D1.3 Data Management Plan (v1)

Deliverable number	**D1.3**
Deliverable title	**Data Management Plan (v1)**
Nature[1]	**ORDP: Open Research Data Pilot**
Dissemination Level [2]	**PU**
Author (email) Institution	**Vassilis Psaltopoulos (vassilispsaltopoulos@iccs.gr)** **Institute of Communication and Computer Systems**
Editor (email) Institution	**Antonis Kalis (antonis.kalis@iccs.gr) Institute of Communication and Computer Systems**
Leading partner	**ICCS**
Participating partners	**All**
Official submission date:	**31/05/2020**
Actual submission date:	**11/06/2020**

[1] **R**=Document, report; **DEM**=Demonstrator, pilot, prototype; **DEC**=website, patent fillings, videos, etc.; **OTHER**=other

[2] **PU**=Public, **CO**=Confidential, only for members of the consortium (including the Commission Services), **CI**=Classified, as referred to in Commission Decision 2001/844/EC

Modifications Index

Date	Version
29/4/2020	0.1 ToC creation
19/5/2020	0.2 First draft circulated to partners
1/6/2020	0.3 Final draft circulated to partners
10/6/2020	0.4 Final document

This work is a part of the HYPERION project. HYPERION has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 821054.

ACRONYMS AND ABBREVIATIONS

EU	European Union
WP	Work Package
GDPR	General Data Protection Regulation
EC	European Commission
NGO	Non-governmental organization

Table of Contents

Executive Summary

HYPERION is a H2020 framework project funded under grant agreement 821054 which aims to introduce a research framework for assessing the risk and resilience of Cultural Heritage sites subject to natural hazards in a climate-change era.

The project will downscale the created climate and atmospheric composition as well as associated risk maps down to the 1x1 km scale, and specific damage functions for Cultural Heritage materials perform combined hygrothermal and structural/geotechnical analysis of the Cultural Heritage sites and damage assessment under normal and changed conditions, based on the climatic zone, the micro-climate conditions, the petrographic and textural features of building materials, historic data for the structures, the effect of previous restoration processes and the environmental/physical characteristics of the surrounding environment.

The project outcomes will be demonstrated to four European historic areas in Norway, Spain, Italy and Greece, representing different climatic zones.

The present deliverable underlines the project's Data Management Plan, describing the procedures for data collection, storage and processing. In principle, the data collected, generated or reused throughout the project's lifecycle will be processed only for scientific reasons. All the data related activities will comply with the requirements of the General Data Protection Regulation.

In parallel, HYPERION is part of the Pilot on Open Research Data so it will make sure that it follows the FAIR principles thus making its data Findable, Accessible, Interoperable and Re-usable.

This deliverable provides a brief description of the data managed within the project's scope. It's worth noting that only pseudonymous data that are necessary for the scientific analyses are expected to be stored in the long term on the databases hosted in secured servers that will be regularly backed up, and accessible only by authorised users.

The consent of the users to the processing of their data will be asked prior to their involvement in any project related activities, and upon being well informed on the project's nature. Additionally, users will have the possibility to easily withdraw their consent and to exercise their rights deriving from the General Data Protection Regulation, such as the right to be forgotten.

The project's outputs in terms of deliverables and/or any project outcomes related scientific publications will be named and indexed with appropriate keywords and will be available via the project website and in research fora.

1. Introduction

1.1 Objectives of the Deliverable

The purpose of this deliverable is to provide the Data Management Plan for HYPERION that follows the Open Research Data Pilot principles. Additionally, it explains how the project activities are complying with Protection of Personal Data (POPD) requirements established by the EC and national regulations. The data management plan describes how data is being collected, stored, documented, shared and reused during and after the project. The deliverable is considered a living document that will be updated over the course of the project, in particular prior to the mid-term and the final reviews or on-demand, when changes in policies by the EU Commission or the consortium synthesis have occurred or when a new innovation potential (e.g. patent) has been identified.

1.2 Participation in the Open Research Data Pilot

HYPERION participates in the Pilot on Open Research Data launched by the EC along with the Horizon 2020 programme. The consortium supports open science, and the large potential benefits to the European innovation and economy stemming from allowing reusing data at a larger scale. Therefore, the majority of data produced by the project may be published with open access – though this objective will obviously need to be balanced with the other principles described below.

The Data Management plan specifies the implementation of the pilot, in particular with regard to the data generated and collected, the standards in use, and the workflow to make the data accessible for use, reuse and verification by the community and defines the strategy of duration and preservation of the data.

The Open Research Data Pilot principles aim to improve, maximize access and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness with the protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions.

Dedicated deliverables on Ethics such as “D11.1 H - Requirement No. 1” and “D11.2 POPD - Requirement No. 2” falling under the scope of WP11 are devoted to explain the procedures and criteria that will be used to identify/recruit research participants and how HYPERION activities are complying with Protection of Personal Data (POPD) requirements established by the EC and national regulations, as well as with GDPR with respect to the privacy of EU citizen.

1.3 IPR management and security

As a Research and Innovation Action (RIA), HYPERION aims at developing close to market technological solutions. The project consortium includes partners from the private sector. Those partners obviously have Intellectual Property Rights on their

technologies and data. Consequently, the consortium will protect data and have the confirmation of concerned partners before every publication of them. Moreover, as data collected through the project's activities are of high value - measures are foreseen to prevent them from being falsified.

All data repositories used by the project will comply with specific security and data protection measures. In addition, protective measures against infiltration will be taken as well as physical protection of core parts of the systems and access control measures. More details on this will be provided in Section 5.

2. HYPERION Data Summary

2.1 End-users' needs

End-users' needs and expectations are gathered through different means (e.g., questionnaires, workshops, combination of previous methods).

Table 1 End-users' needs

Data Category	Data type	Description
Questionnaires	Paper or online form	Physical (paper) or online document with series of questions for the purpose of gathering information from respondents
Interviews	Notes on paper or digital	Private or public conversation with questions asked to elicit information

The purpose of the data collection/generation and its relation to the objectives of the project:

The collection of inputs from the stakeholders, combined with a deep analysis of all technical, regulatory and financial aspects were considered for the development of HYPERION integrated system.

The types and formats of data that the project will generate/collect*:*

Responses to paper-based questionnaires were marked on the paper itself.

Notes from interviews were taken and saved digitally in *.doc file type or physically in paper.

The re-usability of any existing data:

No existing data was reused.

The origin of the data:

Data originated from human personal responses to several questions or where extracted as notes from conversations.

The expected size of the data:

Digital notes on *.doc files were of size ~2MB

Paper-based questionnaires and notes were of about 40 paper sheets.

The data utility:

All information from this process had a positive impact on the definition of the functional and non-functional requirements and the system architecture.

2.2 Pilot area surveys

Table 2 Pilot area surveys

Data Category	Data type	Description
Maps	Digital files (dwg, tiff, png, etc.)	Maps of the pilot study areas
Asset data	Paper and digital forms	Survey forms filled by Hyperion member after visiting each asset (building) designated in the area
Census data	Digital (csv, xls, pdf)	Data from the publicly available sources of the assets and population at risk in the pilot area.

The purpose of the data collection/generation and its relation to the objectives of the project:

The main objective, here, is to collect data as much detailed as possible in order to have an overview of the structural characteristics of the assets belonging to Tier 3. For those building the vulnerability will be quantified using some predefined parameters/information as proxies. The latter will be defined based on the outcome of the surveys conducted in specific areas of the pilot cities.

The types and formats of data that the project will generate/collect*:*

The data will be extracted from online forms structured ad-hoc for the purpose of the project. Data can be also on paper format but will be uploaded online anyway after the survey.

The re-usability of any existing data:

Existing data from municipality or local authorities will be employed to improve the survey quality and coverage.

The origin of the data:

Publicly available data repositories, such as building and population census data, municipality datasets etc. Also, data originated from human personal responses to questionnaires as well as in-person surveying of the study areas via customized asset description datasheets.

The expected size of the data:

- Paper format data: *.pdf files from survey form scan, about 500 KB each.
- Pictures from surveys: *.jpg from survey, about 1 MB each (depending on camera quality)
- Excel spreadsheet including all the digitalized surveys, about 500 KB per pilot city
- Total of 50-200 MB of data at most

The data utility:

Assess the exposure (assets at risk) at each pilot site is a key ingredient of risk & resilience assessment. In particular, those data will be used as input for assessing vulnerability and loss in the framework of HRAP resilience tool.

2.3 Geographic & Climate Data

Based on the requirements gathered via consultation with the users all the available geographic and other data and services (e.g. hydrological, atmospheric, meteorological) related to CC and various geo-hazards were recorded. In addition, an extensive analysis on existing EUROCORDEX climate simulation results was made to identify relevant "episodic" periods. The analysis involved results previously obtained by various Regional Climate Models (RCMs) in resolutions of 0.11 and 0.44 degrees driven by the ERA-Interim input, as well as for one or a few scenarios carefully selected to span the spread of expected uncertainties.

Table 3 Geographic & Climate Data

Data Category	Data type	Description
Episodic Periods	Dataset	Datasets collection for identification of episodic periods.

The purpose of the data collection/generation and its relation to the objectives of the project:

The identification of episodic periods throughout the datasets collected is expected to play a major role in the project's lifecycle and outcomes.

The types and formats of data that the project will generate/collect*:*

Datasets collected from EUROCORDEX contain:

Regional Climate Models are mathematical climate prediction models that can be employed to describe the evolution of the climate over the extended area of interest.

The outcome of the analysis will generate the episodic periods datasets which contain spatially correlated fields of intensities on a daily or 10-min discretization time scale that convey the weather intensities over a time (episodic period) of interest, typically associated with the exceedance of given thresholds of intensities of interest (e.g., high/low temperatures, extreme wind events etc.).

All data will be encoded in csv and netCDF formats

The re-usability of any existing data:

Climate data from existing sources, such as IPCC and EUROCORDEX, and existing climate models, such as the Eulerian Mesoscale models HARMONIE and MEMO, CFD models MIMO and PALM-LES and coupled model MEMICO.

The origin of the data:

Data collected from EUROCORDEX can be found in https://www.euro-cordex.net/

The expected size of the data:

Collected data from EUROCORDEX were of size ~10 GB

Collected data of Regional Climate Models were of size ~100 GB

The generated data containing the episodic periods were of size ~100 MB

The data utility:

Datasets collected combined with the Regional Climate Models will determine the episodic periods. The episodic periods will be used as inputs to the Multi-Hazard Models module, and eventually HRAP.

2.4 Micro-climate stations and Smart Tags

A smart and distributed sensor network based on small, autonomous, unattended, reliable and long-life wireless tags will be developed.

Table 4 Micro-climate stations and Smart Tags

Data Category	Data type	Description
Smart Tags data	Numerical values & Sensor Metadata/ JSON	Datasets derived from Smart Tags on the distributed sensor network
Micro-climate stations data	Numerical values & Sensor Metadata/ JSON	Datasets derived from sensors placed in specific locations of the sites

The purpose of the data collection/generation and its relation to the objectives of the project:

The collected data will provide direct input to the Dynamic Data Assimilation of meteorological and climate data task where data aggregated from the sensor network and delivered through the monitoring system will be engaged with the simulated data and will analyzed and fused (including aggregation, synchronization, calibration and assimilation) from a fully interoperable data management platform. The data will be also used by the deterioration functions, the hygrothermal simulator and the HRAP platform.

The types and formats of data that the project will generate/collect*:*

Smart Tags data: JSON data objects containing time series of measurements.

Micro-climate stations data: JSON data objects containing time series of measurements.

The re-usability of any existing data:

No data will be reused from existing sources.

The origin of the data:

The data originate from on-site deployed sensors and monitoring systems.

The expected size of the data:

The Smart Tags datasets are of size ~ 10MB per sensor per month.

The Micro-climate stations datasets are of size ~ 10MB per sensor per month.

The data utility:

The data generated by the smart tags will be incorporated in the monitoring and mapping tasks and tools.

2.5 Atmospheric Models

Table 5 Atmospheric models

Data Category	Data type	Description
Improved 3D model	Mathematical model	Dynamic Data Assimilation of meteorological and climate data from sensors.
Time-series	Datasets	
impact indicators	Datasets	

The purpose of the data collection/generation and its relation to the objectives of the project:

Topological data, CC scenarios and digital georeferenced maps, in scales ranging from 100km down to <1km including long-term hydrological time-series for each site or a concise set of statistics of historical hydrology data that will be collected will lead to the compilation of very high resolution (better than 1 km) time-dependent land use and land cover maps, including historical data and future land use scenarios. Further to this, maps of thermophysical, biophysical and artificial-surface parameters will be compiled, in cases that such thematic layers cannot be easily derived from the Land Use maps. A 3d model development will also be used in a very-high spatial resolution simulation together with an optimized allocation of temporal study periods for each selected case study.

The types and formats of data that the project will generate/collect:

The time-series and statistics of atmospheric and climatic impact indicators include the following parameters: ambient temperature, precipitation, relative humidity, radiation, atmosphere and the soil, as well as their relevant statistics on medium- and long-term extreme values.

The impact indicators will include parameters as defined in the WMO/ETCCDI specification of Descriptive Indices of Climate Extremes.

The re-usability of any existing data:

No external to the project data will be reused.

The origin of the data:

All related data will come from previous internal project sources.

The expected size of the data:

The Time-series datasets are of size ~10 GB.

The Micro-climate stations datasets are of size ~100 GB.

The data utility:

A qualitative and quantitative assessment on the relevance of primary and secondary impact indicators derived from climate calculations and RT in-situ measurements that will be used in the process of providing RT hazard assessments for the DSS, as well as HT and structural vulnerability assessment.

The time-series indicators will be used as input in the hydrothermal and structural modelling tools.

The impact indicators will be used to dynamically update an independent mapping of atmospheric stressors in the smallest scale that will be explicitly simulated by the modelling cascade.

2.6 Building materials and d-HAM properties

Table 6 Building materials and d-HAM properties

Data Category	Data type	Description
d-HAM properties	Dataset, Lab. measurements	Descriptions, images, numerical values
Building materials	Dataset, Field survey, Lab. measurements,	Descriptions, maps, images, numerical values

The purpose of the data collection/generation and its relation to the objectives of the project:

The data will be collected in such a way to classify and characterize the building materials and their deterioration patterns/mechanisms, and to define the dynamic HAM properties of the materials.

The types and formats of data that the project will generate/collect*:*

The Building materials outcomes data will be collected on .doc, .jpg, .tiff and .xls files.

The d-HAM properties outcomes data will be collected on .doc and .xls files.

The re-usability of any existing data:

No external to project data will be reused.

The origin of the data:

The building materials and the d-HAM properties will be extracted by in-situ inspections and laboratory measurements.

Data will be provided by:

1. Tier 1 building survey: petrographic description and state of damage
2. Petrographic description (macro- and microscopic description)
3. Physical characterisation
    a. Bulk density and porosity (types of analysis: Mercury Intrusion Porosimetry, micro-Computed Tomography, Hydric Tests)
    b. Physical and mechanical properties (ultrasound tests and uniaxial compressive tests)
4. d-HAM properties measurements
    a. Hydric behaviour/capillarity
    b. Relative Humidity (RH) condensation
    c. Specific heat capacity
    d. Thermal conductivity
    e. Water vapour thermal diffusion resistance factor
    f. Hygroscopic sorption properties
5. surface topography (3D profilometry)

When surface patinas are present, specific analyses will be performed (e.g. by XRD, FT-IR, ER-FTIR, μ-Raman, IR-thermography, XPS, Dual VNIR-SWIR).

The expected size of the data:

The building materials properties data are of size ~2Tb.

The d-HAM properties data are of size ~50Gb.

The data utility:

These data will be used to evaluate the effect of micro-climate conditions on deterioration of the different building materials and quantify the linearity of surface recession rate.

2.7 Damage and dose-response functions

Table 7 Damage and dose-response functions

Data Category	Data type	Description
Prediction of decay	Dataset, lab measurements, on-site measurements	Numerical data, images
Validation of Lab tests	Dataset, on-site measurements	Numerical data, images
Dose-response functions	Mathematical model	Numerical data

The purpose of the data collection/generation and its relation to the objectives of the project:

The purpose of data collection is to refine damage and dose-response functions considering the different petrographic and textural features of the building materials, and measuring variations in the surface topography of samples exposed to on-site and lab environmental conditions. Dose-response functions will be validated against long-term recession rate measurements on tombstones from historical war cemeteries.

The types and formats of data that the project will generate/collect*:*

The outcome of these data categories will be collected on .xls, .tiff, .jpg, .doc, files.

The re-usability of any existing data:

No external to project data will be reused.

The origin of the data:

The prediction of decay will be achieved by in-situ inspections and lab measurements, building a database of materials surface features including topography and patinas (3D profilometry IR-thermography, XPS, Dual VNIR-SWIR), and performing aging tests (freeze-thaw cycles and salt crystallization cycles).

The expected size of the data:

The prediction of decay dataset is of size ~2Tb.

The validation of lab tests dataset is of size ~50Gb.

Dose-response functions dataset is of size ~1Gb.

The data utility:

The data generated and collected will provide a better prediction of decay and, integrated in d-HAM models, will allow reliable estimates of CH vulnerability considering validated downscaling climate simulations.

2.8 Materials' physical-mechanical properties

Table 8 Materials' physical-mechanical properties

Data Category	Data type	Description
Physical-mechanical properties	Dataset, lab. measurements	Numerical data, images, descriptions
Deterioration estimation rules.	Dataset, lab. measurements	Descriptions, images

The purpose of the data collection/generation and its relation to the objectives of the project:

This section describes the physical-mechanical properties of the building materials, and develops simplified models/rules for the estimation of their deterioration due to the decay imposed by climate over time. This will lead to a more precise prediction of the possible structural risks (from detachment of parts to collapse) of the cultural asset in a changing climate characterized by an enhancement of extreme events.

The types and formats of data that the project will generate/collect*:*

The data format of these data categories will be of .doc, .tiff, .xls, files.

The re-usability of any existing data:

No external to project data will be reused.

The origin of the data:

The definition of the physical properties and behaviour during stressed climate conditions will be simulated by aging test cycles in the lab (water, sunlight, and thermal cycles).

The expected size of the data:

The physical-mechanical properties data is of size ~20GB.

The deterioration due to decay data is of size ~10GB.

The data utility:

The innovative climate/time dependent mechanical properties of materials will be accessed by the SG simulator's structural analysis tool and by the HRAP model.

2.9 Materials' decay

Table 9 Materials' decay

Data Category	Data type	Description
Materials' decay	On-site measurements	Numerical data, images

The purpose of the data collection/generation and its relation to the objectives of the project:

Determining the effect of extreme events in order to evaluate the vulnerability under climate change scenarios predicting increasing frequency and intensity of such events.

The types and formats of data that the project will generate/collect:

The field tests outcomes will be of .xls, .tiff, file type.

The re-usability of any existing data:

No external to project data will be reused.

The origin of the data:

On-site surveys after extreme events (e.g. m-Raman spectroscopy, Dual VNIR-SWIR, IR-Thermocamera), and analyses of deterioration patterns on assets which have been exposed to documented extreme events in the recent past.

The expected size of the data:

The field tests outcomes data is expected to be of size ~1TB.

The data utility:

Data will be collected to study the expected significant effect of floods, coastal storms and strong wind events on deterioration mechanisms related to salt transport and deposition, in combination with extreme tidal range events and sea-level rise.

2.10 Hygrothermal simulator

Integration of all in-situ and lab measured d-HAM properties of building materials with the proposed HT simulator.

Table 10 Hygrothermal simulator

Data Category	Data type	Description
HT simulator data	Datasets	Tabular data with minimal metadata. Textual data. Documentation and scripts. Image data

The purpose of the data collection/generation and its relation to the objectives of the project:

The HT simulator will integrate all in-situ and lab measured d-HAM properties of building materials and involve a detailed database of building material and their HT properties, as have been identified, collected and characterized. The transient measured data will validate the model of the simulator.

The types and formats of data that the project will generate/collect*:*

The simulator will generate the predictive results of HT performance of building components and constructions in CH sites, in the form of .csv; .rtf and .pdf; .txt and .tiff

The re-usability of any existing data:

No external to project data will be reused that have not previously mentioned.

The origin of the data:

Previously reported data will be integrated and analyzed.

The expected size of the data:

The expected size of the simulation outcomes is expected to be of size ~1TB.

The data utility:

The simulator will be open-access through a webpage to host the databases allowing a realistic prediction of HT performance of building components and constructions in CH sites.

2.11 Multi-Hazard Model

Table 11 Multi-Hazard Model

Data Category	Data type	Description
Natural & human induced disasters	Database	
Hazard model	Dataset	Model results

The purpose of the data collection/generation and its relation to the objectives of the project:

A database will be set containing information layers depicting present hazards and their associated potential to cause harm and a model for a broad spectrum of natural and man-made hazards will be developed

The types and formats of data that the project will generate/collect*:*

The natural & human induced disasters database will contain data in formats .tiff and NetCDF.

The hazard model outcomes will be in NetCDF format.

The re-usability of any existing data:

No further external to project data will be reused that have not previously mentioned.

The origin of the data:

Previously reported data will be integrated and analyzed.

The expected size of the data:

The natural & human induced disasters database is expected to be of size ~1TB

The hazard model outcomes are expected to be of size ~10 GB

The data utility:

All related data will be used to develop an advanced and reliable model for a broad spectrum of natural and man-made hazards, and provide the involved stakeholders, scientists and users with useful information for their specific needs in handling natural and human induced disasters.

2.12 Vulnerability modules

Table 12 Vulnerability modules

Data Category	Data type	Description
Integration software toolkit	numerical values, diagrams, interface software	Tier 1 monitoring interface.
Detailed structure model	Structural response (numerical values)	Tier 1-2 structures model
Hazard simulator	Prognostic results (numerical values)	Vulnerability assessment. MHVM.

The purpose of the data collection/generation and its relation to the objectives of the project:

A software toolkit for approximate Bayesian computation and model updating to integrate sensor output with SG detailed Model of Tier 1 structures to incorporate monitoring data and allow rapid updating of the model properties.

A highly detailed component-by-component Model will be developed for Tier 1-2 structures in the SG simulator, together with corresponding fast-running simplified surrogate Model.

The types and formats of data that the project will generate/collect*:*

The software toolkit will be an .exe file and will generate data in format xslx.

The structural response outcomes will be in JSON or xml format.

The prognostic results will be in JSON or xml format.

The re-usability of any existing data:

The software toolkit will reuse data from various sensing devices, hazard intensity data and local climate data (previously mentioned and described).

The origin of the data:

Previously reported data will be used. The origin of most data is from the project analyses themselves, while some data on climate and asset typologies/configurations come from local surveys or publicly available databases.

The expected size of the data:

The expected size of the toolkit and the outcomes is of maximum 2GB.

The expected size of the structural response outcomes is of several MB depending on the modeled structure.

The prognostic results are of size of several GB depending on the simulation ran.

The data utility:

The software toolkit will orchestrate the multitude of structural analyses under different hazard scenarios to assess the vulnerability of high-importance (geo)structures, post-process the results and incorporate them into the HRAP environment in order to be available to the DSS of HYPERION. It will also produce adaptive (near-real-time) assessment of the health state of (geo)structures. The SG simulator results encoded in software libraries, termed MHVM, will enable a seamless integration of hazard simulators and vulnerability results into the HRAP model of the CH system.

2.13 System Risk Assessment Module

Table 13 System Risk Assessment Module

Data Category	Data type	Description
Holistic risk assessment framework	Dataset (Model results)	Results from the assessment of risk & resilience, including physical and socioeconomic impacts in each pilot-study area

The purpose of the data collection/generation and its relation to the objectives of the project:

The failure of CH and non-CH assets can propagate to affect other interconnected assets or networks. The applied modelling and simulation tools will estimate the state of CH (or its assets) depending on its previous state and/or the states of its interconnected assets. The state of an interconnected asset is thus a result of the nature of the hazard pressure affecting the originating asset, the characteristics of the asset under consideration (risk mitigation, means of immediate response, safety equipment) and the type of interconnection between the assets. The results from millions of hazard/vulnerability/impact realizations will be collected and stored to form the database that will support what-if scenarios, as well as running near-real-time trans-event assessments.

The types and formats of data that the project will generate/collect*:*

The generated data of the holistic risk assessment framework will be digital netcdf, csv, and xml files.

The re-usability of any existing data:

No further external to project data will be reused that have not previously mentioned.

The origin of the data:

Previously reported data will be integrated and analyzed.

The expected size of the data:

The generated data of the holistic risk assessment framework are expected to be of size: 2-10 GB per each pilot study.

The data utility:

The holistic risk assessment framework will pertain to how risk and impacts are propagated. The resulting data will form the basis for all impact/risk/resilience assessment studies of the pilot areas.

2.14 Socioeconomic Resilience Engine

Table 14 Socioeconomic Resilience Engine

Data Category	Data type	Description
Socioeconomic model		Mathematical model
Resilience framework		Software

The purpose of the data collection/generation and its relation to the objectives of the project:

A socioeconomic model of users (residents and visitors), local economy (production and consumption of goods, services), small businesses, and local governance will be generated

The types and formats of data that the project will generate/collect:

Directed surveys, expert opinion elicitation, end user input and literature data will be collected to provide the necessary information for model building.

Existing resilience frameworks will also be assessed in order to propose the format and potential components dealing with the basic parameters of CH-based resilience (anticipation, capacity building, absorption, coping, restoration and adaptation).

Sensor data and observer reports will also be integrated to provide a best-available rapid report of CH core status at the system level.

The re-usability of any existing data:

Existing resilience frameworks found in literature will be reused.

The origin of the data:

Apart from the existing resilience frameworks that will be employed after literature review all other data will be generated by project activities, based on expert opinion, parther and user input, including local surveys

The expected size of the data:

The expected generated data from the Socioeconomic, Community and Organizational Resilience engines are expected to be of size ~10MB per pilot.

The data utility:

The model will offer a hierarchical model of the function and of the CH core community. The overall output will be a resilience framework that will allow CH sites' operators and managers, cultural authorities, policy makers, etc. to assess the complete resilience of an entire CH area, its assets and users.

The overall framework will be encoded on top of the structure/infrastructure resilience software, completing the integrated HRAP engine that will support the simulations and form the basis for the development of the enterprise-level HRAP tool.

2.15 Impact Assessment & Mitigation Tool

Table 15 Impact Assessment & Mitigation Tool

Data Category	Data type	Description
Damage/Loss impact	Dataset	Loss calculation after applying consequence model to damage levels estimated in another module
Damage/Loss impact	Mathematical model	Impact of mitigation strategies on loss estimation

The purpose of the data collection/generation and its relation to the objectives of the project:

The impact estimation is among the final products of the platform; it will quantify the consequences of given perils on CH and non-CH assets in terms of direct losses,

downtime, and casualties. Here we will assess the effects of some scenarios considering or not the influence of a series of mitigation strategies (financial or physical); information coming from the Social Vulnerability and the Business Continuity modules will be integrated.

The types and formats of data that the project will generate/collect*:*

Table summarizing the different outputs from the simulations in csv or xml format.

The re-usability of any existing data:

Literature review and existing data about damage and its link with loss will be considered as reference.

The origin of the data:

Data coming from Socioeconomic Resilience Engine Module, Business Continuity Module and System Risk Assessment Module will be used. Some data will come from surveys and expert opinion, but most will be generated from the analyses run.

The expected size of the data:

Total of about 1 to 5 GB of data.

The data utility: The products of this module will be encoded in the structure/infrastructure resilience software, completing the integrated HRAP engine.

2.16 Drone-based and satellite data

An efficient and comprehensive monitoring system of historic areas including multiscale monitoring ranging from satellite to ground inspection will be developed combining forces of ground, drone based and satellite sensors. The data provided by the aforementioned sensors are listed below.

Table 16 Drone-based and satellite data

Data Category	Data type	Description
Satellite images	Raster dataset	SAR Sentinel 1A and 1B images (open access data) Multispectral very high resolution (World View, QuickBird) images (not publicly available)
Multispectral (RGB) images	Raster dataset	Ground and/or aerial images acquired using a drone
Hyperspectral images	Raster dataset	Ground and/or aerial images acquired using a drone
Thermal images	Raster dataset	Ground and/or aerial images acquired using a drone

The purpose of the data collection/generation and its relation to the objectives of the project:

Satellite imagery will be collected and processed in order to:

- Identify hazards with slow or gradual onset (e.g. subsidence)
- Define the risk factors that have the potential to cause harm (e.g. increase of the impervious surfaces around the historic area)
- Provide warnings for disasters with acute onsets (e.g. cyclones, storms)
- Assess their impact in the area after the event (e.g. ground deformation after earthquakes, flood monitoring) in the broader area of the cultural heritage assets.

Multi-copters equipped with multispectral camera will provide data for monitoring not easily accessible surfaces and enable the 3D monitoring of the historic area. Multi/hyperspectral and thermal sensors will monitor the surfaces of selected structures.

Images generated by passive (multispectral, hyperspectral and thermal cameras) and active sensors (SAR), will be optimally processed and analyzed using time series analysis, photogrammetric analysis, SAR differential interferometry, and advanced ML, such as deep CNN methods.

The types and formats of data that the project will generate/collect*:*

The satellite images will be collected in any of the following formats: Sentinel product (.safe, .zip), JPEG2000 (.jp2), GeoTIFF (.tif).

The multispectral data will be generated in JPEG2000 (.jp2) or GeoTIFF (.tif) format.

The hyperspectral data will be generated in ENVI data product (.hdr), PCIDSK (.pix) or TIFF (.tif) format.

The thermal data will be generated in TIFF (.tif) format.

The re-usability of any existing data:

Sentinel 1 and 2 are publicly available. Due to their high volume they will be kept outside the HRAP system, however they can be reused from sources inside and outside of the project. The very high resolution satellite data hold copyrights that have to be respected. All other data will be project generated and available for reuse from sources inside and outside the project.

The origin of the data:

SAR satellite imagery will be collected from Copernicus EO services while very high resolution multispectral satellite images will be purchased and are subject to public access restrictions. All ground and/or aerial data will be generated within the project's lifecycle.

The expected size of the data:

The expected size of the satellite imagery will be about 10Tb (depending on the time span of the time series).

The expected size of the multispectral camera data is 1Tb per acquisition

The expected size of the hyperspectral camera data is 500Gb per acquisition

The expected size of the thermal camera data is 100Mb per acquisition

The data utility:

The collection of ground, drone-based and satellite data will generate regularly updated reference data to support the risk hotspot identification and to establish a monitoring benchmark on which methods to predict hazard event outcomes can be developed.

2.17 Response Actions

Table 17 Response Actions

Data Category	Data type	Description
Response action data	Text files	Survey forms filled by Hyperion member after interviewing pilot area end users

The purpose of the data collection/generation and its relation to the objectives of the project:

The objective is to collect detailed data on the response actions to be undertaken by public authorities and private stakeholders in the event of emergencies.

The types and formats of data that the project will generate/collect*:*

The data will be extracted from online forms structured ad-hoc for the purpose of the project. Data can be also on paper format but will be uploaded online anyway after the survey.

The re-usability of any existing data:

Existing data from municipality or local authorities will be employed to improve the survey quality and coverage.

The origin of the data:

Publicly available data from local authorities, as well as data originated from human personal responses to questionnaires.

The expected size of the data:

- Paper format data: *.pdf files from survey form scan, about 500 KB each.
- Excel spreadsheet including all the digitalized surveys, about 500 KB per pilot city
- Total of ~50 MB of data at most

The data utility:

It will allow us to form a set of response actions from all involved stakeholders and map the potential paths of response in the near-real-time and post-event recovery phase for each pilot city.

2.18 HRAP Platform

The design and development of HRAP will include authoring tools to design the CH interdependences logic in terms of functional flow block-diagrams, a clearly defined plug-in mechanism where new algorithms/analyses can be added anywhere along the analysis workflow enabling scientists to create new end-to-end analyses or to enhance existing analyses, modelling various hazards impacts on CH, developing risk reduction strategies and implementing adaptation strategies to minimize their impact on societies.

Table 18 HRAP Platform

Data Category	Data type	Description
HRAP validation results	Dataset	Scenario results to validate HRAP and assess its capability to simulate community resilience

The purpose of the data collection/generation and its relation to the objectives of the project:

The overall expected outcome of the HRAP platform is to design the CH interdependences logic for end-to-end analyses resulting in risk reduction strategies and implementing adaptation strategies. HRAP will be validated by running some typical scenarios whose expected outputs can be inferred from simple mathematical models or historical data, where available.

The types and formats of data that the project will generate/collect:

The HRAP will generate risk and resilience assessments per asset, or collection of independent or interdependent assets. The outputs are encoded in csv, xml, and netCDF files.

The re-usability of any existing data:

Available historical data will be employed to act as a benchmark for the HRAP tests

The origin of the data:

Publicly available data will be employed as benchmark

The expected size of the data:

The outputs of HRAP are expected to be of size 100MB per scenario tested

The data utility:

Data that will be used from previously mentioned sources will help in testing and validating the platform.

2.19 Business Continuity Models

HYPERION will introduce a consistent adaptation framework, supported by the HRAP platform, allowing for the definition, authoring, examination and consistent comparison of a large number of different scenarios that will allow CH to be able to better cope with CC, such as on demand planning as mitigation strategy, structural/organizational changes.

A systematic procedure to align different adaptation measures with the needs of the regional stakeholders to assure the optimal use of the systems (tools) will be developed. Therefore, general conditions of the Member States, regional and local authorities, CH operators are matched with the policy targets and optimal adaptation measures.

Table 19 Business Continuity Models

Data Category	Data type	Description
Business continuity model outputs	Digital files (csv, xml)	Model outputs from business continuity simulations for each pilot area
Business continuity historical data	Digital files (csv, xml)	Historical data from past disruptions in the pilot area

The purpose of the data collection/generation and its relation to the objectives of the project:

Data that will be collected will be used in the development of the consistent adaptation framework and the systematic procedure to align different adaptation measures with the needs of the regional stakeholders.

The types and formats of data that the project will generate/collect*:*

Numerous scenarios of business interruption events will be simulated and stored for assessing post-event functionality of the pilot areas. Data from historical disruption events will be gathered and digitized to be compared with the model outputs for reasons of validation.

The re-usability of any existing data:

Any existing data from past events will be collected and stored

The origin of the data:

Outputs from business continuity model simulations under different disruption scenarios. Data from municipality, chamber of commerse, Eurostat and other publicly available sources.

The expected size of the data:

The total size of the files is expected to reach ~50MB

The data utility:

Data will be used for the analysis of a range of different methodologies to identify sector best practices that will maximise Business Continuity (BC) while minimising service disruptions to CH under different hazards pressures.

2.20 Communities' Engagement ICT Tool

The Communities' Engagement ICT tool will be based on PLUGGY's Social Platform and Curatorial Tool and will mostly employ crowdsourcing techniques to collect data at the pilots from local residents.

Table 20 Communities' engagement ICT tool

Data Category	Data type	Description
Images	Jpeg	Images provided by users
Video	Mp4	Video provided by users
Text	JSON	Stories composed by the users
Metadata	JSON	Time and space annotation for the data collected

The purpose of the data collection/generation and its relation to the objectives of the project:

Data will be collected in a way to enable citizens to create stories about the deterioration of CH sites, geo-locate the sites and also provide specific information. The data will support the business continuity models as well as the HRAP platform and the validation of the outputs from the models.

The types and formats of data that the project will generate/collect*:*

Hyperion will generate (collect) crowdsourced data from an extension to PLUGGY's curatorial tool in the form of stories, images, video or text. More specifically the dataset will be composed of geotagged and timestamped images, video and text from the pilot areas.

The re-usability of any existing data:

No data will be reused from existing sources.

The origin of the data:

The data will come from local residents using their smart phones and tablets.

The expected size of the data:

The expected size of the crowdsourced data is expected to be of size GBs. The data include all the images, video, stories and metadata.

The data utility:

The stories that will be collected will be presented in an innovative way to users, in order for them to experience the story and better understand the changes imposed by climate change and extreme events. The data will also be used in the HRAP platform by the scientists and end users for the validation of the outputs from the models.

3. FAIR data

Making research data findable, accessible, interoperable and reusable (FAIR), aims to ease knowledge discovery and innovation as well as to allow data and knowledge integration and reuse.

To this end, HYPERION will ensure to:

i) make the outcomes of the research and innovations of the project openly accessible and findable whenever possible, and

ii) disseminate and communicate the findings of the research available not only through public deliverables as defined but also, with other means as promoted or collaborative events.

The following sections provide the key considerations towards following the FAIR data policies that underpin the usage of HYPERION's data.

3.1 Making data findable, including provisions for metadata

The project strongly focuses on making sure that generated data will be identifiable and easily discoverable. Consequently, it will use established standards for the generated data whenever possible.

An overview the data types that will handled within the context of the project is described in detail in section 2 of the present deliverable and will be refined on project's progress according to the needs. Most of the data are expected to include a unique ID and a timestamp allowing for proper indexing and handling when stored.

Datasets created will follow a unique naming convention that contains: *Date, Dataset Name, Project Acronym and Version Number*. A Digital Object Identifier (DOI), a persistent identifier used to uniquely identify objects, can be created to be used a solid reference to the dataset for future use. Related keywords will be added to the metadata to enhance dataset search ability. Search keywords and clear version numbers will be also provided so as to optimize possibilities for re-use.

Data related to the end users such as pseudonym, email, country, password, GDPR-oriented consent, and farmer's identifier, will be internally kept and not published and shared externally.

Regarding documents/deliverables, clear and harmonized naming conventions will be used. All project deliverables will be named following the following nomenclature.

Project name	"HYPERION"
Space, dash, Space	" - "
Next 3-4 digits following a pattern	"DX.X" + with X.X representing deliverable number according to Description of Action (DoA).
Space	" "

Deliverable title as in DoA	E.g. “Data Management Plan”
Space	“ “
Deliverable version following a pattern (if needed)	"VX.X" with X.X representing revision number of the deliverable.

Example:

“HYPERION D1.1 Project Toolbox v1.0”, means the version 1.0 of D1.1 deliverable, entitled Project Toolbox. Deliverables that have been defined in the DoA as Public will be provided in the project’s website (https://www.hyperion-project.eu/) after their being reviewed and approved by the EC, so that anyone may access them. Apart from the title and their short description, they will include search keywords in their title page. Regarding the deliverables that are confidential and their content is restricted, the executive summary of the deliverable will be placed in the project website after the EC acceptance.

3.2 Making data openly accessible

The end users’ pseudonymous personal data will be stored in dedicated databases, hosted by relevant partners and accessible only to certain authorized users. Any data that will be shared externally to the consortium will be registered in the Registry of Research Data Repositories[3]. Specific scripts will be developed and published for accessing project’s data and offering basic statistics depending on it.

Open datasets will be also uploaded to open access repositories like Zenodo[4]. Links to the open shared datasets will be available in the HYPERION Web site.

Access to data will be enabled through the use of Open APIs and interoperable formats. All legal and other restrictions will be clearly outlined in the metadata.

All project deliverables will be available to authorized consortium members the project through the internal project management tool and document repository, Redmine.

The public project deliverables and the executive summaries of deliverables which are marked as non-public published in the project website and will also be made available through ResearchGate[5].

Moreover, HYPERION will follow the Open Access practice[6] of providing online access to its scientific research articles.

[3] https://www.re3data.org/

[4] https://zenodo.org

[5] https://www.researchgate.net/

[6] https://ec.europa.eu/research/openscience/index.cfm?pg=openaccess

3.3 Making data interoperable

Interoperability aspects have been considered, aiming to enable the maximization of the value of the data provided by the project through the utilization of common systems for transmitting and/or exchanging information. All project developed tools will be based on open source software to facilitate their adoption, expansion, and possible modifications. Interoperability of data will be enabled through the use of standardized data. Database datasets will be exposed in a text format following well-known and established standards (e.g., CSV, JSON or XML).

Any script developed aiming to externally offer data, will make use mainly of the HYPERION's APIs that will be open and thoroughly documented to enable and encourage re-use from every third-party application without forcing any dependencies.

All security mechanisms will rely on information security industry standard format and representations (e.g. for malicious traffic signatures or incident reporting), to allow for interoperability with associated mechanisms and tools widely available in the field. For example, for cybersecurity related data, such formats could include the MISP open standards[7], as well as the Snort rules format[8] which will be used in the tools.

3.4 Increase data re-use (through clarifying licenses)

Appropriate licensing schemes will be applied to project's data. By default, the open access data will be made available to the public for reuse for scientific purposes and they will be licensed under the provisions stipulated in the Grant Agreement.

Data to be shared externally will be offered under Creative Commons License Attribution-Non-Commercial CC BY-NC[9], Open Data Commons[10] and ODbL[11]. The source code of the scripts to be developed for data sharing, will be licensed under Apache v.2 license[12] and will be made available on a public repository along with detailed documentation of usage.

Appropriate licensing project generated data will allow for re-usability by third-party applications. The infrastructure as will facilitate resources utilization in a standardized way as previously mentioned.

Other than the conditions imposed by this license, no other restrictions for re-usage by third parties are envisaged. No embargo period is foreseen. Data is intended to remain reusable for a period of at least 2 years after its publication while also being offered to repositories and portals such as Zenodo for long-term sustainability.

[7] https://www.misp-project.org/

[8] https://www.snort.org/rules_explanation

[9] https://creativecommons.org/licenses/by-nc/4.0/legalcode

[10] https://www.opendatacommons.org/

[11] http://opendatacommons.org/licenses/odbl/

[12] https://www.apache.org/licenses/LICENSE-2.0

Moreover, the data generated from HYPERION will be also made available for re-usability either through the relevant APIs for at least two years after the end of the project.

4. Allocation of resources

To this point, there are no resources specifically allocated for making project's data FAIR.

It is expected that no extra-costs will be incurred. As equipment and the relevant maintenance costs are concerned, these will be, for the lifetime of the project, covered by the consortium.

Regarding personnel costs, we expect that making data FAIR shall not demand any extra calculable time of effort and thus, any personnel hours dedicated to this will be counted within the Person Month dedications to the respective tasks.

In order to proactively mitigate the risk of partners overspending due to the efforts of making any HYPERION data FAIR, WP leaders should be subjected to continuous monitoring by the Project Coordinator through the periodic management reports, the physical meetings, etc.

The Consortium will take advantage of the fact that costs associated with open access to research data can be claimed as eligible costs of any Horizon 2020 grant, and if necessary, reallocation of resources between partners will take place.

Moreover, it should be noted that 3 person months that have been allocated to Task 1.4: "Knowledge and Information Management" which concerns mostly data related matters.

More information regarding the resources for internal data and for data to be shared with external entities are provided in the next sections 5.1 and 5.2 respectively.

4.1 Resources for internal data

The internal data to be exchanged within the consortium are related to deliverables, internal reports, minutes of meetings and teleconferences, agendas, templates, presentations, administrative documents and files of any kind in general. The project will make use of Redmine, a free and open source, web-based project management and issue tracking collaborative tool in order to facilitate information exchange, storage, ordering and retrieval as needed in the project. Redmine will be used as the common document repository of the project and is already configured and runs under the responsibility of ICCS, the project coordinator. All consortium members are actively using this interactive knowledge-sharing platform for information exchange, discussions, news and calendar. This collaborative tool and relevant storage structure are consortium confidential with restricted access.

Since Redmine is free of cost, the only resources consumed concern only the setting and deployment of the collaborative platform. In addition, resources relevant to administrative and maintenance issues with respect to Redmine are covered by the project coordinator, ICCS, who also has the rights for managing users, groups and assigning roles and permissions. To this end, ICCS has already provided credentials to all partners, upon their request for an account, in order for them to be able to login and use it. Further to this, for the needs of continuously upgrading the collaborative tool and back-up its database in order to address unforeseen events (hard disks failures, web attacks, etc.) resources have already been allocated by the consortium.

4.2 Resources for data to be shared with external entities

Any need of personnel hours' allocation for data stewardship is already part of the allocated person months per Work Package, as Work Package leaders are the main responsible for such activities, with the support of all other partners involved in the certain Work Package and in particular with the overall guidance and support of the project's consortium that has budgeted for dissemination costs. HYPERION will follow the Open Access practice[13] of providing online access to scientific information that is free of charge to the end-user and reusable.

This covers both peer-reviewed scientific research articles (published in scholarly journals) and research data (data underlying publications, curated data and/or raw data).

In the cases where publication in a scientific journal shall be selected as a means of dissemination of the project activities, one of the two channels of open access shall be chosen (evaluated per case), i.e. either:

- Self-archiving / 'green' open access – the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online (either institutional, or centralized such as Zenodo) repository before, at the same time as, or after publication.
- Open access publishing / 'gold' open access - an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers. The most common business model is based on one-off payments by authors/partners.

HYPERION will also consider the possibility to make use of free of charge, open data repository services, such as Zenodo. Through such repositories, metadata and digital object identifiers could be assigned to the project's datasets in order for them to be located via search after the end of the project. The datasets in Zenodo will be preserved in line with the European Commission Data Deposit Policy. The data will be maintained indefinitely (minimum 5 years) ensuring no costs for archiving. It should be noted that any unforeseen costs related to open access to research data in Horizon 2020 are eligible for reimbursement for the duration of the project.

[13] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

It is also worth mentioning that the costs for making data FAIR highly depends on the yet to be specified details of the amount data that will be collected and their processing effort. In any case, all costs related to FAIR data management that will occur during project implementation will be covered by the project's budget. However, any other cost that may relate to long term data preservation will be discussed among consortium members, but as stated before the services of free of charge research data repositories will be pursued.

5. Data security

All data will be collected, stored, secured, transmitted, preserved and destroyed in accordance with state-of-the-art security measures and in full compliance with relevant EU legislation, following the requirements of the Protection of Personal Data (POPD) defined by the EC and national regulations. A precise presentation of the POPD requirements in the context of HYPERION is included on deliverable "D11.2 POPD - Requirement No. 2". As a general rule, this data will be stored on paper and/or electronic files secured by state of the art physical and digital protection measures:

- archives containing paper directories will be locked
- digital information will be stored on servers
- hard disks and cloud repositories accessible only to approved staff (within the applicable DIONE partners) via user and password schema identification

Personal data obtained and processed for the purpose of project activities such as interviews, focus groups, etc. will be completely and irrevocably deleted upon completion of the project. However, conclusions deriving from this data will only be included in the project outputs if the individual participant has given his / her free, precise and informed consent.

Data are acquired in conformity of security best practices through protected connections and dedicated state-of-the-art IT infrastructures (LANs, protected servers, firewall).

To guarantee security and reliability of the communication, the most advanced web standards are implemented:

- HTTPS transport level security is applied.
- Authentication and authorization are required both for user interactions with the system and for machine-to-machine transactions.
- Access and interactions are controlled and audited.

All the documents, deliverables and reports generated in the context of the project are deposited in Redmine. The online access to https://redmine.iccs.gr/ is protected using a TERENA (https://www.terena.org) digital certificate. The server hosting the Redmine installation is located in ICCS premises in Athens, Greece in a safe rack in ICCS's server space. Server databases are backed up on a regular basis, although its files are backed up every second day. The server is designed with numerous redundancies, network and disk wise, in order to ensure its continuous operation and network access.

6. Ethical aspects

The project's consortium adheres to ethical rules and complies with European legislation on data protection (Regulation (EU) 2016/679 General Data Protection Regulation[14]), the national legislation applicable in countries where the research will be carried out, as well as recommendations and codes of conduct relevant to research activities.

There have not been identified any specific ethical issues related to the activities of the project that are not already addressed in the Grant Agreement. Ethical procedures have been specified within the project, throughout the two dedicated ethics deliverables (and disseminated between consortium members). These procedures will have to be followed in project activities. The activities related to the user requirements elicitation and the implementation and assessment of the pilot demonstrations will be designed and implemented while taking into consideration the dignity of the participants as well as other fundamental rights and freedoms (freedom, non- discrimination, etc.) and core values will be respected (proportionality, minimization, confidentiality) and will conform with The General Data Protection Regulation (EU) 2016/679 ("GDPR"). As national and EU laws and recommendations on privacy and data protection issues play an important role, the design of activities within the project will involve all partners' engagement in designing, deploying and testing of HYPERION's solutions, which may raise concerns on data sharing and protection issues. The project will assess the execution of the different ethical requirements in accordance with applicable national and European legislation. Ethical standards and guidelines will be rigorously applied, irrespective of the country in which the research is carried out. Obtaining, if needed, the required national authorization and adherence will be the responsibility of the partners in each of the EU states.

The project adheres to the commitment of holding any data in secure conditions, and will make every effort to safeguard against accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data.

Additionally:

- The Consortium will evaluate a secure storage and transfer system/channel
- Research and other data will be stored in a secure and accessible form
- The Consortium will specify procedures for keeping data accessible in terms of migration (conversion of data files from older formats to newer ones) and refreshing (transfer of data from one storage tool to another), if needed
- The Consortium will define procedures for backup and recovery of data
- Access rights and access conditions will be defined taking into consideration the task allocation and the category/type of data

[14] https://ec.europa.eu/info/law/law-topic/data-protection_en

- When defining roles and permissions special attention will be paid to the possibility to track any interaction that entails access, modification and deletion of personal data

Each partner is responsible for informing their own staff involved in the project about the need to comply with the legal principles and provisions with regard to data processing. The Consortium considers all necessary and appropriate measures to mitigate risks, which include:

- The handing over of any data
- The collection of data and its secure storage and transfer
- The confidentiality declaration to be signed by staff
- The process for conducting pilot demonstrations, participant's recruitment and gaining informed consent

6.1 Informed consensus

Since the very beginning of the project, participatory methods will take place, involving the conduction of questionnaires, interviews, focus groups, workshops, and combination of previous methods towards the elicitation of end-users' needs and good practices analysis from relevant stakeholders. Prior to being engaged in all these participatory methods, informed consent forms will be provided and signed by the participants. A GDPR compliant version of the informed consent forms is included on deliverable D11.1 "H - Requirement No. 1".

In the context of HYPERION and according to the GDPR a lawful basis for processing of personal data exists when the individual (the data subject) has given clear consent to process their personal data for a specific purpose. The informed consent forms were designed in compliance with the principles and provisions of GDPR regulation, while emphasizing on:

- The use of concise and clear language
- The option for people to actively opt-in
- Explicitly stating the right for people to withdraw the consent at any time, while describing a clear withdrawal procedure
- The provision of clear and detailed information about the purpose and means of processing of the data (i.e. description of the procedures that will be implemented for data collection, storage, protection, retention and destruction), including the details of the data controller and the data processor

A consent management process has been discussed and established based when and how consent will be acquired from individuals. In case of written consent, a copy of the relevant document is kept. If consent is given online, the records shall include the data submitted as well as a timestamp to link it to the data capture form. Cases where the consent is withdrawn will also be recorded. Regular consent reviews will be established during the pilot demonstration activities.

Furthermore, the consortium shall respect the data subject's right to be informed through the provision of the following information to the participants during their engagement with the project's activities:

- Identity and contact details of the Data Controller (or the controller's representative) and the Data Protection Officer
- Purpose and lawful basis of the processing
- The legitimate interests of the controller or third parties, if applicable
- Any potential recipients of the personal data
- Details and safeguards of transfers to third countries, if applicable
- Data retention period
- The existence of data subject's rights
- The right to withdraw consent at any time, where relevant
- The right to lodge a complaint on a supervisory authority
- Whether the provision of personal data is part of a statutory or contractual requirement or obligation and possible consequences of failing to provide the personal data
- The existence of automated decision making, including profiling and information about how decisions are made, the significance and the consequences, when and if applicable

In addition, participants in the project's activities shall be given the possibility to contact the data controller in order to

- request to access (right of access)
- rectify (right to rectification)
- remove/delete (right to erasure)
- restrict processing (right to restrict processing) of their personal data

The aforementioned process is described in the informed consent form, which will be made also available in the project's website. Participants can also request for a copy of their personal data to be used for their own purposes (right to data portability), while also being able to object, at any time, to the processing of their personal data by contacting the data controller.

7. Conclusions

The primary goal of the present deliverable is to demonstrate the Data Management Plan for HYPERION, given that the project consortium has opted in for the Open Research Data Pilot.

It provides a general overview of the kind of data that will be produced, reused, collected and processed within the project's context. It tackles also topics related specific challenges and constraints that need to be taken into consideration.

Issues such as data types generated and their purpose and relation to the project objectives, as well as how the consortium make sure that it will comply with FAIR data principles were discussed. Moreover, data security and generic ethical issues were also mentioned, along the same guidelines.

Most project partners will eventually be owners or/and producers of data while others will just process generated data. This implies specific responsibilities, described in this report.

The Data Management Plan emphasizes also on the appropriate collection – and publication if the data will be published – of metadata, storing all the necessary information for the optimal use and reuse of those datasets in compliance with all regulations.

This deliverable is considered a living document which will be edited and further completed, after more detailed info on the project activities and outputs will be available. The DOA foresees an official update of the Data Management Plan, deliverable D1.5 “Data Management Plan (v2)” to be submitted on month 42 presenting the final version of the present report on the project’s ending.

8. ANNEX - Public Deliverables

Deliverable number	Deliverable title	WP	Responsible	Nature	Dissemination	Date
D1.3	Data Management Plan (v1)	WP1	ICCS	ORDP: Open Research Data Pilot	Public	M12
D1.4	Societal impact report	WP1	ICCS	Report	Public	M42
D1.5	Data Management Plan (v2)	WP1	ICCS	ORDP: Open Research Data Pilot	Public	M42
D2.1	End-user needs and practices report	WP2	RG	Report	Public	M4
D2.2	Definition of System Requirements, Use Cases and KPIs specification	WP2	ICCS	Report	Public	M10
D2.4	Geographic data and services inventory	WP2	AUTH	Report	Public	M14
D3.1	Report on available climate data and scenario selection	WP3	FMI	Report	Public	M10
D3.2	High-resolution surface parameter maps	WP3	AUTH	Other	Public	M12
D3.3	Report on the dynamical downscaling of climate & atmospheric impacts	WP3	FMI	Report	Public	M24
D3.5	Report on Dynamic Data Assimilation methodology	WP3	AUTH	Report	Public	M28
D3.6	Report on site-specific risk parameters and stressor indicators	WP3	AUTH	Report	Public	M28
D4.1	Classification of building materials, physical-mechanical	WP4	IUAV	Report	Public	M24

	characterization and determination of decay forms and products					
D4.2	Assessment of hygrothermal and textural features controlling building materials decay, and required as input in the hydrothermal simulations	WP4	UNIPD	Report	Public	M24
D4.3	Analysis of microclimatic time series and assessment of systematic deviation from local time series required as input in the hygrothermal simulations	WP4	UNIPD	Report	Public	M27
D4.4	Deterioration of the building materials under extreme events required in the CH vulnerability assessment	WP4	UNIPD	Report	Public	M24
D4.5	Material specific dose response functions and validation required in the CH vulnerability assessment	WP4	UNIPD	Report	Public	M27
D5.1	Advanced and Reliable Models for Natural and Man-Made Hazards	WP5	AUTH	Demonstrator	Public	M24
D5.3	MHVM for CH and non-CH elements	WP5	UGR	Other	Public	M24
D5.4	HRAP software and documentation	WP5	NTUA	Other	Public	M32
D5.5	HYPERION resilience framework	WP5	RG	Report	Public	M32
D6.1	Conceptual framework for remote sensing-based CH monitoring	WP6	NTUA	Report	Public	M10
D6.3	Methodology for routine CH monitoring with multi-type remote sensing	WP6	NTUA	Report	Public	M24
D6.4	Novel methodologies for damage detection and assessment along the CH assets and the surrounding disaster affected area	WP6	NTUA	Other	Public	M24

D6.5	Dynamic link to hazard and resilience assessment	WP6	UGR	Report	Public	M30
D7.3	BC Models and Adaptation Strategies assessment report	WP7	RG	Report	Public	M30
D7.4	Standard Response Procedures Document	WP7	RG	Report	Public	M30
D7.7	Communities' Engagement ICT Tool	WP7	ICCS	Other	Public	M32
D8.6	Reports on pilot testing	WP8	CyRIC	Report	Public	M42
D8.7	Trials assessment and recommendations	WP8	RG	Report	Public	M42
D9.1	Corporate identity and general templates for dissemination material	WP9	IEMC	Other	Public	M3
D9.2	Project Website	WP9	IEMC	Websites, patents filling, etc.	Public	M5
D9.3	Dissemination and Communication Plan (v1)	WP9	IEMC	Other	Public	M6
D9.4	Information Packs for referenced networked communication amplifiers (v1)	WP9	IEMC	Other	Public	M12
D9.5	Annual Magazine issued (v1)	WP9	NTUA	Other	Public	M12
D9.6	Report on Standards and Liaison Activities with relevant organisations	WP9	RG	Report	Public	M42
D9.7	Dissemination and Communication Plan (v2)	WP9	IEMC	Other	Public	M42
D9.8	Information Packs for referenced networked communication amplifiers (v2)	WP9	IEMC	Other	Public	M42
D9.9	Annual Magazine issued (v2)	WP9	NTUA	Other	Public	M42
D10.2	Market Analysis	WP10	RG	Report	Public	M42
D10.3	Workshop Documentation	WP10	IEMC	Report	Public	M40

D10.4	HYPERION Roadmap and Project Handbook	WP10	ICCS	Report	Public	M42