CLaRMaV

Coastal Landscapes and Roman Maritime Villas: A Comparative Regional Study of Architectural Models and Environmental Settings

CLaRMaV. Data Management Plan. Version 1.0

Roberta Ferritto

DOI: https://doi.org/10.6092/unibo/amsacta/7800

CLaRMaV

Contract no. **101106554**

CLaRMaV

Coastal Landscapes and Roman Maritime Villas: A Comparative Regional Study of Architectural Models and Environmental Settings

Deliverable No.	D1.1
Deliverable Full title	Data Management Plan
Lead beneficiary (extended name and acronym)	Alma Mater Studiorum – Università di Bologna UNIBO
Authors (of the deliverable)	Roberta Ferritto
Planned delivery date	30/04/2024
Actual delivery date	26/04/2024
Dissemination level	PU = Public
Project website	soon to be published
Project start date and duration	Start date of project: 01.11.2023 Duration: 36 months

Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
v.01	15.01.2024	Roberta Ferritto	

Scheduled Data Management Plan (DMP) Updates

The DMP is a document that evolves during the lifespan of the project and registers all relevant changes in the life-cycle of all the research datasets. Updated versions of the DMP have already been planned. Moreover, this document will be updated whenever important changes in the data or the data management policy occur.

CLaRMaV

Content

CLaRMaV

The Data Management Plan (DMP)

This DMP provides details on all the research data collected and generated within the CLaRMaV project. In particular, it explains the way research data are handled, organized, licensed and made available to the public, and how they will be preserved after the project is completed.
This DMP reflects the current state of the art of the CLaRMaV project. However, the details and the final number of datasets may vary during the research project. The variations will be recorded in updated versions of this DMP.

1. Data Summary

CLaRMaV focuses on studying the different architectural models of Roman maritime villas developed in the Mediterranean basin, and how they were shaped by, and impacted on, the local environment. The project aims to compare these villa models from architectural, environmental, and socio-economic perspectives. This will involve identifying architectural models, trends, spatial functions, and the villas' role in the local economy. This project will enhance our understanding of the architectural, environmental, and socio-economic history of the Roman world, as well as gain insight into cultural developments and traditions in different regions. Additionally, the project aims to strengthen the connection between ancient and modern heritages, while also improving the management and preservation (or preserve for the future, at least in documented form) of coastal cultural heritage that is at risk of disappearing due to environmental and human activities.

Using a comparative approach, CLaRMaV will compare the architectural models of the Italian peninsula, where typologies are already established, with those of the Roman provinces. By doing so, the project seeks to create a comprehensive overview based on qualitative data derived from secure identification and sufficient documentation of villas found in documentary, literary, and archaeological sources. This comparative methodology will enable the project to build a network of relationships that can help address important unresolved questions. CLaRMaV will utilize archaeological data, documentary sources, GIS technology, and 3-D reconstructions to generate valuable insights.

The project's objective is to compare the villa-models from an architectural, environmental, and socio-economic perspectives, with three specific objectives and three sub-objects. The initial phase of the project involves collecting data from documentary and archaeological analysis, with selected information compiled into a GIS database using QGIS. This will be followed by the examination of villa plans and internal layouts to aid in typological identifications.

In some cases, precise measurements are essential for architectural analysis. When such measurements are not available from documentary sources, they are obtained through digitization and vectorization of plans using AutoCAD. This process involves importing raster files (e.g., JPEG, TIFF) of scanned drawings into AutoCAD, then digitizing and converting them into DWG format (vector file). Additionally, for archaeological remains that are at risk of destruction, 3-D models are created using ArchiCAD. These 3-D models are helpful for analysing the architecture of villas, forming structural hypotheses, providing an overview of the buildings, and calculating volumes. Once typologies are identified, the plans of provincial villas will be compared with those from the Italian peninsula to achieve the project's first specific objective. This comparison will help establish different architectural models, identify Roman influences in provincial domestic architecture, determine if local architectural forms inspired any models, identify regional variations in construction materials and techniques, recognize prevailing trends over time, understand the reasons behind design choices, and identify specific spaces used for important social and cultural activities locally. A GIS database, containing both geographical positions and non-spatial information about each site stored in an Excel table, will be used to differentiate sites, highlight their relationships, and create distribution maps illustrating regional variations and trends.

The second specific objective of CLaRMaV is to evaluate how domestic architecture was influenced by its environment, and vice versa. The impact of maritime villas on the environment (SO_2) is also analyzed with a sub-objective focused on the use of natural resources. To achieve this sub-objective, the study will analyze building materials (both local and imported), décor, and the use of rocky outcrops and marine rock-shelves for construction. This analysis will help estimate the level of investment made in building these villas and their impact on the local economy. Water usage is a key focus in this study, considering its current relevance and potential for waste or recycling measures. The research will identify luxury water facilities such as nymphaea and fishponds, as well as water pipes, canals, collection basins, and drainage systems. The type of water supply (e.g. rainwater, wells, springs, or rivers) will also be examined. 3-D models created in the first phase of the Fellowship will help estimate the average water capacity of selected basins, with height measurements obtained from archaeological data or when not available, postulated based on comparisons with other structures. Field visits will be conducted to address unique structures where measurements are difficult to estimate through comparisons. Case studies will assess total water consumption to determine the impact of maritime villas on water resources in the area. This methodology, which combines archaeological data, 3-D reconstructions, and modern engineering calculations, can be applied to other studies in both private and public architecture.

The third specific objective is to determine the role of maritime villas in the local economy by examining their production activities. Special attention will be focused on one of business in which villa owners could have been involved: fish farming. CLaRMaV will explore the potential labour costs associated with constructing fishponds and their capacity for production. This research goal will be accomplished through an assessment of villa production activities and their impact on the local economy, analysing previously gathered archaeological data for SO1, and studying 3-D models of artificial fishponds. These fishponds were commonly utilized for intensive fish farming, and to measure their production capacity, the researcher will apply a methodology like one used in the study of villas in Southern Latium. Utilizing the 3-D models, water volume within each basin and the construction material used for fishpond walls, typically hydraulic concrete, can be determined. Production potential calculations will be conducted following the approach outlined by Marzano and Brizzi (2009), with labour estimates based on the work of Hohlfelder and Brandon.

The proposal seamlessly integrates elements of architecture (plans, 3-D models), general archaeology, and geography (digital mapping, GIS).

The main datasets of Dr Ferritto's project will consist of textual research, a GIS database, distribution maps generated by the database, graphs/diagrams describing the typology of maritime villas, 3-D models, digitalization of plans and production of 3-D.

Data will be generated in the following formats:

- Shapefiles for spatial data (.shp)
- PNG and SVG for graph data
- DWG files (.dwg) for 2-D drawings and 3-D models
- JPEG and TIFF for graphic image files (.jpg, .tiff)
- Publications, presentations (ppt., txt., pdf files) and research data will be made open access (papers will follow the green or gold path according to funding availability) to increase the dissemination of these data and their use by other researchers interested in related topics.

The project will reuse data from the following sources: scientific publications; archival documentations; archaeological site documentation (e.g. plans, field drawings, photographs, reports), and spatial documentation (Google Earth Pro coordinates). Within the project existing data in .txt, .xlsx, .tiff or .JPEG formats will be used. The data in .txt and .xlsx formats are available in the research and database created by Dr Ferritto for the GIS on the maritime villas of Southern Latium (Italy), which she produced for her PhD thesis, titled: *The Roman Maritime Villas of Southern Latium: Construction, Design, and connection to the Sea (Volume I). Appendix II:*

Catalogue of the Sites (Volume II). Part of these data will be re-used to create a relational geodatabase and to identify existing villa typologies. The archaeological data (plans or images of villas, structures, objects, etc.) in .tiff or .JPEG formats are available in libraries image collections, such as the Penn Image Collections (https://guides.library.upenn.edu/images/penn), the British School at Rome Digital Collections (https://digitalcollections.bsr.ac.uk/collections), the American Academy in Rome Library's Digital Humanities Center (https://dhc.aarome.org/), or in Open Access Database, such as the atlas of Roman building techniques (https://acor.huma-num.fr/projet/). These collections provide subscribers with access to high resolutions files to use in scholarly publication, under the permitted uses. These data will be re-used to create a relational geodatabase, to identify the existing villa typologies, make analysis on dimensions of site spaces, and create 3-D models.

The project will also produce different types of data by using different methodologies:

1. Textual data (.txt) will be generated through the interpretation of archaeological data, existing documentary sources, and from the re-analysis of existing dataset. The project aims to give an overall picture based on qualitative data, i.e. villas with sufficient existing documentation derived from documentary, literary and archaeological evidence.

2. Shapefiles (.shp) will be created in QGIS software, by using data obtained during the collection phase. The GIS will contain both geographical positions and non-spatial information about each site.

3. Coordinates data. These data will be used to create distribution maps (i.e. graphic files) as well as graph data (PNG, SVG). To create distribution maps, the points, i.e. the geographical coordinates of the location of each villa under examination, will be placed on a Google Maps Pro basemap, by using the QuickMapServices (QMS) plugin installed in QGIS software. The QuickMapServices plugin, along with QGIS, is a free and open-source software that enables users to create, edit, visualise, analyse, and publish geospatial information. The QuickMapServices plugin is very useful for CLaRMaV project because it accesses satellite images, OpenStreetMap, and various other repositories directly from GQIS. The QuickMapServices software is distributed under the Creative Commons 3.0 licence, which is open access and open source.

4. Graphic image files (.jpeg, .tiff), such as photographs, will be generated also from surveys.

5. Quantitative data will be generated from the analysis of 2-D plans and 3-D models.

6. Measurements and observational data will be also taken during surveys.

The research team have agreed to convert research data from proprietary formats to well-known and documented open formats in order to facilitate accessibility and reusability (Tab.1).

Table 1 - Summary of data formats

Type of data	Formats used during data processing	Formats for sharing, reuse and preservation
Textual data	MS Word (.doc/.docx)	.txt
Qualitative data	MS Word (.doc/.docx)	.txt, .pdf
Geospatial data	Shapefile (.shp), Geo-referenced TIFF (.tiff)	.shp, .tiff
Vector and raster data	CAD data (.dwg), Geo-referenced TIFF (.tiff) or JPEG (.jpg), Adobe Photoshop (.psd) and Adobe Illustrator (.ai)	.dwg, .tiff
Digital image data	Tiff (.tiff), JPEG (.jpg), Adobe Photoshop (.psd), Adobe Illustrator (.ai), Adobe Portable Document Format (PDF, .pdf), RAW image format (.raw)	.tiff, .jpeg, .pdf
Documentation	PDF (.pdf)	.pdf
Excel table	XLS	.csv
Graph data	PNG, SVG	.jpg or .JPEG

The expected size of the data is currently not known, but likely to be up to 20 GB. Considering the early stage of the project, the effective size may vary with respect to what is declared here. Potential variations will be addressed in further versions of this document.

Data will be a crucial resource for researchers investigating a wide range of topics including architecture, environment, society, economy, water recycling, Roman archaeology, coastal cultural heritages, climate change, and the human behaviour towards nature over time.

2. FAIR Data

2.1 Making data findable, including provisions for metadata

CLaRMaV will make available sources (including the GIS and 3-D models), analytical methods and results, fulfilling in Open Science of the European Commission. These include papers, presentations, and digital data. To improve the findability of research data produced during the project, datasets will be deposited in the AMS Acta, the Institutional Research Repository of the

University of Bologna, if and when appropriate (trusted data repositories if and when appropriate).
The AMS Acta is the official research data repository of the researchers of the University of Bologna maintained by the AlmaDL University of Bologna Digital Library. AMS Acta allows to publish sets of research data in a FAIR way and following the guidelines of the European Open Science Cloud (EOSC), which includes DOIs identifiers in FAIR sharing (a resource on data and metadata standards, related to databases and data policies).
The AMS Acta repository attributes a unique persistent identifier (PID) to the deposited data, i.e. DOI identifiers. The DOIs are then used to cite the datasets within all research publications. All published papers will furnish links to the data employed, and in addition these data will be stored in AMS Acta repository to guarantee their discoverability, access, and preservation beyond the project end. These will be adequately tagged and described using metadata annotations. AMS Acta's objects metadata is indexed in OpenAIRE and, therefore, machine-readable.
In addition, whenever project results are published, the researcher deposits and describe the relative underlying datasets in trusted data repositories to guarantee their discoverability, access, and preservation beyond the project end.
The repository (see section 2.2, Table 2) attributes a unique persistent identifier (PID) to the deposited items (DOI). The unique identifiers are then used to cite the datasets within all research publications.
It also supports standard descriptive metadata to ensure datasets indexing and discoverability.
It is planned to provide search keywords to make the data findable and to follow best practices for naming conventions.
Provisional keywords:

- Maritime villas
- Architecture
- Domestic architecture
- Coastal Landscape
- Roman Provinces
- Ancient Society
- Ancient Economy
- Water recycling
- Classical archaeology
- Network analysis
- 3-D Models

As mentioned above, to improve data visibility, discoverability, citation, and permanent online tracking, naming conventions will be followed, giving dataset titles, e.g.:

- *CLaRMaV. Roman Maritime Villa Typologies in the Mediterranean*

Examples of dataset file names used for the project; are:

- *CLaRMaV.01_DATABASE_MARITIME_VILLAS_04.04.2024_ver01.csv*

- *CLaRMaV.01_DATABASE_MARITIME_VILLAS_04.04.2024_ver01 _README.txt*

2.2 Making data accessible

As a guiding principle, CLaRMaV seeks to make all research data openly available as soon as possible and ensure open access — via the repository — to allow dissemination, validation, and re-use of research results.
To this purpose, all possible and legitimate actions and strategies are adopted to allow data sharing including:

- converting the files to standard open formats;
- providing all relevant documentation and explanation for the data and the datasets;
- in case of copyright on raw data derived, collected, or elaborated from pre-existing databases or from other original sources (i.e. papers, journal articles, book chapters, reports, video and audio sources), collected data will be made available if the reproduction and sharing are allowed by expressed permission of the right holders or by applicable copyright exceptions and exemptions. Otherwise, only aggregate data resulting from the analysis will be openly published.

Restrictions to access are applied only in the following cases:

- collected data belong to third party which have denied permission for sharing on account of confidentiality and proprietary issues;
- data availability would jeopardize the project's main aim.

See Annex I for details on the accessibility of each dataset. In all cases, metadata will be made openly available and licenced under a “No Rights Reserved” CC0 license or equivalent, as per the Grant Agreement, and will contain information on how to access the data.
The data repository chosen (i.e. the AMS Acta) guarantee long term preservation and attribute persistent unique identifiers to the archived datasets. They support open licenses and different access levels. Finally, they adopt descriptive metadata standards such as Dublin Core and

DataCite Metadata Schema, as required by the OpenAIRE guidelines[1], and allow cross-linking between publications and the relevant datasets. Please see the table below for more detail.

Table 2 – Summary of repositories.
The following table shows the repositories for datasets publication and preservation.

Repository name	Type	URL	PID	OpenAIRE compatibility?
AMS Acta	Institutional	https://amsacta.unibo.it/	DOI	Yes

2.3 Making data interoperable

Interoperability will be ensured by furnishing data in open and standard formats that can be automatically incorporated into workflows (CSV for tabular data, Shapefiles for spatial data).
As mentioned above, I will convert all shareable data from proprietary formats to well-known and documented open formats (see section 1, Table 1). This allows data exchange and re-use between different researchers, institutions, organisations, and countries.
Any publication based on the data will be mentioned in later versions of this DPM and will contain a data availability statement. The use of README files (i.e. the textual document that forms part of the dataset and provides informative documentation needed for the understanding of the dataset) is also an important element towards the interoperability. README files will be compiled as required by the AMS Acta repository.
All datasets will be described using standard descriptive metadata, to ensure metadata interoperability for indexing and discoverability. For each deposited dataset, relevant documentation explaining data collection procedures and analysis is made available along with the data, to guarantee intelligibility, reproducibility, and the validation of the project findings.

2.4 Increasing data re-use

Whenever possible, data developed by CLaRMaV will be made open access under the CC-BY license to maximize distribution and economic impact while also protecting author attribution.
As per Grant Agreement, metadata will be open available under a Creative Commons “No Rights Reserved” CC0 licence or equivalent.

[1] OpenAIRE, https://guidelines.openaire.eu/en/latest/

The quality of data will be carefully assured using different approaches, such as:

- the use of procedures for verification, such as network checks and automated simple route modelling
- Data conform to format specification
- Consistency verified with data models and standards.

3. Allocation of resources

No extra storage will be necessary during the research period as the capacity of UNIBO's infrastructure is sufficient. Therefore, there is no need for extra expenditure on storage.
Depositing into AMS Acta is free for the researchers of the University of Bologna and covers the cost of preservation for at least 10 years. There is no additional expense required for long-term preservation or archiving. As part of UNIBO's commitment to ensuring FAIR data, all researchers are responsible for preparing their own Data Management Plans (DPMs) in accordance with UNIBO's Research Data Management Policy. Therefore, there is no need for extra expenditure on data management. Making data FAIR requires an investment of money and researchers' time.

Responsibility for data management sits with dataset creator (see Table 4).

Table 5 identifies all contributors participating in data management activities and specifies their roles.

Table 3 – Summary and contacts of people responsible for data management

Name	ORCID (if available)	E-mail address
Roberta Ferritto	0000-0002-5059-6063	roberta.ferritto@unibo.it

Table 4 – Summary of team members involved in the datasets collection and management.

Name	ORCID (if available)	Role
Roberta Ferritto	0000-0002-5059-6063	Data Collector, Data Curator, Metadata Generator, Storage and Backup (during the Fellowship)
Annalisa Marzano	0000-0002-6485-9143	Supervisor (UNIBO, Host Institution)
Mantha Zarmakoupi	0000-0003-1222-0169	Supervisor (UPenn, Associated Partner)
Roberta Ferritto	0000-0002-5059-6063	Storage and Backup (during the Fellowship)
UNIBO		Long-term preservation (after completion of the project)

See Annex I for details about data management responsibilities related to each dataset.

4. Data security

During the development of the project, research data stored in the researcher's laptop and hard drives are accessible only after logging in with a username and password (periodically modified according to national law provisions for data security) and are protected by updated antivirus software. They are also regularly backed up to prevent accidental losses. None of the project data will ever be left inadvertently available. Data are also saved in the OneDrive space provided by UNIBO. Long term preservation of public data is ensured by the AMS Acta data repository with specific preservation policies. If external devices are utilized to store data files (e.g., backup files), they will be securely stored in a location accessible only to the relevant researchers or encrypted with specialized software.

5. Ethical or legal aspects

All activities carried out under the CLaRMaV project adhere to ethical principles and relevant international, EU, and national laws. This includes compliance with the European Code of Conduct for Research Integrity and all associated fundamental ethical principles, as well as those outlined in the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols. Additionally, the project follows rules for the protection of personal data both inside and outside the EU.

6. Other issues

The CLaRMaV is a HORIZON-TMA-MSCA-PF-GF project funded by the European Union. It is carried out in the Department of Storia Culture Civiltà (DISCI) of the Alma Mater Studiorum - Università di Bologna (UNIBO) in the role of Coordinator, and in the Department of History of Art of the University of Pennsylvania (PENN) in the role of Associated. According to UNIBO's policy, all researchers must prepare a data management plan for the project. Data results must be disseminated according to FAIR and Open data protocols.
UNIBO has a Research Data Management Policy:

https://www.unibo.it/en/attachments/university-policy-on-research-data-management/@@download/file/The-University-Policy-on-research-data-management.pdf

Annex I: Datasets

The analytic description of each expected dataset of the CLaRMaV project is included in this Annex.

Dataset number	**Ready at month of project**	***Dataset title***
Status		
ID [ID type]		
Version		
Creator/s		Family name, given name [TEAM]; …
Contributor/s		Family name, given name [TEAM]; …
Contact Person/s		Family name, given name [TEAM, email]; …
Contents		
Data format		
Data volume		
Accessibility		
Related publication/s		

Annex II: Open Access status of project publications

In the following table we describe the open access status of the project publications and the underlying datasets.

Table 5 – Publications and related datasets.

Publications	
Bibliographic citation of the publication	
Link to copy archived in repository	
Related dataset/s	
Bibliographic citation of the publication	
Link to copy archived in repository	
Related dataset/s	
Bibliographic citation of the publication	
Link to copy archived in repository	
Related dataset/s	
Bibliographic citation of the publication	
Link to copy archived in repository	
Related dataset/s	
[…]	

CLaRMaV

Annex III: "README" file

A "README" file is a document that will be deposited with each dataset, containing relevant information about dataset authorship, terms of reuse and responsibilities, explaining dataset content and structure, collection procedures and analysis (such as file specifics, methodologies, codebooks of variables, data sources, and further necessary notes).

This is a template of the README file that we will use.

README file

Dataset Title: "[insert title as defined in the DMP]"
Dataset Author/s: Name Surname (Affiliation), ORCID (if available);
Dataset Contributor/s: Name Surname (Affiliation), ORCID (if available);
Dataset Contact Person/s: Name Surname (Affiliation), ORCID (if available), email;
Dataset License: this dataset is distributed under a [insert LICENSE]
Publication Year: [insert YEAR]
Project Info: [insert PROJECT ACRONYM] ([project full title], funded by European Union, Horizon 2020 Programme. Grant Agreement num. [insert grant agreement number]; [insert project website url]

Dataset Contents

The dataset consists of:
[EXAMPLE 1

- 1 textual qualitative file saved in .rtf format: "ProjectAcronym_WP3_T3-2_ItalyInterviews_20161221_v01.rtf"
- 1 README file: "README_ProjectAcronym_WP3_T3-2 _ItalyInterviews_20161221_v01.rtf"

EXAMPLE 2

- 1 tabular quantitative file saved in .csv format: "ProjectAcronym_WP7_T7.3_Questionnaire_Sweden_20170905.csv"
- 1 README file: "README_ProjectAcronym_WP7_T7-3_Questionnaire_Sweden_20170905.txt"]

Dataset Documentation

Abstract:
[Insert dataset abstract]

Content of the files:

- file [Insert filename] contains ...
- file [Insert filename] contains ...
- …

File specifics
[Please indicate instruction/technical info in order to allow potential users to correctly visualize and reuse your data (e.g. specific software, …). In case of data converted in open formats it could be useful to provide some further information. For example, if you deposit for long term preservation a .csv file derived from an excel you can describe the conversion. Here is an example of description of conversion using libre office calc software:
To create the .csv files, "LibreOffice Calc" version: 5.1.4.2 (portable) was used, with the following specifics:
•Character set Europa occidentale (Windows-1252/WinLatin1)
•Field delimiter « , » (comma)

•Text delimiter « “ » (quotes)]

Notes
[Related to the whole dataset or to single files of a multi-file dataset (Optional)]

Data sources
[Optional]

Methodologies
[If necessary to understand how to reuse data]

Codebook of variables
[If necessary to understand the meaning of the variables]