

# VAMP Data Management Plan, Version 1

Carlo Andrea Tassinari

DOI: <https://doi.org/10.6092/unibo/amsacta/8107>

This work is licensed under Creative Commons Attribution 4.0 International (CC BY 4.0). To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Contract no. 101106065

**VAMP****Voices from the Anthropocene. Maps and Frameworks for Ecological Conflict**

<b>Deliverable No.</b>	2
<b>Deliverable Full title</b>	Data Management Plan
<b>Lead beneficiary (extended name and acronym)</b>	Università di Bologna, Unibo
<b>Authors (of the deliverable)</b>	Carlo Andrea Tassinari
<b>Planned delivery date</b>	31/03/2024 (M6)
<b>Actual delivery date</b>	31/03/2024 (M6)
<b>Dissemination level</b> (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
<b>Project website</b>	
<b>Project start date and duration</b>	Start date of project: 01/10/2023 Duration: 36 months

*The information in this document reflects only the author's views. 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.*

[insert project logo]

## Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
1	30/03/2025	Carlo Andrea Tassinari	

## 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) .....	1
1. Data Summary .....	1
2. FAIR Data .....	2
2.1 Making data findable.....	2
2.2 Making data accessible.....	3
2.3 Making data interoperable.....	4
2.4 Increase data re-use .....	4
3. Other research outputs .....	5
4. Allocation of resources.....	5
5. Data security.....	6
6. Ethical or legal aspects .....	7
7. Other issues .....	7
Annex I: Datasets.....	8
Annex II: Open Access status of project publications .....	9
Annex III: “README” file template .....	10

## The Data Management Plan (DMP)

This DMP provides details on all the research data collected and generated within the VAMP 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 VAMP 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 is to elaborate a semiotic framework for analysis and resolution of ecological controversies. It focuses on the place and meanings of two main controversial topics in public debate: the uses of the Amazon rainforest during Bolsonaro’s government in Brazil and the future of the territories concerned by the Lyon-Turin railroad infrastructure.

The project will reuse data from the following sources: National Archives (Italian, French and Brazilian), Media Archives (On-line Newspapers, Television Platforms), Institutional and Grassroots Websites. Reuse has been favored as ecological controversies have vocation to deploy through public supports, created by actors involved in the matter of interest. Among these, the project distinguishes dataset created by actors themselves (i.e., a press review on the matter created by a stakeholder, a folder created by an archivist, a column dedicated to the subject by a newspaper) and dataset created by the researcher (i.e. a file or a folder containing a press review reorganizing existing articles, a montage of existing photographs or videos, a collection of graphs and plans related to the controversy).

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

1. Digital Watch Notes
2. Participant observation
3. Semi-directive interviews
4. Audio registration
5. Photography and Video registration

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
Qualitative data. Textual.	MS Word (.doc/.docx).	Adobe Portable Document Format (PDF/A, PDF) (.pdf).
Digital image data.	PEG (.jpeg, .jpg) TIFF (other versions) (.tif, .tiff). Adobe Portable Document Format (PDF/A, PDF) (.pdf). PNG (.png)	PEG (.jpeg, .jpg) TIFF (other versions) (.tif, .tiff). Adobe Portable Document Format (PDF/A, PDF) (.pdf). PNG (.png)
Digital audio data.	MPEG-1 Audio Layer 3 (.mp3)	MPEG-1 Audio Layer 3 (.mp3)
Digital video data.	MPEG-4 (.mp4).	MPEG-4 (.mp4).

[insert project logo]

README files<sup>1</sup> 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 data produced can be of interest to different potential users. They may include researchers involved or interested in environmental humanities, anthropologists, sociologists, semioticians, and philosophers. They may also include a wider audience, such as environmental activists, public administrators, journalists, designers, decision-makers, private R&D departments, and in general any actor involved in societal changes for ecological redirection of practices.

## 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 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. Please note, however, that current datasets do not have a specific PID yet. The unique identifiers are then used to cite the datasets within all research publications.

They also support standard descriptive metadata to ensure datasets indexing and discoverability. These will include project's acronym, the title and a short description of purpose and content of the dataset, the grant project, the date of deposit, author's name, the venue (if applicable), the licensing terms, the DOI and URI identifier attributed by the repository.

Specific keywords will be associated to each dataset to enhance semantic discoverability. Given Keywords for description of dataset's type (i.e. press cut, interview, grey literature, video recording), institutional actors involved (Government, Grassroot associations, Archives, Research Centres), and main discipline of interest (Semiotics, Environment, Cultural Studies) will be chosen preferably from the UNESCO thesaurus (<https://vocabularies.unesco.org/browser/thesaurus/fr/?clang=en>). In absence of a disciplinary thesaurus for Semiotics or Environmental Humanities, I will refer to keywords in use in the most internationally recognized Semiotics and Environmental Humanities Journals, such as *Semiotica*, *VS* or *Environmental Humanities*.

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.

The recommended dataset title structure consists of:

*PROJECT ACRONYM. Task title or description. Additional information (if necessary). Version number*

---

<sup>1</sup> 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).

[insert project logo]

Example:

*VAMP. Lyon-Turin Project History. Courriers parlementaires 1989-1991.  
Version 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).

This DMP also recommends the following rules for file naming:

- for dataset file(s)

*[PROJECT ACRONYM]\_TaskNumber\_Coverage or other content  
specifications\_Date(YYYYMMDD)\_VersionNumber.fileExtention*

Example:

*VAMP\_1\_Lyon-Turin\_Press Cuts 1989\_12-02-2024\_Version 1.pdf*

- for README file(s)

*[PROJECT ACRONYM]\_TaskNumber\_Coverage or other content  
specifications\_Date(YYYYMMDD)\_VersionNumber\_README.fileExtention*

Example:

*VAMP\_1\_Lyon-Turin\_Dossier de presse 1989\_12/02/2024\_Version 1\_  
README.rtf*

## 2.2 Making data accessible

As a guiding principle, VAMP 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;
- anonymizing and aggregating the data;
- 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. 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:

[insert project logo]

- collected data belong to third party which have denied permission for sharing on account of confidentiality and proprietary issues;
- data anonymization is not possible;
- 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 chosen data repositories 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<sup>2</sup>, 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.*

<b>Repository name</b>	<b>Type</b>	<b>URL</b>	<b>PID</b>	<b>OpenAIRE compatibility?</b>
AMS Acta	Institutional/Multidisciplinary	<a href="https://amsacta.unibo.it/">https://amsacta.unibo.it/</a>	DOI	Yes

### 2.3 Making data interoperable

All datasets 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 are related to published works that present/build upon the present dataset, complementary datasets, and readme files including dataset description, collection’s context and methods and appropriate software instructions to access the files. To increase interoperability, files will be registered in widespread and accessible formats as described in table 1.

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 reuse our data.

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

---

<sup>2</sup> OpenAIRE, <https://guidelines.openaire.eu/en/latest/>



[insert project logo]

Tools/software
Mp3 Reader
Mp4 Reader
PDF Reader

## 2.4 Increasing data re-use

VAMP license data under CC BY 4.0, whenever possible. See also section 2.2 above.

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 by using good quality audio-video and photographic equipment; by pre-registration recording in order to ensure interview intelligibility; by selecting carefully, whenever possible, the setting for interviews (in order to maximize audio registrations quality); by always providing transcription of interviews in accessible and interoperable formats (see sections 1. and 2.3 above); by acquiring the best possible quality of existing dataset by asking for appropriate reproduction tools to media and national archives existing dataset are stored.

## 3. Other research outputs

As additional research outputs aside of publication and conference or seminar presentations, results will be diffused by different communicative supports in order to maximize social impact and Open Science practices. Currently, two projects are underway, and three are under study.

The projects already underway are the following:

- The co-writing and co-recording of an episode of a podcast series with a colleague of FILO Department (Unibo), Daniele Farruggia. The episode’s title is *La natura della crisi* [eng. *The nature of the crisis*] and it vulgarizes VAMP’s main thesis on ecological crisis and on the concept of Anthropocene. The episode will be diffused within the series “Idee sbaagliate” [eng. Wrong ideas], financed by the budget of science dissemination at disposal of FILO Department. The series will be produced by a professional studio and released before Summer or for the Fall 2024.
- The co-organization with Universté Paris Cité’s colleague Juan Alonso Aldama of an international Conference on “Fallibility and Maintenance of Meaning”, co-funded by VAMP’s budget in partnership with the University of Urbino. The Conference is scheduled for September 10h 2024.

The projects under study are the following:

- The realization of a card game on the Turin-Lyon railroad project, casting the conflict between the projects and its critiques. Its objectives will be to raise awareness of ecological impacts of infrastructures and to reconceptualise concept as an occasion of enhancement of ecological and technical literacy.
- The organization of a nomad seminar throughout Brazil and Italy at Unesp and Unibo. The seminar will receive international scholars whose work contributes significantly on ecological tension understanding focusing on conflicts of meaning.

#### **4. Allocation of resources**

VAMP doesn't require any specific financial investment to make its data FAIR, since it will use institutional repositories and storage solutions that Unibo provides for free to its fellows for processing, storage, sharing and long-term preservation.

However, an approximate amount of 20 hours per months of the researchers' time is and will be allocated to data management. This time includes: the realization of this document and its further versions; and the cleaning, organization and storage of datasets.

Responsibility for data management usually sits with the fellow, Carlo Andrea Tassinari, ORCID: 0000-0002-3409-5234, e-mail: [carlo.tassinari3@unibo.it](mailto:carlo.tassinari3@unibo.it)

#### **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. If external devices are used to store data files (e.g., backup files), they will be kept in a safe place accessible only to the researchers involved or will be encrypted with ad-hoc software.

The cloud storage solution OneDrive, accessed through researcher's institutional account, will be adopted for data sharing among team members. In this case, too, regular backup of the data will be performed to ensure data recovery.

Long term preservation of public data is ensured by the chosen data repositories that have specific preservation policies.

#### **6. Ethical or legal aspects**

Whenever possible, VAMP collect anonymized data. However, if relevant and/or unavoidable, VAMP may collect personal informations about participants involved in ecological controversies under study. These informations will be conserved and used for research purposes exclusively under the condition of informed consent given by the participant. In this case only, nonanonymized data will be stored in the institutional cloud and repository chosen. Nobody, aside from the researcher, will access these datas before their cleaning and anonymization (when anonymization is required). Before uploading in OneDrive and AMSacta Repository, all data registered in researcher's laptop and smartphone will be protected from databreach with appropriate passwords.

[insert project logo]

## Annex I: Datasets

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

<b>Dataset number</b>	<b>Ready at month of project</b>	Interviews to Lyon-Turin and Amazon stakeholders (divided in anonymous and personal data authorized)
<b>1</b>	<b>9</b>	
<b>Status</b>	<b><i>In progress</i></b>	
<b>ID [ID type]</b>		
<b>Version</b>	1	
<b>Creator/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contributor/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contact Person/s</b>	Carlo Andrea Tassinari, [Unibo, FILO, carlo.tassinari3@unibo.it]	
<b>Contents</b>		
<b>Data format</b>	Mp3	
<b>Data volume</b>	1Gb	
<b>Accessibility</b>	Open access	
<b>Related publication/s</b>	Not available	

[insert project logo]

<b>Dataset number</b>	<b>Ready at month of project</b>	Photographic documentation of impacts and political ecology's practices
1	28	
<b>Status</b>	<i>In progress</i>	
<b>ID [ID type]</b>		
<b>Version</b>	1	
<b>Creator/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contributor/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contact Person/s</b>	Carlo Andrea Tassinari, [Unibo, FILO, carlo.tassinari3@unibo.it]	
<b>Contents</b>	Photographic documentation of strategic sites and practices manifesting the lines of conflict of the chosen case studies	
<b>Data format</b>	Mp3	
<b>Data volume</b>	1Gb	
<b>Accessibility</b>	Open access	
<b>Related publication/s</b>		

<b>Dataset number</b>	<b>Ready at month of project</b>	Press Releases on Lyon-Turin Controversy
1	28	
<b>Status</b>	<i>In progress</i>	
<b>ID [ID type]</b>		
<b>Version</b>	1	
<b>Creator/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contributor/s</b>	Carlo Andrea Tassinari, [Unibo, FILO]	
<b>Contact Person/s</b>	Carlo Andrea Tassinari, [Unibo, FILO, carlo.tassinari3@unibo.it]	
<b>Contents</b>	A selection of Press Releases on Lyon-Turin Controversy and on deforestation issue in the Amazon Rainforest	
<b>Data format</b>	Mp3	
<b>Data volume</b>	1Gb	
<b>Accessibility</b>	Open access	
<b>Related publication/s</b>		

## Annex II: Open Access status of project publications

Even if a detailed publication plan is provided in the first version of the Career Development Plan (Deliverable 3, month 6), at the present stage, it is still soon to determine the underlying datasets for each publication. Therefore, we chose to leave this section uncomplete. However, as soon as possible, and in any case in the second version of this Data Management Plan, we will describe the open access status of the project publications and the underlying datasets in a table structured as follows:

*Table 7 – Publications and related datasets.*

Publications	
<b>Bibliographic citation of the publication</b>	

[insert project logo]

<b>Link to copy archived in repository</b>	
<b>Related dataset/s</b>	
<b>Bibliographic citation of the publication</b>	
<b>Link to copy archived in repository</b>	
<b>Related dataset/s</b>	
<b>Bibliographic citation of the publication</b>	
<b>Link to copy archived in repository</b>	
<b>Related dataset/s</b>	
<b>Bibliographic citation of the publication</b>	
<b>Link to copy archived in repository</b>	
<b>Related dataset/s</b>	
<b>[...]</b>	

[insert project logo]

## 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

Semi-directive interview. Explorative interview

Codebook of variables

[If necessary to understand the meaning of the variables]