MODELLING AND MANAGING BUILT HERITAGE KNOWLEDGE: AN ONTOLOGY-BASED APPROACH FOR MULTI-LAYERED ARCHAEOLOGIES AND HISTORICAL PRODUCTION PROCESS REPRESENTATION

Cassia De Lian Cui, Antonio Fioravanti & Edoardo Currà
Sapienza University of Rome

Davide Simeone
University of Brescia

Stefano Cursi
ISPC Institute of Heritage Science, National Research Council, Montelibretti (RM), 00015, Italy

ABSTRACT: *Classical and industrial archaeologies are a complex cultural field where singularity and uniqueness are expressed through past memory evidence and identity recognition. In order to obtain these values acknowledgement, it is compelling to highlight the material and intangible knowledge by using suitable ICT tools capable of handling complexity and managing large sets of heterogeneous data usually subjected to changes, different interpretations, inconsistencies and sometimes uncertainty. Although the HBIM method has been largely used in the past years, it shows significant limits when dealing with large and heterogeneous information requiring the introduction of advanced methods and tools. In this context, this study presents an approach to the architectural heritage and historical manufacturing activity representation based on integrating the HBIM process with a structured knowledge base, demonstrated through its application to the Sanctuary of Hercules and the former Segrè Papermill case study. The work develops an ontology-based system using existing ontologies for the three domains of interest: architectural artefact, cultural heritage and industrial processes directly connected with the informative model. The intent is to give an overall support system for the complex semantics formalization of these assets to aid the interpretation, intervention and valorization activities.*

KEYWORDS: *Archaeology, Industrial heritage, Knowledge representation, HBIM, Ontologies, Linked data*

1. INTRODUCTION

The widespread of various digital approaches in the built heritage field has raised multiple issues related to the effectiveness of those different methodologies. In this context, it is compelling to comprehend the complex nature of the these specific sites and to define reliable methods to document research and investigation processes for the recovery, intervention and valorization activities. Particular assets distinguished by multiple archaeological stratifications are emblematic cases that perfectly outline the critical aspects of current practices. The knowledge representation and management activities play a crucial role alongside the historical and archival research, the integrated survey and the information modelling of these artefacts, to address the challenges of complex heritage assets related to their uniqueness and singularity/rarity and to solve critical issues of interoperability, alignment of cataloguing systems, in order to make knowledge shareable among the various actors involved in decision-making processes.

Within such peculiar applications, this study tries to apply effective digital knowledge technologies to address the main research question, or rather the knowledge representation in a complex evolutionary scenarios, such as multiple archaeological sites, by integrating information modelling (BIM) and semantic web technologies (ontologies). In particular, with a specific focus on modelling the Segrè papermill historical industrial process, linked to the architectural artefact and its historical evolution. The aim is to highlight comprehensively the industrial archaeology aspects which is defined as an interdisciplinary study field related to both the historical issues of the industry's world and the material culture (processes, machines, workers life). In fact, the industrial archaeology represents a combined form of technological and humanistic culture, whose value extends and matures overtime. Therefore, the experiment focuses on covering the gap in digital data documentation of industrial archaeology and historical manufacturing processes to represent multidisciplinary concepts correlations in a machine-readable way. Furthermore, the study proposes an approach to the knowledge representation and global management for the investigation process of this unique heritage field.

Referee List (DOI: 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)
Cassia De Lian Cui, Antonio Fioravanti, Edoardo Currà, Davide Simeone, Stefano Cursi, *Modelling and Managing Built Heritage Knowledge: An Ontology-Based Approach for Multi-Layered Archaeologies and Historical Production Process Representation,* pp. 768-780, © 2023 Author(s), CC BY NC 4.0, DOI 10.36253/979-12-215-0289-3.77

1.1 Current digital documentation and investigation processes for built heritage

Over the past decade, we have observed a growing focus on Building Information Modeling in the area of heritage architecture also, referred to as Heritage (or Historical) BIM / HBIM. Currently, the most established HBIM workflow consists of surveying the building by point cloud used as a reference to reconstruct simplified geometries employing BIM objects, which are assigned attributes and linked documents, boards and databases helpful in enriching the model with information produced and used by specialists in the field (Logothetis et al., 2015; López et al., 2018; Pocobelli et al., 2018).

If, in the case of the design of new buildings, the definition of accurate and complete documentation of what is to be built goes hand in hand with the progressiveness of the definition of the project itself, and is, therefore, fully consistent with the information practices imposed by BIM environments, in the case of the processes that characterize the activities of investigation and restitution of an existing asset the knowledge we have of the latter, in addition to coming from extremely heterogeneous sources, is subject to continuous changes, interpretations, uncertainties and gaps that must be able to persist until the end of the process and beyond (Bianchini, 2014). In fact, information management in this field is still mainly based on a documentary approach (Moscati, 2021) in which information is stored by presenting it in a linear and orderly manner but in a flat and static form that does not allow the possibility of movement and connection between the information itself. The need to overcome these limitations led to the research of new systems that could allow machines to automatically combine knowledge from different sources and, even better, derive new knowledge from them.

Following linked data principles, ontologies are languages used by semantic web resources to represent knowledge and concepts within a specific domain (Gruber, 1993). In the case of cultural heritage and more in detail built heritage fields, ontologies have been developed to organize and structure information related to buildings, historical monuments, archaeological sites, and other forms of architectural heritage. These ontologies enable better management, search, sharing, and data interoperability in the built heritage field.

The direction in which major international research infrastructures are moving (DARIAH, Etc.) seeks to transform the 'Web of documents' into a 'Web of data', consisting of concrete 'objects' that machines can process (machine understandable). The aim is to provide computers with the ability to combine data to create new knowledge, form new connections and draw new conclusions from indexed data, automating what has hitherto been the exclusive preserve of humans and minimizing the separation between discovery (locating bibliographic news) and delivery (finding the document) so that an ecosystem of metadata can be created. This operation will soon allow global data availability within a larger framework of openness and interconnection of information, enhancing the effectiveness and visibility of information available online.

1.2 Existing ontologies for industrial processes representation

Significant efforts exist to standardize ontologies across industries. Organizations such as Industrial Ontologies Foundry (IOF) (Drobnjakovic et al., 2022) and Manufacturing Enterprise Solutions Association (MESA) (Kazil et al., 2020) are collaborating to develop common ontologies and data models to improve interoperability among manufacturing systems. To date, the mid-level manufacturing ontologies available are the Manufacturing Semantics Ontology (MASON) (Lemaignan et al., 2006) and the Manufacturing Reference ontology (MRO) (Usman et al., 2013). However, they are only able to represent the production and design part of the manufacturing domain (Sanfilippo et al., 2021) and have limited mutual interoperability (Francesconi et al., 2010).

The Supply Chain Reference Ontology (SCRO) (Ameri et al., 2020) is a pilot ontology that extends BFO and IOF Core able to provide the basic ontological constructs needed to represent a supply chain in terms of structure (members and their roles, functions, capabilities, relations, and resources) and operations (processes and flow of material and information). Similarly, the Supply Chain ONTology (SCONTO) (Vegetti et al., 2016) formally describes a supply chain at various abstraction levels. Resources such as workstations, machines, tools and fixtures are formally represented in the Manufacturing Service Description Ontology (MSDL) (Ameri & Dutta, 2006).

Furthermore, with the advent of the fourth industrial revolution (Industry 4.0) and the expansion of the Internet of Things (IoT) into the industrial sphere, ontologies become even more important for organizing and understanding data from an increasing number of connected devices and intelligent systems (Sampath Kumar et al., 2019). Industry 4.0 is mainly based on robotic agents that are responsible for performing the main operations in a smart manufacturing environment. The standardization of the knowledge representation is based on standard ontologies, which are the CORA Ontology for robotics and automation (Prestes et al., 2014) and ROA Ontology (Cheng et al., 2016) that defines the main notions of behavior, function, and goal. Focusing specifically on the manufacturing resources, in literature we can find a variety of standard and ontologies that understand resources differently. Based on industrial standard, resources are defined as " any device, tool and means, except raw material and final product components, at the disposal of the enterprise to produce good and services" (ISO 15531-1; ISO, 2004a). "The types of resources involved in the manufacturing operational management are: personnel, material, equipment" (IEC 62264; IEC, 2013); "Means used by an activity to transform input into output" (ISO 20534; ISO, 2018). As we can see the terminology relies on not aligned vocabularies and we can define three different approaches that have distinct views on manufacturing resources. The first one relates the manufacturing activities occurrences with the resources model (Sarkar & Šormaz, 2019), the second one connects the resource to the activities primarily (Sanfilippo, 2018), the last one presents the resources directly related to the agents goals (Sanfilippo, 2018). The user case needs should rely on adopting one or the other. Furthermore, as we can see in the literature, many studies and approaches exist for contemporary manufacturing processes and Industry 4.0. At the same time, it lakes in the representation of historical industrial processes, therefore in the ontology reuse process, we adapted existing classes, relations and proprieties to our specific use case.

1.3 Existing ontologies for cultural heritage representation

Several knowledge structures have been developed to represent and manage cultural heritage data (Doerr et al., 2020; Hellmund et al., 2018). For instance, the Getty Foundation developed a SPARQL version of its thesaurus (Harpring, 2010). In addition to those mentioned above, one of the most widely used conceptual models to describe cultural heritage in general with extensions possibility to fit built heritage is the CIDOC Conceptual Reference Model (CIDOC-CRM) (Martin Doerr, 2003). CIDOC-CRM provides a conceptual model to describe cultural heritage in general but can be broadened to fit built heritage aspects.

Extensions of CIDOC-CRM include the CRMba model (Ronzino, 2015) conceived to support the archaeological documentation of buildings with an emphasis on recording stratigraphic units and the evolution of the structure over time. Furthermore, other specific ontologies have been developed to address specific issues in built heritage, such as conservation, documentation, and management. For example, Acierno et al. (2017) extend the domains of CIDOC with an ontological structure that covers aspects concerning both the typological and the constructive entities of historic buildings, and the documentation and investigation activities conducted by specialists for the artefact study and preservation. Colucci et. al. (2021) deal with the formal conceptualization to represent the built and architecture domain by proposing an ontological scheme that focuses on connecting semantic and geometric information to generate a parametric and structured model from point clouds.

In a broader context, not confined exclusively to the ontologies field, initiatives such as the EUROPEANA (Europeana, 2017) and ARCHES (Myers et al., 2016) projects aim to create open digital infrastructure to enhance the semantic representation of data related to built heritage. It is clear that the development of ontologies for built heritage is an active field of research and development. New ontologies and approaches continue to be proposed and developed, especially to integrate ontological approaches with 3D and HBIM models (Cursi et al., 2022) to improve the representation and interoperability of information related to this domain.

1.4 Digital approaches for classical and industrial archaeologies

In the archaeological field, the structures and the architectural elements are usually partially visible, restored or modified, often with missing parts and multiple transformations during time. Digital representation is a big challenge since the archaeological reports on the excavations and related documentation are usually the only available data. The models tested in this field primarily deal with virtual reconstruction, starting from different survey techniques to immersive and virtual fruition. Most of the case studies considered are as-built BIM to represent the actual state of the remains with a BIM semantic structures through manual operations or automatic segmentation algorithms (Achille et al., 2015; Bosco et al., 2019; Di & Wu, 2011; Guerrero Vega & Pizzo, 2021; Moyano et al., 2021; Scianna et al., 2021; Trizio et al., 2018). In Garagnani's work (2016), they defined a process of information cataloguing ArchaeoBIM as useful for documentation and consultation and for analytical studies to accompany the reconstruction activity. The main focus of some of them is on the stratigraphic analysis in the HBIM workflows (Diara & Rinaudo, 2020) and the development of the HBIM cloud platform for archaeological analysis and documentation (Diara & Rinaudo, 2021). While in some other cases, these applications have been implemented for preventive archaeological projects, for instance in Banfi et al. study (Banfi et al., 2020) they developed an open-source BIM platform able to merge BIM sensors, monitoring, historic building information modelling and VR, with a specific focus on complex scenarios with heritage sites subjected to flood risks and water level changes. At the same time, Saricaoglu et al. (2022) defined a method for data-driven conservation actions alongside the decision-making process to integrate both levels of geometrical accuracy and the multi-level data for the interventions.

For the industrial archaeology digital process in literature, we can find, from a larger scale, GIS approaches for the creation of databases, spatial analysis and visualization (He et al., 2015), to similar approaches related to the definition of as-built models, from UAV survey to BIM environment (Barrile et al., 2019). In particular, there are some cases where the main focus is the machine apparatus inside the former factories. The integrated survey process of the main components of this architecture type is followed, in one case by an HBIM modelling with a deepening aspect of LOD (Currà et al., 2022), in one other with the creation of a virtual tour (Shults et al., 2019).

Hence, the literature review shows a clear gap in documenting industrial archaeology and the historical manufacturing processes. Furthermore, the applications could be more extensive in the data representation and relation between different domains to address the possibility of making interdisciplinary concept correlations. These gaps represent the starting point for our study which tries to delineate an approach towards industrial archaeology and other archaeologies by considering its valuable aspects, defined in the next paragraph, and through the integration with semantic web technologies to overcome the evident difficulties in the investigation process of a complex heritage site.

2. COMPLEX ARCHITECTURES OF MULTIPLE ARCHAEOLOGIES: RESEARCH, DOCUMENTATION AND INVESTIGATION PROCESSES

Complex architectural sites, defined by the stratification of multiple archaeology, are a peculiar heritage field rooted in the Lazio region. The Sanctuary of Hercules and the former Segrè Paper Mill site in Tivoli – selected case study for our research - are unique evidence of the past Roman empire that crossed centuries up to industrial time. In fact, the actual state is a combination of classical and industrial archaeology. In this case, the main focus of the study regards the semantic definition and knowledge modelling of the industrial processes of the Segrè papermill, which is the latest industrial reuse of the Sanctuary, built for some parts of the former manufacturing structures such as the ironworks and the powder magazines. This area stands mainly between the podium and the northern portico, made of iron and reinforced concrete structures. Although quite invasive, the different reuses over the centuries contributed to preserving the porticos dated II century B.C. A second expansion occurred in the former ironworks, where another paper machine expanded the production cycle (Cairoli & Ten, 2016). The plant had an automated production, and hydroelectric turbines followed later on by an electric cabin serving it. The whole complex was decommissioned in 1956, and today, it is partially reused as a museum.

The factory, machines, objects and documents are explored as part of a system that has historically, socially and economically determined the territory and has shaped the landscape. The evidential value reflects activities that had and continue to have profound historical consequences and is based on this evidence's universal value (TICCIH, 2003). Moreover, the social and cultural aspects are related to an industry, a specific company, an industrial community or a particular trade or skill. It may also carry technological and scientific value in manufacturing, engineering and construction history or have aesthetic qualities deriving from its architecture, design or planning (Douet, 2016). Therefore, documenting the industrial processes represents a way to enhance the possibility of continuously perceiving its values and to highlight, in this specific case, how the multiple stratifications and different industrial reuses have helped maintain the ruins of classical archaeology for centuries. The purpose is to give valid support in the digital processes for built heritage by leaving intact the specific features and historical evidence of different periods, actively working to enlighten these diversities during the intervention and valorization activities. It is not a simple and pure conservation process since the past transmission proceeds through its continuous re-interpretation and through a global and interdisciplinary approach.

3. KNOWLEDGE BASE FOR HISTORICAL ARCHITECTURE AND INDUSTRIAL PROCESSES

3.1 Ontology development methodology

The first step to develop the knowledge-based system is choosing the most appropriate methodology that best addresses the specific case study. As described in paragraph 2, it is necessary to define and consider representation from multiple domains, in particular, it is compelling to highlight how the whole process interacts with the existing and new building spaces, related to different evolution phases and how it is combined with the other archaeology and study fields. Some ontologies already cover parts of the domains of interest, while in other cases, integration of new representations schemas is needed to provide a full coverage of the information for a complete comprehension of the artefact.

In the ontology engineering discipline and, more recently, in its application in the AECO field, the Linked Open Terms (LOT) methodology has recently been introduced and developed from the NeOn methodology (Suárez-Figueroa et al., 2015), which places particular emphasis on the reuse of existing ontologies - both general and domain ontologies - already prevalent in the relevant industry, fostering interoperability and optimizing the definition of concepts, attributes and relationships. In a field as complex as multiple archaeology, the potential benefits of such a methodology are clear: the presence of different disciplines involved, seemingly far apart, with their terms, concepts and general knowledge structures, requires, first of all, an effort of correlation between the various domains and, secondly, the work of building the missing knowledge networks or adapting the existing ones. In the specific object of this study, for example, the reuse and integration of ontologies dedicated to the representation of construction and buildings, those dedicated to the documentation of cultural heritage, and those used to describe industrial processes and/or production facilities are important.

3.2 Ontology requirements specification

The use case and purposes concern the documentation of archaeological and industrial heritage to support the investigation and knowledge processes for the intervention and recovery of complex palimpsest. The data exchange needed for this study is based on different sources. All the information necessary for this study has been gathered from many sources starting from a historical and archival study composed of written sources and cartographic, iconographic and photographic ones from the two central archives (Segrè family and Emo Salvati). Besides these historical studies, it is essential to combine them with the archaeological ones, that is, the analysis of all the remaining on surfaces and excavation necessary to reconstruct the whole industrial process. The data acquisition through an integrated survey is then combined with the information collected through the other sources to get a complete information framework.

The *non-functional requirements*, which refer to the characteristics, qualities or general aspects, can be identified by defining a system able to manage heterogeneous data from multiple domains, to assure flexibility, adherence and coherence. Furthermore, consistency should be considered to avoid duplicate information or over-constrained property assertion. Extendibility can be achieved by using standardized languages and existing patterns and concepts for ontology development (Suárez-Figueroa et al., 2009). For scientific data management, the ontology should also be findable, accessible, interoperable and reusable.

The *functional requirements* are the ones related to the use case. The documentation of industrial processes requires the definition of the activities performed in the process stages and the different resource types used (machines, humans, raw materials, semi-finished products). A detailed description of each production stage is necessary for a comprehensive understanding of the process, the machines and the humans/workers, all valuable and fundamental aspects for documenting and investigating industrial archaeology. Along with defining the manufacturing activities, the ontology needs to determine the spaces or rooms where the activities used to be performed, in terms of architectural and technological aspects through spatial and technological entities and the relation between production process, machines and building components and spaces. Layered archaeologies are complex sites where interpretation is usually a complex activity. Therefore the representation of the evolution phases should concern the whole complex and single building elements to interconnect these three domains of interest in an interdisciplinary way by covering all the knowledge useful for the intended documentation activity.

3.3 Ontology conceptualization

From the sources listed in paragraph 3.1, the first attempt in the conceptualization work was to delineate the complete cycle highlighting classes or concepts necessary to describe the whole production system.

In 1935 the mill was expanded, presumably introducing this cycle. The raw materials used for paper making were rags and wood logs. The process based on the rags, our main focus in this case, started from the arrival at the factory in bales where they were stacked temporarily in large deposits. The processing of rag consisted of cleaning and sorting according to various qualities and cutting them in special cutting machines. Once they were selected and separated, they were cleaned a second time by tumblers and beaters, then shredded and leached under pressure with lime or soda in spherical or cylindrical kettles. These operations were followed by fraying in Hollander piles, sort of tanks filled with water where a cylinder with knives reduced the rags into filaments. After that, the produced “half pulp” was conveyed in special tanks and subjected to bleaching with calcium chloride before being fed into the Hollander refining machines. In the Hollander refiners, the half pulp was beaten by other rotating cylinders and converted to “all pulp”, ready to be transformed into paper. Thus the pulp was coloured and glued with resin soap to give texture and to prevent ink from spreading on it. In the paper machines, the pulp was stretched into a thin layer of thick water on a wire cloth, which was then detached into a veil and it was passed over a long woolen felt around a series of huge drying cylinders, heated by steam, and finally dried, it was rolled up as a raw product .

The long ribbons produced by the paper machines were directly pressed and smoothed by giant calenders and then passed to the rolling machines. In the preparation rooms, the paper was cut into sheets and reams were assembled; in the end, the sheets were sorted one by one to be packed and shipped or for other finishing works.

Space	Activity description	Input components	Activity requirement	Resource in Automatic solution	Resource in manual solution
-	1. wood defibration	wood logs	handling tooling (machines) loading	mechanical defibrators	operator, power tool
rags room	2. rags sorting and preparation	raw rags	handling tooling (machines) workholding	hedge cutter cement sack dusting shredder Lannoye	operator
half pulp processing area	3. under pressure leaching	cleaned rags lime soda	handling tooling (machines)	spherical boiler	operator
half pulp processing area/hollander beater room	4. fraying and pasting	impurity-free fibers resin soap water	handling tooling (machines)	hollander beater for fraying hollander beater for pasting	operator
half pulp processing area/hollander beater room	5. whitewashing	"half" pulp calcium chloride	handling tooling (machines)	hollander beater for whitewashing	operator
half pulp processing area/hollander beater room	6. refining	whitened pulp	handling tooling (machines)	hollander beater for refining	operator
paper machine room	7. unfold the paper	pulp	handling tooling (machines) workholding	paper machine	operator
preparation and finishing rooms	8. pressing, smoothing, cutting, packaging	raw paper	handling tooling (machines) workholding	Haubold rotary cutter winding machines packaging presses calender	operator Diamont paper cutter and cutting table Verny rectangular cutter turntable
Paper sorting room	9. paper selection	commercial paper	handling workholding	-	operator

Fig. 1: Use case data of the latest production process of the Segrè papermill.

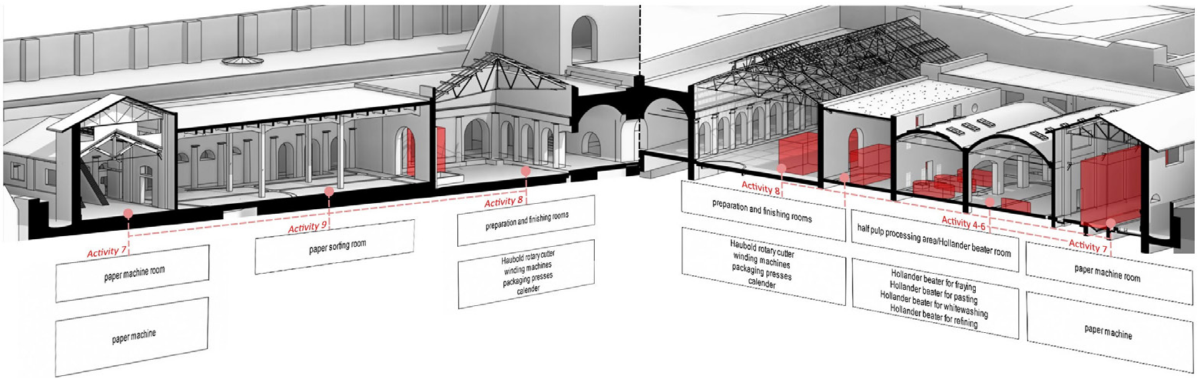

Fig. 2: Spaces, activities and machines in the latest Segrè papermill production process. BIM model carried out by Andrea De Pace and Riccardo Rocchi under the supervision of Edoardo Currà.

Based on the studies conducted by Sanfilippo et al. (2021), we detailed the definition of the process plan, the set of sequential activities decomposing the process, the input component or materials for each activity and the requirements to execute each activity. As shown in the figure, table 1 defines the critical aspects of the use case by listing the activities and their description, input components of each activity, and resources needed to execute each activity, which can be subdivided into manual and automated resources.

After that, it was necessary to identify key concepts to relate the papermill cycle with the building spaces where it used to be performed, which is represented in the first column of Figure 1. The spaces and the building components are required to describe the different halls built in overlapping periods of time. For this reason, it is fundamental to briefly sum up the significant historical stages of the investigated area and, consequently, the evolution of the construction methods related to the different phases. As we already have introduced, the latest industrial production was installed in the northern portico and sacred area of the Sanctuary of Hercules. Over the centuries, many different productions and constructions intersected. In particular, in the 1920s, the mill already had three paper machines, with a layout maintained until its dismission. Between 1923 and 1926, took place the elevation and expansion of the paper machine room in the sacred area. Thanks to the new technologies, a radical change occurred in the 1930s through the widespread use of reinforced concrete structures and prefabricated frames for large roofs.

In this period, the rag room was renovated and expanded, preceding the construction of a paper storage room and a paper sorting room on a two-level building. During the war, the roofing of the sacred area rooms suffered substantial war damage. Three wooden truss roofs were replaced with reinforced concrete structures and two barrel vaults with SAP technology. Overall, there was no real expansion but a simple consolidation of the overall preexisting layout. In Figure 2 some of the spaces/areas, activities performed and machines are represented in the informative model.

Once the main domains, concepts and classes are described, the following step of this study, detailed in the next paragraph, concerns the research and reuse of existing ontologies and the encoding activity.

3.4 Ontology reuse and encoding

The three domains considered for this case study are the *Production Process*, the *Architectural Artefact* and the *Historical Evolution* and they are all connected since the process and the manufacturing tasks are performed in specific spaces that belongs to different areas or rooms of the site which are built in various phases of the papermill production. After reviewing the existing ontologies for all the listed domains we used for the Building spaces and elements the IFC Standard, while for the production process we chose an extension of the IFC and the FA ontology that is mainly focused on the relation between the used resources and activities performed. For the historical evolution documentation we based the whole reconstruction process on the CIDOC-CRM (Fig. 3).

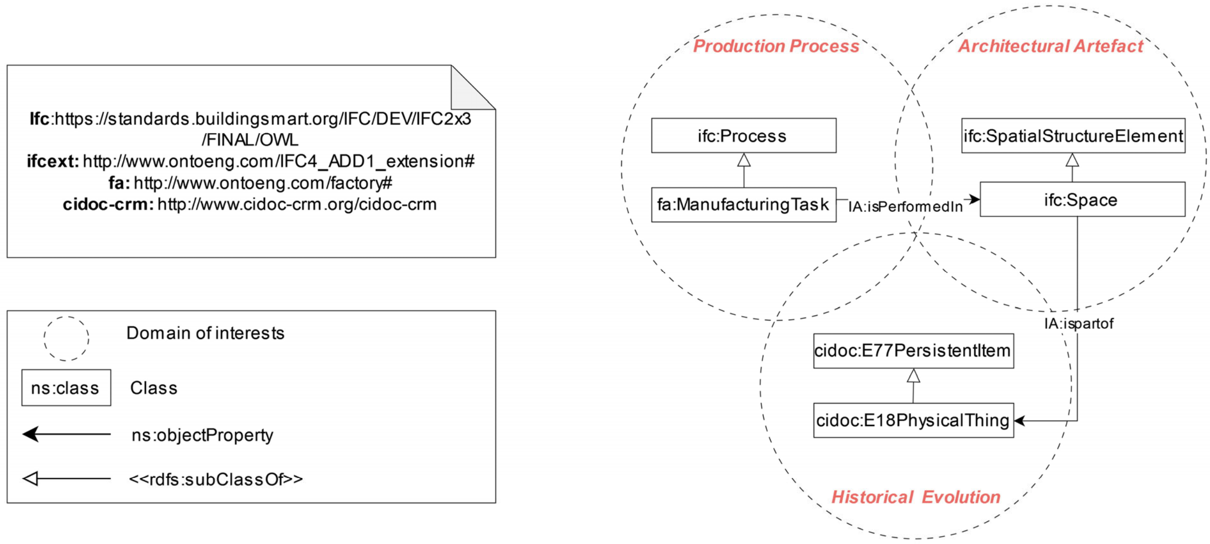

Fig. 3: Knowledge modelling of the three domains considered: the production process, the architectural artefact and the historical evolution.

The chosen ontologies for the production process modelling provide an extendible representation of the factory entities related to production systems, resources and products. The starting point was a simplified version of the IFC based on two main classes. A process (*ifc: IfcProcess*) "is an event that transforms input-output products while making use of resources under specific control rules. A resource (*ifc: IfcResource*) represents an entity that is needed to perform the process. In turn, various objects, such as physical products, people, and materials, can be employed as resources". Then we integrate this approach with the one proposed by Sanfilippo et al. (2021). They started considering three existing high-level approaches and trying to unify them in a single and complete framework. In this way, it was possible to consider and represent manufacturing by taking into account activities, goals and activities occurrences during the entire process since the IFC-based modelling is not able to distinguish alone between the entities and their relations. On the other hand, through this integrated methodology, the properties are explicitly stated, and they can provide a useful explanation of the case study. For instance, the *activity description* is a class that describes the specific manufacturing process description where *whitewashing* is an individual and refers to conveying the *half pulp* in special tanks and subjected to bleaching with calcium chloride. The *input components* for each activity list the components needed for that singular operation and they are detailed through the property *hasComponentReq*. For the activity whitewashing, the input components are *half pulp* and *calcium chloride*, and the output of the same activity, which is expressed through the propriety *hasOutput* is the *whitened half pulp* that is combined with other components and becomes the input resources of the next activity. The activity requirements list the resources and capabilities that are necessary to perform specific tasks. For instance, the whitewashing activity *hasResourcesReq tooling* and *handling* that are individuals of *CapabilityDescr*. The *Resource in Automatic solution* and the *Resource in Manual solution* list the description of resources either in the automatic or manual configuration that can execute the activity; *Whitewashing Hollander machine* and *Operator* satisfy the two resource requirements and are correspondingly individuals of *ArtifactDescr* and *Descr*.

The production process domain and the building artefact domain are connected since the *fa: Manufacturing Task*, a subclass of *Ifc: Process*, *isPerformedin* a specific place of the papermill complex. In fact, the architectural artefact knowledge modelling started from the class *Ifc:IfcSpace* which is defined as "an area or volume bounded actually or theoretically. Spaces are areas or volumes that provide for certain functions within a building". The main classes considered are both subclasses of the *Ifc:IfcProduct*, in detail, the *Ifc:IfcBuildingElement* indicates "all elements that are primarily part of the construction of a built facility, i.e., its structural and space separating system. Building elements are all physically existent and tangible things". The latter class was useful to describe the building components, and it was divided into *Load Bearing Skeleton* itself subdivided into *Steel Frame* and *Concrete Frame*; *Horizontal Closures* divided into *Horizontal Base Closing*, *Horizontal Intermediate Closing* and *Horizontal Top Closing* that can be Plane, Sloped or Curved; and *Vertical Closures*. For instance, considering the manufacturing task *whitewashing,* it was performed in the *Hollander Beater Room*, an individual of *Ifc:IfcSpace*, and it is enclosed in *Industrial floor1* as *Horizontal Base Closing*, and *Barrel Vaults* as *Horizontal Curved Closing*, which replaced the *Horizontal Sloped Closing Wooden Truss Roof* that was damaged during the second world war. For the *Vertical Closing*, we identified two different closings instantiated as *Tuff Wall 1* and *Tuff Wall 2*, probably built in two different periods, and they have incorporated the structures derived from the Hydroelectric and hydraulic canals Construction.

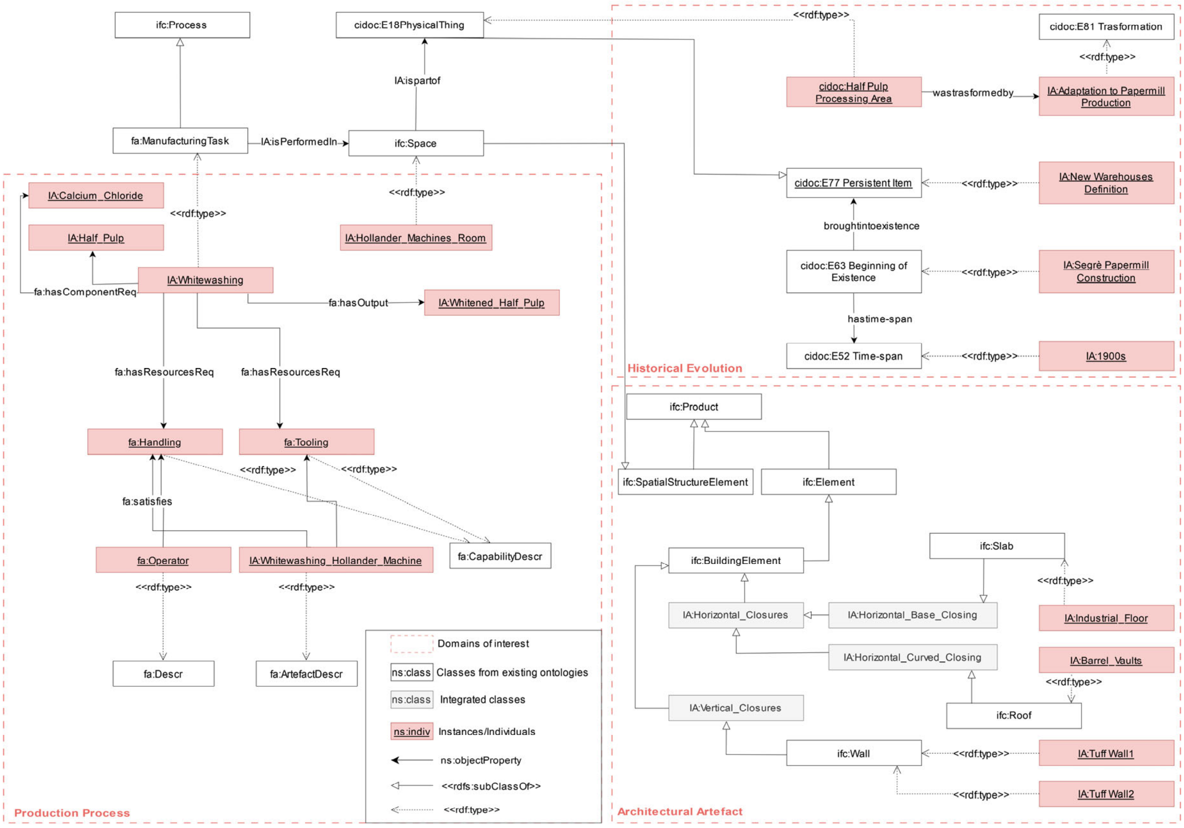

Fig. 4: Instantiation process of the whitewashing activity performed in the Hollander Beater Room.

Moreover, the building construction phases are very intricated due to the multiple overlapping structures, and therefore a clear explanation is necessary for a deep understanding of the artefact elements. For this reason, the Spaces defined in the architectural artefact domain are connected with its historical evolution through the property *ispartof,* that associates the *Ifc:IfcSpace* with the E18 class of the CIDOC-CRM *PhysicalThing*. As explained in paragraph 1.1, the CIDOC-CRM provides a common and extensible semantic framework for evidence-based cultural heritage information integration. The classes used are E18 Physical Thing appropriate for the definition of different rooms where the production process took place, i.e. the *"Half Pulp" processing area*. The E18 is a subclass of the E77 Persistent Item, which defines items with a persistent identity, sometimes known as "endurants" in philosophy. For example, in our case, it could be interpreted as *New Warehouses Definition*, considering the latest papermill production. E66 Beginning of Existence is instantiated as *Papermill*

Construction and it is connected with the E77 Persistent Item through the propriety *hasBroughtintoExistence* and through the property *hasTimeSpan* it is related to E52 Time-Span instantiated as 1900s. Those classes are fundamental to documenting a specific state of the artefact, and we also considered other CIDOC classes to document the transformations that modified the physical structures and also functions. To express the transformation, we considered E81 Transformation, instantiated as *Adaptation to Papermill Production.* Other more detailed classes and proprieties can be helpful, such as E11 Modification and its subclasses E79 Part Addition and E80 Part Removal, which can be directly related to the E18 Physical Thing through the propriety *hasModified.*

In Figure 4 is represented a small portion of the instantiation process which regards one of the activities performed in the Hollander Beater Room. All three domains are represented and are interconnected with each other. Each single manufacturing task is then connected with the following tasks through the input components defined as Component requirements and the outputs which are the input of the subsequent activity. The tasks are then connected with the spaces and the historic evolution domain as explained above. As defined in the picture, we combined classes from existing ontologies with integrated classes, we represented the individuals of each class in red, correlating them with object properties.

4. INTEGRATING KNOWLEDGE-BASED REPRRESENTATION AND HISTORIC BUILDING INFORMATION MODELLING

The knowledge formalization related to the latest papermill production, which includes the disciplines presented above, is then integrated with geometrical and technological aspects of the artefact components. The first connection between the knowledge structure and the informative model has been made through the correlation between the *Ifc:IfcSpace* of the knowledge base and the *Ifc:IfcSpace* of the BIM model. The IfcSpace is defined, for our case study, as the rooms or areas where the activities were performed. The integration of the two models is a critical process. It is necessary to carefully decide which information can better connect these two environments to ensure interoperability and data alignment.

In the BIM environment, each room has a label, for instance, *Paper Machine Room* or *Hollander Beater Room*, individuals of the *IFCSpace*. The label correspondence is between the name of the room in BIM and the individuals of the knowledge base *Paper Machine Room* or *Hollander Beater Room*, instances of the *IfcSpace* Class.

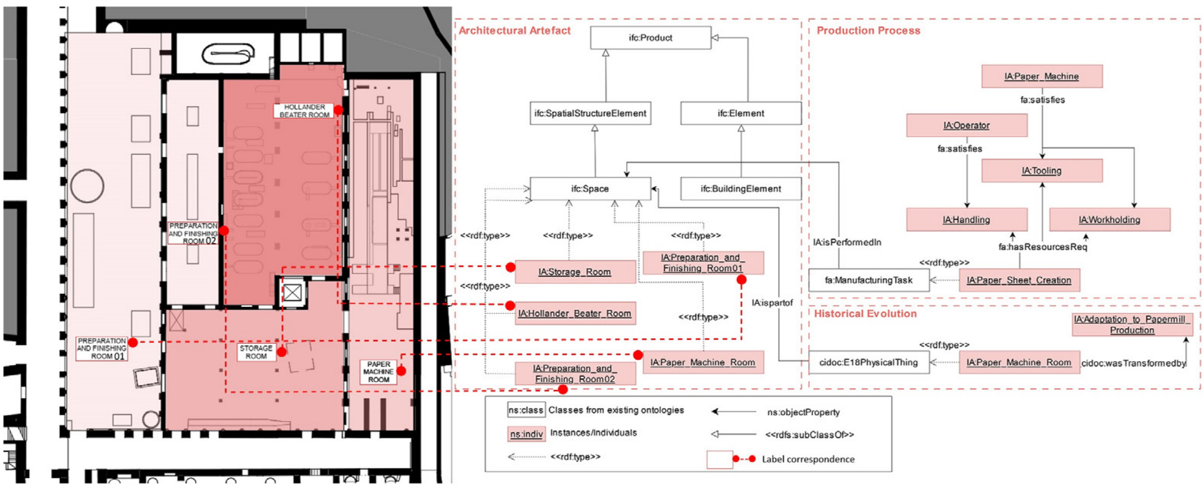

Fig. 5: Label correspondence between the rooms in the model and the individuals in the knowledge base.

The label corresponding process is represented in Fig. 5, through this integration activity is possible to underline a fundamental aspect of this approach. In fact in this way, we can connect the entities with other concepts, abstract or concrete which belongs to different domains in a cross disciplinary way. Using only the BIM representation schema is quite limited since many relations and concepts cannot be fully interpreted and represented. This approach can answer some of the critical issues related to the current practices in the digital built heritage processes. The documentation of the artefact requires the representation of a broader domain of knowledge, which includes extremely diverse and specific aspects such as historical, cultural contexts, construction techniques, and the history of materials which guide its accurate interpretation. Therefore it is possible to represent and manage information from an incremental and recursive perspective and from heterogeneous sources. The complex nature of built heritage fields needs appropriate tools and approaches to better address the multiple and open issues.

While in the first case, the label correspondence was between spaces and rooms, in this second case, we focused on the building elements. The *ifc* schema has a rigorous hierarchy representation, and sometimes to represent the heritage buildings, we are forced to use this schema even if it does not address the singularity and uniqueness of these artefacts. The building components of a historical building may not precisely correspond to specific classes in the *ifc* structure, especially when considering overlapping structures where every element may have acquired different meanings and functions over time, embedding a wide range of historical, cultural, social and technological values. In Figure 6, we can see the label correspondence between the two models by considering some of the building elements. For the architectural artefact representation, it was included in the knowledge base another class level as a subclass of the *ifc:BuildingElement* defined by *Ia:HorizontalClosures* subdivided in *Ia:HorizontalBaseClosing* and *Ia:HorizontalSlopedClosing*, and *Ia:VerticalClosures*. The instantiation process represented within the red boxes in the diagram is conformed to the individuals in the BIM model, which are likewise instances of *ifc:Wall*, *ifc:Slab* and *ifc:roof*.

The possibility of correlating concepts outside the informative model helps us to represent the authenticity of the heritage artefacts, the overlapping structures of the building components, the inconsistencies, and the different interpretations of single or multiple elements. These are all steps of the documentation and investigation processes that all the specialists perform in their field and, through this integrated approach, can be represented consistently in a machine-readable way.

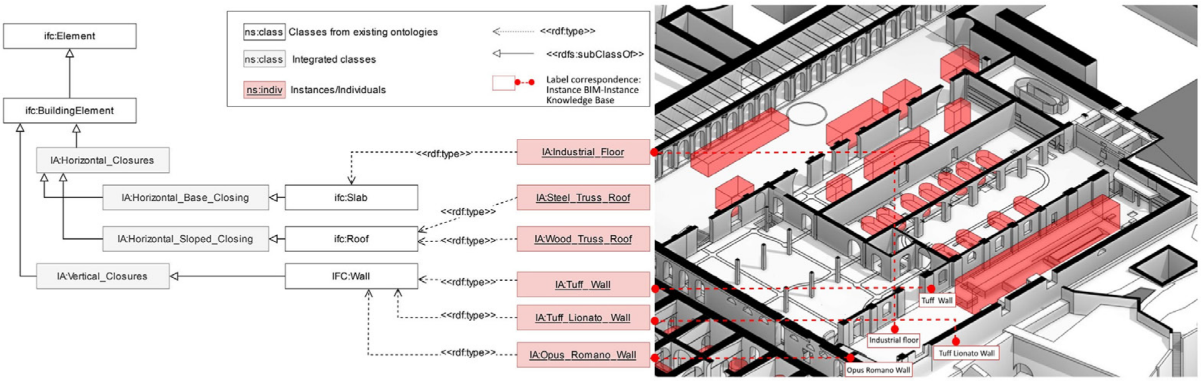

Fig. 6: Label correspondence between building elements in the model and the individuals in the knowledge base. BIM model carried out by Andrea De Pace and Riccardo Rocchi under the supervision of Edoardo Currà.

5. CONCLUSIONS

This research focuses on modelling the main features of industrial archaeology in a multi-layered historical site. We developed a structured knowledge base integrated with the BIM model to document the latest papermill production of the Segrè family. The historical industrial process and the manufacturing machines have shaped over time the architectural assets built on the remains of the Roman sanctuary and the former iron and powder productions. The knowledge base creation started by defining the domain of interests: production process, architectural artefact and historical evolution, followed by the knowledge base creation through an ontology conceptualization, reuse and encoding activities. Finally, the last part deals with integrating the knowledge base with the informative model made through a label correspondence between the same IFC classes.

The novelty of this integrated approach addresses some issues presented in section 1. Starting from a clear gap in the digital documentation process of the industrial heritage field, we proposed a methodology for the knowledge representation that considers all the domains of interest, with a specific focus on the industrial process, to customize the existing ontological approaches for contemporary manufacturing and Industries 4.0. Over time, the aggregated work on multiple industrial heritage sites can extend the queries on a large amount of data and scale to define similarities through the use of computable knowledge, highlighting differences and multiple interventions and recovery actions based on regional and transnational cases of industrial archaeology.

The application and validation of the proposed approach in similar contexts are necessary to highlight improvement aspects, extend the concepts and relations of the considered domains, and implement other valuable domains. Moreover, the difficulties faced during the knowledge acquisition process, caused by data inconsistency and uncertainties, could be managed by new ways of knowledge representation systems able to document and grade multiple interpretations schemas. Indeed, the following step of this research is to overcome the above-presented limits of this approach to better represent and manage the built heritage knowledge.

REFERENCES

Achille, C., Lombardini, N., & Tommasi, C. (2015). BIM and cultural heritage: compatibility tests in an archaeological site. Building Information Modelling (BIM) in Design, Construction and Operations, 1, 593–604.

Acierno, M., Cursi, S., Simeone, D., & Fiorani, D. (2017). Architectural heritage knowledge modelling: An ontology-based framework for conservation process. Journal of Cultural Heritage, 24, 124–133.

Ameri, F., & Dutta, D. (2006). An upper ontology for manufacturing service description. Proceedings of the ASME Design Engineering Technical Conference, 2006(January).

Ameri, F., Wallace, E., & Kulvatanyou, B. (2020). Towards a reference ontology for supply chain management. CEUR Workshop Proceedings, 2900, 0–2.

Banfi, F., Previtali, M., & Brumana, R. (2020). Towards the development of a cloud-based BIM platform and VR apps for complex heritage sites subject to the risk of flood and water level changes. IOP Conference Series: Materials Science and Engineering, 949(1).

Barrile, V., Fotia, A., Candela, G., & Bernardo, E. (2019). Integration of 3d model from uav survey in bim environment. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W11), 195–199.

Bianchini, C. (2014). Survey, modeling, interpretation as multidisciplinary components of a Knowledge System. SCIRES-IT - SCIentific RESearch and Information Technology, 4(1), 15–24.

Bosco, A., D'Andrea, A., Nuzzolo, M., & Zanfagna, P. (2019). A Bim Approach For The Analysis Of An Archaeological Monument. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W9), 165–172.

Cairoli, F. G., & Ten, A. (2016). Santuario Di Ercole Vincitore A Tivoli Iii. L'architettura (Vol. 2016, Issue Serie)

Cheng, H., Zeng, P., Xue, L., Shi, Z., Wang, P., & Yu, H. (2016). Manufacturing ontology development based on industry 4.0 demonstration production line. Proceedings - 2016 3rd International Conference on Trustworthy Systems and Their Applications, TSA 2016, 42–47.

Colucci, E., Xing, X., Kokla, M., Mostafavi, M. A., Noardo, F., & Spanò, A. (2021). Ontology-based semantic conceptualisation of historical built heritage to generate parametric structured models from point clouds. Applied Sciences (Switzerland), 11(6).

Currà, E., D'Amico, A., & Angelosanti, M. (2022). HBIM between Antiquity and Industrial Archaeology: Former Segrè Papermill and Sanctuary of Hercules in Tivoli. In Sustainability (Vol. 14, Issue 3).

Cursi, S., Martinelli, L., Paraciani, N., Calcerano, F., & Gigliarelli, E. (2022). Linking external knowledge to heritage BIM. Automation in Construction, 141(October 2021).

Di, Y., & Wu, C. (2011). Research on the Building Information Model of the stone building for heritages conservation with the outer south gate of the Ta Keo Temple as an example. 2011 International Conference on Electric Technology and Civil Engineering, ICETCE 2011 - Proceedings, 20100470786, 1488–1491.

Diara, F., & Rinaudo, F. (2020). Building Archaeology Documentation And Analysis Through Open Source Hbim Solutions Via Nurbs Modelling. - ISPRS Archives, 43(B2), 1381–1388.

Diara, F., & Rinaudo, F. (2021). Ark-bim: Open-source cloud-based hbim platform for archaeology. Applied Sciences (Switzerland), 11(18).

Doerr, M., Bruseker, G., Bekiari, C., Ore, C. E., Velios, T., & Stead, S. (2020). Volume A: Definition of the CIDOC Conceptual Reference Model.

Douet, J. (2016). Industrial Heritage Re-tooled: The TICCIH Guide to Industrial Heritage Conservation. Taylor \& Francis.

Drobnjakovic, M., Kulvatunyou, B., Ameri, F., Will, C., Smith, B., & Jones, A. (2022). The Industrial Ontologies Foundry (IOF) Core Ontology. CEUR Workshop Proceedings, 3240.

Europeana. (2017). Definition of the Europeana Data Model. 66.

Francesconi, E., Montemagni, S., & Peters, W. (2010). Integrating a Bottom – Up and Top – Down Methodology for Building Semantic Resources. Processing, 95–121.

Garagnani, S., Gaucci, A., & Govi, E. (2016). Archaeobim: Dallo scavo al building information modeling di una struttura sepolta. Il caso del tempio tuscanico di uni a marzabotto. Archeologia e Calcolatori, 27, 251–270.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge Acquisition, 5(2).

Guerrero Vega, J. M., & Pizzo, A. (2021). Análisis arquitectónico y aplicación de metodología bim en el santuario extraurbano de tusculum. Archeologia e Calcolatori, 32(1), 99–116.

Harpring, P. (2010). Development of the Getty Vocabularies: AAT, TGN, ULAN, and CONA. Art Documentation: Journal of the Art Libraries Society of North America, 29(1), 67–72.

He, J., Liu, J., Xu, S., Wu, C., & Zhang, J. (2015). A gis-based cultural heritage study framework on continuous scales: A case study on 19th century military industrial heritage. - ISPRS Archives, 40(5W7), 215–222.

Hellmund, T., Hertweck, P., Hilbring, D., Mossgraber, J., Alexandrakis, G., Pouli, P., Siatou, A., & Padeletti, G. (2018). Introducing the heracles ontology—semantics for cultural heritage management. Heritage, 1(2), 377–391.

Kazil, J., Masad, D., & Crooks, A. (2020). Utilizing Python for Agent-Based Modeling: The Mesa Framework. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 12268 LNCS (Issue May 2021). Springer International Publishing.

Lemaignan, S., Siadat, A., Dantan, J. Y., & Semenenko, A. (2006). MASON: A proposal for an ontology of manufacturing domain. Proceedings - DIS 2006: IEEE Workshop on Distributed Intelligent Systems - Collective Intelligence and Its Applications, 2006, 195–200.

Logothetis, S., Delinasiou, A., & Stylianidis, E. (2015). Building information modelling for cultural heritage: A review. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(5W3).

López, F. J., Lerones, P. M., Llamas, J., Gómez-García-Bermejo, J., & Zalama, E. (2018). A review of heritage building information modeling (H-BIM). Multimodal Technologies and Interaction, 2(2).

Manacorda, D. (2014). Prima lezione di archeologia. Editori Laterza.

Martin Doerr. (2003). The CIDOC CRM – An Ontological Approach to Semantic Interoperability of Metadata. AI Magazine, 24(3), 75–92. http://cidoc.ics.forth.gr/docs/ontological_approach.pdf

Moscati, P. (2021). Digital Archaeology: From Interdisciplinarity to the 'Fusion' of Core Competences Towards the Consolidation of New Research Areas. Magazen, 2(2), 253–274.

Moyano, J., León, J., Nieto-Julián, J. E., & Bruno, S. (2021). Semantic interpretation of architectural and archaeological geometries: Point cloud segmentation for HBIM parameterisation. Automation in Construction.

Myers, D., Dalgity, A., & Avramides, I. (2016). The Arches heritage inventory and management system: a platform for the heritage field. Journal of Cultural Heritage Management and Sustainable Development, 6(2), 213–224.

Pocobelli, D. P., Boehm, J., Bryan, P., Still, J., & Grau-Bové, J. (2018). Building information models for monitoring and simulation data in heritage buildings.- ISPRS Archives, 42(2), 909–916.

Prestes, E., Federal, U., Grande, R., Rama, S., Pontif, F., Cat, U., Gra, R., Luis, J., Universidade, C., & Grande, R. (2014). Core Ontology for Robotics and Automation 1 st Standardized Knowledge Representation and Ontologies for Robotics and Automation , Workshop Chicago , Illinois , USA. September.

Ronzino, P. (2015). Cidoc Crmba A Crm Extension For Buildings Archaeology Information Modeling.

Sampath Kumar, V. R., Khamis, A., Fiorini, S., Carbonera, J. L., Alarcos, A. O., Habib, M., Goncalves, P., Howard, L. I., & Olszewska, J. I. (2019). Ontologies for industry 4.0. Knowledge Engineering Review, 34, 1–14.

Sanfilippo, E. M. (2018). Feature-based product modelling: an ontological approach. International Journal of Computer Integrated Manufacturing, 31(11), 1097–1110.

Sanfilippo, E. M., Terkaj, W., & Borgo, S. (2021). Ontological modeling of manufacturing resources. Applied Ontology, 16(1), 87–109.

Sarıcaoğlu, T., & Saygı, G. (2022). Data-Driven Conservation Actions of Heritage Places Curated With Hbim. Virtual Archaeology Review, 13(27), 17–32.

Sarkar, A., & Šormaz, D. (2019). Ontology model for process level capabilities of manufacturing resources. Procedia Manufacturing, 39(2019), 1889–1898.

Scianna, A., Gaglio, G. F., & Guardia, M. La. (2021). HBIM data management in historical and archaeological buildings. Archeologia e Calcolatori, 31, 231–252.

Shults, R., Levin, E., Habibi, R., Shenoy, S., Honcheruk, O., Hart, T., & An, Z. (2019). Capability of matterport

3d camera for industrial archaeology sites inventory. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W11), 1059–1064.

Suárez-Figueroa, M. C., Gómez-Pérez, A., & Fernández-López, M. (2015). The Neon Methodology Framework: A Scenario-Based Methodology for Ontology Development. Applied Ontology, 10(2), 107–145.

Suárez-Figueroa, M. C., Gómez-Pérez, A., & Villazón-Terrazas, B. (2009). How to write and use the ontology requirements specification document. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5871 LNCS(PART 2), 966–982.

TICCIH. (2003). The Nizhny Tagil Charter for the Industrial Heritage. The Nizhny Tagil Charter for the Industrial Heritage, July, 1–6.

Trizio, I., Savini, F., & Giannangeli, A. (2018). The Building Information Modelling for the Documentation of an Archaeological Site. 2018 IEEE International Conference on Metrology for Archaeology and Cultural Heritage, MetroArchaeo 2018 - Proceedings, 204–210.

Usman, Z., Young, R. I. M., Chungoora, N., Palmer, C., Case, K., & Harding, J. A. (2013). Towards a formal manufacturing reference ontology. International Journal of Production Research, 51(22), 6553–6572.

Vegetti, M., Böhm, A., Leone, H., Henning, G., Conicet-utn, I., & Fe, A. S. (2016). SCONTO : A Modular Ontology for Supply Chain Representation.