LEONSEGS

LARGE EARTH OBSERVATION
NEW SPACE ECOSYSTEM
GROUND SEGMENT

D1.4 –DATA MANAGEMENT PLAN (DMP) V1 (REVISED)

31/01/2025

Dissemination Level: **Public**

LEONSEGS

Grant Agreement No.: 101082493
Call: HORIZON-CL4-2022-SPACE-01
Topic: HORIZON-CL4-2022-SPACE-01-13
Type of action: HORIZON Innovation Actions

D1.4 DATA MANAGEMENT PLAN

Work package	WP 1
Task	T1.1
Due date	31/01/2025
Submission date	30/01/2025
Deliverable lead	GMV
Version	V1.1
Authors	Stan Markelov (GMV)
Reviewers	Cátia Figueiredo (F6S) Viktoriya Dimova (F6S) Luis Saturnino (AISTECH) Dirk Tiede (PLUS)
Consortium Internal Code	D02_D1.4_DataManagementPlanV1 GMV internal code: GMV 26450/23 V2/25

Dissemination Level: **PUBLIC**

Co-funded by the European Union

Keywords	EO, New Space, flexible multi-mission Earth Observation ground segment, Data Management

Document Revision History

Version	Date	Description of change	List of contributors(s)
V1	30/01/225	Updated version based on feedback from EC	

DISCLAIMER

COPYRIGHT NOTICE

The LEONSEGS Consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	GMV AEROSPACE AND DEFENCE SA	GMV	Spain
1.a	GMV GmbH	GMV	Germany
1.b	GMV Romania	GMV	Romania
2	PARIS-LODRON-UNIVERSITAT SALZBURG	PLUS	Austria
~~3~~	~~SATELLOGIC SOLUTIONS SL~~	~~SATL~~	~~Spain~~
4	F6S NETWORK LIMITED	F6S	Ireland
5	**AISTECH SPACE SL**	**AISTECH**	**Spain**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. EXECUTIVE SUMMARY

1.1 PROJECT INTRODUCTION

LEONSEGS is a federated environment (called Multi-mission Earth Observation Ground Segment Service Platform) of EO (Earth Observation) data providers that collaborate all together through harmonized interfaces as depicted in the diagram below.

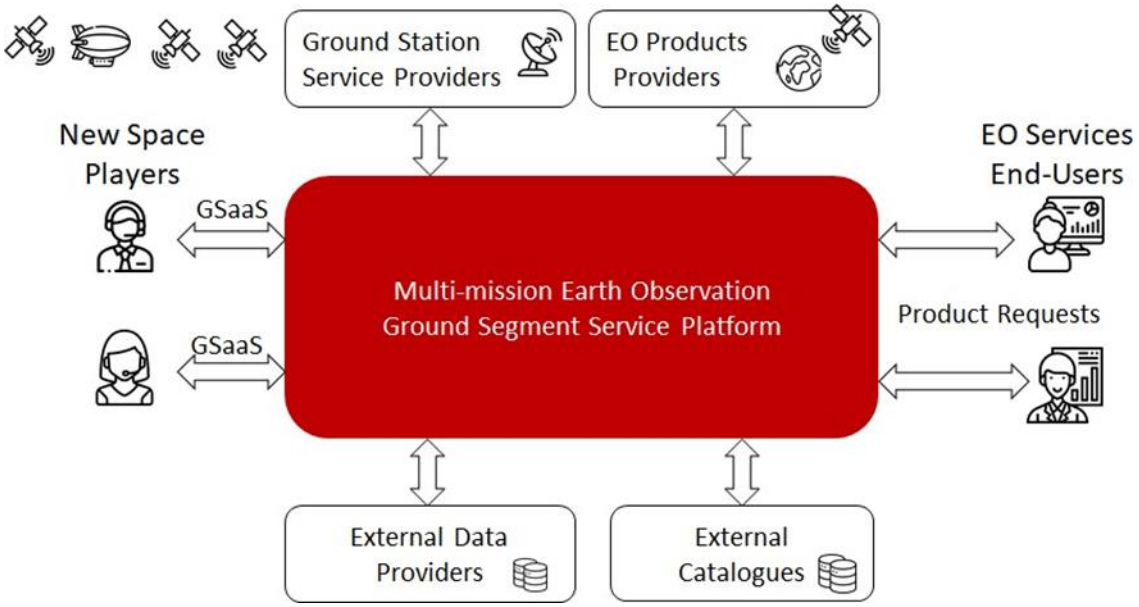

FIGURE 1-1: LEONSEGS HIGH-LEVEL ARCHITECTURE

The objective of LEONSEGS project is to prototype a flexible multi-mission EO ground segment able:

- to federate European New Space players through its GSaaS (Ground Segment as a Service) paradigm widening their access to a larger market whose complex requests could not be served in an isolated manner, and
- to offer optimized and sophisticated EO-based products/services to End-Users on the basis of intelligent search and best combination of heterogeneous datasets from the different federated providers and other external providers and archives.

LEONSEGS will contribute to cover the EU needs to maintain a strong industrial and technological presence in key parts of digital industrial and other supply chains, in industrial ecosystems while safeguarding its ability to access and use space. This will be a critical effort for the EU to be able to compete globally, as well as to protect its citizens, deliver services and products of the highest quality, and preserve its values and socio-economic model. LEONSEGS will master space technologies contributing to position Europe as a competitive, secure and trusted leader.

LEONSEGS project is executed by a multidisciplinary consortium which consists of the participants listed in the Table 2-1.

PARTICIPANT number	Participant organisation name	Short name	Country
1	GMV AEROSPACE AND DEFENCE SA	GMV	Spain
1.a	GMV GmbH	GMV	Spain
1.b	GMV Romania	GMV	Spain
2	PARIS-LODRON-UNIVERSITAT SALZBURG	PLUS	Austria
4	F6S NETWORK LIMITED	F6S	Ireland
5	**AISTECH SPACE SL**	**AISTECH**	**Spain**

TABLE 1-1: LEONSEGS CONSORTIUM COMPOSITION

1.2 PURPOSE AND SCOPE OF THE DOCUMENT

This LEONSEGS Data Management Plan (DMP) has been developed to be compliant with Horizon 2020 DMP guidelines. The primary goal of this DMP is to describe an overall approach to how LEONSEGS project will collect, store and process **research data.**

This document describes the mechanisms to handle research data during the project execution as well as after the project completion. It describes how data will be created or collected, stored and backed-up, who owns it and who is responsible for the different types of data. Specifically, the present Data Management Plan (DMP) contains information about:

- the handling of research data during the project

- what data will be collected, processed and/or generated
- which formats these data have
- which methodology will be applied in their collection, process or generation
- whether data will be shared/made open access, and
- how data will be stored and preserved

The document structure and contents are based on the Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 (v3.2, March-21 2017). Specifically, sections 3-2 are dedicated to the data organisation and data description with the level of detail appropriate to the maturity of the project and version of this DMP.

Therefore, it is not required to provide all detailed information in the first version of the DMP (i.e. this version). Rather, the DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and/or when significant changes occur.

Considering the afore mentioned, the DMP release timeline is as follows:

Deliverable No	Deliverable Name	Dissemination Level	Due Date	Comments
D1.4	Data Management Plan v1	PU - Public	31/01/2025	Per GA a contractual deliverable
D1.5	Data Management Plan v2	PU - Public	30/04/2025	Per GA a contractual deliverable
NA	Data Management Plan v3	PU - Public	None	Optional update of the DMP if new information becomes available

2. APPLICABLE AND REFERENCE DOCUMENTS

2.1 APPLICABLE DOCUMENTS

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority.

Ref.	Title	Code	Version	Date
[AD.1]	Grant Agreement Nº 101082493 - LEONSEGS	HORIZON-CL4-2022-SPACE-01	V1.0	14/06/2023
[AD.2]	LEONSEGS Consortium Agreement,	GMV 30700/23	Ver. 1	May 2023

TABLE 2-1. APPLICABLE DOCUMENTS

2.2 REFERENCE DOCUMENTS

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.X]:

Ref.	Title	Code	Version	Date
[RD.1]	Guidelines on Implementation of Open Access to Scientific Publications and Research Data	https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf	Ver 1.1	21 Apr 2017

TABLE 2-2. REFERENCE DOCUMENTS

2.3 ACRONYMS AND DEFINITIONS

The following acronyms have been used across this document:

Acronym	Full term
AD	Applicable document
AOI	Area of Interest
CFS	Certificate on the financial statement
CCSDS	Consultative Committee for Space Data Systems
CI	Configuration Item
CIDL	Configuration Items Data List
CM	Configuration Management
CM	Configuration Management
CMMI	Capability Maturity Model Integration
CN	Change notice
CR	Change request
DMP	Data management plan
DO	Dataset Owner
DOI	Digital Object Identifier
DWH	European Space Agency Data Warehouse
DWP	Dataset Work Package
EEA	European Environmental Agency
EO	Earth Observation

EU	European Union
FTP	File Transfer Protocol
GA	Grant Agreement
GDPR	General Data Protection Regulation
GIS	Geographical Information System
GMV-SGC-PRO-005	Project Management Procedure
GSaaS	Ground Segment as a Service
IT	Information Technologies
ISO	International Organization for Standardization
KOM	Kick-off meeting
KPI	Key Performance Indicator
LIDAR	Light Detection and Ranging
LSA-SAF	EUMETSAT Land Surface Facility – Satellite Application Facility
MOCC	Massive Open Online Course
MOM	Minutes of Meeting
MTR	Mid Term Review
NA	Not Applicable
OGC	Open Geospatial Consortium
ORDP	Open Research Data Pilot

OS	Open source
OTH	Other items (WRT the project's documentation nomenclature)
PDF	Portable Document Format
PDPA	Payload Data Processing and Applications Business Unit at GMV
PoC	Point of contact
POC	Proof Of Concept
PR	Progress Report
RD	Reference document
REA	EC Research Executive Agency
RS	Remote sensing
RTD	Research Technology and Development
SB	Stakeholders Board
STAC	SpatioTemporal Asset Catalog
SWOT	Strengths, Weaknesses, Opportunities, and Threats Analysis Method
TBC	To be confirmed
TBD	To be defined
ToR	Terms of reference
TRL	Technology readiness level
USGS	United States Geological Survey

VHR	Very High Resolution
WBS	Work Breakdown Structure
WKS	Workshop
WP	Work Package
WPD	Work Package Description

TABLE 2-3. ACRONYMS

Concept/Term	Definition
List of configuration items	Catalogue of configuration items in a reference line or delivery (Configuration Item Data List, CIDL).
Configuration	Product Description provided from the amount, nature, and interconnections of its parts, allowing thus to know exactly and completely how the product has been conceived, designed, and created.
Configuration Management/Control	Configuration management item consisting in the assessment, coordination, approval, or disapproval and implementation of changes in configuration items after having established formally their configuration identification.
Configuration Status Accounting	Registering and reporting the data needed to manage effectively a configuration.
Life cycle of a Product	All phases in the life of a product, from needs identification through disposal.

TABLE 2-4. DEFINITIONS.

3. DATA SUMMARY

The purpose of this section is to provide a detailed information related with the LEONSEGS project's data collection and creation, definition of the datasets, their purpose, types, formats, and the origin of the datasets. In this first version of the document, not all identification of the data is included. This information will be improved upon in subsequent versions of the DMP.

LEONSEGS' datasets described below are related only to research data. Thus, a dataset was grouped by scientific and/or research topic with the following naming convention:

LDS-x_Y.zz , where

LDS – LEONSEGS DataSet;

x is a sequential numbering for a dataset;

Y – is data set owner ID that can take these values:

G (for GMV), P (for PLUS), S (for SATL), A (for AISTECH), M (for multiple owners) ;

zz is a version of dataset that will increment sequentially as updates are made to the dataset (e.g. LDS-1.1).

The above-mentioned datasets organisation and grouping are presented in the Table 3-1 below.

Data set ID	Related WP Description	Dataset Owner	Dataset ID (without version #)	General Description of Data	Related WBS task
LDS-1	System and Ground Segment Engineering	GMV	LDS-1_G.01	Data related to Prototyping the Multi-mission Control Center and interfaces	T3.2
				Data related to Prototyping the Automated Multi-mission EO Service and Interfaces	T3.4

Data set ID	Related WP Description	Dataset Owner	Dataset ID (without version #)	General Description of Data	Related WBS task
		PLUS	LDS-1_P.01	Data related to Developing a Fitness-for-Use Index and Semantic Querying of Multi-mission Archives	T3.3
LDS-2	End-to-End Trials Preparation, Coordination and Execution	PLUS	LDS-2_P.01	Data related to POC Assessment	T4.4
		GMV, PLUS, EC, SATL, AISTECH	LDS-2_M.01	Data related to POC Validations	T4.2

TABLE 3-1. DATASET ORGANISATION.

As described in Table 3-1 and at the time of preparation of this first version of the Data Management Plan, LEONSEGS project has 4 datasets that will be described in the following sections.

3.1 DATA COLLECTION AND CREATION

The main purpose of the data collection and management is the adequate implementation of the project, as defined in the Grant Agreement (GA).

With regard to the **type,** the data could be:

- Observational: this type of data is collected in situ and cannot be recreated. Monitoring data or results from surveys can be categorised in this group.
- Experimental: these data are generated under controlled conditions and could be recreated.
- Simulated data: these data are generated by the execution of a simulation model that recreate the behaviour of a theoretical system, fed with specific input data.

With regard to the **origin**, the data could be:

- Collected from previous activities (previous projects, public data from public sources, scientific papers…).
- Generated in the project (from experimentation, simulation, or formal demonstration).

The **formats of the data** could be very diverse depending on type or origin of the dataset (different data formats and different file formats). Non-proprietary, non-compressed and open formats are preferred. The standard file format for reports will be in PDF. However other file formats could be used, such as .xls or .csv, in the case of experimental data or Microsoft Office standard (.pptx, .docx…) in the case of dissemination activities. Large datasets (like images,) will be organised and stored with more appropriate tools such as databases or file formats that will allow more efficient access to the data. Unless documentation is protected under intellectual property rights, such as copyright, it will be made freely available under Creative Commons License.

Partners confirmed that the data used are publicly available and do not involve a further processing of previously collected personal data (including use of pre-existing data set or sources, merging existing datasets). If an unexpected further processing of previous personal data (also known as secondary use) arises, an adequate lawful basis for processing that data and complying with data minimisation principle will be used.

If the original consent provided by the data subject does not allow the use of their personal data in the LEONSEGS project, such information will not be used. As a mitigation measure, we will gather the permission of the data subject, allowing their personal data to be used within the specific purposes of our project.

Regarding the **utility of the data**, LEONSEGS project data will be useful for:

- European Commission can benefit from the project datasets to be able to assess EU technological capabilities, its associated impacts, and future opportunities.
- Researchers can obtain data for analysis and research
- Private sector/stakeholders can benefit from datasets finding new business opportunities.
- Citizens/general public can benefit from the project datasets to be aware of EU technological capabilities, its associated impacts and future opportunities.

3.2 DESCRIPTION OF DATASET

In the following subsections four (4) datasets identified in section 3 (Data summary) are presented, with information about the dataset's relationship with:

- the WP / WP Leader,
- Dataset owner,
- WBS task,
- expected data exploitation, and
- data preservation and archiving.

The dissemination of each dataset will be decided by the Dataset Owner (DO) responsible for the management of the respective dataset. The datasets will be as open as possible taking into account the restrictions due to protection of personal data or intellectual property rights.

3.2.1 DATASET LDS-1_G.01

Data Management Topic	Description
Description of the datasets generated or collected	Data related to Ground Segment Engineering, and Prototyping of the MOCPS and AMEOS (T3.2, T3.4). The datasets to be generated will include any innovative approaches, designs and algorithm development that are related to the definition of the LEONSEGS architecture, interfaces and EO services to be provided by LEONSEGS prototype.
Dataset Responsibilities and Dependencies	Responsibilities: The Dataset owner is GMV. All partners contribute to the design of the LEONSEGS high level architecture. Dependencies: the design of the LEONSEGS architecture has a strong dependency on the design of each subsystem.
Data sharing and exploitation	The dataset generated is partially public. Only data defined as non-proprietary will be shared and open to public access. At the current stage, the data are shared within the consortium for the purpose of the implementation of all other WPs (WP2-WP4). The final format of this dataset will be in Microsoft word or PDF formats.

Data Management Topic	Description
Archiving and preservation of the data	At the current stage, the datasets generated described above will be stored in SharePoint/Teams platform that is shared with all partners; in addition, information will be stored on GMV internal servers. GMV servers comply with ISO 20000 and ISO 27001 standards as well as the General Data Protection Regulation (GDPR). Internal procedures will be followed for long term preservation and curation.

3.2.2 DATASET LDS-1_P.01

Data Management Topic	Description
Description of the datasets generated or collected	Data related to Developing a Fitness-for-Use Index and Semantic Querying of Multi-mission Archives (T3.3): For the demonstration of the semantic querying, information/data will be derived from dataset LDS-2.M-01 (Dataset 2 (Sentinel-2) and combination of Dataset 4 (Satellogic) and Dataset 6(AISTECH)) for the selected AOIs and the selected time frames. This will result in semantic enrichment of the initial data (= labelling of each pixel with spectral categories). The Fitness-for-Use Index will work on existing metadata of the dataset LDS-2.M-01.
Dataset Responsibilities and Dependencies	Responsibilities: The Dataset owner is PLUS. Dependencies: The semantic enrichment is based on the dataset LDS-2.M-01 (Dataset 2 (Sentinel-2) , Dataset 4 (Satellogic) and Dataset 6 (AISTECH)) and associated with these datasets within a data cube approach.
Data sharing and exploitation	Data will be shared within the consortium for demonstration and POC. STAC will be used for catalogue-based access. Subsets will be accessible for the webinar/MOOC implementation. The data can be

Data Management Topic	Description
	shared with interested parties for outreach and promotion of the project results.
Archiving and preservation of the data	As long as an infrastructure grant is in place (cloud resources), the data will be preserved. After that: The workflows are documented and repeatable. Dataset can be reproduced based on the open and freely available Sentinel-2 data as well as on the Satellogic and AISTECH data sets acquired within this project.

3.2.3 DATASET LDS-2_P.01

Data Management Topic	Description
Description of the datasets generated or collected	Data related to POC assessment (T4.4). Data generated within this task encompass AOI specific implementations/use case definitions, end-user requests / feedback, as well as scientific results regarding validation and user involvements.
Dataset Responsibilities and Dependencies	Responsibilities: The Dataset owner is PLUS. All partners contribute within their POC involvement. Dependencies: POC execution is strongly dependent on the developed components of WP 2 and 3.
Data sharing and exploitation	The dataset generated under this WP/task is sensitive, especially in respect to user/stakeholder involvement. The information is shared within the consortium for the purpose of planning / execution / validation of the POCs. Data will be aggregated into deliverables (D4 .1-4.4 including a SWOT analysis), which are sensitive. Further sharing of the dataset is not applicable.

Data Management Topic	Description
Archiving and preservation of the data	At the current stage, the datasets generated described above will be stored in SharePoint/Teams platform that is shared with all partners. In addition, information will be stored on GMV internal servers. GMV servers comply with ISO 20000 and ISO 27001 standards as well as the General Data Protection Regulation (GDPR). Internal procedures will be followed for long term preservation and curation.

3.2.4 DATASET LDS-2_M.01

<table>
<tr><th>Data Management Topic</th><th>Description</th></tr>
<tr><td>Description of the datasets generated or collected</td><td>
The datasets to be collected and generated within tasks 4.2 and 4.3 cover the satellite data that will be collected through the LEONSEGS system and the derived EO services that will be generated. These datasets will be used for the POC validation and demonstration according to the use cases defined in WP2.

A summary table is included below with the potential satellite missions identified so far to be integrated within LEONSEGS and an image dataset for the EO services. Further information about the satellite missions and the derived EO indexes will be provided in D3.1.
<table>
<tr><th>Image Dataset ID</th><th>Image Dataset name</th><th>Image Data type</th><th>Origin</th><th>Source</th></tr>
<tr><td>1</td><td>Sentinel-1</td><td>SAR HR data</td><td>Public data from public sources</td><td>Copernicus programme</td></tr>
<tr><td>2</td><td>Sentinel-2</td><td>Optical HR data</td><td>Public data from public sources</td><td>Copernicus programme</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Data Management Topic</th><th>Description</th></tr>
<tr><td></td><td>
<table>
<tr><td>3</td><td>Landsat-8</td><td>Optical HR data</td><td>Public data from public sources</td><td>USGS</td></tr>
<tr><td>4</td><td>Satellogic</td><td>Optical VHR data</td><td>Project data</td><td>Satellogic</td></tr>
<tr><td>5</td><td>EO services</td><td>Derived EO indexes</td><td>Developed by LEONSEGS</td><td>LEONSEGS</td></tr>
<tr><td>6</td><td>AISTECH</td><td>Optical MR, NIR, LWIR</td><td>Project data</td><td>AISTECH</td></tr>
</table>
</td></tr>
<tr><td>Dataset Responsibilities and Dependencies</td><td>Responsibilities: Datasets from Sentinel and Landsat are open data publicly available by their providers. The Satellogic image datasets' owner is LEONSEGS project . The dataset owner of the EO services generated by GMV is GMV. The dataset owner of the EO services generated by PLUS is PLUS.

Dependencies: The datasets to be collected and generated within tasks 4.2 and 4.3 have a strong dependency on WP2 (use cases definition), WP3 (EO services definition and Satellite missions' integration) and the planning of the POC validation and demonstration defined in Task 4.1.</td></tr>
<tr><td>Data sharing and exploitation</td><td>The dissemination level of the related deliverables is sensitive. However, if possible, non-sensitive information such as satellite images, EO services and validation/demonstration results (and their metadata) will be publicly shared in certified repositories such as Zenodo following the procedures described in Section 8.</td></tr>
<tr><td>Archiving and preservation of the data</td><td>The datasets generated described above will be stored in SharePoint/Teams platform that is shared with all partners. In addition, information will be stored on GMV internal servers. GMV servers comply with ISO 20000 and ISO 27001 standards as well as the General Data</td></tr>
</table>

Data Management Topic	Description
	Protection Regulation (GDPR). Internal procedures will be followed for long term preservation and curation.

4. DATA ACCESS

The final objective of the LEONSEGS DMP is to provide a roadmap for research related knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

It is important to note that the majority of data created in each LEONSEGS WP cannot be publicly available with two exceptions:

1. Formal deliverables to EC that designated with a public dissemination level.
2. Images and image-related products that are identified in Sharing/Exploitation section of each dataset (section 3.2).

Scientific publications and other material will be published under a Digital Object Identifier (DOI) alphanumerical code, which allow the unique identification in a consistent way. These publication will be stored in Zenodo repository as described in Section 8 of this document.

The use of keywords is another mechanism that could help the findability of the data. Keywords are usually required by scientific journals, so partners are encouraged to use this mechanism adding the most pertinent words according to the research. In the case keyword frameworks are not required, the partners are encouraged to use relevant words in the abstract or introduction of the documents.

Wherever possible, consortium partners will identify and use standard data and metadata vocabularies to maximise the data interoperability. The consortium follows the defined terms established by the Rules for participation and dissemination in Horizon Europe 2020 established by Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe (and repealing Regulation (EU) No 1290/2013 of the European Parliament and the Council of 11 December 2013 [1].

The scientific and research types of data generated within the project will be made re-usable through the creation of documentation that helps others interpret and analyse it.

Due to the early stage of the project, it is still unclear how data will be licensed to allow the widest re-use possible and when it will be made available for re-use, but unless documentation is protected under copyright, or other kinds of intellectual property rights, all publications will be made freely available under a Creative Commons License (CC attribution 4.0 international license being the preferred one).

5. ALLOCATION OF RESOURCES

Each partner is responsible for managing their research data. The project coordinator will centralise and survey that data management. In addition, all partners are responsible for data generation, metadata production and data quality and they will have specific responsibilities depending on the data and the internal organisation within the WPs and tasks where research data is created or used.

The responsible partner for implementing the DMP in the LEONSEGS project is GMV, as project coordinator. Complimentary to this, each Dataset Owner (DO) is also responsible to follow-up data management for their respective dataset.

In general, each partner is responsible for collecting/generating data/ metadata adequately. Specific responsibility is to be assigned depending on the data and the internal organisation within the WP and tasks where such data is created. The resources for long term preservation (such as costs and potential value) will be discussed inside the consortium, on a case-by-case basis, by the partners involved with the advice of GMV.

Open access for publications (as per HE2020 standards), are an obligation foreseen in article 17 of the GA (Open Science) : "*The beneficiaries must ensure open access to peer-reviewed scientific publications relating to their results*.". Furthermore, the GA includes the obligation of the partners to deposit, at the same time as the publication, the research data needed to validate the results presented in the deposited scientific publications ('underlying data'), considering the storage in a public data repository a best practice. This is governed under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a license with equivalent rights; for monographs and other long-text formats, the license may exclude commercial uses and derivative works (e.g. CC BY-NC, CC BY-ND).

As the management of the data following this DMP is aligned with the commitments of the LEONSEGS project, costs incurred by the partners have to be eligible and comply with the GA rules. Partners are aware that costs related to open access to research data in HE2020 are eligible under the conditions of article 17 of the GA.

6. DATA SECURITY

Data security focuses on ensuring the confidentiality, integrity and availability of data when it is being stored, transmitted or used. Several safeguard security measures will be described in this section.

The partners will store their research data in the project common repository held in Microsoft SharePoint/Teams. The D1.1. LEONSEGS Project Management Plan describes the approach of sharing of data using Teams groups and channels. In line with article 25 of the GDPR, a data protection by design and by default approach must be adopted. An access control via logging key to the common repository will be implemented. Requesting further authentication of the subject avoids that an indeterminate number of natural persons could have access to the information, avoiding unauthorised access. The data will only be handled by directly involved personnel of the partners, and no one will have access to the data unless this is necessary to carry out the project work. Encryption of the information (encoding messages so only those authorised can read them) will provide an extra layer of security to the data stored. The cloud storage and cloud computation resources providers guarantee the adoption of the adequate data security to prevent unauthorised access.

Special provisions will be made for the management of sensitive confidential data throughout the project. Research data will be publicly shared unless there are justifiable reasons for keeping datasets confidential. The main reasons for considering information contained in datasets as confidential or sensitive are existence of intellectual property rights titles/trade secrets (because the data could confer a competitive advantage to the owners) or personal data collected compliant with the data protection regulation (GDPR).

LEONSEGS project will maintain protection of personal data and compliance with Data Regulations as per national and European legislation regarding the protection of personal data.

If possible, non-sensitive information (and their metadata) will be publicly shared in certified repositories such as Zenodo (https://zenodo.org/), and will follow their internal procedures for long term preservation and curation.

In relation to internal security measures, each partner must be responsible for establishing and implementing technical and organisational measures according to the type of data that they are going to process. Article 32 of the GDPR deals with the security of processing, following data protection by default and by design principle, mentions that it is mandatory to implement adequate protection measures, that it could be categorised in a technical (section 5.2) or organisational way (section 5.3).

7. ETHICAL ASPECTS

LEONSEGS has a dedicated WBS task (WP1/T1.2 – Research Ethics, Gender, Legal and Societal Aspects) to ensure that ethical requirements are met for all research undertaken in the project, including data management aspects, in compliance with HE2020 ethical standards. Additionally, LEONSEGS partners are committed to follow the ethical principles as set out in GA Article 14 — Ethics and Values.

Data that will be part of the DMP will be considered as non-personal data, applying Regulation EU 2018/1807, allowing that this information could be processed freely throughout the EU. In this way, GDRP guidelines will be adopted to minimise the risk of breaching of personal data regulations.

Ethical considerations will be integrated throughout all stages of data handling, encompassing data collection, storage, transfer, and access control. The project will work to ensure that management of personal data is compliant with GDPR and other applicable legal frameworks related to personal data protection. During the project, each partner will consider the standards, treaties and laws regarding data protection and privacy in both EU and national level legislation.

8. OPEN DATA MANAGEMENT: ZENODO

The LEONSEGS consortium has plans to establish its first steps towards the open data management application. To this end, an account in the trusted repository Zenodo has been opened, establishing the LEONSEGS community (link here: https://zenodo.org/communities/LEONSEGS). This action will enable the sharing of not only the scientific publications generated throughout the project, but also the open data (designated as "PUBLIC" dissemination level) that has been utilised in the project span.

Until now, data are being collected for internal purposes, therefore no content has been published so far in Zenodo. However, the data will be uploaded and updated regularly throughout the project. Instructions concerning uploading the data, as well as the practical implementation of the ORDP using Zenodo repository will be produced by GMV. Since GMV is having overall responsibility for the DMP, GMV will coordinate the whole data management strategy, but every Dataset Owner (DO) (as mentioned in section 3 of this document) is responsible for ensuring that the open data uploaded in Zenodo follows the procedures established in the previous sections.

9. CONCLUSIONS

This document constitutes the first version of the DMP of the project. The next version of this document will build upon this first version by providing necessary updates.

The data that is integrated within this DMP must be considered non-personal data, being excluded from the application of the GDPR and other related personal data protection laws. A public dissemination of the datasets that has been collected/generated will be assessed case-by-case, limiting their access at only internal level due to intellectual property or privacy reasons under well-founded justification of a partner. No additional ethical or legal issues have been identified as impacting data sharing.

Partners are aware that this document contains only the first version of LEONSEGS' DMP, committing to update this content at least once during the course of the project (as D1.5 DMP v2). Additionally, it can be updated whenever significant changes arise, including, but not limited to:

- new data
- changes in consortium policies (e.g. new innovation potential, decision to file for a patent)
- changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving)

END OF DOCUMENT

Co-funded by the European Union