springer International Publishing. [https://doi.org/10.1007/978-3-030-21395-4\\_10](https://doi.org/10.1007/978-3-030-21395-4_10)
- Berners-Lee, T., Hendler, J., Lassila, O., 2002. The semantic web: a new form of web content that is meaningful to computers will unleash a new revolution of possibilities. *Scientific American*, 5(1).
- Bonduel, M., Oraskari, J., Pauwels, P., Vergauwen, M., Klein, R., 2018. The IFC to linked building data converter - Current status. *CEUR Workshop Proceedings*, 2159(August), 34–43.
- Cursi, S., Martinelli, L., Paraciani, N., Calcerano, F., Gigliarelli, E., 2022. Linking external knowledge to heritage BIM. *Automation in Construction*, 141(June). <https://doi.org/10.1016/j.autcon.2022.104444>
- Dublin Core Metadata Initiative. 2012. Dublin Core Metadata Element Set, Version 1.1: Reference Description. Retrieved from <http://dublincore.org/documents/dces/>
- IFC BuildingSMART. Retrieved from <https://www.buildingsmart.org/standards/standard-summaries/ifc-summary/>
- Moyano, J., Carreño, E., Nieto-Julián, J. E., Gil-Arizona, I., Bruno, S., 2022. Systematic approach to generate Historical Building Information Modelling (HBIM) in architectural restoration project. *Automation in Construction*, 143(March). <https://doi.org/10.1016/j.autcon.2022.104551>
- Nieto-Julián, J. E., Antón, D., Moyano, J. J., 2019. Implementation and Management of Structural Deformations into Historic Building Information Models. *International Journal of Architectural Heritage*, 14(9), 1384–1397. <https://doi.org/10.1080/15583058.2019.1610523>
- Python Software Foundation. 2022. Python (version 3.11). Retrieved from <https://www.python.org/>
- RDFLib. (n.d.). RDFLib: A Python Library for Working with RDF. Retrieved from <https://rdflib.readthedocs.io/>
- Schema.org. (n.d.), 2023, from <https://schema.org/>
- Touré, V., Krauss, P., Gnodtke, K., Buchhorn, J., Unni, D., Horki, P., Raisaro, J. L., Kalt, K., Teixeira, D., Crameri, K., Österle, S., 2023. FAIRification of health-related data using semantic web technologies in the Swiss Personalized Health Network. *Scientific Data*, 10(1), 1–11. <https://doi.org/10.1038/s41597-023-02028-y>
- World Wide Web Consortium. (n.d.). Home page. Retrieved April 30, 2023, from <https://www.w3.org/>