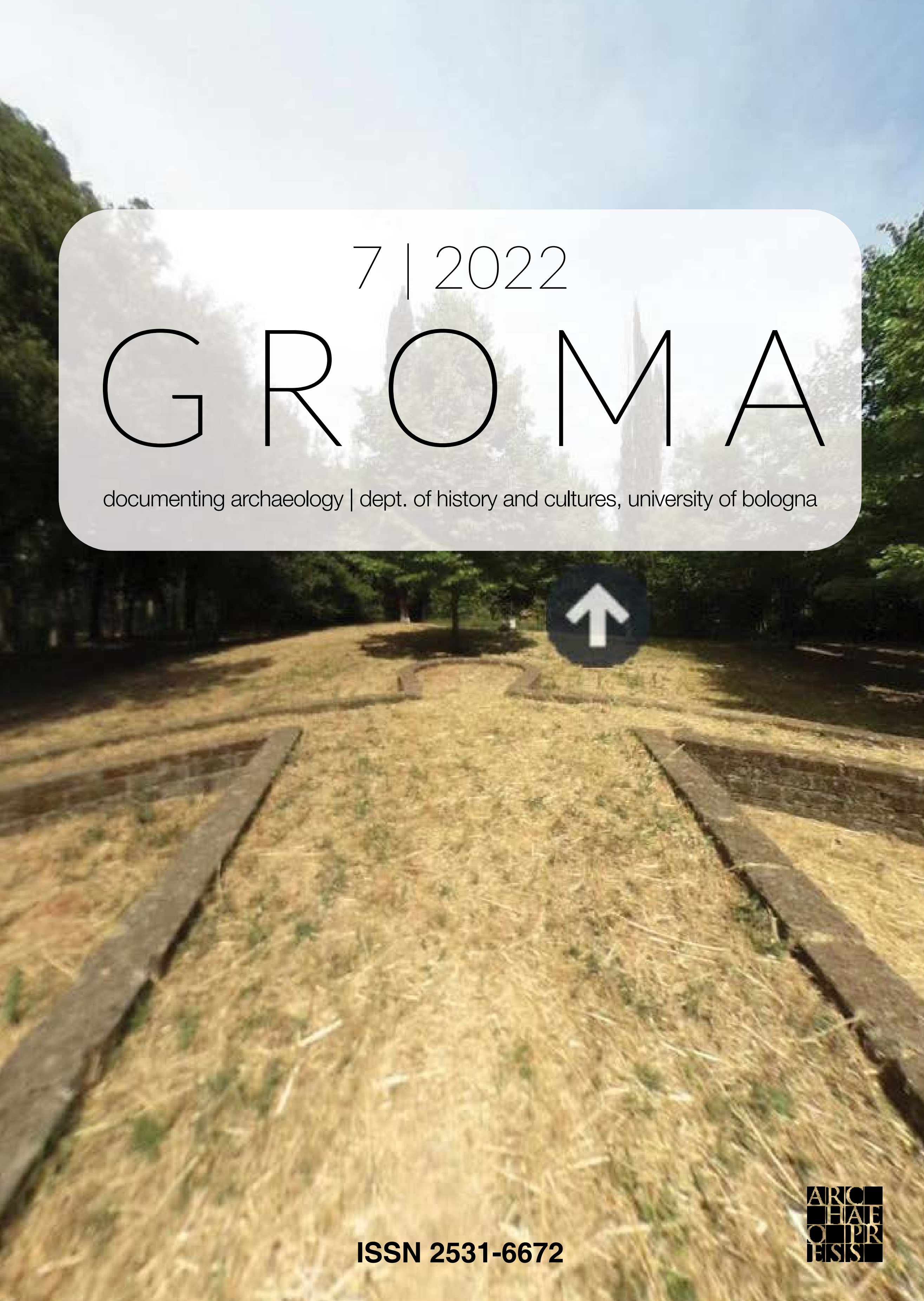

7 | 2022

GROMA

documenting archaeology | dept. of history and cultures, university of bologna

ISSN 2531-6672

ARCHAEOPRESS

GROMA. Documenting Archaeology
An open access, peer-reviewed journal founded by Enrico Giorgi and supported by the University of Bologna.

To submit an article visit https://archaeopresspublishing.com/ojs/index.php/groma and/or contact us at disci.groma@unibo.it

GROMA

Issue 7 2022

Proceedings of ArchaeoFOSS XV 2021

Archaeopress Journals

Archaeopress Publishing Ltd
Summertown Pavilion
18-24 Middle Way
Summertown
Oxford OX2 7LG
www.archaeopress.com

ISBN 978-1-80327-698-4
ISSN 2531-6672

Contents

Preface

Already with the previous, the seventh issue of *Groma. Documenting Archaeology*, many novelties were introduced, the most evident being the change of the publishing house. Archaeopress provided a new home to our open-access journal after the passage from Badypus which has played a fundamental role in the creation and improvement of this editorial project. Bradypus gave voice in 2016 to a small group of archaeologists and researchers of the University of Bologna under the guide of Enrico Giorgi, engaged in experimenting with innovative methodologies and technologies for the documentation of the archaeological contexts and data and more in generally dealing with the topics of landscape archaeology. The project rapidly evolved into a journal indexed as A-class by the Italian National Agency for the Evaluation of the University System (ANVUR) and it is now based in two of the most active Italian universities for archaeological research: the Bologna University and the Sapienza University of Rome, with a renovated international scientific committee.

The seventh issue marks these important changes and at the same time represents a return to the origins, since it publishes the research endeavours of young researchers strongly believing that methodological and technological evolution in archaeological research (and not only) should be based on strong ethical ground, i.e., on the paradigm of Open Science and on the free and open-source software and processes. Domizia D'Erasmo (LAD, Sapienza University of Rome), Cristina Gonzalez-Esteban (University of Southampton), Paolo Rosati (LAD, Sapienza University of Rome), Matteo Serpetti (Univertsità degli Studi della Tuscia) and Livia Tirabassi (University of Ghent) are the organisers of the 15th edition of the ArcheoFOSS International conference on open software, hardware, processes, data and formats in archaeological research (https://archeofoss.org) and the editors of the papers collected in this volume. As editors of Groma, we look with great interest and gratitude to their work, for their competencies and most importantly for their enthusiasm which is by far the most important factor of the success of Groma.

Julian Bogdani
Enrico Giorgi

An open source platform addressing structural stability risk assessment in historical centres

Renzo Carlucci[1], Charilaos Maniatakis[2], Philip Fayad[1], Andrea Di Iorio[1], Nikolaus Sketakis[1], Costantine Spyrakos[2], Haris Saroglou[2], Alessio Di Iorio[1], Alexandros Paraskeuas[3] and Nikos Papadopoulos[4]

[1]Alma Sistemi srl Guidonia (RM) Italy; [2]National Technical University of Athens Greece; [3]Geosystem Hellas, Greece; [4]FORTH Hellas

*Correspondence: rca@alma-sistemi.com

Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract:

The management of stability risk assessment in historical centres in a project that aims to design and develop an open platform where blocks of buildings and large structures may be represented in damage maps before the event occurs, addressing damage forecasts for seismic movements impacting on the structural stability of the CH. The structural stability is calculated using a Matlab code developed at the National Technical University of Athens, translated to an open source programming language (Python) using the software tool GeoPandas to enable operations that would normally require spatial database software (PostGIS). Furthermore, the use of open source technologies like QGIS desktop software will provide open and free access to view and contribute in the future.

Keywords: cultural heritage; open source; european project; risk assessment.

FOSS software used and licence: QGIS (*GNU General Public License*), PostGIS (*GNU General Public License (GPLv2 or later)*)

Introduction

STABLE (STructural stABiLity risk assEssment, 2018-2022, Horizon2020, MSCA - Marie Sklodowska Curie Actions - RISE - Research and Innovation Staff Exchange - grant agreement nº. 823966) has the goal of developing a digital platform to forecast earthquake damage to Historical City Centers in Europe combining different techniques and methodologies: structural stability models, earthquake simulation tools combined with geotechnical data, remote sensing and in situ monitoring technologies.

The STABLE team is composed of seven different partners from Italy (Sapienza University, Rome; University of Tuscia, Viterbo), Greece (National Technical University and Geosystem Hellas from Athens and FORTH from Rethymno) and Cyprus (Frederick University and System Space Solution Ltd), with specific expertise, coordinated together by ALMA Sistemi Srl (Italian SME), and with a plan of collaborations and staff exchanges in order to allow interdisciplinary research.

Some emblematic case studies of buildings from different countries were selected as test cases: a few blocks in the city centre of Rieti, Italy; city blocks in Nafplion and the Lavrion area, Greece; a few buildings in the central region of Strovolos, a district close to Nicosia, Cyprus; and, recently, some buildings in the historic centre of Rethymno in Crete, Greece.

The goal of the project is to define a risk modelling system through the integration of different data, from SAR satellite images to geological and geotechnical information, in order to create deformation soil maps. Additionally, we use free and open source solutions, as far as possible, to achieve all the above.

The risk map generation

The study aims to investigate the vulnerability of the architectonic heritage and to combine the results with hazard scenarios, as a first fundamental step in the seismic risk mitigation process. The two key elements of a vulnerability analysis are the capacity (strength and deformation) of a structure and the seismic demand. From these elements it is possible to predict expected vulnerability functions of possible damage to the structure as an effect of the seismic input. Corresponding fragility curves will be determined, expressing the probability of a structure, belonging to a certain class, reaching or exceeding a particular damage grade given a deterministic estimate of the spectral displacement caused by an expected hazard (e.g. Kappos et al. 2008).

The expected risk is determined according to the following expression:

Risk = *f* (Hazard, Vulnerability)

where:

Hazard: seismicity of the region and accepted probability of exceedance of the destructive action

Vulnerability: capacity of the structure as described below.

The *f* function is dependent on the various methods for the assessment of the structure's vulnerability and the hazard of the soil. The type of method chosen depends mainly on the objective of the assessment, but also on the availability of data and technology. In general, a vulnerability function is a relationship which defines the expected damage on a structure or a class of structures as a function of the ground motion (fig. 1). The two key elements of a vulnerability analysis are the capacity of the structure and the seismic demand. In order to estimate the damage (*D)*, the ability of the structure to resist constraints (capacity of the structure) must be compared with the constraints on the structure due to the earthquake ground motion (seismic demand).

To express the seismic demand, until very recently, the "intensity" was used nearly exclusively. This is a descriptive parameter of an earthquake based on observations of the effects of the earthquake on the environment. It has the advantage that historical data on earthquakes are available. However, information on the real ground movement is lost, and empirical relationships between intensity and peak ground acceleration vary a lot. Some methods use peak ground acceleration as the parameter defining the earthquake. Nevertheless, in that case, not only the information on the duration of the earthquake is lost, but also the information on the frequency content. Thus, a better parameter is the spectral acceleration (*Sa)* or the spectral displacement (*Sd)*.

The estimation of damage probability will be performed by means of a simplified mechanical methodology (Lagomarsino and Giovinazzi 2006). This method provides results of satisfactory accuracy for relatively simple constructions, while the assessment of structural capacity for more complex constructions, such as monuments, requires more elaborated methodologies (Spyrakos et al. 2015). Initially, the required input parameters will be collected, including:

- *Structural data* (typology, construction period, use of the building, post-earthquake damage data, detection of different construction phases, etc.)
- *Geometrical data* (total height, number of storeys, perimeter, footprint area, average bearing wall thickness of masonry structures, plan view of at least the ground floor, etc.)
- *Satellite data* (coordinates of each building, number of building blocks, number of buildings in each building block)
- *Vulnerability data* (estimation of vulnerability by rapid assessment methods related to the evaluation of conservation of the structure and materials used)

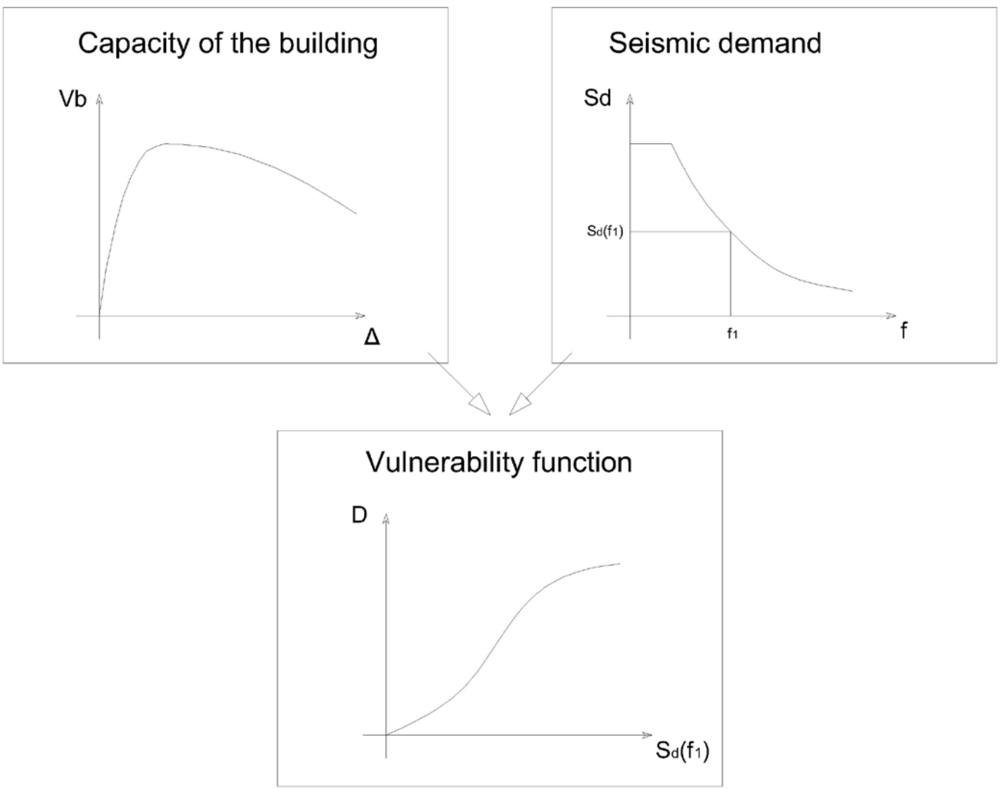

Figure 1. Principle of a Vulnerability function

The Structural Model Processor

This program calculates the risk map of the building stroke of an area for a given seismic hazard by applying a simplified mechanical method. The risk is defined as the probability of a certain level of damage being exceeded. Four levels of damage are examined, namely P1, P2, P3 and P4. The level of damage Pi increases with increasing i = 1 to 4. The probability takes values from zero (totally unlikely) to one (certainty).

All the calculation process has initially been done in MATLAB, a programming platform designed to allow expressions of computational mathematics, which for our experimental process means facilitating the use of mathematical complex functions. With this, it has been possible to analyse data, develop algorithms and create models and applications, with the final aim of defining the capacity curve and the exact displacements related to specific damage states, which may be derived through sophisticated analysis of a selective building case-study.

The code from MATLAB (wrote by C. Maniatakis) has been translated to an open source programming language (Python) using the software tool GeoPandas (translation by N. Schetakis). GeoPandas enables the user to easily do operations in Python that would normally require spatial database software. PostGIS is used for the webGIS representation of the final risk map. By translating the code to Python we further ensure the open source capability.

The code requires as inputs both Structural data and seismic spectral data.

The structural data

The structural data related to the building to be examined are included in a Geopackage file containing information on the geometry, average height (Z) and different typologies of the building, stored in the parameters Type 1, Type 2, Type 3, Type 4, Type 5, Type 6, Use, WT. GeoPackage is an open, standards-based, platform-independent, portable compact format for transferring geospatial information.

The following tables better explain the meaning of each parameter:

Typo 1 - general classification

Typologies	**Building types**	**Type 1**
Unreinforced Masonry	Rubble stone	1
	Adobe (earth bricks)	2
	Simple stone	**3**
	Massive stone	4
	U Masonry (old bricks)	5
	U Masonry – R.C. floors	6
Reinforced/confined masonry	Reinforced/confined masonry	7

Reinforced Concrete	Concrete Moment Frame	8
	Concrete Shear Walls	9
	Dual System	10

Typo 2 - type of horizontal structure (diaphragm/ slab/ vault)

Typologies		**Type 2**
Unreinforced and/or reinforced Masonry	wood slabs	1
	masonry vaults	2
	composite steel and masonry slabs	3
	reinforced concrete slabs	4

Typo 3 - presence or not of a soft first story in reinforced concrete structure.

Typologies		**Type 3**
Reinforced Concrete	soft first storey - pilotis	1
	regular along height	2

Typo 4 - presence or not of infill walls in reinforced concrete structure.

Typologies		**Type 4**
Reinforced Concrete	infill walls	1
	no infill walls	2

Typo 5 - ductility class.

Ductility class		**Type 5**
	without ductility class	1
	low ductility class	2
	medium ductility class	3
	high ductility class	4

Typo 6 - irregularity along height for every structure related to a reinforced concrete structure.

Typologies		Type 6
Reinforced Concrete	irregular	1
	regular	2

Use - main use of the structure.

Typologies		Use
Reinforced Concrete	residential	1
	public	2

WT – The average bearing wall thickness for masonry structures.

If the WT field is unavailable, a value of WT = 0.60 is assigned by the code, assuming an average thickness of 60 cm for bearing walls.

The Seismic hazard data

The seismic hazard data are included in a Geopackage file that includes different seismic acceleration spectra as a grid of points within the area of interest. The side of the grid is proposed to be 5.0 m x 5.0 m. Each point corresponds to an acceleration spectrum which resulted from a seismic hazard analysis, considering the seismic sources and the local soil effect (e.g. Tselentis et al. 2010). In case of active faults at a short distance from the buildings, a more detailed approach would include the examination of the particular phenomena that take place in the nearby field of active faults, the so-called near-fault area (Spyrakos et al. 2008; Maniatakis and Spyrakos 2012).

The Geopackage file includes information on the Geometry and the Spectrum type defined for each point and included in separate **"Spectrum" TXT** files, where the "name" is the corresponding value included in the relevant field of acceleration shapefile referred to geohazard. For example, if four areas with different seismic hazards are defined, the TXT files 1, 2, 3 and 4 should also be provided. The content of the TXT files is formatted into two columns, with the first including period values and the second spectral acceleration values in g units, without any text, as shown below:

0.01	0.3241
0.02	0.3243
...	
...	
4.0	0.0203

The acceleration values should be provided in the range of 0 ~ 4.0 seconds.

The final output

The code may be executed, provided that the above-mentioned input data are available in the same folder with the code and four figures in TIFF format; namely P1, P2, P3 and P4.

Every figure depicts with colours the value of probability Pi for the given seismic scenario. In figure 2, indicative results of the code are presented for probabilities P1 and P2. The results regard a seismic scenario that assumes peak ground acceleration equal to a_{gR}=0.30 g on rigid bedrock.

The GIS platform

All the data derived will be stored in the STABLE GIS platform. More specifically, and concerning the processing of the Structural Model processor outputs, automated workflows are used to convert the Georeferenced Fragility curves to actual classified Structural Vulnerability features in an open format.

The basic subject of analysis is a polygon feature covering each building of study. For the total sum of buildings, a classification process should be implemented in order to export the final

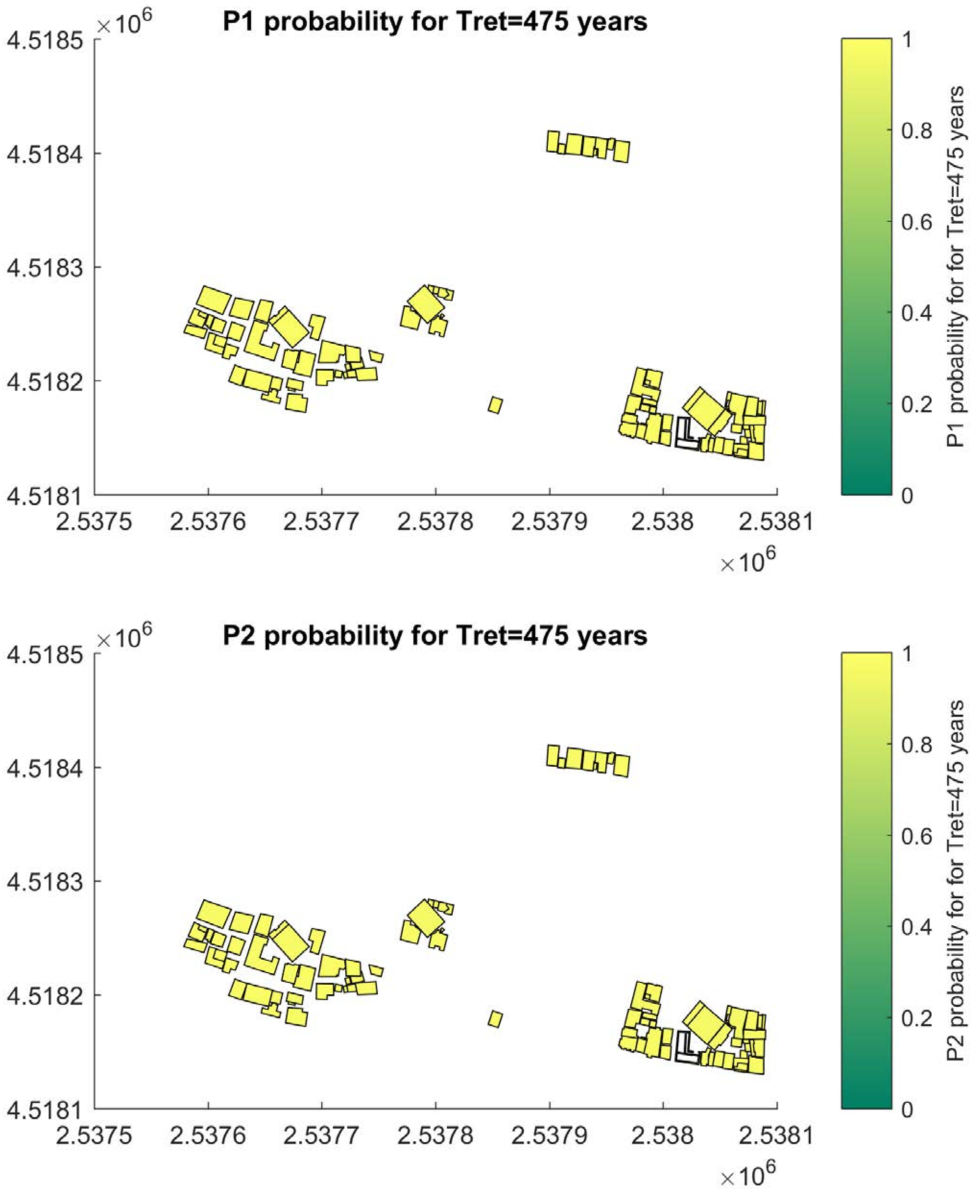

Figure 2. Indicative output of the code for the test case in Nafplion (Greece)

Structural Vulnerability map. Essentially, each building/feature is classified according to its assigned fragility value from the previous Structural Model processor step.

As illustrated in figure 3, the automated process through the SML operator takes as input building polygons, deriving from the Structural Model analysis, and extracts different classes of these buildings according to their vulnerability level.

The case studies

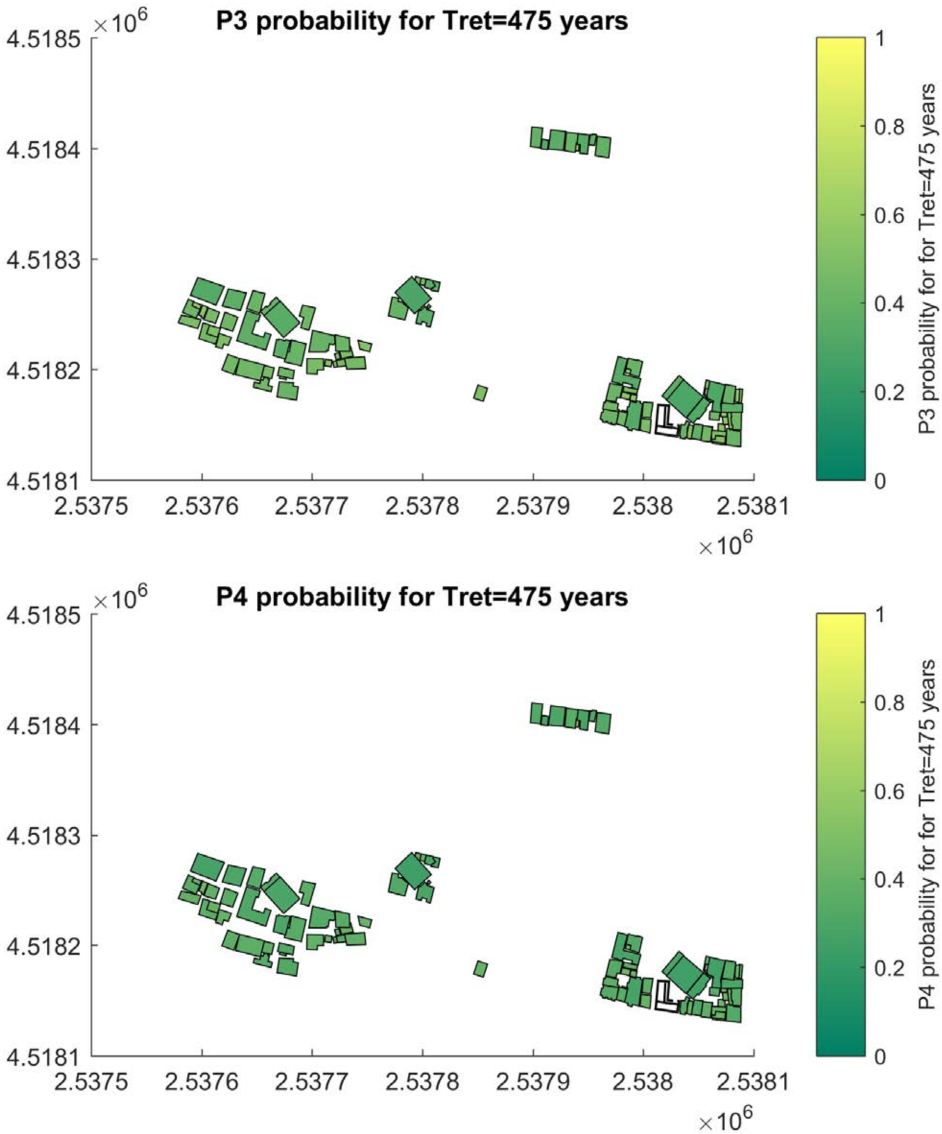

Figure 3. Result of the Spatial Model for building vulnerability classification

In the first phase of the project, the STABLE platform will contain within it selected sample cases. The cases are historical centres located in Cyprus, Greece and Italy.

Nafplio and the historic centre of Rethymno (case study recently included in the research) in Greece, Strovolos in Cyprus and Rieti in Italy (fig. 4). The aim is to start from individual buildings (monuments, historic buildings, etc.) and then hypothesise a future extension to the entire fabric that constitutes the historic centre.

To study how the structure of the building reacts in the presence of a seismic event of lesser or greater intensity, it is necessary to start from a careful study of mere basic maps, such as cadastral map, statistical data (age of the buildings, H maximum, number of floors), geological studies, aerial photos and historical maps. This will become part of a database of buildings that will allow you to select, investigate and observe all the information relating to the structure.

Details on the latest case study in Rethymno

The study area of the research is focused on the old town of Rethymno, on the island of Crete. The old town of Rethymno is one of the best-preserved towns of the Renaissance. Lying in the heart of modern Rethymno, it combines the oriental features of the Turkish period with Renaissance-style Venetian architecture. The plan of the old town of Rethymno has remained almost completely unchanged since the XVI century and it definitely needs a plan for the preservation of its architectural and historical evidence.

The main aim of our approach will be to assess the feasibility of Persistent Scattered Method (PSI) in estimating deformations and displacements in the broader area of the Old Town.

For this study, more than 100 S1B images were downloaded from the Copernicus Sentinel 1- Hub, concerning the monitoring period from January 2015 to June 2018. The Sentinel-1 mission comprises a constellation of two polar-orbiting satellites, operating day and night, and performing C-band synthetic aperture radar imaging, enabling them to acquire imagery regardless of the weather.

Afterwards, a combination of open source software such as the Sentinel Application Platform (SNAP), the Stanford Method for Persistent Scatterers (StaMPS), and the Statistical Cost Network Flow Algorithm for Phase Unwrapping (SNAPHU) package were installed in a Linux environment in order to carry out the PSI processing. The PSI requires at least 20 SAR images (Mora et al. 2013) to perform the analysis in C-band and measure surface deformation over a time period of months or years, removing atmospheric, topographical, and signal noise effects (fig. 5). In PSI, only the coherent pixels with stable phase or amplitude are used in the processing pipeline (Mancini et al. 2021). In this concept, the PSI exploits interferograms with a single master scene. In our case, 107 images in total were incorporated into the study (tab. 1):

Table 1. Analysed Tracks

Period (yyyy-mm-dd)	Master Science Acquisition date (yyyy-mm-dd)	Pass	Image Number

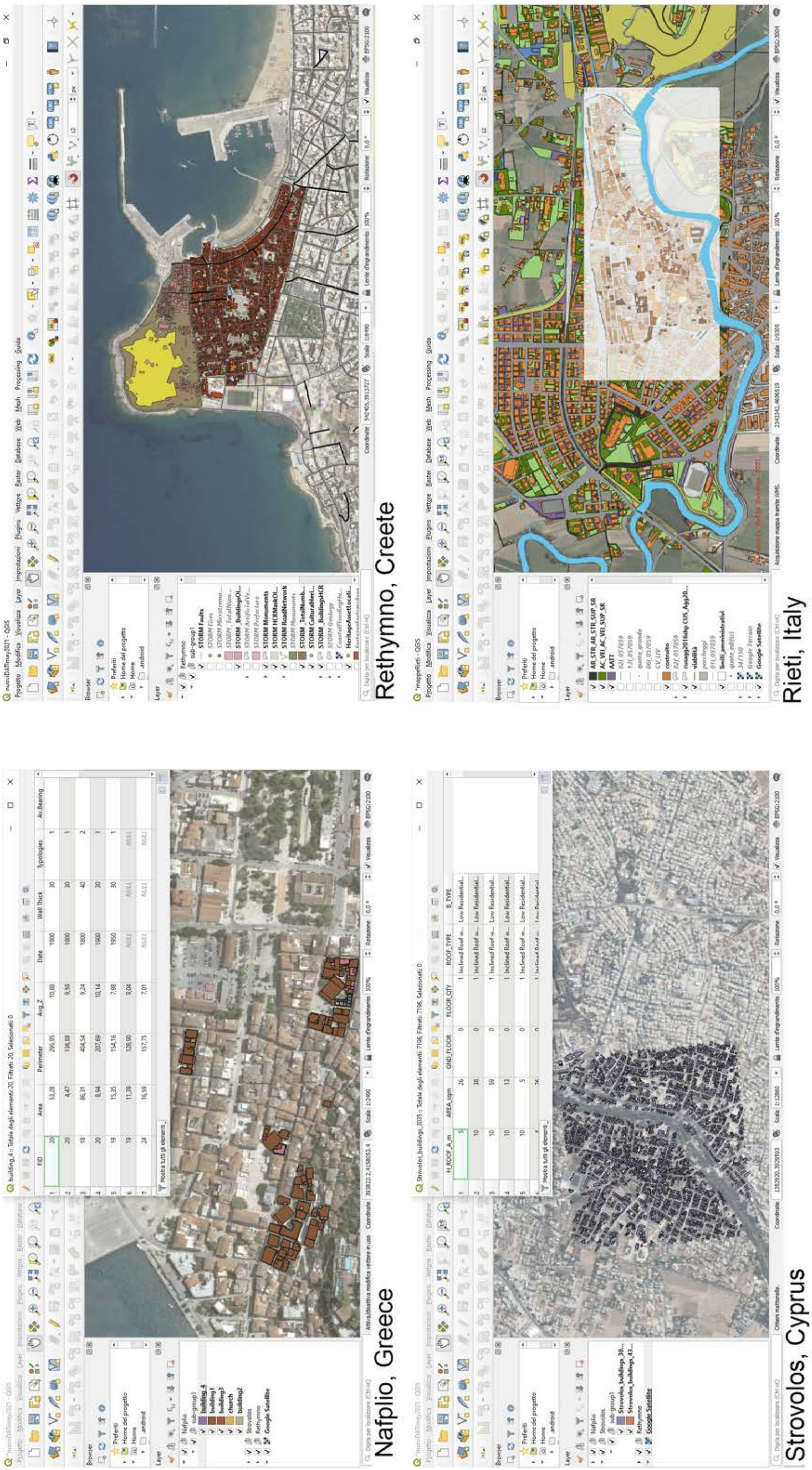

Figure 4. The test case location for STABLE project

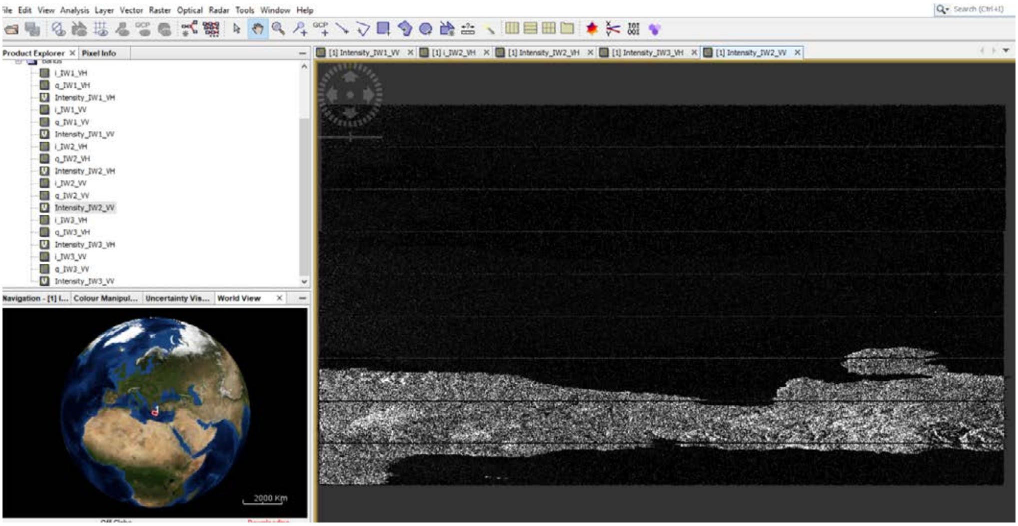

Figure 5. Processing of Sentinel-1 image in SNAP open source software environment

2015-01-05 / 2018-06-30	2017-23-02	Ascending	107

At the final step, the interferometric processing of S1 SLC images will generate a geolocated velocity map over the study area. This output will offer a ground deformation map which will provide us with useful knowledge about deformation phenomena occurring in the study area as a further contribution to geohazard analysis.

The methodology presented here will be applied also to the site of the Fortezza, a Venetian fort built in the XVI century on a rocky hill known as Paleochora. It is a site of particular interest since its walls (total length: 1.370 m) are well preserved, but at the same time, they are exposed to extreme climate events which endanger their integrity.

Actual achievements

The project is based on the application of the simplified structural stability model to the whole area of the city centre under analysis. The simulation of different earthquake scenarios, also considering the geological information available, and the response of the structures involved in the scenario will provide detailed damage maps of the area at building level, enabling preparedness for the seismic events. Public authorities, urban planners and cultural heritage responsible will have the capability to address preventive maintenance and consolidation of the structures that will be most damaged in future earthquakes.

This approach will have the twofold objective to drastically reduce the reconstruction and restoration costs as well as reduce the death toll associated with seismic activity.

The project has been running for almost one year and a half, mostly spent on the definition of the overall system, and will last until October 2023.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 823966.

References

Del Soldato M., Solari L., Poggi F., Raspini F., Tomás R., Fanti R. and Casagli N. 2019. "Landslide-Induced Damage Probability Estimation Coupling InSAR and Field Survey Data by Fragility Curves." in *Remote Sensing*, 11: 1486, doi:10.3390/rs11121486.

Fárová, K., Jelének, j., Kopacková-Strnadová, V., and Kycl, P. 2019. "Comparing DInSAR and PSI Techniques Employed to Sentinel-1 Data to Monitor Highway Stability: A Case Study of a Massive Dobkovicky Landslide, Czech Republic", in *Remote Sensing*, 11: 2670. doi:10.3390/rs11222670.

Hanssen, Ramon F. 2001. *Radar Interferometry: Data Interpretation and Error Analysis*, Kluwer Academic, ISBN 9780792369455.

Lagomarsino, S., and Giovinazzi, S. 2006. "Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings." in *Bulletin of Earthquake Engineering*, 4 (4): 415-443.

Kappos, A. J., Panagopoulos, G., and Penelis, G. G. 2008. "Development of a seismic damage and loss scenario for contemporary and historical buildings in Thessaloniki, Greece." in *Soil Dynamics and Earthquake Engineering*, 28 (10-11): 836-850.

Kottke A.R. and Rathje E.M. 2008. "Technical manual for Strata, PEER Report 2008/10", in *Pacific Earthquake Engineering Research Center College of Engineering*, University of California, Berkeley.

Mancini, F., Grassi, F. and Cenni, N. 2021. "A Workflow Based on SNAP–StaMPS Open-Source Tools and GNSS Data for PSI-Based Ground Deformation Using Dual-Orbit Sentinel-1 Data: Accuracy Assessment with Error Propagation Analysis", in *Remote Sensing,* 13: 753. https://doi.org/10.3390/rs13040753.

Maniatakis, C. A., and Spyrakos, C. C. 2012. "A new methodology to determine elastic displacement spectra in the near-fault region." in *Soil Dynamics and Earthquake Engineering*, 35: 41-58.

Massonnet, D. and Feigl, K. L. 1998. "Radar interferometry and its application to changes in the earth's surface", in *Rev. Geophys.* 36 (4): 441–500, Bibcode: 1998 RvGeo. 36.441M, doi: 10.1029/97RG03139.

Mora, O., Mallorqui, J.J. and Broquetas, A. 2003. "Linear and Nonlinear Terrain Deformation Maps from a Reduced Set of Interferometric SAR Images. IEEE Trans. Geosci." in *Remote Sensing*, 41: 2243–2253.

Nakamura Y. 1989. "A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface." in *Quarterly Report of Railway Technical Research Institute* (*RTRI*), 30 (1): 25-33.

Spyrakos, C. C., Maniatakis, C. A., Kiriakopoulos, P., Francioso, A., and Taflampas, I. M. 2015. "Performance of a post-Byzantine triple-domed basilica under near and far fault seismic loads: Analysis and intervention." in *Handbook of Research on Seismic Assessment and Rehabilitation of Historic Structures,* IGI Global: 831-867.

Spyrakos, C. C., Maniatakis, C. A., and Taflambas, J. 2008. "Evaluation of near-source seismic records based on damage potential parameters: Case study: Greece." in *Soil Dynamics and Earthquake Engineering*, 28 (9): 738-753.

Tselentis, G. A., Danciu, L., and Sokos, E. 2010. "Probabilistic seismic hazard assessment in Greece–Part 2: Acceleration response spectra and elastic input energy spectra." in *Natural Hazards and Earth System Sciences*, 10 (1): 41-49.

MOvEIT: a Proof of Concept of a Road Graph for Late Antique Egypt

Julian Bogdani[1*]

[1]Dipartimento di Storia Antropologia Religioni Arte Spettacolo and Laboratorio di Archeologia Digitale (https://lad.saras.uniroma1.it), Sapienza University of Rome
*Correspondence: julian.bogdani@uniroma1.it

Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract: In the context of the fruitful collaboration between PAThs, an ERC-funded project, and LAD, the Laboratory for Digital Archaeology at Sapienza University of Rome, an experimental attempt to link between them places being described by the Archaeological Atlas of Coptic Literature developed into a road graph for Late Antique Egypt.

Desktop GIS software has been used to build and update the graph by following different sources and methodologies, and the resulting data have been published online as open access (https://paths-erc.eu/moveit/). A single-page application was built and is available to provide a graphical user interface to access the data and use the graph to calculate directions from one place to another. The project is opened and shared with the broad community of scholars for further development.

Keywords: digital archaeology, communication graph, late antique Egypt, landscape archaeology, Web technologies.

FOSS software used and licence: QGIS (GNU General Public License)
Open dataset and licence: https://github.com/paths-erc/moveit (AGPL-3.0 license)

Introduction

This paper is about MOvEIT, a first and still rough attempt aimed at interconnecting the archaeological sites being cataloged and described in the Archaeological Atlas of Coptic Literature by PAThs, a project directed by Paola Buzi and based at the Sapienza University in Rome (Buzi 2017; Bogdani 2019; https://paths.uniroma1.it; https://atlas.paths-erc.eu). This proof of concept has been conceived and developed in the context of LAD: Laboratorio di Archeologia Digitale alla Sapienza (Laboratory for digital Archaeology at Sapienza University of Rome), directed by Julian Bogdani (https://lad.saras.uniroma1.it). The project was strongly inspired, in its initial phases, by the highly stimulating ORBIS: the Stanford Geospatial Network Model of the Roman World (Scheidel 2015; 2014; https://orbis.stanford.edu/). Progressively, as explained in the next paragraphs, it developed in a quite different direction.

MOvEIT is both (1) an open access data set about a road graph connecting Egyptian places dated to Late Antique and Medieval period and (2) a Single Page Web Application (SPA) offering a graphical user interface to easily build least-cost paths between archaeological sites. Although these two products are tightly connected, the road graph being used as the base for all geospatial analysis by the SPA (Single Page Application), they can be used (and are distributed) separately by serving different scopes. The road graph is a continuouslyimproving model to enable spatial comprehension of Egyptian landscape and geography. It is also a centralised multi-temporal database integrating connection data from various sources. The Web application, on the other hand, is a very intuitive way to explore the model by simply retrieving directions, as a sort of Google Maps for late antique Egypt. It is evidently mostly focused on public engagement with archaeological and geospatial data, even if through further customisation it can turn into an easy-to-use spatial analysis tool for people not very familiar with GIS software.

Finally, the name MOvEIT is a play on words between English and Coptic, ⲙⲟⲉⲓⲧ in Coptic meaning “road, path” and even “place”[1] and the addition of the minuscule “v” shapes the English expression “move it”[2]. The strong idea that both expressions push forward is the movement between archaeological places and the underpinning connection network as means to explore the landscape.

The road graph

The road graph of Late Antique and Medieval Egypt is created and maintained in a GIS environment, as a very simple two-layers project. The first layer contains PAThs places, automatically retrieved in real-time from PAThs central database, rendered as a point layer and providing the nodes of the graph. The second theme, a polyline layer, provides connections between these nodes, producing the arcs of the graph. As far as the archaeological sites are concerned, many scientific contributions have already been published explaining the methodology followed to locate and fully describe each place of interest for the Archaeological Atlas of the Coptic Literature (Buzi 2020; Buzi et al. 2018; Colonna 2020). The implementation work is still ongoing and from time-to-time new sites are being added to the Paths atlas, or positions of already available sites are refined. It must be underlined that the Archaeological Atlas of Coptic Literature has entered a very stable phase of implementation and the bulk of the current work consists in providing descriptions of the already listed sites and/or geographical coordinates (i.e., positioning) of toponyms still lacking a proper geographical contextualization. In any case, changes in site location should trigger edits on the arcs of the graph, and therefore it should be considered as continually evolving.

Different criteria have been followed to create the arcs connecting archaeological sites and defining the communication network (fig. 1). While the precise position of a site can be in most cases clearly defined thanks to a long tradition of archaeological and literary studies, the same cannot be said for the communication network. Two primary sources have been used to build the graph: the Barrington Atlas of the Greek and Roman World (Talbert and Bagnall 2000) and the hydrographic network of Egypt. The Barrington Atlas was used in its already digital, GIS-ready, format developed and openly distributed by the Ancient World Mapping Center[3] (fig. 1, red lines).

1 TLA lemma no. C2072 (ⲙⲟⲉⲓⲧ), in: Coptic Dictionary Online, ed. by the Koptische/Coptic Electronic Language and Literature International Alliance (KELLIA), https://coptic-dictionary.org/entry.cgi?tla=C2072 (accessed 2022-05-24).
2 Credits for the name of the projects go to Domizia D’Erasmo.
3 All data sets are distributed as open data with Open Database License 1.0 in the Shapefile format (http://awmc.unc.edu/awmc/map_data/, accessed 2022.05-24). For our purpose, the ba_roads data set was downloaded and processed.

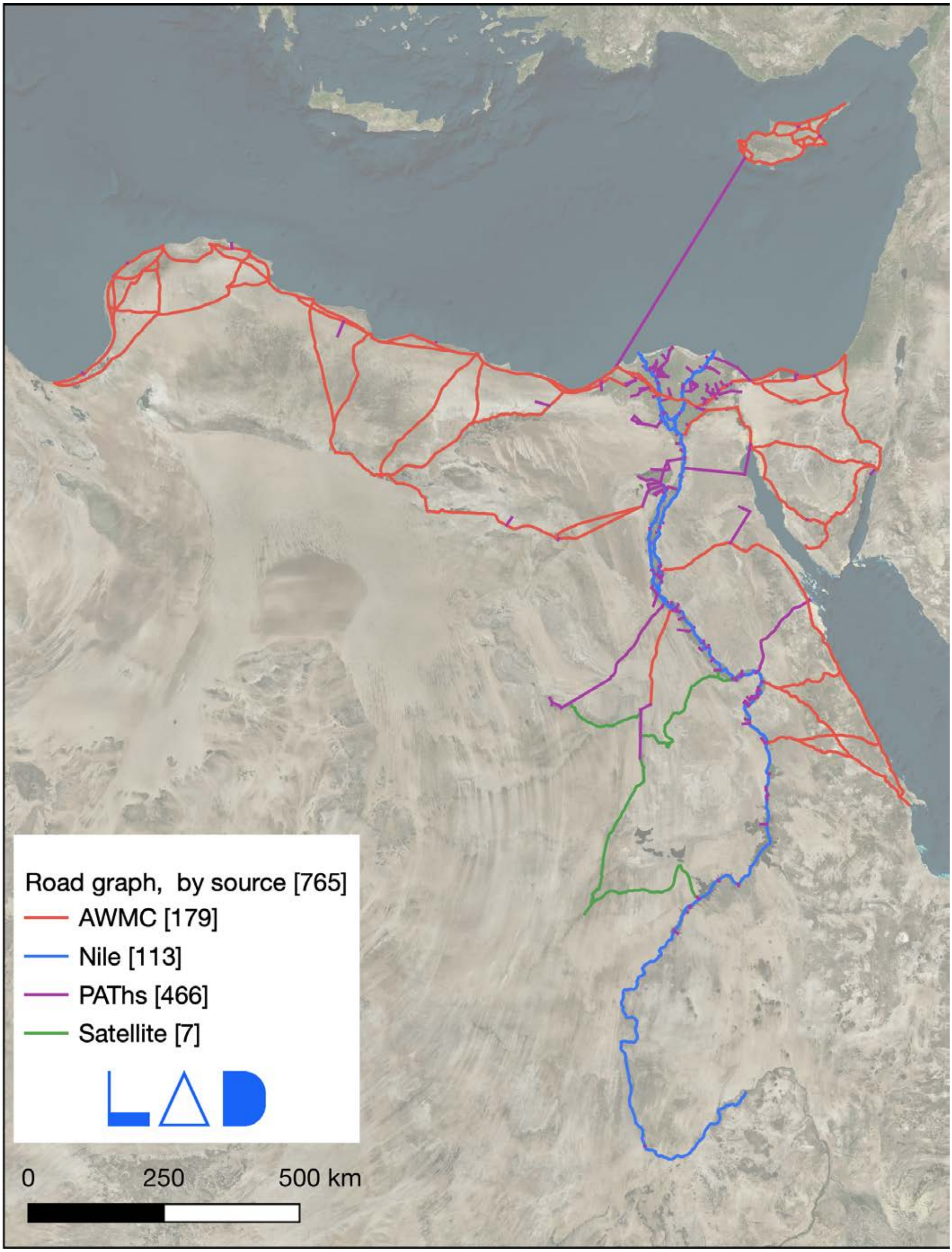

Figure 1. General view of the Egyptian road graph; in different colours are represented the different sources of information for each feature.

A second important source of information is the course of the Nile itself and its main navigable branches: the Nile, the Rossetta and Damietta branches of the Delta and the Bahr Yussef, connecting the Nile with the Fayyum (fig. 1, blue lines). It is well known that the Egyptian hydrological network has undergone important changes in the course of history and radical ones in the last one hundred years. For this reason, and at least for the main water courses, a frequent use of historical cartographical sources has facilitated the survey. Some of these notable documents, such as the Napoleonic map of Egypt (Jacotin and Jomard 1828; Jacotin 1824) have been made objects of deep analysis at LAD (Bogdani et al. 2022; D'Erasmo 2020). It must be clear enough, nevertheless, that when historical maps are considered, the state of the art of the road graph is only a little scratch on the surface of the potential data set. A huge quantity of information encoded in these maps is still waiting to be recovered (see paragraph Future implementation).

These two important sources of communication do provide an important backbone but are far from being exhaustive and are not able to connect all places mapped by PATHs. The next implementation step involved connecting the remaining sites to the graph by straight and direct arcs (fig. 1, violet lines). This is a highly abstract and not realistic implementation of the communication graph from the archaeological point of view. Nevertheless, it is plausible, considering the geomorphology of the country, where the Nile acts as an arterial road and direct connections to it are facilitated by the generally flat surface of the country (fig. 2). Interconnections through the sites are also created using straight lines, trying to privilege direct communications. Much more complex is the situation of the area of the Delta where

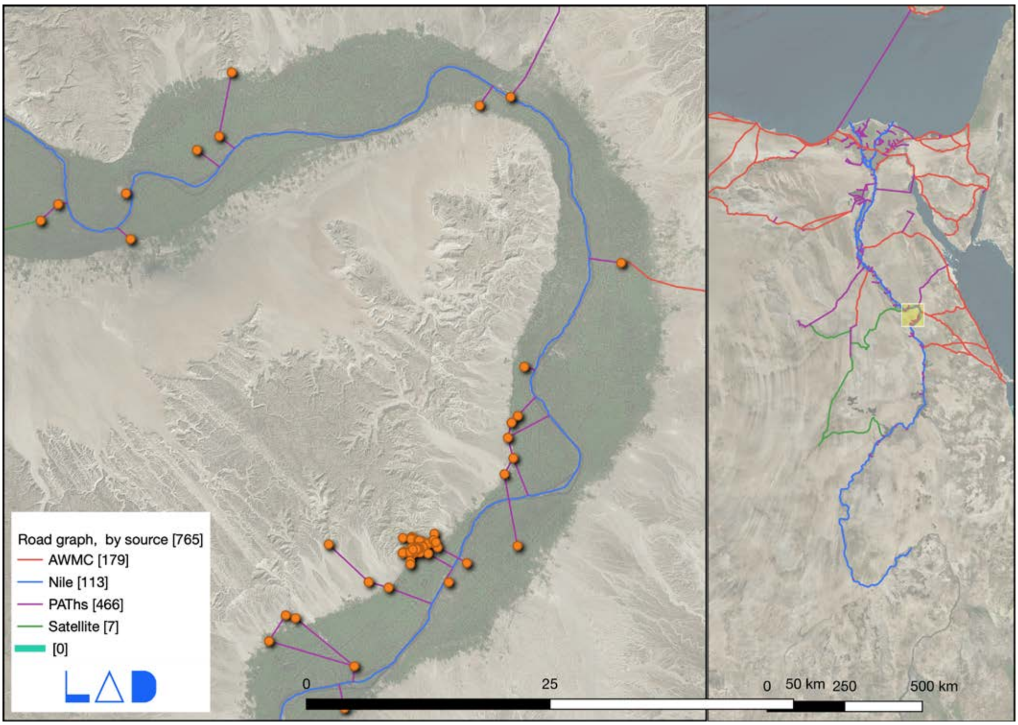

Figure 2. Detail of the meander of the Nile between modern Qena and Luxor.

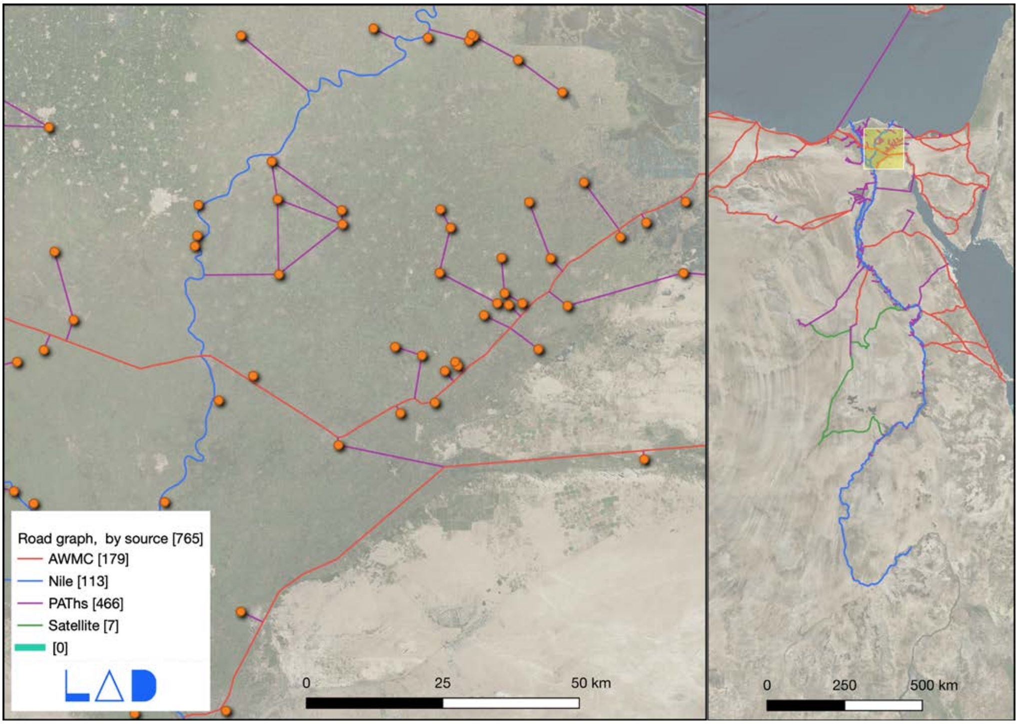

Figure 3. Detail of the Delta area around modern Zagazig.

connections are much more schematic and highly uncertain (fig. 3). The base assumption, as much as arbitrary, is that sites were connected to each other if no natural barrier forbade the direct passage. Straight lines have been manually adapted in a few cases to avoid natural barriers, such as hills, or ridges that are difficult to cross.

Finally, a few desert tracks, mostly located in the Western desert, have been digitised based on traditional paths still visible on the satellite imagery of commercial providers such as Google Maps or Bing Maps (fig. 1, green lines; fig. 4). These tracks, as is well known, are strongly conditioned by the difficult morphology of the terrain and their trace should be considered as highly conservative (Paprocki 2019: 190–275).

It is clear from the above paragraph that the different data sets integrated into this road graph are highly uneven, considering the resolution, adherence to archaeological methodology and general reliability of the information. Moreover, even if a general agreement is put on the soundness of the single tracks, these expect quite different transportation means and times, or simply put, costs. A boat trip of 100 km along the Nile can be a quite different experience compared to an equal linear distance in the rough Western desert of the wetlands of the Delta. These different parameters cannot be mixed accidentally.

Detailed metadata are provided for each feature of the graph with the aim of clearly stating the source, the resolution and reliability of the information, and to make it possible to assign to each vector a different weight in the calculation of the least cost path (fig. 5). Information

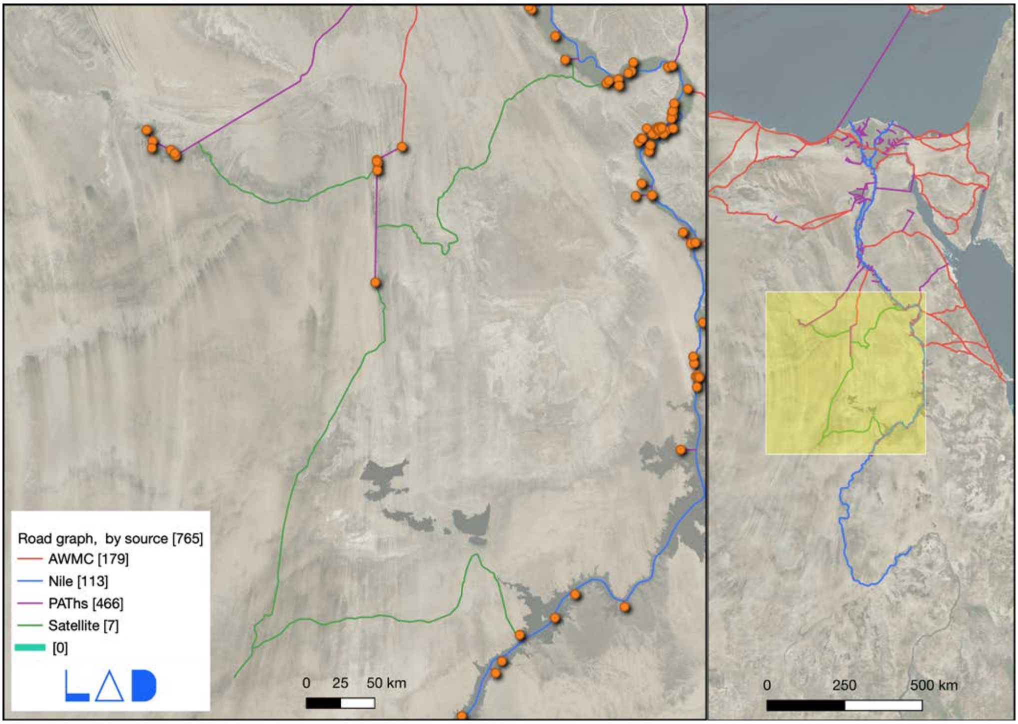

Figure 4. Detail of the tracks in the Western desert.

about the operator who created the feature and the date it was created permits tracking successive edits of the graph. In conclusion, this data set should not be considered as a final point of arrival, but as a pilot project testing methodologies and opening the road towards a collaborative and continuously developed effort, able to collect studies about connections and movements with a multi-temporal paradigm. As such, each single arc of the main graph can be made the object of further studies, discussions, and reviews.

The Web application

A Web application has been made available as a SPA built with HTML and JavaScript, serving as a graphical user interface to the graph-model[4]. This application, basically, calculates least-cost routes from one site to another, manually selected from the map by clicking markers or selected from drop-down menus containing the available PAThs places. The cost is calculated considering only two parameters: the linear overall distance between the two points and the type of path for each travelled segment. In fact, to each path-type a multiplying factor has been associated, as shown in the table.

Path type	Multiplying factor
Rivers	0.8
Channels	0.8
Sea routes	1
Along the coast (land) paths	1
Valley	1
Desert	2

[4] https://paths-erc.eu/moveit/ (accessed 2022-05-25).

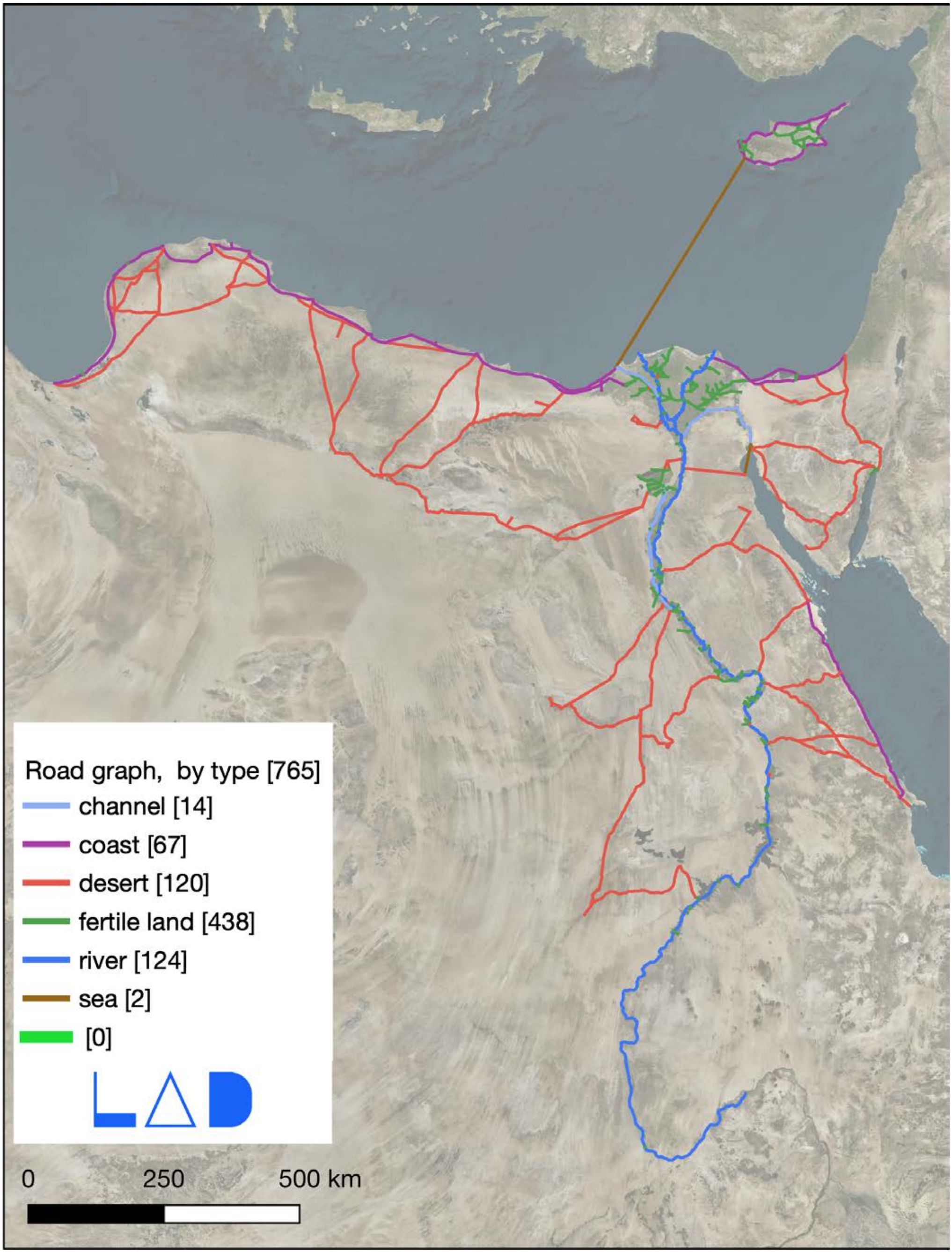

Figure 5. General view of the Egyptian road graph; the different typologies of paths are classified in different colours.

These coefficients listed above should be considered as a faint attempt to assign different weights to diverse track typologies and can (and should) be made the object of further consideration. An example of the different results obtained by taking into account the path type, an optional parameter, is shown in fig. 7.

The routing functionality — i.e., the calculation of the shortest path between two nodes — is handled by the GeoJSON Path Finder[5] an open source, serverless, offline routing JavaScript library for the browser. Since this library doesn't integrate any UI (user interface) tool, this has been built using Leaflet, the well-known and open source web-mapping library[6]. The application embeds the road graph and pulls the list of the PAThs Places from the PAThs web database based on BraDypUS, by using the available read-only REST API (Bogdani 2022)[7]. When a new site is added to the database and the road graph is not promptly updated, an error will be thrown and the path to/from the newly added site will not be calculated. Each time a route is calculated, the URL of the page in the address locator will be updated, easily obtaining shareable URLs containing routes.

GeoJSON Path Finder is a NodeJS[8] package, which means that NodeJS is required for further development or update, even when the road graph is updated. The Web application, nevertheless, integrates a few scripts to render plain and fast, both developing and building the application.

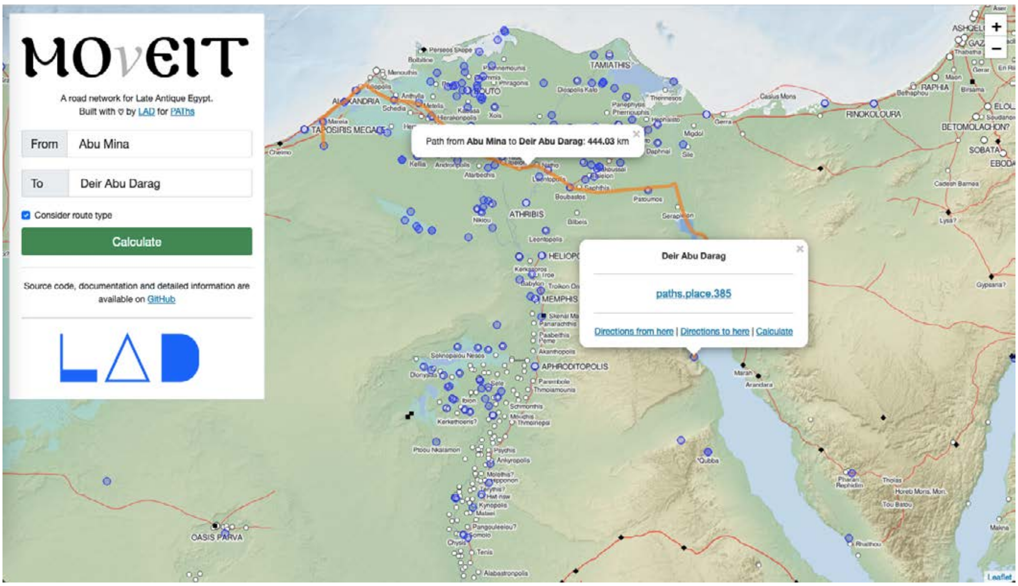

Figure 6. Screen capture of the Web application, available at https://paths-erc.eu/moveit (accessed 2022-05-25).

[5] https://www.liedman.net/geojson-path-finder/ (accessed 2022-05-25), ISC License.
[6] https://leafletjs.com/ (accessed 2022-05-25), BSD-2-Clause License.
[7] https://docs.bdus.cloud (accessed 2022-05-25)
[8] https://nodejs.org/ (accessed 2022-05-25)

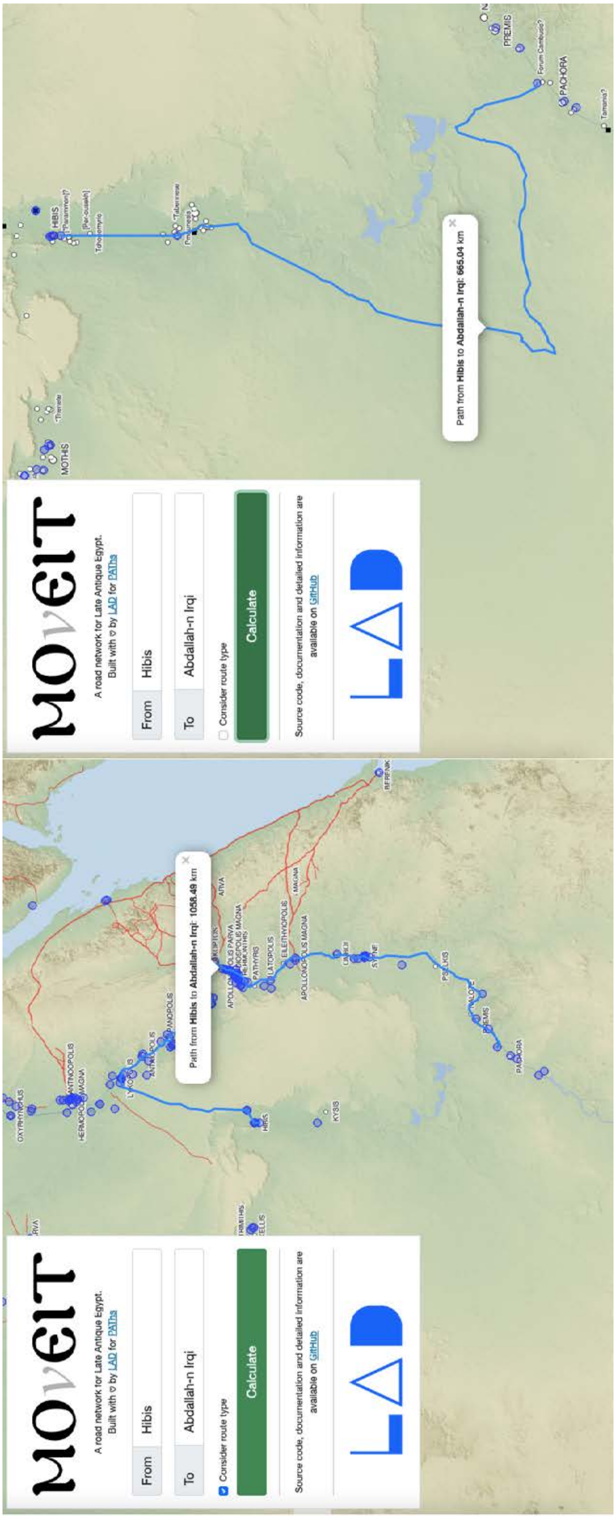

Figure 7. Different results of least-cost path analysis travelling from Hibis (paths.place.185) to Abdallah-n Iqri (paths.place.360) by applying the cost coefficient associated with path type on the left (desert road are much more "expensive" in relation to water/valley paths) and by considering only the linear distance on the right.

Further implementation

As already mentioned, the road graph and the graphical application should be considered as a proof of concept, aimed at defining methodologies and means to future development. None of these products should be considered as a finite research object ready to be deployed but as a starting point for the integration of information about movement from different sources.

Many further developments can be considered by deepening the analysis of the already considered sources, on one hand, and by considering other sources.

An eloquent example of the first case is the rich information of the Napoleonic map of Egypt, at present considered only for the course of the Nile and its main branches. Yet, this precious source can be further examined since it provides much more movement-related information, such as the late 17-century road network, medium and minor natural and artificial canals, dams, bridges, river-crossing, etc. (Bogdani et al. 2022). Also specific studies related to movement analysis through different means and in different periods (for Roman times desert roads, cf. Paprocki 2019; for mediaeval period waterways, cf. Cooper 2014) can be transformed into data and integrated into the road graph. The same is true for specific, large-scale archaeological, topographical, or geographical studies dealing with specific areas or contexts.

It is far beyond the scope of LAD to build and maintain a complete road graph of Late Antique and Medieval Egypt; our modest aim with this project is mostly to investigate how multi-scale, multi-temporal and multi-sourced data can be integrated, without flattering the general complexity.

Access to code and availability, collaboration, Acknowledgements

MOvEIT is created and maintained by LAD as an open source project. The GIS files and the SPA code is made available in Github[9] and deposited for long-term preservation on Zenodo[10]. The entire project has been released with the GNU Affero General Public License Version 3, 19 November 2007, also known as AGPL 3.0, which extends the "share alike" statement also for derived versions running on a remote server. To facilitate collaboration, detailed documentation is maintained online, providing useful resources and guides to enhance the graph, test it, and install and maintain a local copy of the web application[11].

This project is the fruit of a collaborative work in the LAD context, and special thanks go to Paolo Rosati and Domizia D'Erasmo.

[9] https://github.com/paths-erc/moveit (accessed 2022-05-25), AGPL-3.0 License.
[10] DOI: https://dx.doi.org/10.5281/zenodo.6583305.
[11] https://github.com/paths-erc/moveit/blob/master/README.md (accessed 2022-05-25)

References

Bogdani, J. 2019. "Linking Coptic Literary Manuscripts to the Archaeological Context by Means of Digital Humanities: The Case of 'PAThs' Project". in *Adamantius,* 24: 200–210.

Bogdani, J. 2022. "Archaeological Documentation as a Service. Archaeological Information Systems in the Cloud Era: The BraDypUS Case-Study". in *Archeologia e Calcolatori* 33 (2): 115–34.

Bogdani, J., D. D'Erasmo, C. De Leone, and A. Giordano. 2022. "Digitising the Napoleonic Map of Egypt. The Creation and Publication of An Open Data Set for the Reconstruction of the Ancient Egyptian Landscape". in *Vicino Oriente* XXVI: 77–97.

Buzi, P. 2017. "Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in Their Geographical Context. Production, Copying, Usage, Dissemination and Storage ("PAThs"). A New Project on Coptic Literary Texts". in *Early Christianity* 8: 507–16.

Buzi, P. 2020. "The Places of Coptic Literary Manuscripts: Real and Imaginary Landscapes. Theoretical Reflections in Guise of Introduction". In *Coptic Literature in Context (4th-13th Cent.). Cultural Landscape, Literary Production, and Manuscript Archaeology. Proceedings of the Third Conference of the ERC Project Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature : Literary Texts in Their Geographical Context ('PAThs')*, edited by Paola Buzi, 7–26. Percorsi Di Archeologia 5. Rome: Quasar.

Buzi, P., F. Berno, and J. Bogdani. 2018. "The 'PAThs' Project: An Effort to Represent the Physical Dimension of Coptic Literary Production (Third–Eleventh Centuries)". in *COMSt Bulletin* 4: 47–66. https://doi.org/10.25592/uhhfdm.253.

Colonna, A. 2020. "Contextualising Northern Egypt in Late Antiquity: An Archaeological Perspective from Western Delta". In *Coptic Literature in Context (4th-13th Cent.). Cultural Landscape, Literary Production, and Manuscript Archaeology. Proceedings of the Third Conference of the ERC Project Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature : Literary Texts in Their Geographical Context (PAThs')*, edited by Paola Buzi, 243–60. Rome: Quasar.

Cooper, J. P. 2014. *The Medieval Nile: route, navigation, and landscape in Islamic Egypt*. Cairo, New York: The American University in Cairo Press.

D'Erasmo, D. 2020. "Georeferencing Napoleonic Cartography to Reconstruct Ancient Egypt Landscapes: Methods in Comparison and the Case of the Island of Iw-Rd in the 16th Nomos of Upper Egypt". in *Groma. Documenting Archaeology 4: 1-18.* https://doi.org/10.12977/groma22.

Jacotin, P. 1824. "Mémoire Sur La Construction de La Carte Del l'Égypte". In *Description Del l'Égypte Ou Recueil Des Observations et Des Recherches Qui Onte Été Faites En Égypte Pendant l'expédition Del l'armée Française. Seconde Édition Tome Dix-Septième. État Moderne.*, edited by Charles Louis Fleury Panckoucke, 437–652. Paris: Imprimerie de C. L. F. Panckoucke.

Jacotin, P., and E. François Jomard. 1828. *Carte Topographique De L'Égypte Et De Plusieurs Parties Des Pays Limitrophes; Levée Pendant L'Expédition De L' Armée Française, Par Les Ingénieurs-Géographes, Les Officiers Du Génie Militaire Et Les Ingénieurs Des Ponts Et Chaussées; Assujettie Aux Observations Des Astronomes, Construite Par M. Jacotin, Colonel Au Corps Royal Des Ingénieur-Géographes Militaires, Chevalier de l'Ordre Royal et Militaire de Saint-Louis et de la Légion d'Honneur, Membre de l'Institut d'Égypte; Gravée Au Dépôt Général De La Guerre, à l'Échelle de 1 Millimètre pour 100 Mètres. Publiée Par Ordre Du Gouvernement*. Paris.

Paprocki, M. 2019. *Roads in the Deserts of Roman Egypt: Analysis, Atlas, Commentary*. Oxford, UK: Oxbow Books. https://doi.org/10.2307/j.ctvkjb2wr.

Scheidel, W. 2014. "The Shape of the Roman World". in *Journal of Roman Archaeology* 27: 7–32. https://doi.org/10.1017/S1047759414001147.

Scheidel, W. 2015. "Orbis: The Stanford Geospatial Network Model of the Roman World". in *Princeton/Stanford Working Papers in Classics*. https://orbis.stanford.edu/assets/Scheidel_64.pdf.

Talbert, R. J. A., and R. S. Bagnall. 2000. *Barrington Atlas of the Greek and Roman World*. Princeton, N.J. : Princeton University Press.

Palaeo-landscape feature identification: a FOSS cloud-based Python approach through Google Earth Engine (GEE)

Guillem Domingo-Ribas [1,*] and Filippo Brandolini [1,*]

[1]McCord Centre for Landscape - School of History, Classics and Archaeology, Newcastle University, UK, Newcastle upon Tyne, NE4 5HP, UK; *
*Correspondence: g.domingo-ribas2@newcastle.ac.uk; filippo.brandolini@newcastle.ac.uk
Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract:

Landscapes are the result of complex long-term interactions between natural and cultural forces, hence requiring techniques that enhance past landscape mapping and sustainable development perspectives. This paper presents a FOSS cloud-based protocol that employs Sentinel-2 imagery in GEE to identify palaeo-landscape features. Developed in Google Colab, the procedure is tested in the Po Plain riverscape (Italy), and it can be easily adapted and replicated for any area of the world.

Keywords: multispectral analysis, Sentinel-2, Google Earth Engine, riverscape, FOSS

FOSS software used and licence: , Google Earth Engine (*Google Earth Engine Licence Agreement*), Python Programming Language (*PSF Licence Agreement*), QGIS, (GNU Public Licence (GPL) Version), Google Colab (Modified BSD Licence).

Open dataset and licence: Copernicus Sentinel-2 satellite data (*Open Access compliant Creative Commons CC BY-SA 3.0 IGO licence*).

Introduction

Contemporary landscapes materialise from complex long-term interactions between natural and cultural forces. Human populations reorganise the land, adapt its use and its spatial structure to meet their needs. Their use of the land has altered the natural environment at least since the late Pleistocene and, over recent millennia, these processes have accelerated. Therefore, given the threat of human activities as active geomorphological agents that modify the physical landscape, it becomes paramount to develop techniques that enhance past landscape reconstruction while aiding future landscape planning and sustainable development perspectives.

In this line, interdisciplinary GIS and remote sensing techniques have provided significant enhancements in landscape research. These tools have improved the analysis of landscape dynamics over different spatial and temporal scales, enabling the identification of land use change, inferring population dynamics, economic development and, in turn, providing examples of human resilience through past sustainable development practices. Thus, GIS and remote sensing tools also permit the assessment of human impact on natural environments, hence requiring an approach in which the natural and cultural heritage domains are perceived together (Harrison 2015: 36). At the same time, other scholars highlight that the customizability and the potential for streamlining and automatising feature identification processes offered by remote sensing tools significantly reduce time-consuming manual mapping tasks, notwithstanding the need for analysis on the accuracy of the results (Traviglia and Torsello 2017).

However, considering that archaeological research or commercial archaeological companies do not normally receive a high investment, software companies' licences can hinder the usage of GIS. In this line, remote sensing software licences, powerful working stations with large storage space and certain satellite images can also require higher funding efforts (Parcak 2009). Instead, free and open source software (FOSS) constitute an alternative that does not only imply saving money, but also benefits from a vast active online community as well as installing, modifying and sharing the software when necessary (Ducke 2012). At the same time, in order to move towards a good scientific practice, it is necessary to publish the data, the source code and the software environment that generated the results, which is a growing tendency amongst open source projects and enables other users to reproduce the same process (Donoho et al. 2009).

Therefore, the advent of free-ware cloud computing services such as Google Earth Engine (GEE) represents a step forward towards sustainable and cost-effective landscape monitoring through the processing of more than 40 years of free and open satellite imagery and earth observation tools. Moreover, the possibility of employing GEE through its Python API in Google Colaboratory (commonly known as Colab) allows the use of a Python development

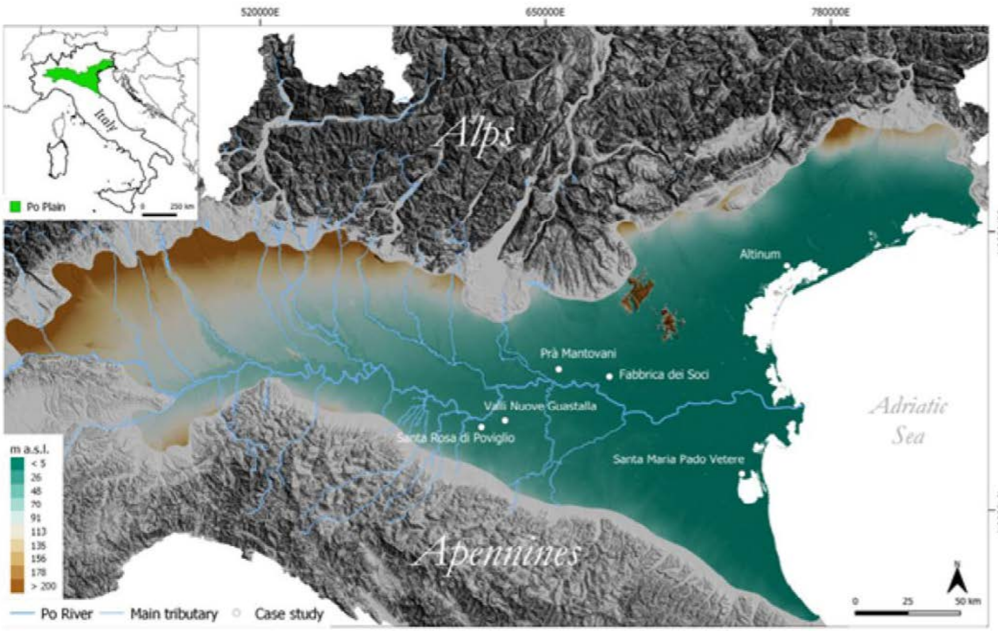

Figure 1. Test case area and locations of the selected case study locations (Brandolini et al. 2021)

environment that runs in the browser using Google Cloud, facilitating the creation of extensions, packages and scripts that broaden the applications of FOSS, particularly for processing more than 40 years of free and open satellite data, whose integration within archaeological projects is expected to continue increasing (Orengo 2015). In this line, while the capabilities of Python in modelling landscape dynamics are widely known, few research revolves around the potential of the GEE Python API (Vos et al. 2019), with a few exceptions that present the potential of the tool and its data catalogue in archaeological research (e.g. Liss, Howland, and Levy 2017; Fattore et al. 2021). Thus, we propose a complete FOSS-cloud approach to identify palaeo-landscape features through the use of GEE Python API in Google Colab. The present paper summarises our presentation at the ArcheoFOSS 2021 Conference, and further information about this methodological approach is available at (Brandolini et al. 2021).

Test case area

In order to test the potential of our FOSS approach, riverine landscapes have been selected for assessing the applicability and effectiveness of the methodology that we propose. Fluvial/alluvial environments have played a key role since prehistory, and archaeological research has shown that human agency has altered deeply the spatial configuration and rate of fluvial processes. Therefore, riverine landscapes are the result of complex relationships between human activities and environmental factors (Fryirs and Brierley 2012). Moreover, the large scale of palaeo-riverscape features has triggered their identification through remote sensing and satellite imagery, hence becoming a solid example for testing this type of approach.

In this line, the Po Plain (Northern Italy) (fig. 1), which is the largest floodplain in Italy, is rich in field and remotely sensed geomorphological data. Thanks to its complex settlement and land-management history, this test case area represents an ideal space in which to test the possibilities of our FOSS-cloud approach to detect riverscapes' palaeo-features. Thus, in this contribution we will introduce the first Python application of Sentinel-2 data for heritage research in a European riverscape, while illustrating the possibility of detecting and interpreting buried anthropogenic landscape features originating in different periods.

Materials and methods

Several applications of remote sensing techniques and GIS to record past landscape settings have been applied in the Po Plain over the last decades. Nevertheless, the current state-of-the-art in FOSS software and the availability of open source satellite data enable the development of protocols that require less computational power and are more user-friendly for non-experienced researchers.

Our approach tests the potentiality of Sentinel-2 (S2) satellite constellation, which was developed by the European Space Agency (ESA) and provides data in 13 separate bands with a spatial resolution up to 10 metres.

In this line, it is worth mentioning that buried features can sometimes be visible through crop marks and soil marks on the surface, given the difference in moisture that they retain in comparison to their surrounding area. However, even with high-resolution satellite sensors,

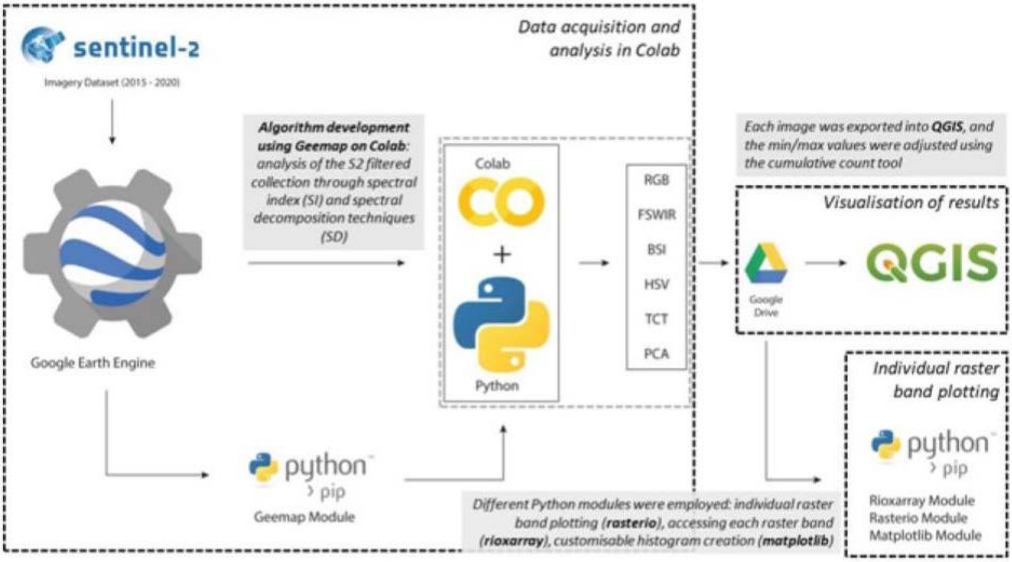

Figure 2. Methodological workflow with the FOSS tools employed. All the steps from data acquisition, processing, analysis and plotting are RGB: Red-Green-Blue colour composite; FSWIR: False colour composite; BSI: Bare Soil Index; HSV: Hue, Saturation and Value; TCT: Tasselled Cap Transformation; PCA: Principal Component Analysis. Based on Brandolini et al. 2021

crop marks are often dependent on various aspects, amongst which the phenological stage of the crops stands out. Therefore, in order to overcome the problematic nature of crop mark identification, a multi-temporal approach has been taken by calculating the mean values of bands in the most significant periods, depending on drought and ploughing episodes, for the identification of crop/soil marks between the years 2015 – 2020.

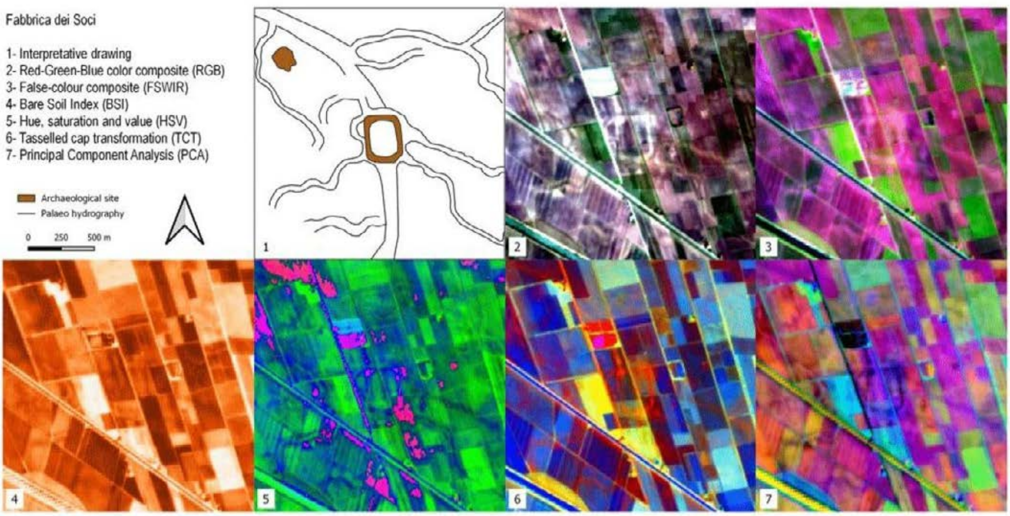

Figure 3. Outputs of the Fabbrica dei Soci site (Brandolini et al. 2021)

The S2 satellite data were accessed through the Python module geemap in Colab, a serverless Jupyter notebook computational environment for interactive development. The native GEE Python API has relatively limited functionality for visualising the results, hence the geemap Python module was created specifically to fill this gap. The Python code developed enables the analysis of the S2 filtered image collection through spectral index (compositions RGB (bands 4-3-2), false short wave infrared colour (FSWIR, bands 12-8-4) and Bare Soil Index (BSI)), as well as spectral decomposition techniques (hue, saturation and value (HSV), tasselled cap transformation (TCT) and principal component analysis (PCA)). The Python module rasterio was also used to create individual plots for each raster band. Additionally, the Python packages rioxarray and matplotlib were employed, respectively, to access each raster band and create customisable histograms of their values (fig. 2).

Results

As presented in fig. 1, different sites in the Po plain were employed to test the presented methodological approach and the output of the algorithm after the processing and analysis of the satellite data. Fig. 3 shows the results for Fabbrica dei Soci, a relevant and well-documented settlement for the Terramara culture in the selected area, which consisted of a square-shaped village and a hydraulic system that administered water coming from a palaeo-channel river to the surrounding fields.

RGB and FSWIR outputs permit an easy identification of palaeo-channels, moats and canals, as well as the square-shaped habitat area in the middle, while HSV highlights the buried earthworks. At the same time, the PCA and, to a certain extent, the TCT outputs also display the shape of the buried features.

Discussion

Sensing tool and its applications in archaeological research. First, geemap's inspector and plotting tools enable the user to assess the spectral signature value of different points, such as the buried features and their surrounding area, showing how detecting different values can indicate the location and shape of the earthworks.

Secondly, the second half of the protocol is dedicated to plotting each output band separately. By accessing the data of each band and plotting them against their histogram, it is possible to assess the performance of single bands. As depicted in fig. 4, some bands seem to offer a more relevant contrast between the features and their adjacent areas, performing better than the combined outputs. The most suitable bands for the identification of features can vary according to the characteristics of each study area.

However, it is also worth mentioning the limitations of this approach. Beyond the resolution that GEE outputs can reach, which may not allow to identify buried palaeo-features of smaller size, our approach requires the user to choose the right timespan for the study case. Crop rotation and meteorological conditions must be considered when selecting the best period of visibility, as soil moisture is generally correlated with crop/soil marks.

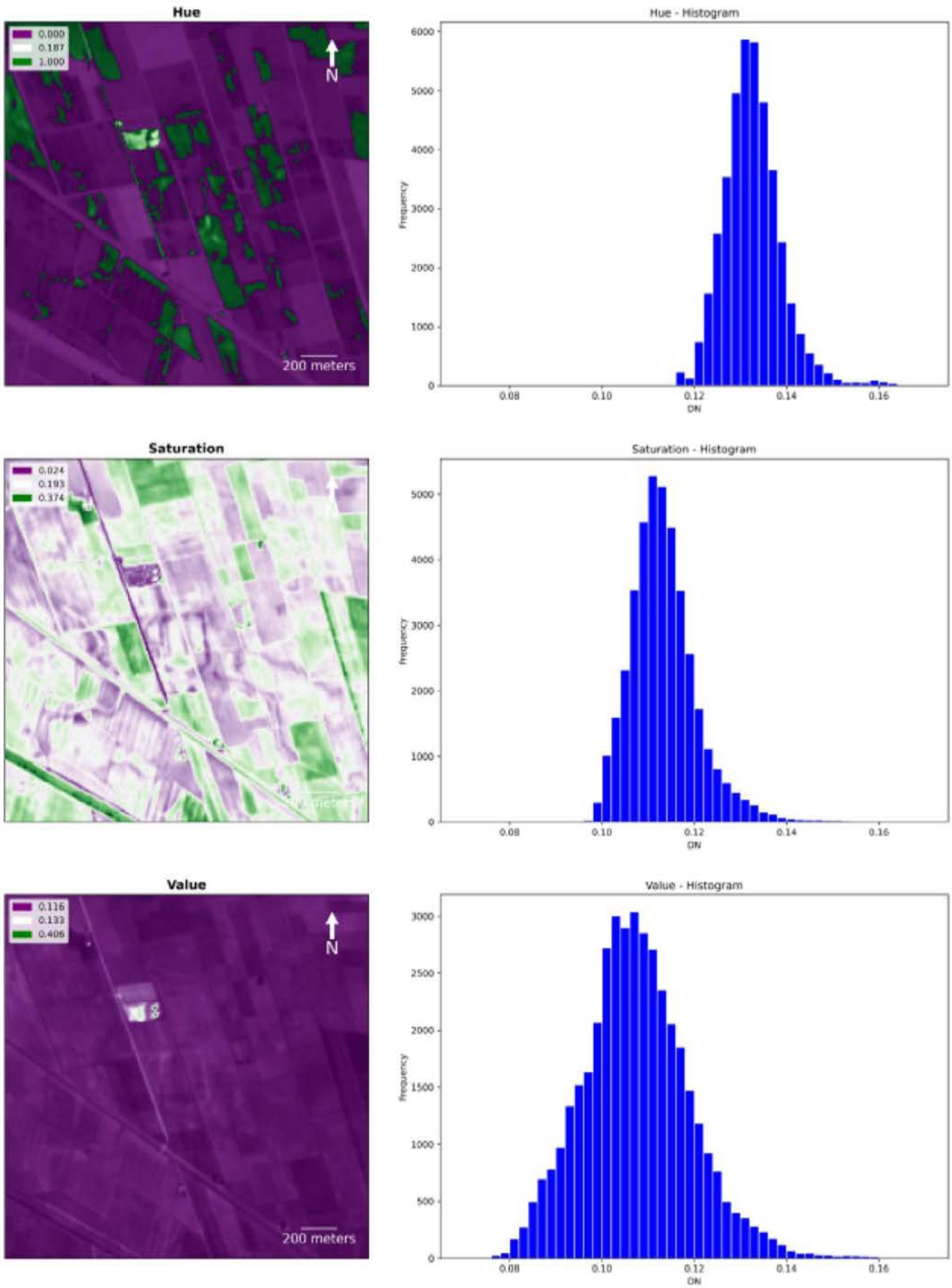

Figure 4. Hue, Saturation and Value plots and histograms of the Fabbrica dei Socci site (Brandolini et al. 2021)

Final remarks

While the choice of timespan is the only part of the protocol that needs to be customised by the final user according to the environmental conditions of each study area, the FOSS-cloud protocol presented here offers significant advantages regarding the mitigation of specialist software, data licensing and computational power. Overall, the outputs generated for the test locations of the Po Plain show some of the potentialities and limits of the GEE Python API in Colab as an alternative remote sensing tool to identify buried natural and anthropogenic palaeo-riverscape features.

Supplementary Materials: Link to the Python script (DOI: 10.5281/zenodo.5235030): https://doi.org/10.5281/zenodo.5235030#.YUJu13AIslk.link

Funding: This research was funded by the NUAcT PhD fellowship, Newcastle University (UK) (G. Domingo-Ribas), and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No [890561], Historic Landscape and Soil Sustainability (HiLSS) (F. Brandolini).

References

Brandolini, F., G. Domingo-Ribas, A. Zerboni, and S. Turner. 2021. "Engine-enabled Python approach for the identification of anthropogenic palaeo-landscape features [version 2; peer review: 2 approved, 1 approved with reservations]." in *Open Research Europe,* 1: 22. DOI: https://doi.org/10.12688/openreseurope.13135.2

Donoho, D. L., A. Maleki, I. Ur Rahman, M. Shahram, and V.Stodden. 2009. "Reproducible research in computational harmonic analysis." in *Computing in Science & Engineering,* 11 (1): 8–18. DOI: https://doi.org/10.1109/MCSE.2009.15

Ducke, B. 2012. "Natives of a connected world: free and open source software in archaeology." in *World Archaeology*, 44 (4): 571–579. DOI: https://doi.org/10.1080/00438243.2012.743259

Fattore, C., N. Abate, F. Faridani, N. Masini, and R. Lasaponara. 2021. "Google Earth Engine as Multi-Sensor Open Source Tool for Supporting the Preservation of Archaeological Areas: The Case Study of Flood and Fire Mapping in Metaponto, Italy." in *Sensors,* 21 (5): 1791. https://doi.org/10.3390/s21051791

Fryirs, K. A., and G. J. Brierley. 2012. "Geomorphic Analysis of River Systems: An Approach to Reading the Landscape". In *Geomorphic Analysis of River Systems: An Approach to Reading the Landscape*, edited by Kristie A. Fryirs and Gary J. Brierley, 1-8. Chichester: Wiley-Blackwell. DOI: 10.1002/9781118305454

Harrison, R. 2015. "Beyond "Natural" and "Cultural" Heritage: Toward an Ontological Politics of Heritage in the Age of Anthropocene." in *Heritage & Society,* 8 (1): 24-42. DOI:https://doi.org/10.1179/2159032X15Z.00000000036

Liss, B., M. D. Howland, and T. E. Levy. 2017. "Testing Google Earth Engine for the automatic identification and vectorization of archaeological features: A case study from Faynan, Jordan." in *Journal of Archaeological Science: Reports,* 15: 299-304. http://dx.doi.org/10.1016/j.jasrep.2017.08.013

Orengo, H. 2015. “Open source GIS and geospatial software: towards their integration into everyday archaeological practice”. In *Open source archaeology – Ethics and Practice,* edited by Andrew T. Wilson and Ben Edwards, 64-82. Warsaw: De Gruyter. DOI: https://doi.org/10.1515/9783110440171-006

Parcak, S. H. 2009. *Satellite Remote Sensing for Archaeology.* London: Routledge. https://doi.org/10.4324/9780203881460

Traviglia, A., and A. Torsello. 2017. “Landscape Pattern Detection in Archaeological Remote Sensing.” in *Geosciences,* 7: 128.

Vos, Kilian, K. D. Splinter, M. D. Harley, J. A. Simmons, and I. L. Turner. 2019. “CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery.” in *Environmental Modelling and Software*, 122: 104528. DOI: https://doi.org/10.1016/j.envsoft.2019.104528

Integrating pyArchInit and BraDypUS for field and academic archaeological research

Giuseppe Guarino[1]* and Paolo Rosati[2]

[1] Dipartimento di Storia Culture Civiltà, Università degli Studi di Bologna, Bologna, Italia;
[2] DigiLab, La Sapienza University of Rome, Rome, Italy
*Correspondence: giuseppe.guarino8@unibo.it, paolo.rosati@uniroma1.it
Presented at ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract:
This article presents an attempt to integrate into a single workflow two important tools for the management of archaeological data: pyArchInit and BraDypUS. Although each system has its proper applications and capabilities, their combined use could have brought several advantages. However, it is worth noting that the work currently has some structural limitations that can be resolved only by the modification of the source code. This work summarises and reports on the current methodologies (as of 2021) for operating on both desktop and web platforms simultaneously using specific tools.

Keywords: FOSS archaeological documentation, Archaeological protocol documentation, QField, PyArchInit, BraDypUS

List of FLOSS software used/ data repository and licence: QFIELD (*GNU Public Licence (GPL) Version 2*), BraDypUS (*MIT licence*), PyArchInit (*GNU Public Licence (GPL) Version 2*);
Open dataset and licence: Repository available at https://github.com/Heryx/cfg_from_pyarchinit_to_bradypus, (*MIT licence*).

The contribution of "Open-Source Archaeology" for exchanging data and ideas between researchers and archaeologists.

Within the domain of archaeological research, and particularly in informatics archaeology, the relevant costs associated with software licences might present a concrete obstacle, particularly in situations where research funds are scarce. Furthermore, beyond their proprietary nature, numerous prevalent software applications in archaeology lack explicit design and development for archaeological purposes, thereby falling short of meeting the specific requirements of the discipline. Additionally, certain archaeology-specific software programs are no longer receiving support. As Benjamin Ducke noted:

> One possible solution to this problem lies in embedding software development and research within larger, open-source projects that are not confined to the scope and lifetime of a single research project. This provides a two-fold advantage. On the one hand, paying programmers to implement functionality in open-source software delivers a solid return on investment. On the other hand, it offers countless opportunities for the sharing of investments and resources, reducing the burden on the individual stakeholders. (Ducke 2012, 574)

Open-source software is distributed as readable program source code written in a (usually high-level) programming language (Wilson and Edwards 2015, 1). The availability of source code allows the end user not only to run the software but also to manipulate, change, redevelop and understand how it works.

We use different software tools for our academic research which are often not interoperable with each other. Nowadays, an important objective is finding a meeting point to allow groups and people with different backgrounds to quickly consult the archaeological data, thus increasing the exchange of resources and ideas. Our brief contribution provides an attempt to integrate some of these tools, pyArchinit and BraDypUS to allow for accessible software communication.

The software used

Initiated in 2005, the project "pyArchInit - python for archaeology" was designed to facilitate the management of data derived from archaeological contexts within the well-known and widely used QGIS platform (Cocca and Mandolesi 2013). The overarching goal of the developers was to establish a novel application for archaeologists, seamlessly adaptable to both field and laboratory settings. The envisioned application aimed for user-friendliness and seamless integration with management systems utilised by public administrations. Luca Mandolesi spearheaded the development of Pyarchinit, and Enzo Cocca further implemented it (Mandolesi and Cocca in 2022). The development is based on the commitment to maximum system portability, ensuring ease and immediacy of use, leveraging open-source software for development, creating modules for comprehensive archaeological data management, integrating existing and robust data management systems, and maintaining flexibility for the accommodation of alphanumeric, cartographic, and multimedia data.

BraDypUS stands as an application dedicated to the creation and management of archaeological databases online (Bogdani 2021). Originating from Julian Bogdani's initiatives in 2008 to develop bespoke web-based software solutions for the management of archaeological data, has since expanded its scope, managing, and providing services for significant Italian and international cultural heritage projects. Within the realm of digital humanities, BraDypUS emerges as a valuable tool, particularly in the field of archaeology. The software expedites the processes, methods, surveys, and documentation involved in excavations, often constrained by site-specific requirements. The WebGIS solutions integrated into BraDypUS offer practical assistance in inventorying, understanding, and networking cultural heritage within a given area (Bogdani 2019).

(P.R.)

From BraDypUS to PyArchInit and vice versa

Both pyArchInit (https://github.com/pyarchinit/pyarchinit) and BraDypUS (https://github.com/lab-archeologia-digitale/BraDypUS) are specifically tailored for archaeological applications. PyArchInit is predominantly developed in Python and relies on a predetermined database schema primarily designed for spatial data management. BraDypUS employs various web languages such as PHP, Javascript, and CSS, and boasts extensive versatility, offering complete customization for relational databases. The installation of BraDypUS is typically conducted on a web server and is accessible from any device running a modern web browser, while pyArchInit runs on a local machine and requires QGIS. In addressing the challenge of different structures, a resolution was achieved by establishing a connection between the two systems through the utilisation of a Spatial Relational Database Management System (RDBMS). Both applications are capable of interfacing with PostgreSQL and PostGIS, providing a collaborative solution for multiple individuals to concurrently engage in a shared project. Given the rigid database schema inherent in PyArchInit, the process necessitated constructing the database using PyArchInit's dedicated tools.

PyArchInit requires many tables to be created, each specifically designed to adhere to the stringent standards set by the Central Italian Institute for Cataloguing and Documentation (ICCD).

Given that both software packages were originally designed for distinct purposes and implemented using different programming languages, encountering incompatible components is inevitable. One prominent challenge lies in certain database conventions that hinder a straightforward and immediate connection between the systems. For instance:

- BraDypUS needs two mandatory fields (named "**id**" defined as the primary key, and "**creator**") to be added to every PyArchInit table to properly work;
- BraDypUS requires each table name to have a unique prefix separated by a double underscore;
- Yet, pyArchInit is very restrictive on the table naming system, and names such as '**site_table**' and '**US_table**' cannot be changed

To enable a connection between the two platforms, sharing a common database, we recommend the following procedural steps. As a premise, a PostgreSQL/PostGIS, PHP and Apache web server must be available; we will not cover here in detail the setup of a proper environment, since the BraDypUS official guide offers some support for this[1].

In QGIS, the pyArchInit plugin must be used to create and initialize the PostgreSQL database to strictly adhere to the required schema. Subsequently, the QGIS built-in function named DB Manager or an external tool (such as PgAdmin4) can be used to access the database generated by pyArchInit and add to each table the two columns required by BraDypUS, namely 'id' and 'creator'. Both columns should be configured as type INTEGER[2]. This is a critical step to align the PyArchInit database to BraDypUS. After these first steps, a new BraDypUS application can be created by pointing to the already-created database by selecting "pgsql" as the database type and entering the same connection credential used during the creation of the PostgreSQL database in pyArchInit[3]. Consistency between BraDypUS and pyArchInit data is crucial for establishing a reliable connection and ensuring seamless interoperability.

[1] https://docs.bdus.cloud/environment/setup-windows (accessed 20/12/2023)
[2] More information is available at: https://docs.bdus.cloud/conventions (accessed 20/12/2023)
[3] Further details and thorough instructions are available at the official BraDypUS documentation site: https://docs.bdus.cloud/create_app (accessed 20/12/2023).

To ensure easy access and manipulation of data within the BraDypUS environment, it is essential to meticulously create the configuration for each of the pyArchInit tables. BraDypUS integrates a graphical user interface (GUI) for this, which is simple and safe to use.[4] that will be stored in the "Project" folder of BraDypUS. To facilitate this process, we have made available a first draft of the configuration files required to run BraDypUS on a pyArchInit database schema.[5]

As a result, some aspects remain unsolved, the main one being the issue related to the primary key of each table. BraDypUS mandatory requires a field named "id" to act as a primary key, a rather diffused practice in relational databases. On the other hand, PyArchInit uses a proper, uncommon, way of naming the primary key, by prefixing "id_" to the name of the table (e.g., us.id_us, sito.id_sito, tafonomia.id_tafonomia, and so on).

pyArchInit and BraDypUS use also different methods for representing stratigraphic relationships. PyArchInit files all related contexts as an array in a single field, while BraDypUS employs related tables by taking full advantage of the relational model. In the first case, heavy application logic is required to grant data consistency and avoid recursive relationship, something that is solved in the second case by few database constraints that prevent, for example, do define twice a relationship between two contexts.

These are only few preliminary notes that we wanted to share with a broader public of users and developers, in the attempt to facilitate further integration between software solutions that are spreading rapidly in the Italian context, at least.

Finally, the database behind pyArchInit and BraDypUS seamlessly integrates with QField, a GIS software for mobile devices, facilitating the direct storage of archaeological data gathered during fieldwork (Bernasocchi 2022a, 2022b). QField empowers users to visualise and manage GIS projects on smartphones or tablets, leveraging QGIS project themes, labels, and styles, and is currently becoming a rather common practice in the archaeological domain (Montagnetti and Guarino 2021, Montagnetti and Rosati 2019).

(G.G.)

[4] https://docs.bdus.cloud/setup/
[5] https://github.com/Heryx/cfg_from_pyarchinit_to_ bradypus (accessed 20/12/2023).

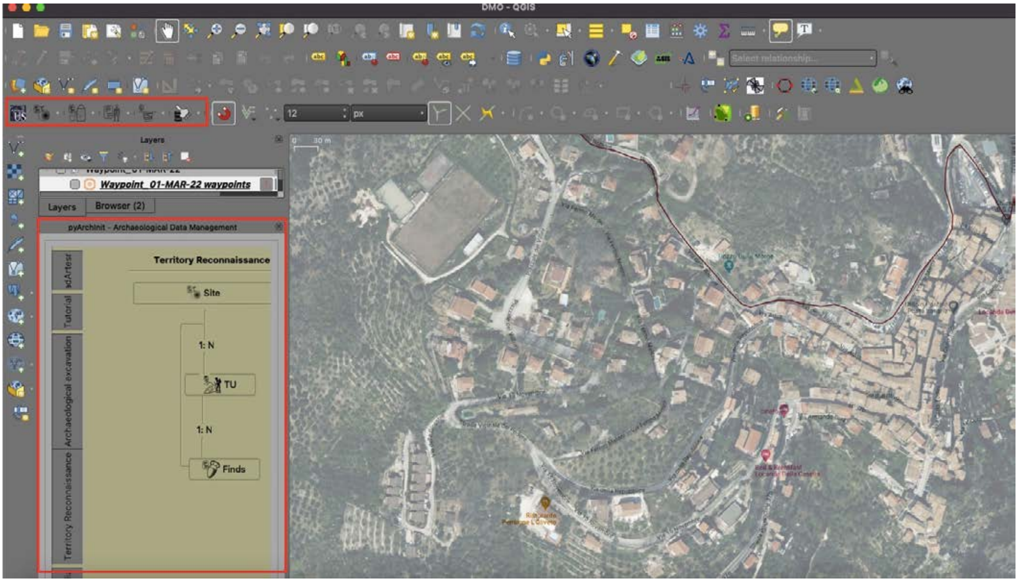

Figure 1. Pyarchinit on Mac, in the red squares the toolset of pyarchinit plugin and the presentation/shortcut bar under the layers panel.

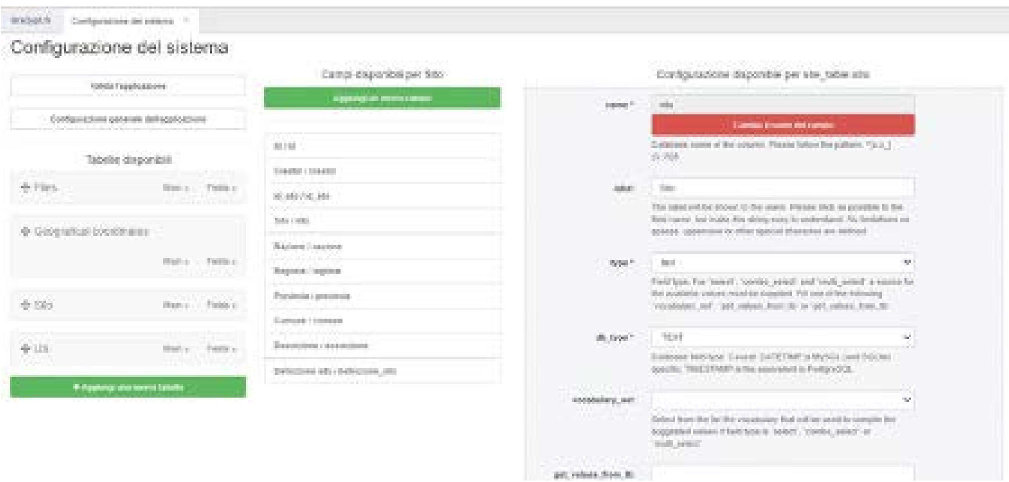

Figure 2. The BraDypUS configuration systems page

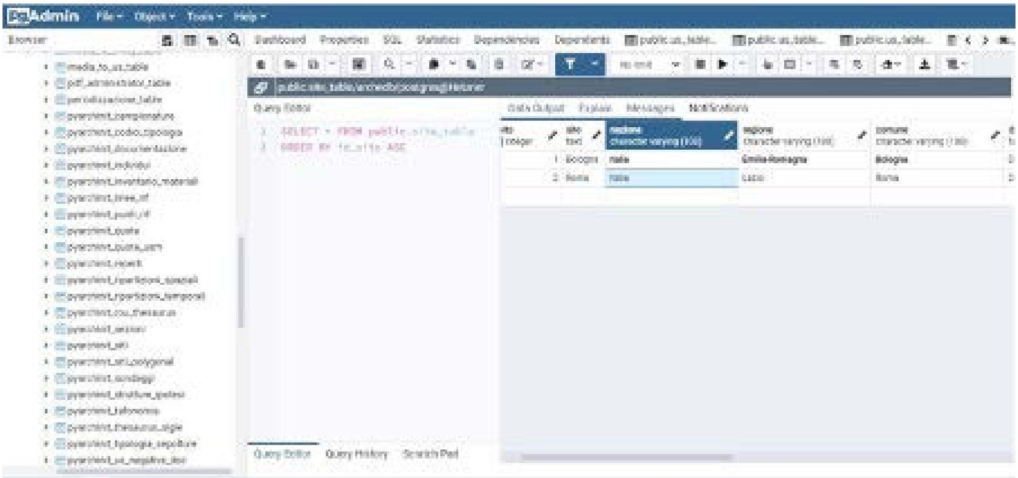

Figure 3. The pgmyadmin main page with the config folder

References

Bernasocchi M. 2022a. *QField* (version 1.10.00). https://github.com/opengisch/QField.

Bernasocchi M. 2022b. *Qfield Cloud* (version 0.13.2). https://github.com/opengisch/qfieldcloud.

Bogdani, J. 2019. "Archeologia e tecnologie di rete. Metodi, strumenti e risorse digitali". in *I Cardini di Groma*, 4. Roma 1st ed.: BraDypUS Communicating Cultural Heritage

Bogdani, J. 2021. *BraDypUS* (version 4.1.2). https://github.com/lab-archeologia-digitale/BraDypUS.

Ducke B. 2012. "Natives of a Connected World: Free and Open Source Software in Archaeology". In *World Archaeology*, 4 (44): 571–79.

Cocca E. and L. Mandolesi L. 2013. "Pyarchinit: gli sviluppi dopo Archeofoss 2009". In *Archeologia e Calcolatori* Suppl. 4:128–38.

Montagnetti, R., and G. Guarino. 2021. "From Qgis to Qfield and Vice Versa: How the New Android Application Is Facilitating the Work of the Archaeologist in the Field" *Environmental Sciences Proceedings* 10, no. 1: 6. https://doi.org/10.3390/environsciproc2021010006

Mandolesi L. and E. Cocca. 2022. *PyArchInit* (version 3.9.12). https://github.com/pyarchinit/pyarchinit.

Montagnetti R. and P. Rosati. 2019. "Georiferire la stratigrafia archeologica". *Archeologia e Calcolatori* 30: 463–66

Edwards, Benjamin and A.T. Wilson. "1 Open Archaeology: Definitions, Challenges and Context" In *Open Source Archaeology* edited by Andrew T. Wilson and Ben Edwards, 1-5. Warsaw, Poland: De Gruyter Open Poland, 2015. https://doi.org/10.1515/9783110440171-002

From micro-regional to intra-site analysis: the GIS of the Italian Archaeological Expedition in the Erbil Plain (Kurdistan Region of Iraq)

Valentina Oselini[1]*, Michael Campeggi[1], Luca Forti[2], Elisa Ginoli[1], Andrea Pezzotta[2], Agnese Vacca[1] and Luca Peyronel[1]

[1] Department of Literary Studies, Philology and Linguistics, University of Milan, Milan, Italy;
[2]Department of Earth Sciences "Ardito Desio", University of Milan, Milan, Italy;
*Correspondence: valentina.oselini@unimi.it

Presented at ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021

Abstract:

The Italian Archaeological Expedition in the Erbil Plain (MAIPE https://orienteantico.unimi.it/it/maipe/) of the University of Milan is experimenting with a GIS project based on the QGIS system. The project aims to combine data from survey, excavation, and geomorphological investigations carried out at the two sites of Helawa and Aliawa and their surrounding landscape. The MAIPE GIS is intended as a valuable tool for predictive, distributive and cross-data analyses considering the complex diachronic development of the area and the results of paleoenvironmental studies at a micro-regional and intra-site scale.

Keywords: QGIS, cross-data analyses, fieldwork, archaeology, geomorphology

FOSS software used and licence: QGIS (*GNU General Public License*), QField (*GNU General Public License v2.0*), GRASS GIS (*GNU General Public License*)

Introduction (LP, AV)

Since 2013, the Italian Archaeological Expedition in the Erbil Plain (MAIPE, University of Milan), directed by L. Peyronel with A. Vacca as deputy director, has been carrying out archaeological investigations at Helawa and Aliawa, two sites located ca. 27 km SW of the city of Erbil, in the Kurdistan Region of Iraq. The investigated area (ca. 20 km^2) lies in a strategic position of the plain, just north of the Awena Dagh, a range of low hills parallel to the Qara Chuq mountains, the latter representing the border with the Makhmour Plain. The sites are also situated on two tributaries of the Kordara, one of the rivers forming the Erbil drainage system with the Siwasor and the Bastora Chai, south of the Greater Zab (Peyronel *et al.*, 2019, 1-3, figs 1-2). The two settlements, located just 3 km apart, have a typical Mesopotamian site morphology. They feature a central mound resulting from the superimposition of earlier strata created by the decomposition of previous mudbrick structures, and an occupation of the surrounding area close to favourable natural resources. These represent ideal conditions for a micro-scale regional case study and the analysis of the local cultural and historical development.

MAIPE agreed with the Directorate of Antiquities of Erbil to investigate the two sites following a well-defined research program that aims at building a site-based chronology constructed on stratigraphy, pottery seriation and C14 dating. Moreover, it strives to frame the Helawa-Aliawa occupational history in the broader regional context and to reconstruct the human-environment interaction in the micro-region through time by applying a multidisciplinary approach combining archaeological excavations, micro-stratigraphy, bio-archaeological analyses, and geo-archaeological investigations.

As a result of this integrated approach, during seven years of actual fieldwork (2013, 2015-2019, 2021), MAIPE collected a large amount of data that documents the long-standing occupation of the area, spanning from the Late Neolithic to the Islamic period. All data have been organised to be stored in a GIS environment which is continuously updated, named the MAIPE GIS, based on the open-source QGIS which is increasingly used as an effective medium for managing archaeological and paleoenvironmental data in projects in Western Asia (Ardissone *et al.* 2008; D'Andrea 2003; Frigeri 2012; Kolinski 2019, 12; Pescarin 2006).

The present contribution shows how information from an ongoing research project is processed using a QGIS system and illustrates how this tool helps combine various types of data and carry out multi-scalar and cross-correlated analyses (Orengo 2015).

A Multidisciplinary Approach: Data Management (MC, VO, AV)

The use of GIS for the management and analysis of archaeological data allows the integration of information at different scales within a single spatial framework (Boyd *et al*. 2021; Scianna and Villa 2011; Dell'Unto and Landeschi 2022, 5-17). Thus, MAIPE's choice of an open system (QGIS) is intended to improve data accessibility, encourage result sharing and foster collaboration between the project teams, thanks to the elimination of licence constraints and to the reduction of software costs, all without affecting the work quality (Marchetti *et al*. 2018). QGIS is, in fact, available for all popular operating systems, and layers in shapefile format are compatible with other paid-for software, facilitating the exchange of data between different projects (Casagrande *et al.* 2012).

The MAIPE GIS was initially developed in 2013 through the collection of spatial data on the field, the realisation of a 3D model – a digital elevation model (DEM) of Helawa–, and the digitalisation of the general features of this site and its surroundings. Moreover, an orthorectified WorldView2 satellite image acquired in 2011 was used as a base-map for the 5x5 km area, including Helawa, Aliawa and the surrounding territory (Peyronel, Bursich and Di Giacomo 2016). The first DEM of Aliawa, used as a base map for the survey, was provided in 2016 by the Erbil Plain Archaeological Survey (EPAS), (Ur *et al.* 2021). Further Digital Surface Models (DSM) of the main mound at Helawa and Aliawa were realised by the MAIPE team in 2021.

A correlated table system elaborated in a .XLSX format was designed to organise the information collected by the project's research teams. The choice of using .XLSX spreadsheets instead of a .DB server system was initially determined by the need to work remotely from several devices, even without an internet connection.

These tables are linked through shared fields and, together with the spatial data, constitute the MAIPE GIS' relational geodatabase which is built as a Geopackage (.GPKG format). Data collected during the excavation activities conducted from 2016 at Helawa and 2019 at Aliawa have been progressively georeferenced in the GIS and digitalised as shapefile layers.

This system allows an effective workflow that enables team members to record their data in a standardised format compatible with a GIS environment (see also Ardissone *et al.* 2008).

The data processing is organised in various phases and is structured as showin in fig. 1:

1. Collection of base cartography, satellite and historical images of the investigated area;

2. Acquisition and processing of the topographical information by using Unmanned Aerial Vehicle (UAV) imagery;

3. Manual recording of features in the field through Stratigraphical Units (SUs) sheets, Loci sheets, and sample lists, and in the laboratory through pottery and small finds records, photographs and drawings;

4. Digitalisation of the data into the Survey Database, Excavation Database, and Geoarchaeological Database. These databases are organised as spreadsheets and then saved as .CSV files which are uploaded into the MAIPE GIS as attribute tables linked to their respective shapefiles within the Geopackage;

5. Performance of spatial and cross-correlated analysis through the GIS according to specific research questions.

Data Processing: The GIS for Geomorphological and Archaeological Research

Geomorphological Mapping and Preservation of the Archaeological Landscape (LF, AP)

Geomorphological analysis and mapping are fundamental tools to understand the interplay between natural and anthropogenic landscape shaping processes. In this setting, comparisons between satellite and historical imagery through remote sensing and GIS-based analysis are necessary to discriminate between old and recent modifications of anthropogenic and natural landforms (Hritz 2014; Grabowski and Gurnell 2016). In the Kurdistan Region of Iraq, the interplay between tectonics, climate and human agency shaped the landscape during the Late Quaternary, influencing the evolution of geomorphological processes and landforms (Fouad 2015; Abdulnaby 2018; Zebari *et al.* 2019; Forti *et al.* 2021). The area between Helawa and Aliawa is the subject of a detailed geomorphological mapping, conducted both with remote sensing and field survey, to outline the environmental landscape surrounding the sites. The fieldwork was carried out during the 2019 and 2021 excavation campaigns.

Remote sensing was performed through QGIS and GRASS GIS software to process the elaboration and comparison of Google Earth™, Landsat, WorldView2 satellite data and historical imagery derived from the Corona Atlas of the University of Arkansas 1967-1968

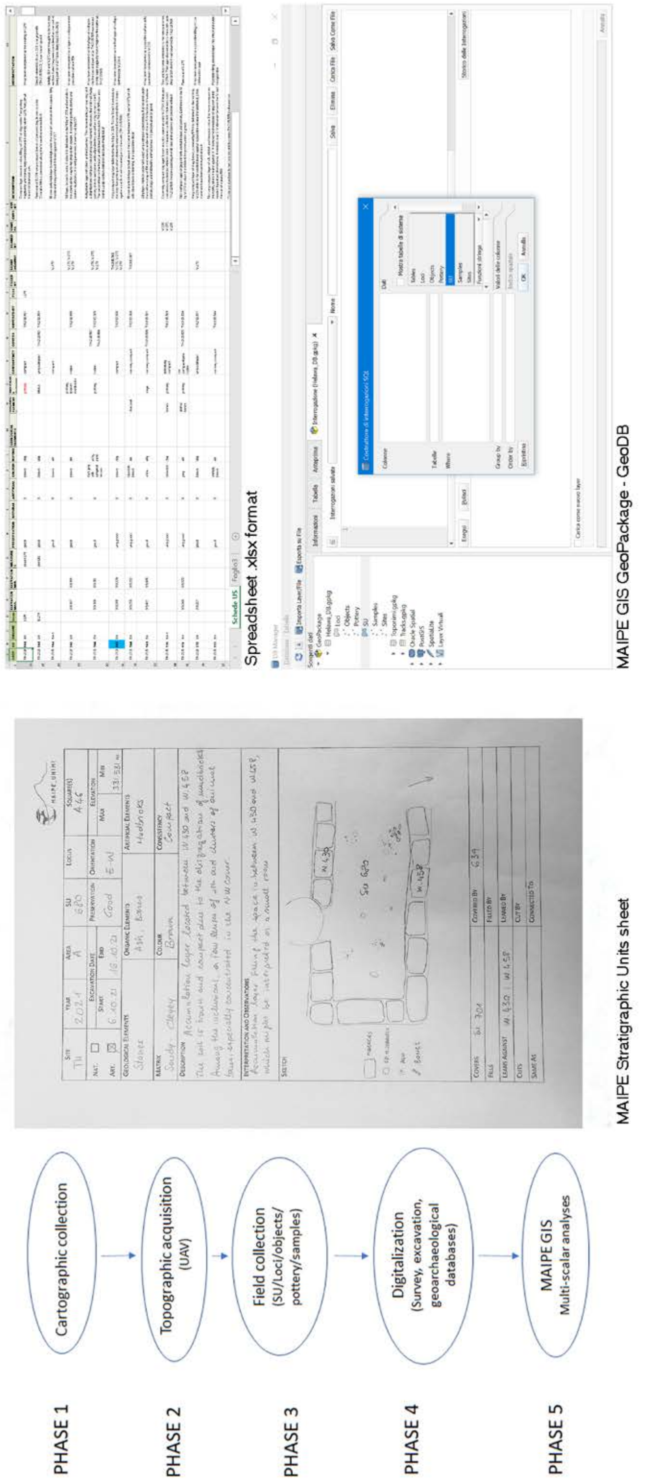

Figure 1. Organization of the MAIPE workflow. © MAIPE – University of Milan

(https://corona.cast.uark.edu/). In addition, a DSM with 1° horizontal resolution (~30 m resolution at the equator) was extrapolated from Alos Jaxa and processed for further analysis such as slope aspect, asperity and hillshade, to highlight specific geomorphological features like escarpment, fluvial landforms, wadi valleys and paleo-valleys related to "underfitted streams model" (Dury 1964). Through this approach, we reconstructed the evolution of the landscape during the Holocene, intending to understand how variations of the local fluvial network influenced settlement dynamics and land use around the two sites. The area is characterised by "underfitted streams", defined by the interplay between the large former valley, whose riverbeds are incised by extant *wadis*, and ancient and recent human exploitation of the channel network and the floodplain.

The comparison of field data and remote imagery analysis was useful to place the two sites in the environmental context existing at the time of their occupation and to identify the natural and anthropogenic surface processes currently underway, which are influencing the preservation of the archaeological record (Forti *et al.* 2020) (fig. 2).

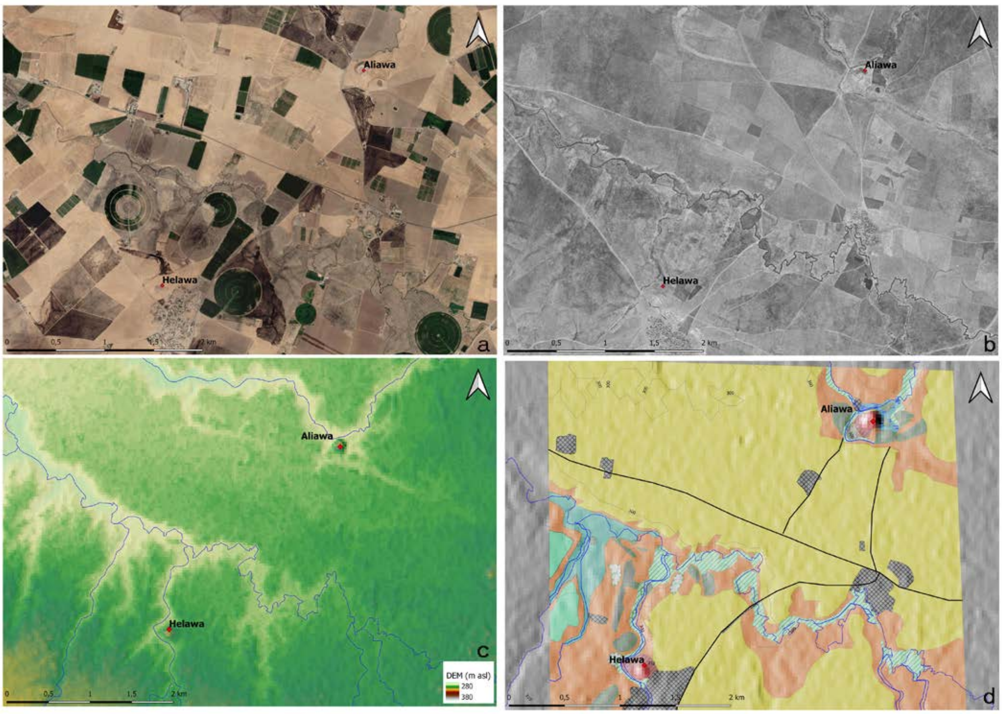

Figure 2. QGIS display of various WMS, raster, and shapefiles from the MAIPE GIS. a) Recent satellite Imagery (GoogleTM Earth). b) Historical satellite Imagery (1967 Declassified CORONA; USGS, 1968). c) 30m (ALOS) DEM (JAXA, 2021) with superposed slope and hillshade (extrapolated through GRASS GIS). d) Geomorphological vector layers (polygons and line geometries). © MAIPE – University of Milan.

The Archaeological Survey: Predictive Intra-site Analyses (MC, VO)

In 2013 and 2015, MAIPE carried out an integrated archaeological survey of Helawa and Aliawa to establish the sites' chronological occupations, to make predictions on their morphologies and extensions within the different periods, and to assess the effects of natural phenomena on the archaeological deposit (Peyronel and Vacca 2015; Peyronel, Vacca and Zenoni 2016).

In order to gain statistically relevant information on the spatial distribution of surface finds for GIS-based analyses, georeferenced Collection Areas (CAs) and Collection Units (CUs) were established on the mounded area of Helawa (c. 1.2 ha), while the lower sector, more disturbed by modern ploughing, was investigated through an extensive survey. These areas were then digitised within the MAIPE GIS as polygon shapefiles, and the finds collected from the field were assigned to the specific CUs thanks to the "Insert Random Point" function in QGIS which allowed us to visualise the different phases of occupation at the site from the Late Neolithic to the mid-2nd millennium BC and to identify the sporadic findings dating up to the Islamic period (Oselini 2020, fig. 4; Peyronel, Vacca and Zenoni 2016; Vacca, Moscone and Rosati 2020).

Similarly, at Aliawa, the site was initially divided into 14 different topographically relevant CAs, departing from the central mound and extending towards the lower sector. Artefacts were collected in each of these areas. Also, in this case, the pottery collected was recorded in QGIS following the method used for Helawa which showed that the site was occupied during the Early Bronze Age (EBA), the Middle and Late Bronze Ages (MBA, LBA), and from the Iron Age (IA) until the Hellenistic/Seleucid and Parthian periods, with the latest settlement dated to the Late Islamic period (Peyronel and Vacca 2020).

In order to assess the extension of the various occupational phases at the sites, the MAIPE GIS was then used to run a combination of spatial analyses such as Kernel Density Estimation (KDE), heatmaps, and the interpolation functions, all of which are available in QGIS. The aim of KDE, which is commonly used in GIS-based archaeological research, is to provide a smoother visualisation of the various degrees of distribution of items around a specific core, given a predetermined radius which may either be set up manually or automatically through the GIS (Bonnier, Finné and Weiberg 2019, 72-73).

Moreover, the integration of survey and geomorphological data enabled us to carry out multi-scalar and predictive modelling on the interplay between natural phenomena and the archaeological deposits at the sites. In the case of Aliawa, for instance, the combination of the geomorphological analysis of natural and anthropogenic processes, the chronological data obtained from the collected sherds, the density clusters detected through the KDE, and the slope analysis run through the QProf plugin available in QGIS, suggested that the southern side of the central mound of Aliawa was potentially of great interest for further research due to the likely presence of an underlying multiphase stratigraphy which would allow a chronological reconstruction of the site's occupation.

Based on these predictions, Step Trench A was therefore excavated in 2019 and 2021 in the south-eastern sector of the mound, where the site's topography was easily accessible (fig. 3).

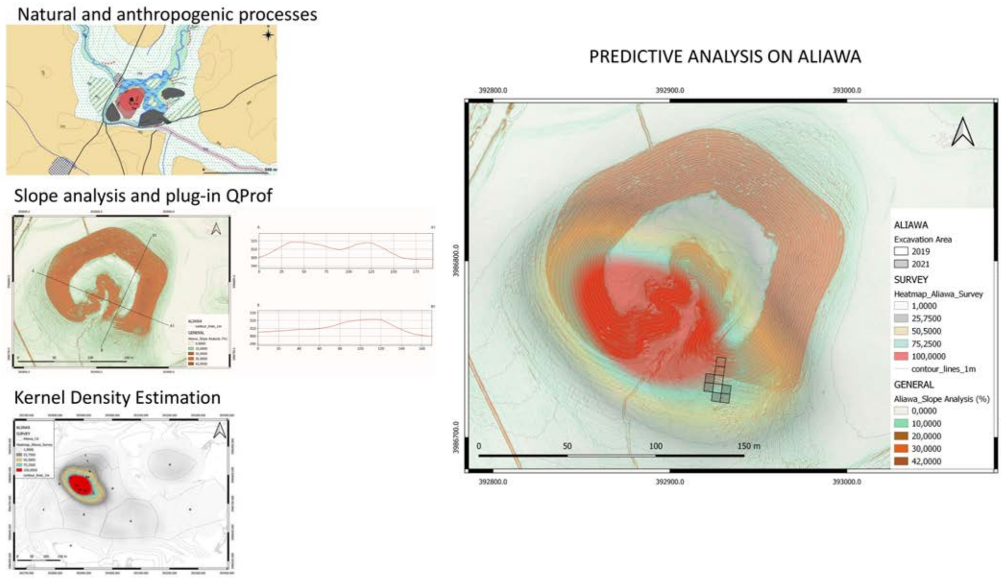

Figure 3. Predictive analysis on Aliawa, after analysing the natural and anthropogenic processes, the slopes and by applying the Kernel Density Estimation. © MAIPE – University of Milan.

QGIS for the Archaeological Excavation (MG, EG, VO)

In addition to its use for large-scale operations such as survey and landscape investigations, the MAIPE GIS is also being employed as the project's repository for the data collected at a stratigraphic and micro-morphological level from the ongoing excavations at Helawa and Aliawa.

Topographic plans of the excavation areas are obtained from photographs. The workflow consists of two stages: the first phase is carried out in the field, while the second may also be performed remotely. During the first phase, control points are positioned within the area of interest and their absolute elevation is recorded through the total station. This operation is necessary to scale and position the photomosaic correctly within the excavation area so that it may be used later in the documentation process. Numerous nadiral photos are then taken through a camera mounted on a telescopic rod or using a UAV (Fig. 4a-b).

Later, during the second phase of the process, the partial photos of the excavation area are assembled both through the photomosaic and photogrammetric techniques to obtain a total image of the desired area with realistic proportions and as less distortion as possible. Once this "base" image is produced, it is imported into QGIS and processed through the "georeferencer" plugin, which links it to the absolute coordinates of the control points previously positioned in the field. This way the photomosaic is correctly scaled, oriented, and positioned within the excavation area and is ready to be digitised in a shapefile format through polygons and polylines in QGIS and further characterised based on the typology of the architectural elements (wall, bench, floor, etc.) for visualisation purposes (Fig. 4c-d).

The shapefile layers corresponding to SUs (stratigraphic unit) and Loci are subsequently linked to their respective attribute tables, derived from the .XLSX spreadsheets transformed into several .CSV, and which contain information regarding the occupational phase, the stratigraphic relation to other layers, dimensions, elevation, and associated materials. Micromorphological samples were also taken from the excavation sections at Helawa (Step Trench B) to investigate the archaeological record at a microscopic scale and were positioned by the Total Station, uploaded to the GIS project, and linked to the metadata in the attribute table associated with the SUs shapefile (Fig. 4e).

Shapefiles and relative attribute tables are all correlated to one another thanks to the relational structure of the MAIPE GIS Geopackage database which allows users to rapidly filter and perform queries for the excavated sequence, carry out distribution analyses of the materials found in the archaeological context and examine the associations between the diverse types of artefacts.

Discussion and future perspective (MC, VO, LP, AV)

The use of QGIS to assemble information from the different teams involved in the MAIPE project is advantageous for the development of analyses and multi-scalar interpretations of the complex relationship between human communities and the natural landscape of the southwestern Erbil plain in a *longue durée* perspective. The combination of geomorphological data and the analyses on the distribution of the archaeological material within a multidisciplinary relational database in QGIS, in addition to the use of various plugins, made it possible to identify specific points of interest and create new insights for further investigation at Helawa and Aliawa.

In terms of future perspectives, MAIPE aims to speed up the work in the field to simplify the data management stage and reduce post-processing and publishing times. In this regard, during the 2021 campaign, MAIPE experimented with the app QField for processing data directly on the field (Montagnetti and Guarino 2021). The geomorphological and archaeological survey in the Aliawa surroundings was implemented to obtain insight into the settlement's spatial development and its geomorphological setting in the different phases of occupation. QField was run on Android tablets and was used as a direct recording tool for surface findings and observations on the territory by incrementing the shapefile layers previously created on the GIS project and corresponding to the survey transects and collection areas, the archaeological and geomorphological features, and the pottery concentrations. The use of this app for the management of the archaeological excavation is still experimental, but it could be used to streamline the documentation of lists during the excavation and for the positioning of pottery and small finds.

MAIPE's choice of an Open Source software like QGIS provides solutions to both practical and research issues and agrees with its pursuit of a policy of open access data for scientific research, a practice which we hope will be increasingly adopted by a wider community. In fact, in addition to being an easily accessible medium in a multidisciplinary research project, QGIS can be transformed into a web-GIS and can be made available to a wider audience of users, both for academic and dissemination purposes.

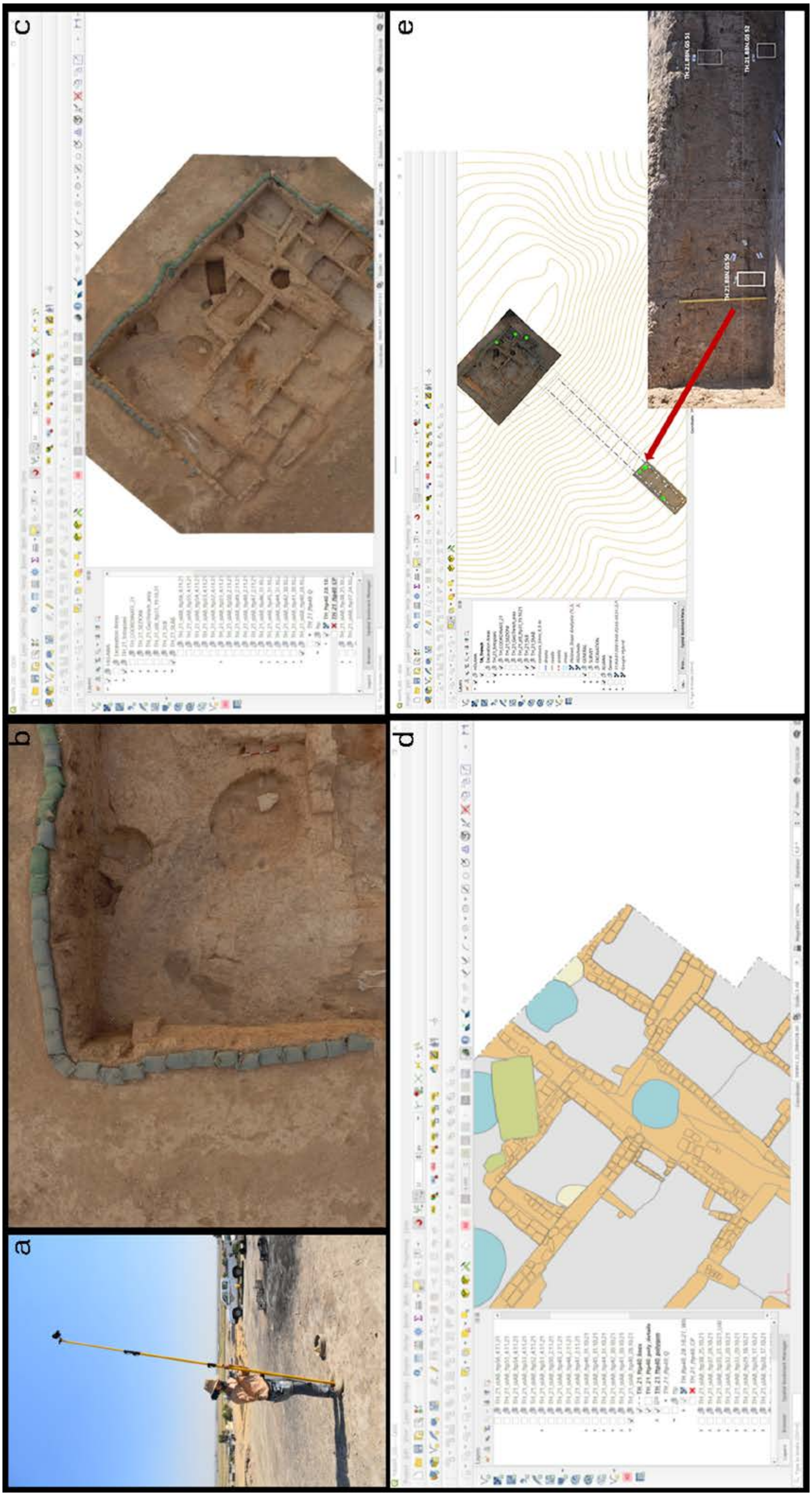

Figure 4. The various steps of the topographic process: a) acquisition of nadiral images on the field; 2) example of a UAV image of the excavation trench at Helawa; c) orthorectified photomosaic of the entire excavated area uploaded and georeferenced in QGIS; d) digitalization of the SUs and Loci in the GIS in geometric shapefiles; e) geolocation of micromorphological samples. © MAIPE – University of Milan

References

Abdulnaby, W. 2018. "Chapter 4 - Structural Geology and Neotectonics of Iraq, Northwest Zagros." In *Developments in Structural Geology and Tectonics Vol. 3,* edited by Ali Farzipour Saein, 53-73. Amsterdam: Elsevier. DOI: 10.1016/B978-0-12-815048-1.00004-4

Ardissone, P., F. Balossi Restelli, E. Borgia, and E. Brienza. 2008. "Web-GIS Solutions for the Analysis and Valorisation of Archaeological Sites in the Mediterranean Basin." In *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA). Berlin, Germany, April 2-6, 2007 (Kolloquien zur Vor- und Frühgeschichte, Vol. 10),* edited by Posluschny A., Lambers K. and Herzog I., 1-7, Bonn: Habelt.

Bonnier, A., M. Finné, and E. Weiberg. 2019. "Examining Land-Use through GIS-based Kernel Density Estimation: A Re-evaluation of Legacy Data from the Berbati-Limnes Survey". In *Journal of Field Archaeology,* 44 (2), 70-83.

Boyd, M. J., R. Campbell,, R.C.P. Doonan, C. Douglas, G. Gavalas, M. Gkouma, C. Halley,B. Hartzler, J. A. Herbst, H.R. Indgjerd, A. Krijnen, I. Legaki, E. Margaritis, N. Meyer, I. Moutafi, N.P. Iliou, D.A. Wylie, and C. Renfrew. 2021. "Open Area, Open Data: Advances in Reflexive Archaeological Practice." In *Journal of Field Archaeology,* 46 (2): 62-80. DOI: 10.1080/00934690.2020.1859780

Casagrande, L., P. Cavallini, P. De Rosa, P. Napolitano, and F. Iancobellis. 2012. "5. Quantum GIS in GIS Open Source." In *GRASS GIS, Quantum GIS e SpatiaLite. Elementi di software libero e applicato al territorio,* edited by Casagrande L., Cavallini, P., Frigeri, A., Furieri, A., Marchesini I., and Neteler, M., 33-40. Palermo: Dario Flaccovio Editore.

D'Andrea, A. 2003. "Analisi Spaziali Intra-Site. Soluzioni GIS per lo Scavo Archeologico." in *Archeologia e Calcolatori,* 14, 329-335.

Dell'Unto, N. and G. Landeschi. 2022. "Geographical information systems in archaeology." in *Archaeological 3D GIS*, 5-17. London and New York: Routledge.

Di Giacomo, G. 2016. "Sperimentazione di un sistema GIS cloud open source per la condivisione e la valorizzazione del patrimonio archeologico." in *Archeologia e Calcolatori,* Supplemento 8, 181-187, Firenze: All'insegna del Giglio.

Dury, G.H. 1964. "Principles of underfit streams." in *Geological Survey Professional Paper, no.* 452, (US Government Printing Office, 1964), 1-67.

Frigeri, A. 2012. "2. Software libero, licenze e dati geografici." in *GIS Open Source. GRASS GIS, Quantum GIS e SpatiaLite. Elementi di software libero e applicato al territorio,* edited by Casagrande L., Cavallini, P., Frigeri, A., Furieri, A., Marchesini I., and Neteler, M., 13-18. Palermo: Dario Flaccovio Editore.

Forti, L., E. Regattieri, A.M. Mercuri, I. Mazzini, A. Pezzotta, A. Florenzano, C. Conati Barbaro, L. Peyronel, D. Morandi Bonacossi, and A. Zerboni. 2020. "Geoarchaeological and paleoenvironmental reconstruction of the Late Quaternary climate- environmental-human nexus in the Kurdistan region of Iraq." in *EGU General Assembly 2020, Online*, (4–8 May 2020): 545. DOI: 10.5194/egusphere-egu2020-545

Forti, L., A. Perego, F. Brandolini, G.S. Mariani, M. Zebari, K. Nicoll, E. Regattieri, C. Conati Barbaro, D. Morandi Bonacossi, H.A. Qasim, M. Cremaschi, and A. Zerboni. 2021. "Geomorphology of the northwestern Kurdistan Region of Iraq: landscapes of the Zagros Mountains drained by the Tigris and Great Zab Rivers." in *Journal of Maps*, 225-236. DOI: 10.1080/17445647.2021.1906339

Fouad, S.F.A. 2015. "Tectonic map of Iraq, scale 1:1000,000, 3rd edition." in *Iraqi Bulletin of Geology and Mining*, 11, 1–7.

Grabowski, R.C., and A.M. Gurnell. 2016. "Chapter 4. Using historical data in fluvial geomorphology." In *Tools in Fluvial Geomorphology, Second Edition*, edited by G. Mathias Kondolf and Hervé Piégay, 56-75. Hoboken: JohnWiley & Sons.

Hrtiz, C. 2014. "Contribution of GIS and Satellite-based Remote Sensing to Landscape Archaeology in the Middle East." in *Journal of Archaeological Research,* 22, 229–276. DOI: 10.1007/s10814-013-9072-2

Koliński, R. 2019. *Catalogue of Archaeological Sites. Navkūr Plain, Al- Ḫazir River Basin. Settlement History of Iraqi Kurdistan Vol. 3.* Wiesbaden: Harrassowitz Verlag.

Marchetti, N., I. Angelini, G. Artioli, G. Benati, G. Bitelli, A. Curci, G. Marfia, and M. Roccetti. 2018. "NEARCHOS. Networked Archaeological Open Science: Advances in Archaeology Through Field Analytics and Scientific Community Sharing." In *Journal of Archaeological Research,* 26, 447-469. DOI: 10.1007/s10814-017-9112-4

Montagnetti, R. and G. Guarino. 2021. "From QGIS to QField and Vice Versa: How the New Android Application Is Facilitating the Work of the Archaeologist in the Field." in *Environmental Sciences Proceedings,* 10 (1), 6. DOI: 10.3390/environsciproc2021010006

Orengo, H. 2015. "5 Open-Source GIS Geospatial Software for Archaeology: Towards its Integration into Everyday Archaeological Practice." In *Open-Source Archaeology,* edited by Andrew T. Wilson, 64-82. Warsaw: De Gruyter Open.

Oselini, V. 2020. "Defining the MB-LB Transition in Northern Mesopotamia: Some Archaeological Considerations on the New Data from the Erbil Plain and Neighbouring Regions." In *From the Prehistory of Upper Mesopotamia to the Bronze and Iron Age Societies of the Levant. Proceedings of the 5th "Broadening Horizons" Conference (Udine 5-8 June 2017), Vol. 3,* edited by Costanza Coppini and Francesca Simi, 206-220. Trieste: EUT.

Pescarin, S. 2006. "Open Source in archeologia. Nuove prospettive per la ricerca." In *Archeologia e Calcolatori*, 17, 137-155.

Peyronel, L. and Vacca. A. 2015. "Northern Ubaid and Late Chalcolithic 1-3 Periods in the Erbil Plain. New Insights from Recent Researches at Helawa, Iraqi Kurdistan." In *Origini,* 37, 89-127.

Peyronel, L., D. Bursich, and G. Di Giacomo. 2016. "QuantumGis per la gestione dei dati della survey 2013 a Helawa nella piana di Ebil, Kurdistan, Iraq." In *Archeologia e Calcolatori,* Supplemento 8, 88-96.

Peyronel, L, A. Vacca, and G. Zenoni. 2016. "Helawa: A New Northern Ubaid/Late Chalcolithic Site in the Erbil Plain." In *The Archaeology of the Kurdistan Region of Iraq and Adjacent Regions,* edited by Kostantinos Kopanias and John MacGinnis, 309-321. Oxford: Archaeopress.

Peyronel, L., C. Minniti, D. Moscone, Y. Naime, V. Oselini, R. Perego, and A. Vacca. 2019. "The Italian Archaeological Expedition in the Erbil Plain, Kurdistan Region of Iraq. Preliminary Report on the 2016-2018 Excavations at Helawa". In *Mesopotamia,* 54: 1-104.

Peyronel, L. and A. Vacca. 2020. "The Italian Archaeological Project in the Erbil Plain (2013-2017)". In *Proceedings of the 11th International Congress on the Archaeology of the Ancient Near East: 03-07 April 2018, Munich. Vol. 2. Field Reports, Islamic Archaeology,* edited by Adelheid Otto, 317-330, Wiesbaden: Harrassowitz Verlag.

Scianna, A. and B. Villa. 2011. "GIS applications in Archaeology." In *Archeologia e Calcolatori, 22*: 337-363.

Ur, J., N. Babakr, R. Palermo, P. Creamer, M.N. Soroush, S. Ramand, and K. Nováček. 2021. "The Erbil Plain Archaeological Survey: Preliminary Results, 2012–2020". In *IRAQ,* 83: 1-39.

Vacca, A, D. Moscone, and P. Rosati. 2020. "Mapping Survey Data from Helawa, Erbil Plain (Kurdistan Region of Iraq)". In *From the Prehistory of Upper Mesopotamia to the Bronze and Iron Age Societies of the Levant. Proceedings of the 5th "Broadening Horizons" Conference (Udine 5-8 June 2017), Vol. 1,* edited by Marco Iamoni, 107- 124. Trieste: EUT.

Zebari, M., C. Grutzner, P. Navabpour, and K. Ustaszewski. 2019. "Relative timing of uplift along the Zagros Mountain front flexure (Kurdistan Region of Iraq): constrained by geomorphic indices and landscape evolution modelling". in *Solid Earth*, 10: 663–682. DOI: 10.5194/se-10-663-2019

A digital ecosystem for the knowledge, conservation and valorization of the medieval archaeological site of *Satrianum* (Tito, PZ). FOSS instruments

Giorgia Dato[1], Eugenio Saccà[2] and Alessandro Spadaro[3]*

[1] *Ricercatrice indipendente.*
[2] *Dipartimento delle Culture Europee e del Mediterraneo, Università degli Studi della Basilicata, Matera, Italia.*
[3] *Dipartimento Interateneo di Scienze, Progetto e Politiche del Territorio, Politecnico di Torino, Torino, Italia.*
*Correspondence: giorgia.dato90@gmail.com
Presented at ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract:
The case study concerns the fortified mediaeval site of *Satrianum* in Basilicata. Topographical, archaeological and architectural analysis permitted the creation of a spatial database with QGIS: further developments involve the survey of the site and the elaboration of 3D digital models to be georeferenced in the space. The complete data set obtained for the knowledge of the site allows to design projects for conservation, valorisation and usage of the site itself at several levels.

Keywords: *Satrianum,* medieval archaeology, landscape archaeology, archaeology of architecture, heritage architecture, digital humanities, QGIS, 3D modelling, digital survey, structure-from-motion

FOSS software used and license: WebODM (*GNU Affero General Public License v3.0*), QGIS, (*GNU Public License (GPL) Version 2*), Blender (GNU General Public License v3.0)
Open data set and license: Tarquini et al. 2007, *TINITALY*, *a digital elevation model of Italy with a 10 meters cell size (Version 1.0) [Data set].* Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/*TINITALY*/1.0.; (*CC BY 4.0*)

Introduction

The *Satriaunum* archaeological project is under the scientific direction of Francesca Sogliani and SABAP-MIC. A particular mention goes to Simona Di Gregorio, who has allowed this work to be carried out.

The study of the archaeological site of *Satrianum*, on concession from SABAP - MIC, began in 2000 under the direction of Massimo Osanna, in collaboration with the Superintendence for Archaeological Heritage of the Basilicata and it is part of a multidisciplinary research project of the Postgraduate School in Archaeology of Matera. Since 2005, the investigation has focused on the mediaeval settlement under the scientific direction of Francesca Sogliani (Osanna 2011; Colangelo 2011; Osanna, Capozzoli 2012).

Satrianum is currently included in an enhancement project aimed at defining the physiognomy of the Open Museum through paths of public archaeology (Format "Festivalia. Archaeology tells itself"), digital storytelling and virtual reconstruction. It is also one of the SmartLabs included in the Basilicata Heritage Smart Lab project coordinated by the Basilicata Creative Cluster.

The area of the archaeological site of *Satrianum*[1] used to be a mediaeval fortified settlement. It is distributed on a hill dominating the surrounding area and is located between the present-day cities of Tito and Satriano di Lucania in Basilicata (fig. 1).

Particularly important for its strategic position along axes of territorial connection that link the Ionian, Adriatic and Tyrrhenian coasts and for the abundance of resources, the site of *Satrianum* has been occupied extensively since the VIII century BC, as indicated by the important results of the archaeological investigations that were carried out. In particular, the data that emerged from archaeological excavations allowed the reconstruction of the settlement dynamics of the site during the Archaic period, from the VIII century BC to the V-IV century BC. The most important phase of growth of the pre-Lucan indigenous settlement started in the VI century BC; then, with the arrival of Lucanian people in the IV-III century BC, great transformations took place in this area (Osanna and Capozzoli 2012).

Figure 1. (a) overview of the tower and the area of the cathedral (2021): the picture was taken during a drone flight for the survey; (b) area of the cathedral during the excavation (SOGLIANI 2011); (c) General view of the hill (2020)

[1] We want to thank the scientific director of the excavation, Francesca Sogliani, and SABAP-MIC, with particular reference to Simona Di Gregorio, who have allowed this work to be carried out.

Traces of transformations during the Roman period (III-II century BC) give back the physiognomy of a much less populated area, at least until the Middle Ages when the town of *Satrianum* was born. Built in the last decades of the XI century, it is a particularly strategic area, also for its topographical characteristics, which constitutes a case study of particular importance: it represents the result of a process of hierarchisation and restructuring of spaces and structures willed by the noble initiative to control territories and the ways of exploiting the countryside (Sogliani 2017).

The progressive decline of the area led to the abandonment of the settlement in the XV century.

The mediaeval settlement of *Satrianum*, built using local limestone found in situ, represents an exemplary case of an abandoned village that, having not been affected by post-medieval and modern renovations, has maintained its architectural features and original plans sufficiently intact (Sogliani and D'Ulizia 2008).

The internal organisation of the site consists of different architectural complexes, and it is structured in three main areas:

A summit area that was enclosed by a wall with a square tower –the seat of lay power–, and a cathedral with an adjoining episcope, –the seat of ecclesiastical power–.

a first village along the western slopes.

a second village on the southern side.

The course of the research and study led to the analysis of the state of conservation of the fortified settlement of *Satrianum*, and the design of a feasibility project for its restoration through the programming of a dedicated digital archiving system (a database on QGIS).[2]

The fact-finding survey was related to an in-depth work of archival research as well as plans for intervention on buildings.

In the case of *Satrianum*, a detailed survey of the site was carried out aiming at assessing all problems rigorously, and at reducing any risks of approximate and/or superficial assessments.

The aim of the database is to obtain a uniform corpus connecting different kinds of information – architectural, technical, typological, etc. – to deduce new data for interpretation.

The hierarchical organisational model of the Reference Units, developed by Gian Pietro Brogiolo, was chosen as a reference for the digital archive system of recording and relating the structural and functional data of the archaeological-architectural components, both with

[2] The archaeological conservation project for *Satrianum* represented the undersigned's thesis entitled "Archeology and restoration of monuments. Prospects for research in the medieval fortified site of *Satrianum* in Basilicata" (Supervisor: Francesca Sogliani; co-supervisor: Filiberto Lembo). Furthermore, the project was introduced to the Superintendence of Archeology, Fine Arts and Landscape of Basilicata. The project particularly concerned the cathedral and some buildings of the first and second village of the archaeological site.

typological information – building typology, type of materials – the state of conservation and the relative forms of decay (Mangialardi and Sibilano 2011, 77-78).

The database consists of a preliminary section containing the personal data; another one with the archaeological-architectural data; a third part is dedicated to recording the various types of decay and the state of conservation; finally, a fourth part, the last one, proposes the recording of the structural vulnerability: this analysis is based on qualitative methods for the valuation of the masonry.

The collected information represents a solid basis for the elaboration of synthetic visions. Also, the analysis of the architectural artefact as the result of several constructive actions allows to evaluate both the state of conservation and the material characteristics and structural consistency.

Underlining the importance of creating a heterogeneous working group that can generate contributions from various points of view to the conservation project, it would be desirable that this preliminary work could be a starting point for an even more in-depth future work on *Satrianum*.

Working on a complex and structured analysis from an urban, functional and geomorphological point of view, relating to the existing structures and the routes to the level curves (for example, to identify and prevent the behaviour of the site during meteorological events of both moderate intensity and daily frequency as well as particularly violent ones) could increase the knowledge of the site under investigation.

The study of the state of conservation, especially for the possible degradation processes starting from the post-excavation phase, is a mandatory step in thinking and planning the future of archaeological sites.

In conclusion, it is our opinion that only adequate wide-ranging planning can constitute the essential first link of an integrated system. Nontheless the various issues, the needs and the available technologies may be considered in a multidisciplinary perspective.

(G.D.)

QGIS for the structure and development of spatial databases

To catalogue, organise and manage the data, we used a GIS platform, an essential tool for carrying out these types of operations interconnected with spatial analysis (Forte 2002; Bogdani 2009; Montagnetti and Rosati 2019; Noti 2021).

The focus of the project is the creation of a geographic database that allows documenting and systematising information through geolocation.

For the development of the project, we used the open-source software QGIS version 3.16 “Hannover” (http // wwwqgis.org) with GPL licence (GNU General Public License, commonly referred to by the acronym GNU GPL or simply GPL, is a licence for free software). QGIS, through various functions and a simple graphical interface, allows managing various input layers such as raster, vectors, and alphanumeric information. Furthermore, it allows expanding the workflow through different plugins.

The reference system used is Monte Mario / Italy Zone 2 (zone E) - Datum: Rome 40 - Projection: Gauss-Boaga the EPSG 3004 (European Petroleum Survey Group http://www.epsg.org); a cartographic system which uses the decimal metre to define coordinates, unlike geographic systems which apply coordinate expressions in degrees (https://docs.qgis.org/3.22/en/docs/gentle_gis_introduction/coordinate_reference_systems.html)

The construction of the database for the management of the USM of the *Satrianum* (PZ) site consists of two different types of data: rasters and vectors.

For the rasters, we used a cartographic base from the Google Satellite web tiles services, available through the Quickmap Service plugin (Licence CC-BY- SA, Creative Commons Attribution – Share-Alike 3.0). The other rasters were orthophotos taken by drone in jpg format with a 10×10 m pixels resolution, relating to the three main areas of the site: the cathedral (CF1, CF2, and CF3); the area of the first village (CF 23 and CF 104); the area of the second village (CF 59 and CF 75 and CF 79).

For the vectors, we chose a polygonal geometric layer within a geopackage; this format, based on OGC (The Open Geospatial Consortium – OGC) standards, is an information package that can be used as a database, an open compact format.

For this case study, a database was created in a GPKG file, which presents a single table organised to easily interact with the data and make queries using the SQL language (Ferrero 2004).

The first practical operation consisted of autoptic georeferencing the greater detail orthophotos of the three areas, using known points on the background map.

Through the **"Georeferencer**" tool, we used the transformation type "Helmert" (QGIS, like other GIS software, which has algorithms that manage the georeferencing of raster images. The algorithm is chosen based on the number of points we use, the quality of the graphic file, and the distortion or error in the final result. https://docs.qgis.org/3.22/en/docs/user_manual/working_with_raster/georeferencer.html), which offers the possibility to make roto translations and scale variations on the map with the "nearest neighbour " resampling method.

After this georeferencing activity, we continued with the GPKG database, creating the polygonal vector layers and attribute tables. To organise the table's field, we used, as a source, the criteria for the masonry techniques description of the ICCD (Criteria for the description of masonry techniques for the preparation of coded planning forms, ICCD – Central Institute for Catalog and Documentation, 2013 – translation from ICCD "Criteri di descrizione delle tecniche murarie per la predisposizione di moduli schedografici codificati" – and we have organised the items into four tabs:

#	Nome	Tipo	Nullo	#	Nome	Tipo	Nullo
0	fid	INTEGER	N	27	Presenza di intonaco	TEXT	Y
1	geom	POLYGON	Y	28	Elementi di reimpiego	TEXT	Y
2	Località	TEXT	Y	29	Interventi pregressi di consolidamento	TEXT	Y
3	Sito	TEXT	Y	30	Descrizione	TEXT	Y
4	Area	TEXT	Y	31	stato_conservazione_malta	TEXT	Y
5	Funzione	TEXT	Y	32	stato_conservazione_materiali lapidei	TEXT	Y
6	Entità	TEXT	Y	33	stato_conservazione_intonaco	TEXT	Y
7	Foto	TEXT	Y	34	tipologia/e degrado/i	TEXT	Y
8	Rilievo	TEXT	Y	35	Rischio principale	TEXT	Y
9	Esposizione	TEXT	Y	36	Rischio/i secondario/i o derivato/i	TEXT	Y
10	Elemento_architettonico-strutturale	TEXT	Y	37	Motivo della schedatura	TEXT	Y
11	Materiali lapidei_Litotipo	TEXT	Y	38	Elementi sensibili_qualità muraria	TEXT	Y
12	Materiali lapidei_Lavorazione	TEXT	Y	39	Elementi sensibili_presenza diatoni	TEXT	Y
13	Materiali lapidei_Pezzatura	TEXT	Y	40	Elementi sensibili_forma elementi resistenti	TEXT	Y
14	Materiali lapidei_Colore	TEXT	Y	41	Elementi sensibili_dimensioni elementi resistenti	TEXT	Y
15	Materiali lapidei_Provenienza	TEXT	Y	42	Elementi sensibili_sfasamento giunti verticali	TEXT	Y
16	Malta_Tipo	TEXT	Y	43	Elementi sensibili_orizzontalità filari	TEXT	Y
17	Malta_Funzione	TEXT	Y	44	Elementi sensibili_resistenza degli elementi	TEXT	Y
18	Malta_Colore	TEXT	Y	45	Valutazione complessiva	TEXT	Y
19	Apparecchiatura muraria	TEXT	Y	46	Peggioramento nel tempo	TEXT	Y
20	Posa degli elementi	TEXT	Y	47	Documentazione esistente	TEXT	Y
21	Zeppe e scaglie	TEXT	Y	48	Livello di gravità generale	TEXT	Y
22	Ricorsi e listature	TEXT	Y	49	Elementi di gravità particolari	TEXT	Y
23	Diatoni	TEXT	Y	50	Necessità di approfondimento	TEXT	Y
24	Nucleo	TEXT	Y	51	Valutazione complessiva e note	TEXT	Y
25	Tipologia	TEXT	Y	52	Riferimenti ad altre schede, allegati	TEXT	Y
26	Spessore dei giunti	TEXT	Y	53	Compilatore/i	TEXT	Y
				54	Data	DATETIME	Y

Figure 2. The four different sheets: sheet 1 "Dati Anagrafici" goes from field 1 to 9; sheet 2 "Dati archeologico-architettonici" goes from field 10 to 30; sheet 3 "Stato di conservazione" goes from field 31 to 34; sheet 4 "Vulnerabilità strutturale" goes from field 35 to 54.

the first group is related to personal data, like the names of the entity to be catalogued and digitised,

a second group contains the archaeological-architectural analysis data,

the third group is related to the USM conversation state,

the fourth and last to the structure vulnerability, as it is possible to see in fig. 2.

The last part related to the structural determination referred to the analysis based on qualitative evaluations of the masonry to obtain quantitative results of the mechanical characteristics (Borri *et al.* 2011). Therefore, it was worthwhile and necessary to consider the hierarchy of the referential realities, the Reference Units, that compose the concept of historic building developed by Gian Pietro Brogiolo (Brogiolo 1988).

To better define and organise the data groups, we used tabs within the table and widgets, as it is possible to see in fig. 3.

We can set these functions from the layer properties through the Attributes Module GUI, which works directly on the structure of the fields. To organise the tabs, we used the "Drag and Drop Designer" function of QGIS 3.16. Afterwards, we used the "value map" widget to facilitate the compilation of the table. The latter allows the formation of drop-down menus with different values, according to the database values, that can be entered in the respective fields. This is a modality that facilitates and limits the possibility of error in the structure rafting. Lastly, we used the "attachment" type of widget to insert the photos.

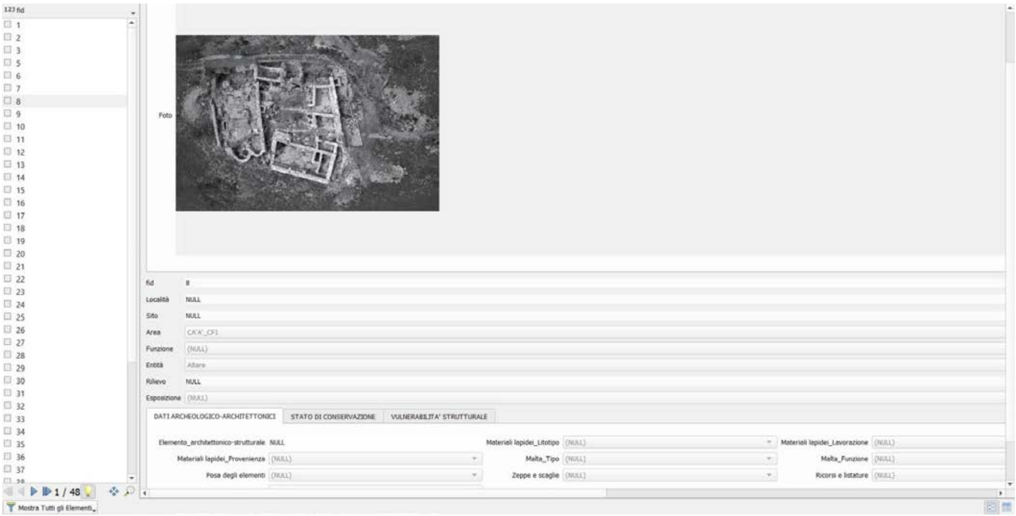

Figure 3. Structure of the attribute table.

After structuring the table, we continued with the digitisation operations through the polygonal vector of the USM from the orthophotos. In the last phase, we continued with the categorisation of the entities based on the "Entità" field for printing and drafting tables. In this way, we can differentiate the geometries by colour, name and functionality.

We have printed three tables relating to the different architectural complexes with relative examples of a database query and with the "Atlas" function, a tool that offers the possibility of overcoming the static nature of the print composer and making composite prints in series.

The structuring of the workflow in QGIS has enabled extensive data management and precise organisation of the design operations.

(E.S.)

Survey and 3D GIS modelling

As it has already been said, by choosing this particular workflow – Luca Mandolesi, "OpenDroneMap: fotogrammetria open source | WebODM + QGIS + Blender", May 17, 2021, video, https://www.youtube.com/watch?v=El4F38gcav0) – we utterly believed in the multidisciplinary approach to get an integrated system aimed at the valorisation of the archaeological site. We tried to obtain a foundation for a complex analysis that could serve different purposes: urban planning study, geomorphology, and architectural function (for a deeper insight into the role of 3D modelling in archaeological research, we would like to cite Hemon, Joanna 2008).

For these reasons, as a first fundamental step, we decided to obtain the geometrical acquisition of the archaeological site with its structures and its geomorphology. To do so, we proceeded to a 3D survey to generate a digital replica of the site. After having obtained permissions from the University of Basilicata and the office of the Superintendence of Archeology, Fine Arts and Landscape of Basilicata, part of the work team travelled to the site and carried out the operation on the field. An Unmanned Aerial System model DJI Mavic Pro was used to perform the acquisition of the photographs data set from the zenith perspective with two crawls. The UAS camera had a 12.71-megapixel matrix equipped with a 26 mm (35 mm equivalent) lens and permitted high-quality images to be captured. The survey was carried out by one operator who took care of covering the entire area and recorded sufficient overlaps for the photos; a second operator assisted as an observer. The drone was piloted at approximately 20 m to get a proper resolution of the whole site: two flights and more than 30 minutes were needed for the acquisition. The survey was carried out at noon to get the best lighting possible. For future campaigns, however, it would be desirable to continue with the terrestrial completion measures and concentrate on individual monuments.

As a second step, the images were processed with WebODM, an open source photogrammetry software for 3D spatial data generation in absolute coordinates. The software allows importing an image data set to be computed by an algorithm of structure from motion; afterwards, it matches bi-univocal correspondences among them to reconstruct the camera position and, therefore, the orientation in space. Finally, the software can elaborate a sparse point cloud, a dense point cloud, and a mesh. In our case, the geospatial references were obtained by UAS GPS which can already georeference the pictures – the average error consists of a value of 2 m ca. – nonetheless, it would be preferable for future developments to match the photogrammetry with Ground Control Points coming from a more precise topographical survey on the field. The interface allows us to regulate different parameters and obtain several kinds of outputs: in our case, we wanted to generate a Digital Surface Model, an orthophoto in raster format, a mesh in OBJ format, and the relative texture. To do so, we used the standard parameters that the software suggests, with a few exceptions: we set the parameters to get the orthophoto from the 3D texture and to raise the mesh size.

The DSM had a resolution of c. 0,24 m per cell and was imported in QGIS to elaborate contour lines (1 m wide) and save them as a shapefile., A section of the Digital Terrain Model was downloaded from the online dataset from TINITALY DEM by the National Institute of Geophysics and Volcanology (INGV) to validate the correctness of the elaboration. The raster was imported in QGIS and used to elaborate contour lines as well, then exported in shapefile.

All these data sets were lastly imported into the open source software for 3D modelling, Blender. We immediately turned on the add-on BlenderGIS, which allows working in an absolute coordinate system 3D space. At this point, it was possible to import the OBJ for the archaeological site and its texture and the shapefile of its contour lines. We also imported the shapefile of the TINITALY DEM contour lines, which were then interpolated to obtain a mesh through the Delaunay method: the two altimetry data sets matched satisfyingly, so our elaboration was validated.

Afterwards, we converted the polygons related to the walls from the geopackage to the shapefile format, and then we imported them: these were organised inside the outliner in different "collections" (groups of layers) according to a chronological principle to allow a more handful consultation of the SU. Figure 4 allows a general overview of the final outcome of the work inside Blender.

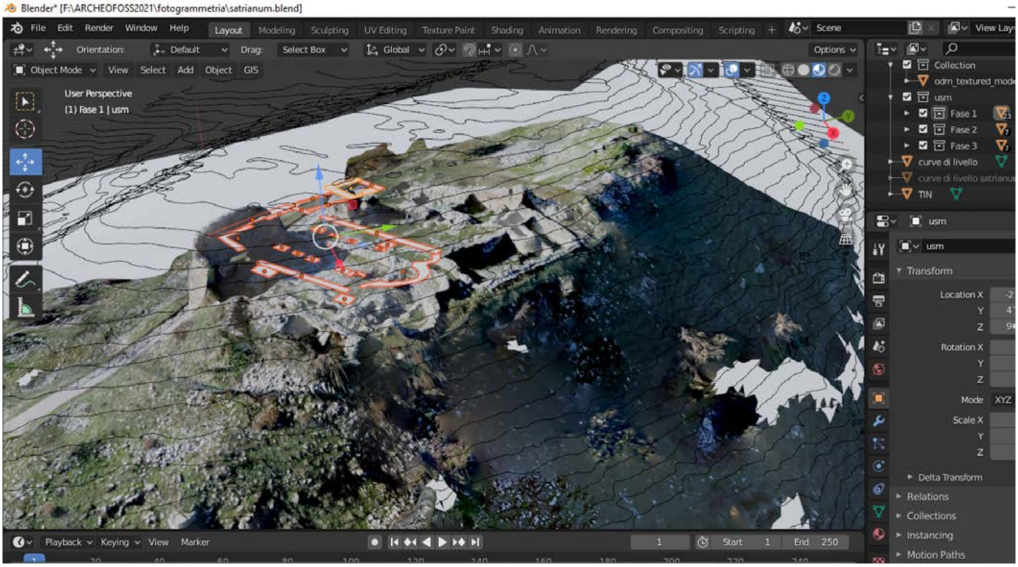

Figure 4. Screenshot from the Blender 3D viewport at the end of the elaboration. It is possible to appreciate the 3D modeling of the site with its texture; the contour lines; the polygons imported as shapefiles from QGIS and related to the different UUSSMM of the buildings; the organization of the polygons inside the outliner in the collections according to chronological principles.

In our opinion, the data set that we obtained on the field and organised in a GIS database inside a 3D modelling software could reveal as utterly useful from different and multidisciplinary perspectives, for which we hope the works on this beautiful site could be enhanced and carried on in the future.

First, it gives a geometrically and geographically precise model to be used for architectural and landscape project design, e.g. by architects and engineers teams to improve the fruition of the site for the audience. It would be extremely helpful for designing better visiting ways all along the site, preventing geomorphological instabilities for the remains, or designing new museum elements for the didactic of the site.

Secondly, the dataset provides precious information for a topographical and orographic study of the site to better understand the ancient centre and its relationship with the territorial surrounding context. It would be desirable to go further in this way, enlarging the scale and having better and more interesting developments in this landscape archaeology.

Thirdly, the digital output could be useful to be used for the enhancement of the site and its digital communication, for example, from a museum or a tourist perspective. Again, it would be preferable for the future to put this case on a larger scale in a regional network.

Finally, the digital analysis, along with the site inspection and its organisation in an interactive virtual database, gives a fundamental stratigraphic comprehension of the complex structures and the ancient buildings: this could be fundamental for accurate restoration design projects and the reinforcements of the structures as well. Nevertheless, such a data set could be likely implemented in further digital and scientifically accurate elaborations: e.g., the case study would be perfect for development according to the Extended Matrix paradigm (Demetrescu 2018), which would give a scientifically precise reconstruction of the fortified site in its historical development.

(A.S.)

References

Bogdani, J. 2009. "GIS per l'Archeologia". In *Atti della Tavola rotonda Bologna 12 aprile 2008- Groma 2. In profondità senza scavare. Metodologie di indagine non invasiva e diagnostica per l'archeologia*, edited by E. Giorgi, 421-438, Roma: BraDypUS. .

Borri, A., G. Cangi, A. De Maria, and C.Donà. 2011. "Metodi qualitativi per la valutazione della qualità muraria". In *Manuale delle murature storiche. Analisi e valutazione del comportamento strutturale*, vol. 1, edited by C. Dona, 237-294, Roma.

Carpenter, J. 2008. "Open Source GIS for archaeology", in *Archaeology in New Zealand,* 51 (4): 257–263

Brogiolo, G. P. 1988. Archeologia dell'edilizia storica. COMO.

Colangelo, L. 2011. "Lo scavo dell'insediamento medievale di *Satrianum*. Le indagini 2006-2009", *in Gli spazi del potere civile e religioso dell'insediamento fortificato di Torre di Satriano in età angioina,* edited by Lucia Colangelo, Massimo Osanna, Anna Rita Parente, and Francesca Sogliani, 230-232. Firenze.

Demetrescu, E. 2018. "Virtual Reconstruction as a Scientific Tool: The Extended Matrix and Source-Based Modelling Approach", in *Digital Research and Education in Architectural Heritage: 5th Conference, DECH 2017, and First Workshop, UHDL 2017, Dresden, Germany, March 30-31, 2017, Revised Selected Papers,* 102-116. Cham: Springer Publishing.

Ferrero, M. 2004. *SQL: la guida tascabile al linguaggio di interrogazione dei database.* Milano.

Forte, M. 2022. *I Sistemi Informativi Geografici in Archeologia*. Roma .

Hermon, S and N. Joanna. 2008. "3D Modelling as Scientific Research Tool". In *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA):* 1-6, edited by Axel Posluschny, Karsten Lambers, and Irmela Herzog. Berlin

Mangialardi, M. and M.G. Sibilano. 2011. "ArchiDb: uno strumento *open source* per l'analisi degli elevati antichi" in *ARCHEOFOSS. Open Source, Free Software e Open Format nei processi di ricerca archeologica:* 75-85.

Montagnetti, R., and P. Rosati. 2019. "Georiferire la stratigrafia archeologica", in *Archeologia e Calcolatori*, 30: 463-466.

Noti, V. 2021. *Gis Open Source per geologia e ambiente: analisi e gestione dei dati territoriali e ambientali con QGIS*. Palermo.

Osanna, M. 2011. "Un progetto di archeologia globale a Torre di Satriano", in *Colangelo, Lucia, Massimo Osanna, Anna Rita Parente, Francesca Sogliani, Gli spazi del potere civile e religioso dell'insediamento fortificato di Torre di Satriano in età angioina:* 229-230.

Osanna, M, and V. Capozzoli. 2012. "Lo spazio del potere II. Nuove ricerche nell'area dell'anaktoron di Torre di Satriano". In *Atti del terzo e quarto convegno di studi su Torre di Satriano- Tito (16-17 ottobre 2009; 29-30 settembre 2010),* 9-11. Venosa.

Virtual Tour realizzato con Pannellum ed integrazione LeafletJS

Valerio De Luca[1] and Martina Frau[1]*

[1]Una Quantum inc;
*Correspondence: frau.marti@tiscali.it
Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract:

A virtual tour is a tool for knowledge, conservation, enhancement, inclusion and accessibility. The images, taken with a spherical camera, are processed with Pannellum that uses Python and Hugin software to create "tiles" of spherical images, and the whole process takes place locally. Photos are linked in order to simulate the route within the virtual tour. On the virtual tour to the traditional "hotspot" button (to link photos) we added two bars: "titlebar" (the top bar for the title) and "footerbar" (the bottom bar for the credits), in addition to buttons for image rotation, full screen view and to open the personalised map created with Leaflet, useful for user orientation.

Keywords: Virtual tour, Pannellum, immagini equirettangolari, foto 360, LeafletJS, accessibilità culturale.

FOSS software used and license: GIMP, GNU LGPL v3+ e GNU GPLv3+; G'MIC, GNU GPL v2; Pannellum, MIT license; Hugin, GNU GPL v2; LeafletJS, BSD license.

Introduzione

Il virtual tour sta conoscendo sempre maggiore applicazione e diffusione nel settore culturale. A tal proposito è bene sottolineare che si è imparato a conoscere e apprezzare questo strumento specialmente in tempo di pandemia, momento storico in cui ha rivelato maggiormente le sue potenzialità e versatilità.

Nell'ambito dello sviluppo del Virtual Tour di Pannellum, i cosiddetti periodi di *lock down* durante la pandemia COVID-19 sono stati, per gli autori, il massimo momento di sperimentazione nell'arco di vita del suddetto progetto. Difatti, erano già in uso strumenti gratuiti e open source per la creazione di virtual tour ma si è deciso di ampliare le conoscenze e testare nuovi mezzi che fossero totalmente open source: fino a questo momento infatti si adoperava solo Marzipano tool che, per quanto open source, è parzialmente proprietà di Google per le sue funzionalità online (https://www.marzipano.net/). L'obiettivo del progetto è dunque quello di dimostrare che non occorrono strumenti eccezionali per giungere alla realizzazione di un ottimo virtual tour che sia, soprattutto, di facile utilizzo e accesso per tutti.

Origine del VT

Se attualmente il virtual tour viene ampiamente adoperato in campo culturale, un esempio è il catalogo "Gran tour virtuale. Viaggio nel patrimonio" realizzato dal Ministero della Cultura, all'epoca Ministero dei beni e delle attività culturali e del turismo, che raccoglie i progetti che hanno come soggetti i luoghi della cultura ministeriali (https://www.beniculturali.it/virtualtour). Occorre però immaginare che non è sempre stato così: è infatti esempio di una delle tantissime tecnologie nate con diverse funzionalità e poi acquisite dal mondo dei beni culturali.

I virtual tour nascono per scopi commerciali, principalmente per la promozione del settore immobiliare e turistico per poi approdare nella nostra quotidianità attraverso Google Maps. Dopo una lunga fase di sperimentazione, nel 2007, venne introdotto il servizio Google Street View (https://www.google.com/intl/it_it/streetview/). La copertura inizialmente ha riguardato solo 5 città statunitensi per poi allargarsi in breve tempo ad altri centri degli USA e a diverse città canadesi, australiane, giapponesi ed europee per poi estendersi, in breve tempo, alla maggioranza del globo. Un virtual tour si compone di fotografie sferiche scattate a distanza ravvicinata che consentono all'utente di esplorare e visionare le strade, proprio come se si trovasse sul luogo, con una vista completa a 360 gradi.

Dopo che Google ha portato questa tecnologia all'interno delle attività commerciali per accrescerne la visibilità si ha, nel 2013, la prima apertura al settore della cultura con il progetto Google Art Project, poi divenuto Google Arts&Culture (https://artsandculture.google.com/?hl=it), tramite il quale sono stati trasposti in forma di virtual tour ad altissima risoluzione i maggiori musei del mondo.

Ad oggi si contano migliaia di luoghi della cultura digitalizzati da Google, dimostrando così che la loro virtualizzazione non sostituisce la visita fisica ma, al contrario, costituisce un valore aggiunto, una forma di apertura e di superamento delle barriere.

Caratteristiche

Un virtual tour propriamente detto si compone di una sequenza di immagini sferiche. Tali immagini sono dette equirettangolari per il loro particolare dimensionamento: si tratta infatti di immagini rettangolari il cui rapporto base-altezza è di 2:1. Tale proporzione consente la trasposizione della sfera sul piano (https://onix-systems.medium.com/how-to-use-360-equirectangular-panoramas-for-greater-realism-in-games-55fadb0547da; https://www.casa360.net/formato-foto-panoramiche-equirettangolari/). Le immagini sferiche, anche dette in gergo "sfere", sono ordinate in modo tale che la loro sequenza, nel virtual tour, simuli la successione degli spazi nella realtà. Per ottenere tale risultato devono essere adoperati appositi software tra i quali spiccano per notorietà Matterport (https://matterport.com/it), 3D Vista (https://www.3dvista.com/it/) e We Book (https://webobook.com/it/virtual-tour-software.html) e che consentono di realizzare un progetto completo in cui le immagini siano arricchite da ulteriori contenuti, oltre che collegate reciprocamente mediante apposite icone grazie alle quali l'utente finale avrà totale libertà di movimento.

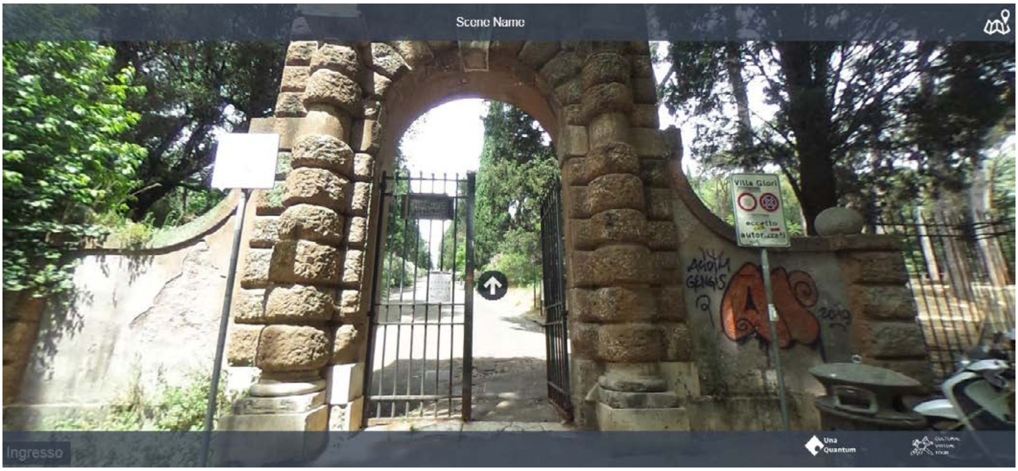

Figure 1. L'immagine presenta una schermata tipo del progetto in modalità desktop*. Si evidenziano tutti gli elementi costituenti:* titlebar *e* footerbar, hot spot *e pulsante di apertura/chiusura della mappa*

Realizzazione

La creazione di un virtual tour si compone di cinque fasi.

Campagna fotografica

La prima fase è quella dell'acquisizione delle immagini sferiche. Allo scopo è stata adoperata una macchina fotografica a 360 gradi, Ricoh Theta S (https://theta360.com/it/about/theta/s.html). Si tratta di uno strumento di semplice utilizzo, commercializzata principalmente per scopi ricreativi, ma che garantisce una qualità di immagine tale da renderla adatta anche per un uso in campo di studio e/o lavorativo.

Il corpo macchina, piccolo e leggero, la rende estremamente maneggevole e può essere utilizzata a distanza grazie all'applicazione apposita per dispositivi mobili.

Lo strumento deve essere posizionato su un sostegno che sia il meno invasivo possibile, preferibilmente su un'asta o monopiede. La macchina infatti è dotata di due obiettivi che, all'atto dello scatto, producono ognuno un'immagine di circa 200 gradi di apertura così da avere un margine di sovrapposizione. Le due immagini vengono unite automaticamente dalla macchina mediante un'operazione detta di *stitching.*

L'immagine finale presenterà, in corrispondenza del nadir, una zona d'ombra dovuta al corpo macchina in corrispondenza del quale si avranno inevitabili deformazioni per l'incapacità dello strumento di compiere una compensazione completa. Dunque, minore è l'impatto del supporto e minori saranno le deformazioni dell'immagine in corrispondenza del nadir.

Postroduzione

La fase di postproduzione è un passaggio importante per la correzione di parametri non controllabili in fase di acquisizione.

Il programma adoperato, GIMP (https://www.gimp.org/) è un software di fotoritocco completo tramite il quale è possibile modificare, integrare e personalizzare le immagini.

Nel caso specifico è stato arricchito dalle funzionalità del plugin G'MIC (GREYC's Magic for Image Computing (https://gmic.eu/). Il plugin si compone di centinaia di filtri tra i quali si individua il filtro "Equirectangular to Nadir-Zenith", fondamentale per la lavorazione delle immagini sferiche poiché permette l'estrapolazione dalle immagini equirettangolari di due porzioni, coincidenti con zenith e nadir appunto, che vengono restituiti su due piani quadrati. L'isolamento del nadir, in particolare, è indispensabile per intervenire sulle deformazioni dovute alla compensazione che viene eseguita automaticamente dallo strumento e che permette di effettuare delle ricostruzioni o di procedere con l'inserimento di un logo.

Creazione delle tiles

Le sfere che costituiscono il virtual tour vengono collegate tramite l'applicativo Pannellum, (https://pannellum.org) un visualizzatore di foto sferiche, gratuito e open source per il Web, realizzato utilizzando HTML5, CSS3, JavaScript e WebGL.

Per comporre il virtual tour tramite l'applicativo Pannellum (https://pannellum.org/documentation/examples/multiresolution/) occorre preliminarmente trasformare le sfere in formato "tiles", ovvero procedere alla frammentazione dell'immagine equirettangolare in un elevato numero di immagini di dimensione inferiore e forma quadrata, che risultano gestibili più facilmente dai motori di ricerca.

Questa operazione non è indispensabile, ma fortemente consigliata per avere una visualizzazione fluida delle sfere in Pannellum, altrimenti denominata "multi-risoluzione".

Ciò richiede la conversione di una immagine equirettangolare nel formato "multi-risoluzione" di Pannellum utilizzando lo script `generate.py`.

Per poter creare tiles in "multi-risoluzione", è necessario avere installato il programma Nona, che è disponibile come parte di Hugin (http://hugin.sourceforge.net/) così come Python3 con i pacchetti Pillow e NumPy.

Creazione del virtual tour

È possibile creare un collegamento tra più sfere in un tour virtuale utilizzando il linguaggio in modo opportuno come dimostrato nella porzione di codice seguente adottata come esempio (https://pannellum.org/documentation/examples/tour/).

```
"scenes": {
        "circle": {
            "title": "Mason Circle",
            "hfov": 110,
            "pitch": -3,
            "yaw": 117,
            "type": "equirectangular",
```

```
            “panorama”: “/images/from-tree.jpg”,
            “hotSpots”: [
                {
                    “pitch”: -2.1,
                    “yaw”: 132.9,
                    “type”: “scene”,
                    “text”: “Spring House or Dairy”,
                    “sceneId”: “house”
                }
            ]
        },
        “house”: {
            “title”: “Spring House or Dairy”,
            “hfov”: 110,
            “yaw”: 5,
            “type”: “equirectangular”,
            “panorama”: “/images/bma-0.jpg”,
            “hotSpots”: [
                {
                    “pitch”: -0.6,
                    “yaw”: 37.1,
                    “type”: “scene”,
                    “text”: “Mason Circle”,
                    “sceneId”: “circle”,
                    “targetYaw”: -23,
                    “targetPitch”: 2
                }
            ]
        }
    }
```

Come è possibile riscontrare, nel codice sono presenti due sfere identificate dalle variabili circle” e “house”, a loro volta contenute in una variabile più generale “scenes” che gestisce tutte le eventuali sfere del virtual tour.

Ogni sfera è caratterizzata da una serie di parametri (https://pannellum.org/documentation/reference/) descritti di seguito.

`titolo` (stringa): Se impostato, il valore viene visualizzato come titolo della sfera. Se non si desidera alcun titolo, non è necessario impostare questo parametro;

`hfov` (numero): Imposta il campo visivo orizzontale iniziale della sfera in gradi. Il valore predefinito è 100;

`pitch` (numero): Imposta la posizione di inclinazione iniziale della sfera in gradi. Il valore predefinito è 0;

`yaw` (numero): Imposta la posizione di imbardata iniziale della sfera in gradi. Il valore predefinito è 0;

`type` (stringa): Specifica il tipo di hotspot;

`panorama` (stringa): Imposta l'URL dell'immagine equirettangolare.

Per ogni sfera posso esserci più hotspot che consentono il collegamento ad altre sfere a partire da essa. Gli hotspot sono caratterizzati dai seguenti parametri.

`pitch` (numero): Specifica la porzione dell'hotspot, in gradi.

`yaw` (numero): Specifica l'imbardata della posizione dell'hotspot, in gradi.

`type` (stringa): Specifica il tipo di hotspot.

`text` (stringa): Questo specifica il testo che viene visualizzato quando l'utente passa sopra l'hotspot.

`sceneId` (stringa): Specifica l'ID della scena cui l'hotspot si collega.

Inserimento della mappa

Nel virtual tour è stata integrata una mappa interattiva. Tale strumento consente di arricchire l'esperienza dell'utente durante la fruizione della visita virtuale.

La mappa è realizzata in LeafletJS, (https://leafletjs.com/) una libreria JavaScript open source per mappe interattive ottimizzate per dispositivi mobili.

Figure 2. Lo screenshot *mostra il progetto in conseguenza all'apertura della mappa. Punti salienti sono i pulsanti per aumentare e diminuire il livello di zoom e i pulsanti, rossi e numerati, per il collegamento guidato alle altre immagini*

Tale contenuto può essere derivato da un'elaborazione CAD, successivamente trasformata in formato vettoriale GIS. Per l'inserimento di tale elemento all'interno di LeafletJS è necessario infine trasformare il vettoriale GIS nel formato GeoJSON, particolarmente indicato per applicazioni web.

Il virtual tour: il caso di Villa Glori

La metodologia di lavoro adoperata è applicabile su larga scala e può essere utilizzata indifferentemente su qualunque sito culturale.

Possono essere individuate, nel campo dei beni culturali, le più disparate funzionalità e utilità.

Nel settore più specifico dell'archeologia il virtual tour può diventare un utile strumento di accompagnamento ai metodi di rilievo e documentazione grafica e fotografica più tradizionali dal momento che l'immagine sferica permette di avere, con un unico scatto, la visione integrale dell'area ripresa.

Il virtual tour realizzato si presenta con una schermata occupata quasi esclusivamente dalla foto sferica, alla cui apertura, essendo attiva la modalità "autorotazione", inizia appunto un movimento automatico che deve invitare il fruitore ad assumere il controllo della visita.

La schermata in alto è delimitata dalla "titlebar", una sottile barra destinata ad ospitare il nome identificativo dell'immagine che si sta visualizzando; all'estremità destra è poi ospitato il pulsante per l'apertura e la chiusura della mappa interattiva. Nella parte bassa, la "footer bar", una barra gemella a quella superiore, è utile per l'inserimento di loghi dei soggetti responsabili della realizzazione del virtual tour e, al contempo, si consente l'apertura dei rispettivi siti internet.

All'interno di ogni immagine si individuano delle frecce che funzionano da "link hotspot", ossia dei pulsanti per il collegamento reciproco delle sfere che vengono così ordinate simulando il percorso di visita. Al passaggio del mouse compare un piccolo popup, il "tooltip", che contiene il nome della sfera collegata.

L'elemento saliente del progetto è indubbiamente la già citata mappa interattiva la cui apertura è possibile grazie al pulsante posto in alto a destra. Nella rappresentazione schematica della planimetria del sito i punti, visibili tramite immagine sferica, sono identificati da un numero: anche questi sono dei pulsanti che consentono il passaggio da un'immagine all'altra, alternativi ai "link hotspot". La mappa può essere, all'occorrenza, leggermente ingrandita o rimpicciolita.

Il virtual tour è stato progettato in modo tale da consentire la perfetta adattabilità su qualsiasi tipo di dispositivo. Nel momento in cui viene visualizzato su apparecchi mobili, quali smartphone e tablet, gli elementi che incorniciano le immagini sferiche aumentano leggermente di dimensione per facilitare la digitazione; la mappa invece, a seguito della sua apertura, viene visualizzata a schermo intero così da agevolarne la fruizione.

La sperimentazione di Pannellum per la realizzazione di virtual tour è stata condotta su fotografie sferiche realizzate nel parco romano di Villa Glori.

La scelta è stata dettata dalla necessità di individuare uno spazio liberamente accessibile nel maggio 2020, periodo non facile per le numerose chiusure e limitazioni dettate dalla pandemia COVID- 19. Gli autori avrebbero preferito lavorare su materiale prodotto in ambito archeologico, nel contesto di una campagna di scavo ma, per le ragioni sopra indicate, ciò non è stato possibile.

Il Parco è uno spazio di notevole importanza per il suo stretto legame con la storia, è stato infatti sede di battaglie durante il Risorgimento. Attualmente è un luogo della memoria, da qui la sua denominazione di Parco delle Rimembranze, ma anche sede di una collezione di arte contemporanea a cielo aperto che conta molteplici installazioni *site specific* oltre che di alcuni siti archeologici.

I punti di scatto sono stati selezionati per dare un'immagine completa della Villa oltre che per evidenziare i punti maggiormente significativi quali quelli coincidenti con le installazioni e i monumenti a memoria delle battaglie e dei caduti. Per quanto ampi i vari settori sono stati immortalati da un'unica fotografia per evitare ridondanze.

Figure 3. La figura riprende la Figura 2 ma visualizzata in modalità mobile*. L'immagine permette di evidenziare come la mappa vada ad occupare interamente lo schermo al fine di facilitare la sua fruibilità tattile su uno schermo di piccole dimensioni*

La realizzazione della campagna si è protratta per qualche ora, date le difficoltà nel posizionare correttamente la strumentazione nel terreno irregolare, cui si sono aggiunti tempi di attesa dovuti alla necessità di acquisire le immagini in assenza di persone, sia per questioni legate alla privacy che per rendere le fotografie più leggibili.

La fase di acquisizione del dato fotografico è l'unica che viene svolta direttamente *in situ*, le fasi successive, di pura elaborazione del dato fotografico, vengono portate avanti lontano dal sito oggetto.

Conclusioni

Allo stato attuale il virtual tour non comprende contenuti aggiuntivi fatta eccezione per la mappa interattiva, la cui presenza si ritiene un buon risultato dal momento che altre applicazioni consentono, di norma, l'inserimento di una sola mappa statica.

La mappa dinamica consente una migliore navigazione e la possibilità di riportare una maggiore e più completa quantità di informazioni del sito culturale.

Si sottolinea inoltre la scelta di LeafletJS per la realizzazione della mappa interattiva, applicativo ampiamente utilizzato nell'ambito della cartografia digitale su web. Contrariamente sarebbe stato necessario l'uso di altri applicativi cartografici meno diffusi o la realizzazione di un applicativo *from-scratch*. Tali ultime due scelte avrebbero vanificato gli indubbi vantaggi delle funzioni di LeafletJS, ormai riconosciute dalla comunità di sviluppatori e utilizzatori in ambito di applicazioni cartografiche online.

Il progetto è però facilmente implementabile con ulteriori pulsanti, detti "info hotspot" da collocare in corrispondenza di punti salienti delle immagini sferiche e che permetterebbero l'apertura di popup informativi all'interno del quale porre ulteriori contenuti, anche multimediali.

Nel settore archeologico una campagna fotografica cadenzata nel tempo e strutturata in diversi virtual tour, rappresentativi ognuno di una specifica fase di scavo, permettono di avere un'ulteriore memoria del procedere dell'attività di studio da affiancare agli altri strumenti di documentazione normalmente adoperati. Diviene inoltre un utile strumento di divulgazione delle conoscenze per i "non addetti ai lavori".

Figure 4. L'immagine propone una visualizzazione del progetto in modalità mobile *a seguito della chiusura della mappa. Lo schermo è interamente occupato dall'immagine sferica al fine di favorire la sua navigazione. Fanno da cornice, anche nella modalità di visualizzazione per schermi di dimensioni ridotte,* titlebar *e* footerbar

Grazie alle caratteristiche che il lavoro assumerà mediante le implementazioni future che puntano all'integrazione delle immagini con popup informativi, renderanno lo strumento utile sia per la divulgazione delle informazioni ma anche, e soprattutto, per la documentazione della campagna di scavo. All'interno dei popup possono essere visualizzate, o anche richiamate nella forma di contenuto esterno, tutti i documenti prodotti durante le attività: schede, documentazione fotografica, rilievi tradizionali e laser per fare alcuni esempi. Un sapiente posizionamento dei popup consentirebbe di mantenere memoria, inoltre, dell'esatta collocazione dei reperti rinvenuti in fase di scavo creando un collegamento con oggetti che, dopo lo scavo, perdono inevitabilmente la loro posizione originaria.

Bibliografia

Albrizio, P., F. De Virgilio, G. Panzarino, and E. Zambetta. 2013. "WebGIS e divulgazione del dato archeologico con software open source. Il progetto "Siponto aperta". In *Proceedings of ArcheoFOSS VIII: Free, libre and open source software e open format nei processi di ricerca archeologica*, edited by Filippo Stanco and Giovanni Gallo, 101-114. Oxford: Archeaopress.

Communication Strategies Lab, 2012. "Realtà aumentate. Esperienze, strategie e contenuti per l'Augmented Reality". Milano: Apogeo.

Conti, F. 2018. "Tecniche e strumenti software per la produzione di visite virtuali". Torino. https://webthesis.biblio.polito.it/8147/1/tesi.pdf

De Luca, V., C. Conati Barbaro, C. La Marca, M.L. Savino, and P. Rosati. 2019. "L'open source per i musei: il tour virtuale del Museo delle Origini (Sapienza Università di Roma)" in *Archeologia e Calcolatori*, 30: 479-482.

Orlandi, M., Zambruno, S., Vazzana, A. 2014. "Tecnologia, Beni Culturali e Turismo: i Tour Virtuali (Virtual Tours) come strumento per una corretta comunicazione dei Beni Culturali" in Storia e Futuro, *Rivista di Storia e Storiografia Contemporanea online*, 34.

Kennicott, P. 2011. "National Treasures: Google Art Project Unlocks Riches of World's Galleries" Washington: The Washington Post.

Petroff, M. 2019. "Pannellum: a lightweight web-based panorama viewer"in *Journal of Open Source Software,* 4 (40): 1628. https://doi.org/10.21105/joss.01628.

Waters, F. 2011. "The Best Online Culture Archives". Londra: The Telegraph.

The Roman Villa of l'Albir (Alicante, Spain). The use of Blender and the Extended Matrix in the virtual reconstruction

Jaime Molina[1], Javier Esclapés[1], Carolina Frías[3], Javier Munoz[1], Alejandro Martín[1], Mónica Sánchez[1], Laia Fabregat[1] and Daniel Tejerina [1*]

[1] Patrimonio Virtual, University Institute of research in archaeology and heritage, University of Alicante.

[2, 3] Museo al Aire Libre Villa Romana de l'Albir

***** Correspondence: daniel.tejerina@ua.es

Presented at ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract: In this paper we describe the workflow followed to produce a 360-degree digital short film to virtually recreate the Roman Villa of L'albir (L'alfàs del Pi, Spain), from the preliminary steps and production stages, to the use of the Extended Matrix tool to document the entire virtual reconstruction and recreation process.

Keywords: Archaeology; Virtual Reconstruction; Digital Photogrammetry; Virtual Reality; 360-degrees; Blender; Extended Matrix; Meshroom.

FOSS software used and licence: Blender, GNU General Public Licence v3.0; Extended Matrix Blender Tools, GNU General Public Licence v3.0; Meshroom, Mozilla Public Licence v2.0.

Introduction

The Villa Romana de l'Albir is a Late Imperial *vicus* (4th-7th centuries AD) located on the coast of l'Albir and discovered after several archaeological seasons carried out in the 80s and 90s of the last century. This important heritage site includes an extensive *necropolis*, a private *mausoleum* and the villa where the *vicus* administrator lived (Morote and Fernandez Rojo 1989; Chavarria, Brogiolo and Arce 2006, Chavarria 2007).

Enhancement of the thermal baths, undertaken between 2008 and 2010, led to its designation as the Villa Romana de l'Albir Open Air Museum by the regional government. Once the intervention in this area (sector I) had been completed, work focused on the villa itself, having so far located the villa's *pars urbana*, a *necropolis* dating from the 6th century AD, a Visigothic stage of occupation (7th century) with working areas and garbage dumps and a final, late mediaeval (14th century) stage of occupation with productive structures and large garbage dumps (Frías Castillejo 2010).

Based on the results obtained throughout the different research seasons and with the aim of promoting and disseminating this heritage site, a 360-degrees, short animated film has been produced, to show visitors a virtual recreation of the village through an immersive experience.

As can be seen in the following diagram, the process to create this short film has been divided into two main stages (pre-production and production) which will be described throughout this paper, with special emphasis on the open source tools used in the project: Meshroom, Blender and the Extended Matrix.

Pre-production

During this preliminary stage it is important to define the initial idea, which should include aspects such as: basic plot, number of characters, year in which the story will take place, number of scenes and total length. Once these details were defined, three basic working documents were created: the script, the shooting script and the storyboard. Together, these three documents should be detailed enough to provide all the information that will be needed during the production process.

Initial idea, story, plot

The plot centres around a married couple from the countryside of Alicante who visit the Villa of l'Albir - located on the seaside- to close trade agreements. Their hosts - the villa owners - receive them and throughout the three scenes they go through the different rooms of the villa while talking to each other. This plot has allowed us to introduce aspects of daily life in a rural settlement in the 5th century AD, but also others related to local and regional trade, or references to the political, social and even religious situation of the time.

The story is set in the year 412 AD and is divided into 3 scenes lasting 10 minutes, and structured as follows:

Scene 1. Reception of guests in the *peristylium*. 3 minutes.

Scene 2. Tour through the different rooms of the *balneum*. 3 minutes.

Scene 3. Dinner in the *oecus*. 3 minutes.

Script

Once the plot was defined, the next task was writing the script, which was divided into scenes and sequences, each of which was preceded by a brief introduction to then write the dialogue for every character, describing also the feelings of each of them through their body language and facial expressions.

Shooting script

The shooting script is a text file that defines technical aspects that do not appear in the script, complementing it. It describes in detail every action that will take place in the scene, the elapsed time (in seconds) and the exact position of each character, asset or camera throughout the scene. Together with the script, the shooting script will be essential for the CGI (Computer Generated Imagery) department during the virtual production stage, since all the camera and character animation work will be based on it.

Storyboard

The storyboard consisted in a set of sequenced and ordered illustrations, used as a visual guide to the story, to preview an animation but also to structure the film before it is made. It is the pre-production mode used in the film industry.

Its main purpose is to tell the story with a few sequenced images, including some details about the film shot number, its length and a brief description of what happens in it. All this information will serve as a guide once in the production stage. It is not essential, like the script or the shooting script, but it is still a very useful tool and constitutes the first visual preview (although not very detailed at the moment) of the short film itself.

Production

As with any film project, this stage is the most time and resource-intensive. It involves not only the elaboration of the visual and sound contents of the short film, but also the development of an app that will allow users to enjoy the whole experience.

Sources

In this type of project, candidates for virtual reconstruction hypotheses are always preceded by collecting and researching the available data sources. In the specific case of the Roman Villa of l'Albir, these can be classified according to their nature and their direct or indirect relationship with the archaeological site itself: 3D survey, archaeological reports, and other sources.

3D survey

The preserved archaeological remains were surveyed using digital photogrammetry, following a well-known methodology (Historic England 2017), widely used and fully integrated in today's archaeological fieldwork. The input usually consists of structured sets of photographs with high overlap between them, together with the acquisition of spatial data through the use of appropriate surveying instruments, such as total stations and satellite positioning data. The output is an oriented, textured 3D model. To process the images we used Meshroom, an open source software which offers a simple but powerful nodal approach (Alice Vision, n.d.).

As a result, a textured 3D model of each of the different areas of the archaeological site was obtained, ready to be used in any 3D modelling software.

Archaeological reports

Archaeological excavation is a destructive process, so detailed, accurate reports generated during each excavation season are an essential data source for virtual reconstructions. Both written descriptions and graphic documentation (technical drawings and photographs) will be useful data sources.

Other sources

Where archaeological reports and 3D photogrammetry models are not enough, other data sources are needed, such as other references or, in this case, other Roman villas. This, although valid, must be done with special attention and preferably looking for other similar examples based on criteria of contemporaneity and geographical area. Also, taking into account that they may have undergone alterations, such as modern restorations, not always documented or easily identifiable.

The same procedure can be used in the case of written sources or artistic representations. In virtual reconstructions of Roman architecture and artefacts, wall paintings and mosaics are particularly useful as a source of information.

Hypothesis

From the study of the different sources of information, several hypotheses were elaborated, tested and validated until one of the candidate hypotheses was selected as the working hypothesis.

Ideally, validating and selecting the working hypothesis requires the participation of as many specialists as possible and from different fields, not only Archaeology, but also History, Art History or Architecture, to ensure that the selected hypothesis is technically and aesthetically coherent with the available scientific information. Therefore, each part must be analysed and tested according to a scientific methodology.

Preliminary sketches

At this point and from the available data, several sets of preliminary drawings and views were created at a minimum level of detail to show, compare and validate the different proposals.

Technical drawings

Once the proposal has been selected, a set of highly detailed, technical drawings (plans, sections and elevations) was generated, based both on the photogrammetric 3D model and the preliminary drawings. Since it is possible to design the architecture from zero in the 3D modelling software, this step is not essential. However, we consider that this intermediate step eases the subsequent 3D modelling work, while at the same time providing useful extra information about the Villa's architecture.

Virtual production

Within the production stage, virtual production refers specifically to all computer-generated elements, including architecture, natural and artificial objects, characters and animations that will be shown in the short film (all of which are commonly referred to as set, assets and characters).

Architecture and other assets

Modelling. Technical drawings —plans, sections and elevations— were imported into Blender and placed in their correct position. Then, 3D volumes were generated for each entity (floors, walls, ceilings and other architectural elements), subdividing and extruding the original geometric mesh, adapting it to the limits previously defined in the technical drawings and using the snap tools to ensure modelling accuracy.

Other types of assets, such as sumptuary objects and jewellery, were also created with a similar approach, in this case using archaeological drawings as references. But because of its higher complexity, the modelling process was in some cases slightly different and required the use of other modelling tools and techniques, such as bèzier curves and digital sculpting. In order to optimise the process —keeping the number of polygons in the scenes as low as possible while showing a high level of detail—, two copies of the same object were created: the first one, more detailed, with a higher number of polygons and a second, lighter copy.

At this point it is particularly important to mention the importance of using a proper nomenclature for each collection and object within Blender, which is especially useful during the subsequent stage of creating and applying textures. To facilitate this task we used the suffixes _HIGH and _LOW respectively.

UV unwrapping and texturing. Blender's ability to handle UDIM textures, which allow working with different textures and resolutions within the same uv map, helped to optimise each of the 3D models, drastically reducing render times. In order to do that, the basic workflow involved dividing the UV map into different parts, assigning each part a specific resolution. Once the UVs were generated, the models were exported in FBX format to then work with them in the texturing software (Adobe Substance Designer).

Baking. Low resolution models were imported into Substance Painter using custom settings to be able to work with the UDIM textures. Prior to texture painting, the first step was baking, that is projecting the high-resolution model features onto the low-resolution model in order to achieve as much information as possible while using the optimised (lighter) low poly model.

Texture painting. The use of PBR (Physically based rendering) maps is essential at this point, as they accurately simulate a realistic look of materials and textures, as well as the behaviour of light over the object's surface. Priority was given to photorealism, creating subtle superficial imperfections by applying base materials or textures and superimposing several layers with Substance Designer tools (Smart Materials and Smart Masks). After adjusting each of these layers settings, different brushes were used to manually create small details and imperfections across the model's surface.

Once this process was finished, textures were exported at the required resolution to finally import them into Blender, where they would continue working with the pieces in terms of distribution, lighting and final rendering within the scene. Thus, again in Blender, we worked with the Shader Editor and a specific system of nodes that allows working with PBR textures through a unique system known as Principled BSDF, which combines multiple layers or configuration of materials (fig. 1).

Figure 1. 3D modelling. (a) solid mesh; (b) uv unwrap; (c) textured mesh.

Characters

This stage has been mainly done with commercial software and therefore we will limit to briefly indicate the workflow followed in this project to create, clothe and animate the virtual characters.

Modelling. A total of 8 characters have been created with Character Creator, according to the detailed description (age, height, ethnic features or body mass volume, among others) previously provided for each of them during the pre-production stage.

Animation, editing and retargeting. To perform the different characters' actions, full body Motion Capture systems were used (capable of recording the movement of the body, hands and face), and Rokoko Studio was used for data management. After editing and optimising the generated animations, the armatures and their corresponding actions were exported in FBX format. Once in Blender, the actions and movements were assigned to their corresponding characters in a process known as retargeting.

Clothing. Although Blender has a clothing solver able to accurately simulate textile material behaviour, it is not particularly efficient at simulating complex interactions between clothing and an animated mesh, often generating artefacts and unrealistic results. Therefore, all the character's clothing was created using Marvelous Designer. This process included two stages: designing patterns and materials for the clothes and simulating their physical behaviour on the character's animated mesh (fig. 2).

Figure 2. Motion capture. (a) dressed virtual character in T pose; (b) virtual acting test.

Integration, lightning, previsualization and rendering

Once the different assets (architecture and objects) and animated characters were generated, they were distributed in 3 scenes within Blender, following the indications described in the shooting script and storyboard generated during the pre-production stage. From this point on, the modelling and animation work was completed and it was lighting that was going to play the main role. At this point it is fundamental to consider which lighting strategy is the most appropriate and the one that best suits the initially proposed ambience. In our case, the scenario is at midday, so only natural light was used. In any case, and in projects of this scale where it will take many days to generate the final images, rendering the images at low resolution to preview the general appearance of the scene is key to avoid wasting resources later on.

The different elements of the scene were distributed in layers to allow further compositing.

The last step is rendering the scenes, exporting each frame in individual EXR image files, at a 8K resolution and setting a high number of samples (the unit by which Blender measures the noise filtering in the final image, based mainly on the number of light's bounces on each objects' surface) and the denoise tool, which drastically reduces rendering times. This tool is very useful with static meshes, but when it comes to animated meshes, it must be combined with a high sampling value to avoid artefacts and undesired results.

Colour correction, colour grading, video and audio editing

Once the rendering process is finished, the result is basically a sequence of images. The next step is to compose the whole sequence, modifying aspects related to light and colour to give the final footage the desired look. Finally, the audio tracks with voices and ambient sound were added and synchronised in the final video (one for each language), ready to be integrated in the 360-degrees app.

Extended Matrix

Parallel to the development of the work described so far, another task has been of particular importance: recording and tracking the entire decision-making process, which is key in projects involving heritage virtual reconstruction. The tool used to carry out this process is the Extended Matrix, developed by Emanuel Demetrescu (CNR - ISPC). "*The EM allows to record the sources used and the processes of analysis and synthesis that have led from scientific evidence to virtual reconstruction. It organises the 3D archaeological record to make the 3D modelling steps smoother, more transparent and scientifically complete*" (Demetrescu 2022a http://osiris.itabc.cnr.it/extendedmatrix/).

Basic workflow

The workflow with the Extended Matrix has been widely described by its author in the resources available on the official web site (Demetrescu 2022b http://osiris.itabc.cnr.it/extendedmatrix/index.php/learn-em/; 2022c http://osiris.itabc.cnr.it/extendedmatrix/index.php/nodes-of-the-em/) and in scientific publications (Demetrescu 2015, Demetrescu 2021), so we will briefly describe the steps that we have followed in applying this tool to the virtual reconstruction of the Roman Villa of l'Albir.

Graph editor. As its name indicates, this tool is based on the same approach as the Harris Matrix commonly used in modern field archaeology. It is therefore an ordered sequence of units. So the first step consisted in the elaboration of an extended matrix with a graph editor capable of exporting files in XML. A set of nodes and connectors specifically designed to work within the Extended Matrix can be easily imported in the graph editor.

Blender & EMTools. One of the many advantages of this innovative tool is its seamless integration into Blender as a plug-in. Once the extended matrix is completed, the resulting graph can be imported into blender using the EMTools plugin. Each of the different units is automatically recognised in Blender, so we could work within them to create proxies of the different entities according to their features and properties.

The application of the Extended Matrix in the virtual characters' recreation

While elaborating the extended matrix for this project, we faced the issue of how to apply it to virtual characters, since each of them has different, source - based attributes (physical characteristics, hairstyle, clothing, jewellery). Even the arrangement of each participant at the *stibadium* of the *oecus* followed a strict protocol in a culture as ritualised as that of Rome and well documented in historical references. So the same procedure was applied to the virtual characters together with their attributes and positions at the *stibadium* (fig. 3).

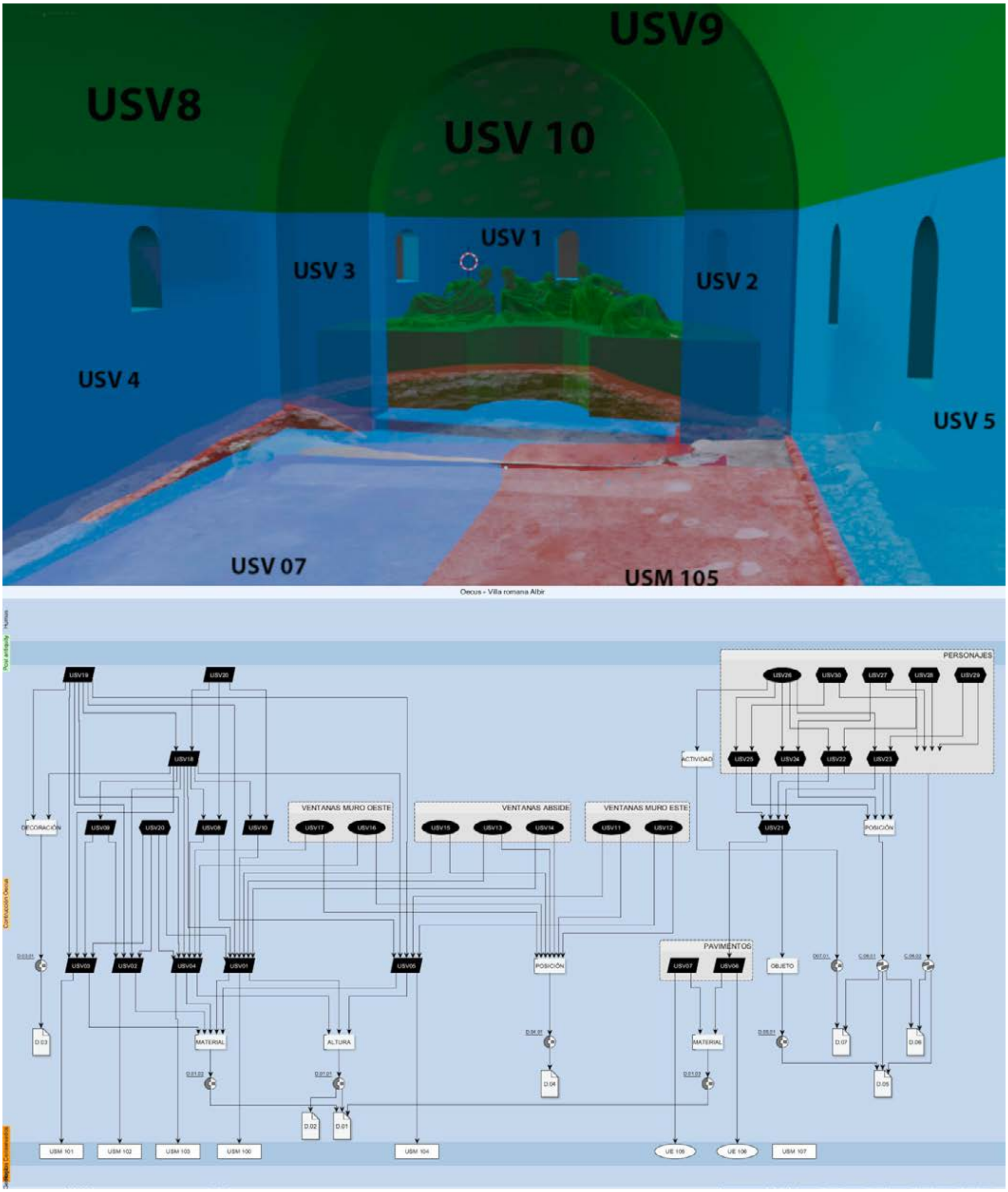

Figure 3. (a) Extended Matrix proxies from the oecus*; (b) Extended Matrix graph from the* oecus*.*

App design

Being panoramic, this animated short film is designed to be displayed in virtual reality devices, so it was necessary to design and develop an app that met the following requirements:

Allow synchronisation between the different devices in the room and from another, master device (in this case, an Ipad), in order to efficiently handle groups of people.

To be available in 3 languages: Spanish, Valencian and English.

Graphic design and User Interface

UI has been designed to be simple and easy to follow, which is common to similar immersive projects where the interface and the input device is not a computer screen and a keyboard, but a headset device that fully covers our field of vision. To show users how to proceed, one of the screens is a simple video tutorial explaining what a 360-degrees panoramic image is, how it works and how users can interact with it.

Software development

The system consists of an ipad, 8 Oculus Quest 2 VR headset devices and a router, which is used to deploy a private local network to which the computer and VR devices are connected.

The application is divided into two versions, administrator and user.

Administrator. Is the one who manages every session, creating new ones, controlling the status of each device, launching the video or closing the session.

In the admin mode, and after the start screen, the device connects to the private local network previously deployed and creates a new session, to then proceed to the standby screen, showing a list of every virtual reality device and its current status (that is, the screen that each user is on at that moment: connected/disconnected, language selection, tutorial, or waiting for the rest of participants).

When everyone is on the standby screen, the administrator launches the content and automatically every headset device starts displaying the video. Once it has finished, the app returns to the home screen.

User. Once users have put on the VR device, they will enter the home screen, where they will remain while the device detects a new session created by the administrator and joins it. When users connect to the session, they are taken to the language selection window so that they can choose the language (English, Valencian and Spanish). In addition and for those with hearing disabilities, voices and sounds will be described using text if the accessible content mode is on.

The next screen the user goes to once the language has been chosen is a brief tutorial, where the content to be displayed is introduced and explained.

*Figure 4. (a) scene 01,*peristylum*; (b) scene 02,* balneum*; (c) scene 03,* oecus.

Once the tutorial is finished, users will enter the standby screen until the rest of participants have joined the session, chosen the language and completed the tutorial. When all users are logged in and the administrator has executed the start playback command, the video will start playing in parallel on all VR headsets.

At the end of the video, credits are shown and the VR headsets automatically return to the standby screen.

Results

The output of the project described is an animated, 360-degrees short film, approximately 10 minutes long, which takes the owners and their guests through the different rooms of a late Roman villa, showing visitors aspects of Roman daily life, but also about the situation of the Roman Empire in the 5th century AD.

These contents, ready to be visualised on virtual reality devices and in different languages, are currently available at Villa Romana de l’Albir Open Air Museum (Alicante, Spain) (fig. 4).

Conclusions

Making an animated, immersive short film focused on virtual reconstruction and recreation of the cultural heritage is a challenging and complex creative process that requires specialised human, technological and financial resources. We believe that it is a unique and engaging product, able to provide visitors with a singular experience to gain knowledge about history and archaeology.

In order to guarantee a solid scientific basis for the contents shown, it is essential to justify and record each and every decision taken along the process,so that it can be accessed and evaluated. In this regard, the Extended Matrix tool represents a leap forward to efficiently manage the large volumes of data generated during a process of this kind.

In projects involving virtual reconstructions of cultural heritage, the focus is usually on architecture. However, virtual characters are also powerful knowledge and information providers. Therefore, aspects such as their physical appearance, attributes and actions should also be equally researched and documented. And in this sense, we believe that the Extended Matrix is also capable of handling this type of data.

References

Alice Vision. n.d. “Alice Vision. Photogrammetric Computer Vision Framework.” Accessed October 23, 2022. https://alicevision.org/#

Aparicio, P. 2021. “Guía para la documentación de reconstrucciones virtuales de bienes patrimoniales.” in *Revista Schema,* 2: 129-146.

Chavarria, A., G. P. Brogiolo, and J. Arce (eds). 2006. *Villas en Hispania durante la Antigüedad Tardía.* Madrid: CSIC.Chavarria, A. 2007. *El final de las villae en Hispania (siglos IV-VII D. C.).* Paris: Brepols Publishers.

Demetrescu, E. 2015. "Archaeological Stratigraphy as a formal language for virtual reconstruction. Theory and practice." in *Journal of Archaeological Science,* 57: 42–55. https://doi.org/10.1016/j.jas.2015.02.004.

Demetrescu, E. and D. Ferdani. 2021. "From Field Archaeology to Virtual Reconstruction: A Five Steps Method Using the Extended Matrix." in *Applied Sciences,* 11 (11). https://doi.org/10.3390/app11115206.

Frías Castillejo, C. 2010. *El poblamiento rural de Dianium, Lucentum, Ilici y la ciudad romana de La Vila Joiosa (siglos II a. C.-VII d. C.). Bases para su estudio.* Alicante: Universitat d' Alacant (Publicacions).

Historic England. 2017. *Photogrammetric Applications for Cultural Heritage. Guidance for Good Practice*. Swindon: Historic England.

Morote, J. G., and A.J. Fernández Rojo. 1989. "Resumen de los trabajos realizados en los años 1988-1989." in *Revista de Fiestas Patronales de l'Alfàs del Pi*: 50-55.

Webliography

Demetrescu, E. 2022a. "Extended Matrix in a nutshell" Accessed January 16, 2022. http://osiris.itabc.cnr.it/extendedmatrix/

Demetrescu, E. 2022b. "Learning EM" Accessed January 16, 2022. http://osiris.itabc.cnr.it/extendedmatrix/index.php/learn-em/

Demetrescu, E. 2022c. "Nodes of the EM" Accessed January 16, 2022. http://osiris.itabc.cnr.it/extendedmatrix/index.php/nodes-of-the-em/

Combination of standardised methods to create a detailed source-based reconstruction of the Tepidarium at L'Alcudia de Elche, Alicante, Spain

Cristina Gonzalez-Esteban[1]*

[1] Department of Archaeology, University of Southampton.
*Correspondence: cristina.ge14@gmail.com

Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; online; November 23rd-26th, 2021.

Abstract: This study proposes a combined method for Virtual Archaeological Reconstructions (Extended Matrix, Reconstructive Units and Evidence Graphic Scale) to ease understanding, recording and display of archaeology. It aims to create new questions and challenges to increase knowledge. The *Tepidarium* project at L'Alcudia de Elche, Spain, showed the need to reduce duplicity and quicken the workflow; however, it increases the possibility of comparative studies due to similar results in other projects.

Keywords: FLOS Software; Archaeology; Source-based Virtual Reconstructions [V-R]; L'Alcudia de Elche; Roman *Tepidarium;*

FOSS software used and license:
1. Blender, GNU Public License (GPL) Version 2.83 and 2.90;
2. yEd, Freeware Version 3.20.1;
3. LibreOffice, MPLv2.0 Version 7.1.5;
4. Inkscape, GNU Public License (GPL) Version 1.0.1.

Introduction

This project started as a MA Virtual Heritage study (2019) when Prof. Molina and Prof. Muñoz (University of Alicante) shared data from the site of L'Alcudia de Elche, Spain, to create a source-based Virtual Reconstruction [V-R] of the Eastern *Thermae Tepidarium*, testing different techniques of metadata and paradata standardisation (fig. 1).

The main objective was to create an exhaustive and easy-to-read 3D archive of the site and its reconstruction process. This would boost Virtual Archaeology [VA] as a scientific discipline and promote V-R as new tools characterised by historical rigour and scientific transparency. They could help further archaeological research and bring accurate and interactive narratives of the past to the public.

This article innovates through the combinations of three standardisation techniques: Extended Matrix [EM] by Dr Demetrescu (CNR); Reconstructive Units [RU] from Prof. Molina and Prof. Muñoz; and Graphic Scale of Historic-Archaeological Evidence [GS] by Aparicio and Figueiredo. Diverse projects have used these techniques, as can be seen in Table 1.

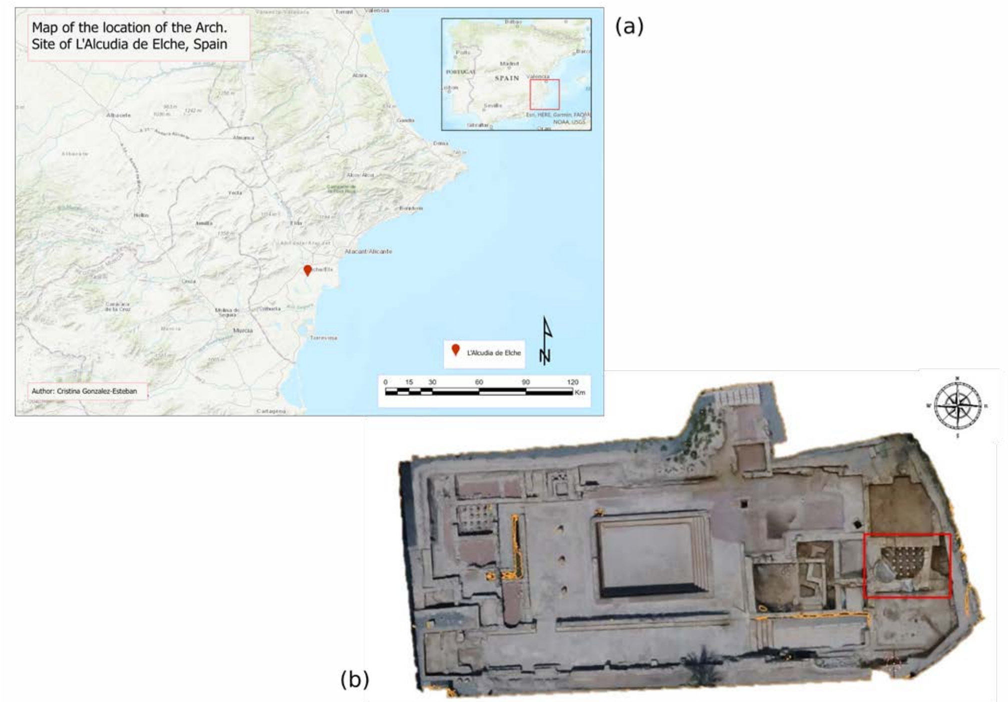

Figure 1. Location of the study site: (a) Map of Eastern Spain showing the location of the site of L'Alcudia of Elche; (b) Zoomed view of the Eastern Thermae *of the site of L'Alcudia of Elche with a square over the* Tepidarium. *(Images: Cristina Gonzalez-Esteban)*

Nevertheless, no one had combined the three of them. This unification is the main significance of this project, pushing to combine styles of virtual standardisation to cover different needs and show how important it is to integrate upcoming perspectives. This avoids diversifying standards for virtual reconstructions, which would hinder the core purpose of creating a standardisation system: comparable, compatible and reproducible data.

Standardisation techniques for Virtual Archaeological Reconstructions

The 20th century came with advances in technology coupled with increasing usage of computers and data management, which helped boost growth in recordings and displays of archaeological remains (reality-based models/Digital Twins). These new tools gave birth to the new subdisciplines of Virtual Archaeology (Reilly 1990) and the later CyberArchaeology (Forte 2010) to encompass the possibilities and potential of new technologies for historic and archaeological studies. However, the purpose of archaeology is to interpret and learn of past societies through their material culture, which does not get fully covered by these digital documentations, hence the need of promoting representations of archaeological interpretative components (i.e. Virtual Reconstructions [V-R]).

Nonetheless, these models are developing much slower, especially hindered by the fact that hypotheses are based on the researcher's analysis (subjective). Therefore, the accuracy of reconstructions cannot be objectively measured, as is the case of reality-based models,

Table 1: Examples of published or conference-presented projects that have reported the usage of some kind of metadata recording methodology for their reconstructions. Note that this is a partial list of projects, just some examples.Table: Cristina Gonzalez-Esteban.

Project	Reconstructive Units (Molina Vidal and Muñoz Ojeda 2015)	Graphic Scale (Aparicio Resco and Figueiredo 2016)	Extended Matrix (Demetrescu and Ferdani 2021)	References
Tepidarium of the Eastern Thermae of L'Alcudia de Elche	YES	YES	YES	Gonzalez-Esteban 2019
Roman Villa of Aiano (San Gimignano, Italy)			YES	Demetrescu and Ferdani 2021
Roman Villa at L'Albir (Alicante, Spain)	YES		YES	*In prep*
Roman Villa out of Porta Marina (Ostia, Italy)			YES	Demetrescu and Ferdani 2021
Temple A of Illeta dels Banyets (El Campello, Alicante, Spain)		YES		Aparicio Resco and Figueiredo 2016
Imperial Forum of Augustus (Rome, Italy)			YES	Demetrescu and Ferdani 2021
German WW2 bunker H669 (*no location*)	YES	YES		Aparicio Resco and Figueiredo 2016
Onde Marine, Necropolis of Banditaccia (Cerveteri, Italy)			YES	Demetrescu and Ferdani 2021
1st Phase of the Santiago Domus (Bracara Augusta, Braga, Portugal)		YES		Aparicio Resco and Figueiredo 2016
Amba Aradam (Rome, Italy)			YES	Demetrescu *et al.* 2021
Door-Tower of Bejanque (Guadalajara, Spain)	YES	YES		Aparicio Resco and Figueiredo 2016
Great Temple (Sarmizegetusa, Romania)			YES	Demetrescu and Ferdani 2021
Craft Building (Montebelluna, Italy)			YES	Demetrescu and Ferdani 2021
Amphitheatre of Cartagena (Murcia, Spain)			YES	Garcia 2021

whose level of detail is represented by a metric value. This brought the issue of downgrading virtual reconstructions to graphic design products since their historical rigour and scientific background could not be checked.

In order to overcome this obstacle, a series of Charters were written defining new concepts and requisites of these new Archaeological tools: the 2009 London Charter (Denard 2009) and the 2011 Seville Principles (IFAV 2012). From their principles, the most important ones for this research are “Documentation”, “Authenticity”, “Scientific Transparency”, and “Historical rigour”. The above-mentioned experts have attempted to fulfil these obstacles of Archaeological Virtual Reconstructions, leading to the creation of source-based V-R to include metadata (sources) and paradata (thoughts) needed for creating interpretive models from reality-based ones possible to validate and reuse.

This is the line of research that this article follows, focused on making the model’s information accessible to read, enjoy and reuse through standardisation techniques that guide the creation of source-based models and help the user understand its content. Many techniques have been developed along the years, but this paper focuses on three: the Reconstructive Units [RU] (Molina Vidal and Muñoz Ojeda 2015), the Extended Matrix [EM] (Demetrescu 2015), and the Graphic Scale of Historic-Archaeological Evidence [GS] (Aparicio Resco and Figueiredo 2016) (fig. 2).

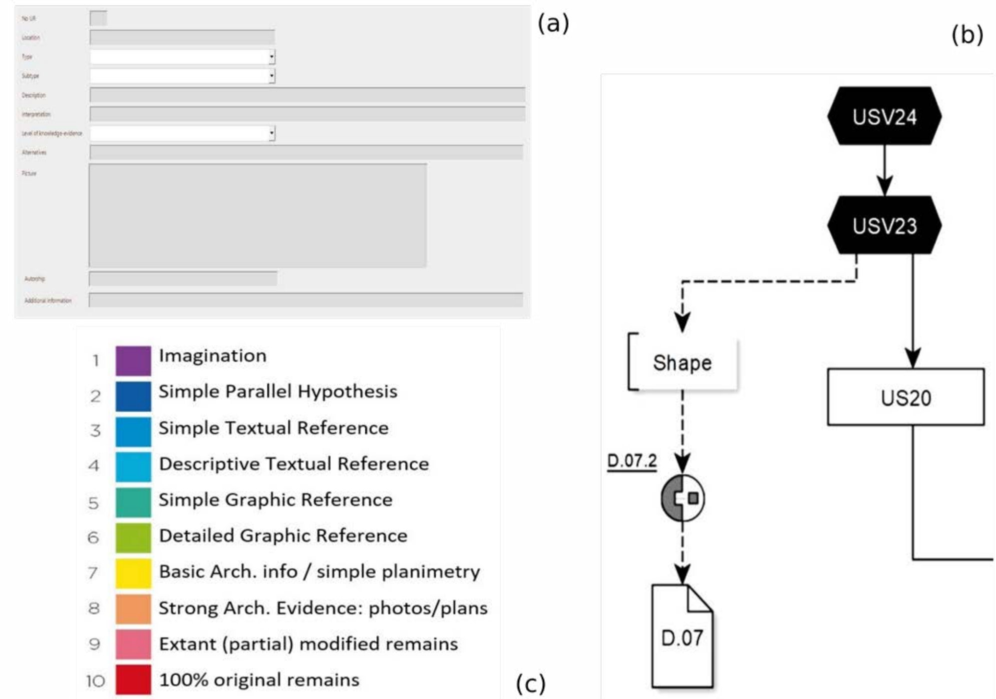

Figure 2. Examples of the three standardisation techniques applied in this study to record the metadata and paradata of V-Rs: (a) clean RU template in English; (b) Example of the combination of “found” nodes (HM) and reconstructed nodes to create an EM; (c) Fixed colour ramp of the GS together with its relation to a type of source. (Images: Cristina Gonzalez-Esteban)

Reconstructive Units [RU]

The RUs (fig. 2a) are context sheets whose template was designed in Access by Prof. Molina Vidal and Prof. Muñoz Ojeda in 2015 but which is easily transferable to a FLOS software such as LibreOffice. The files use the same parameters to record the metadata and descriptions of the archaeological units found on the field and the ones reconstructed in labs, making it easy to complete, publish, share and access them.

They are simple and integral, collecting the fundamental information to reproduce the reconstruction using limited options; but at the same time, having an open space in case the author needs to add extra information.

Extended Matrix [EM]

The EM (fig. 2b) was first published in 2015 by its founder Dr Demetrescu. This method was born from the necessities and complexities of the archaeological discipline and it has been defined as "a visual node-based formal language grounded on a stratigraphic approach designed for virtual archaeology and on the theory of knowledge graphs" (Demetrescu and Ferdani 2021, 2).

Even though it is still in development, its core uses the semantic tools applied for recording reality-based models together with the stratigraphic principles of the Harris Matrix to create a new language able to document and validate Archaeological 3D Reconstructions. These "extra" nodes (Demetrescu and Ferdani 2021, 29) represent the "not found" parts of Virtual Reconstructions and allow to easily connect the archaeological field finds with its lab reconstruction.

Furthermore, the visual differentiation of the nodes makes it not dependent on any language (i.e. universally understandable) and its synoptic (all the elements and their relationships are in the same file) and synchronic (all chronological phases are in one file) characteristic eases the reading of archaeological data. (A detailed use of the EM and its tool can be read in Demetrescu and Ferdani 2021)

Graphic Scale of Historic-Archaeological Evidence [GS]

The idea of using a colour ramp to display data granularity, authenticity and reliability of a V-R was first presented during the project "Byzantium 1200" (Ihsan-Tunay and Berger 1994, http://www.byzantium1200.com/). However, it would be in 2016 when the concept was published (Aparicio Resco and Figueiredo) establishing the currently most popular colour ramp for Archaeological V-R (fig. 2c). Establishing a series of fixed colours to present source type and reliability helped eliminate the "Black Box" effect of V-R: warm colours represent high levels of authenticity (being the highest reserved to context found on the field), while cold show lower reliability (leaving the coolest colours to the imaginative additions).

Even though assigning the colours to each reconstructed unit is still a manual job, creating subjectivity on the interpretation, the use of this fix GS has proven to have positive results to boost the importance of the transparency principle. It is not dependent on language (it is visual), and it can be adapted to different audiences, reaching benefits beyond the academic world into the public engagement scenarios as new tools to bring the public to understand the reconstruction process and evidence support of the models.

Tepidarium of L'Alcudia de Elche, Spain

The site is located outside Elche, Alicante, on an elevated platform; hence its Arabic name, "L'Alcudia". Human settlement date back to the 4th millennium BC, through the Iberian, Roman and Visigoth/Byzantine periods, until the founding of the current city during the Muslim Era (8th century AD). Afterwards, this land was farmed and looted, as rich remains were unburied. Archaeological excavations started in 1933, but it would not be until the 21st century when the University of Alicante buys the land and, together with local authorities, establishes the archaeological site.

The reconstructed room (Ambient 28) is part of the Roman Eastern *Thermae* complex of the site (1st to 3rd century AD) (see fig. 1b). The presence of different construction styles indicates at least two phases and a series of targeted repairs. During the 4th century AD, the site lost its function and split into different rooms for habitation. During the 7th century AD, the Visigoth Era, most structures were semi-ruined, and the area became a dumpsite (Tendero Porras *et al.* 2014).

Ambient 28 was initially catalogued as the *Calidarium* due to its small size. However, further excavation revealed that the heating ovens were separated by another room, probably making the water arrive lukewarm. Therefore, it was renamed *Tepidarium.* The room shows its history of repairs, reuse, and abandonment/looting until becoming a dumpsite *c.* 7th century AD.

Methodology

The project methodology was developed in 8 steps:

Documentation and historical study of the site, building or object. This stage included a ***detailed*** study of reports, grey literature, comparative examples (e.g. *Thermae* of Pompey, Herculano, Carranque, Olmeda, etc.) and complementary subjects (architectonic studies, talks and debates of the area, etc.). A collection of the final sources used for the V-R was done using a table that can be found in Appendix1.

Collection of reality-based data: photogrammetry/ 3D scanner. In this case, a photogrammetric survey of the whole Eastern *Thermae* had already been done by the University of Alicante (see fig. 1b). Therefore, my task was only to crop and re-mesh the *Tepidarium* area to be used as a base for the next steps.

Designing the hypothesis. The design was possible thanks to the information gathered and the site plan developed using the photogrammetry model. The hypothesis was discussed and reviewed by the archaeologists in charge of the excavation of the *Thermae* before continuing the process of creating the source-base V-R.

3D modelling. The modelling was done using Blender 2.83 and 2.90.

6. and 7. These steps encompass the **creation of the metadata and paradata** of the model to make it accessible, validated and reusable data. In the beginning, these steps were thought of as independent; however, as the project went on, it was clear that the production of one set of data helped to better understand and quicken the writing of the rest. Therefore, it was decided to group them in the article. A total of 36 RU were written for this V-R (fig. 3a), each one associated with one GS colour (fig. 3c). These sheets also match the 36 nodes (including field and reconstructed) that ended up creating the EM (fig. 3b), which was organised using "Virtual Activities" to ease its reading.

Creating the outcomes. The resulting model opened the possibility of developing a wide range of outcomes for academics and the public. In this project, the visualisations were limited to rendering and images displaying the different elements of the reconstruction, both showing the metadata and paradata recorded (see fig. 3), but also through a photorealistic texturing of the 3D (fig. 4).

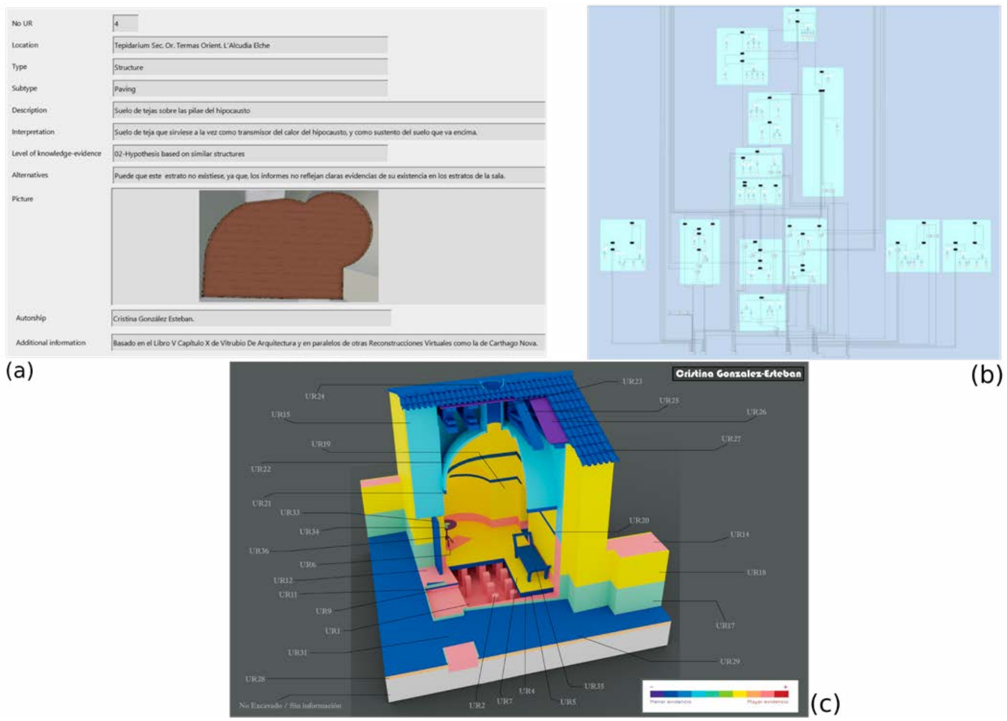

Figure 3. Standardisation techniques applied to the Case Study of the Tepidarium *of the Eastern* Thermae *at L'Alcudia: (a) original RU 4, tiled floor level of the hypocaust, written in Spanish, showing a clear language barrier when using RUs; (b) EM of the* Tepidarium *during its usage as a Roman* Thermae, *separated into Virtual Activities to organise and ease the understanding of the graph; (c) 3D model of the Source-based V-R of the* Tepidarium *textured using the GS colours and completed by linking the RUs to each modelled element. (Images: Cristina Gonzalez-Esteban)*

Discussion

The results reached the main objective, achieving a deep understanding of the room, its construction and usage thanks to a detailed study of the remains, comparative examples and other sources. Due to limited space and thanks to the visual impact of this project, the architectonic and historic results are visible in fig. 4.

Figure 4. Photorealistic texturing of the Source-based V-R representing the interpretation of the Tepidarium *of the Eastern* Thermae *complex of L'Alcudia of Elche, Spain during the Roman Period (1st – 3rd century AD). (Image: Cristina Gonzalez-Esteban)*

Nevertheless, a series of limitations evidence the need to polish the process in future works:

Lack of versatility. The method separates each step and suggests finishing them before moving on. In reality, the workflow could be simplified (Demetrescu and Ferdiani 2021; Gonzalez-Esteban 2021) and should be more fluid and understandable of changes and new interpretations, as they are widespread in Archaeology (e.g. the *Caldarium* switched into a *Tepidarium*).

Novelty. Despite the techniques being from the mid-2010s, their application has not expanded yet, meaning that there are not many source-based models for comparative studies.

Incipient discipline. There are still not many experts on the techniques, hindering cross-referencing the use of the process.

Despite these obstacles, the outcomes are very favourable towards the advantage of using standardisation techniques in virtual archaeological projects to validate them and reinforce their scientific base. The main aspects benefited are:

Time/Effort invested in the source-based reconstruction compared to "traditional" undocumented 3D models of heritage sites.

Potential of sourced models as interactive, transparent and historically accurate tools for public engagement and as references themselves to further academic research.

Furthermore, it would be possible to add the stratigraphic layer of the site, recreating the history of the area (4D model).

Range of outcomes, including both the academic/research environment and as a public engagement tool, thanks to being a great asset to easily display the sources and processes developed in order to reach the resulting V-R.

Conclusions

This project was helpful in understanding that the techniques are compatible, not comparable and that their joining would highly benefit the recording and also the possible outcomes of these source-based models.

Even though this is just one case example, and it would be necessary to test the limits and possibilities of each technique in more case studies (e.g. Gonzalez-Esteban 2021) to consolidate the process in a discipline as heterogeneous as Archaeology, this study has reinforced the value of source-based models for both research and public engagement purposes. Reconstruction models can be scientific and be validated and repeated thanks to the recording of their sources (metadata) and chain of thoughts (paradata) that developed the hypothesis. This makes a new way of creating archaeological assets accessible, comparable and contrastable in an easy and user-friendly manner. It creates a new baseline to start rejuvenating archaeological documentation and bringing it into the increasingly international and virtual world.

Funding: The project that gave rise to these results received the support of a fellowship from "La Mutua Madrileña" Foundation.

Acknowledgements: I would like to thank my supervisor, P. Aparicio Resco, from PAR and the University of Alicante, especially Prof. Muñoz Ojeda, for sharing the data of the site of L'Alcudia with me, making this project possible.

Appendix A

This Appendix presents the table developed to collect all the sources used in the reconstruction study of the *Tepidarium* of the Eastern Roman *Thermae* at L'Alcudia of Elche, Spain. It is based on an initial template of the "Dossier Comparatif" proposed by Dr. Demetrescu in 2018 (tab. A1).

Table A1: Sources used in the Tepidarium *source-based model detailing: its related EM node, the type of document, its descriptions and its reference. This table contains images that have had to be erased due to limits on the current publication; however, they have always been substituted with a website reference to an online copy of the image. (Table: Cristina Gonzalez-Esteban).*

# EM Node	Type of Document	Description	Reference
D00	Wildcard	Represents the absence of archaeological finds	Absence
D1	Text	Vitruvio, Book V Chapter X, ***De Architecture***	Vitruvio Polion, Marco. ***Ten books on Architecture.*** Ediciones Akal, 1992.
D2	V-R	San Juan de Maliaño	Balawat: https://zonacachonera.wordpress.com/2014/02/06/asi-serian-las-termas-romanas-de-san-juan-de-maliano/
D3	V-R	Villa de la Olmeda	Balawat: https://www.balawat.com/reconstruccionesvirtuales/#&gid=1&pid=1
D4	V-R	Carranque	http://www.lacerca.com/noticias/reportajes/parquearqueologico-carranque-anos-126559-1.html
D5	V-R	Eastern ***Thermae*** of L'Alcudia of Elche	Digital Report of Elche 2014 and 2015: http://www.elche.me/video/las-termasorientales-de-la-alcudia-29-de-junio-de-2014 http://www.elche.me/video/virtualizacion-de-las-termas-orientales-de-la-alcudia-23-de-septiembre-de-2014 http://www.elche.me/video/fotogrametriaaerea-y-reconstruccion-virtual-en-arqueologia-la-alcudia-24-de-marzo-de-2015
D6	V-R	V-R of the site of Carthago Nova	Madrid, Maria Jose, Noquera Celdran, Jose Miguel and Pavia, Marta "Las termas del Puerto de' Carthago Nova': un complejo augusteo de larga perduración". ***Tarraco Biennal: actes, 2on Congrés Internacional d'Arqueologia i Món Antic: August i les províncies occidentals, 2000 aniversari de la mort d'August: Tarragona, 26-29 de novembre de 2014*** (2015): 15-22.

D7	Text	Detail description, including measurements, of the Alhange ***Thermae***	Alvarez Martinez, Jose Maria "Las termas romanas de Alange", ***Habis*** 3 (1972): 267-290.
D8	Image	Alexandre de Laborde 1811: Baños de Alhange. It shows a room with a lantern, exedras and a circular dome.	https://commons.wikimedia.org/wiki/File:1806-1820,_Voyage_pittoresque_et_historique_de_l%27Espagne,_tomo_I,_Ba%C3%B1os_de_Alhange_(cropped).jpg
D9	Image	Painting by L. Alma-Tadema 1881: The ***Tepidarium***	https://commons.wikimedia.org/wiki/File:Tepidarium_Lawrence_Alma-Tadema_(1836-1912).jpg
D10	Paint	Painting by L. Alma-Tadema 1879: Strigils and Sponges. Tipical objects in the ***Thermae***	https://commons.wikimedia.org/wiki/File:Tepidarium_Lawrence_Alma-Tadema_(1836-1912).jpg
D11	Arch. Remains	Remains of the ***Thermae*** at Pompeii	https://commons.wikimedia.org/wiki/File:Frigidarium._Termas_del_foro._01.JPG
D12	Arch. Remains	Remains of the ***Thermae*** at Pompeii: ***Caldarium***	https://commons.wikimedia.org/wiki/File:Forum_Baths,_the_apse_of_the_caldarium_(hot_bath_room)_containing_a_labrum_or_(marble_basin),_Pompeii_(14832295629).jpg
D13	Image	Example of a Roman ***Caldarium*** at a public forum of Pompeii, crown by a ¼ dome.	https://commons.wikimedia.org/wiki/File:Caldarium_of_the_Old_Baths_at_Pompeii_by_Overbeck.png
D14	Arch. Object Catalogued	Wooden double doors found in the 2nd floor of the ***Decumanus Maximus*** house at Herculaneum.	Digital Report of Elche 2014: http://www.elche.me/video/virtualizacion-de-las-termas-orientales-de-la-alcudia-23-de-septiembre-de-2014
D15	V-R	V-R of the ***Thermae*** at Herculaneum.	MAV (Museo Archeologico Virtuale de Erconalo) 2014: https://www.youtube.com/watch?v=klre-67UI1Q
D16	Image	Polychrome mosaic of reddish and ochre tones, located on the wall, creating garland diamond shapes. Provenance: Villa del Romeral (Albesa, Lleida).	Museu d'Arqueologia de Catalunya

References

Alvarez Martinez, Jose Maria. 1972. "Las termas romanas de Alange", *Habis* 3 : 267-290.

Aparicio Resco, Pablo, and Cesar Figueiredo. 2016. "El grado de evidencia histórico-arqueológica de las reconstrucciones virtuales: hacia una escala de representación gráfica." *Otarq* 1 : 235-247. http://dx.doi.org/10.23914/otarq.v0i1.96.

Demetrescu, Emanuel. 2015. "Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice." *Journal of Archaeological Science* 57: 42-55. https://doi.org/10.1016/j.jas.2015.02.004.

Demetrescu, Emanuel. 2018. "Virtual Reconstruction as a Scientific Tool". In *Digital Research and Education in Architectural Heritage*, edited by Sander Münster, Kristina Friedrichs, Florian Niebling and Agnieszka Seidel-Grzesińska, 102-116. Cham: Springer. ISBN: 9783319769912.

Demetrescu, Emanuel, and Daniele Ferdani. 2021. "From Field Archaeology to Virtual Reconstruction: A Five Steps Method Using the Extended Matrix." *Applied Sciences* 11, no. 11 : 5206. https://doi.org/10.3390/app11115206.

Denard, Hugh. (ed.). 2009. *The London Chapter for the computer-based visualisation of cultural heritage.* London: King's College London.

Forte, Maurizio. (ed.). 2010. *Cyber-archaeology*. Oxford: Archaeopress.

Garcia, Juan Simon. 2021. "Extended Matrix aplicada al Anfiteatro de Cartagena". MA Diss. *Unpublished*. University of Alicante.

Gonzalez-Esteban, Cristina. 2019. "Estandarización metodológica Internacional de la Reconstrucción Virtual en Arqueología. Caso práctico del Tepidarium de la Alcudia de Elche". MA Diss., *Unpublished*. University of Alicante.

Gonzalez-Esteban, Cristina. 2021. "Testing the archaeological importance of metadata-orientated Virtual Reconstructions. Case Study of a Roman Wreck from the Black Sea". MSc Diss. *Unpublished*. University of Southampton.

IFAV (International Forum of Virtual Archaeology). Principles of Seville. International Principles of Virtual Archaeology. Accessed Feb. 20, 2022. https://icomos.es/wp-content/uploads/2020/06/Seville-Principles-IN-ES-FR.pdf

Ihsan-Tunay, M., and Albrecht Berger. 1994. Proyect Byzantium 1200. Accessed Feb. 20, 2022. https://www.byzantium1200.com/

Madrid, Maria Jose, Jose Miguel Noquera Celdran, and Marta Pavia. 2015. "Las termas del Puerto de' Carthago Nova': un complejo augusteo de larga perduración". *Tarraco Biennal: actes, 2on Congrés Internacional d'Arqueologia i Món Antic: August i les províncies occidentals, 2000 aniversari de la mort d'August: Tarragona, 26-29 de novembre de 2014* : 15-22.

Molina Vidal, Jaime, and Francisco J. Muñoz Ojeda. "Ficha UR, 2015". Accessed Feb. 14, 2022. http://rua.ua.es/dspace/handle/10045/46205.

Reilly, Paul. 1990. Towards a virtual archaeology. In *CAA90: Computer Applications and Quantitative Methods in Archaeology 1990,* edited by Kris Lockyear and Sebastian Rahtz (eds.), 133–139. Oxford: BAR International Series, 565.

Tendero Porras, Mercedes, Ana M. Ronda Femenia, Rafael Ramos Fernández, Alejandro Ramos Molina, Diego Peña Domínguez, and Lorenzo Abad Casal. 2014. "Ciudades Romanas Valencianas. Ilici". In *Ciudades Romanas Valencianas. Actas de las Jornadas sobre Cuidades Romanas Valencianas. Actualidad de la investigación historicoarqueológica, celebradas en el MARQ los días 3 y 4 de diciembre 2013,* edited by Manuel H. Olcina Doménech, 225-249. Alicante: MARQ, Dip. Alicante,. ISBN 978-84-15327-47-9.

Vitruvio Polion, Marco. *Ten books on Architecture.* Ediciones Akal, 1992.

Using Programming Environments for Academic Research and Writing

Morgan Lemmer-Webber[1]*

[1]Art History, Executive Director of the World History Association,
Director of FOSS & Crafts Studios LLC
*Correspondence: morganlemmerwebber@gmail.com

Presented at the ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; on-line; November 23rd-26th 2021.

Abstract: Developer tools, such as code editors, markup languages, and revision control have a greater range of functions than word processors. As a scholar engaged in both Digital Humanities and the FOSS community, I have become increasingly interested in how these tools can be applied to research workflows. I wrote my dissertation in the editor Dr. Racket using Scribble, allowing me to incorporate code directly into my document. In this paper, I discuss the benefits and pitfalls of this decision.

Keywords: Markup languages; Revision control; Digital humanities

FOSS software used and license:
Scribble, MIT license and Apache License v 2.0;
Racket and Dr. Racket, MIT license and the Apache License version 2.0 (some components distributed under the GNU Lesser General Public License, version 3);
git-annex, AGPL version 3 or higher (parts of git-annex are licensed under the GPL, BSD, and other licenses);
GNU Emacs, GPL;
LibreOffice, Mozilla Public License v2.0

Open dataset and license:
N/A

Repository and license:
https://mlemmer.org/git/dissertation/, Apache License v2 (non-Scribble racket files) and Creative Commons Attribution-Share Alike 4.0 International (document files).

Introduction

I have been an active user of free and open source technology for about fifteen years and have run Linux distributions for the operating system on my primary computer for over a decade. For most of that time, my free software advocacy and use has run in parallel with little overlap to my academic studies as an art historian focusing on the art and archaeology of the Roman Empire. However, when I started getting involved in digital humanities projects, I had more tangible reasons to incorporate that technology into my academic projects. This article will discuss my own experiences using programming tools in my research workflow and outline how these methods could be more broadly applied (fig. 1).

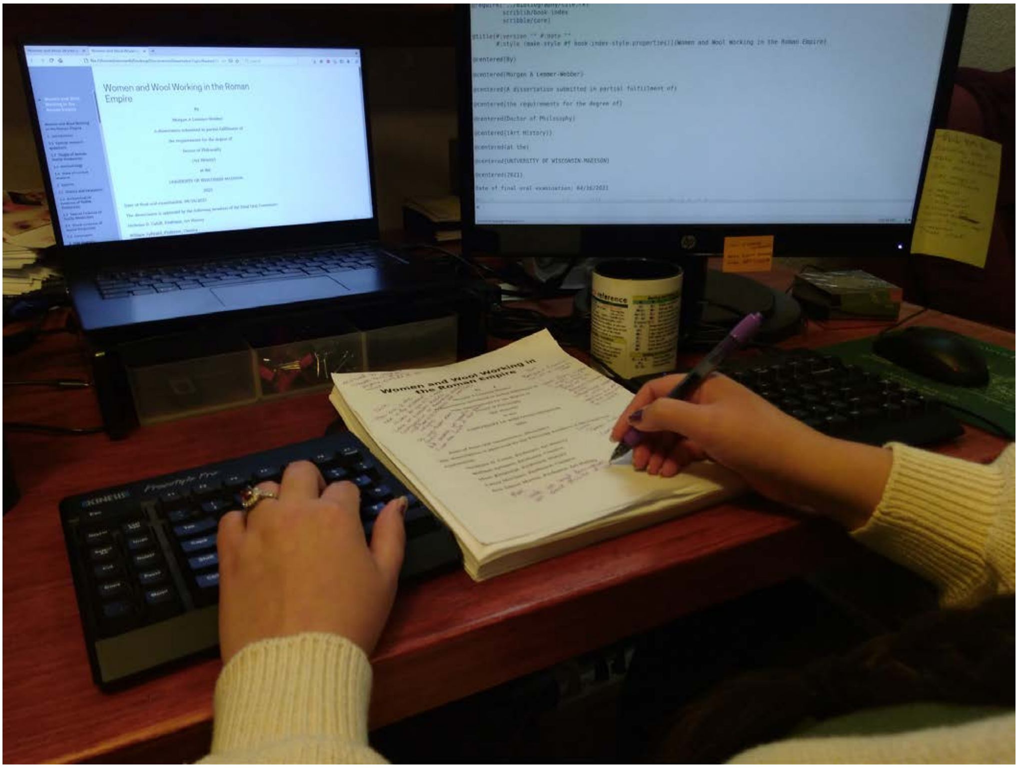

Figure 1. Morgan Lemmer-Webber's dissertation workflow as she incorporates notes written on the physical draft of her dissertation into the scribble sourcecode (right monitor), which is displayed via html in the browser (left monitor).

Digital Humanities Workshops

I co-developed and ran a series of digital humanities workshops with my wife, Christine Lemmer-Webber, to teach the basics of programming to students with no math or computer science background (Lemmer-Webber and Lemmer-Webber 2018). For these workshops, we decided to use the programming language Racket (Racket 2022) because it had a code editor, Dr. Racket, and markup language, Scribble (Scribble 2022), built in. Based on my own experience learning Python, as well as conversations with other humanities students about their anxieties around learning programming, we made a conscious effort to tailor the workshops to minimize those anxieties. The objective of the three-hour workshops was to give a basic introduction to the concepts of programming with a deliverable skill of writing an academic paper using Scribble. Christine wrote a programming tutorial that introduced the basic principles of programming using Racket's picture language to build a snowman. We found that using basic geometric shapes as our universal language eased the fears of some students who thought they wouldn't be able to code because they did not have a math background. Since the number of code projects that are accessible within a basic tutorial that will also have a deliverable outcome relevant to students are fairly limited, I created a tutorial for using Scribble to write an academic paper. This continued the code elements of the workshops using a method that is relatively easy to pick up and had the potential for immediate application.

Dissertation Workflow

Scribble

When it came time to write my dissertation (Lemmer-Webber 2021), I decided to practice what I preach and use Scribble with Dr. Racket as my primary writing environment. This technology provided many advantages over standard word processors. As with any markup language, Scribble made it easier to maintain consistent formatting throughout my dissertation. There was no risk of pasting in a quote from another source and corrupting the formatting for the rest of the section, for example. I was able to export to multiple file-formats including HTML and PDF, from the same source file.

One of the basic features of writing in a code editor that I found most helpful was the ability to comment out text. This means that the text remains in the source document but does not export into the final versions of the text. In programming, this is often used for writing documentation about how your code words or keeping earlier drafts of your code so that you know what you have already tried. When writing academic work, I used this feature to make to-do items for myself, add notes to myself, include feedback from my committee members, or to indicate text that was not immediately relevant but could be useful elsewhere. In a screenshot of my commercial chapter (fig. 2), you can distinguish the sections that are commented out with @; and appear in tan. At the bottom of that page is a note from one of my committee members reminding me to keep the focus on my topic. Near the middle of the page is a table and paragraph which I subsequently commented out to heed that advice. Having the ability to keep all of this text in one location allowed me to remain organized without cluttering the drafts I was sending out to my advisors.

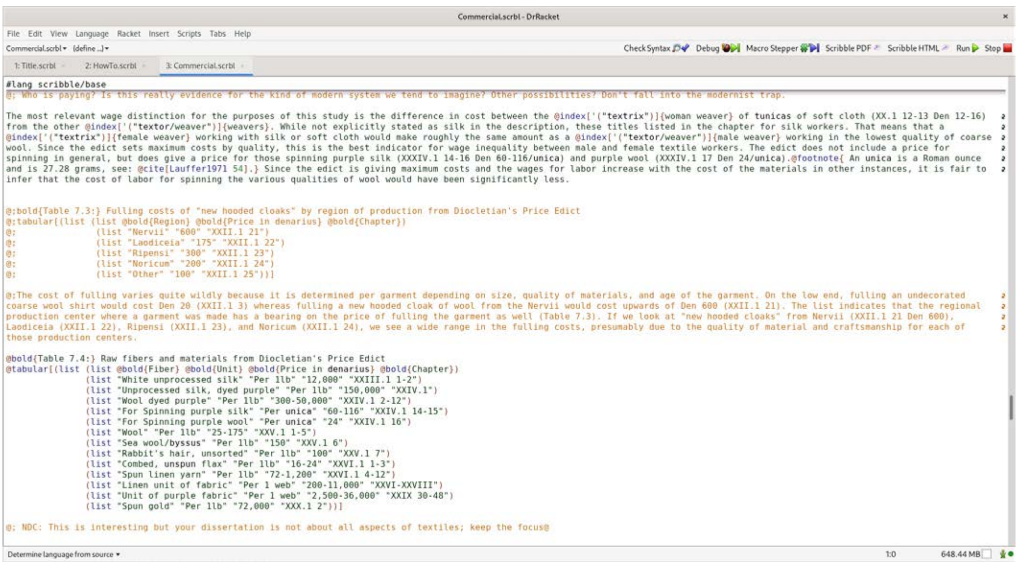

Figure 2. Screenshot of the commercial chapter of my dissertation showing Scribble formatting including commented out notes, and tags for the index, citations, and tables.

Since Scribble is associated with the programming language Racket, I was able to code functions directly into the source document. The document is still effectively Racket code but when you set the programming language within Dr. Racket to `#lang Scribble/doc`, it inverts the typical structure of the programming environment to assume that everything you input is a string unless you call a function with the `@` symbol. Scribble has a wide array of functions out of the box including bibliography, footnotes, images, and figures. However, it was intended for students and researchers within the discipline of computer science and therefore did not have all of the nuances that I needed as a PhD candidate in the humanities.

Luckily, the association with Racket allows for custom code to be built into Scribble documents. Working with Christine Lemmer-Webber we were able to automate certain functions in ways that were relevant to my field. The bibliographic style options for Scribble were limited to those typical of computer science journals and did not have an option appropriate to Art History or Archaeology. The way the bibliography for Scribble operates relies on creating a new function for each citation and manually entering the bibliographic information. However, I also had an existing database of bibliographic metadata for my dissertation references from the citation manager Zotero and wanted to avoid reduplicating work unnecessarily (Zotero 2022). Since we had to re-create the bibliographic system to meet the requirements of my field, we developed a Racket program that compiled the bibliography from the XML database into an output that followed the style guide of the *AJA*. Like the bibliography, the boilerplate structure for creating figures includes manually entering the image descriptions. Since I wanted to be as consistent in the image identifications as possible, we wrote a piece of custom racket code to draw that information from a CSV database and procedurally generated the descriptions.

Scribble has an existing structure for compiling a multi-chapter document. In a standard word processor, this would likely be achieved either by using one large document for the entire dissertation, which can become cumbersome or by creating individual documents for each chapter and then manually pasting them into a full draft in the end stages of the process. In Scribble, you can use the `@include-section` command to append additional Scribble files on one master document. In the case of my dissertation, I used the title page as my master document and appended the individual chapters as well as call other functions that will be procedurally generated. Functionally, this means that exporting the title page compiles the whole dissertation including the Scribble files for individual chapters as well as calls functions to generate the image list, bibliography, and index (fig. 3).

The built-in option to export to PDF is formatted based on conventions in computer science and did not incorporate an easy way to reset the formatting parameters of the output. Formatting LaTeX to a specified style guide is already a difficult feat on its own. Formatting generated LaTeX from a markup language adds a further layer of difficulty. Ultimately, we were only able to get about 90% of the formatting to output to PDF correctly on export. In the end, Christine built a separate export option for the Open Document Format or ODF (Durusau and Svante 2021). Once exported to ODF, I was able to open the document and make the final small adjustments using the open source word processor LibreOffice (The Document Foundation 2022).

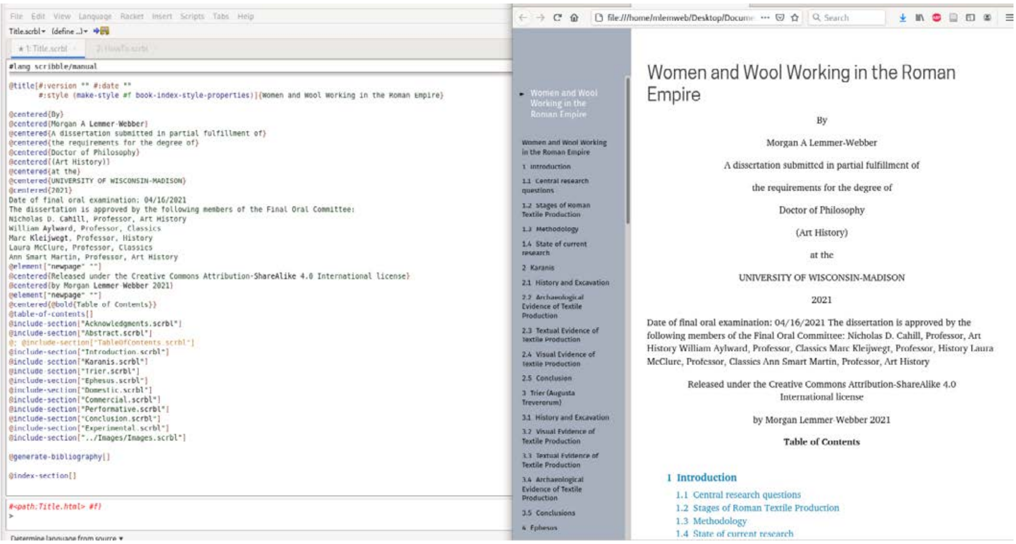

Figure 3. Screenshots of the title page of my dissertation (a) in the Scribble format this serves as the master document with the separate sections included (b) The HTML output of my dissertation with the title page visible and the hyperlinked table of contents both below and to the left of the title page.

Revision Control

Since I was already using programming tools for my dissertation, I further enhanced my workflow by using revision control. This process saves changes to a file incrementally over time, effectively allowing the user a time-traveling view of document history. For many academic writers, saving incremental or even periodic changes requires saving each version of the document with a slightly different file name. This clutters your file system and can be difficult to navigate to find the specific change you are searching for. Git, the most common free and open source system for revision control, saves these versions in a separate repository, leaving only one copy of the document in your file system and prompts you to explain what changes you've made in a commit message (Git 2022). This is not a feature built into Dr. Racket, nor most standard word processors. It is most typically accessed either through the command line, or through a platform such as GitHub that offers a graphical user interface. Git can be used with most text-based file types, so revision control is accessible if you are writing using a standard word processor. Since Git on its own does not handle images well, I added Git Annex, a free software revision control system built for large file types, in order to include all of the files necessary to export my dissertation in the same repository (Git-annex 2022).

Access to Git repositories can be shared with collaborators and track who made changes to a document and when. This makes them indispensable for team projects. While co-developing the custom Racket functions for my dissertation it was necessary to commit the code regularly so that Christine and I were both working with the most recent version. It is worth noting that this is a very foreign workflow for most people and the command line interface is notoriously difficult even for many experienced developers, let alone a humanities student. The majority of work on a dissertation is done alone. When not coding on a regular basis or having another person reliant on the most recent update, it is much easier to forget to commit your changes even if it is still good practice. In order to compensate for this user oversight, I used a program called Git Annex Assistant to automatically commit my dissertation files to Git and Git Annex periodically (Git-annex/assistant 2022).

Obstacles

Unfortunately, achieving these benefits came with its own share of obstacles. While Scribble on its own has a relatively low barrier to entry for the boilerplate options, it is still higher than standard word processors. All markup languages are an adjustment, particularly for users who are acclimated to graphical user interfaces for formatting. With few exceptions, making a single error in formatting in a standard word processor is unlikely to crash your entire document. When using a markup language, a single missed closing parentheses or a bracket instead of a curly bracket will result in an error message and the inability to export your document until it is fixed. The interface of Dr. Racket does syntax highlighting on functions to indicate which information is covered within that function and the line where the error occurs will be highlighted in pink on the right side of the screen. These indicators help locate such bugs as they occur and are reasonably easy to identify if you are mindful of them as you work. The real issues arise when you do not catch a bug soon enough, particularly if you have multiple bugs at once. The best preventative measure against this, in my experience, is to export the document to HTML once per paragraph written. This way you have a constrained area of text to search for bugs if the export fails making it more manageable than searching through pages of text.

While novice programmers like myself can absolutely debug Scribble errors, define new functions and even incorporate simple code into their source documents, for the higher-level functions and customization I did have to rely on assistance from a more experienced developer. Without access to such guidance this process would have been stalled and I would have either been limited to the boilerplate options available in Scribble or given up and returned to LibreOffice. Before Scribble can be widely adopted by non-programmers, I believe more infrastructural work needs to be laid out to make the options more adaptable to the needs of other fields.

Incorporating git and git-annex as someone with minimal development experience likewise had a steep learning curve. The command line is not an intuitive interface even for many experienced programmers. For people who have minimal command line experience, committing to git is more labor and time investment than it is to programmers who commit code several times daily. Programmers are also more likely to have the command line open as they test their code, run virtual environments, or perform other routine tasks. When your daily work does not require the command line, it is difficult to remember to commit regularly, thus the necessity for a program to auto-commit my dissertation. This workaround was not without flaws, particularly since the auto-commits do not include commit messages. I, therefore, sacrificed having an easily navigable repository with messages explaining what changes were but with those commits happening for a repository with regular commits in short intervals that are likely to catch more changes but without context clues as to what those changes were.

While writing my dissertation in Scribble had many benefits, it also added layers of stress to an already difficult situation. Like most people within academia, I had years of experience writing papers using a word processor. While a lot of the tasks in that workflow are tedious, they are at least familiar. Using a markup language introduced unexpected issues that I wasn't always able to fix in a timely manner. It is difficult explaining to your committee members that while italicizing words in foreign languages is a trivial task in a word processor, it was an export error for my dissertation. Even though none of the text in the exported draft that I submitted had italics, behind the scenes the correct words were still marked as italic in the source document but a solution we applied to fix one formatting error somehow corrupted other areas of formatting. This issue arose because debugging LaTeX output is difficult at the best of times but debugging auto-generated Latex as an intermediate step between the markup language (Scribble) and the final document format (PDF) is even more difficult. The level of complexity here can be inferred by the fact that it was easier, in the end, to write an entirely new export method to ODF than debug the existing Scribble to LaTeX to PDF exporter. With a lot of work, we were able to find workarounds for all of these issues by the time I needed to submit my dissertation but troubleshooting these issues averted my attention away from writing the content of my dissertation. That being said, this article is an overview of my own personal experience using these technologies and therefore has a sample set of one. I have only written one dissertation, and without a control sample, it is hard to say whether I spent more time trying to debug my encoded dissertation than I would have individually formatting every image, manually compiling an index and all of the other minutiae that were successfully automated in my workflow.

Gnu Emacs

When I started working on my dissertation, Dr. Racket and Scribble were the obvious choices of platform because they were the most familiar programming environment to me. I knew that there were more powerful programming environments out there but did not anticipate the number of obstacles that would arise from scaling the tools I knew from a seminar paper to a dissertation. We realized midway through the process that using a more robust editor such as Gnu Emacs would have eliminated some of the more difficult obstacles (Gnu Emacs 2022). However, at that point, my dissertation deadlines were looming too close to completely change my workflow. I began learning Emacs after I submitted my dissertation and had more free time to invest.

The Emacs programming environment has been consistently used and developed since it appeared on the scene in 1976. This means that it has accrued a powerful assortment of features that can adapt to most computing needs (Lemmer-Webber and Lemmer-Webber 2022). This long history, however, also creates its own barriers. The keyboard shortcuts for Emacs were developed before the standardized shortcuts we all know today. Instead of using Ctrl-c and Ctrl-v to 'copy' and 'paste', for example, you use Ctrl-w and Ctrl-y to 'kill' and 'yank' (Free Software Foundation 2016). It has gone through various iterations and I am using Gnu Emacs specifically. While the program does have a graphical user interface, it was designed to be navigated using keyboard shortcuts and therefore is less intuitive than many other programs.

Whereas Dr. Racket was created as a programming environment purpose-built for Racket and Racket-based languages, Emacs has support for most standard and many obscure programming languages. Therefore you are similarly able to write custom code into your source document, but you are able to choose your programming language of choice (i.e. python, R, Racket). This versatility expands to markup languages as well and Emacs is compatible with markdown, LaTeX, and HTML, among others and can export to PDF, HTML, LaTeX and ODT. When working on the Scribble ODF exporter, Christine reverse-engineered and adapted the Emacs ODF exporter as a template.

Emacs-Org Mode is an organizational system that uses a simple but versatile markup system that can be used for outlining, task assignment, project planning, and to write text documents (fig. 4). Since I am currently not writing code on a regular basis, the majority of my tasks in Emacs have been using either Org Mode or markdown. These tasks vary but include documents to track my job application process, write updates and manage my personal website, write tutorials and educational materials, collaborate on outlines for projects, and manage contracting clients.

The other main aspect of my workflow that Emacs has tremendously improved is Magit, a git porcelain that is integrated into the Emacs interface (Magit 2022). This means that I can access a much more intuitive interface for git that is accessible through the same program I am writing in and doesn't require a separate terminal or command line interface (fig. 4). It is far easier to remember to stage and commit changes to a document when it only requires a handful of keystrokes and a commit message rather than switching interfaces entirely.

Given the nature of this article as a retrospective of my experiences, it felt fitting to write the article itself in Emacs to expand that experiential aspect. While this brief writing exercise is far less complex than a dissertation as a whole, it has been an interesting comparison. The main limitation to this workflow in this particular instance is that instead of a style guide for the formatting, the editorial board of the ArcheoFOSS conference provided a .docx template with formatting built in. This means that I am writing the document in Org Mode using Emacs, exporting it to ODF, then assembling it into the template in LibreOffice.

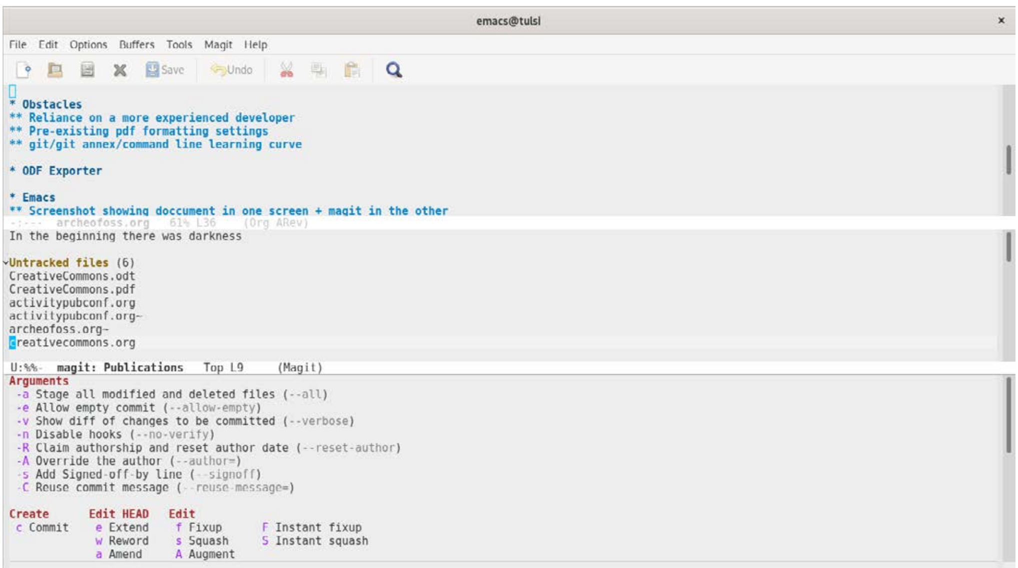

Figure 4. A screenshot of the outline for this article in Emacs Org Mode visible at the top of the screen, and the Magit interface to commit the document to Git at the bottom of the screen.

Conclusions

My experience using Scribble to write my dissertation was unique and, in many ways, experimental. There were learning curves along the way that exceeded most dissertation experiences. I believe that incorporating these types of programming tools into my research workflow has greatly enhanced my experience. I was able to hone skills that were applicable both to my academic career and more broadly marketable while minimizing the list of tedious tasks that typically overwhelm students at the end stages of the dissertation process. If I were to go back to the start, I would still begin the learning process with Dr. Racket and Scribble. Emacs itself is too overwhelming and its user interface too foreign to be a useful entry point to using programming environments. Having an intermediate stage that has a comparatively low barrier to entry allowed me to gain the confidence I needed to later take on learning Emacs. Though, given the limitations we hit with Scribble for a complex document with strict formatting regulations with my dissertation, I would advise anyone who is contemplating this approach to take the time to learn a more robust code editor prior to beginning a dissertation- or book-length project.

I believe that incorporating these types of programming tools into research workflows has significant merit and broad application. While the custom features written into my dissertation were primarily intended to automate the processes, which are tedious and time-consuming, the ability to write custom code into a source document has an infinite number of applications for projects whose needs exceed standard word processors. Projects that involve large amounts of data analysis, for example, could write these functions directly into their source file rather than compiling the information elsewhere and then incorporating it. A project that compiles data from a changing or fluctuating pool of information, such as annual reports for an excavation, could create a custom template that auto-generates the annual statistics into a consistent format. Any project which requires multiple users to edit a document could benefit from revision control which monitors who made changes, when, and preserves older copies of the document in the event that an error is made. Furthermore, having a publicly available git repository where the data, source code, methods, and reports are available greatly increases reproducibility. While digital tools and processes have increasingly revolutionized the way that data and research is interpreted, visualized, and shared, the writing workflow for many scholars has remained relatively unchanged. Imagine how much more we could achieve if we think outside of the .docx.

Supplementary Materials: As the hosts of the podcast *FOSS and Crafts*, Christine and I have recorded episodes about many of the themes covered in this article. The following are available online at https://fossandcrafts.org, Podcast S1: Digital Humanities Workshops, Podcast S2: Scribble and the Open Document Format, Podcast S3: Learning Emacs.

Funding: This research received no external funding.

Acknowledgments: I am deeply indebted to my wife, Christine Lemmer-Webber, for her assistance in developing the custom code embedded into my dissertation as well as her work co-developing the digital humanities workshops and the podcast FOSS and Crafts.

References

The Document Foundation. 2022. "LibreOffice." Accessed February 28, 2022. https://www.libreoffice.org/.

Durusau, P. and S.Schubert (eds). 2021. "OASIS Open Document Format for Office Applications (OpenDocument) TC v1.3." Accessed February 28, 2022. https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office.

Free Software Foundation. 2016. "GNU Emacs manual." Accessed February 28, 2022. https://www.gnu.org/software/emacs/manual/emacs.html.

"Git – fast-version-control." 2022. Accessed February 28, 2022. https://git-scm.com/.

"Git-annex." 2022. Accessed February 27, 2022. https://git-annex.branchable.com/.

"Git-annex/assistant." 2022. Accessed February 27, 2022. https://git-annex.branchable.com/assistant/.

"Gnu Emacs." 2022. Accessed February 27, 2022. https://www.gnu.org/software/emacs/.

Lemmer-Webber, C. and M. Lemmer-Webber. 2020. "Scribble and the Open Document Format" Podcast episode. *FOSS and Crafts,* November 5, 2020. https://fossandcrafts.org/episodes/15-Scribble-and-the-open-document-format.html.

Lemmer-Webber, C. and M. Lemmer-Webber. 2022. "Learning Emacs." Podcast episode. *FOSS and Crafts,* February 5, 2022. https://fossandcrafts.org/episodes/41-learning-emacs.html.

Lemmer-Webber, M.. 2021. "Women and Wool Working in the Roman Empire." Doctoral Dissertation, University of Wisconsin Madison. https://mlemmer.org/dissertation/

"Magit, A Git Porcelain inside Emacs." 2022. Accessed February 27, 2022. https://magit.vc/.

"Racket, the Programming Language." 2022. Accessed February 27, 2022, https://racket-lang.org/.

"Scribble: The Racket Documentation Tool." 2022. Accessed February 27, 2022. https://docs.racket-lang.org/Scribble/.

"Zotero, Your personal research assistant." 2022. Accessed February 27, 2022. https://www.zotero.org/.

re3dragon – A Research Registry Resource API for Data Dragons

Florian Thiery [1,*] and Allard W. Mees [1]

[1]Leibniz-Zentrum für Archäologie, Mainz, Germany
*Correspondence: florian.thiery@leiza.de

Presented at ARCHEO.FOSS XV 2021: Open software, hardware, processes, data, and format in archaeological research; online; November 23rd-26th 2021.

Abstract: The re3dragon (REsearch REsource REgistry for DataDragons) FLOSS tool envisages two aims: first, publication of an archaeology related open extendable LOD resource catalogue including authority-data, thesauri, gazetteers, (space-) time gazetteers as well as typologies and domain-specific resources. 2nd, re3dragon offers an API for requesting distributed LOD resources, returning resources in a standardised JSON format based on JSKOS. The re3dragon is coded in JAVA and Open Source published on GitHub.

Keywords: Linked Open Data, SKOS, JSKOS, JAVA, Thesauri Catalogue

FOSS software used and license:
1 re3dragon, MIT License;
2 Netbeans, Apache 2.0 License;
3 Apache Maven, Apache-License, Version 2.0;
4 Apache Jena, Apache-License, Version 2.0;
5 Eclipse Jersey, Eclipse Public License 2.0
Repository and license:
re3dragon on GitHub: https://github.com/leiza-rse/re3dragon
JSKOS+ on GitHub: https://github.com/leiza-rse/jskos-plus

Introduction

Cartographers in historical maps used the phrase "*Hic sunt dracones*" (historically translated as: here be dragons) to describe areas which were beyond the known world of the map creator. Today, the digital data universe is full of unknown data, which have to be made FAIR (Findable, Accessible, Interoperable and Reusable) to integrate them in archaeological research. The World Wide Web offers researchers the possibility to share research data and enables the community to participate in the scientific discourse in order to generate previously unknown knowledge. However, much of this data is neither directly comparable and alignable nor easily findable or accessible, thus resulting in so-called modern unknown Data Dragons (Thiery et al. 2019). These Data Dragons lack connections to other data sets, which leads to a lack of interoperability and, in some cases, unusability. To overcome these shortcomings, a set of techniques, standards and recommendations can be used: Semantic Web and Linked Open Data (LOD) (Berners-Lee 2006; Schmidt et al. 2022) and consequently Linked Open Usable Data[1] (LOUD). To tame and unveil the modern data dragons, the CAA Special Interest Group (SIG) on Semantics and LOUD

[1] https://linked.art/loud/ (accessed 25/02/2022)

in Archaeology (SIG-DataDragon[2]) was established in 2020. Data dragons require a safe location, what we call Dragon Lair, where they can be clustered and made accessible. This data dragon LOD lair location and its machine-readable accessibility are combined in the re3dragon tool - the REsearch REsource REgistry for DataDragons. The FLOSS[3] re3dragon tool envisages two aims. First, publication of an open extendable archaeology related LOD resource catalogue (the Dragon Lair) including authority data (e.g. Integrated Authority File - GND), thesauri (e.g. Getty AAT, Heritage Data Vocabularies), controlled vocabularies, gazetteers (e.g. GeoNames, Pleiades), space-time gazetteers (e.g. ChronOntology, PeriodO), as well as typologies and domain specific resources (e.g. Roman Open Data, Nomisma, Linked Open Samian Ware). 2nd, offering an API for requesting distributed LOD resources (so-called Dragon Items) providing resources in a standardised JSON format based on the JavaScript Object Notation for Simple Knowledge Organization Systems: JSKOS (Voß et al. 2016; Voß 2021). The tool is currently being implemented as part of the LEIZA digital research data infrastructure. It plays an essential role in the design of an overarching keyword registry (Meta-Index), which enables aggregating external distributed data resources in order to qualify internal data (e.g. the Meta-Index term Mainz may refer to Roman, mediaeval[4] or modern Mainz[5], resulting in different aggregation references). This process to transform a vocabulary term into a label by linking it to reference thesauri concepts has been described in Piotrowski et al. (2014). This approach enforces interoperability and reusability, enabling mapping on external major union systems like Europeana. Moreover, these aspects play a pivotal role in the planned German National Research Data Infrastructure (NFDI) consortium NFDI4Objects[6]. Based on creating a JSKOS-enhancement (JSKOS+), re3dragon supports the furtherance of interdisciplinary major existing knowledge hubs such as the Basic Register of Thesauri, Ontologies & Classifications (BARTOC)[7] registry and the accompanying SKOS based mapping tool Cocoda[8].

Implementation

The re3dragon API is coded in JAVA using Maven and Apache Jena and is published Open Source on GitHub[9] (Thiery 2021). The re3dragon is based on the Labeling System approach (Piotrowski et al. 2014; Thiery and Engel 2016; Thiery and Mees 2023) and the re3cat API[10].

The basis of re3dragon is an online catalogue of Linked Open Data resources (the Dragon Lair), which includes norm- and authority data resources (e.g. Getty AAT, Wikidata) as well as domain specific typologies and archaeological data (e.g. Terra Sigillata production centres, distribution sites, potters and ceramic typologies), stored as RDF in a RDF4J triplestore. Dragon Lairs are semantically modelled using the so called Linked Archaeological Data Ontology (LADO) which is based on the Research Software Engineering (RSE) Tools Ontology (Thiery 2019). Figure 1 demonstrates exemplary the Getty AAT Dragon Lair using the prefixes lado[11] and wd[12]:

[2] https://caa-international.org/special-interest-groups; http://datadragon.link (accessed 25/02/2022)
[3] https://www.gnu.org/philosophy/floss-and-foss.en.html (accessed 25/02/2022)
[4] https://pleiades.stoa.org/places/109169; http://imperium.ahlfeldt.se/places/3 (accessed 25/02/2022)
[5] http://sws.geonames.org/2874225; https://www.wikidata.org/entity/Q1720 (accessed 25/02/2022)
[6] https://osf.io/4t29e/wiki/home/ (accessed 25/02/2022)
[7] https://bartoc.org/ (accessed 25/02/2022)
[8] https://coli-conc.gbv.de/ (accessed 25/02/2022)
[9] https://github.com/RGZM/re3dragon (accessed 25/02/2022)
[10] https://github.com/mainzed/re3cat (accessed 25/02/2022)
[11] lado prefix: http://archaeology.link/ontology#
[12] wd prefix: http://www.wikidata.org/entity/

property	value
rdf:type	lado:DataDragonLair
owl:sameAs	wd:Q611299
rdfs:label	Getty AAT
lado:description	controlled vocabulary used for describing items of art, architecture, and material culture
lado:author	Getty Research Institute
lado:sparqlendpoint	http://vocab.getty.edu/sparql
lado:hasLegalType	lado:ResearchInstitution
lado:language	lado:en

Figure 1. Dragon Lair example as RDF (Getty AAT). [Florian Thiery, CC BY 4.0]

The re3dragon offers three types of API services. 1st, a search API service for LOD resources with string and distance similarity; 2nd, an API resolving service for LOD resources related to specific URIs. 3rd, a catalogue API for Dragon Lairs. Service types 1 and 2 offer results according to the JSKOS+ format and HTML, and service type 3 provides an interoperable output integrable in research applications.

/re3dragon/rest/search?repo=gettyaat&q=lion


```
[{
        "displayLabel": {
                "en": "Panthera leo (species)"
        },
        "type": ["http://www.w3.org/2004/02/skos/core#Concept"],
        "lair": {
                "scheme": "Getty AAT",
                "legaltype": "lado:ResearchInstitution",
                "id": "ULBU3XXM",
                "type": "Thesaurus",
                "wikidata": "wd:Q611299",
                "quality": "lado:qualityHigh",
                "group": "lado:chreferencethesauriGroup"
        },
        "uri": "http://vocab.getty.edu/aat/300310388",
        "displayDesc": {
                "en": "Large, powerful species of cat that is well-muscled, with a large head, short legs, size and
appearance that varies considerably between the sexes, and is unique among the cats in living in family groups or
prides. [...]"
        },
        "similarity": {
                "levenshtein": 20.0,
                "dameraulevenshtein": 20.0,
                "jarowinkler": 0.52,
                "normalizedlevenshtein": 0.91,
                "searchstring": "lion",
                "lairstring": "Panthera leo (species)"
        }
}]
```


Figure 2. JSKOS+ output for a Getty AAT search string “lion”. [Florian Thiery, CC BY 4.0]

The JSKOS format defines a JavaScript Object Notation (JSON) structure to encode Knowledge Organisation Systems (KOS) (Voß 2021). JSKOS+ provides a standardised (Geo-)JSON(-LD) data model according to JSKOS which can be used in response to an API request. Figure 2 demonstrates the JSKOS+ JSON results for Getty AAT requests, showing the original JSKOS items in red and the added JSKOS+ items in blue.

The re3dragon applied in a research application

The re3dragon tool has been applied in the ARS3D project[13] (Thiery and Rokohl 2021; Thiery et al. 2023), in which it enriched data related to late Roman African Red Slip Ware (ARS). As an example, an ARS vessel is generically described as a "bowl" with decorative motives of "Hercules" and "Victoria" (O.39446[14]), which can be annotated using Getty AAT, IconClass and Wikidata items. The potform can be described as a bowl using Getty AAT item 300203596 and Wikidata Item Q15398[15], cf. fig. 3. The two features can be described as follows: Hercules: IconClass item 94L[16] (cf. fig. 4) and Wikidata item Q240679 according to (Zu Löwenstein 2015) "B / FT III" (Q110892402), (Armstrong 1993) "8.109" (Q110892540), and (Atlante 1981) "135" (Q110892520); Victoria: IconClass item 96A5 (VICTORIA) and Wikidata item Q308902 according to (Zu Löwenstein 2015) "N / FT III (Victoria)" (Q110892417) and Armstrong (1993) "8.100" (Q110892537).


```
{
  "displayLabel": {
    "en": "bowl"
  },
  "endDate": "",
  "prefLabel": {
    "de": "Schüssel",
    "ast": "concu",
    "en": "bowl",
    "hr": "Posuda",
    "eo": "bovlo",
    "it": "ciotola",
    "fr": "bol",
    "ta": "கிண்ணம்",
    "hu": "tál",
    "es": "cuenco",
    "hsb": "škla",
    "pl": "miska",
    "da": "skål",
    "ur": "[illegible]"
  },
  "narrower": [ … ], // 316 items
  "type": [ … ], // 1 item
  "lair": { … }, // 7 items
  "uri": "http://www.wikidata.org/entity/Q153988",
  "displayDesc": {
    "en": "round, open-top container frequently used as tableware"
  },
  "location": {},
  "broader": [ … ], // 2 items
  "startDate": "",
  "scopeNote": {
    "de": "Typ von Gefäßen",
    "ast": "caciu redondu",
    "en": "round, open-top container frequently used as tableware",
    "hr": "Predmet različitih oblika i veličina od raznih materijala u kojem se drži, priprema ili poslužuje hrana",
    "eo": "speco de ronda ujo",
    "it": "Contenitore per il cibo di forma semisferica",
    "fr": "élément de vaisselle",
    "ta": "அகன்ற மூடியில்லாத ஒரு பாத்திரம்",
    "hu": "ételek feltálalására alkalmas széles edény",
    "es": "recipiente semiesférico y sin asas",
    "hsb": "družina sudobjow",
    "pl": "naczynie",
    "da": "beholder",
    "ur": "[illegible]",
    "ca": "peça de vaixella"
  },
  "depiction": []
}
```


Figure 3. left: JSON+ Output for Wikidata item "Q15398" (bowl); right: scanned bowl O.39446 with Hercules and Victoria. [left: Florian Thiery, CC BY 4.0; right: ARS3D Project / RGZM / i3mainz, CC BY-SA 4.0]

[13] https://ars3d.rgzm.de (accessed 25/02/2022)

[14] https://ars3d.rgzm.de/object.htm?id=ars3do:eab38a5a-aaa2-41a1-b17b-0b91cbab006c (accessed 25/02/2022)

[15] [host]/re3dragon/rest/item?uri=http://www.wikidata.org/entity/Q153988

[16] [host]/re3dragon/rest/item?uri=http://iconclass.org/94L

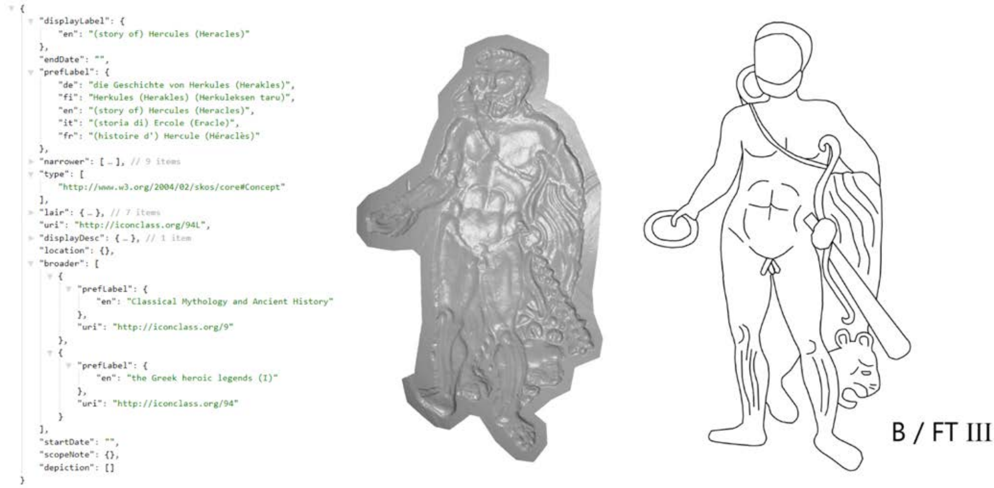

Figure 4. left: JSON+ Output for IconClass item "94L" (Hercules); middle: extracted Hercules feature from bowl O.39446; right: Figure B / FT III according to S. von Löwenstein [left: Florian Thiery, CC BY 4.0; middle ARS3D Project / RGZM / i3mainz, CC BY-SA 4.0; right: Sophie zu Löwenstein (2015, p.459)]

Outlook

During its usage in archaeological data enrichment projects, the re3dragon tool can be easily and demand driven enhanced by adding more Dragon Lairs. The tool is already embedded in the ongoing development of the so-called Meta-Index of the Römisch-Germanisches Zentralmuseum Mainz (Thiery et al. 2018; Thiery and Mees 2021).

This procedure enables future enhancements of the JSKOS standard, e.g. semantic location descriptions, mapping properties as well as space and time typologies (Dimensionally Extended 9-Intersection Model, Allen's Interval Algebra, etc.).

References

Armstrong, M.A., ed., 1993. *A Thesaurus of Applied Motives on African Red Slip Ware.* New York University: ProQuest.

Berners-Lee, T. 2006. "Linked Data - Design Issues". Last modified 18/06/2009. https://www.w3.org/DesignIssues/LinkedData.html.

Istituto della Enciclopedia Italiana. 1981. *Atlante delle forme ceramiche: Ceramica fina romana nel bacino mediterraneo.* Roma: Istituto della Enciclopedia Italiana.

Piotrowski, M., Colavizza G., Thiery F. and Bruhn K-C. 2014. "The Labeling System: A New Approach to Overcome the Vocabulary Bottleneck". In *DH-CASE II: Collaborative Annotations on Shared Environments: Metadata, Tools and Techniques in the Digital Humanities - DH-CASE '14, (1–6 September 2014).* Fort Collins, CA, USA: ACM Press. https://doi.org/10.1145/2657480.2657482.

Thiery, F. 2019. "The RSE Tools Ontology". Zenodo. https://doi.org/10.5281/zenodo.3541559.

Thiery, F. 2021. *Re3dragon - REsearch REsource REgistry for DataDragons* (version v0.1). Zenodo. https://doi.org/10.5281/zenodo.5338740

Thiery, F. and Engel T. 2016. “The Labeling System: The Labelling System: A Bottom-up Approach for Enriched Vocabularies in the Humanities”. In *CAA2015. Keep the Revolution Going. Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology.*, edited by Campana S., Scopigno R., Carpentiero G. , and Cirillo M., 259–68. Oxford: Archaeopress.

Thiery, F. and Mees A. 2021. “RGZM Meta-Index and Community-Driven Vocabularies in the LOD Cloud: Examples from the Department of Scientific IT”. https://doi.org/10.5281/zenodo.5764570.

Thiery, Florian and Allard Mees. 2023. “Taming Ambiguity - Dealing with doubts in archaeological datasets using LOD” in Human History and Digital Future. Proceedings of the 46th Annual Conference on Computer Applications and Quantitative Methods in Archaeology. http://dx.doi.org/10.15496/publikation-87762

Thiery, F., Mees A. and Heinz G. 2018. “RGZM-Meta-Index: A Central Linked Data Hub for Aligning Distributed Databases”. https://doi.org/10.5281/zenodo.2222237.

Thiery, F. and Rokohl L. 2021. “African Red Slip Ware Digital (ARS3D) - The Portal”. https://doi.org/10.5281/zenodo.5646897.

Thiery, F., M. Trognitz, E. Gruber, and D. Wigg-Wolf. 2019. “Hic Sunt Dracones! The Modern Unknown Data Dragons”. https://doi.org/10.5281/zenodo.3345711.

Thiery, F., , J.Veller, L. Raddatz, L. Rokohl, F. Boochs and A. W. Mees. 2023. “A Semi-Automatic Semantic-Model-Based Comparison Workflow for Archaeological Features on Roman Ceramics” *ISPRS International Journal of Geo-Information* 12, no. 4: 167. https://doi.org/10.3390/ijgi12040167

Schmidt, S. C., F. Thiery and M. Trognitz. 2022. “Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata” Digital 2, no. 3: 333-364. https://doi.org/10.3390/digital2030019

Voß, J. 2021. “JSKOS Data Format for Knowledge Organization Systems”. Github. https://gbv.github.io/jskos/jskos.html.

Voß, J., Ledl A. and Balakrishnan U. 2016. “Uniform Description And Access To Knowledge Organization Systems With Bartoc And Jskos’. https://doi.org/10.5281/zenodo.438019.

Zu Löwenstein, S. 2015. *Mythologische Darstellungen Auf Gebrauchsgegenständen Der Spätantike: Die Appliken- Und Reliefverzierte Sigillata C3/C4*. Kölner Jahrbuch 48.