Journal of Cultural Heritage 68 (2024) 225–236

Contents lists available at ScienceDirect

Journal of Cultural Heritage

journal homepage: www.elsevier.com/locate/culher

Original article

Ontology and collaborative data annotation for the standardisation of documentation processes of underground heritage spaces

Andrés Galera-Rodríguez[a,*], François Guéna[b], Mario Algarín-Comino[c]

[a] Departamento de Expresión Gráfica Arquitectónica, Instituto Universitario de Arquitectura y Ciencias de la Construcción, Universidad de Sevilla, Seville, Spain
[b] UMR MAP 3495 CNRS/MCC. Modélisations pour l'Assistance à l'Activité Cognitive de la Conception (MAACC) ENSA Paris La Villette. Paris. France
[c] Departamento de Proyectos Arquitectónicos. Instituto Universitario de Arquitectura y Ciencias de la Construcción, Universidad de Sevilla, Seville, Spain

ARTICLE INFO

Article history:
Received 15 May 2023
Accepted 3 June 2024

Keywords:
Semantic annotation
Documentation
Collaborative
Ontology
Underground Built Heritage (UBH)
Grottoes
Carmona

ABSTRACT

Progress in the field of documentation and safeguarding of heritage buildings has evolved considerably since the implementation and use of information models. However, these models are being refined and particularised in an attempt to cover the full range of elements to be protected.

There is one case that has not yet been sufficiently studied, namely the specific typology of underground built heritage spaces, where the procedures designed for more conventional buildings are not applicable. A lack of consensus has been detected in the standardisation of data processing, which is fundamentally since these spaces do not respond to the classification nor as a building or as an archaeological site.

In addition, the morphology of these spaces, -with irregular surfaces of complex geometry-, makes the three-dimensional representation indispensable to record concise and exact documentation.

This research proposes a specific methodology for processing heritage data from CIDOC-based models for underground built heritage spaces that fills the identified gap. The research also proposes solutions to the ontological problem of linking information and annotated data, establishing a methodology for performing this kind of work directly on three-dimensional models using web applications. This simplification of the procedure facilitates the reading and multidisciplinary access to the information of the agents involved in safeguarding heritage, replacing two-dimensional projections representation with the introduction of three-dimensional tools.

1. Introduction

1.1. Contextualisation of the case study. Geographic location

Las Cuevas de la Batida are located in the municipality of Carmona, Spain. Morphologically, they are located in Los Alcores ledge [1], an elevated plateau overlooking the territory adjacent to the Guadalquivir River depression. At the geological level, it is composed of a large layer of calcarenite rock, a porous rock that allows the filtration of water and an important escarpment in its Eastern slope. (Figs. 1 and 2).

1.2. Historical and archaeological context

The material extracted from these Cultural heritage quarries has historically been used since the 15th century for the construction of buildings both in the municipality of Carmona and the surrounding area. This is the case of the church of Santa María in Carmona [2]. In these quarries there are various inscriptions by stonemasons, such as the one inscribed on the rock with the name of the master Antón Gallego, dated in 1518 [3].

The material extracted is calcarenite rock. It is sedimentary type in nature. The different densities and consistencies of the extracted material in each geological stratum compel the stratigraphy of the extracted rocks, to be divided into two large groups [4]. On the one hand, there is the deeper and less porous extractable variety that can be used as building material, and on the other hand, the non-extractable variety, used for backfilling and for obtaining aggregates. This fact conditioned the exploitation of the site when it

* Corresponding author.
E-mail address: agalerar@us.es (A. Galera-Rodríguez).

https://doi.org/10.1016/j.culher.2024.06.003

Fig. 1. Cuevas de la Batida.

Fig. 2. Territorial plan and relationship of the historic centre of the town of Carmona (lower) with Las Cuevas de la Batida in purple (upper).

was in operation, configuring the current morphology of the complex, generating a series of caves (Fig. 3) when the thickness of the non-extractable and residual material stratum was substantial.

Of all the existing caves in the *Batida* area, the Antón Gallego cave was chosen as the one with the greatest variety and diversity of elements to be described. The aim is that the working methodology applied to this case study could be extrapolated directly to the rest of the caves, and also, indirectly, to other underground heritage sites. This cave has a surface area scarcely up 100 m^2etres, a maximum height of 6.70 metres and the maximum length of its main nave is 21 metres.

The result of this productive activity, which has been carried out since ancient times, is the creation of a site of important heritage value. This activity has transformed the landscape by means of open-cast excavation, creating large vertical surfaces and a series of cavities in some of these areas, which penetrate horizontally into the ground, creating spaces of considerable dimensions, with a great geometric complexity, which can sometimes be indistinguishable from natural geological formations.

1.3. Problem analysis

1.3.1. Underground built heritage, difficult access

This contribution is the result of the need for documentation, analysis and management of the preventive conservation of these spaces through data linked to the traces of the construction processes that have produced them, or the identification and monitoring of perceptible pathologies in their walls. The information obtained is often difficult to structure and this problem is compounded by the fact of the standardisation of data applied to heritage buildings, especially in underground spaces.

This is compounded by the difficulty of objectively comparing several similar heritage objects that, despite having a similar typology, are documented in different ways, which makes working in multidisciplinary teams very complex.

This problem is accentuated for the establishment of underground spaces typologies, for several reasons:

a) As the underground spaces are the result of material extraction processes, they differ from the usual tectonic processes in architecture.
b) Nor do they correspond to the characteristics of archaeological sites, since they are spaces conceived as subterranean *per se*, in many cases reversing the stratigraphic order.
c) The difficulty of locating and contextualising them. In many cases, they are hidden beneath the contemporary city and usually have limited and precarious access due to their size and lack of natural light.
d) The geometric complexity of this type of heritage. The irregularity of their morphology makes it difficult to produce a complete and exhaustive graphic representation. Two-dimensional representations are imprecise because they depend on the choice of sections since they are not flat or associated with a specific geometry. This type of representation makes semantic annotation of the data difficult, as much data may be hidden and undocumented.

If the typology establishment is not easy, the encoding of the information is even less so. The underground spaces own nature means that the information fields can be very different from those of any other type of heritage, which has given rise to a specific term, Underground Built Heritage, used by various authors [5], and also requires 3D documentation of these architectures [6].

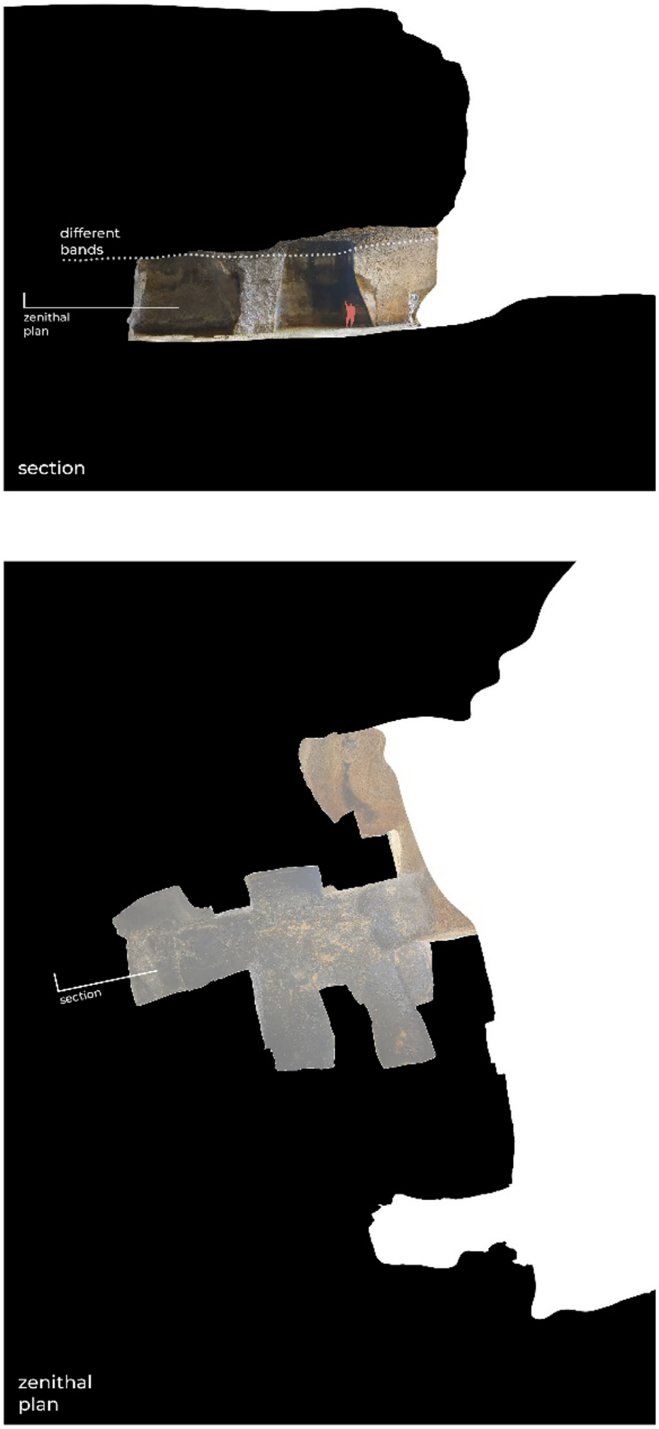

Fig. 3. Plan and section of the Cave of Antón Gallego. Carmona. Spain.

One solution to obtain a visualisation of this heritage typology is three-dimensional surveys using point clouds. However, modelling complex surfaces from point clouds requires an arduous optimisation process to make them compatible with and manageable by most graphics software [7,8].

2. Research aim and methodology

The input of this research presents the first results of a methodological proposal to solve these problems, with the aim of contributing to the processes of documentation and conservation of this type of heritage, in two work areas: The standardisation of information through a data ontology, and a strategy for the management and visualisation of alphanumeric information linked to a three-dimensional model through web applications.

Both lines of work will make use of reference models that have already been experimented and applied to cultural heritage, such as the CIDOC CRM entity systems, a basic ontology that allows the exchange of information from different heterogeneous sources on cultural heritage [9], and its CRM dig extension, which allows the collection of documentation procedures and the digital acquisition of heritage information [10].

In addition, virtual platforms and applications for accessible and multidisciplinary three-dimensional digital modelling will be used to provide the necessary data for correct heritage documentation.

With this objective in mind, the data developed in the ontology structuring process will be used to give order and meaning to the mass of information in the three-dimensional digital model, searching for a systematic way of working that can help the interdisciplinary group involved in the documentation and management: the technicians who capture the metric data by means of point clouds, the archaeologists and historians who record the information, the technicians in charge of preventive conservation or consolidation interventions and, at a later stage, the general public, although the latter would initially be outside the scope of this work proposal.

The first results of the attempt to bring together these two work lines in order to respond to the proposed objectives are presented below, beginning each section with a brief background description.

3. Materials and methods

3.1. Data ontology

Data ontologies are the definition of the relationships between a network or data system. For each domain or area of knowledge there is a specific ontology. The reference ontology for the interchange of cultural heritage information is the CIDOC–Conceptual Reference Model (CRM) contained in the ISO 21,127:2014 standard [11]. This data ontology was originally developed for the exchange of information managed by museums, libraries and archives, specifically promoted by the International Council of Museums (ICOM). The similarities of these exchanges with the management of heritage information led to the spread and implementation of its use also in the field of cultural heritage [12]. Currently, the CIDOC–CRM ontology is used in most of the ontology diagrams proposed by research related to cultural heritage, creating an extensible and adaptive categorisation to other complementary ontologies [13].

Reading ontology schemas is not always an easy task as it requires not only advanced knowledge of the subject so to understand the information encoding, but also to take into consideration that the ontology design depends on multiple factors such as the type of heritage elements or the process to be described in each semantic data schema.

For this purpose, a review of the current state of research in the field of study is carried out using graphical means. The CIDOC–CRM system is used as the basis for most of the ontology designs for heritage data. In order to be able to compare one research with another on the basis of common data, a study of the most commonly used entities is carried out (Fig. 4). The results indicate which are the most common entities in the ontology schemas that can be used for a basic and standard description of heritage assets, such as the basic entities *E21 Person, E39 Actor, E52 Time-Span, E53 Place or E55 Type*. On the other hand, the less used ones serve for a specific description of particular heritage data.

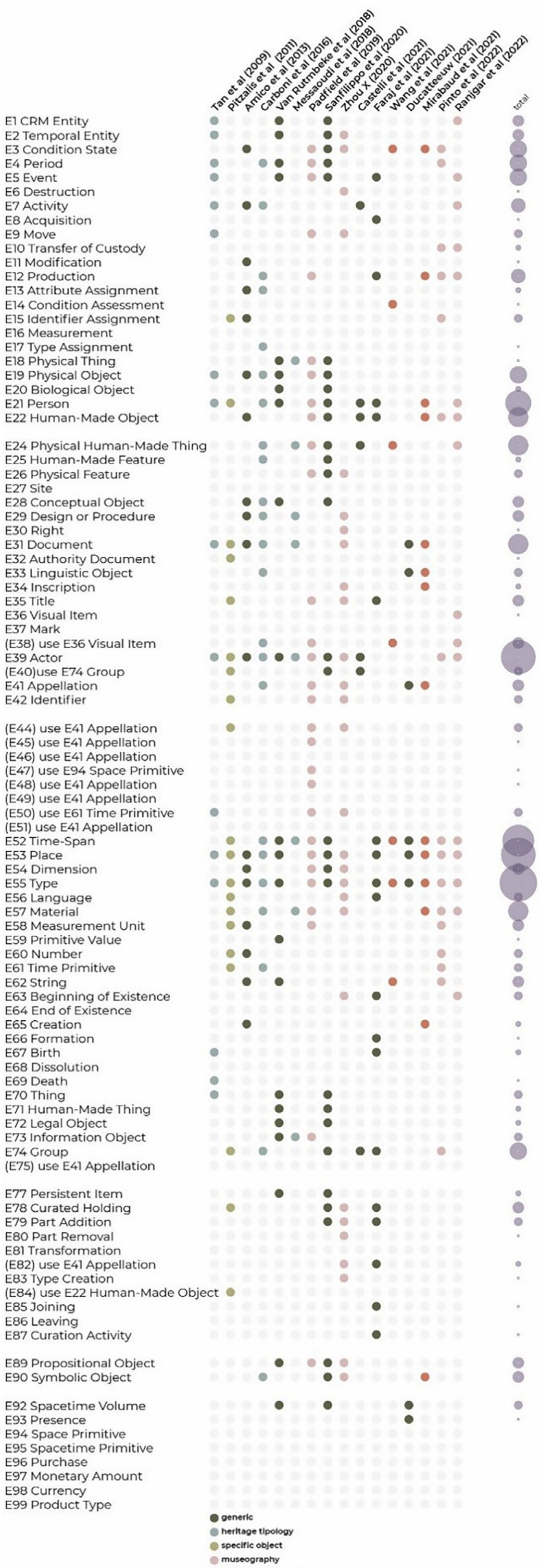

Fig. 4. Comparative diagram of the CIDOC CRM entities used in the different investigations. The ontology typology is differentiated by colour. In dark green the generic ontologies, in light blue the ontologies designed for a specific heritage typology, in light green the ontologies dedicated to defining a specific heritage object, in pink the ontologies that describe museum processes and in orange the ontologies on underground heritage spaces.

A colour classification of the type of ontology is also carried out (Fig. 4). According to the research analysed, ontology designs are of a generic nature, i.e. they do not describe a specific heritage element, but can be applied to any type of heritage [14-19]. Research focused on the conservation state [20]. Other research designs ontology diagrams specific to a heritage typology, such as intangible heritage [21], or even to a very specific singular element, such as the "Cylinder Seal of Ibni-Sharrum" [22]. Of particular interest [23-26] are the ontologies created for museographic purposes.

If we focus on the specific field of underground heritage spaces, we can find far fewer cases, but the following research is worth mentioning: Generation of ontological diagrams for the heritage inventory of one of the caves on the Silk Road in China, in this case the ontology is focused on studying its state of conservation [27]. Research on the Ethiopian rupestrian churches, generating a series of ontological models, in this case on a representative fragment of mural painting. A first model is designed for the description of these visual representations and their iconographic analysis. The second model describes the painting materials and their conservation conditions, as well as including information on samples taken and the results of their physico-chemical analysis. Finally, a model is designed for the digitisation of the heritage element, focusing on the elaboration of three-dimensional digitisation processes from a point cloud, also including the orthophotos taken. For this model, the CIDOC–CRM ontology is complemented with the CRMdig extension for heritage digitisation [28].

3.2. Adaptation of the data ontology for underground spaces

For the correct management of the data and graphic documentation of underground heritage spaces, once the state of the art has been analysed, 3 levels of data ontology corresponding to the different heritage management processes are proposed, using the CIDOC CRM and CRM dig codes:

a) Standardise the process of obtaining and classifying alphanumeric data on underground built heritage spaces.
b) Standardise the complete and comprehensive metric capture of underground built heritage spaces.
c) Standardise the generation of three-dimensional models of underground spaces.

This specific ontology (Fig. 5) seeks to standardise the documentation process of underground built heritage spaces from a complete description including data of different typology.

The ontologies proposed below have been designed to describe specific processes in the documentation of underground built heritage spaces. They are based on existing ontology diagrams described in the existing research previously described. Since these existing ontologies do not describe the complete heritage documentation processes that this research aims to achieve for this unique and specific heritage typology, the methodology for generating these diagrams has consisted of a partial selection of other diagrams, either because they describe similar documentation processes or because they are investigations on the same heritage typology.

The heritage element under study, in this case a particular cave from *Las Cuevas de La Batida* (described in 1.1), is defined as *Physical Thing* (E18), and documented as *Information Object* (E73)

On the one hand, entity E73 collects the alphanumeric descriptive information of specific data of the heritage asset. For this purpose, a series of specific information fields are established. Firstly, each underground built heritage space is coded with *Identifier* (E42) according to the heritage typology, using codes of type Xx00, in this case Ca01, where Ca corresponds to the heritage typology "cave" and the numbers, to the inventory. In addition to the identifier, other descriptors are used. *Period* (E4), *Place* (E53), *Type*

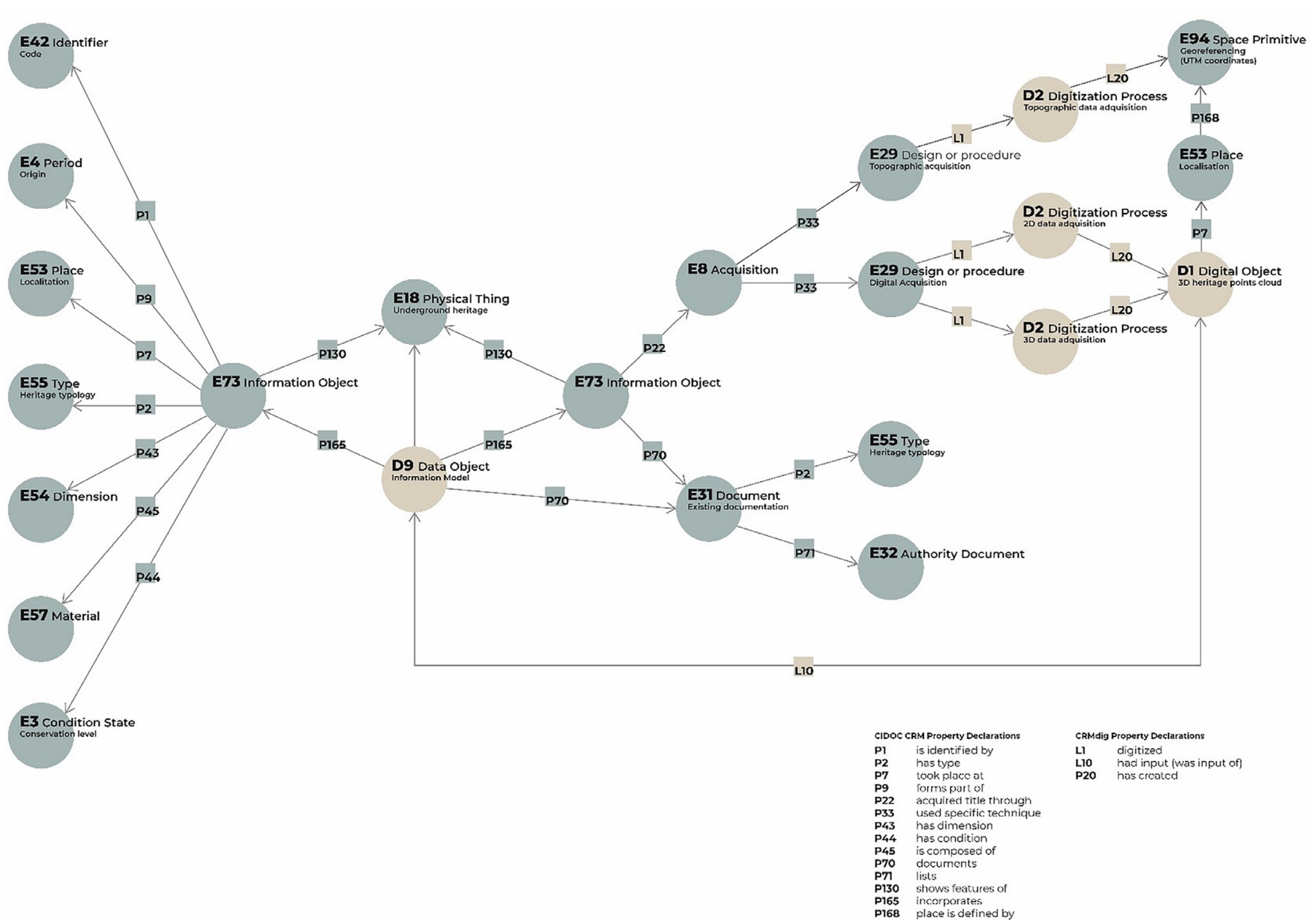

Fig. 5. Ontology diagram for the documentation of underground built heritage spaces.

(E55), *Dimension* (E54), *Material* (E57) and *Condition State* (E3). This systematics are already common in many sources, and it does not seem necessary to discuss them further.

However, for the description of another object information entity (E73) of a graphic nature, related to the existing graphic documentation of the heritage asset, the prescriptor *Documen*t (E31) is used, and for the acquisition of new data, *Acquisition* (E8). Using these prescriptors, published experiences are more limited, so we will develop them in the following subsections.

We will start from the existing documentation (E55) which is in turn classified with the heritage typology fields *Type* (E55) and author of the document *Authority Document* (E32).

Also, for the acquisition of new data (E8), involving metric data collection and modelling, a digital survey procedure *Design or procedure* (E29) is specifically created, which is composed of two digitalization processes: *Digitalization Process* (D2), one for the acquisition of two-dimensional information and the other one for three-dimensional information. Both processes are encompassed in a digital object *Digital Object* (D1), corresponding to a point cloud.

In parallel, for the geolocation of the point cloud, a parallel procedure for the topographic acquisition *Design or procedure* (E29) has been generated from a topographic information acquisition process *Digitalization Process* (D2) that creates the georeferencing of the 3D heritage points cloud from UTM coordinates *Space Primitive* (E94).

This process is detailed in a schema of specific ontology model for data acquisition and three-dimensional modelling of heritage built underground spaces (Fig. 5).

All the information generated and processed in the object (E73) and digital model (D1) information entities is collected in a single information model, *Data Object* (D9).

3.2.1. Ontology for data acquisition and three-dimensional modelling

As mentioned above, to carry out a complete and exhaustive graphic documentation of underground built heritage spaces, which by definition are characterised by being composed of irregular surfaces of complex geometry, it is essential to develop three-dimensional models.

This work is usually carried out by a specialist team which provides a point cloud using a variety of techniques that the rest of the multidisciplinary team are not usually aware of.

It is for this reason that it is necessary to generate a specific ontology -that could be recognized by all the coworkers- for heritage digitisation from three-dimensional surveys made up of point clouds (Fig. 6). To obtain a point cloud of the heritage element (E18), two digitisation processes *Digitisation Process* (D2) are proposed. Depending on the characteristics of the space to be documented, the most appropriate solution will be selected to use: point clouds generated by laser scanning or clouds generated using the photogrammetry technique. Each of the digitisation processes incorporates the information of the date on which the process is carried out, denominated *Time-Span* (E52) and the author of the procedure, *Actor* (E39).

The process of generating point clouds using laser scanning equipment is carried out of a series of stages, starting with the planning and procedure for the specific data acquisition, *Design of*

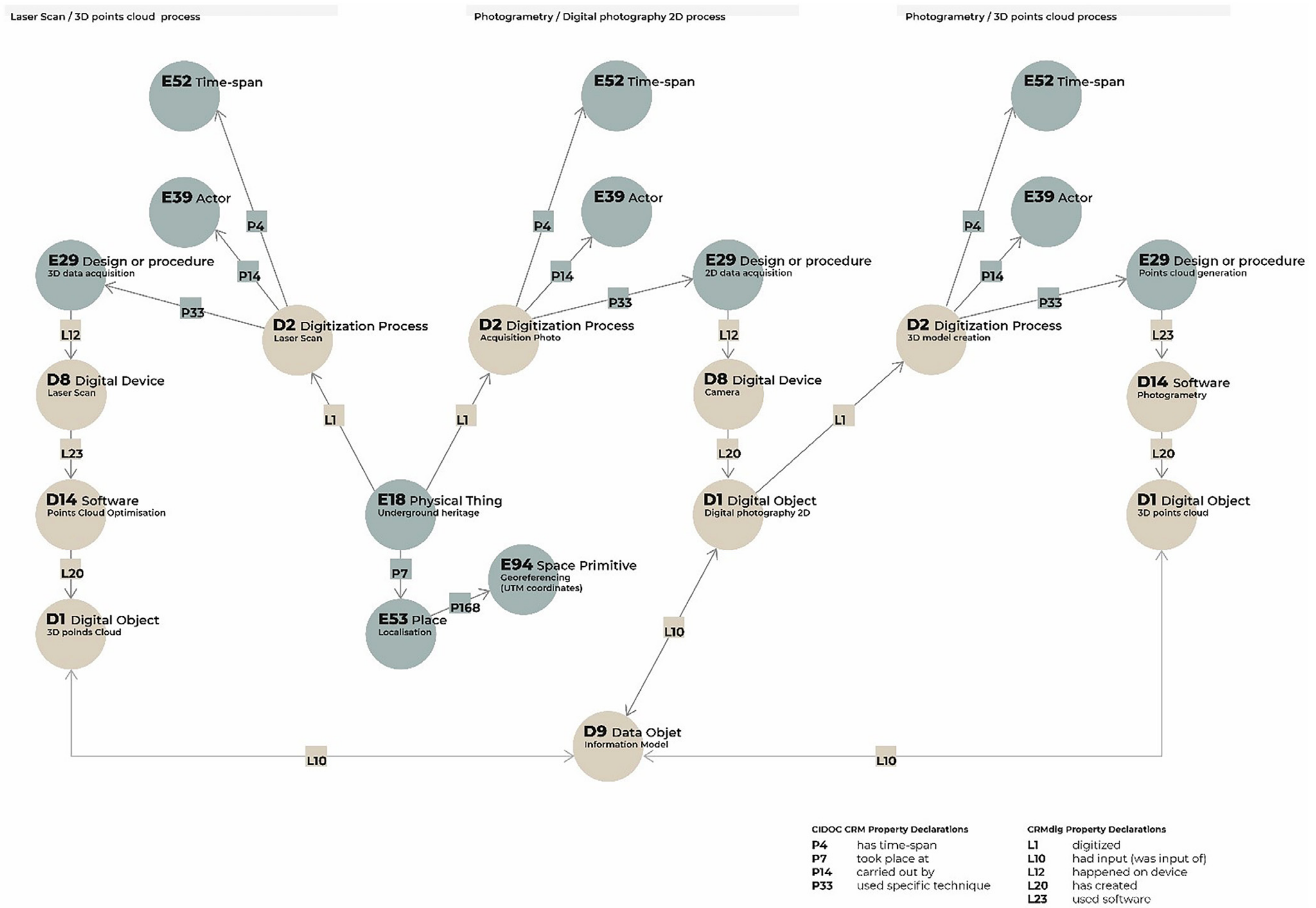

Fig. 6. Ontology diagram for data collection and three-dimensional modelling.

procedure (E29), the data acquisition is carried out using a specific laser scanning equipment, *Digital Device* (D8). During the process and afterwards, the cloud generated by the scanner is optimised by means of specific software, *Software* (D14). Finally, a three-dimensional model of the heritage asset is obtained, consisting of an optimised point cloud, *Digital Object* (D1).

The methodology for the acquisition of point clouds with photogrammetry techniques is mainly divided into two digitisation processes (D2). The first process is the photographic data acquisition of the heritage asset (E29), using specific photographic equipment (D8) to generate a collection of images (D1). Once the data have been collected, the next process begins, the creation of the three-dimensional point cloud model (E29) using specific software (D14), obtaining a three-dimensional model (D1).

All the collected information resulting from the different three-dimensional data acquisition processes, such as point clouds generated by laser scanning or photogrammetry, including the collection of images, are compiled in a unitary information scheme, *Data Object* (D9).

3.2.2. Ontology for annotation of heritage data on three-dimensional models of underground spaces

Once the data ontology for the documentation -obtained from metric capture and generation of three-dimensional models- of underground built heritage spaces have been established, the next level is to define a ontology for the annotation of data on these three-dimensional models generated from web applications (Fig. 7). The objective is to unify and encompass all the information from the two previous ontologies in a single standardised and complete model, through web applications.

The procedure to be executed consists of the following: on the three-dimensional point cloud *Digital Object* (D1) of the heritage element *Physical Thing* (E18), a selection of the geometric regions, *Area* (D35), of the point cloud to be defined according to each category is made. The set of points included in each region is classified in different annotation categories, producing an association *Annotation Event* (D30) between the areas (D35) and the categorised data. The information in each annotation category, *Information Object* (E73), is defined in terms of the CIDOC CRM and CRM dig (E1) entities. A categorised annotation model, *Annotation Object* (D29), is generated thanks to the use of the web application *Software* (D14) *Aïoli platform* that allows interoperability between the different types of data described. With all this, a unique and specific proposition is designed for the annotation of information in three-dimensional models for underground built heritage spaces of complex geometry, *Propositional Object* (E89).

3.3. Semi-automatic annotation of heritage semantic data in three-dimensional models from point clouds using web applications

Once the specific ontology for underground built heritage spaces has been designed, we address the second objective of the work: how to make visible a heritage typology characterised by being hidden underground. To achieve this, the 3D annotation platform Aïoli, developed by the MAP research unit of the CNRS in France, was used to build semantically enriched descriptions from photogrammetric reconstruction and spatial labelling. This

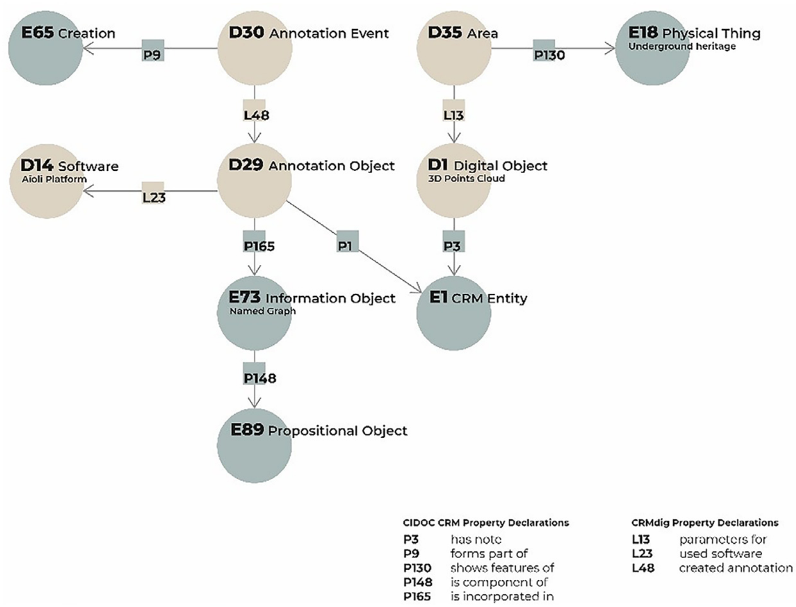

Fig. 7. Ontology diagram for the annotation of data on surfaces of three-dimensional models.

web application is of particular interest as it allows the linking of image-based modelling, 2D-3D diffusion and correlation of semantic annotations, and analysis of multiple attribute layers [29,30]. One of the most significant scientific applications of the Aïoli platform is its use as the main tool within the Digital Information Working Group for the post-fire reconstruction of Notre-Dame de Paris Cathedral. This tool was used to collect data on the current state of conservation of the various elements after the fire. The different phases of the reconstruction were also periodically recorded [31]. It has also been used to make a comparative approach between the semantic annotation and information management processes carried out with this platform and those that can be developed from HBIM information models [32].

In addition to the information management of heritage buildings, it has also been used to perform semantic annotations on movable assets, in this case as part of the digital diagnosis process of the Autumn statue, mainly indicating the alterations suffered by the carving at different times and the pathologies on its surface [33].

For the annotation of information on digital models, workflows based on HBIM tools and models could also be used [34]. However, the Aïoli Platform resource was chosen for this research because, in addition to allowing the development of different strategies that improve the documentation, conservation, enhancement and dissemination of heritage, it allows these data to be displayed and made accessible to a wider range of agents; facilitating access to information is essential for improving the conservation of heritage assets.

For this reason, once a specific hierarchy of data adapted to underground built heritage spaces has been designed, we propose to be able to visualise and locate them in the heritage asset, in this case on a three-dimensional model representing it.

The first step is to carry out a data collection campaign, performing a photogrammetric survey of the cave, similar to those carried out in other heritage buildings [35,36] or in underground spaces [37-39]. For this specific scenario, it is not possible to use a point cloud generated by 3D scanner, although this method would be more accurate, as the semantic annotation is generated directly by means of graphical regions on the images.

The photogrammetric elevation has been carried out from a photographic data collection with Sony ILCE-7RM3 equipment both inside the cave and its surroundings. Once the data had been collected, these images were processed and positioned using point cloud processing and generation software.

Once the point cloud is generated, it is introduced into the web application.

3.3.1. Structure of the information

The data are structured in different categories according to the ontological model specifically designed for this type of heritage elements, based on the CIDOC CRM and CRM DIG entity systems [40]. On one side, the heritage asset as a whole, with the general information of the entities -*physical thing, identifier, site, type, condition state, resignation, period, acquisition, design of procedure, actor, information object, document, authority document*-.

On a specific level, the following categories are annotated on the three-dimensional model: discontinuities and densities of materials, spatial composition, defining the structures that conform it, stages and differentiation between original spaces and extensions. Annotation of stonemasons' marks, inscriptions, pickaxe marks and ashlar cut marks is also included, considering that a prior analysis of the object to be analysed is necessary from different approaches: archaeological, architectural-constructive, historiographical, etc.

Finally, alterations such as fire marks, contemporary graffiti, cracks, lime stains, detachment of material and paint are also noted.

The items (of information) covered by this information model are structured and hierarchical at different levels (Fig. 8).

A first level of information comprises the general information levels into which the categories describing this heritage site are grouped. It is composed by categories of materiality, spatial composition, stonemasonry marks and alterations.

A second information level includes different subcategories of the previous level, established as descriptors. The descriptors have the objective of making an advanced and specific classification that allows a correct identification of this heritage typology. The category *Materiality* includes the descriptor of discontinuity of materials. The category *Spatial Composition* contains the descriptors of *Naves* (cave structures or sections), *Stages and Phases of excavation* and *Cave Extensions*. The category *Stonemasonś marks* include the descriptors of *Stonemasons' Inscriptions*, *Pickaxe Markings* and *Ash-*

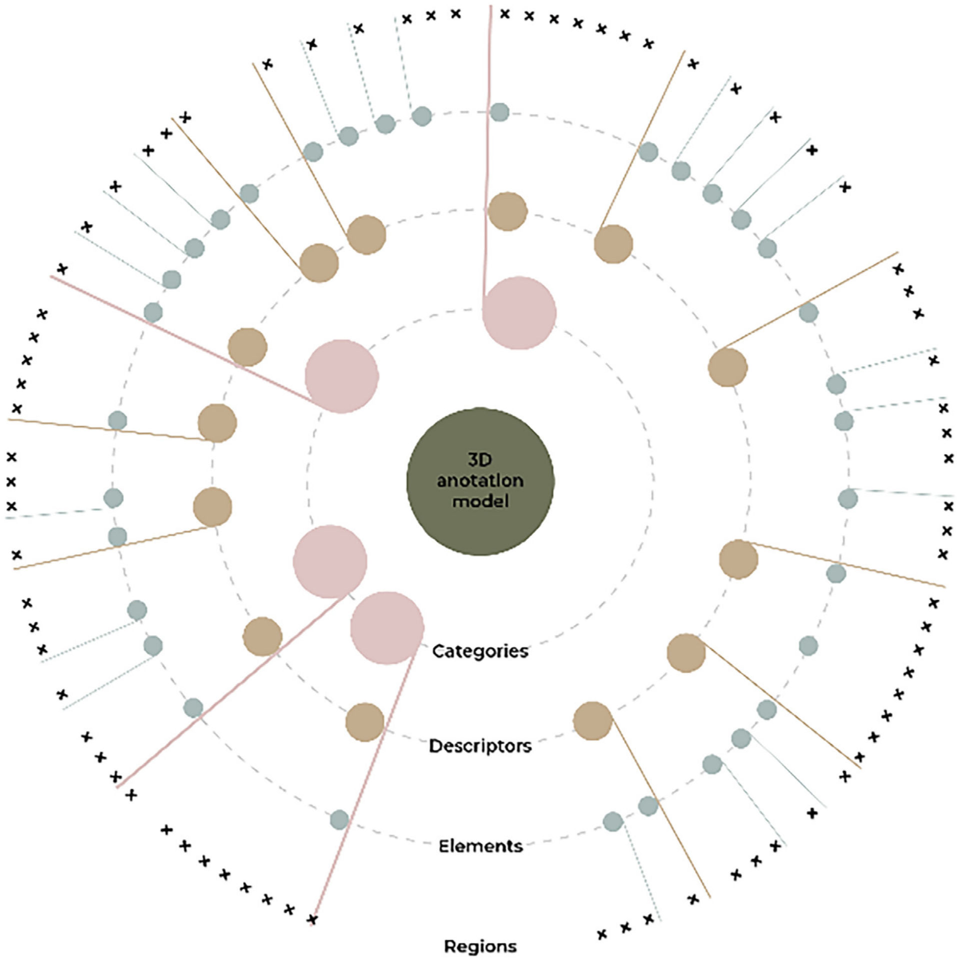

Fig. 8. Diagram of the structure of the different levels of information in the annotation model. The level of information corresponding to categories is represented in pink colour, descriptors in yellow, individual elements in blue and polygonal regions in black.

lar Cuttings. Finally, the *Alterations* category is constituted by the descriptors *Fire Marks, Cracks, Lime Stains, Detachments, Areas Covered with Paint* and *Contemporary Graffiti.*

The third information level includes each of the elements highlighted and described by the annotation that appear on the asset. Depending on the degree of presence of each descriptor, the number of elements of each of them, will vary.

The fourth and last information level is composed of each of the polygonal areas that have been drawn on the different images to identify each of the descriptors. The polygonal areas differ from the elements, as the latter can be composed of several regions. This is due to the geometric complexity of the surfaces of the space to be documented, as there are certain elements that cannot be seen as a whole in a single image, so several polygons have to be drawn to cover the entire element to be described.

All this information allows us to approach the knowledge of this type of space, which in the case of *La Batida* cave is related to a productive system that has had a great impact on the region, Lower Andalusia, previously studied through the application of GIS and Graph systems, given its territorial scale. It therefore represents an attempt to approach a more detailed and specific level of information on one of the key elements in the productive activity of a region, the quarries [41].

The database that has been generated, in addition to proposing a structure of specific information levels (Fig. 8), also designs a coding of the concrete and clear information for this heritage typology (Fig. 9) with the precise objective of facilitating the reading of the data by the agents involved. The coding is structured on the basis of the descriptors of the second information level, with levels 3 and 4 being associated to this level, thus generating a tree like (or branched) structure of data that allows the heritage information to be easily traced.

3.3.2. Web-based annotation platform

The Aïoli web application mentioned above processes the point cloud of the cave previously generated by a specific software and at the same time, locates and orients the images that have been generated by the point cloud. These are the images that will be used later to perform the data annotation.

This platform has been used in previous research for the management of heritage information of movable and immovable assets, but it has not been used for the description and documentation of Underground Built Heritage (UBH), this heritage typology cannot be classified as a building as they are really spaces generated by the extraction of material, which makes it difficult to standardise the processes of documentation and management of heritage information. This research proposes the standardisation of heritage information of this typology based on the Aïoli Platform.

In the image where some of the relevant features for annotation are detected, a graphical area associated with a particular feature is marked and the area is automatically extrapolated to the three-dimensional point cloud. The regions are stored in layers.

data framework		
Categories		
descriptors		
	UBH descriptors based on CIDOC-CRM	Type Domaine Aioli
Information general		
	E42 Identifier	Text
	E27 Site	Text
	E55 Type	Text
	E18 Physical Thing	Text
	E7 Activity	Text
	E3 Condition State	Text
	E54 Dimension	Decimals
	E4 Period	
	E8 Acquisition	Date
	E28 Design of Procedure	Text
	E39 Actor > (E74 Group E21 Person)	Text
	E73 Information Object	Text
	E31 Document	Hyperlink
	E32 Authority Document	Text
Materialité		
	E42 Identifier	Text
	E55 Type	Text
	E57 Material	Text
Domaine		
Discontinuité de matériaux	Discontinuité de matériaux	
	E37 Mark	Text
	E53 Place	Text
	E54 Dimension	Decimals
Composition Spatiale		
Nefs	Composition Spatiale_Nefs	
	E42 Identifier	Text
	E55 Type	Text
	E54 Dimension	Decimals
Étapes	Composition Spatiale_Étapes	
	E4 Period	Decimals
Extension	Composition Spatiale_Extensions	
	E79 Part Addition	Boolean
	E54 Dimension	Decimals
Marques des carriers		
Inscriptions	Marques de Carriers_Inscriptions	
	E24 Physical Human-Made Thing	Text
	E34 Inscription	Text
	E39 Actor	Text
	E56 Languaje	Text
	E54 Dimension	Decimals
	E4 Period	Text
Marques de pioche	Marques de Carriers_Marques de pioche	
	E24 Physical Human-Made Thing	Text
	E37 Mark	Text
	E54 Dimension	Decimals
	E4 Period	Text
Découpage des blocs d'extraction	Marques de Carriers_Découpage des blocs d'extraction	
	E24 Physical Human-Made Thing	Text
	E37 Mark	Text
	E54 Dimension	Decimals
	E4 Period	Text
Alterations		
Taches de Feu	Alterations_Taches de Feu	
	E24 Physical Human-Made Thing	Text
	E37 Mark	Text
	E54 Dimension	Decimals
	E4 Period	Text
Graffittis Contemporain	Alterations_Grafittis	
	E24 Physical Human-Made Thing	Text
	E25 Human-Made Feature	Text
	E37 Mark	Text
	E54 Dimension	Decimals
	E4 Period	Text
Fissures	Alterations_Fissures	
	E37 Mark	Text
	E54 Dimension	Decimals
Taches de calcaire		
	E37 Mark	Text
	E54 Dimension	Decimals
Détachements		
	E37 Mark	Text
	E54 Dimension	Decimals
Pintures		
		Text
		Text
	E24 Physical Human-Made Thing	Decimals
	E37 Mark	
	E54 Dimension	
	E4 Period	Text

Fig. 9. Chart with the categories, descriptors and elements defining the annotation model database.

Each layer and sublayer (Fig. 9) correspond to the different levels of information described above (Fig. 8).

Each outstanding element is defined by a series of fields based on CIDOC CRM entities and designed specifically for the heritage typology and for each of the descriptors. In this way, the alphanumeric information in the database is directly associated with the three-dimensional model of the point cloud (Fig. 10).

By applying the described procedure through web applications, it is possible to graphically visualise the annotation of the data of this complex typology of heritage assets (Fig. 10). Therefore, a more intuitive and complete alternative for reading the data is proposed, enriched in this paper by the application of a specific ontology that should contribute to its better integration in the processes of documentation and management of this type of heritage.

4. Discussion and results

The work carried out aims to go beyond obtaining a point cloud in order to obtain detailed information on the complex geometric heritage object of *La Batida* Cave, but also to advance in the definition of the concept of Survey or *Rilievo*, understood as a process of knowledge. The technological advances resulting from the numerous published experiences offer procedures and methods that are usually limited to very a specific disciplinary field, which are difficult to integrate in the daily management related to the protection of the heritage.

This contribution aims to show a first approach to the standardisation of the documentation process and the structuring of descriptive data on underground heritage, which will allow future research to adapt it to other types of underground heritage which, as we have explained, add to the usual difficulties inherent in the representation of built heritage, the complexity of its irregular geometry and the generation of its interior, solving or reducing its limitations with the proposed workflow.

The methodology used for the visualisation and localisation of specific data on the three-dimensional model aims to improve the workflow previously experimented on the web platform that the MAP research unit of the CNRS in France has allowed us to share. To achieve this, a specific data structure was designed based on CIDOC CRM and CRM dig, the 3D point cloud generated in Agisoft Metashape and the use of the Aïoli web application for direct annotation on the point cloud has been resorted to. This data flow has simplified the process of semantic annotation and the difficult reading of data ontologies that require specific prior knowledge. However, this workflow currently has the weakness that the annotation can only be done on 2D images that are then reflected in the 3D point cloud, and it is not possible to use point clouds generated by a 3D scanner. A future working line will be the investigation on different workflows that make data acquisition more flexible and broaden the range of technical solutions.

It is also intended to show the possibilities of specific ontological diagrams to describe specific aspects of underground cultural heritage spaces, within international standards to future research. For the design of ontology diagrams of the specific processes described on *documentation, data collection* and *annotation,* it has been possible to carry out a complete description only with the CIDOC CRM and CRM ontologies. However, it should be noted that the adopted cataloguing may also be a limitation for the ontological description of other specific processes, so it may be necessary to complete the diagrams by adding and specialising the ontologies with other additional ontologies, thus extending the scope and allowing to alleviate this limitation.

From the information obtained as a result of the standardisation process of heritage documentation processes, cataloguing has been

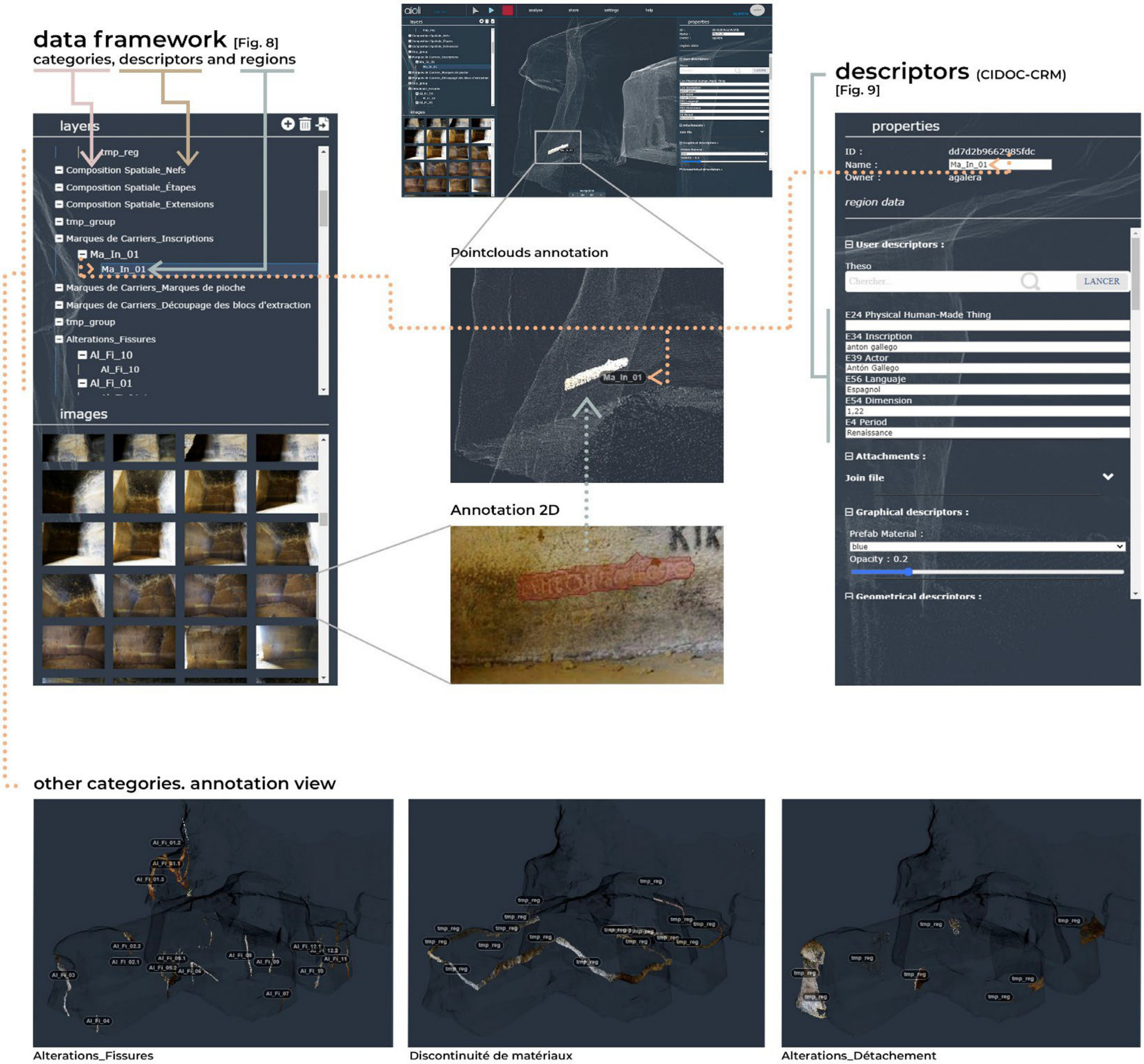

Fig. 10. Three-dimensional web model which shows all the elements of one of the descriptors, in this case that of the "facing cracks" of the heritage asset. And Annotation properties of a particular area following the information structure described by the research, in this case, for one of the elements of the descriptor "stonemark". Examples of annotation of the categories: cracks, material discontinuity and spalling are also shown. Model created by the authors in the collaborative web application Aïoli platform®.

carried out, but further progress will be necessary in order to be able to generate a more comprehensive and common database for different underground heritage typologies, including other cases of diverse or specific studies, creating a common information network, contributing to improving the processes of heritage conservation and enhancement.

5. Conclusions

This work specifies and defines the standardisation of the information for the specific typology of "Heritage Underground Spaces" in different categories of the ontological model. These elements, halfway between the typology of a heritage building and that of an archaeological site, have a special difficulty in the standardisation of their information, for which a specific data ontology has been designed based on CIDOC CRM and CRM dig, proposing and following several staged processes: the definition of a specific methodology for the documentation of this type of heritage, the acquisition of specific characteristics and three-dimensional models, and the annotation of data on them. This research creates a workflow using previously described ontology diagrams, which makes it possible to standardise a complete heritage documentation and data compilation process, without the main objective of this work being the design of new ontology diagrams.

The work project advances in a wide field of work within the possible typologies of underground heritage, addressing and including very different elements of underground infrastructure, such as cisterns, crypts, burial sites and other engineering and architectural productions, where the cavities are morphologically shaped by surfaces of complex geometry, the result of the extraction of material from the subsoil. This documentation process improves the management of all heritage information, even if it is diverse or of a different nature, allowing it to be collected in a single digital reference.

The use of the Aïoli web application, in addition to unifying the storage of all types of heritage information, facilitated the legibility of the data, thanks to its web design and easy access, and simplified its reading, which was often complex even for all the agents involved in the heritage. This specific strategy reinforces the protection and safeguarding of heritage, laying the foundations for future research into the enhancement of the various underground built heritage spaces.

Acknowledgements

Research supported by FPU Program of the Spanish Ministry of Science, Innovation and Universities (FPU19/05432). It is also funded by the Research for the Internationalisation Aid (IUACC 2022) from the VII PPIT-US of the University of Seville. The corresponding author also acknowledges MAP-MAACC laboratory (ENSA Paris-La Villette), for the research period and the support provided.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.culher.2024.06.003.

References

[1] F. Amores Carredano, Carta arqueológica de los alcores (Sevilla)., diputación provincial de sevilla, 1982.
[2] A.L. Ampliato Briones, J.C. Rodríguez Estévez, La obra nueva de Santa María, Crónica de una construcción, La Obra Gótica St. María Carmona Arquit, y ciudad en la transic, a La Edad Mod 2017 (2017) 173–212 ISBN 978-84-472-1895-0, Págs173–212.
[3] J.C. Rodríguez Estévez, A. Ojeda Barrera, La historia de Santa María a través de los documentos, in: La Obra Gótica St. María Carmona Arquit (Ed.), y Ciudad En La Transic. a La Edad Mod. 2017, Universidad de Sevilla, 2017, pp. 69–119. ISBN 978-84-472-1895-0, Págs. 69-119.
[4] H. Herrero Fernández, J. García Talegón, Á. Justo, J.L. Pérez Rodríguez, A.C. Íñigo, La Calcarenita de Carmona, Cueva de la Batida: estratigrafía, mineralogía y sistema poral | DIGITAL.CSIC, cardel carmona, Rev. Estud. Locales. 6 (2008) https://digital.csic.es/handle/10261/17144. (accessed 14 March 2022).
[5] R. Varriale, Underground built heritage": a theoretical approach for the definition of an international class, Heritage 4 (2021) 1092–1118, doi:10.3390/HERITAGE4030061.
[6] G. D'Agostino, M. Figuera, G. Russo, M. Galizia, P.M. Militello, Integrated 3d survey for the documentation and visualization of a rock-cut underground built heritage, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. (2022) 167–174 XLVI-2/W1-2022, doi:10.5194/ISPRS-ARCHIVES-XLVI-2-W1-2022-167-2022.
[7] E. Acosta, F. Spettu, F. Fiorillo, A procedure to import a complex geometry model of a heritage building into bim for advanced architectural representations, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 46 (2022) 9–16, doi:10.5194/ISPRS-ARCHIVES-XLVI-2-W1-2022-9-2022.
[8] A. Galera-Rodríguez, R. Angulo-Fornos, M. Algarín-Comino, Survey and 3D modelling of underground heritage spaces with complex geometry: surface optimisation for association with HBIM methodology, SCIRES-IT - Sci. Res. Inf. Technol. 12 (2022) 177–190, doi:10.2423/I22394303V12N1P177.
[9] M. Doerr, The CIDOC CRM, an ontological approach to schema heterogeneity, DROPS-IDN/35. (2005). https://doi.org/10.4230/DAGSEMPROC.04391.22.
[10] M. Doerr, M. Theodoridou, CRM dig: a generic digital provenance model for scientific observation, TaPP'11, 3rd USENIX Work. Theory Pract, Provenance, Heraklion, Crete, Greece, 2011.
[11] M.T. Biagetti, An Ontological Model for the Integration of Cultural Heritage Information, CIDOC-CRM, 2016.
[12] A. Tibaut, S.G. de Oliveira, A framework for the evaluation of the cultural heritage information ontology, Appl. Sci. 12 (2022), doi:10.3390/app12020795.
[13] G. Bruseker, N. Carboni, A. Guillem, Cultural heritage data management: the role of formal ontology and CIDOC CRM, (2017) 93–131. https://doi.org/10.1007/978-3-319-65370-9_6.
[14] N. Amico, P. Ronzino, A. Felicetti, F. Niccolucci, Quality management of 3D cultural heritage replicas with CIDOC-CRM, Conf. Theory Pract. Digit. Libr. (TPDL 2013), 2013.
[15] M. Van Ruymbeke, P. Hallot, G.A. Nys, R. Billen, Implementation of multiple interpretation data model concepts in CIDOC CRM and compatible models, Virtual Archaeol. Rev. 9 (2018) 50–66, doi:10.4995/VAR.2018.8884.
[16] E.M. Sanfilippo, B. Markhoff, P. Pittet, Ontological analysis and modularization of CIDOC-CRM, Front. Artif. Intell. Appl. (2020), doi:10.3233/FAIA200664.
[17] L. Castelli, A. Felicetti, F. Proietti, Heritage science and cultural heritage: standards and tools for establishing cross-domain data interoperability, Int. J. Digit. Libr. 22 (2019) https://link.springer.com/content/pdf/10.1007/s00799-019-00275-2.pdf. (accessed 14 March 2022).
[18] G. Faraj, A. Micsik, Representing and validating cultural heritage knowledge graphs in CIDOC-CRM ontology, Futur. Internet 13 (2021) 277, doi:10.3390/FI13110277.
[19] V. Ducatteeuw, Developing an Urban gazetteer: a semantic web database for humanities data, in: Proc. 5th ACM SIGSPATIAL Int. Work. Geospatial Humanit. GeoHumanities 2021, 2021, doi:10.1145/3486187.3490204.
[20] T. Messaoudi, P. Véron, G. Halin, L. De Luca, An ontological model for the reality-based 3D annotation of heritage building conservation state, J. Cult. Herit. 29 (2018) 100–112, doi:10.1016/J.CULHER.2017.05.017.
[21] G. Tan, T. Hao, Z. Zhong, A knowledge modeling framework for intangible cultural heritage based on ontology, in: 2009 2nd Int. Symp. Knowl. Acquis. Model. KAM 2009, 1, 2009, pp. 304–307, doi:10.1109/KAM.2009.17.
[22] D. Pitzalis, F. Niccolucci, M. Cord, Using LIDO to handle 3D cultural heritage documentation data provenance, Proc. - Int. Work. Content-Based Multimed. Index. (2011) 37–42, doi:10.1109/CBMI.2011.5972517.
[23] J. Padfield, K. Kontiza, A. Bikakis, A. Vlachidis, Semantic representation and location provenance of cultural heritage information: the national gallery collection in London, Heritage 2 (2019), doi:10.3390/heritage2010042.
[24] A. Pinto, Y. Cardinale, I. Dongo, R. Ticona-Herrera, An ontology for modeling cultural heritage knowledge in urban tourism, IEEE Access 10 (2022) 61820–61842, doi:10.1109/ACCESS.2022.3179664.
[25] B. Ranjgar, A. Sadeghi-Niaraki, M. Shakeri, S.M. Choi, An ontological data model for points of interest (POI) in a cultural heritage site, Herit. Sci. 10 (2022), doi:10.1186/S40494-021-00635-9.
[26] T. Wang, L. Zhou, Application of oblique photography and gis technologies in the integrated conservation and development of historic cities in China: practices in Shigatse, Tibet and Quanzhou (Zayton), Fujian, in: Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch, International Society for Photogrammetry and Remote Sensing, 2019, pp. 1203–1208, doi:10.5194/isprs-archives-XLII-2-W15-1203-2019.
[27] X. Wang, Y. Gong, D. Myers, S. Wang, Arches dunhuang: heritage inventory system for conservation of grotto resources on the Gansu section of the silk road in China, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 46 (2021), doi:10.5194/isprs-archives-XLVI-M-1-2021-837-2021.
[28] S. Mirabaud, A. Stoleru, C. Bosc-Tiessé, N. Carboni, F. Guéna, Vers la formalisation d'un modèle ontologique pour la description des églises rupestres peintes d'Éthiopie, Hist. l'art. (2021).
[29] L. de Luca, Aïoli, une plateforme d'annotation sémantique 3D pour la documentation collaborative d'objets patrimoniaux, Cult. Rech. (2018) https://hal.archives-ouvertes.fr/hal-03626588. (accessed 27 November 2022).
[30] A. Manuel, V. Abergel, Aïoli, a reality-based annotation cloud platform for the collaborative documentation of cultural heritage artefacts, Un Patrim. Pour l'avenir, Une Sci. Pour Le Patrim, 2022 https://hal.science/hal-03819923 (accessed 10 January 2023).
[31] R. Roussel, L. De Luca, An approach to build a complete digital report of the notre dame cathedral after the fire, using the aioli platform, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. (2023) 1359–1365 XLVIII-M-2–2023, doi:10.5194/ISPRS-ARCHIVES-XLVIII-M-2-2023-1359-2023.
[32] V. Croce, G. Caroti, L. De Luca, A. Piemonte, P. Véron, Semantic annotations on heritage models: 2d/3d approaches and future research challenges, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch (2020), doi:10.5194/isprs-archives-XLIII-B2-2020-829-2020.
[33] R. Roussel, M. Bagnéris, L. De Luca, P. Bomblet, A digital diagnosis for the « Autumn » statue (Marseille, France) : photogrammetry, digital cartography and construction of a thesaurus, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42 (2019) 1039–1046, doi:10.5194/ISPRS-ARCHIVES-XLII-2-W15-1039-2019.
[34] R. Angulo-Fornos, M. Castellano-Román, HBIM as support of preventive conservation actions in heritage architecture. experience of the renaissance quadrant Façade of the Cathedral of Seville, Appl. Sci. 10 (2020) 2428 2020, Vol. 10, Page 2428, doi:10.3390/APP10072428.
[35] D.F. García-Molina, R. González-Merino, J. Rodero-Pérez, B. Carrasco-Hurtado, Documentación 3D para la conservación del patrimonio histórico: el castillo de Priego de Córdoba, Virtual Archaeol. Rev. 12 (2021) 115, doi:10.4995/var.2021.13671.
[36] P.M. Millán-Millán, Investigaciones en el proyecto de restauración de la antigua sala capitular y panteón del monasterio de Santa Clara de la Columna de Belalcázar (Córdoba), Arqueol. La Arquit. (2021) e121–e121, doi:10.3989/ARQ.ARQT.2021.013.
[37] P. Clini, R. Nespeca, R. Angeloni, R. Mammoli, The integrated survey of narrow spaces and underground architecture: the case study of campana caves bas-reliefs, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. (2019), doi:10.5194/isprs-archives-XLII-2-W9-277-2019.
[38] A. Galera Rodríguez, Documentación de elementos patrimoniales subterráneos de geometría compleja, a partir de modelos de información, in: F. Pinto Puerto (Ed.), Model. Digit. Inf. En La Tutela Sosten. Del Patrim. Cult. Contrib. Al Conoc. e Innovación Soc, Universidad de Sevilla, 2020, pp. 228–240. ISBN 9788447230631.
[39] E. Nocerino, F. Menna, E. Farella, F. Remondino, 3D virtualization of an underground semi-submerged cave system, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42 (2019) 857–864, doi:10.5194/ISPRS-ARCHIVES-XLII-2-W15-857-2019.

[40] A. Manuel, A.A. M'Darhri, V. Abergel, F. Rozar, L. De Luca, A semi-Automatic 2D/3D annotation framework for the geometric analysis of heritage artefacts, in: Proc. 2018 3rd Digit. Herit. Int. Congr. Digit. Herit. 2018 - Held Jointly with 2018 24th Int. Conf. Virtual Syst. Multimedia, VSMM 2018, 2018, doi:10.1109/DIGITALHERITAGE.2018.8810114.

[41] P. Ferreira-Lopes, F. Pinto-Puerto, GIS and graph models for social, temporal and spatial digital analysis in heritage: the case-study of ancient Kingdom of Seville Late Gothic production, Digit. Appl. Archaeol. Cult. Herit. 9 (2018) e00074, doi:10.1016/J.DAACH.2018.E00074.