



# PANTHEON

Community-Based Smart City Digital Twin Platform  
for Optimised DRM operations and Enhanced Community  
Disaster Resilience

## D1.5

### DATA MANAGEMENT PLAN - M6



The project has received funding from the European Union's Horizon Europe programme under Grant Agreement N°101074008.

## DOCUMENT INFO

<b>Deliverable Number</b>	D1.5
<b>Work Package Number and Title</b>	WP1 Project & Quality Management
<b>Lead Beneficiary</b>	UPC
<b>Due date of deliverable</b>	30/06/2023 (M6)
<b>Deliverable type<sup>1</sup></b>	DMP – Data Management Plan
<b>Dissemination level<sup>2</sup></b>	PU – Public
<b>Author(s)</b>	Cristina Barrado (UPC) Esther Salamí (UPC)
<b>Internal reviewer(s)</b>	Dimitrios Petridis (EPSILON), Julie Auerbach (EPSILON), Constanze Geyer (JOAFG), Ana-Maria Dumitrescu (SIMAVI), Mike Karamousadakis (THL), Anna Tsabanakis (THL), and Daoiz Zamora (MION).
<b>Version - Status</b>	V1.5

## TASK ABSTRACT

The main goal of task *T1.5 - Data Management* is to specify the data management life cycle for all data that will be collected, processed or generated by the project. It will provide an outline of the data types the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. The Data Management Plan (DMP) of PANTHEON will be realised in accordance with the *Guidelines on Open Access to Scientific Publications and Research Data in Europe Horizon* and will consider the issues to be addressed in order to be compliant with the *Guidelines on FAIR Data Management*.

<sup>1</sup> Please indicate the type of the deliverable using one of the following codes:

R = Document, report

DEM = Demonstrator, pilot, prototype, plan designs

DEC = Websites, patents filing, press & media actions, videos

DATA = data sets, microdata

DMP = Data Management Plan

ETHICS: Deliverables related to ethics issues.

OTHER: Software, technical diagram, algorithms, models, etc.

<sup>2</sup> Please indicate the dissemination level using one of the following codes:

PU = Public

SEN = Sensitive

## REVIEW HISTORY

Version	Date	Modifications	Editor(s)
1.1	13/04/2023	First draft	Cristina Barrado (UPC) Esther Salamí (UPC)
1.2	06/06/2023	Updated draft after contributions from Dimitrios Petridis and Julie Auerbach.	Cristina Barrado (UPC) Esther Salamí (UPC)
1.3	13/06/2023	Updated draft after contributions from Dimitrios Petridis, Constanze Geyer, Ana-Maria Dumitrescu, and Lola Vallès	Cristina Barrado (UPC) Esther Salamí (UPC)
1.4	16/06/2023	Clean document for approval. Dataset catalogue embedded.	Cristina Barrado (UPC) Esther Salamí (UPC)
1.5	27/06/2023	Final document after WP leaders review.	Cristina Barrado (UPC) Esther Salamí (UPC)

## DISCLAIMER

The document is proprietary of the PANTHEON consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

## TABLE OF CONTENTS

LIST OF FIGURES .....	4
LIST OF TABLES .....	5
LIST OF ABBREVIATIONS .....	6
EXECUTIVE SUMMARY .....	7
1 Data Summary.....	8
1.1 Purpose of data collection and generation .....	8
1.2 Existing input data .....	9
1.4 Types and formats of data.....	10
1.4 Data size .....	10
1.5 Data origin .....	10
1.6 Data utility .....	11
2. FAIR data .....	12
2.1 Making data findable, including provisions for metadata .....	12
2.2 Making data accessible.....	13
2.3 Making data interoperable.....	14
2.4 Increase data re-use .....	14
3 Other research outputs.....	16
4 Allocation of resources.....	17
5 Data security .....	18
5.1 Data preservation.....	18
5.2 Data access control.....	18
5.3 Personal data.....	18
6 Ethics .....	19
7 Other issues.....	20
8. Dataset catalogue .....	21
9. References.....	22

## LIST OF FIGURES

Figure 1: PANTHEON platform will be combined with data from multiple sources [Image from PANTHEON Description of the Action (DoA) Part B] .....	8
Figure 2: Screenshot of the dataset catalogue .....	21

## LIST OF TABLES

Table 1: Categories used to describe each dataset in the data catalogue .....	12
---	----

## LIST OF ABBREVIATIONS

Abbreviation	Meaning
<b>CCL</b>	Creative-Commons-License
<b>CEMS</b>	Copernicus Emergency Management Service
<b>DMP</b>	Data Management Plan
<b>DoA</b>	Description of the Action
<b>DOI</b>	Digital Object Identifier
<b>DRM</b>	Disaster Risk Management
<b>EGNOS</b>	European Geostationary Navigation Overlay Service
<b>EMODnet</b>	European Marine Observation Data Network
<b>EO</b>	Earth Observation
<b>EOSC</b>	European Open Science Cloud
<b>ESA</b>	European Space Agency
<b>FAIR</b>	Findable, Accessible, Interoperable and Reusable
<b>GA</b>	Grant Agreement
<b>GB</b>	Gigabytes
<b>GIS</b>	Geographic Information System
<b>GNSS</b>	Global Navigation Satellite System
<b>GDPR</b>	General Data Protection Regulation
<b>GTAP</b>	Global Trade Analysis Project
<b>IoT</b>	Internet of Things
<b>OAI-PMH</b>	Open Archive Initiative-Protocol for Metadata Harvesting
<b>OGC</b>	Open Geospatial Consortium
<b>OIH</b>	Ocean InfoHub
<b>RIO</b>	Research and Innovation Objectives
<b>SCDT</b>	Smart City Digital Twin
<b>TB</b>	Terabytes
<b>UAV</b>	Unmanned Aerial Vehicle

## EXECUTIVE SUMMARY

This document provides the draft version of the Data Management Plan (DMP) for the PANTHEON project. The purpose of the DMP is to support the data management life cycle for all research data that will be collected, processed or generated by the project. The initial version is delivered in month 6, but the DMP is a live document and shall be continuously updated during the project.

The document contents are based on the current legislation (Regulation EU2016/679 of the European Parliament and of the Council, and the Repealing Directive 95/46/EC on General Data Protection Regulation). In addition, it follows the H2020 programme guidelines: EC Directorate-General for Research & Innovation *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020*, and *H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research*. Finally, public metadata tools such as RDA, Metadata Directory from Common European Research Information Format (CERIF) and DCC, Disciplinary Metadata from Digital Curation Centre will be used.

The DMP shall foster the transparency of the project and its results by sharing project data, promoting their visibility and make it findable using the FAIR principles (*Findable, Accessible, Interoperable and Reusable*).

## 1 DATA SUMMARY

### 1.1 PURPOSE OF DATA COLLECTION AND GENERATION

The PANTHEON project focuses on the creation of knowledge and tools to develop and implement an extensive, realistic and detailed Community-based Digital Ecosystem for Disaster Resilience using Smart City Digital Twin (SCDT) technology and complex interdependent visual and data analytics. PANTHEON platform and technologies will be combined with IoT infrastructure, satellite and in situ data (Figure 1).

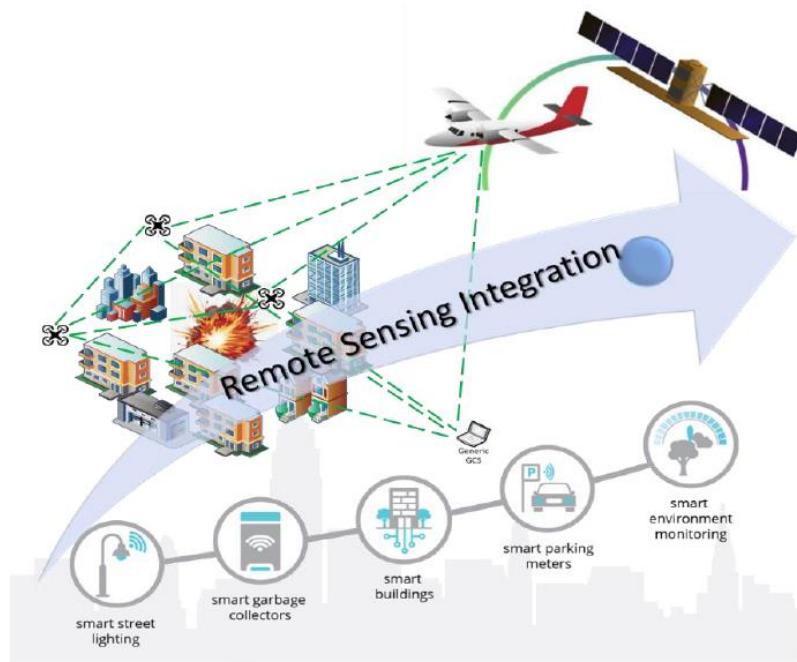


Figure 1: PANTHEON platform will be combined with data from multiple sources [Image from PANTHEON Description of the Action (DoA) Part B]

Eight research and innovation objectives (RIO) have been defined in the Grant Agreement (GA):

- RIO-1. PANTHEON Disaster-Resilient Communities Model
- RIO-2. Decentralised and Participatory Governance & Tools
- RIO-3. Development of a Smart City Digital Twin for Community DRM
- RIO-4. Development of a synthetic Environment
- RIO-5. Interactive User Interface for a Holistic Simulation Environment
- RIO-6. Intelligence, CB-SCDT Processes and Optimal DRM Operations
- RIO-7. Data Delivery Scheme and Remote Sensing for Community-based SCDT
- RIO-8. Integration of all the modules and Hosting Platform

Different datasets will help to reach the above project objectives. In the initial stage of the project, knowledge from the community and stakeholders will be collected (mainly in the form of interviews, questionnaires, and online workshops) to draft a model for Disaster-Resilient communities (RIO-1). The consortium will analyse

regional Disaster Risk Management (DRM) policies, assess regional multi-hazards and community vulnerability and capacity, and investigate the use of scientific and technical capacities and innovation. It will elaborate a participatory governance model as well as recommendations on how to involve, mobilize, inform, and organise the community to effectively participate in all aspects of preparedness, response, and recovery (RIO-2).

Key findings of the initial stage will be used to define the conceptual model of an SCDT for community disaster management (RIO-3). The SCDT will include a suite of tools that use data as well as simulation modelling to provide decision-makers with predictive information about future conditions (RIO-4, RIO-5). Operational area characteristics such as topography or meteorology will be considered, as they have a significant impact on different types of systems. The operational capabilities will also be enhanced by an integrated intelligent subsystem (RIO-6) and a collaborative infrastructure where information from various sources (satellite, in-situ UAVs, Smart City IoT sensors, GIS sources, and social applications) will be fused to provide updated and valid information during all the phases of the disaster management (RIO-7). The different software components will be integrated, and the PANTHEON platform will be tested on dedicated scenarios (RIO-8).

Overall, PANTHEON will generate new data products through modelling and collect a wide variety of data that can be classified into the following categories:

- a) Experts' opinions, used to define requirements and recommendations;
- b) Spatial and Geographical data, used in the SCDT and the UAVs flights;
- c) In-situ sensors data and UAVs flight logs, used to populate the SCDT;
- d) Models and programs, used to develop the platform and to provide intelligence.

## 1.2 EXISTING INPUT DATA

We aim at re-using existing data from satellites, UAVS and in-situ sensors monitoring platforms. For the moment (month 6) we have not disregarded any dataset.

Navigation and positioning data from the European Global Navigation Satellite System (GNSS), EGNOS and Galileo, and data and services offered by the European Earth Observation (EO) programme Copernicus, will be integrated into the system for simulation and training purposes. Collected data will be partially processed before being used.

In particular, the Copernicus Emergency Management Service<sup>3</sup> (CEMS) uses satellite imagery and other geospatial data to provide mapping services in cases of natural disasters, human-made emergency situations and humanitarian crises throughout the world. The general service is free of charge for everyone, but only the national competent authority has access to higher resolution images due to the sensitivity of the data. Other Copernicus Services for Ocean EO data and forecasting have higher resolution. Additionally, only civil authorities have real-time access, otherwise there is a delay which must be considered in terms of responding during disaster events.

For now, PANTHEON will rely upon the available CEMS data. In case higher resolution or real-time data is needed, data from real sensors and UAVs will be collected if possible. Contact with the authorities is also considered as a last option if necessary.

---

<sup>3</sup> <https://emergency.copernicus.eu/>

Data will also be collected through existing in-situ platforms, such as the EU funded platform known as EMODnet<sup>4</sup> (European Marine Observation Data Network) which deals with marine and coastal data and includes some river networks within Europe, the Caribbean and in the Baltic since last year. The UNESCO platform Ocean InfoHub<sup>5</sup> (OIH), which is a full repository of available information and data resources related to the ocean, can be searched to find other useful data sources.

Additional data providers to be considered in PANTHEON include Eurostat<sup>6</sup>, GTAP database<sup>7</sup>, ESA Earth online<sup>8</sup>, and weather and climate simulation hydrological and environmental databases.

## 1.4 TYPES AND FORMATS OF DATA

PANTHEON is expected to generate the following types and formats of data:

- a) Experts' opinions will be collected in the following formats:
  - a. Online workshops: standard mp4.
  - b. Interviews: standard mp3.
  - c. Questionnaires: csv, xls, and pdf.
- b) Spatial and Geographical data will be stored as raster, vector (shape files), csv and pdf files.
- c) In-situ sensors' data and UAV flight logs will be stored as csv, xml, and xlsx files.
- d) Models and programs. Code will be stored in the respective programming language.

## 1.4 DATA SIZE

The overall size of the data is not currently known, but it is expected to be large volumes of data. Most datasets will be in the order on gigabytes (GB), with a total size in the order of terabytes (TB).

## 1.5 DATA ORIGIN

PANTHEON will collect and process data, focused mainly on the Paris and Athens sites, from various sources and devices:

- a) Experts' opinions will be collected by video or audio recordings, and by General Data Protection Regulation (GDPR) compliant anonymous surveys.
- b) Spatial and Geographical data will be obtained from open access products of the Copernicus program or other open access products.
- c) In-situ sensors' data will be collected through different sources:
  - a. In-situ platforms, such as EMODnet.
  - b. On-site observations and flight records collected by the PANTHEON UAV swarm system.
  - c. Real-time data from ground-based IoT sensors gathered from local infrastructure.
- d) Models and programs will be developed by the project partners.

<sup>4</sup> <https://emodnet.ec.europa.eu/>

<sup>5</sup> <https://oceaninfohub.org/>

<sup>6</sup> <https://ec.europa.eu/eurostat>

<sup>7</sup> <https://www.gtap.agecon.purdue.edu/databases/>

<sup>8</sup> <https://earth.esa.int/eogateway>

## 1.6 DATA UTILITY

The PANTHEON project will develop a Community-based Digital Ecosystem for Disaster Resilience based on SCDT technology to advance the existing risk assessment methods, mitigate potential vulnerabilities, and build a strong community disaster resilience foundation.

The SCDT will be useful to the end users such as public authorities, emergency responders and disaster managers to improve the anticipation and management of disasters and to promote community engagement and the use of smart technologies.

PANTHEON research data will also be of interest to other research groups developing projects related to SCDT technology, Disaster Risk Management (DRM), multi-hazards assessment, and community engagement.

Finally, small and mid-size enterprises or infrastructure providers may be interested in developing, testing and exploiting the final product.

## 2. FAIR DATA

The FAIR principles describe how research outputs should be organised in order to be *Findable, Accessible, Interoperable* and *Reusable*, thus facilitating knowledge discovery and reuse.

Open research data generated in this project are intended to be deposited in Zenodo<sup>9</sup>, the OpenAIRE's trusted repository hosted by CERN, which supports FAIR principles. To this end, PANTHEON has set up its own **Zenodo Community**:

[https://zenodo.org/communities/horizoneurope\\_101074008/](https://zenodo.org/communities/horizoneurope_101074008/)

This also ensures that all uploads to the community collection are validated by the community curator before being published.

### 2.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Data is made findable by providing rich metadata. Metadata must be assigned a globally unique and persistent identifier and registered or indexed in a searchable resource.

PANTHEON research data will be organized in datasets. All datasets will be **named** uniformly when stored, based upon the following criteria:

- Naming convention: PANTHEON\_<partner>\_<description>\_V<#>.<#>
- No special characters such as “ / \ : \* ? < > [ ] & \$ will be used in names.
- Underscores (‘\_’) will be used to separate terms, not spaces.
- Names will be 30 characters or less in length.
- Names will be descriptive of what information they contain.

To make the data findable by the consortium partners during the project lifetime, a dataset **catalogue** will be prepared and deployed in the BOX<sup>10</sup> cloud-based collaborative content management platform. The catalogue will be implemented as an excel table that gives an overview of all the datasets. All partners are asked to update the table with information of the research data they will collect, reuse or generate during the project. Each entry corresponds to a dataset that will be described with specific information given in Table 1. The catalogue serves as an inventory of data and provides information to other users in the project.

**Table 1: Categories used to describe each dataset in the data catalogue**

Category	Content
<b>Name</b>	Name of dataset according to filename rules in section 2.1 of DMP
<b>Description</b>	Short description of the dataset
<b>Status</b>	Completed (C), In-progress (P), Not started (N)
<b>Task</b>	Related project task
<b>Partner</b>	Partner in charge
<b>Contact</b>	E-mail of contact person for questions/remarks
<b>Origin</b>	Collected (C), Reused (R) or Generated (G)

<sup>9</sup> <https://zenodo.org>

<sup>10</sup> <https://app.box.com/>

Type	Type of data (interviews, measurements, code, models, etc.)
Gathering/Sources	How the data is gathered (audio recording, anonymous survey, public repositories, etc.) and sources used to prepare the dataset
Sources	Sources used to prepare the dataset
Format	Data format (csv, pdf, mp3, etc.). Provide software to read data, if not standard
Storage	Where the data is stored
Size	Expected required disk space
Confidential	Specify if the dataset is confidential
Personal	Specify if dataset contains personal data (name, e-mail, etc.)
Sensitive	Specify if dataset contains sensitive data (political affiliation, criminal history, etc.)
License of reused data	If data is reused, specify the license that allows you to reuse it
Access	Terms of data access/usage (owner only, consortium partners, public, license, etc.)
Metadata	Specify which metadata standard will be used to describe the data
Vocabulary	Specify any controlled vocabulary used in the metadata
Notes	Specific remarks (e.g., restrictions, gaps, etc.)
DOI/URL	DOI or URL if available

It will be discussed within the PANTHEON consortium which datasets can be made public during or after the completion of the project (always considering the licenses from reused data, if any). **Metadata** standards specific to the discipline will be used. **Keywords** for open data from Thesauri and controlled vocabularies that are suitable for the specific type of data will be selected. A **DOI** will be assigned by Zenodo and included as a mandatory field in the metadata. Metadata of each record will be indexed and searchable directly in Zenodo's search engine immediately after publishing.

## 2.2 MAKING DATA ACCESSIBLE

Data is made accessible by ensuring it is retrievable online using standardised protocols.

### **Repository**

During the development of the project, data will be deposited on the partners' home servers and on the PANTHEON BOX cloud repository. A unique URL is assigned to each file in BOX for unequivocally identification.

The long-term data trusted repository proposed for PANTHEON is Zenodo. The selected repository is free to upload and free to access. It uses Persistent Identifiers (DOIs) and generic or discipline-specific metadata, which promote data visibility.

### **Data**

The specification of data which will be made openly available is still under discussion. It will be decided among the project partners on a case-by-case basis and this section will then be updated. As a guiding principle, PANTHEON seeks to make data openly available whenever possible. Restrictions on data access will be considered in cases where the collected data belongs to third parties with their own confidentiality and proprietary rules, and also when protection of personal data is required.

## Metadata

Metadata are intended to be publicly accessible and licensed under public domain. No authorization will be ever necessary to retrieve it. In Zenodo, metadata is harvestable using the OAI-PMH protocol. Metadata is also retrievable through the public REST API.

At this time, all data is expected to use standard formats and no special software is required to access or read the data. Some file formats, such as tabular ones, will be automatically converted to open preservation formats. Data and metadata will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years, at least.

## 2.3 MAKING DATA INTEROPERABLE

Metadata interoperability is ensured by following metadata standards. Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML. Widely used standard formats and protocols such as OAI-PMH, ensure datasets exchange and reusability between researchers.

For geographic data, the INSPIRE Directive 2007/2/EC and related standards, such as the Open Geospatial Consortium (OGC) standards, will be adopted. In all the other cases, community standards will be searched. Otherwise, metadata compliant to general purpose standards such as Dublin Core, will be adopted.

For the time being, the following standard vocabularies have been identified in PANTHEON:

- The ASIS&T Thesaurus of Information Science, Technology, & Librarianship.
- For stakeholders' datasets, the standardized vocabulary EU Country Named Authority List (<https://data.europa.eu/euodp/en/data/data/aset/country>) is used.

In case it is unavoidable to use uncommon or project specific ontologies or vocabularies, PANTHEON will openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them.

## 2.4 INCREASE DATA RE-USE

The data must have a clear and accessible data usage license so others know what types of reusing are permitted. The Creative-Commons-License (CCL) option, is a good start to initiate the discussion within the Consortium, but this will be decided when a full picture of the data is available. The license for each dataset will be one providing the widest re-use possible. It will also be made clear which data in the project can be useable by third parties.

All relevant documentation needed to facilitate data reuse will be made available along with the data, mostly in the form of readme text files or pdfs. This can include users' manuals, data collection and cleaning procedures, methodology description, etc.

In the case of reused data, the provenance of the data will be thoroughly documented. Most data providers have a policy of encouraging free re-use of its data, provided that the source is indicated as Eurostat, and any modification or processing to the data is clearly stated.

In addition to making data findable, accessible and interoperable as described in the previous sections, to increase data reuse, special attention will be given to enhancing data quality and interpretation, and to dissemination and communication activities.

### **3 OTHER RESEARCH OUTPUTS**

For the moment there are no more research outputs foreseen, other than the ones listed in Section 1.4.

## **4 ALLOCATION OF RESOURCES**

During the project life, the BOX cloud-based collaborative content management platform is adopted to share data and documentation among the partners. The cost to activate and maintain it for the duration of the project is covered by the project budget.

There are no costs for long-term deposit and preservation of public data, because the chosen repository Zenodo is free of charge and allows long-term preservation of data for the lifetime of the repository, defined for the next 20 years.

Although UPC is the leader in the task of drafting the DMP, it is up to all partners to regularly contribute and review the plan and meet the objectives of the plan throughout the life cycle of the project.

Responsible for the management of individual datasets are the dataset creators, who are generally the team leaders directly involved in the task. The data catalogue provides the project task, the partner in charge, and the e-mail of a contact person for each dataset.

## **5 DATA SECURITY**

Data security is supported by the barriers that are set in a storage system against the potential misuse through non-authorised access and also against the potential loss or damage of data.

### **5.1 DATA PRESERVATION**

To avoid the loss or damage of the research data, we will have up to three levels of redundancy of the data storage. All the data collected or generated by a partner will be stored at their premises, and additionally at BOX, a collaborative platform with shared storage for all group partners.

Moreover, storage systems (BOX included) have a backup program to avoid losses of data.

To further secure collected or generated research data for future use, it is recommended to store the data in certified repositories for long-term preservation and curation like Zenodo, a trusted public repository proposed for PANTHEON. Zenodo is a multi-disciplinary open repository hosted by CERN. In particular, Zenodo allows long-term preservation of data for the lifetime of the repository, defined for the next 20 years at least.

### **5.2 DATA ACCESS CONTROL**

The storage systems used by the partners and the collaborative BOX platform have a first security barrier consisting of access through a username and password. Usernames and Passwords as set at personal level, have sufficient length to not be easy to break, and passwords are hidden and never shared. Additionally, BOX keeps a historical record of modifications and data updates. These historical records can help to assess the correct access to data. Storage systems have physical protection, deployed in a lock room, accessible only to authorised users.

### **5.3 PERSONAL DATA**

Personal data is not considered research data and thus are not part of the DMP. Lists of persons participating with the project in workshops, interviews or social networks, as well as consent forms signed by volunteers, are out of the scope of the DMP. However, it is worth noting that any system storing personal data will have the encrypt functionality or be off-line and hence not freely accessible through the internet. Only safe storage at the premises of the partner collecting the personal data will be used, together with the consent forms. A copy of the consent forms will be also stored in BOX in a separate folder that is accessible only by the involved partners. Exceptionally, the content including the personal data can be shared with another project partner if their processing capabilities are necessary. This partner must delete them after returning the processed data to the source.

## 6 ETHICS

The PANTHEON Description of Action sets the need to take vulnerable groups (children, elders, persons with disabilities, refugees, etc.) into consideration when managing emergencies. Motivation is supported by the Convention on the Rights of Persons with Disabilities (United Nations, 2006) and the Charter of Fundamental Rights of the European Union (European Union, 2000).

As PANTHEON is intended to work with vulnerable groups, a special protocol will be considered to deal with it. When preparing and submitting consent forms for vulnerable persons, PANTHEON researchers will use the checklist set in Table 1 of D10.1 to ensure compliance with all relevant aspects of participation, ethics, personal data processing and privacy of participants. These requirements have been created in accordance with the three basic principles of informed consent: information, voluntariness and competence (European Commission, 2010 & 2013). Section 3.3 of the Ethics deliverable D10.1 provides details on when a group is considered as vulnerable for the PANTHEON project.

Data stored for the long term will be anonymised. All the data gathering processes including personal data will, in advance, be supported by a consent form, signed by all persons participating. Signed consent forms will be stored by the partner collecting the data and the project coordinator will securely store a copy on the dedicated BOX folder. Consent forms will be preserved at least for the same period of time in which the non-anonymised data is stored. More details are given in the Data Security (section 5) and in the Ethics deliverables of the project.

## **7 OTHER ISSUES**

The PANTHEON data management will follow the European FAIR principles, detailed as explained in this document, with no additional procedures set at any other level (National, partner, etc.).

## 8. DATASET CATALOGUE

The first request for data information has identified 19 data sets from 10 partners (FINT, THL, JOAFG, ISEMI, ISPC, ENAC, CMSA, EPSILON, KEMEA and SIMAVI) and 21 tasks. Currently (M6), none of these datasets are finished yet, thus the current version of the catalogue (Figure 2) shows only the structure but no dataset yet.

#	Name	Description	Status	Task	Partner	Contact	Origin	Type	Gathering/Sources	Format	Storage	Size	Confidential	Personal	Sensitive	License of reused data	Access	Metadata	Vocabulary	Notes	DOI/URL
1																					
2																					
3																					
4																					
5																					
6																					
7																					
8																					
9																					
10																					
11																					
12																					
13																					
14																					
15																					
16																					
17																					
18																					
19																					
20																					
21																					
22																					
23																					
24																					
25																					

Figure 2: Screenshot of the dataset catalogue

## **9. REFERENCES**

European Commission, Directorate-General for Research & Innovation (2016). *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0. Available at: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf) [accessed 28 May 2023]

European Commission, Directorate-General for Research & Innovation (2017). *H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*, Version 3.2. Available at: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf) [accessed 28 May 2023]

RDA | Metadata Directory (2021). Available at CERIF (Common European Research Information Format): <http://rd-alliance.github.io/metadata-directory/standards/cerif.html> [accessed 28 May 2023]

DCC | Disciplinary Metadata (2021). Available at DCC: <https://www.dcc.ac.uk/guidance/standards/metadata> [accessed 28 May 2023]

*Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*. OJ L 108, 25.4.2007, p. 1–14, ELI: <http://data.europa.eu/eli/dir/2007/2/oi> [accessed 28 May 2023]

*Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. OJ L 119 04.05.2016, p. 1, ELI: <http://data.europa.eu/eli/reg/2016/679/oi> [accessed 28 May 2023]

UN General Assembly, *Convention on the Rights of Persons with Disabilities : resolution / adopted by the General Assembly*, 24 January 2007, A/RES/61/106. Available at: <https://www.refworld.org/docid/45f973632.html> [accessed 28 May 2023]

*Charter of Fundamental Rights of the European Union*. OJ C 326, 26.10.2012, p. 391–407, ELI: [http://data.europa.eu/eli/treaty/char\\_2012/oi](http://data.europa.eu/eli/treaty/char_2012/oi) [accessed 28 May 2023]