DOMES

**Architectural Technology Transfer on the Silk Road:
Iranian Double-Shell Domes and the West: 14th to 20th Century**

Data Management Plan

Lorenzo Vigotti

Version 1.0
1 December 2023

https://doi.org/10.6092/unibo/amsacta/7516

Deliverable No.	1
Deliverable Full title	Data Management Plan
Lead beneficiary (extended name and acronym)	Architectural Technology Transfer on the Silk Road: Iranian Double-Shell Domes and the West: 14th to 20th Century (DOMES)
Authors (of the deliverable)	Lorenzo Vigotti
Planned delivery date	01/02/2024 (M6)
Actual delivery date	dd/mm/yyyy (M)
Dissemination level **(PU = Public; PP = Restricted to other program participants; RE = Restricted to a group specified by the consortium; CO = Confidential, only for members of the consortium)**	PU
Project website	www.iraniandomes.eu
Project start date and duration	Start date of project: 1 September 2023 Duration: 36 months

The information in this document reflects only the author's views. The Horizon Europe of the European Commission is not responsible for any use that may be made of the information it contains. The user therefore uses the information at its sole risk and liability.

Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
1.0	01/12/2023	Lorenzo Vigotti	New file

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.

Content

The Data Management Plan (DMP)

This DMP provides details on all the research data collected and generated within the DOMES 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 DOMES 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

The aim of the project DOMES focuses on the transmission of two types of architectural knowledge: technical knowledge and preservationist knowledge. This project first proposes to investigate the transmission of the technological knowledge of building double-shell brick domes from Iran to Italy. The data generated will be retrieved from on-site architectural analysis (3D laser scanners, photogrammetry, thermographic analysis) and archival research (documents about past restorations). They will be crucial components for the full understanding of double-dome brick structures in medieval monuments in Iran.

The project will reuse data from the Sanpaolesi archive at the Iranian Centre for Architectural Studies and Documentation at the School of Architecture, Shahid Beheshti University in Tehran, Iran. Reuse has been favoured by the director of the archive, prof. Hadi Safaeipour, who will grant access to the documents and drawings held there.

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

1. Digital images of original documents from 1960s-1979 held at the Iranian Centre for Architectural Studies and Documentation at the School of Architecture, Shahid Beheshti University in Tehran, Iran;

2. Graphical data from 3D laser scanners (DWG models), photogrammetry, and thermographic cameras used on-site on selected documents;

3. Textual data from the comparison of archival documents and on-site 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
Digital images	JPEG	JPEG
Scanned documents	PDF	PDF
3D laser scans	LGS	E57
3D architectural models	DWG	DWG
Photogrammetric data	TIFF	TIFF
Textual data/documentation	DOC	PDF

README files[1] explaining all relevant details regarding data collection, processing methodologies and quality assurance will be deposited along with the datasets in .odt, .rtf or .pdf format.

The expected size of the data is 1 Tb. 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.

The generated data holds potential interest for various user groups, encompassing architects, conservators, and engineers engaged in the planning and execution of restoration projects for similar historical monuments in Iran. National and local politicians, along with UNESCO administrators, will utilize the data in shaping policies for the tourism and cultural development of these sites. The produced data will serve as a valuable resource for architectural historians, art historians, historians of science, as well as other scholars and students, supporting their respective projects and providing a means to further test the new methodology introduced by DOMES in the context of global history. Additionally, the general public will find value in gaining a new perspective on cross-cultural exchanges along the Silk Road during the Middle Ages.

2. FAIR Data

2.1 Making data findable, including provisions for metadata

To improve the findability of research data produced during the project, dataset will be deposited in trusted data repositories such as University of Bologna's institutional data repository, AMSacta (https://amsacta.unibo.it/) if and when appropriate. In addition, whenever project results are published, the team members deposit and describe the relative underlying datasets in trusted data repositories in order to guarantee their discoverability, access and preservation beyond the project end.

The chosen repositories (see section 2.2, Table 2) attribute a unique persistent identifier (PID) to the deposited items. The unique identifiers are then used to cite the datasets within all research publications.

They also facilitate the implementation of standard descriptive metadata to guarantee the indexing and discoverability of datasets. The descriptive metadata encompasses details such as the author's name, the date and place of file creation, the file's subject, and carefully chosen keywords to streamline the search for pertinent files. Furthermore, it includes information such as the project name and the name of the sponsor (EU). This metadata is designed to be accessible for harvesting and indexing purposes.

Specific keywords will be associated to each dataset to enhance semantic discoverability. For example, keywords referring to architectural details will be provided in three languages (English, Farsi, Italian), because different languages often developed a distinctive way to describe these, not directly translatable.

Research data are organized in datasets, which are named collections of data units with the same focus and scope. In this DMP we set out common rules for dataset naming, in order to improve data visibility, discoverability, citation and permanent online tracking.

---

[1] A "README" file is a document 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). (See Annex II to visualize the suggested README file template).

The recommended dataset title structure consists of:

DOMES. Task title or description. Additional information (if necessary). Version number

Example:

DOMES. 3D scan. Dome in Soltaniyeh. 1.1

The version number of the dataset will be added at the end of the title in case of data revisions to help identifying the dataset updates especially in repositories that do not track versioning automatically (see Annex I).

The following rules for file naming will be followed:

- for dataset file(s)

DOMES_TaskNumber_Coverage or other content specifications_Date(YYYYMMDD)_VersionNumber.fileExtention

Example:

DOMES_01_3Dscan_Soltaniyeh_20240101_01.e57

- for README file(s)

DOMES_TaskNumber_Coverage or other content specifications_Date(YYYYMMDD)_VersionNumber_README.fileExtention

Example:

DOMES_01_3Dscan_Soltaniyeh_20240101_01_README.txt

2.2   Making data accessible

As a guiding principle, DOMES seeks to make all research data openly available as soon as possible and ensure open access — via the repository — in order 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;
- obtaining copyright permissions from third party data owners to be allowed to re-use, reproduce and distribute the collected data. Specifically, Shahid Beheshti University, the owner of the right of the archival collections I will be using in Tehran will allow for taking the pictures to be used for the scheduled publications; discussions are ongoing to open the possibility to deposit the acquired images of photos and documents on the open access depository;
- in case of copyright on raw data derived, collected or elaborated from pre-existing databases or from other original sources (i.e. archives, 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. When the sources are freely available on-line in their original repositories, but direct reproduction is not allowed, a detailed account on how the dataset was created from the original data will be provided, together with the specification of open repositories from where the original datasets are available. Raw data consisting in full texts will not be made available without the copyright holders' permission.

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.

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 repositories chosen 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[2], 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?
AMSActa	Institutional	https://amsacta.unibo.it/	DOI	YES

2.3   Making data interoperable

All datasets produced by DOMES will be described using standard descriptive metadata, in order 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, in order to guarantee intelligibility, reproducibility and the validation of the project findings.

As mentioned, the team 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.

Datasets produced during DOMES are new, but they often referred to the following previous publications on medieval Persian architecture, which include architectural surveys of monuments included in the project:

---

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

Mehrdad Hejazi, *Persian Architectural Heritage. Architecture* (WIT Press, 2014)
Arthur Pope, *Persian Architecture: the Triumph of Form and Color* (Braziller, 1965)
Arthur Pope, *A Survey of Persian Art from Prehistoric Times to the Present* (Oxford University Press, 1938-2005)

Piero Sanpaolesi, *Progetto di restauro del mausoleo di Olgietu a Soltanieh* (Teheran: National University of Teheran, ca. 1972).

Additional publications based on the data will be mentioned in later versions of this DMP and will contain a data availability statement.

If specific software is used during data processing, full explanation and instructions will be included in the deposited documentation. See Table 3 below for a summary of the tools and software necessary to access and/or reuse our data.

Table 3 – Summary of tools and software for enabling re-use of the datasets

Tools/software
Open source versions of Google's Chrome and Firefox browsers can be used to read PDF files. For downloaded files, **xpdf** can be used.
Open source **FreeCAD** software can be used to read Autocad's DWG files.
Open source **GIMP** can be used to access and edit all image files (JPG, JPEG, TIFF, etc…)

2.4 Increasing data re-use

DOMES license data under CCBY 4.0 license.

As per Grant Agreement, metadata will be open available under a Creative Commons "No Rights Reserved" (CC0) license or equivalent.

The quality of the data will be carefully assured using different approaches. The principal investigator will put in place quality control processes for: multiple representations, duplicate records, redundant or combined data, and other inconsistencies or erros. Control processes will include systematic steps to clean the data, check the data against previous versions, and reconcile problems and errors that arise in the data. This will be done on a monthly basis while the data is collected and then performed at the end of the study with the full data set. Throughout the project, historical versions of the data set will be kept in order to maintain a record of the control processes.

3. Other research outputs

No other research outputs are expected.

4. Allocation of resources

Making data FAIR requires an investment of money and researchers' time.

Data processing: An Iranian doctoral student from the School of Architecture at Shahid Beheshti University will provide transcription and translation for the documents in handwritten Farsi fund

in archives in Tehran. Expected cost is €15/hour for an expected 26 hours/month for the first year (€4,000-5,000).

Data storage: An estimated 1Tb (1,000 Gb) of data will be stored, for an expected cost of €3,500.

Data long-term preservation: Long term preservation of the data will be ensured by the longevity of UniBo's open access platform AMSActa. The data that will be kept on the AMSActa platform will include the qualitative and quantitative datasets created by the project. Because AMSActa is free, it is not expected that the preservation of the dataset on the AMSActa platform will require any additional costs.

Responsibility for data management usually sits with each dataset creator (generally the team leader, but see Table 4 below). Table 5 identifies all contributors participating in data management activities and specifies their roles.

Table 4 – Summary and contacts of people responsible for data management

Name	ORCID (if available)	E-mail address
Lorenzo Vigotti	0000-0002-3573-0949	lorenzo.vigotti@unibo.it

Table 5 – Summary of team members involved in the datasets collection and management.

Name	ORCID (if available)	Role
Lorenzo Vigotti	0000-0002-3573-0949	Data collector and producer
Doctoral student (to be hired)		Researcher (transcription/translation)

5. Data security

During active data management (e.g., during data collection and analysis), research data stored in computers, laptops, intranets or hard-drives are accessible only after logging in with username and password (periodically modified according to national law provisions for data security) and are protected by updated antiviruses. They are also regularly backed-up in order avoid accidental losses. None of the project data will ever be left inadvertently available. At any moment, two external devices are used to store data files (backup files) and they are be kept in two separate safe places accessible only to the researchers involved.

Because the severe limitation in accessing and editing files in cloud storage websites, none will be use during the outgoing phase in Iran.

Long term preservation of public data is ensured by the AMSActa data repository that has specific preservation policies.

6. Ethical or legal aspects

No ethical or legal issues are present because no personal data is included in the project, and materials are not copyrighted because Iran is not a signatory to the Berne Convention for the Protection of Literary and Artistic Works or the WIPO Copyright Treaty, or a member of the World Trade Organization (WTO).

7. Other issues

I will not make use of other national/funder/sectorial/departmental procedures for data management.

Annex I: Datasets

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

Dataset number	Ready at month of project	Dataset title
1	**12**	***Sanpaolesi archive at Shahid Beheshti University***
Status		***In progress***
ID [ID type]		
Version		
Creator/s		Vigotti, Lorenzo
Contributor/s		Vigotti, Lorenzo
Contact Person/s		Vigotti, Lorenzo [lorenzo.vigotti@unibo.it]
Contents		Documents, architectural drawings, photos, negatives, publications by Piero Sanpaolesi and his team produced in Tehran during 1969-1979.
Data format		JPG, PDF
Data volume		5Gb
Accessibility		Not open available yet (embargo awaiting publication scheduled 2025)
Related publication/s		Book project scheduled in 2025, title TBA

Dataset number	Ready at month of project	Dataset title
2	24	*Architectural surveys of Iranian monuments*
Status		*Not yet available*
ID [ID type]		
Version		
Creator/s		Vigotti, Lorenzo
Contributor/s		Vigotti, Lorenzo
Contact Person/s		Vigotti, Lorenzo [lorenzo.vigotti@unibo.it]
Contents		3D scans, photogrammetry, laser scanner surveys of five Iranian monuments with double-shell brick domes
Data format		TIFF, E57, DWG
Data volume		1Tb (1,000 Gb)
Accessibility		Not open available yet (embargo awaiting publication scheduled 2026)
Related publication/s		Book project scheduled in 2026, title TBA

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 7 – 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	
[...]	

Annex III: "README" file template

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]