Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782 Start Date of Project: 01/01/2019 Duration: 40 months

Deliverable 5.18

Report on the archaeological case study

Dissemination Level	PU
Due Date of Deliverable	31/03/2022, M39
Actual Submission Date	02/05/2022
Work Package	WP5 - Innovations in Data Access
Task	Task 5.7 Open Linked Data. Archaeology Case Study
Type	Report
Approval Status	Waiting EC approval
Version	V1.0
Number of Pages	p.1 – p.66

Abstract:

This deliverable is the final report on the archaeological case study in SSHOC Task 5.7. A virtual reconstruction of the Roman theatre in Catania has been created as a case study for the transition of archaeological data to the cloud, i.e. from data silos on individual computers to webservices. The case study is based on a unified workflow that starts with the archaeological documentation and results in a virtual reconstruction, embedded in a Linked Open Data scenario.

SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the EU Horizon 2020 Research and Innovation Programme (2014-2020); H2020-INFRAEOSC-04-2018, under the agreement No. 823782

History

Version	Date	Reason	Revised by
0.1	15/04/2022	First draft version	all
0.2	29/04/2022	Second draft version	all
0.3	29/04/2022	Peer review	Holly Wright (ADS)
1.0	30/04/2022	Address peer review comments	Wolfgang Schmidle

Author List

Organisation	Name	Contact Information
DAI	Wolfgang Schmidle	wolfgang.schmidle@dainst.de
CNR (ITABC)	Emanuel Demetrescu	emanuel.demetrescu@cnr.it
CNR (ITABC)	Bruno Fanini	bruno.fanini@ispc.cnr.it
CNR (IBAM)	Alberto Bucciero	alberto.bucciero@cnr.it
CNR (IBAM)	Alessandra Chirivì	alessandra.chirivi@ispc.cnr.it
CNR (IBAM)	Ivan Ferrari	ivan.ferrari@cnr.it
CNR (IBAM)	Francesco Giuri	francesco.giuri@cnr.it
Politecnico di Torino	Elisabetta Caterina Giovannini	elisabettacaterina.giovannini@polito.it

Executive Summary

In SSHOC Task 5.7 (Open Linked Data. Archaeology Case Study), a virtual reconstruction of the Roman theatre in Catania has been created as a case study for the transition of archaeological data to the cloud, i.e. from data silos on individual computers to webservices. The case study is based on a unified workflow that starts with the archaeological documentation and results in a virtual reconstruction. With this workflow, data manually acquired during an excavation and traditionally stored on paper can now be stored in the cloud and used for 3D visualisations of the site.

Part of the data for the Roman theatre comes from earlier projects, where it was transferred from proprietary formats to open formats and open systems, and part of the data has been created in this task. This represents a useful exemplar of using both new and legacy data in a single project.

The workflow uses tools that have been developed by the task partners, in particular the Extended Matrix for 3D reconstruction. The task partners have documented the existing systems and their individual workflows and applied them to existing as well as newly created data on the Roman theatre.

The case study also relies on systems for normative data for place and time information, and based on the case study the task partners have evaluated additional data standardisation strategies to enable data sharing and re-use of archaeological data. To this end, the underlying data models have been aligned with CIDOC CRM. In addition, the 4D virtualisation has been embedded in a Linked Open Data scenario, which creates several entry points for examining, comparing and searching the 3D model.

Abbreviations and Acronyms

API	Application programming interface, used by computer programs to communicate with each other
CIDOC CRM	CIDOC Conceptual Reference Model
CRMarchaeo	an Extension of CIDOC CRM that models the archaeological excavation process
EM	Extended Matrix, developed by ISPC Rome
GraphML	Graph Markup Language, an XML-based data format for storing semantic graphs
LOD	Linked Open Data (aka Linked Data, Open Linked Data)
RM	representation model, contains 3D modelling details
SFM	Structure from Motion
SSH	Social Sciences and Humanities (i.e. the SSH in SSHOC)
SSHOCro	SSHOC Reference Ontology
STV	spacetime volume
UAV	Unmanned Aerial Vehicle (aka drone)
XML	Extensible Markup Language, used for saving structured data in plain text files
yED	an editor for drawing diagrams

Table of Contents

1.Introduction

Archaeological field work is a complex process in which many scientific resources and competences are allocated for long periods of time, sometimes spanning years or even decades. However, the data is often either not available in digital form or not public. For this reason, the sustainability of archaeological research would benefit immensely from optimised and effective systems of digital documentation, analysis and visualisation of archaeological sites.

The SSHOC task 5.7 (Open Linked Data. Archaeology Case Study) presents an example of a transition of archaeological data to the cloud. A workflow from archaeological documentation to virtual reconstruction is described, i.e. how data manually acquired during an excavation and traditionally stored on paper can be stored in the cloud by every archaeologist working on the same site and subsequently used for 3D visualisations of the site. For this workflow, the ancient Roman theatre in Catania was chosen as a case study.

The tools for this workflow, in particular the Extended Matrix (EM) for 3D reconstruction, have been developed by the task 5.7 partners before SSHOC. The task's objective was to combine these tools and develop them further, make the tools easier to use and offer a cloud perspective for sharing and re-using data. New survey data of the Roman theatre and new authority records for contextualising the Roman theatre were created for this task, and the EM was complemented with the *EMviq* online tool.

One focus was on data transparency. The graphic representation of reconstructive hypotheses suffers from the so-called Black box effect, i.e. a user's cognitive difficulty to appreciate which parts of a reconstructive hypothesis are based on concrete data (archaeological remains) and which on less reliable sources, and which parts have been added with a merely evocative purpose. Solving this cognitive semantic ambiguity of construction models can drastically increase the communicative and educational strength of virtual reconstructions for Cultural Heritage. The focus on transparency also motivated the implementation of the FAIR principles, as well as the CIDOC CRM mappings, especially tracking the data provenance with the CIDOC CRM extension SSHOCro. The task work was used as a case study for SSHOCro. In addition, the 4D virtualisation has been embedded in a Linked Open Data scenario. This creates several entry points for examining, comparing and searching the 3D model. It is possible to search for concrete objects in the model, for other 3D models based on the area, the people involved or the time period, as well as CIDOC CRM classes, properties and modelling paths.

The task partners are the German Archaeological Institute (DAI) in Berlin and two institutes of the Istituto di Scienze del Patrimonio Culturale (CNR-ISPC) in Rome and in Lecce. They will be referred within this document as DAI Berlin, ISPC Rome and ISPC Lecce.[1] The work in this task required domain-specific knowledge about archaeology and archaeological excavations, IT knowledge for

[1] In October 2019, i.e. early on in the SSHOC project, ISPC Rome (CNR-ITABC) and ISPC Lecce (CNR-IBAM) were merged with two other CNR institutes to form the CNR-ISPC Institute of Heritage Science.

creating domain-specific software and computer-based workflows, experience with 3D modelling and reconstruction as well as knowledge about using and creating authority records (aka norm data) and long-term preservation of data. Based on this expertise, task 5.7 brings together the workflow of ISPC Rome for generating 3D visualisations from excavation and survey data, the norm data systems and the experience with Linked Open Data and ontologies of the DAI Berlin and the excavation and survey data and archaeological expertise of the ISPC Lecce.

This task has evolved signifianctly since its initial description in the SSHOC grant agreement, which we think is a positive. The initial plan was to identify an archaeological community and to create a software module to help them move their existing data into the cloud, as is evident in the original description of the first deliverable (D5.17): "Plan for the software module for entering and storing the data, based on community input and findings and implementations from project partners, and new functionality". The idea that the DAI Berlin group had in mind when formulating the task description was to improve and standardise the workflow for archaeometric lab data on bone analysis. However, early discussions made it clear that the task would be more efficient if we focussed on the task member competencies to achieve a new goal that the individual groups couldn't have achieved on their own. Consequently, the focus of the first deliverable as well as the whole task had shifted from creating a software module from scratch to connecting and improving selected existing systems. This change in direction was described in D5.17.

In D5.17, several concrete work steps were listed. Work step 1 was documentation and support. DAI Berlin has created an extensive documentation on the DAI's ChronOntology temporal gazetteer, and ISPC Rome has documented its EM system. We have also presented our work on various occasions, including an online workshop in May 2021. Work steps 2 and 3 were to prepare the data from the Roman theatre and to create the virtual reconstruction. This was the main body of work in this task. The data, as well as the process of preparing the data, is described in chapter 3, the virtual reconstruction is described in chapter 4. Work step 8 was a cloud-based multi-user EM system. This has been reached with the development of the EMviq tool, which is described in chapter 4.3.

Work steps 4 and 5 were the use of systems for authority records and "LOD, interoperability and standardisation". These two topics have merged into the Linked Open Data scenario described in chapter 5 (contextualisation) and 7 (implementation). Work step 6 was to connect the data models via their CIDOC CRM mappings. We have started aligning the mapping (see chapter 6), but as is typically the case, it is an ongoing process. We will continue our efforts after the end of the SSHOC project.

Work step 7 was to connect idai.field and EM. The integration with excavation data in iDAI.field was one thing that we did not pursue. For now, we have decided against creating fictitious data in idai.field. However, the have started a mapping between the idai.field output format and EM input format that forms the basis for exchanging data.

Covid-19 took a heavy toll on the task. During the whole project the project partners were never able to meet in person. In particular, the DAI partner in Berlin was not able to travel due to restrictions imposed by the DAI, as well as nursing a relative for whom Covid had to be strictly avoided. We hope to finally meet in person later in 2022.

The present document is the second of two deliverables (D5.18) in task 5.7. It reports on the final implementation of the work that was outlined in D5.17. An earlier version of the Linked Open Data was already outlined in Milestone MS36 report, and an earlier version of the alignment of the different CIDOC CRM mappings was described in Milestone MS37 report.

2. The Roman theatre in Catania

2.1 Selecting the case study

The Roman theatre in Catania was selected for this case study for several reasons. For one, we decided that a certain level of preservation is preferable for testing the workflow we wanted to create. The Roman theatre is certainly one of the best-preserved examples of this type of Roman architecture, and even here the characteristics represent peculiarities that have still to be fully documented. The problem faced by an interpretative and reconstructive study essentially lies in the absence of accurate and up to date graphic documentation. To remedy this lacuna, ISPC Lecce has undertaken a survey campaign using different techniques and tools (see below).

In addition, this archeological site presents evidence of different construction phases, which allows a focus on temporal data:

- Greek phase: wall structures built with large sandstone blocks with Greek letters in the V-VI century BC;
- first Roman phase: This period sees a remaking of the previous wall structures, the *cavea* and the construction of the *scena* using square lava blocks;
- second Roman phase: Between the Antonine and Severan periods, the structure was extended and underwent transformations that gave the entire complex a monumental aspect. The stage building was substantially modified by the insertion of curving exedra into the rear wall of the side doors and the enrichment of the decorative scheme, while the *cavea* was enlarged by the addition of the third walkway and creation of a new seating tier. The theatre was decorated with marble columns, statues, and decorative reliefs celebrating public events associated with mythological themes.

The quantity and variety of available data, derived from the use of different techniques, tools and methods, were integrated to obtain an exhaustive representation of architectural geometry and its specific materiality, also contributing to making this case study particularly interesting. Thus, the Roman theatre provided the chance to fully test and validate the Extended Matrix scientific approach as well as the DAI authority record systems.

The selected case study also offered the opportunity to reflect on the crucial issue of the so-called scientific transparency of virtual archaeology projects. To achieve a high scientific rigour in a reconstructive archaeology project, it is essential to collect documentation and transparently present the entire work process. This allows the validation of the reconstruction and enables other researchers to review the results without necessarily starting a new study from scratch.

Finally, the case study of the Roman theatre in Catania allowed the ISPC Lecce team to compare the 3D reconstruction already carried out by the Lecce Laboratory with the reconstructive hypothesis that they developed starting from the sources organised in the EM, offering a term for comparison and validation of the work done.

2.2 History of the Roman theatre

The Graeco-Roman theatre of Catania stands in the heart of the historic centre, on the south-eastern slopes of the hill of Montevergine, within an area that played a central role in the dynamics of the city's urban development. The original structure seems to date to the Greek period, as suggested by the remains of the walls in limestone blocks identified in several sectors of the monument.

The building visible today was built during the Julio-Claudian period as part of a programme that saw the rebuilding of the monument, which probably reused structures and materials from the earlier Greek theatre. Between the Antonine and Severan periods, the structure was extended and underwent transformations that gave the entire complex a monumental aspect. The stage building was substantially modified by the insertion of curving exedra into the rear wall of the side doors and the enrichment of the decorative scheme, while the *cavea* was enlarged by the addition of the third walkway and creation of a new seating tier. The theatre was decorated with marble columns, statues, and decorative reliefs celebrating public events associated with mythological themes. During the final period of use, which can be dated to around the 4th century AD, the building was probably adapted for aquatic shows, with the creation of channels and systems for filling it and draining water, work which also involved the repaving of the orchestra.

Between the 5th and 6th centuries AD the theatre lost its original function, as attested by the building of structures in the orchestra and gradual filling of the monument's lower part. From this moment onwards, the building slowly deteriorated, beginning with the robbing that took place throughout the Medieval period and then with the gradual obliteration of the ancient structures caused by the unstoppable building activity, which between the late Medieval and modern periods determined the formation of a residential quarter known as "Grotte" in the theatre area (Malfitana et al. 2016).

2.3 Earlier projects

The Lecce ISPC already participated in previous projects on the same issues that paved the way for this task. Some of the main initiatives are summarised here.

2.3.1 The OpenCiTy project

The case study of the Roman theatre of Catania was addressed for the first time in 2015 within the OpenCity project. This project was part of the wider PON Smart Cities DICET-INMOTO[2].

The OpenCiTy project, on which the former Institute for Archaeological and Monumental Heritage of the National Research Council worked (now part of ISPC), aimed at creating a platform capable of producing, storing, managing and sharing a heterogeneous mass of information relating to the city, favouring the development of shared knowledge with the entire community, which is increasingly called upon to play an active role in decision-making processes.

What OpenCiTy set out to create was therefore a powerful and versatile tool to support the needs of research, protection and enhancement of the territory. The core of the project consisted of a relational database, specifically designed to interface with a GIS platform, through which it became possible to manage, analyse and interpret, on a geospatial basis, the data produced. All sectors and disciplinary fields, which contribute to the definition of a complex historical organism, such as Catania, have been considered, with the collection and full integration, at a very high level of detail, of archaeological, historical-artistic, geological datasets, urban planning.

To this mass of data, coming from archival research, was added that derived from specific campaigns conducted by the former IBAM-CNR in sectors of the city and monumental contexts with high informative potential, investigated with the use of instruments and methodologies not invasive (Georadar, Geoelectric) or with high-precision survey techniques (laser scanner, photogrammetry from the ground and via drone), preparatory to the reconstruction of 3D models and immersive galleries for study and enhancement. All the data produced within the OpenCiTy project have been made usable through a WebGIS platform, which in addition to access and download will provide the user with powerful query tools, predictive analysis and decision support.

Under the OpenCiTy project, the former IBAM-CNR conducted and finalised the reconstructive study on one of the most representative monuments of the city of Catania (Sicily), which constitutes a very precious document of the ancient splendour of city and at the same time the direct witness of the complex its history, constantly crossed by great cultural changes and urban transformations.

[2] See: https://www.cnr.it/it/evento/13344/workshop-progetto-pon-smart-cities-cultura-e-turismo-dicet-inmoto [April 2022]

The Roman theatre of Catania stands on the southern side of the Montevergine hill in the centre of a district rich in archaeological remains that still represents the beating heart of the city today. The original building probably dates back to the Greek age but only a few wall remains are preserved.

The researches of the former IBAM CNR were focused on the most monumental architectural phase, that of the II-III century. A.D., and made it possible to reveal, between hypotheses and confirmations, an unprecedented face of the ancient building, in its moment of greatest splendour.

2.3.2 The “PON Energia Smart City” project

Within the "PON Energia Smart City" project the ITLab (Information Technologies Lab) of IBAM CNR Institute implemented visualisation output for mobile devices adopting technologies based on Augmented Reality (AR). (This section is adapted from Gabellone et al., 2014.)

These technologies allow creating an overlap between the real experience and the virtual information (multimedia information, geo-located data, analytical data, and so on) in an environment wherein the multimedia elements “increasing” information on reality can be progressively added, superimposed, and displayed through a “direct vision” approach.

There are several ways to view in AR mode. The most classic form presents a simple superimposition of information directly displayed on the framed object. Texts, images, and other information appear directly on the framed object, but in some implementations of graphics libraries, it is possible to overlap simplified three-dimensional models that help to understand the archaeological structures within the urban fabric.

The idea of linking 3D models to the real environment has been demonstrated for a long time, even in the manufacturing , automotive, and fashion industries. Moreover, many research groups have developed solutions that allow contextualising 3D objects directly using a smartphone. The use of these resources, however, is strongly affected by the limitations of the various computational devices, primarily due to poorly performing equipment. One solution to this limitation is given by a hybrid mode of AR, in which ultra-realistic, three-dimensional reconstructions are mixed with high-resolution spherical VR panoramas. Many sceptics and lovers of performing technologies lose sight of these old solutions, well supported in HTML5, which have effective communication, ease of use and excellent visualisation quality.

The basic idea is very simple. The three-dimensional model of the historical structure is placed in the centre of the spherical panorama, then linking the attachment points to the points detected on the ground and reported in the 3D scene. These attachment points must exactly match those in the landscape. In this way, we can show the 3D object perfectly anchored to the real scene.

Within this process, special attention has been devoted to the lighting of the scene and to the implementation of a set-up that recreates the same environmental conditions present in reality, to provide a convincing result that can be perfectly superimposed on the site. This solution allows the various monuments to be viewed in their original context, in an “optimised” and efficient way, even in particularly complex scenarios. Specifically, this is used in historic structures within urban environments, where the adoption of a simplified 3D model without shadows, without radiosity and with low texture resolution provides poor integration and an unacceptable level of quality in the final result, not to mention the issue of items located at various depths, which partially obscure the reconstruction and are partially obscured by the reconstruction in turn. Using AR, a 3D object solves these problems, plus the problems related to the real-time visualisation. In the solution shown in these pages, the reconstructed three-dimensional model is integrated into the urban fabric through simple masking using various objects placed in-depth, the global illumination, and, last but not least, the point of view of the observer.

In the archaeological district of Neapolis, the work has been focused on the Greek theatre in Siracusa and mainly on the reconstruction of its Hellenistic scene. The structures comprising the great theatre, built by Hieron II, have almost entirely disappeared, except for some fragments of sculpture and architectural elements. On the rocky level where the scene was built, several cuts and holes are visible today, but their exact chronology is difficult to establish. Although extensive literature has dealt with the Greek theatre of Syracuse – which is one of the most famous in the world – we can say that until now a graphic rendering of the site based on scientific data has never been attempted.

The reconstruction hypothesis is based on previous studies of Sicilian minor theatres, probably inspired by the great theatre of Syracuse. So, probably the *proskenion* was characterised by revolving walls spaced by pillars and the scene included two orders, Doric in the lower part, Ionic in the upper part, with walls punctuated by doors and windows. In the two foreparts on the sides of the scene, called *paraskenia*, the architectural sculptures were recovered and are now preserved in the archaeological museum Paolo Orsi in Syracuse, where they have been virtually reconstructed and positioned. Two pairs of Satyr-Telamons have been placed in the lower part of *paraskenia* on the side of the gates that used to allow access to the Orchestra, following the example of the neighbouring and *coeval* altar of Hieron II. The other preserved sculpture is the upper part of a Maenad-Caryatid, which is sculpted in-the-round. Consequently, unlike the Satyr-Telamon, we cannot presume that this Maenad- Caryatid was set very close to a wall. The most reasonable solution was to imagine a loggia crowning the *paraskenia* where a pair of maenads were placed in line with the underlying Satyrs, although this represents an unpublished solution for this kind of monument.

In this work, the whole process of 3D reconstruction of sculptural fragments (the biggest approximately 160x80x50 cm) required only a photographic campaign performed with a high-resolution reflex camera (Canon 5D Mark II, 24 Mpx). In the shots, we maintained a constant focal length (24 mm) and a constant sampling step, to cover the entire surface of the objects and ensure a sufficient overlapping of the images (about 70-80%): a fundamental condition to obtain the tracking points in space and their consequent position in 3D. The frames (a total of 60 shots for sculpture) have been processed with the software PhotoScan of Agisoft, proceeding with the alignment of the shoot, the creation of the point cloud, the meshing, and the processing of the textures.

The software uses flexible algorithms that ensure the correct orientation of the photos, even in the absence of the classic procedures of digital photogrammetry, even without previously calibrating the camera and without any substantial contribution by the human operator. All operations have therefore been automated, making it possible to set the parameters according to the desired quality by defining the number of polygons and the size of the textures.

The 3D models obtained have been optimised with respect to the number of polygons needed (approximately faces 2.000.000) without any meaningful loss of resolution of the textures to be imported into the modelling software. For the restoration of missing parts, in a first phase techniques of polygonal modelling point-to-point and subdivision surfaces were used in order to precisely control the process of creation.

Afterward, techniques of advanced digital sculpting for the characterisation of surfaces have been applied, especially for the reconstruction of the drapery present on one of the two caryatids. The texturing of surfaces has been achieved with digital painting techniques for the creation of UV-maps able to realistically simulate the materials and the original colours.

3. The Data from the Roman Theatre

As already discussed in Deliverable D5.17, to overcome the absence of accurate and updated graphic documentation about the Roman theatre, indispensable for developing the reconstructive hypothesis of the site, a documentation and analysis campaign was conducted using these tools and methodologies:

(A) 3D model by laser scanning;
(B) digital photogrammetry images;
(C) photos;
(D) video shooting with a UAV (Unmanned Aerial Vehicle) and from the ground;
(E) scientific literature.

We start with a detailed description of (A) to (E).

3.1 Description of the data

(A) **3D model created using laser scanning**, with a Leica ScanStation P20. The work *in situ* was set up to cover the entire complex, which due to its substantial size required the planning of a series of stations that would cover the entire external perimeter, the *cavea* and each individual walkway including the single flight of steps connecting them and those of the *vomitoria*. The coverage of all wall surfaces required the use of 39 stations, each set at a resolution of 6 mm on a dome of 10 m. Due to the quantity of data created by the fusion process for each individual scan within a single point cloud, (about 3 billion points – 100 giga .pts) per extrapolation of the mesh, it was necessary to subdivide the files into four parts and work on each one separately (Figure 1).

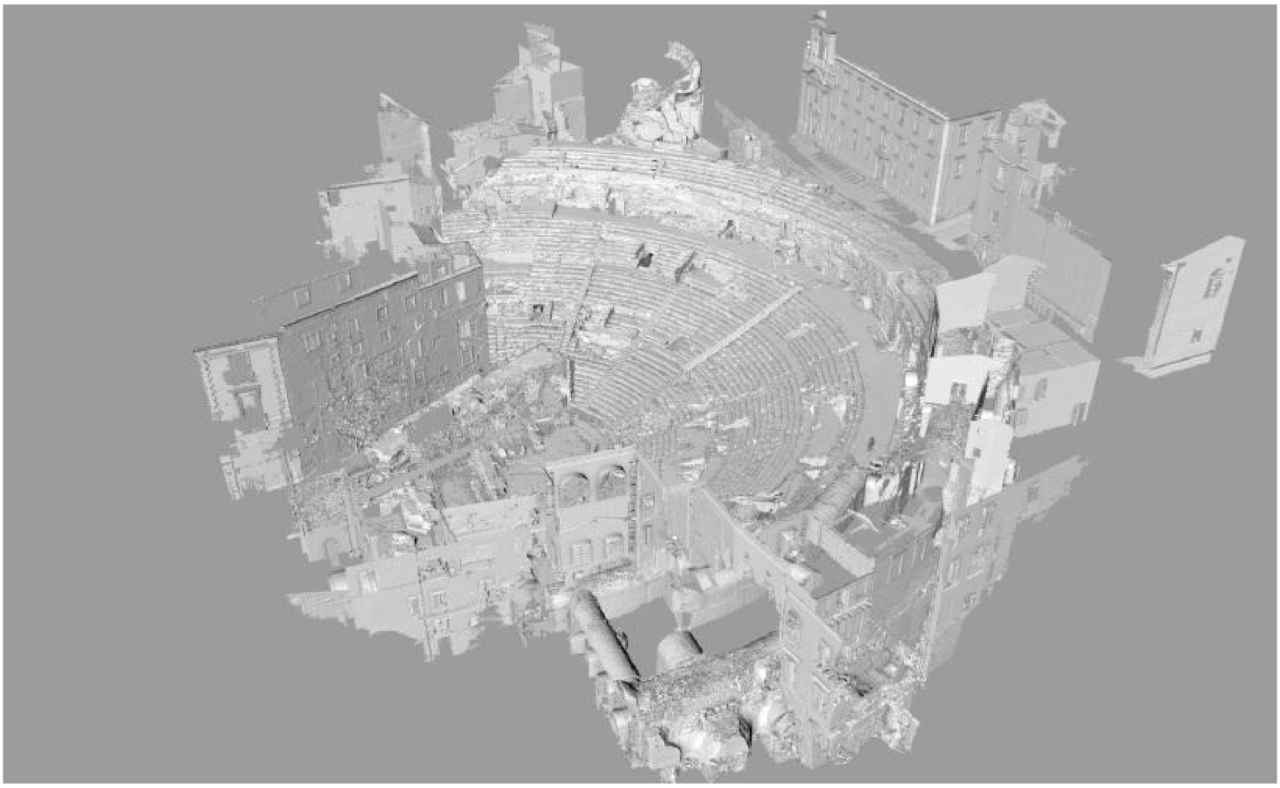

Figure 1: 3D Model from laser scanner

(B) 3D graphic reproductions of of about 300 representative decorative architectural elements preserved within the structure, such as capitals, columns, cornices, pedestals, friezes, and statues, using **digital photogrammetry** (Camera canon 5D Mark II, software for data processing Agisoft Photoscan, Figure 2). Photogrammetry is a technique that uses multiple images of the same area to rebuild the 3D geometry of the objects in an entirely automatic manner with the help of complex algorithms from computer vision, called SFM (Structure from Motion)[3].

[3] Structure from Motion: for introductions see e.g. https://alicevision.org/; [24 March 2022]; https://peterfalkingham.com/blog/; [24 March 2022]

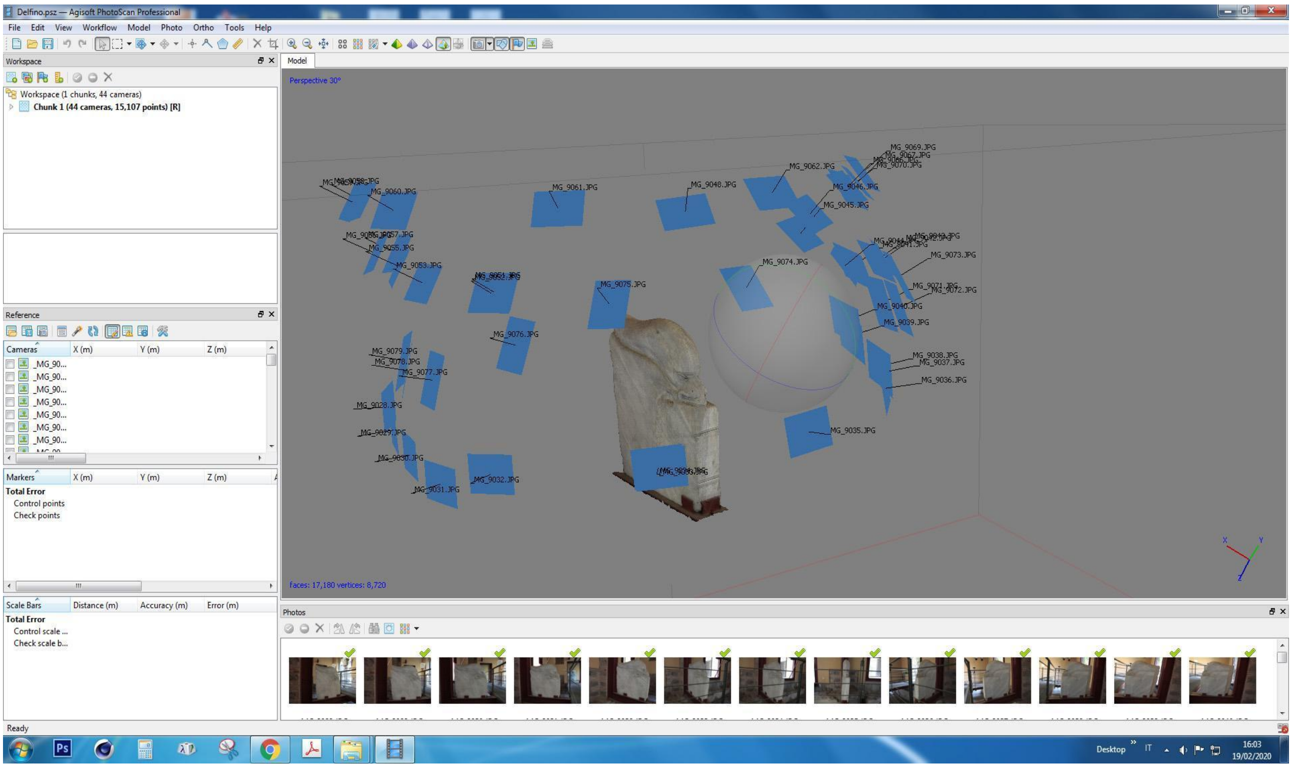

Figure 2: 3D model from digital photogrammetry

(C) **Photographic survey** performed with a Canon Mark II 24 mpx (photos of architectural, decorative, construction techniques, incorporating about 700 images, Figure 3).

Figure 3: Photo of the Roman theatre in Catania (Source: CNR-IBAM, 2015 survey conducted by DiCeT project)

(D) **Video capture** with a UAV (Unmanned Aerial Vehicle) and at ground level (together about 30 videos, Figure 4).

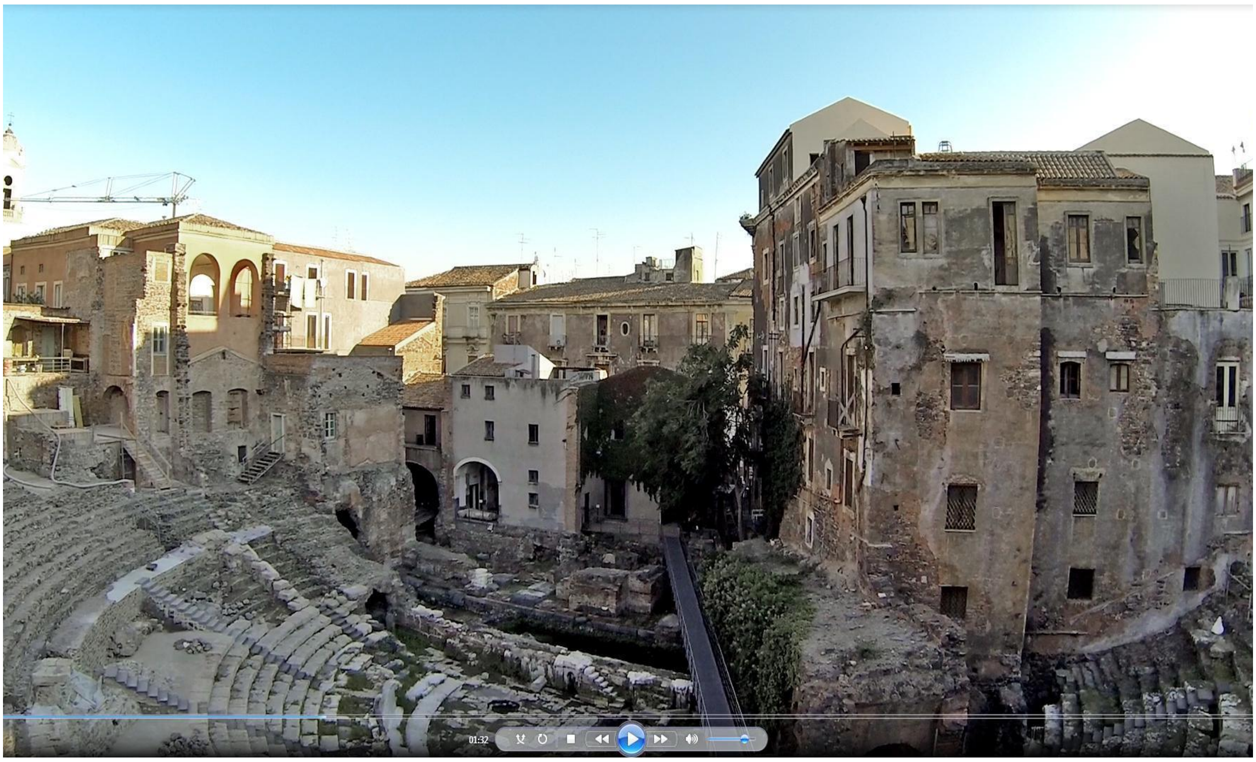

Figure 4: Video of the theatre from UAV; screenshot (Source: CNR-IBAM, 2015 survey conducted by DiCeT project)

(E) In order to advance a reconstructive hypothesis for the theatre, all the documentation produced and analysed was integrated with the data and information inferred from the **scientific literature** on the site (see Malfitana et al., 2016), and with typological comparisons with other contemporary or otherwise worthy monuments of interest.

3.1.1 The Existing Virtual Reconstruction

Within the PON project (see above), the problem related to the absence of accurate and updated graphic documentation of the Catania theatre was addressed. To remedy this lacuna, it was decided to undertake a survey campaign using a laser scanner. The work *in situ* was set up with the aim of covering the entire complex, which due to its substantial size required the planning of a series of stations that would cover the entire external perimeter, the *cavea* and each individual walkway including the single flight of steps connecting them and those of the *vomitoria*. The coverage of all wall surfaces required the use of 39 stations, each set at a resolution of 6 mm on a dome of 10 m. Due to the quantity of data in the fusion process for each individual scan in a single point cloud, (about 3 billion points – 100 giga .pts) per extrapolation of the mesh, it was necessary to subdivide the files into four parts and work on each one separately. In this way, it was possible to generate a 3D model that for practical reasons had to undergo a polygonal decimation, which facilitated the extrapolation of a series of sections, put together in technical images, which then represent the base for the process of interpreting the monument and its three-dimensional reconstruction. A first and important phase was the study and definition of the *cavea's* slope and exact positioning of the heights of the floor levels relating to the walkways and intermediate passageways (Figure 5, Figure 6, Figure 8). No less important was the understanding of the levels at which the spectator access points from the exterior to interior were positioned, and this in an urban morphological context where, precisely in the area of the Roman theatre, considerable differences in height. A third and just as important element was the graphic reproduction of some representative decorative architectural elements such as capitals, columns, cornices, pedestals, friezes, and statues, present within the structure, using an Image Based 3D survey with algorithms from Structure from Motion.

Once the metric characteristics were acquired, and the data cross-checked with the published material, it was thought opportune to link the hypothetical reconstruction the monument's most monumental and richly decorated phase. This construction phase has been dated across a relatively long period, which began in the Antonine period and probably continued until the Severan period (Pensabene 2005, p. 200; Branciforti 2010, p. 208; Buscemi 2012, p. 128). During this phase, the theatre underwent radical restoration involving the enlargement of the entire structure, with the building of the third walkway and creation of a new external façade. Given the monument's technical characteristics, with the *cavea* abutting the hill slope behind it, the new annular walkway could only exploit the natural support provided by the slope in its central part, making it necessary to create two different substructures that from the two sides of the *scaena*, supporting the walkway as far as its central section. This intervention necessitated the modification of one side and a series of interventions on the other side to make the complex comply with Roman architectural canons. In fact, the rebuilding of the *cavea*, attested by the superimposition of the limestone blocks of the first *diazoma* on an earlier *diazoma* presumably dating to the Julio-Claudian period (Buda 2015, p. 269), was probably a direct consequence of the theatre's enlargement. The creation of a new *scaenae frons*, narrower than the preceding one, and a new third seating tier above the pre-existing ones, was probably part of the same restoration project (Figure 7). Based on the archaeological data, the intervention modified the opening of the lateral doors through the creation of a curved *exedra* interrupted by two narrow radial corridors communicating with the rear of the stage (Branciforti-Pagnano 2008, p. 61; Branciforti 2010, p. 204). Regarding the decorative elements, the repertory was notably enriched by the insertion of statues, decorative friezes, and columns in different marbles (Pensabene 2005, pp. 193-212; Branciforti-Pagnano 2008, p. 56), which gave the entire complex a monumental aspect.

The laser scan revealed the presence of a semicircular *cavea* with a maximum width of 97 m, depth of 51 m and overall height that from the orchestra must presumably have reached 26 m, with the seating tiers subdivided into nine *cunei* (wedge-shaped divisions) with eight radial flights of steps. Access to the seating tiers was provided by three orders of *vomitoria*, whose flights of steps departed from the three annular walkways that distributed the flow of spectators within the structure. Another problem to emerge during the reconstruction process related to the points of access to the theatre, strictly related to the various external ground levels. It emerged from the 3D study that most of the spectators must have entered from the north side. In fact, there was direct access from here to the external walkway via a series of arched openings each with a short flight of steps. The *summa cavea* could also be entered from this side, owing to the presence of two distinct masonry-built foreparts abutting the structure, each with a double flight of steps, symmetrically divergent, able to overcome a difference in height of about 8 m. The existence of a third entrance is attested by the presence of a large flight of steps in the north-western sector, providing access to the second walkway (Figure 1).

There was further access to the theatre in the eastern sector, at a notably lower level, 13 m, than those described above. The flow of spectators from the eastern part of the theatre had another external forepart available to it that covered a jump in height of about 11 m to lead into the third walkway. Alternatively, a number of arched passageways led directly into the so-called eastern atrium on the same level. This in turn was linked via a radial passageway directly to the eastern addition for access to the orchestra floor, crossed in the central portion of the lower walkway.

Particular attention was paid to the reconstruction of the stage building, of which only the eastern part is preserved, while the western side remains unexplored as an imposing 19th century palace stands on top of it (Branciforti-Pagnano 2008, p. 56). In the proposed three-dimensional hypothesis – based on the Antonine-Severan period – the foreparts with short longitudinal flights of steps that form the orchestra led up onto the stage, have been omitted as they belonged to a later phase. The orchestra has been re-proposed with its original flooring of bi-chrome marble slabs, shaped to create geometric motifs with inscribed circles inside squares. The study also included the re-contextualisation of the Roman theatre within the urban fabric also characterised by the presence of the Odeon, another Roman building of considerable importance whose remains are still visible in the west area of the theatre (Figure 9).

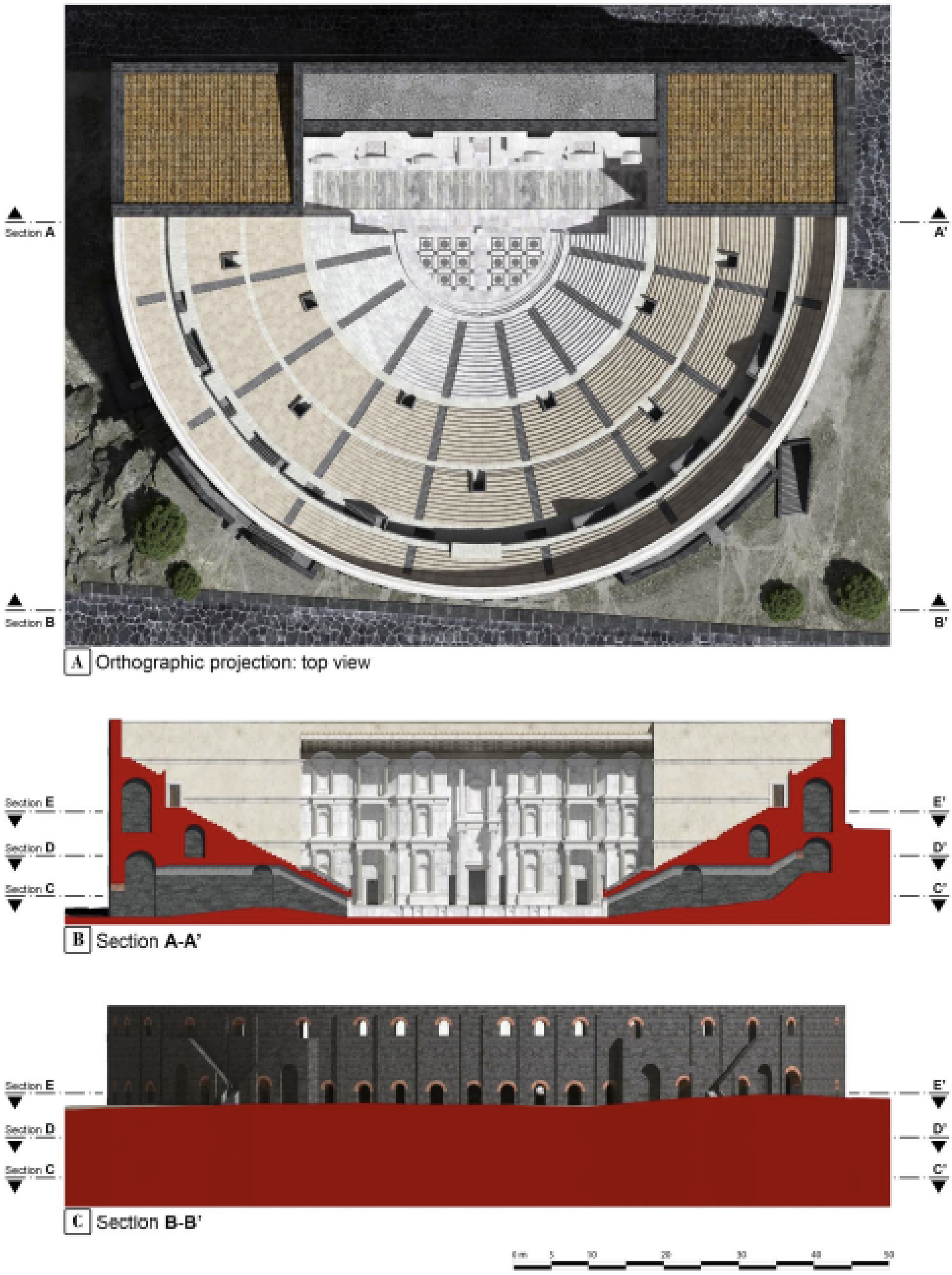

Figure 5: 3D model of the hypothetical reconstruction of the Roman theatre of Catania: Orthographic views (A) and sections (B, C)

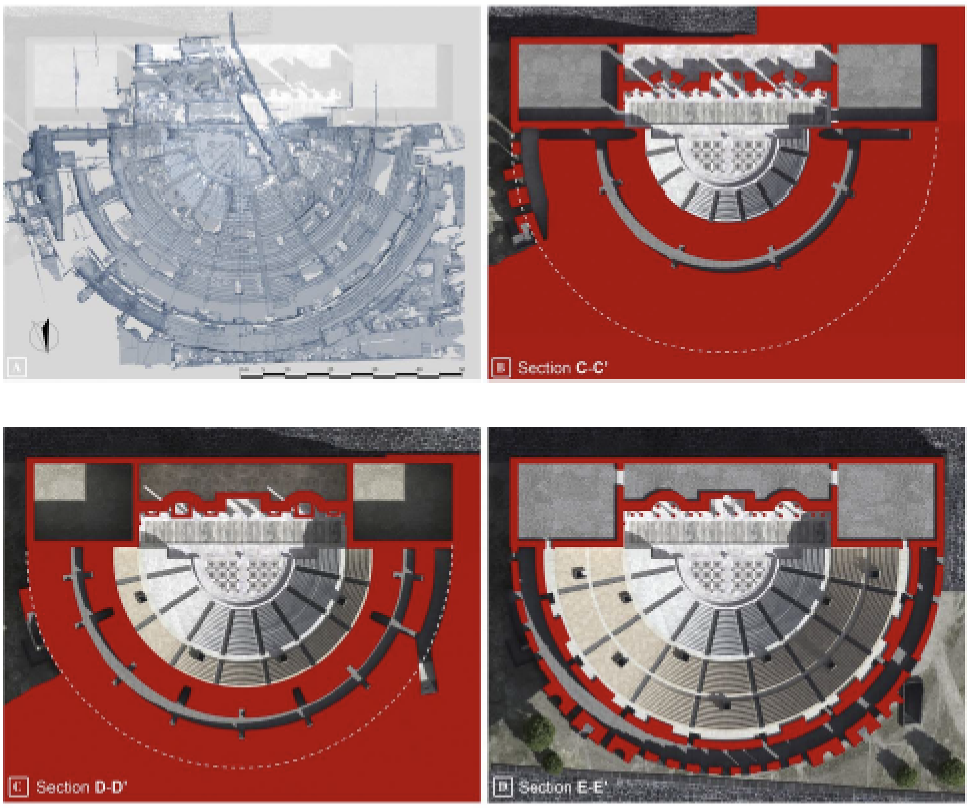

Figure 6: Orthographic views of the laser survey (A), planar sections of the Roman theatre 3D model (B, C, D)

Figure 7: Detail of the scaenae frons *3D reconstruction*

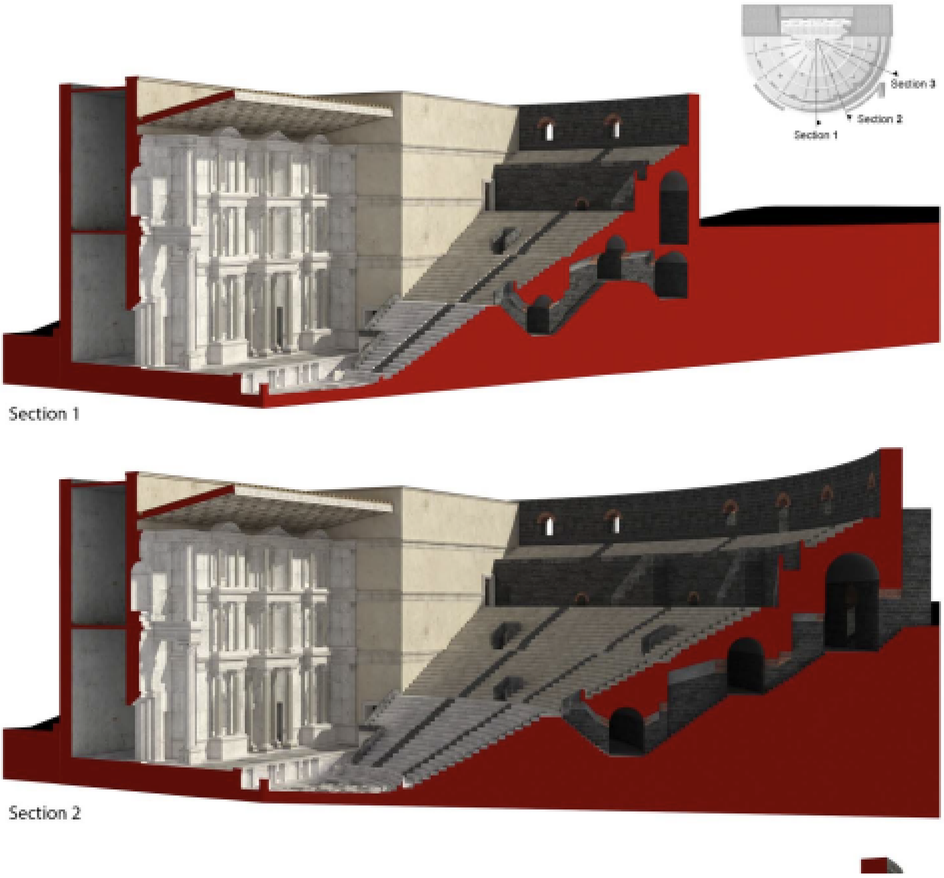

Figure 8: Vertical sections of the 3D model hypothetical reconstruction

Figure 9: 3D reconstruction of the ancient Roman city of Catania: panoramic view of the theatre and Odeon

3.1.2 Problems using the previous process

Even if the result of the previous process of surveying, reconstruction and documentation, described in the previous section, was excellent and the quality of the virtual reconstruction was undoubtedly very high, the whole process was affected by a range of problems:

- the raw data collected during the survey campaign was stored on various media (in physical and virtual drives) without taking care to create a proper data structure to preserve their logic and semantic meaning over time;
- the intermediate data produced during the virtual reconstruction process was not appropriately labelled;
- the entire reasoning process was described in a single scientific publication, which, although of value, describes the decisions taken in a narrative and non-specific way;
- no metadata has been created to contextualise the data collected and resulting from the various post-processing stages;
- some of the technological tools used are obsolete today.

The main problem was the loss of the intermediate steps leading to the final result, as well as the scientific "narrative": unfortunately several reconstructive processes remained in the mind of the archaeologists.

3.2 Activities carried out in ISPC Lecce

At ISPC Lecce, the research activities have been focused on digital workflows and the case study of the Roman theatre in Catania. The research group in Lecce (1) has shared with the task 5.7 partners the different data used and produced by the former ITLab (now DHILab) to present its existing reconstructive hypothesis of the Roman Theatre in Catania; (2) has Export 3d former model and texture from C4D to Blender in OBJ and Jpg format in order to make the data more FAIR; (3) has used the Extended Matrix approach to formulate a revised virtual reconstruction proposal for the archaeological site; (4) has documented the different sources used in order to ensure scientific accuracy, which will also allow testing and validating of EM solutions. The initial idea of trying to enrich the dataset with the acquisition of new data *in situ* and in collaboration with ISPC Rome has been abandoned due to covid pandemic situation.

3.2.1 The data repository and Raw Data sharing

The initial activity tackled by the ISPC Lecce research unit was to find, rationalise and make available all the material previously acquired during the field survey phase conducted in 2016 and all the files and resources subsequently subjected to the various post processing phases.

For this reason, a shared repository was created, made available by the research unit of ISPC Rome and accessible by all team members, where a folder structure sorted by type was created:

#0 EXTERNAL SOURCES
#1 PTS
#2 OBJ
#3 C4D+texture
#4 DXF-DWG
#5 PSD-JPG render
#6 JPG-RAW Survey Photo
#7 PSZ+texture
#8 HDRI
#9 DIGITAL FRAME ANIMATION
#10 MOV aerial survey drone
#11 Blender Model

#0 - The folder contains all the bibliographic information useful for the elaboration of a 3D reconstructive proposal of the Catania theater. There are Historical photos, metric surveys, excavation documentation, restoration interventions, and reconstructive hypotheses of the scaenae frons. The file format is JPEG and TIFF.

#1 - The folder contains the cloud of points obtained from the 3D metric survey made with the Leica P20 laser scanner, essential for documenting the status quo of the monument. The file format is PTS, OBJ, PLY, FBX and DXF.

#02 - The folder contains the polygon mesh obtained from the point cloud. It was used as metric support for 3D modeling. The file format is OBJ.

#03 - The folder contains the textured 3D model of the reconstructive proposal of the Roman theater of Catania made with Maxon C4D and Vray (external render engine) software. The file format is C4D, JPEG, TIFF, PSD, TGA, HDRI and PNG.

#4 - The folder contains the cross-sections of the 3D polygonal model, carried out at different planimetric heights and along the radial axes of the auditorium. It was used as metric support for 3D modeling. The file format is JPEG and DXF.

#5 - The folder contains the final renderings of the 3D reconstruction proposal of the monument. The file format is JPEG and PSD.

#6 - The folder contains the photographic documentation acquired in situ and relating to the structures and individual architectural elements. It was used as support for 3D modeling. The file format is JPEG and RAW.

#7 - The folder contains the photographic data sets and the digital models created by Agisoft Metashape software and relating to the photogrammetric surveys about architectural elements of the scaenae frons. The file format is PSX, WRL and JPEG.

#8 - The folder contains the High Dynamic Range Image (HDRI) files for lighting the 3D scene using the Vray render engine. The file format is HDRI.

#9 - The folder contains all the rendered frames edited for the digital animations of the 3D model. The file format is JPEG, PSD and MP4.

#10 - The folder contains the aerial videos acquired by unmanned aerial vehicle (UAV). The videos have been used as metric support for 3D modeling The file format is MP4.

We are not going into more detail about #1 to #11 here. Making the data public is unfortunately still not straightforward due to ongoing copyright issues with Italian heritage science data. However, we will give a glimpse into #0: The reconstructive hypothesis produced concerning the Theatre of Catania was mainly based on a series of published scientific articles that have been collected and analysed and organised in this folder. The main writings examined were the four articles Pensabene (2005), Branciforti (2010), Branciforti and Pagnano (2008) and Wilson (1996). Through the study of these articles it was possible to determine the typology and characteristics of a series of architectural details of the site. Below is a list, albeit not exhaustive, of how each individual source has contributed to the determination and clarification of reconstructive aspects.

The irony that this structure with non-descript numbers and raw file names repeats exactly the data management mistakes of the previous project has not escaped us. We will document the structure and contents of the whole repository as soon as we can make it public.

3.2.2 Data FAIRification

We assume that the reader already has some knowledge of the FAIR principles and the *Findable – Accessible – Interoperable – Reusable* steps. For an excellent introduction we refer to D5.15 (Wright et al., 2022).

One of the main limitations to the shareability of the data and the scientific approach used to perform the first virtual reconstruction of the Roman Theatre of Catania is due to the massive use of commercial software solutions for data acquisition and processing. This resulted in the generation of a series of files in proprietary format that could only be opened and modified by having the corresponding licensed software.

In order to overcome this limitation and to make the dataset at the basis of the virtual reconstruction work open and interoperable, in accordance with FAIR principles, the first technical activity that was tackled was to export the 3D models from the proprietary Cinema4D format to the open Blender format.

In order to obtain this goal, the following activities have been achieved:

- 3D modelling and texturing of the theatre in Maxon C4D. Maxon Cinema 4D software has been used for polygonal modelling, texturing and lighting. V-ray is the engine used for rendering frame. The final model has about 20 million polygons and 138 MB of texture (Figure 10).

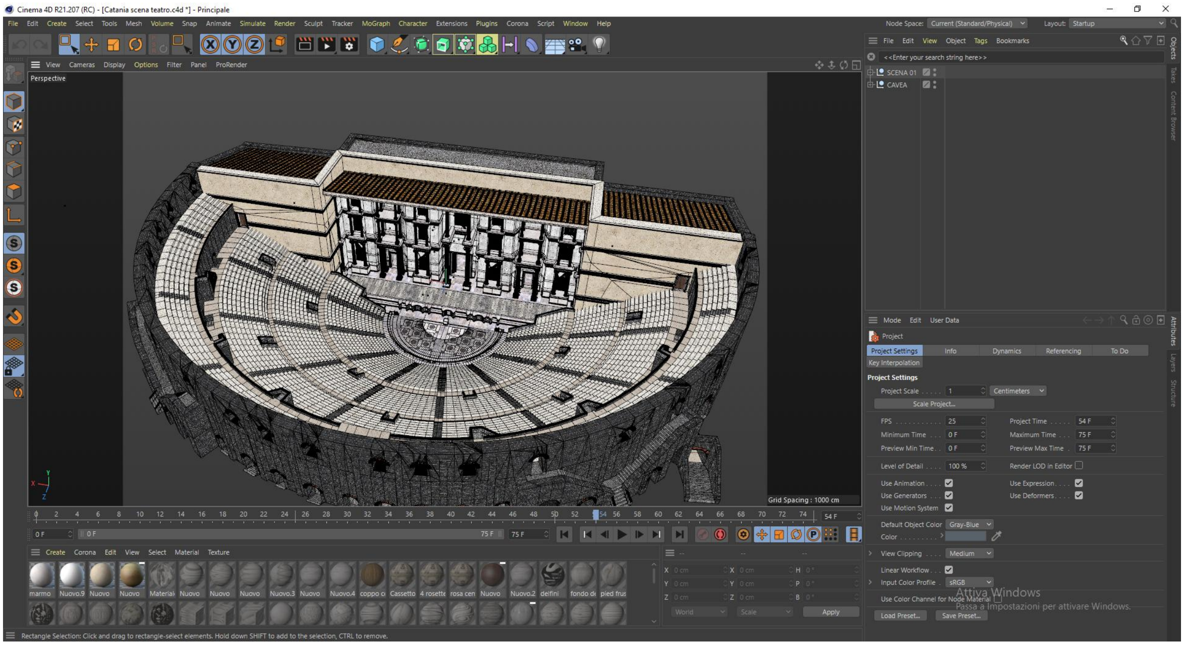

Figure 10: Modelling and texturing C4D

- 3D model optimization by polygon reduction and bake texture in Maxon C4D before exporting to Blender. Before exporting the 3D model from C4D, it was necessary to reduce the number of polygons and optimise the textures with the bake texture tool (Figure 11).

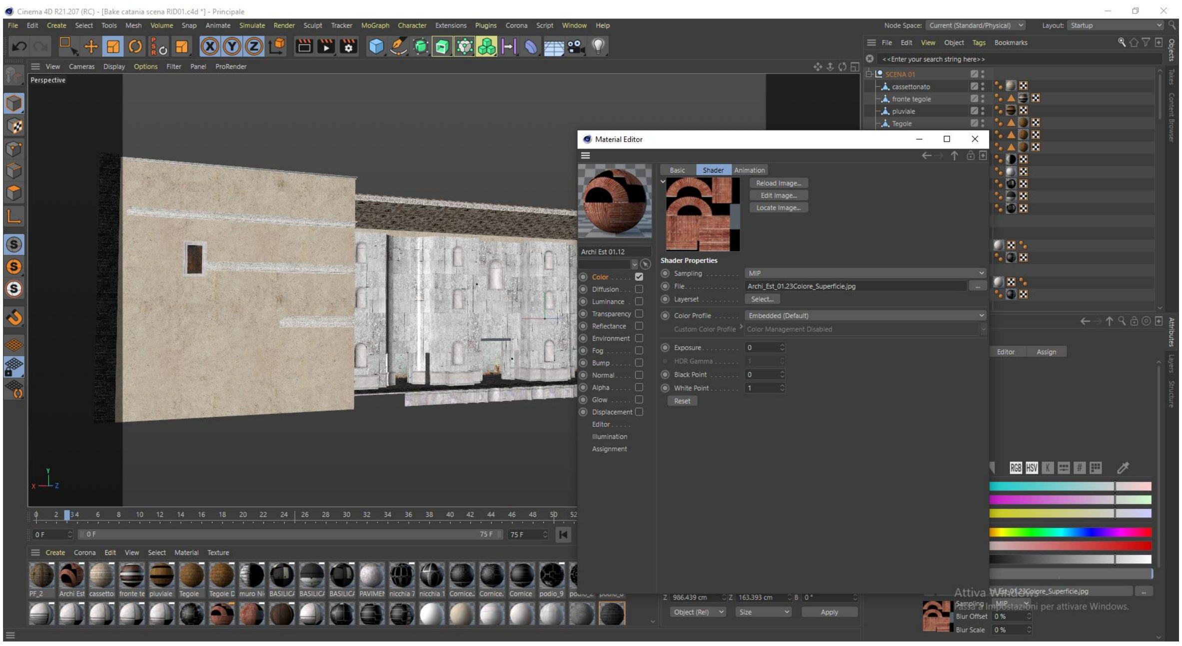

Figure 11: Optimization of the c4d model

- Export 3d model and texture from C4D to Blender in OBJ and Jpg format. The optimised 3D model and the textures have been exported in obj. and jpg. essential to be imported in Blender (Figure 12).

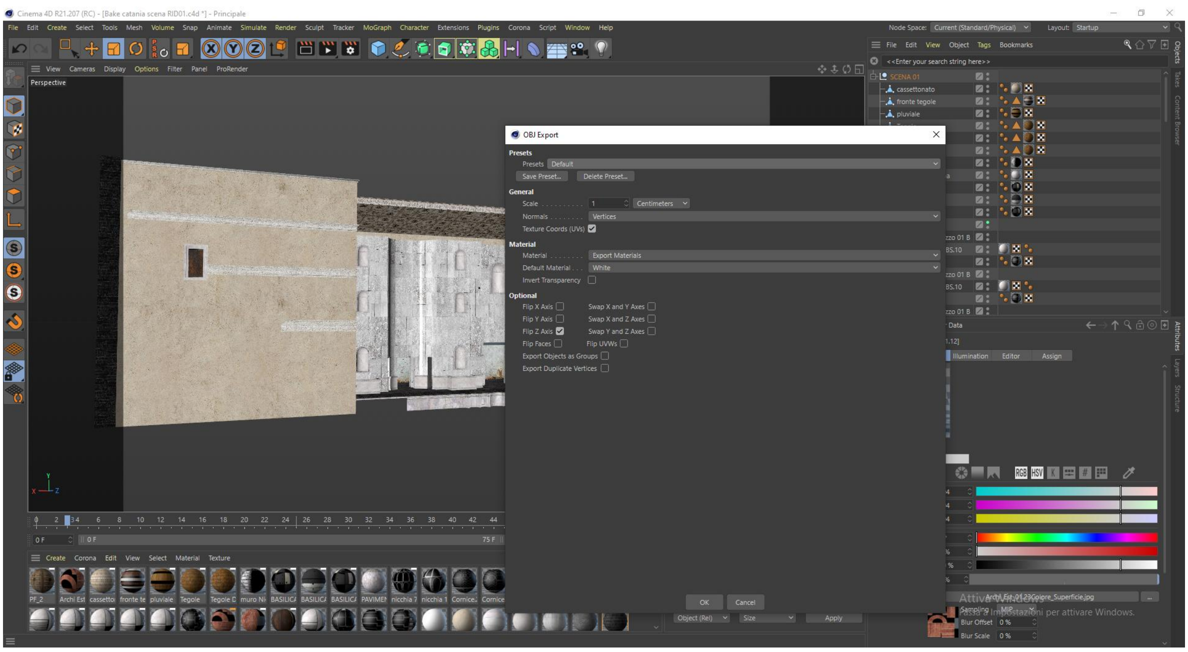

Figure 12: Export from C4D to Blender in OBJ format

3.2.3 Roman theatre: applying data to EM

ISPC Lecce has documented the virtual reconstruction process of the Roman theatre in Catania, using the Extended Matrix approach (which will be desribed in the next chapter), following a specific workflow that included the following two operative steps (see also Figure 13):

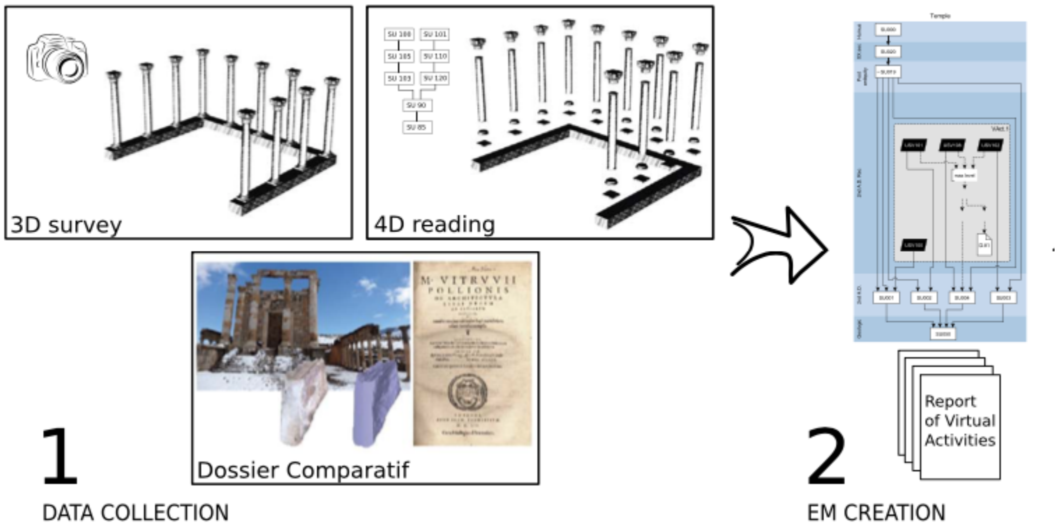

Figure 13: Production work-flow, from data collection to EM creation (Source: Demetrescu and Fanini 2017)

1. **Data collection**: 3D survey of the remains and of the special findings (i.e. non in situ blocks of stone from the architectural apparatus), stratigraphic reading (4D analysis of the *palimpsest*), creation of a *Dossier Comparatif* (comparative study);
2. **EM model creation to organise data within the Extended Matrix**: writing down of the Extended Matrix and the Report of Virtual Activities;

These steps weren't be separate phases of the reconstruction project: during data collection; ISPC Lecce has started populating the EM with interpretative elements as well as sketching out some visual representations of a given epoch. This synchronous workflow has enabled a cross-fertilisation of the relationships between the archaeological elements, the external sources, and the hypotheses.

1. Data collection has included diversified methodologies and technical solutions. Different contexts required different approaches. Some scenarios cannot involve a 3D survey due to disparate limitations (physical, legal, etc.) but can make use of legacy data (drawings or blueprints). Furthermore, when the reconstruction project involves a completely lost context, all the data collection is focused on the sources organised within the *Dossier Comparatif* (see step 1C below).

1A) *3D Survey of the context*: The Roman theatre in Catania is a palimpsest of different epochs in which each of these epochs must be purged of non-coeval elements in order to highlight the preserved ones and the overall shape of each chronological phase. Furthermore, a clear distinction between the different stratigraphies (grouped by epoch of belonging) has been a crucial phase to propose a valid reconstruction hypothesis. Alongside the *in situ* elements, the 3D acquisition will include the survey of all the non *in situ* objects. The 3D measurements resulted in accurate 3D models or digital replicas with colour information.

1B) *4D reading: the stratigraphic sequence and the Harris Matrix*: After the 3D survey, a stratigraphic reading has been performed in order to highlight the temporal sequence (4th dimension) of the actions. The Harris Matrix visually represents the stratigraphic elements that allow archaeologists to understand the sequence of deposition at a site.

1C) *Dossier Comparatif*: All the sources and the comparisons with other contexts have been stored in the Dossier Comparativ (Gros 1995): The Dossier Comparatif (also called DosCo) is a collection of documents which follows a specific nomenclature, a composite known as "D." (Document), plus a sequential numbering (i.e. D.01). The number is set according to the chronological sequence of data ingestion. All the documents are linked inside the EM and used to validate the reconstruction hypothesis. These are represented inside the EM through the *source node* (Demetrescu 2015).

2. Creation of the Extended Matrix and the Report of Virtual Activities: In archaeological practice, a virtual reconstruction is usually created at the end of an excavation project or 3D survey and is not part of the research itself. The 3D model, in this sense, is intended as ex-post work which synthesises different hypotheses made during the investigations (Medri 2003). Generally, these are not formally annotated. In some cases, experts write down intermediate reports to fix certain general ideas, but the collection of the archaeological record lacks a precise way to store the rich connections that are recognised between pieces of evidence during the fieldwork.

ISPC Lecce used the EM approach to fix this issue. In this case study, virtual reconstruction has been part of the archaeological investigation starting from the bibliographic research about this archeological site and the data from the documentation and analysis of the campaign carried out. The reconstruction hypothesis has been formally "documented" along with the archaeological record (even if taken from an already published archeological survey) from which it is derived. "Early collection" simplifies the management of the reconstructive record and "drives" the subsequent source-based modelling. This approach has made it possible to publish a thorough account of the scientific process, highlighting the connection between the archaeological and the reconstructive records and enabling other researchers to actively check, modify, and consciously reuse the virtual reconstruction.

Report of Virtual Activities: For real stratigraphic units, Virtual Reconstruction Units (USV) are combined into Virtual Activities (VActs, see Fig. 14) using rectangular shapes in the EM canvas. VActs are described in the Report of Virtual Activities (Demetrescu 2015), a textual report that acts as an intermediate output and is useful for quickly sharing a reconstruction hypothesis. A Report of Activities has been written down for each reconstructed epoch.

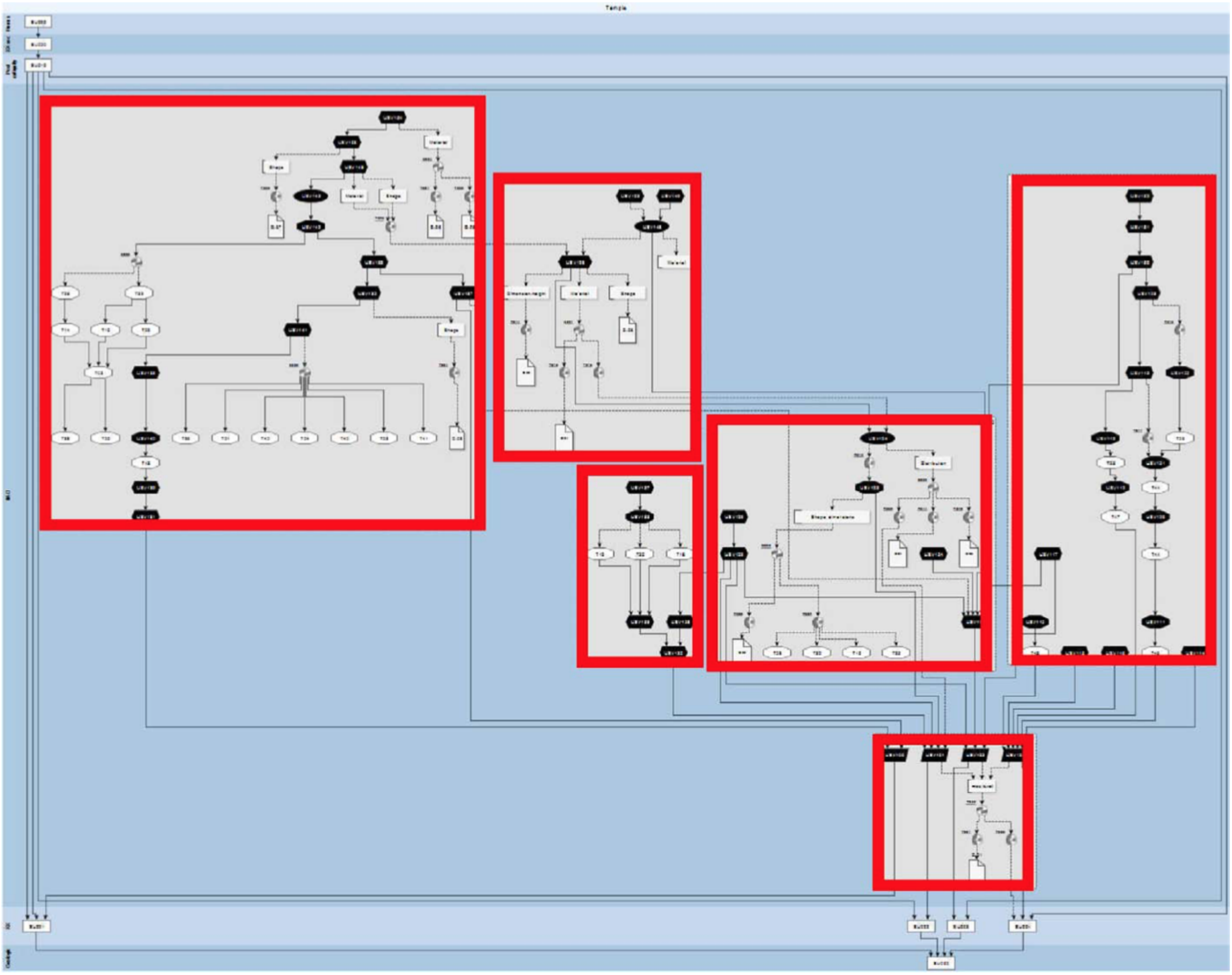

Figure 14: Extended Matrix represents the Great Temple in all its epochs and hypotheses, making explicit all the elements involved in the reconstruction and their relationships. Marked in red are the Virtual Activities. (Source: Fanini and Demetrescu 2017)

3.2.4 Virtual reconstruction re-processing

Creation of 3D models (source-based modelling): Starting from both the sources organised in the EM (see Fig. 15) and the already existing virtual reconstruction, ISPC Lecce has performed the 3D modelling, with different levels of representation from small details to a broad perspective:

- digital restoration of digital replicas;

- digital anastylosis of the elements in order to restore their original position and spatial relationships;
- virtual reconstruction of a given epoch starting from all the sources available and the analyses performed. It is important to take into account that from a reality-based model it is possible to make several reconstruction hypotheses: ideally at least one for each period formalised in the Harris Matrix.

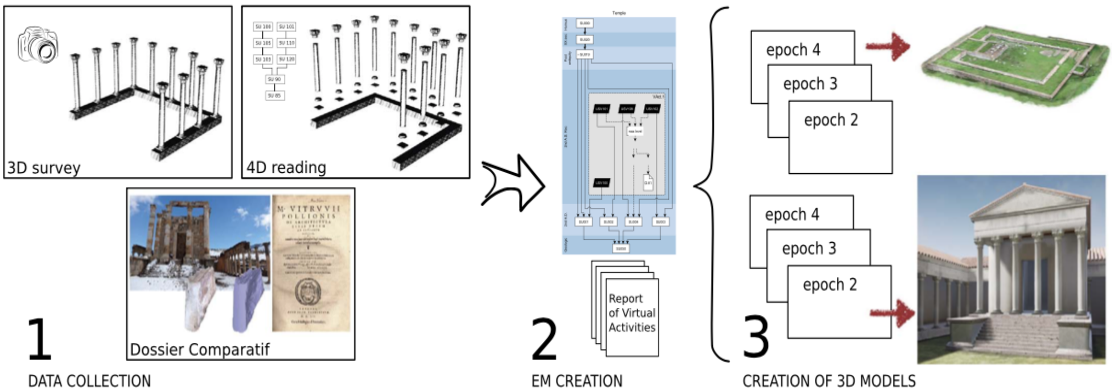

Figure 15: Complete production work-flow, from data collection to EM creation and creation of 3D models (Source: Demetrescu and Fanini 2017)

The proxy-representation model approach: The creation of 3D content has followed two levels of abstraction: proxy models which are simplified representations of the reconstruction through basic geometrical shapes (cylinders, boxes, spheres, etc.) and representation models which are focused on fine geometries, colour, and material simulations resulting in the final, aesthetic depiction of the reconstruction hypothesis. The Proxy level has made it possible to highlight portions of the representation model and to interact with it. Furthermore, it also represents the first draft model of the reconstruction and is useful, along with the Report of Virtual Activities and the EM, for the sharing of intermediate results with colleagues and other experts and for providing feedback.

Graphical documentation for the reconstruction hypothesis: As mentioned before, the reconstruction hypothesis of a context has two depictions: a proxy model and a representation model. These models make it possible to explore and experience the research results in a virtual space. However, the EM approach is not focused only on digital media, but it pertains, first of all, to the archaeological documentation for a site's interpretation and scientific publication. Each stratigraphic unit - virtual or real - has a proxy representation. Making sections and plans out of proxy models helps provide technical documentation along with the description of virtual activities. Extended Matrix, Report of Virtual Activities, and graphical documentation published together has allowed a coherent formal representation of the reconstruction hypothesis.

Validation of the reconstructed model: The construction of a 3D representation model of the archaeological site of the Roman theatre in Catania (or part of it) has been the final part of the practical application of the virtual reconstruction approach based on Extended Matrix and has been carried out in strict collaboration with the research unit of ISPC Rome. Among other reasons, the Roman theatre was chosen because, in addition to having already an important series of raw data (bibliographic, laser scans, photogrammetric and photographic surveys), it already has a well-documented virtual reconstruction that can certainly represent the first and most direct yardstick to evaluate the effectiveness of the EM approach.

In fact, although the new and the old representation are both 3D representations of the same archaeological site, they are the result of work processes that have different approaches, objectives and practices: the former more streamlined, expeditious, oriented to a broad audience, is particularly focussed to the high quality photo-realism of the final product; the second, the EM in fact, aims to link the sources to the corresponding reconstructive hypotheses and therefore also more onerous in carrying out the entire process previously analysed also because of a much finer granularity of the elements.

4. The Extended Matrix

4.1 The stratigraphic reading approach

The Extended Matrix approach, defined by the ISPC Rome research team, is based on the *stratigraphic reading* approach[4] and aims to create a common framework connecting archaeological documentation and virtual reconstruction in the earliest stages of the excavation of a site or a 4D survey of a monument. The relationships between the units found during an excavation, e.g. one unit being above another unit, result in a stratigraphic matrix, for example a *Harrix matrix*. However, stratigraphic units are not used only in archaeological excavations: they are also the base element in similar disciplines like building archaeology, restoration and landscape archaeology. Stratigraphic units indeed do not exclusively describe the information that derives from the excavation phase, which goes to document all those destructive activities that are typical of the phases of the removal of soil layers with uniform characteristics, but also map elements above ground (such as wall structures, plasters, decorations, frescoes). Through the population of the extended matrix, both objective information and interpretations derived from deductive, contextual or analogical information (i.e. derived by analogy) are documented (see Figure 16).

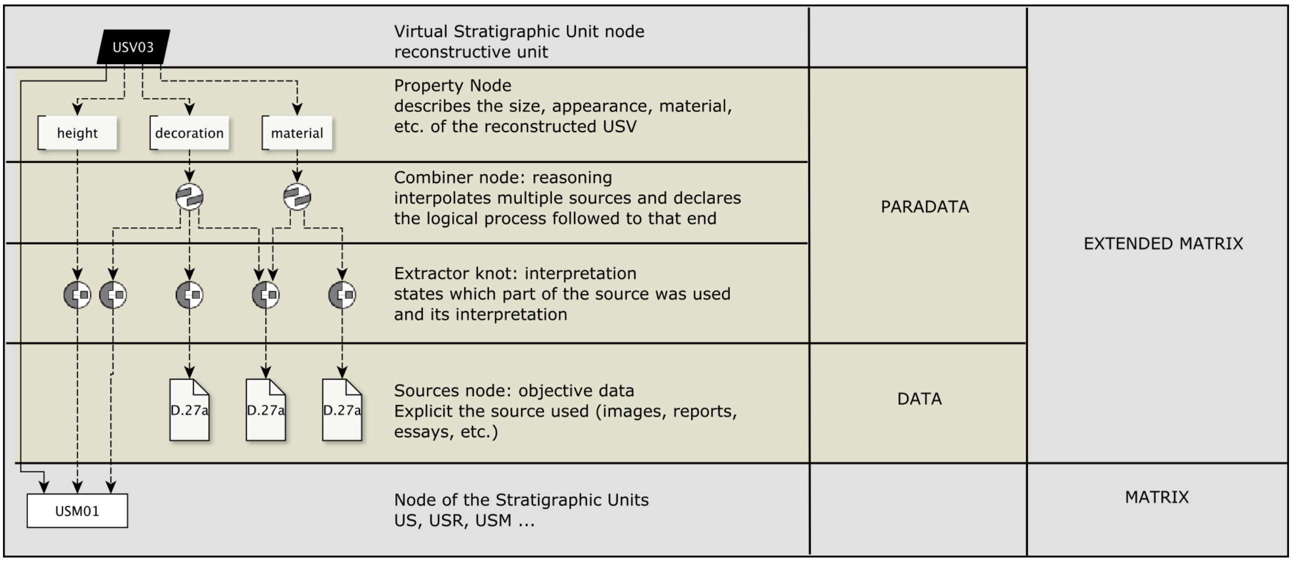

Figure 16: EM node structure showing the relation within Virtual stratigraphic units, property nodes, combiners, extractor and source data

[4] OSIRIS website, details on statigraphic approach:
http://osiris.itabc.cnr.it/extendedmatrix/index.php/why-a-stratigraphic-approach/; []

The Extended Matrix is populated through the free editor yED[5]. The yED editor is able to design Harris Matrices correctly by including some graphic elements that are often complicated to reproduce manually, such as bridges to avoid the crossing of connecting lines between stratigraphic units. yED also has an effective automatic layout system useful to keep the dataset orderly and readable. The GraphML[6] XML file produced by yED is imported into Blender's[7] EMtools plugin[8], specifically designed to read the file and connect it with the geometries - simplified shapes called *proxies* - which represent the volumes associated with the various stratigraphic units, on which the virtual reconstruction will take shape.

Figure 17 illustrates the process of creating a virtual reconstruction through five canonised steps: it starts with the collection of data that includes the instrumental survey of the chosen context. In a second step the data are analysed from a chronological and stratigraphic point of view to model the known data within the 3D space. In the third step the data are implemented with the sources that have been found proposing a first reconstructive hypothesis with volumetric character and with standardised canonised colours that express the degree of certainty of the single reconstructive elements (virtual stratigraphic units). In the fourth step a representation model is developed, complete with textures and materials, useful to visually represent the context both for experts and for the general public. Finally, in the fifth step the publication and the dissemination through the development of Rendering, video computer graphics, applied games or virtual museums takes place.

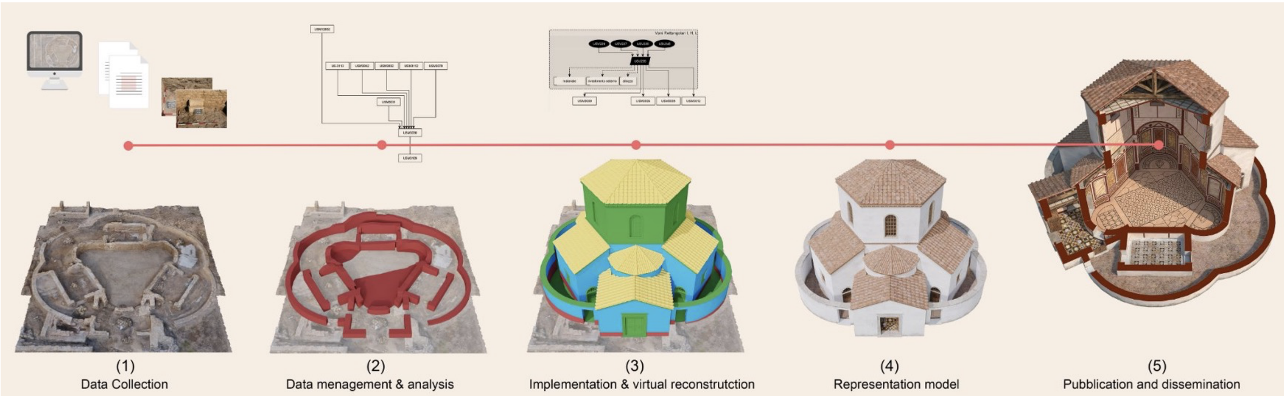

Figure 17: The Extended Matrix process in five standardised steps

[5] yED editor for drawing diagrams: https://www.yworks.com/products/yed; [24 March 2022]

[6] GraphML, an XML-based data format for storing mathematical graphs: http://graphml.graphdrawing.org/; [24 March 2022]

[7] Blender website: https://www.blender.org/; [24 March 2022]

[8] OSIRIS website, details on EMtools: http://osiris.itabc.cnr.it/extendedmatrix/index.php/extended-matrix-framework-emf/embt/; [24 March 2022]

4.2 Applying the FAIR principles in the EM

The starting point to apply the FAIR principles (for an introduction we once again refer to D5.15 / Wright et al., 2022) to the use case of Catania was the conversion of data from closed and proprietary formats to open formats encoded according to containers and whose specifications are public (for example the 3D scene file .blend or the .gltf). This task required lab test work to transform the data present inside the closed software cinema 4D into the chosen open formats without loss of information. The activity required several iterative steps in order to retain the information. Despite this, some elements were not convertible because they were created outside the standards of 3D models.

Considering that the starting data were heterogeneous (see previous section), to make the data accessible and integrated between them, the metaphor of the knowledge graph was used and in particular the formalism in which it is possible to connect the data to their three-dimensional representations.

Another key aspect for dissemination was the development of an online tool capable of reading the EM database in open format and make it searchable and viewable within a browser. This tool, named EMviq was developed within the framework ATON 3.

The work of data interchange is made possible if the service platforms are able to talk to each other in an effective way. For this reason, particular effort has been made in the design and execution of tools for data exchange between ATON 3 and Matrix and iDAI collaborative tools.

In order to make the EM knowledge graph approach compatible with the other databases present in the iDAI tools, a first mapping of EM with CIDOC CRM has been developed. This is the culmination of an effort that lasted several months in collaboration with colleagues at the Polytechnic of Turin.

4.3 EMviq

Once the reconstruction phase is finished, the files generated by EMTools can be exported in a 3D format (the glTF standard) which in turn can be imported into the *EMviq* tool[9]. EMviq is a complete, interactive 4D visualisation and interrogation tool for Extended Matrices. By using this tool, the 3D reconstruction together with all related stratigraphic information can be interactively queried and inspected, allowing users to virtually explore an entire archaeological context and spatially perform 3D queries. Once a proxy is hovered over/selected through user input, a source-graph is presented to the user that validates that specific proxy.

[9] EMviq website: http://osiris.itabc.cnr.it/scenebaker/index.php/projects/emviq/; [24 March 2022]

In order to deploy an interactive 4D inspection context with which the user can interact, three different extraction steps are defined from an Extended Matrix stored in a graphDB[10] (Fanini and Demetrescu, 2018):

1. *Timeline extraction*: This step will extract from the GraphDB a finite number of time-periods, identified by ID, including beginning and end values.
2. *Proxy-graph extraction*: This will create a graph of proxies for real-time queries and interrogation, arranging semantic 3D shapes (proxies) into the 3D space including procedural rules (e.g. so-called seriation nodes).
3. *Source-graph extraction*: This will generate an internal runtime representation of EM sources relationships (paradata).

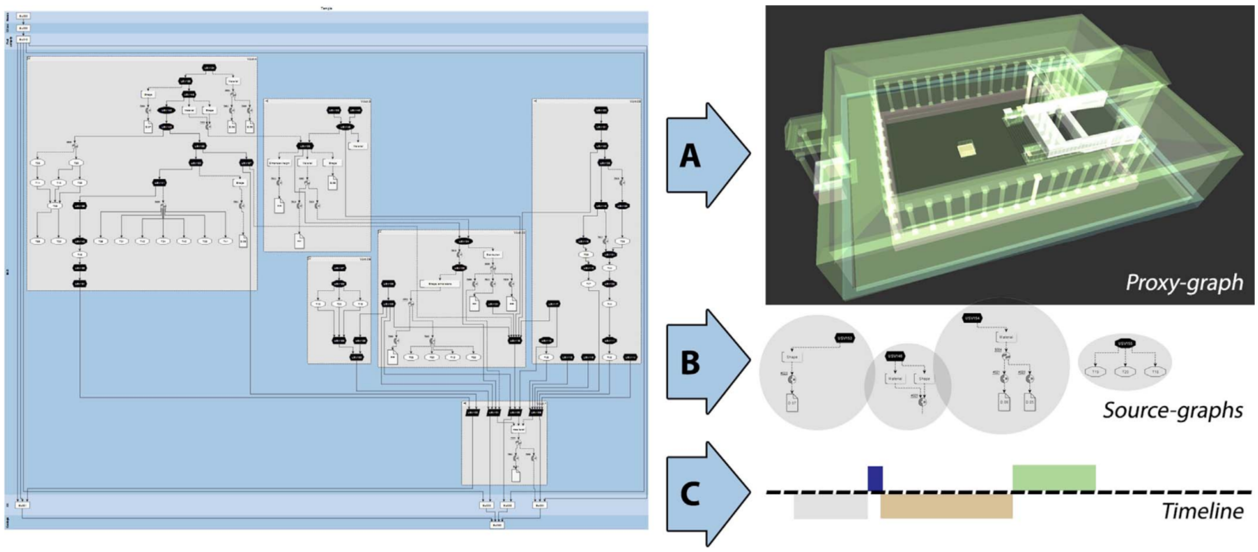

Figure 18: Three different extraction steps from an Extended Matrix

These computational steps need to be performed only when the Extended Matrix (the GraphML file) is modified/updated by the user. This is particularly important for cloud-based scenarios, where remote users are operating on the same Extended Matrix (GraphML file) through the yED editor, or other tools able to visually manipulate GraphML files.

Within the SSHOC project, the EMviq tool was completely redesigned and developed as open-source Web3D online tool[11], based on the open-source ATON framework by CNR ISPC (Fanini et al., 2021) accessible through a common web browser, without any third-party installation required. The web-based tool targets real-time Web4D, offering users interactive inspection of multi-temporal,

[10] GraphDB website: http://graphdb.ontotext.com/; [24 March 2022]

[11] https://github.com/phoenixbf/emviq [April 2022]

semantically enriched virtual environments through the EM formalism. The project was designed with ease-of-use in mind and cloud-based integration to establish a fast and robust pipeline within multi-disciplinary teams.

The framework was the optimal choice to develop the online version of EMviq since it provides:

- Full embracement of modern web standards and open specifications
- Cross-device presentation (mobile, desktop and XR devices) and responsive UIs
- Efficient multi-dimensional (3D/4D) queries at runtime
- Scalable deployment
- Integration with external services/tools through REST API

The tool inherits several features from the ATON framework, offering cross-device consumption (mobile and desktop devices), including support for immersive VR inspection on head-mounted displays (WebXR). EMviq consumes a scene ID (mapping an ATON scene descriptor, as JSON file) and a matching Extended Matrix (a GraphML file). The scene descriptor defines all representation models for each period, matching the ones defined in the EM file. This approach allows the online tool to load different projects through a single identifier (in this case, the scene ID) through an inline parameter, thus providing a way to address not only the tool, but also the specific project.

More specifically the Web4D tool offers three main tasks to its users:

- **Explore**: users can interactively move through space and time. The space can be explored leveraging on ATON built-in navigation system (orbit and first-person modes), while time can be explored using a timeline automatically extracted and presented at interface level
- **Query**: users can interactively query the EM proxies into the 4D space, automatically extracted from the EM leveraging on ATON built-in routines for semantic shapes. The tool also offers options related to spatial occlusion (representation and semantic model) and visibility settings, alongside proxy-specific routines (focus, etc.)
- **Search**: users can semantically search into the virtual 4D space through a UI field. This is useful to interactively ad progressively filter and highlight a proxy or set of proxies by ID or term (e.g. “walls”, “USV100”, “groud”, ...) extracted from proxies description fields, enabling users to perform a semantic navigation of the 4D space exploiting automatic viewpoints

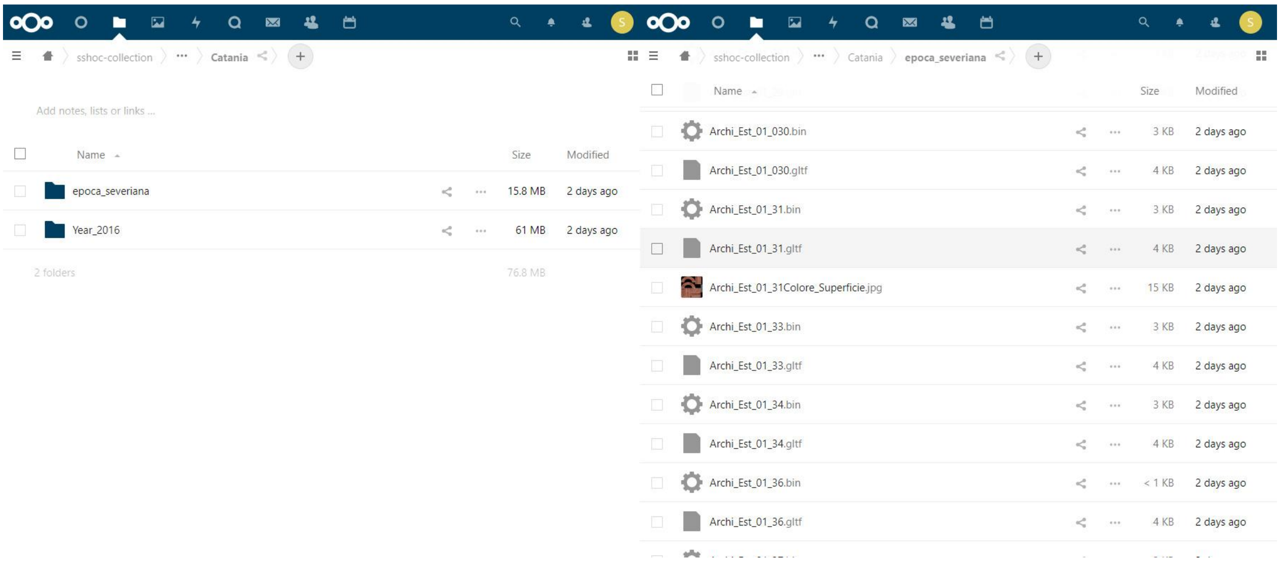

Figure 19: The Web4D tool

Furthermore, a cloud integration of EMviq was explored allowing remote and distributed workflow. Existing open-source solutions like NextCloud[12] offer also fine-grained access policies already explored within the ATON framework and previous projects (Fanini et al. 2019). This enables professionals to operate remotely on the EM associated with a specific case study and dynamically update the interactive 4D presentation. Such cloud-based workflow is also powerful to support validation and visual debug of EM by remote professionals.

[12] See: https://nextcloud.com/ [April 2022]

Figure 20: Dynamically updating the 4D presentation with EMviq

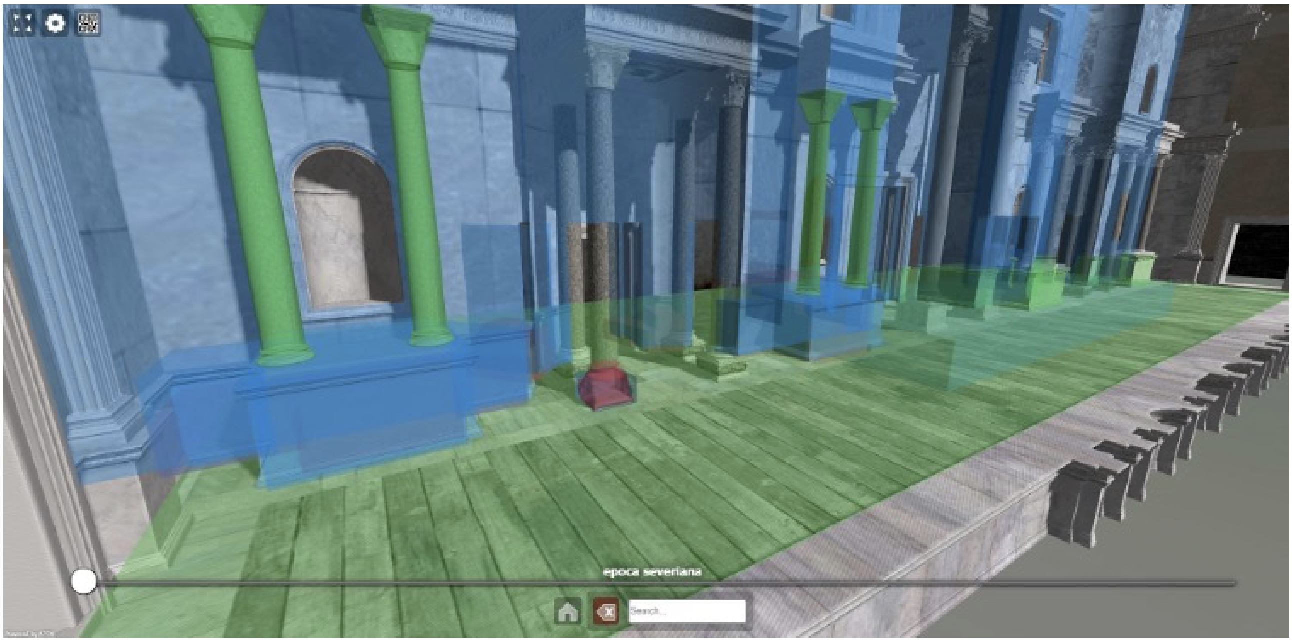

Figure 21: Immersive VR inspection in EMviq

Finally, EMviq tool offers immersive VR inspection on 3 and 6 DoF (Degrees of head-mounted displays (HMDs) including VR controllers, exploiting ATON XR features (through WebXR) and functionalities (see Fanini et al. 2021). More specifically, regarding spatial UI targeting XR, a set of reusable elements offer

3D panels with optimised font rendering, toolbars, buttons and labels that can be arranged into the 3D space and customised to suit EMviq requirements. For immersive VR exploration, a well-known and established locomotion technique based on teleport is provided. The component automatically adapts to 3-DoF and 6-Dof head-mounted displays, switching pointing/selection methods accordingly to query the proxy-graph, also depending on the presence of VR controllers, thus reaching a wide range of HMDs devices.

5. Contextualising the Roman theatre

People gave the order to extend the theatre or carried out the order. Places can mark the area where the theatre is situated or point to other places such as quarries for the building material. Periods may represent the general context of some building activity or the reign of a local ruler. Any building activity is an event. Objects may be part of the structure of the theatre or may have been placed there for ornamental reasons. In addition, the digital objects in the 3D virtualisation have physical counterparts with object types and dimensions.

In the following we describe two concrete examples how the Roman theatre can be contextualised. In both examples, the starting point is some physical evidence found at the site of the Roman theatre. In chapter 6, these examples will be mapped to CIDOC CRM. This contextualising scenario (we will use *contextualisation* and *scenario* synonymously) connects the Extended Matrix of the Roman theatre with objects, people, places, periods and events that are associated with it, as well as the literature on the Roman theatre. (An earlier version of the scenario was described in Milestone MS36.)

5.1 Example 1: architrave block

The first example is a marble block belonging to an architrave (i.e. a lintel that rests on the capitals of columns), documented in Wilson (1990) and DAI's Arachne database of archaeological objects[13] and part of the EM of the Roman theatre. A drawing of a part of the architrave had already been included in volume V of "Le antichità della Sicilia esposte ed illustrate per Domenico Lo Faso Pietrasanta, duca di Serradifalco … Antichità di Catana", published in 1842 and documented in Arachne[14] and the Heidelberg Digital Library[15]. However, this marble block had fallen in the *orchestra* of the theatre (where the chorus sings and dances) and was only discovered in the 1980s. The block contains an ornamental relief and can be likely dated to the early 3rd century CE. This corresponds to the second major building phase of the Roman theatre during the Severan dynasty (193 to 235), in particular to emperor Elagabalus (218–222) or Alexander Severus (222–235).

[13] See: https://arachne.dainst.org/entity/1183759 [April 2022]
[14] See: https://arachne.dainst.org/entity/1642763 [April 2022]
[15] See: https://digi.ub.uni-heidelberg.de/diglit/serradifalco1842bd5 [April 2022]

This example can be formalised as follows:

Block

is made of marble
contains an ornamental relief
is documented in Arachne, based on Wilson (1990)
was discovered in the 1980s in the *orchestra* of the theatre
which implies that it had fallen down
Is part of an architrave
(has the same dating as the architrave as a whole)

Architrave

probably early 3rd century CE, according to Wilson
i.e. likely during the reign of emperor Elagabalus (218–222) or Alexander Severus (222–235)
which makes it part of the second major building phase of the Roman theatre
depicted in a book by the Duca di Serradifalco

Second major building phase of the Roman theatre

during the Severan dynasty (193 to 235)

Book

published in 1842 by the Duca di Serradifalco
documented in Arachne and the Heidelberg Digital Library

5.2 Example 2: inscription

The second example demonstrates how to provide additional context: An inscription, documented in Arachne[16] and the Last Statues of Antiquity database[17], was found at the site of the Roman theatre. The inscription belongs to a column that depicts the guardian heroes of Catania. It first stood at the Campus Piorum near Catania and was later moved to the Roman theatre. This was not unusual especially in the second building phase: The theatre was decorated with marble columns, statues and decorative reliefs celebrating public events associated with mythological themes (see also chapter 2.1).

Even later, a consular governor of *Sicilia* named Merulus, about whom nothing else is known, either restored or replaced the monument and ordered the addition of an inscription. The inscription calls him “spectabilis”, a title that (according to the Last Statues of Antiquity database) was not given to consular governors before 434. Alternatively, Arachne cites Wilson (1990) who claims that the term

[16] See: https://arachne.dainst.org/entity/1185927 [April 2022]
[17] See: http://laststatues.classics.ox.ac.uk/database/detail-base.php?record=LSA-2057 [April 2022]

originated at the end of the reign of Valentian III (425-55) or during the reign of von Leo I (457-74). In both cases this is well after the two main building periods of the Roman theatre. The Last Statues of Antiquity database suggests that Merulus might even have been consular governor in the Ostrogothic period, i.e. the reign of the Germanic Ostrogoths in Italy and neighbouring areas from 493 to 553.

One can formalise the second example as follows:

Statue

depicts (mythical) persons: the guardian heroes of Catania
was at some point situated at the Campus Piorum near Catania
was moved to the Roman theatre
was either restored or replaced by Merulus
only the base still exists

Base

is part of the statue
was discovered in 1951, in the central sector of the lower seat rows
Is now in the courtyard of the Museo Civico at Catania
has a width of 56 cm and a height of 26,5 cm
is documented in Arachne and the Last Statues of Antiquity database
bears an inscription

Inscription

was ordered by Merulus
calls Merulus “spectabilis”, which implies a *terminus post quem* for the inscription of 434

Merulus

was consular governor of Sicilia
is called “spectabilis”, which implies a *terminus post quem* for (the end of) his reign of 434
may or may not have been consular governor in the Ostrogothic period (493 to 553)

6. The CIDOC CRM mappings

There is a growing consensus in the SSH that CIDOC CRM is a good choice for semantically modelling its data. In this chapter we first briefly introduce CIDOC CRM and its extensions and then describe the mappings of the EM and the contextualisation to CIDOC CRM and an alignment of the mappings. We also discuss the difficulties we faced when using CIDOC CRM.

6.1 The CIDOC CRM core ontology

In this section, we introduce some basic concepts of CIDOC CRM, as far as they are relevant for the case study. However, understanding CIDOC CRM and its extensions takes time and we can only describe the tip of the iceberg here. For more details about CIDOC CRM we must refer to other places, for example Bekiari et al. (2021) or D4.20 (Bekiari et al. 2022).

CIDOC CRM a high-level ontology for the humanities. The "CRM" stands for "conceptual reference system". It sees the world as a series of events, i.e. meetings of people, ideas and objects, which interact with each other in limited areas of space, time and periods, and bring about noteworthy changes. CIDOC CRM consists of **classes** such as *Actor* (which is a person or a group) or *Activity* (which is an event that is intentionally carried out by an Actor, as opposed to e.g., a volcanic eruption) and **properties** that connect these classes: for example, the property *carried out by* connects an Activity with an Actor, as in "the painting of the Sistine Chapel *carried out by* Michaelangelo Buonaroti".

It employs a strict bottom-up approach, which means it is always based on real data to be modelled. For example, when the concept of acqusition was introduced early on in the development of the CIDOC CRM, a use case was the acquisition of an object by a museum. This explains the restriction to physical things, i.e. Acquisition cannot be used to describe for example, the copyright of a song.

Each class or property has a number and a name, for example *E8 Acquisition* or *P24 transferred title of*, which names the *Physical Thing* that is being acquired. The name should give a hint about its meaning, but the correct way to understand the class or property is to read the **scope note** that explains which things belong to it, and to look at examples. It is a common mistake to use classes and properties whose names sound as if they describe what one is looking for but where the scope note makes clear that they don't. For example, when a property is called *consists of*, it is not clear from the name whether it applies to things or periods.

Classes and properties each form a **hierarchy** where more general classes and properties are higher up and more specific ones are further down. For example, *E96 Purchase* is a **subclass** of *E8 Acquisition* that only applies when money changes hands in the acquisition. *P11 had participant*, which connects *E5 Event* with an *E39 Actor*, has a **subproperty** *P14 carried out by*, which only applies when the Event is in fact an Activity and the Actor actively participates in it. As a general rule when using CIDOC CRM, be as specific as you can, but don't hesitate to use a less specific class or property that is sure to fit. For example, if it is not clear whether an Actor actively participated in an Activity or was "just there", using the more general *P11 had participant* is fine.

When *P14 carried out* by connects an E7 Activity with an E39 Actor, then the E7 Activity is called the **domain** and the E39 Actor the **range**. The domain and range are comparable to the subject and object in a simple "subject verb object" sentence. Virtually all properties have **inverses** where the domain and range are swapped. For example *P14 carried out by* has the inverse property *P14i performed*, which connects the E39 Actor with the E7 Activity.

All statements consist of one or more simple "subject verb object" sentences, where the object of one sentence is the subject of the next sentence. For example:

E22 Human-Made Object "**Architrave**"
P108i was produced by E12 Production
P4 has time-span E52 Time-Span
P86 falls within E52 Time-Span "**probably early 3rd century CE**"

actually consists of several "subject verb object" sentences:

subject	**verb**	**object**
E22 Human-Made Object "**Architrave**"	P108i was produced by	E12 Production (of the architrave)
E12 Production (of the architrave)	P4 has time-span	E52 Time-Span (denoting the exact time, which is unknown)
E52 Time-Span (the exact time, which is likely unknown)	P86 falls within	E52 Time-Span "**probably early 3rd century CE**"

6.2 CIDOC CRM extensions

CIDOC CRM has a set of extensions[18] that model domain-specific knowledge. We give a very brief overview of CRMarchaeo, CRMba, CRMinf and CRMdig, which are used in the EM and LOD mappings.

CRMarchaeo models the excavation process and the resulting knowledge. It takes into account the fact that the process of excavation always destroys the archaeological resource and cannot be repeated, which makes proper archaeological documentation a necessity. CRMarchaeo introduces classes which are compatible with stratigraphic units, and the relationships that can be inferred from the excavation documentation. For example, if an object is found in one layer of soil and another object is found in a

[18] All CIDOC CRM extensions are documented here: https://cidoc-crm.org/collaborations [April 2022]

deeper layer of soil, then one can infer that the second object was deposited before the first object. The scenario mapping uses *S19 Encounter Event* for archaeological finds (which is referred to in CRMarchaeo because it is a fundamental part of an excavation but is actually from the CRMsci extension for scientific observation).

CRMba is an ontology to encode metadata about the documentation of archaeological buildings.[19] The EM mapping uses its classes *B1 Built Work*, its subclass *B2 Morphological Building Section* and its subclass *B3 Filled Morphological Building Section*, together with *B4 Empty Morphological Building Section*, which – despite its name – is not another subclass of B2 but is a feature of a Filled Morphological Building Section ("feature" in CIDOC CRM means a physical thing that is physically attached in an integral way to particular physical objects).

CRMinf defines an *Argumentation Model*. It is "intended to be used as a global schema for integrating metadata about argumentation and inference making in descriptive and empirical sciences".[20] The EM mapping uses its classes *I5 Inference Making* and *I7 Belief Adoption*. The scenario mapping uses *I2 Belief* and *I7 Belief Adoption*.

CRM Digital, or CRMdig for short, is "an ontology to encode metadata about the steps and methods of production ("provenance") of digitisation products and synthetic digital representations such as 2D, 3D or even animated Models created by various technologies."[21] The EM mapping uses its classes *D1 Digital Object* and *D2 Digitization Process* and the connecting property *L11 had output*.

SSHOCro, which stands for "SSHOC Reference Ontology", is a CIDOC CRM extension that was developed within SSHOC by FORTH in task 4.7 (Modelling the SSHOC data lifecycle). It models the lifecycle of SSH data and extends the reach of CIDOC CRM into the social sciences. SSHOCro has a few basic concepts:

- SSH Project (with a project status and a work plan)
- Knowledge Workflow Activity (KWA), which models a common order of work steps in scientific research:
    1. data collection,
    2. data preparation/processing and connecting different datasets,
    3. interpretation, including writing reports.
- auxiliary activities: data storage, publication of data, reports etc.
- Dataset: the main inputs and outputs of the Knowledge Workflow Activity steps
- Service and Tool.

[19] See: https://cidoc-crm.org/crmba/ [April 2022]
[20] See: http://www.cidoc-crm.org/crminf/ [April 2022]
[21] See: https://cidoc-crm.org/crmdig/ [April 2022]

6.3 The mapping of the EM

We have created a first version of a mapping of the EM to CIDOC CRM and its extensions. This is a work in progress, and still only on the conceptual level. It has not been applied to the concrete situation of the Roman theatre yet. However, we have conceptualised the use of CRMba for the extended matrix (see Figure 22).

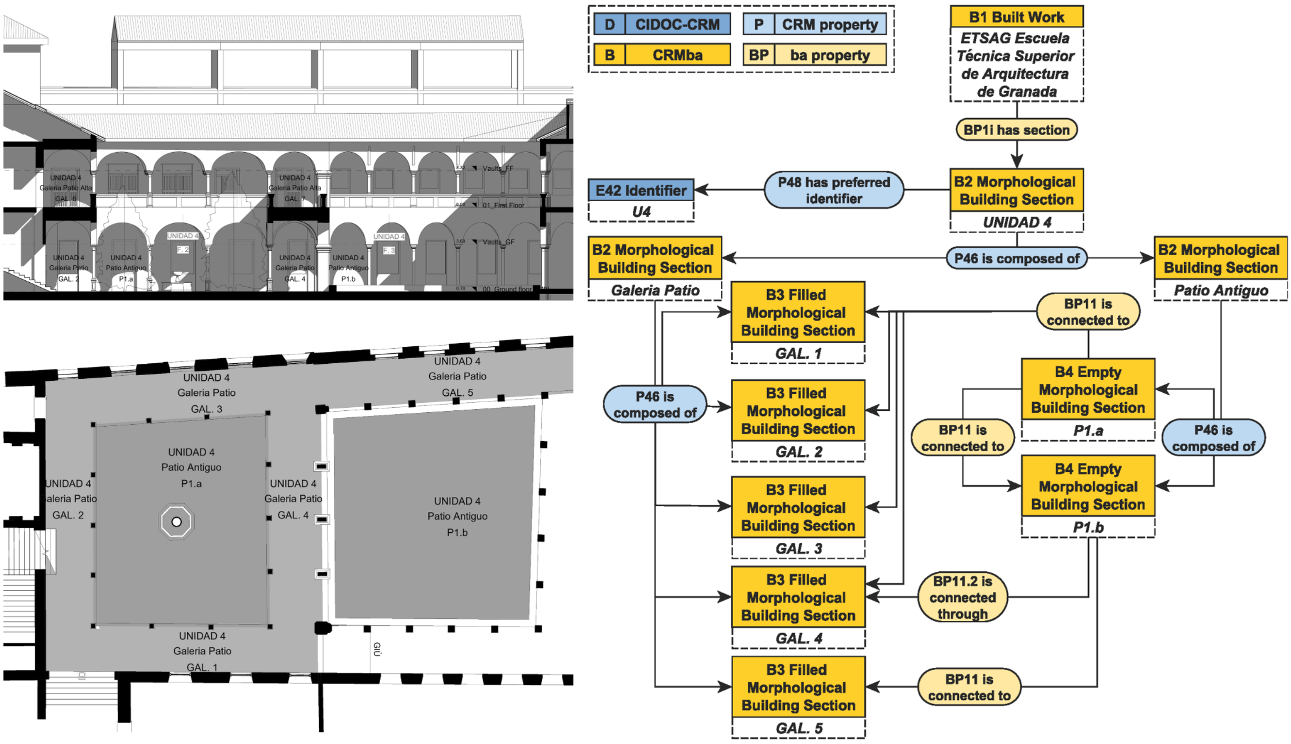

Figure 22: Using CRMba for the EM

The main hurdle for modelling the EM were the *virtual stratigraphic units* (see chapter 4). CRMarchaeo models the stratigraphic units that are documented in an archaeological excavation and documented in a Harris matrix. The EM adds virtual stratigraphic units that have not been documented in an excavation and are only assumed or inferred to have existed. These virtual stratigraphic units are not foreseen in CRMarchaeo, and it was not straightforward to extend CRMarchaeo to model these assumed parts. Thus, the virtual stratigraphic units had to be modelled from scratch. We model them as D1 Digital Objects from CRMdig, see Figure 23.

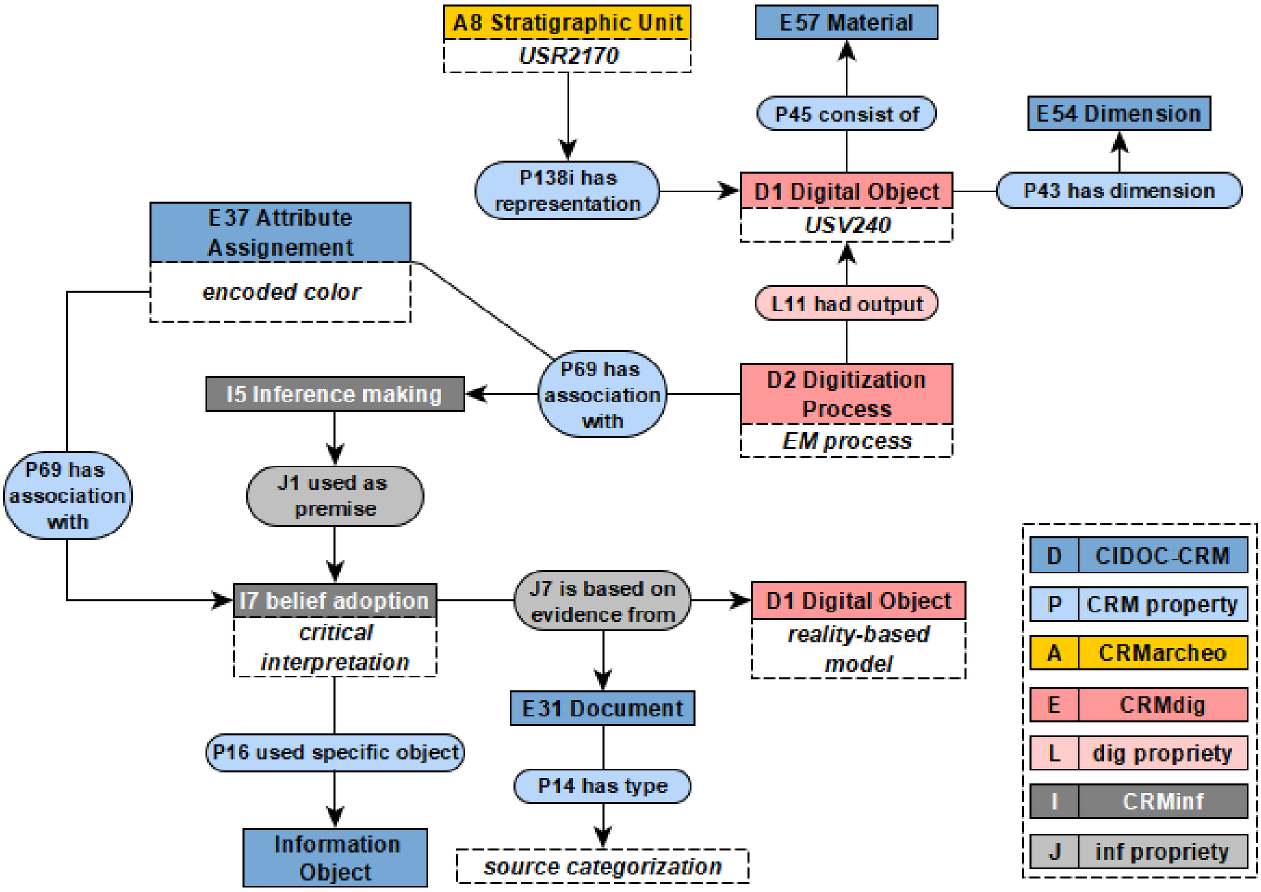

Figure 23: Relationship between virtual stratigraphic units and their representation

6.4 The contextualisation mapping

In this section we describe mapping the contextualisation examples to CIDOC CRM. The mapping of the architrave block (example 1) looks like this:

E22 Human-Made Object "**Block**"
- P45 consists of E57 Material "**marble**"
- P56 bears feature E25 Human-Made Feature "**ornamental relief**"
- P70i is documented in E73 Information Object "**Arachne page**"
  - P165i is incorporated in E73 Information Object "**Arachne**"
  - P165 incorporates E31 Document "**Wilson, Sicily**"
- P53 has former or current location E53 Place "**the orchestra of the theatre**"

P12i was present at E5 Event "**block falls down into the orchestra**"
P46i forms part of E22 Human-Made Object "**architrave**"

E22 Human-Made Object "**Architrave**"
P108i was produced by E12 Production
P4 has time-span E52 Time-Span
P86 falls within E52 Time-Span "**probably early 3rd century CE**"
P70i is documented in E31 Document "**Wilson, Sicily**"
P9i forms part of E7 Activity "the **second major building phase** of the Roman theatre"
P62i is depicted by E31 Document "**Le antichità della Sicilia … Antichità di Catana**"

E8 Activity "the **second major building phase** of the Roman theatre"
P9i forms part of E4 Period "the **Severan dynasty**"
P4 has time-span: E52 Time-Span
P86 falls within: E52 Time-Span "**193 to 235**"

E31 Document "**Book**"
"**Le antichità della Sicilia … Antichità di Catana**"
P70i is documented in E73 Information Object "**Arachne**"
P70i is documented in E73 Information Object "**Heidelberg Digital Library**"

The mapping of the inscription (example 2) looks like this:

E22 Human-Made Object "**Statue**"
P62 depicts E28 Conceptual Object "**the (mythical) guardian heroes of Catania**"
P53 has former or current location E53 Place "**Campus Piorum near Catania**"
P25i moved by E9 Move
P26 moved to E53 Place "**the Roman theatre**"
P31i was modified by E11 Modification "**either restored or replaced**"
P15 was influenced by E21 Person "**Merulus**"
P44 has condition E3 Condition State "**only the base still exists**"

E22 Human-Made Object "**Base**"
P46i forms part of E22 Human-Made Object "**statue**"
P53 has former or current location E53 Place "**the central sector of the lower seat rows**"
P54 has current permanent location E53 Place "**courtyard of the Museo Civico at Catania**"
P43 has dimension E54 Dimension "**width 56 cm**"
P43 has dimension E54 Dimension "**height 26,5 cm**"
P56 bears feature E25 Human-Made Feature

P128 carries E34 Inscription "**Inscription**"

E34 Inscription "**Inscription**"
P70i is documented in E73 Information Object "**Arachne**"
P70i is documented in E73 Information Object "**Last Statues of Antiquity database**"
P15 was influenced by E21 Person "**Merulus**"
P165 incorporates E33 Linguistic Object "**spectabilis**"
J7i is evidence for I7 Belief Adoption
J6 adopted I2 Belief "**inscription is dated not before 434**"

E21 Person "**Merulus**"
P14i performed E7 Activity "**consular governor of Sicilia**"
P4 has time-span E52 Time-Span
P173i ends after or with the start of E2 Temporal Entity "**the year 434**"
P86 falls within E52 Time-Span "**may or may not have been consular governor in the Ostrogothic period (493 to 553)**"

Note that for now we simply ignore uncertainties in statements such as "probably early 3rd century CE" or "may or may not have been consular governor in the Ostrogothic period (493 to 553)". This will be addressed in another modelling round.

6.5 Aligning the mappings

The mappings of the scenario and the EM complement each other. Rather than having to align different modelling decisions, it was sufficient to model the transitions between the mappings. An alignment of the vocabulary for describing the different building phases and its parts turned out to be unnecessary as well. (An earlier version of the alignment was described in Milestone MS37.)

The alignment is based on the LOD mapping of the LOD, which can be seen as a wrapper around the EM. Leaving out the parts of the mapping that are not relevant here, the LOD mapping of the architrave block (example 1) looks like this:

E22 Human-Made Object "**Block**"
P53 has former or current location E53 Place "**the orchestra of the theatre**"
P12i was present at E5 Event "**block falls down into the orchestra**"
P46i forms part of E22 Human-Made Object "**architrave**"

E22 Human-Made Object "**Architrave**"
P108i was produced by E12 Production
P9i forms part of E7 Activity "the **second major building phase** of the Roman theatre"

E7 Activity "the **second major building phase** of the Roman theatre"
P9i forms part of E4 Period "the **Severan dynasty**"

The marble block and the architrave can now be connected to the EM by modelling (1) the discovery of the block in the orchestra of the theatre as an S19 Encounter Event and (2) the relationship between the architrave as a physical object and as a virtual stratigraphic unit:

E22 Human-Made Object "**Block**"
P53 has former or current location E53 Place "the orchestra of the theatre"
O19i was object found by S19 Encounter Event
O21 has found at A2 Stratigraphic Volume Unit "the orchestra of the theatre"
P4 has time-span E52 Time-Span
P86 falls within E52 Time-Span "the 1980s"
P12i was present at E5 Event "block falls down into the orchestra"
P46i forms part of E22 Human-Made Object "architrave"

E22 Human-Made Object "**Architrave**"
P108i was produced by E12 Production
P9i forms part of E7 Activity "the second major building phase of the Roman theatre"
P165i is incorporated in D1 Digital Object "Architrave" (a virtual stratigraphic unit)

E8 Activity "the **second major building phase** of the Roman theatre"
P9i forms part of E4 Period "the Severan dynasty"
P70i is documented in SHE1 Dataset "The EM of the Roman theatre"

We can do the same exercise for the inscription (example 2). The relevant parts of the mapping are this:

E22 Human-Made Object "**Statue**"
P25i moved by E9 Move
P26 moved to E53 Place "**the Roman theatre**"
P44 has condition E3 Condition State "**only the base still exists**"

E22 Human-Made Object "**Base**"
P46i forms part of E22 Human-Made Object "**statue**"
P53 has former or current location E53 Place "**the central sector of the lower seat rows**"
P43 has dimension E54 Dimension "**width 56 cm**"
P43 has dimension E54 Dimension "**height 26,5 cm**"

It can be connected to the EM like this:

E22 Human-Made Object "**Statue**"
P25i moved by E9 Move
P26 moved to E53 Place "the Roman theatre"
P44 has condition E3 Condition State "only the base still exists"

E22 Human-Made Object "**Base**"
P46i forms part of E22 Human-Made Object "statue"
P53 has former or current location E53 Place "the central sector of the lower seat rows"
O19i was object found by S19 Encounter Event
O21 has found at A2 Stratigraphic V. Unit "central sector of the lower seat rows"
P4 has time-span E52 Time-Span
P86 falls within E52 Time-Span "1951"
P43 has dimension E54 Dimension "width 56 cm"
P43 has dimension E54 Dimension "height 26,5 cm"

A2 Stratigraphic Volume Unit "central sector of the lower seat rows"
P165i is incorporated in D1 Digital Object (a stratigraphic unit in the EM)

6.6 Mapping issues

Part of the alignment in a broader sense is the correction of modelling mistakes, including formalities. Sometimes properties sound right but cannot actually be used in the right circumstances. For example, BP11 "is connected to" from CRMba has the domain and range B2 Morphological Building Section, so it seems alright to use it like this:

BP4 Empty Morphological Building. BP11 is connected to: B3 Filled Morphological Building Section

However, *B3 Filled Morphological Building Section* is a subclass of *B2 Morphological Building Section* but *BP4 Empty Morphological Building* is not.

Another case was the connection between a virtual stratigraphic unit and its representation (D1 Digital Object). *P138i has representation*, which connects any *E1 CRM Entity* with a *E36 Visual Item*, did not seem correct. Perhaps *P165 incorporates* is more fitting, but it connects *E73 Information Object* with *E90 Symbolic Object*. This is still undecided.

Getting rid of issues like this is an ongoing process. Unless the mapping is available as RDF and checked against the RDF representation of CIDOC CRM as well as the relevant extensions, all involved mappings will contain mistakes like this, they are virtually unavoidable. Hence, the next step would be to create a Linked Data representation using a serialisation of RDF.

We were relieved to see that even the SSHOCro showcase mappings as well as SSHOCro itself did change after our discussions. For example, *SHE2 Knowledge Workflow Activity* used to be the superproperty, not only of the concrete work steps *SHE5 Data Collection*, *SHE6 Data Preparation and Connection* and *SHE7 Data Interpretation*, but also of the auxiliary activities *SHE8 Publication* and *SHE9 Data Storage*. We convinced them to restrict Knowledge Workflow Activity to the three work steps. The semantics of the connection between *SHE2 Knowledge Workflow Activity* and the two auxiliary activities is in fact, still not completely clear.

While mapping a dataset to CIDOC CRM tends to bring great clarity about the semantics of the data, there are many decisions that need to be made and pitfalls to avoid. One example from the ChronOntology period gazetteer highlights this: Periods in the sense of CIDOC CRM are always associated with a Spacetime Volume (STV), which consists of all points in space and time where the period happened. An STV has an associated place, namely the area where it happened at least at some point in time. For example, modern-day England belongs to the associated place of the period “Roman empire” as well as the sub-period “reign of Roman emperor Hadrian” because it was conquered by Hadrian, but not to the associated place of the sub-period “reign of Roman emperor Augustus” because England was not a part of the Roman empire yet. Thus, one can model:

E7 Activity “reign of Roman emperor Hadrian”
    P161 **has spatial projection** E53 **Place**
        P89i **contains** E53 **Place** “England”

In general, it is of course ahistorical to describe the extent of the Roman empire, or of any historical period, with modern-day countries, but it is very convenient because every modern geodata system such as Google Maps or OpenStreetMap can find them.

ChronOntology refers geodata to the idai.gazetteer. If a record in ChronOntology documents a culture that existed within the boundaries of modern-day Sudan but for which no specific entry exists in the gazetteer, this would be modelled in ChronOntology as:

“Meroë”
    spatiallyPartOfRegion: “Sudan”

which corresponds to this mapping in CIDOC CRM:

E4 Period “Meroë”
    P161 **has spatial projection** E53 **Place**
        P89 **falls within** E53 **Place** “Sudan”

Now, a common situation in ChronOntology is a record documenting a culture that existed in more than one country, for example in parts of modern-day Chad and Sudan. However, the statements in ChronOntology:

“Kingdom of Kush”
spatiallyPartOfRegion “Egypt”
spatiallyPartOfRegion “Sudan”

are interpreted as additive (meaning the combined area of Chad and Sudan) and do not translate directly into their CIDOC CRM counterpart, where they would be interpreted as subtractive (Kingdom of Kush is part of the area where both statements are true, which is nowhere).

SSHOCro, a CIDOC CRM extension still in the making, has some issues as well. For example it assumes that a cycle of Knowledge Workflow Activities, namely 1. Data Collection, 2. Data Preparation and Connection and 3. Data Interpretation, actually involves an interplay of the interpretation with the earlier steps. For example, when the data is a list of mentionings of person names in a text, the preparation can sort the mentionings by person. The interpretation can then find that two different people are in fact the same:

- Preparation: "Vodowice" and "Wodowice" are found in the text and and sorted as mentionings of two different persons
- Interpretation: the realisation that both spellings in fact denote the same person
- Preparation: from now on, "Vodowice" and "Wodowice" are treated as one person

These disparities are modelled by a string of properties called “follows” which are subproperties of *P134 continued*, with the scope note: “This property associates two instances of *E7 Activity*, where the domain is considered as an intentional continuation of the range. A continuation of an activity may happen when the continued activity is still ongoing or after the continued activity has completely ended. The continuing activity may have started already before it decided to continue the other one. Continuation implies a coherence of intentions and outcomes of the involved activities.”

While this makes perfect sense for the Knowledge Workflow Activities, CIDOC CRM also assumes that the continuing activity must start later than the continued activity. Logically, this implies that one can never get back to exactly the same instance of a Knowledge Workflow Activity. CIDOC CRM simply didn’t have cases in mind where two activities continue each other. Sorting out these issues without introducing any new issues is a very slow process. In this case, either CIDOC CRM can embrace this new use case of two activities continuing at the same time as each other and drop the requirement that the continuing activity must start later than the continued activity, or SSHOCro can simply declare the “follows” properties are no longer subproperties of *P134 continued* but of the less specific *P15 was influenced by*.

6.7 Modelling the data provenance

A central goal of SSHOCro is to model the provenance of data. The creators of SSHOCro have analysed the research workflow in the present task as follows (taken from D4.20):

> One of the issues the partners of SSHOC task 5.7 have decided to address is the transparency of virtual archaeology projects. Documenting all action points undertaken in the context of a 3D reconstruction allows validating the results of the reconstruction without having to duplicate the effort from scratch.
>
> The scientific research workflow identified by SSHOC task 5.7 documents the following steps:
>
> 1. Documentation: the literature review, which in this case extends to both written and iconographic sources related to the object of investigation, as well as structured data –such as archaeological excavation data.
> 2. Survey: the collection of graphic documentation data using techniques like laser-scanning, digital photogrammetry, photo and video shooting of the area of interest. Instrumentation is also specified.
> 3. Data Processing: establishing connections among collected data points, f.i. enhancing 3D models of in situ elements with 3D models of non-in situ objects, implementing stratigraphic readings using Harris Matrix relations (through the EM tool), etc.
> 4. 3D-Reconstruction Hypothesis: optimisation of 3D models, texturing, illuminations etc.; validation of the reconstructed models.

The exact relationship to the three Knowledge Workflow steps is not explicated in D4.20, but roughly speaking, step one and two translate to different instances of *SHE5 Data Collection*, step three translates to *SHE6 Data Preparation and Connection* and step four translates to *SHE7 Data Interpretation*. However, we are still in discussions about how well the Knowledge Workflow steps can capture the similarities and differences between handling physical objects (such as the Roman theatre) and the digital data that results from examining the physical objects, and where the exact boundary between data processing and data interpretation lies. SSHOCro draws the line between descriptive statistics (e.g. calculating a mean value) and inference statistics (e.g. calculating a p-value, which can be used to decide whether a statistical result is likely a real effect or just an artefact that would disappear when looking at more data). Interpretation further includes everything from the decision that a roof must have had a specific shape up to preparing a report for publication.

Thus, we are faced with two major obstacles: One, it is not straightforward to model the Knowledge Workflow steps, and two, quite often neither the intermediate datasets nor the Knowledge Workflow steps leading to them are documented. This makes modelling the data provenance quite difficult, and it is still work in progress. However, we have started modelling the relationship of the data that was gathered before SSHOC and the data that was gathered in SSHOC. In the case of the Roman theatre, there have been data collections and preparation steps in earlier projects that led to a first 3D virtualisation. In the present task, the data was prepared again and interpreted to result in a 3D model. This can be roughly modelled like this:

Earlier Project "**OpenCiTy**"

collected dataset **DC1 (no longer available)**

had dataset DC1 as input and prepared dataset **DP1 (available, but copyright issues)**

used dataset DP1 for interpretation to create **virtual reconstruction 1**

SSHOC task 5.7

had dataset DP1 as input and prepared dataset **DP1a (transfer to EM)**

collected dataset DC2 and prepared **DP2**

interpreted dataset DP2 to create **virtual reconstruction 2**

Compared the two virtual reconstructions

In CIDOC CRM / SSHOCro, it looks like this:

SHE3 SSH Project "**Earlier project OpenCiTy**"

SHR1 consists of SHE5 Data Collection

SHR10 produced dataset SHE1 Dataset **DC1 (no longer available)**

SHR13i is followed by SHE6 Data Preparation Connection

SHR2 consists of SHE6 Data Preparation and Connection

SHR31 had input for preparation SHE1 Dataset DC1

SHR11 prepared dataset SHE1 Dataset **DP1 (still available, but copyright issues)**

SHR16i is followed by SHE7 Data Interpretation

SHR3 consists of SHE7 Data Interpretation

SHR12 used dataset for interpretation SHE1 Dataset DP1

SHR30 created outcome SHE1 Dataset "**virtual reconstruction 1**"

SHE3 SSH Project "**SSHOC task 5.7**"

SHR2 consists of SHE5 Data Preparation and Connection

SHR31 had input for preparation SHE1 Dataset DP1

SHR10 produced dataset SHE1 Dataset **DP1a (transfer to the EM)**

SHR16i is followed by SHE7 Data Interpretation B

SHR1 consists of SHE5 Data Collection

SHR10 produced dataset SHE1 Dataset **DC2 (new data)**

SHR13i is followed by SHE6 Data Preparation Connection

SHR2 consists of SHE6 Data Preparation and Connection

SHR31 had input for preparation SHE1 Dataset DC2

SHR11 prepared dataset SHE1 Dataset **DP2**

SHR16i is followed by SHE7 Data Interpretation A

SHR3 consists of SHE7 Data Interpretation A

SHR12 used dataset for interpretation SHE1 Dataset DP1a

SHR12 used dataset for interpretation SHE1 Dataset DP2

SHR30 created outcome SHE1 Dataset "**virtual reconstruction 2**"

SHR3 consists of SHE7 Data Interpretation B
SHR24 based on SHE5 Data Interpretation A
SHR12 used dataset for interpretation SHE1 Dataset “virtual reconstruction 1”
SHR12 used dataset for interpretation SHE1 Dataset “virtual reconstruction 2”
SHR30 created outcome “**comparison of the virtual reconstructions**”

This produces a provenance chain for the datasets and interpretations:

OpenCiTy

(DC1) → DP1 → using DP1 for virtual reconstruction 1

SSHOC task 5.7

↳ DP1a
→ using DP1a and DP2 for virtual reconstruction 2
DC2 → DP2
↳ comparison of virtual reconstructions 1 and 2

Tracking the data provenance is still a work in progress. And even on the level of the ontology itself it is an ongoing discussion as to how the data provenance should be tracked: Is the knowledge about the data provenance encoded in the “follows” properties, or is it encoded in the chain of output/input connections between datasets?

Another issue is incomplete knowledge: One can only model what is known. SSHOCro, and CIDOC CRM in general, is designed to accommodate incomplete knowledge. For example, if the data only exists in a processed form, then one can either model

SHE5 **Data Collection** [which may have e.g. a place and a time]
produced SHE1 **Dataset** [about which nothing is known]
SHR31i **was input by** preparation SHE6 **Data Preparation and Connection**
SHR11 **prepared** dataset SHE1 **Dataset** “refined data”

or model the collection and preparation steps as a single *SHE2 Knowledge Workflow Activity*:

SHE2 **Knowledge Workflow Activity**
SHR5 **had output** SHE1 **Dataset** “refined data”

If the data provenance is encoded in the output/input change and if there should be the possibility to encode incomplete knowledge, then there should be SSHOCro properties that fulfil this encoding even for generic Knowledge Workflow Activity steps. SSHOCro only offers a generic property *SHR5 had output*, which is not constricted to *SHE1 Dataset*. So far, SSHOCro does not have a property "had dataset output" that would be more specific than *SHR5 had output* but less specific than *SHR10 produced dataset* and *SHR11 prepared dataset*, which can only be used with the specific Data Collection and Data Preparation and Connection classes.

The position in the hierarchy of the hypothetical new properties is shown in the following table, which is adapted from the SSHOCro Properties Hierarchy in D4.20:

P15 was influenced by (influenced)				E7 Activity	CRM Entity
—	P16 used specific object (was used for)			E7 Activity	E70 Thing
—	—	SHR4 had input (was input of)		SHE2 Knowledge Workflow Activity	E70 Thing
—	—	**[hypothetical property] SHRxx had dataset input (was dataset input of)**		**SHE2 Knowledge Workflow Activity**	**SHE1 Dataset**
—	—	—	SHR12 used dataset for interpretation (dataset was used by interpretation)	SHE7 Data Interpretation	SHE1 Dataset
—	—	—	SHR31 had input for preparation (was input by preparation)	SHE6 Data Preparation and Connection	SHE1 Dataset
—	—	SHR5 had output (was output of)		SHE2 Knowledge Workflow Activity	E70 Thing
—	—	—	SHR30 created outcome (outcome was created by)	SHE7 Data Interpretation	E73 Information Object
—	—	**[hypothetical property] SHRxx had dataset output (was dataset output of)**		**SHE2 Knowledge Workflow Activity**	**SHE1 Dataset**
—	—	—	SHR10 produced dataset (dataset was produced by)	SHE5 Data Collection	SHE1 Dataset
—	—	—	SHR11 prepared dataset (dataset was prepared by)	SHE6 Data Preparation and Connection	SHE1 Dataset

7. Implementation as Linked Open Data

The contextualisation (also called scenario) of the Roman theatre will made available as Open Data, i.e. it will be openly accessible and licensed under an open license. Furthermore, it will be implemented as Linked Data (LD) based on the CIDOC CRM mappings from the last chapter. Taken together, it will be available as Linked Open Data (LOD). An implementation as LOD in a triplestore that can be queried via SPARQL is under construction.

This implementation creates several new entry points for examining the 3D model. One can search for concrete objects in the model, but also for object types and object sizes. One can also search for CIDOC CRM classes, properties and modelling paths as well as the types of the building parts used in the EM or its contextualisation. If more 3D models are embedded in this way, these models can also be searched and compared at a semantic level. One can search for 3D models based on place, the people involved or the time period. If one has found a 3D model, one can search for other models that share connections to the same authority record, such as a place record in a gazetteer, or a time period or event in a time gazetteer. And it is a starting point for a semantic comparison of the first and second virtual reconstructions of the Roman theatre.

The contextualisation relies heavily on the *iDAI.world,*[22] a system of open access Web Services for archaeological data delevoped by the DAI as open source software. Two categories of services are relevant here, namely data services including *Arachne*, is the overarching object database of the DAI (which we have already linked to in the examples) and authority record services for data interoperability and standardisation, notably *idai.gazetteer* for place-related data and idai.chronontology aka *ChronOntology*[23] for temporal terms.

New authority records for place will created in the the idai.gazetteer for places, and in ChronOntology for periods and Arachne for objects. People do not have an implentation yet. However, if they can be represented by a period, such as the “the reign of Merulus” (but not “the life of Merulus”, and even adding his reign is dubious as there is virtually nothing known about it apart from the *terminus post quem* for the end of his reign), they are added to ChronOntology as well. We may also add a catalogue of inscriptions.

Early versions of the authority records that were newly created for the scenario have been added to the Gazetteer and ChronOntology, but so far only in test versions, and in a static version in the DAI-SSHOC GitHub repository. Since other projects may link to these authority records and will expect the record to be available indefinitely, there is virtually no way to remove data that later turns out to fail to meet a certain quality standard. For this reason, we will have another curation round before publishing the data. While the latest versions of the data will always be hosted on GitHub, the publication of the data in the authority control (aka norm data) systems is planned for May/June 2022.

[22] iDAI.world website, overview of all DAI webservices: https://idai.world/ [24 Mar 2022]

[23] iDAI.chronontology website: https://chronontology.dainst.org/ [24 Mar 2022]

The progress is documented in the DAI-SSOH GitHub repository[24], which all newly created authority records, the documentation (including the modelling) and complete data of the LOD triple store, and a documentation of the ongoing work on the CIDOC CRM extension for temporal gazetteers based on the ChronOntology data model.

It also contains the ChronOntology documentation that was announced in D5.17. This work has been done in collaboration with the Entangled Africa project[25]. We would like to end this chapter with an overview how Chronontology incorporates concepts from CIDOC CRM, which makes it well suited for handling the period data of the scenario.

ChronOntology is a gazetteer for temporal terms, i.e. a system for storing, managing, mapping and publishing temporal terms (i.e. periods such as Middle Helladic). ChronOntology allows a rich semantic modeling of various terminological systems for cultural periods. The data model can describe their semantic relations and is based on CIDOC CRM. A central tenet is the separation of a temporal term definition from any dating information about the term. For example, one can denote the reign of Augustus without stating or agreeing on strict start and end dates. Only the definition determines whether one is talking about the same thing, and temporal information is meaningful only in connection with the definition. However, since strict period definitions are often not the focus of the research literature, ChronOntology currently represents them by a system of period types such as “political” or “material culture” or subtypes such as “pottery”. The view of periods as concepts with separate definitions and dating information complements other approaches to period gazetteers where spatiotemporal information is seen as part of the definition of the terms themselves.

Each defined temporal term is automatically associated with a phenomenological *spacetime volume* (STV), i.e. the area in space and time where it happened, regardless of what is actually known about its extent. STVs allow for periods such as “Neolithic” to take place at different times in different regions. (Note, however, that the usage of terms is often not consistent. For example, “Neolithic in the Levant” might seem to denote the same as “Neolithic, limited in a strictly geographic sense to the Levant”, but has in fact a different definition.) Any concretely given spatial and temporal information is an approximation of the phenomenological spacetime volume. ChronOntology uses widgets for space and time to visualise the concretely given information. ChronOntology itself concentrates on the temporal aspect, while the geographical aspect is handled by DAI’s gazetteer for place names idai.gazetteer[26].

The documentation of a complex system is always inherently a work in progress, and efforts will continue. After having a very close look at SSHOCro, we decided to create a similar ontology for modelling time periods.

[24] https://github.com/dainst/sshoc
[25] https://www.dainst.blog/entangled-africa/
[26] iDAI.gazetteer website: https://gazetteer.dainst.org [24 March 2022]

8. Cooperation in SSHOC and Dissemination

The application of the FAIR principles to the case study was discussed with Holly Wright (task 5.6, ADS branch). Joe Padfield's (task 5.6, NG branch) work on CIDOC CRM, triples stores and visualisation helped us shape our own mapping. We have discussed the LOD scenario with him and are using his tools in our own triple store.

The mapping of the EM was created in collaboration with Elisabetta Caterina Giovannini from the Politecnico di Torino. We also had discussions with the SSHOCro creators Chrysoula Bekiari, Athina Kritsotaki and Eleni Tsouloucha (FORTH), which led to several changes of the SSHOCro and improved documentation. At an early stage of the project we discussed the possible use of Dataverse with Cesare Concordia. We will continue this discussion once the LOD triple store is deployed.

The work of the task was presented on several occasions. The first occasion was a workshop by task 5.7 ("SSHOC archaeological case study Workshop – The Roman theatre in Catania from survey to interactive 4D visualisation") in May 2021. In the first part of the workshop we presented the case study as a work in progress. We looked at the data and systems of the three partners, including the existing and the new visualisation of the Roman theatre and the EM system. We concluded the first part with the integration process of systems and data leading to a unified workflow from survey data to an interactive 4D visualisation. In the second part we situated the case study in the broader context of the SSHOC project and beyond. Parallel presentations reported on the FAIR principles in archaeology (Holly Wright), SSHOCro (Athina Kritsotaki) and the EMViq tool that accompanies the EM system (Bruno Fanini). The workshop ended with a discussion where to go from here and especially how to integrate the FAIR principles and the reference ontology in our work. The workshop is described in more detail in Milestone MS36.

The work has also been presented at other workshops and conferences. One was a workshop of the French ministry of Culture in Paris in March 2022. The CIDOC CRM integration in EM was presented during the Extended Matrix Conference held online on February 4th 2022.

A planned face-to-face workshop towards the end of the project had to be cancelled due to the still ongoing pandemic, and because the Berlin partner was not able to travel.

9. Conclusion

Our work in this task has demonstrated that working towards freely available and re-usable data according to the FAIR principles is a time-consuming but worthwhile pursuit. Several issues had to be resolved, most notably the "data rot" of the earlier projects about the Roman theatre, the transition from proprietary data to open data and systems and copyright issues for Italian heritage data. We would therefore advise other projects to resolve any copyright issues as soon as possible. We would also advise other projects to take the FAIR principles seriously from the beginning to the end of their project. Sticking to these principles is not just another burden but really helps liberate your data.

There were experts on CIDOC CRM available in SSHOC, namely the creators of SSHOCro, but we hesitated in showing them our half-finished mappings. In retrospect, we realise that we should have shared our mappings much earlier. We would advise other projects to overcome this hesitation and present their work in progress early. It also became clear that we need to create an RDF representation of our mappings using a tool such as 3M[27] and validate it against the RDF representations of CIDOC CRM and its extensions. As for tracking the data provenance, we started with a first mapping of the processing steps and the resulting datasets to SSHOCro, and we will continue to model it in more detail and discuss the modelling issues with the creators of SSHOCro.

Although the overall structure of the task remained the same, the exact workflow of the case study partially changed as a result of our increased knowledge about the resource. We have fulfilled most of the work as planned in D5.17, but there are a few areas we did not pursue, most notably creating excavation data in iDAI.field. However, we are planning a physical meeting as soon as Covid permits – hopefully later in 2022 – with a hackathon to tie up loose ends. One topic would be to create sample data of the Roman theatre in iDAI.field and to connect it with the EM. This would truly create a workflow from excavation data via archaeological survey data, to 4D visualisation.

At the same time, we created some additional steps to the task. For example, looking in detail at SSHOCro has motivated the Berlin partner to create an ontology for modelling time periods. This will be based on ChronOntology but will accommodate the data models of other time gazetteers, such as PeriodO[28]. This work has already begun and will be documented in the DAI-SSHOC GitHub repository.

[27] See: https://isl.ics.forth.gr/3M/ [April 2022]
[28] See: https://perio.do [April 2022]

10. References

Bekiari, C., Bruseker, G., Doerr, M., Ore, C.E., Stead, S., Velios, A. (2021). Definition of the CIDOC Conceptual Reference Model (vol.A). https://CIDOC CRM.org/sites/default/files/cidoc_crm_v.7.1.1_0.pdf [also cited as CIDOC CRM]

Bekiari, C., Kritsotaki, A., Tsouloucha, E. & Theodoridou, M. (2022). D4.20 SSHOCro. Zenodo.

Branciforti, M.G. (2010), Da Katane a Catina, in: M.G. Branciforti, V. La Rosa (eds.), Tra lava e mare. Contributi all'archaioloqhia di Catania, Catania, pp. 183-209.

Branciforti M.G., Pagnano, G. (2008), Il complesso archeologico del Teatro e dell'Odeon di Catania, Palermo.

Buda G. (2015), Teatro antico di Catania. Lavori tra il 2014 e il 2015. In: F. Nicoletti (ed.), Catania antica. Nuove prospettive di ricerca, Palermo, pp. 247-279.

BUSCEMI F. (2012), Architettura e romanizzazione della Sicilia di età imperiale: gli edifici per spettacoli, Palermo. CARRERA P. 1639, Delle memorie historiche della città di Catania, Catania.

Demetrescu, E. (2015). Archaeological Stratigraphy as a formal language for virtual reconstruction. Theory and practice. J. Archaeol. Sci. 57, 42–55.

Demetrescu, E. & B. Fanini (2017), A White-Box Framework to Oversee Archaeological Virtual Reconstructions in Space and Time: Methods and Tools. Journal of Archaeological Science: Reports 14: 500–514, 2017.

Fanini, B. & Demetrescu, E. (2018). Carving Time and Space: A Mutual Stimulation of IT and Archaeology to Craft Multidimensional VR Data-Inspection. In International and Interdisciplinary Conference on Digital Environments for Education, Arts and Heritage (pp. 553-565). Springer, Cham.

Fanini, B., Pescarin, S., & Palombini, A. (2019). A cloud-based architecture for processing and dissemination of 3D landscapes online. Digital Applications in Archaeology and Cultural Heritage, 14.

Fanini, B., Ferdani, D., Demetrescu, E., Berto, S., & d'Annibale, E. (2021). ATON: An Open-Source Framework for Creating Immersive, Collaborative and Liquid Web-Apps for Cultural Heritage. Applied Sciences, 11(22), 11062.

Gabellone, F., Ferrari, I. & Giuri, F. (2014), Methodological Approaches and ICT Solutions for Smart Cities, in Cultural Heritage and New Technologies November 3–5, pp. 1-13.

Gros, P.P. (1995). Hercule à Glanum: Sanctuaires de transhumance et développement urbain. Galliapp. 311–331.

Malfitana, D., Gabellone, F., Cacciaguerra, G., Ferrari, I., Giuri, F., Pantellaro, C. (2016), Critical reading of surviving structures starting from old studies for new reconstructive proposal of the Roman theatre of Catania, in Proceedings of the 8th International Congress on Archaeology, Computer Graphics, Cultural Heritage and Innovation 'ARQUEOLÓGICA 2.0' in Valencia (Spain), Sept. 5 – 7, 2016. Available online: http://itlab.ibam.cnr.it/new_itlab/wp-content/uploads/2017/12/Arqueologica-Critical-reading-of-surviving-structure-Theatre-of-Catania.pdf;

Medri, M.M. (2003). Manuale di rilievo archeologico. Roma-Bari.

Pensabene, P. (2005). La decorazione architettonica del Teatro di Catania, in: R. Gigli (ed.), Megalai Nesoi. Studi dedicati a Giovanni Rizza per il suo ottantesimo compleanno, Catania, pp. 187-211.

Schmidle W., Pescarin S., Demetrescu E., Fanini B., Gabellone F., Bucciero A., Chirivì A., Ferrari I., & Giuri F. (2020). D5.17 Implementation plan for the archaeological case study. Zenodo. https://doi.org/10.5281/zenodo.4558290

Wilson, R.J.A. (1996), La topografia della Catania romana, in: B. Gentili (ed.), Catania Antica. Atti del convegno della SISAC, Catania 23-24 Maggio 1992, Pisa-Roma, pp. 157-163.

Wilson, R. J. A. (1990), Sicily under the Roman Empire: the archaeology of a Roman province, 36 BC-AD 535, Warminster.

Wright, H., Moore, R. & Evans, T. (2022). D5.15 Report on opening access to research data in the archaeology domain. Zenodo.

List of Figures