Planning a digitally integrated future: GIS and BIM for roman architecture, the case of Pompeii Archaeological Park

Angrisani Giovanni[1], Borriello Giovanni[2], Bosco Angela*[2], Cera Valeria[1] Forte Francesca[2], Fregonese Luigi[3], Rosignoli Olga[3], Scandurra Simona[1]

[1] Università degli Studi di Napoli Federico II – Naples, Italy
[2] Università di Napoli L'Orientale – Naples, Italy
[3] Politecnico di Milano – Milan, Italy

*Angela Bosco
Correspondence: abosco@unior.it

ABSTRACT

The ongoing wave of digitization across scientific domains is reshaping the fields of cultural heritage and archaeology. Nevertheless, as demonstrated by previous research, the exceptional nature of heritage historic assets often conflicts with software architectures, such as data schemas and communication protocols, that are not conceived with those specific requirements in mind. The archaeological remains preserved within the Pompeii Archaeological Park constitute an extraordinary case study, where architectural features have been preserved almost in their entirety. Given this configuration, any digital management tool designed to support the site's conservation must integrate the functionalities of both Building Information Modelling (BIM) and Geographic Information Systems (GIS), ensuring their seamless interoperability.

This study offers a quantitative evaluation of this interoperability challenge-frequently mentioned in articles but not demonstrated yet-by identifying the precise information loss that occurs during the digitization process and the following alignment with established data standards.

Specifically, a semantic mapping between the entities involved in the process is presented: the open-format standards IFC (for BIM) and CityGML (for GIS) are compared with a purpose-built taxonomy developed for Vesuvian Roman architecture.

The resulting correspondence matrix indicates the conceptual relationships across schemas, assessing their instrumental relevance for archaeologists, conservators, and site managers.

To further enrich the semantic representation of archaeological data, the research explored the possibilities given by ontologies and thesauri. This approach allowed for the definition of domain-specific concepts and relationships that are not adequately captured by existing BIM and GIS standards, thereby aligning the data more closely with archaeological practices.

All this constitutes a foundational step toward the development of an integrated digital system that bridges the BIM, GIS and ontological domains, while remaining responsive to the methodological principles and operational needs articulated by archaeological experts.

Keywords: BIG_SMAART, BIM, GIS, semantics, Pompeian architecture

Introduction

This contribution is part of the broader context of the digital framework for the digitisation of archaeological assets. Nowadays, there are many studies that highlight the exceptional nature of historical heritage and how it clashes with software, schemas, or data protocols that have not been specifically constructed with archaeology's needs in mind (Bosco et al., 2021; Bosco et al., 2019; Garagnani et al., 2020).

During the last decades -as well as in the past- archaeology has looked tirelessly for creative and satisfying solutions to represent both the complex status quo and the development of archaeological sites. Its complicated long relationship with digital representation tools is part of this on-going research, one that started several years ago, mainly with GIS (Geographic Information Systems) representation. Then, with the popularity of BIM increasing in the sector, the demand for its integrations to GIS drastically grew, and the first experiments in this regard started almost simultaneously (Delpozzo & Balletti, 2023)

Within this context, the aim of this investigation is to define a theoretical structure necessary for the organisation of the heterogeneous data that contribute to the study of archaeological evidence, identifying in the relevant available schemas the points of connection between representation systems such as BIM, GIS and the archaeological domain.

The BIG_SMAART project, funded by the Italian Ministry of Universities and Research, involves the co-operation of three universities: the University of Naples L'Orientale, the University Federico II of Naples and the Politecnico di Milano.

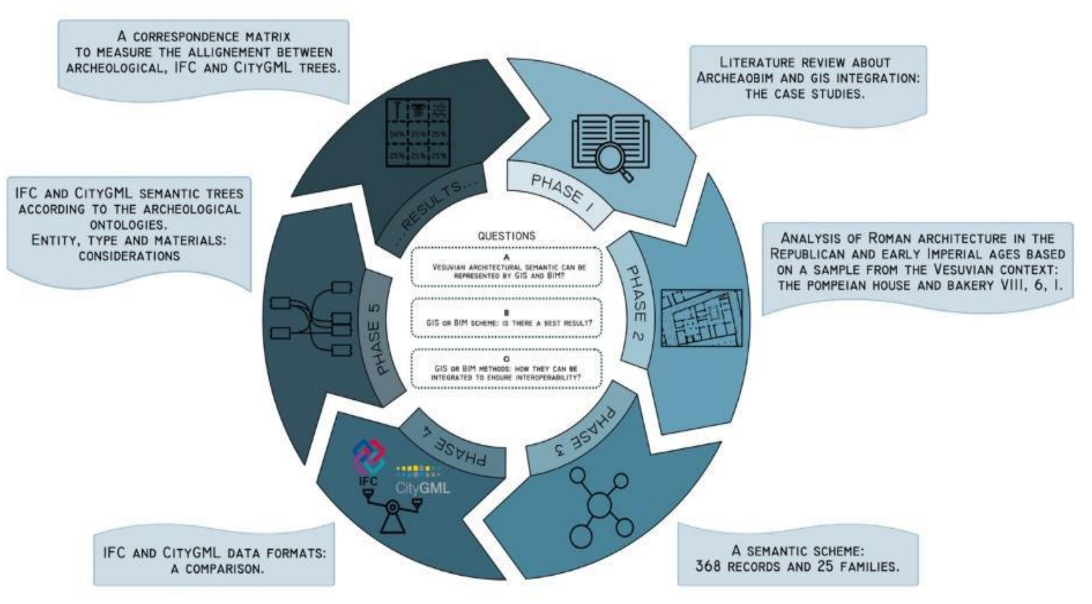

Figure 1 - Project workflow with division of the various phases

The aim of this project is to provide a quantitative assessment of the interoperability between these systems. Although the topic is widely known in the literature, it is still rare to find research demonstrating what information is lost during digitisation processes and the compliance between the most widely used data schemas in BIM and GIS.

Considering these assumptions, to define the specific workflow (Fig. 1), it was decided to define three main research questions:

(a) do current GIS and BIM data schemas or existing ontologies dedicated to archaeology meet the requirements for the representation of ancient architecture and especially Vesuvian architecture?

(b) does a data schema support these requirements better than its equivalent?

(c) how can BIM and GIS methods be integrated in an archaeological context?

The research aims to address these questions through the analysis of a concrete case study in which data formats have been applied and evaluated.

To this purpose, the project's methodology is based on the analysis of how the semantic data provided by these systems overlap with the archaeological needs. Since the archaeological world is so wide-ranging, the experimentation focuses on the Vesuvian area, which is unique because it was crystallised in 79 AD. Thanks to this condition, the context reveals extensive information about life and technology in the Roman world of the Republican and early Imperial periods.

More specifically, we propose a semantic mapping between the entities that are part of the process, starting from an analysis of the most widely used open format standards for BIM and GIS, namely IFC and CityGML. We then proceeded to compare the semantic possibilities offered by these formats with the semantic tree created specifically for Roman architecture in the Vesuvius area.

Digital Data Management in the Pompeian context

The archaeological analysis of evidence and construction techniques was extended to allow the creation of a semantic tree (a schematic hierarchical model) useful for breaking down the single elements of the complex archaeological reality of Pompeii. The choice of the site is linked to various reasons, which can be summarised in the following points:

- Extensive, well-preserved archaeological evidence, on which to test as many cases as possible;
- Evidence for which there is a good level of previous documentation, even when it concerns excavations carried out in past centuries;
- A complex reality to manage, characterised by modern infrastructure that adapts to ancient evidence;
- The volume of evidence requires an informatic system capable of providing effective management support.

Over time, the Park has implemented, through the Great Pompeii Project, research actions focused on the collection and systematisation of data with a precise view to preventive maintenance. In particular, the so-called “Knowledge Plan”[1] has led to the creation of SiPompei, a georeferenced and relational platform that describes and catalogues the entire ancient city, joined by Open Pompeii (https://open.pompeiisites.org/index.html), a digital archive with free access aimed at facilitating the dissemination of data that is integrated with SIAV (information system of the Vesuvian area) and Tolomeo for archival materials and historical documentation.

[1] The data from the GPP_Knowledge Plan has been kindly provided by the Archaeological Park of Pompeii.

The Knowledge Plan has also produced new 3D survey documentation with laser-scanning and photogrammetric methodologies covering the entire surface of the Park as of 2015. This three-dimensional documentation, however, is unrelated to the systems mentioned above.

It therefore follows that the conservation and management requirements for such a complex site would necessitate the combined use of BIM and GIS functionalities.

Based on these assumptions, the initial phase of the research involved an in-depth review of the state of the art regarding the integration of BIM and GIS applied to archaeology. The analysis of the literature made it possible to identify the data formats most frequently attested in previous studies, the software most adopted in this field of investigation, as well as to define the main problems met by other research groups with the aim of setting up an efficient mapping between the entities belonging to these knowledge structures we are aiming to communicate.

The choice of the case study

To verify the theoretical model defined at the outset of the project, an application case was identified and selected.

The choice of the House - Bakery VIII, 6, 1-9/11 (Fig. 2), is linked to several factors: the location of the complex between two public areas, the Civil forum and the Triangular forum; a complex stratigraphic sequence with various modifications and restorations, including modern ones; the type of building structure that was transformed from a private house into a productive activity with the addition of peculiar installations such as the oven and millstones.

The archaeological complex was already discovered at the end of the 19th century, but it is still almost unstudied[2]. Therefore, a census and cataloguing of the architectural and functional elements of the complex, as well as the building techniques and materials used for their construction, was carried out (Borriello et al., 2025).

Although not widely known to the general public, the Domus — in addition to occupying a strategic position within the topographical framework of the ancient city — is perfectly suited to the experimental needs of the project due to its structural complexity and its multi-layered chronological and functional stratification.

[2] The first excavations took place between 1881 and 1882.

Figure 2 - Photo of the oven installed in place of the cubiculum east in the atrium of the House - Bakery VIII, 6, 1-9/11 (Courtesy of the Archaeological Park of Pompeii)

The semantic breakdown of Vesuvian constructions

To make the management of the evidence as suitable as possible for its use with BIM and GIS platforms, a semantic breakdown of private construction elements in Pompeii was carried out (Fig. 3). The semantic scheme was based on the HBIM model, where entities are described as: Class, Category, Type and Instance, attempting to adapt it within the limits of archaeological requirements. The main components of the Roman building were divided into three disciplines (structural, architectural and plant engineering), based on the function of the object in the Roman building complex. Each discipline was then divided into classes or sets, i.e. elements that share the same formal and functional characteristics (wall, coin, pillar), which can be divided into subclasses or subsets, based on the variation of certain characteristics (e.g. rectangular pillar, square pillar, oval pillar). This is followed by categories, which allow the class to be broken down into constituent elements (e.g. column, which is subdivided into capital, shaft and base). Likewise, the subcategory allows the order or building technique used to be defined (e.g. *opus incertum* shaft). Finally, the type indicates the material used (e.g. limestone); the instance, on the other hand, constitutes the individual object created (Borriello et al., 2025).

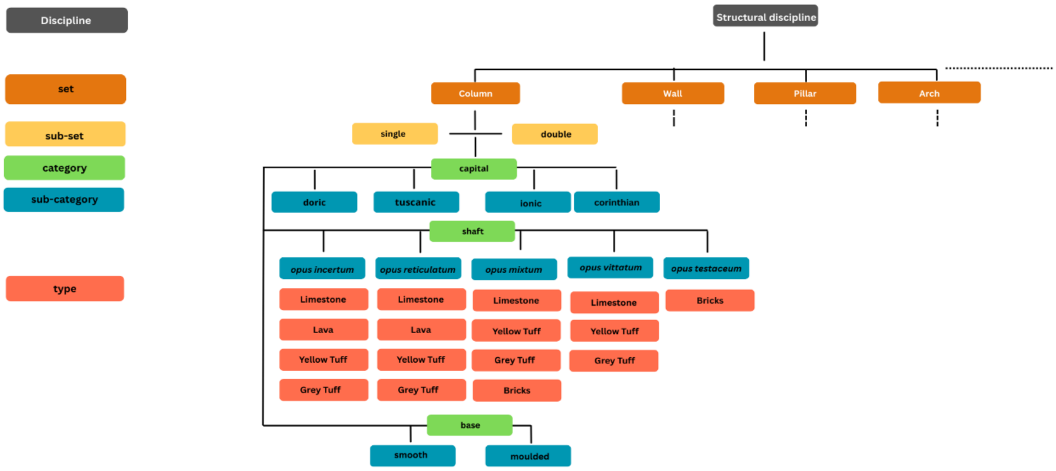

Figure 3 - Semantic structure of Pompeian architecture

In addition to being assigned an ID during modelling, each instance was accompanied by a set of data that allows for the management of archaeological information in various ways. The modelling phase and subsequent implementation of BIM objects were structured to combine, on the one hand, a broad pattern of archaeological requirements and, on the other hand, to avoid overloading the BIM system. All this required operational choices to be made (Carpentiero - D'Auria 2023): the addition of numerous families available locally would make it possible to compensate for the lack of archaeological elements within the BIM platform; however, this could lead to the possible loss of essential data during export.

Among other necessary choices, it was useful to refer to the data management system of the Archaeological Park of Pompeii for the nomenclature of the walls. This system involves dividing a single wall into two different sides, so in the information set for each instance, it was necessary to add a field that would allow the two separate sides to be joined into a single wall entity.

Another choice involved the introduction of a code to identify elements other than walls, such as floors, ceilings and cladding.

Lastly, necessary choices concerned the reconstruction of the quoins, identified as sections of masonry, and the doorways, which were modelled as empty spaces. These examples are just a few of the difficulties that are evident in the modelling phase. For this reason, it was clear that there was a need to compare the semantics of Roman architecture with IFC formats and the main ontologies of the archaeological domain.

The semantic comparison between data schemas

IFC[3] is both the schema and exchange format more widely used in the BIM domain. The hierarchical structure that arranges classes in descending order and connects them through specific relationships is actually very complex and at the same time descriptive: at the point of development it reached nowadays (with the version 4x3 of the standard) it can be argued that IFC, of the available and approved-on data schema is the more advanced in the semantic

[3] BuildingSMART Alliance 2024, Industry Foundation Classes Release 4.3 ADD2 (IFC 4.3 ADD2) (https://standards.buildingsmart.org).

description of built assets; it was put to the comparison with the created “pompeian” classes and a mapping (connection) was performed to the more semantically correspondent. The correspondence was registered both in a spreadsheet and graphically.

As an example, in Fig. 4, we position the class ifcWall, its declination in “Opera Vittata” and the materials that make it up and determine the type. The various wall “instances” (entities in the project) will inherit these characteristics.

In a comparison work of complete mapping of entities, a missing percentage relationship was estimated for several elements such as quoin, arch or threshold, for which there are no IFC classes.

Going into the specifics of the class “Wall”, the declination in the typology standard was also analysed. As can be seen graphically from the mapping connections in Fig. 5, the “Archaeo-Wall” class (on the left) corresponds to the ifcWall class of the 4x3 standard (on the right), the latest official version. The typologies, on the other hand, flow into an ifcWallType unique class where the various specifications due to building works and materials can subsequently be created inside the project. Optionally, the various instances can take on an additional Enumerative Type which, however, does not help to better specify the semantics at the level required by the archaeological needs.

The same mapping process was carried out with CityGML , the de-facto semantic standard in the geospatial sciences. Designed for representations at the urban scale, which includes morphological data for built and natural environments, similarly to IFC has a dually structured nature (with geometry and information combined) but it is operates in the Geography Markup Language (GML).

Due to the number of classes (named “features”) available, CityGML is subdivided in interconnected modules, that connect through the same hierarchy. This standard is flatter and less descriptive than IFC, and customisation mechanisms have to be employed that deviate from the predefined classes to reach good. The definition of types is less complete than the IFC because it must be reproduced with a customised mechanism called “CodeList”. Using the example illustrated before, specifically for the class “WALL”, this flows into the “Building Constructive Element”, a very generic concept, while the types of flow into the “CodeList”, which is not globally approved (Rosignoli et al., 2025).

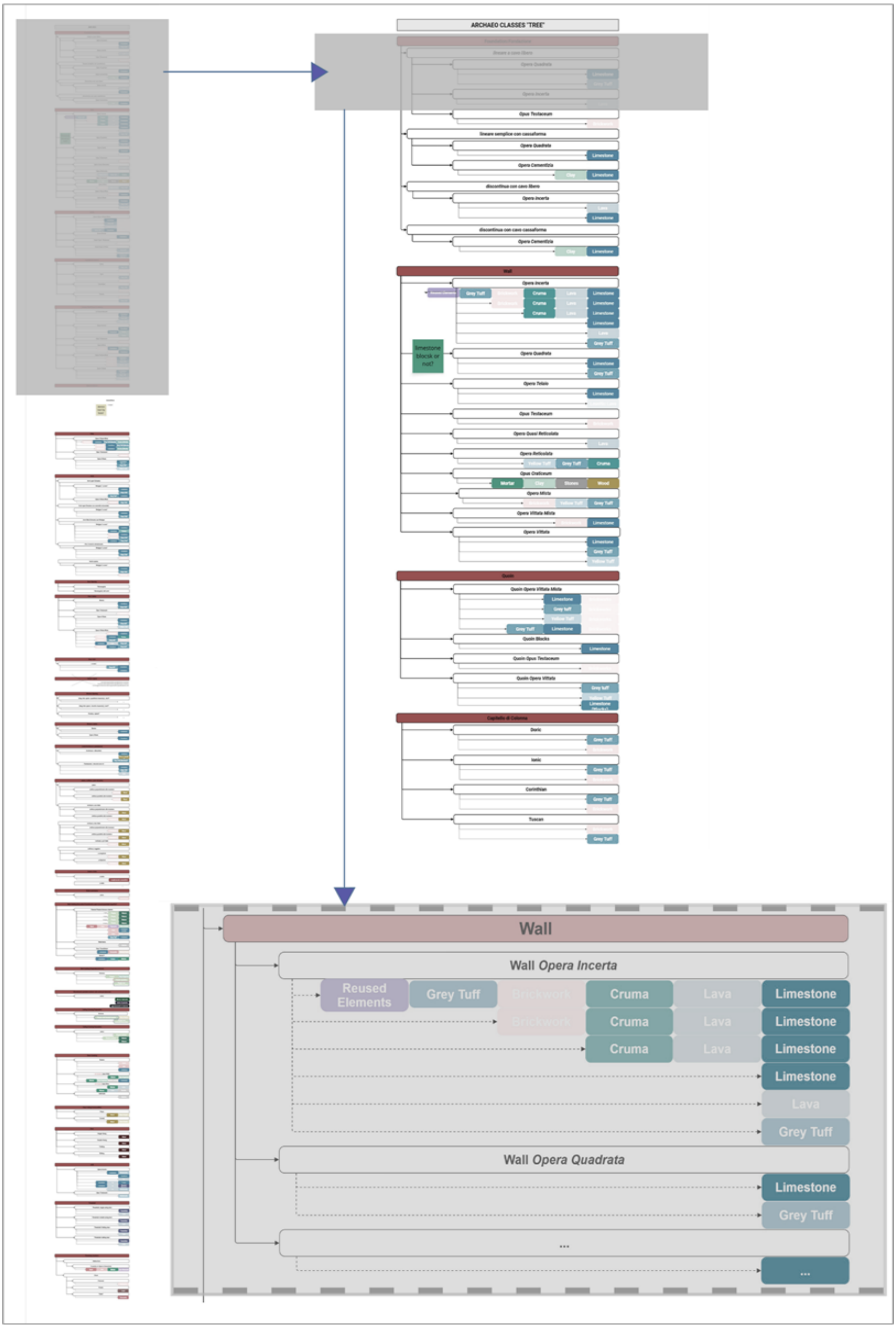

Figure 4 - the complete archaeological classes "tree" on the left and a zoom-in on the entity wall in "opera incerta"

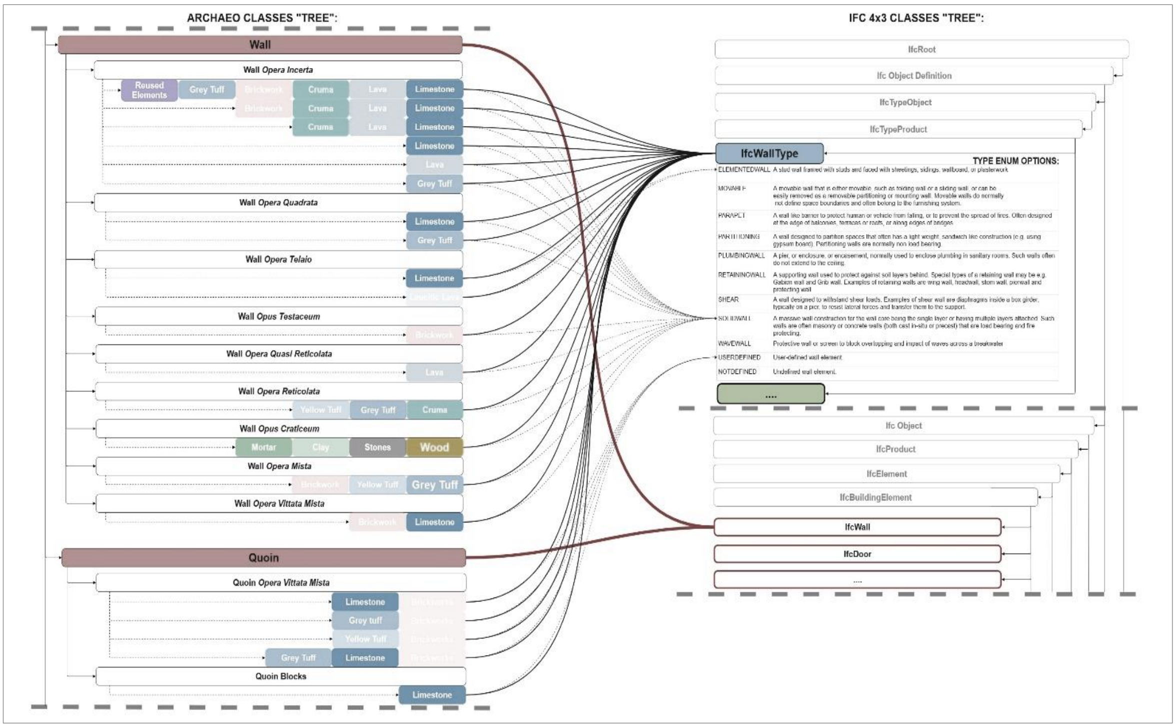

Figure 5 - in the image, an extract from the visual mapping performed between the classes of the ex-novo hierarchical "tree" constructed for the archaeological entities and the partial hierarchical schema of IFC 4X3. The extract focuses on wall entities. It is clear from the number of arrows that often converge in the same direction towards IFC classes that there is a loss of semantic granularity.The complete archaeological classes "tree" on the left and a zoom-in on the entity wall in "opera incerta"

The possibilities provided by ontologies: an overview

The need to align this semantic decomposition with the most popular standard formats required the integration of data with additional systems capable of supporting this type of description. This led to a reflection on conceptual representation tools and semantic interoperability closer to the archaeological domain, directing an analysis of what is already available in the current landscape and what alternatives they presented for the description of the entities of the archaeological semantic tree. Evidence of this can be found in ontologies, i.e. conceptual models that provide a formal structure to describe information in a systematic and interoperable way. The system of ontologies has found wide application in cultural heritage as it provides a common language for the description of the asset, allows a standardization of terms and concepts to ensure a uniform description, favours interoperability between systems and supports more advanced and precise research, since it enables data to be queried based on concepts and their relationships, not only on textual terms. In addition, these systems allow for providing adequate support in the identification of classes and properties to be included in the semantic tree of the archaeological building, where the conceptual structure offered by the BIM and GIS systems does not sufficiently meet the needs of archaeological documentation.

In particular, the research involved the analysis of international and national ontologies and thesauri:

- On the international scene, the best-known and most widely used domain ontology is the CIDOC-CRM[4], an ISO standard that provides the basic classes and relationships designed for the world of cultural heritage. It has a "core" version and a series of modular extensions, designed to support different types of specialized research and documentation queries, which can be integrated to further enrich the descriptions; of particular interest for the aims of this research is the CRMba (Buildings Archaeology) extension[5], whose semantic model allows to identify the structure of which an archaeological building is composed and the relationships between the parts, supporting the process of recording evidence and material discontinuities on archaeological buildings. Very interesting is the possibility of defining empty spaces created with intention (e.g., a door or a window), giving us the possibility to distinguish it from a structural collapse (Ronzino, 2016). Another extension of particular importance for this project is represented by CRMarchaeo[6], a semantic model developed to support the archaeological excavation process and the activities related to it. The richness of the stratigraphic relationships provided by CRMarchaeo has proven to be particularly effective in capturing and representing the complex network of connections between the various elements of the archaeological semantic tree. Furthermore, the possibility of harmonizing the CRMba and CRMarcheo extensions enables the integration of complementary perspectives and terminologies, thereby enhancing the overall semantic depth and expressiveness of the resulting model (Ronzino, 2017).
- ArCo[7], on the other hand, is the Knowledge Graph of the Italian Cultural Heritage, developed to provide a comprehensive and interoperable semantic infrastructure for the representation, integration, and dissemination of cultural heritage data. The system is structured as a network of thirteen interrelated ontologies, each addressing specific conceptual aspects of the cultural heritage domain and data from the General Catalogue of the Italian Ministry of Cultural Heritage and Activities (MiBAC), published as RDF. This articulated network builds upon and extends well-established conceptual frameworks, indirectly reusing and aligning classes and properties that have been validated in other reference ontologies such as DOLCE, CIDOC CRM, Europeana Data Model (EDM), and BIBFRAME, among others. Such reuse ensures semantic coherence and interoperability within the wider ecosystem of cultural heritage knowledge representation.

Among the various ArCo modules, several are of particular interest for the project: in addition to the *core* ontology, two extensions stand out for their relevance to the description of built heritage, that is the ‘Construction Description’ ontology which introduces a specific semantic framework for representing the structural and material aspects of architectural artefacts, including information about construction techniques, building components, and phases of development; complementarily the ‘Denotative Description’ ontology provides mechanisms for the refinement and contextualization of descriptive attributes, enabling a more detailed and expressive representation of cultural entities. Taken together, these modules significantly enhance the semantic richness and expressive capacity of ArCo, making it a robust and versatile framework for representing archaeological building heritage within a Linked Open Data environment.

The need to rely on multiple ontologies arose from the fact that each of them provides specific classes and properties suited to address the different features identified within the archaeological semantic tree. Nevertheless, in many cases, particularly concerning the highly specific subcategories, a fully satisfactory correspondence could not be found among the examined ontologies. This limitation

[4] https://cidoc-crm.org/ the core module and the extensions were last accessed on May 2025
[5] https://cidoc-crm.org/crmba
[6] https://cidoc-crm.org/crmarchaeo
[7] http://wit.istc.cnr.it/arco, last accessed May 2025

highlighted the inherent difficulty of representing the full complexity and heterogeneity of the archaeological record within a single conceptual framework.

To achieve a more appropriate level of semantic granularity, the research was therefore extended to include the use of thesauri, which are controlled vocabularies that enable a coherent and standardized organization of information related to cultural resources. By integrating thesauri, it becomes possible to refine and enrich the descriptive capacity of the model, providing additional terms and hierarchical relations that can complement the formal logic of the ontologies.

Moreover, the use of thesauri facilitates semantic alignment and interoperability across different systems and datasets. Through their publication in SKOS[8] (Simple Knowledge Organization System), thesauri can be seamlessly incorporated into ontological environments, allowing for the linking of concepts, the reconciliation of terminological variations, and the enhancement of data discoverability. In this way, the combined use of ontologies and thesauri contributes to building a more flexible, expressive, and interoperable semantic ecosystem for the representation of archaeological knowledge.

The thesaurus that best meets our needs was the AAT, i.e *the Art and Architecture Thesaurus*, developed by the Getty Research Institute[9]. It is the thesaurus that offers the greatest specificity for description, following a controlled structure of elements from ancient architecture. In the AAT thesaurus, it was possible to find a very detailed semantics of specific elements such as threshold, sacrarium, aedicule, and impluvium.

In some cases, the correspondence was as detailed as that provided by IFC format; an example could be the case of the class “wall”: for instance, in ArCo’s hierarchy, the most detailed class is similarly represented by “wall”. Instead, in the CIDOC ontological network, the description of architectural building elements stops at the class B3 ‘Filled Morphological Building Section’ provided by CRMba extension. To further specify the concept, the CRMba has to be connected to the thesaurus entry through a property. In this way, the concept taken from the thesaurus can be treated as a valid class for comparison mapping. Similarly, the thesaurus provides terms to define materials, which can likewise be used as classes within the model.

Thesauri, by offering a wide range of lexical entries, have proven essential in instances where the correspondence with the IFC format was inadequate and where existing domain ontologies lacked the granularity required for full semantic alignment with archaeological needs (Fig.6). This issue emerged, for example, in relation to architectural elements such as *cantonali* (quoins) and specific material like the yellow tuff, where precise correspondence was found in the AAT Thesaurus.

[8] https://www.w3.org/TR/skos-reference/
[9] https://www.getty.edu/research/tools/vocabularies/aat/

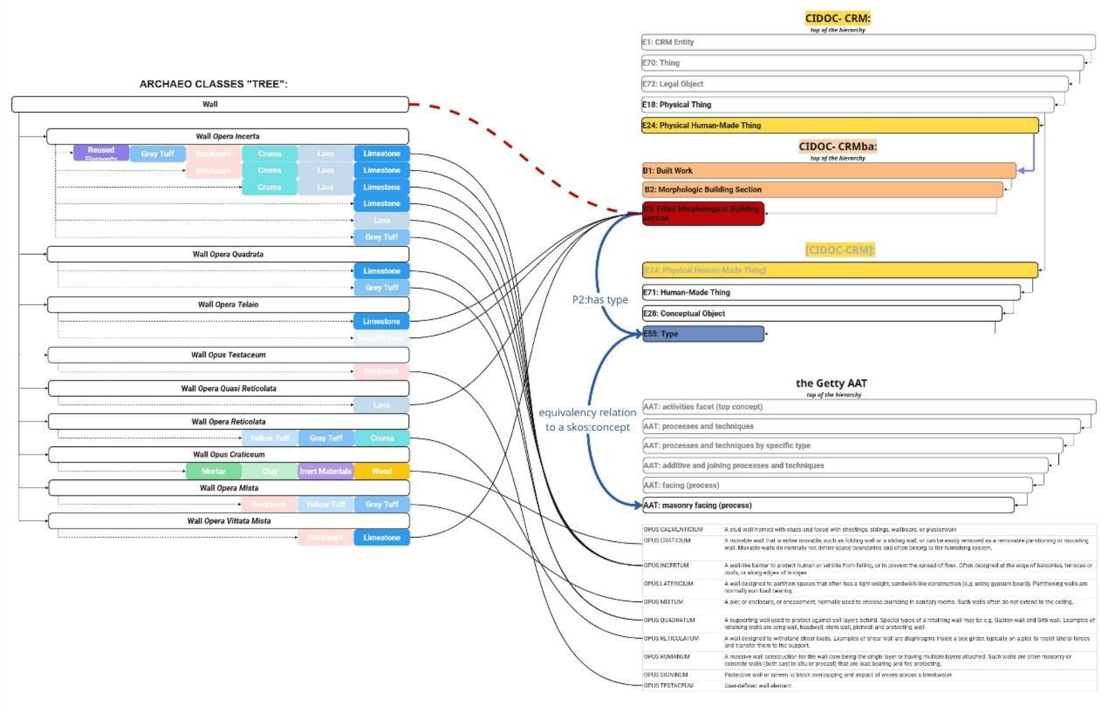

Fig.6: the configuration of the mapping between the archaeo-classes and the CIDOC-CRM ontology combined with Getty AAT; in the example, wall masonry technologies.

Conclusions

Starting from the semantic decomposition of the Vesuvian architecture, all elements were compared with the most used formats in the field of HBIM, GIS and Ontologies. Findings from the case study demonstrate that current specialized formats remain insufficient to fully represent the material and structural complexity of archaeological sites, indicating the necessity for further optimization or the creation of new interoperable standards.

The Big_SMAART project therefore defined the differences in the representation of architectural entities and materials, which are fundamental for archaeologists for both systems. A specific semantic mapping for Vesuvian architecture has been developed and it has been created a correspondence matrix between specific archaeological elements and the solutions offered by existing digital formats. In this regard, the national and international ontological models already available for the cultural heritage domain, which guarantee the formalisation of basic conceptual structures, provided fundamental support.

Future work therefore points towards the construction of better performing data schemas and ontologies for the archaeological domain. Another possible approach to be tested on the case study would be to construct a custom-made ADE (the “Application Domain Extension”, official extension mechanism of CityGML) to include all the Pompeian classes without compromises. This would hinder interoperability with other existing CityGML datasets, but at the same time provide an environment without the many semantics compromises produced by IFC adaptation or codelist mechanism. Lastly, we remember that this mapping analysis performed on semantic classes of the standards is not complete without considering a thoroughly mapping of the relationships between entities and without considering the relation of semantics to the model geometries.

Acknowledgements

We would like to thank Gabriel Zuchtriegel, Director of the Archaeological Park of Pompeii, architect Raffaele Martinelli and Dr Giuseppe Scarpati for their availability and support for this project.

Funding

This project, titled “BIG_SMAART - BIM & GIS for Spatial and Multidimensional Archaeological Artefacts and Techniques”, is financially supported by the National Recovery and Resilience Plan (PNRR), as a Research Project of National Interest (PRIN), funded by the European Union-NextGenerationEU - CUP C53D23001900001 - Grant Assignment Decree No. 961 adopted on 30/06/2023 by the Italian Ministry of University and Research (MUR).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

References

Bosco A, Carpentiero L, D’Andrea A, Minucci E, Valentini R (2021) Developing an ABIM system: a new prospective for archaeological data management. *Archeologia e Calcolatori* ***32.2***,167-176. doi: 10.19282/ac.32.2.2021.15

Bosco A, D'Andrea A, Nuzzolo M, Zanfagna P (2019) A BIM approach for the analysis of an archaeological monument. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLII-2/W9**, 165-172. https://doi.org/10.5194/isprs-archives-XLII-2-W9-165-2019

Borriello G, D’Auria D, Forte F (2025) Verso un modello semantico per l’archeologia: BIG_SMAART e il caso della casa-panificio VIII 6, 1/9-11 di Pompei. *Vesuviana*, **16**, 65-88.

Carpentiero L, D’Auria D (2023) Operative tools for BIM in archaeology: libraries of archaeological parametric IFC objects. *Proceedings of the 16th International Conference on Open Software, Hardware, Processes, Data and Formats in Archaeological Research (Archeofoss 2022). Archeologia e Calcolatori*, **34 (1)**, 69-76. https://doi.org/10.19282/ac.34.1.2023.08

Delpozzo E, Balletti C (2023) Bridging the gap: An open-source GIS+BIM system for archaeological data. The case study of Altinum, Italy. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, **48** (M-2–2023), 491–498. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-491-2023.

Garagnani S, Gaucci A, Moscati P, Gaiani M (2021) *Archaeobim. Theory, Processes and Digital Methodologies for the Lost Heritage*. Bononia University Press.

Ronzino P, Niccolucci F, Felicetti A, Doerr M (2016) CRMba: a CRM extension for the documentation of standing buildings. *International Journal on Digital Libraries Antiquity*, **17**, 71-78. https://doi.org/10.1007/s00799-015-0160-4

Ronzino P (2017) Harmonizing the CRMba and CRMarchaeo models. *International Journal on Digital Libraries Antiquity* **18**, 253–261. https://doi.org/10.1007/s00799-016-0193-3.

Rosignoli O, Fregonese L, Angrisani G, Cera V, Scandurra S, D'Auria D, Bosco A (2025) BIM and GIS integration: Planning a digital approach to archaeological site management. *Archeologia e Calcolatori*, **36(1),** 171–190. https://doi.org/10.19282/ac.36.1.2025.09.