This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

EXCELSIOR Project

H2020-WIDESPREAD-2018-2020 | Grant Agreement No 857510	
Project full title:	**ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment**
Project acronym:	**EXCELSIOR**
Work Package:	**WP10 Dissemination and Exploitation**
Deliverable:	**D10.8 Data Management Plan**
Version:	**Final D10.8**
Dissemination level:	**Public**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

	**H2020-WIDESPREAD-2018-2020/ H2020-WIDESPREAD-2018-01** **Grant Agreement No 857510** This project is funded by the **EUROPEAN COMMISSION** in the Framework Programme for Research and Innovation (2014-2020).
Call / Topic:	**H2020-WIDESPREAD-2018-01 /** **WIDESPREAD-01-2018-2019** **Teaming Phase 2**
Project full title:	**ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment**
Project acronym:	**EXCELSIOR**
Work Package (WP): Task (T):	**WP10 Dissemination and Exploitation** **T10.4 Data Management Plan**
Deliverable (D):	**D10.8 Data Management Plan**
Due date of deliverable:	**31 March 2020** **(Month 06 of the project)** / **Version: Final D10.8**
Author(s):	CUT: **Kyriacos NEOCLEOUS, Kyriacos THEMISTOCLEOUS, Christiana PAPOUTSA**
Contributor(s):	CUT: **Diofantos HADJIMITSIS** DLR: **Gunter SCHREIER** NOA: **Maria PAPAKONSTANTINOU, Themistoklis HEREKAKIS, Mihalis TSOUTSOS** TROPOS: **Albert ANSMANN, Johannes BÜHL, Patric SEIFERT** DEC-MTCW: **Georgios KOMODROMOS, Stelios TZIORTZIS**
Start date of project:	**1 October 2019** / Duration: **84 months**
Dissemination Level:	**Public**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Document Sign-off

Nature	Name	Role	Partner	Date
DRAFT	Kyriacos NEOCLEOUS Kyriacos THEMISTOCLEOUS Christiana PAPOUTSA	WP10 Participant Technical Manager WP10 Leader	CUT	02/03/2020
REVIEWED	Haris KONTOES	WP10 Participant	NOA	18/3/2020
REVIEWED	Gunter SCHREIER	WP10 Participant	DLR	19/03/2020
REVIEWED	Diofantos HADJIMITSIS	Project Coordinator	CUT	19/03/2020
REVIEWED	Albert ANSMANN	WP10 Participant	TROPOS	22/03/2020
DRAFT	Kyriacos NEOCLEOUS	WP10 Participant / Quality Assurance	CUT	23/03/2020
APPROVED	All partners			30 /03/2020

Work Package 10: Dissemination and Exploitation

D10.8: Data Management Plan

Sections to be protected	Description	Owner	Access Rights	
			Period	**Type***
None			**-**	**PD**

*PD: Public dissemination CA: Confidentiality Agreement required for disclosure

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Executive Summary

The aim of the Data Management Plan is to outline the data management policy of the EXCELSIOR Phase II Teaming Project, and elaborate the appropriate framework for access, use and re-use of the four types of data which would be collected or generated by the EXCELSIOR Phase II Teaming Project.

The Data Management Plan of the EXCELSIOR Project initially provides a summary of the four types of datasets (i.e. Communication, Scientific Publications, Deliverables and Trainings), followed by the procedures adopted to make the data discoverable, accessible, interoperable and re-useable (FAIR). The resources available to the Data Management Plan are also outlined as well as the provisions made for data security. The Data Management Plan also includes a template for describing the data, according to the H2020 recommendations.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Table of Contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

List of Figures

List of Tables

Abbreviations

CA	Consortium Agreement
CUT	Cyprus University of Technology
DAS	Data Acquisition Station
DMP	Data Management Plan
ECoE	Eratosthenes Centre of Excellence
EMMENA	Eastern Mediterranean Middle East and North Africa
EO	Earth Observation
EXCELSIOR	Eratosthenes: Excellence Research Centre for Earth Surveillance and Space-based Monitoring of the Environment
GBS	Ground-based Remote Sensing Station
GDPR	General Data Protection Regulation
NRL	Near Real Time
IPR	Intellectual Property Rights
RS	Remote Sensing

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

1 Introduction

Deliverable 10.8 aims to define the framework for the Data Management Plan (DMP) of the EXCELSIOR Phase 2 Teaming project. The DMP describes the management life cycle for the data that will be collected and generated by the EXCELSIOR Phase II Teaming Project. These data are primarily documentation, input used and output generated in the framework of the EC funded Project to support the establishment of the Eratosthenes Centre of Excellence (ECoE), as described in the following chapter.

In a wider framework and for the operations of the ECoE, the "data management" identifies the different data sources consisting of geospatial information, satellite datasets, modelled data, active remote sensing data, meteorological products, questionnaires, studies, etc, and consolidates licensing and approaches to make research data findable (i.e. discoverable), accessible, interoperable and reusable (FAIR) for fostering knowledge discovery and regional innovation (according to the H2020 Data Management Manual and section 2.2.2). As part of making these data findable, accessible, interoperable and re-usable (FAIR), the DMP includes information on the handling of the data during and after the end of the Project, and the methodology, standards and resources applied for making the data FAIR. Data curation and preservation (including after the end of the project) are also discussed. The DMP includes a template for the management of data, which is based on H2020 recommendations.

Therefore, this task is related to the appropriate infrastructures and procedures to be set up by ECoE to acquire, manage, store and disseminate such science data (through the activities of T5.6 "ECoE "Big Data" Infrastructure" and T7.2 "ECoE Data Cube"). The DMP will be regularly updated, and CUT will lead this task, with contributions made by all Project partners.

1.1 GDPR

To comply with the General Data Protection Regulation (GDPR) (European Parliament and the Council, 2016) as well as of the Consortium Agreement (CA) of the EXCELSIOR Project, the following provisions apply to the Project's Data Management Plan. As a general rule, all acquired data will be treated and handled under the guidelines of the EU's GDPR.

Security and Confidentiality:

- Personal Data will be treated as confidential information and will be treated according to section 10 of the EXCELSIOR Project CA.
- Technical and organizational security measures will be taken by the data controller that are appropriate to the risks, such as unauthorized access, presented by the processing. Any person acting under the authority of the data controller, including a processor, will not process the data except on instructions from the controller.
- The Project will notify within seventy-two (72) hours the Parties and/or Personal Data Subjects concerned with potential Persona Data breach.

Accountability, Transparency and Rights of Access:

- Personal Data Subjects will be provided with information as to the purposes of the processing and the identity of the data controller, and other information insofar as this is necessary to ensure fair processing.
- Data Subjects will have a right of access to all data relating to them that are processed and, as appropriate, the right to the rectification, erasure or blocking of data the processing of which does not comply with the principles set out in the GDPR, in particular because the data are incomplete or inaccurate. They will also be able to object to the processing of the data relating to them on compelling legitimate grounds relating to their particular situation.

Accuracy, Quality and Proportionality:

- Personal Data will be accurate and, where necessary, kept up to date. The data will be adequate, relevant and not excessive in relation to the purposes for which they are collected and further processed.
- Personal Data will be processed and subsequently used or further communicated only for the specific purpose of carrying out its relevant part of the EXCELSIOR Project.

Storage Limitation:

- Personal data will not be kept longer than necessary for the purposes for which they are collected and processed. All remaining personal data will be deleted at the end of the EXCELSIOR project. However, if personal data is required after the end of the project (for further scientific research purposes), Articles 89 of the GDPR will be applied.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

2 Data Summary

2.1 Purpose of data collection and generation

The EXCELSIOR Phase 2 Teaming Project aims to upgrade the ERATOSTHENES Research Centre, established within the Cyprus University of Technology (CUT), into a sustainable and viable ERATOSTHENES Centre of Excellence (ECoE) for Earth Surveillance and Space-Based Monitoring of the Environment. Furthermore, the EXCELSIOR Project will establish, within the ECoE, an Earth Observation (EO) Satellite Data Acquisition Station (DAS) and an Atmospheric Ground-based Remote Sensing Station (GBS) Supersite for Aerosol and Cloud Monitoring (which will also be used for calibration and validation of satellite data). The Satellite DAS will be able to directly receive data from EO satellite missions, which will allow Near Real Time (NRL) monitoring and thereby provide time-critical information for science and products within the receiving cone of the station, namely, over the EMMENA region (Eastern Mediterranean Middle East and North Africa).

The DMP of the EXCELSIOR Project covers the datasets collected or generated in response to the activities of the ten Work-Packages of the Project. These datasets comprise of four main categories: Communication datasets, Scientific Publications datasets, Deliverables datasets and Trainings datasets.

The datasets produced by projects and other activities of the ECoE will be dealt with in a separate DMP, which will be developed once the ECoE is legally established and the framework of the Digital Innovation Hub is formalised by the Project Consortium.

2.2 Types and formats of data

The following table provides a synoptic view of the data expected to be collected or generated during the lifetime of the EXCELSIOR Phase 2 Teaming Project.

Table 1: Types and format of data to be collected and processed by the EXCELSIOR Project

Description	Type	Form	Format
Communication Dataset	Material generated for collected by the consortium for communication events, including stakeholder workshops	Documents, Brochures Annual Editions Presentations Questionnaires Photos Video	DOCX, ODT, RTF, PDF PPTX, PPSX, ODP, PDF DOCX, XLSX, ACCDB, CSV, XML, ODS JPG, TIFF MP4, MOV, WMV

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Description	Type	Form	Format
Scientific Publications Dataset	Material generated or collected by the EXCELSIOR consortium to prepare scientific publications, including articles published in conferences, workshops and peer-reviewed journals.	Documents	DOCX, ODT, RTF, PDF DOCX, XLSX, ACCDB, CSV, XML, ODS
		Sensor and Satellite Data (Multispectral, optical and radar data sets, RGB, Infrared, Hyperspectral and LIDAR images) and Metadata, Binary Data of Atmospheric State, Statistics, Historical Data, Third Party Scientific Data, Modelled Data.	Sentinel Standard Archive Format for Europe (SAFE), JP2, XML (for metadata), TIFF, NC, GEOTIFF, CCR, HDF, DOQQ, CCM, ASCII, Fast-L7A, EOSAT Fast Format, EEF, NetCDF, TIFF, TFW, RPC, HDR, SHP, SHX, DBF, GEOTIFF, XML, N1, E1, E2, NTIF, NetCDF,
		Aerial and UAV RBG, infrared, Hyperspectral and LIDAR images and data	PLY, FBX, DXF, OBJ, PDF, SHP, GEOTIFF, LAS, LAZ, XYZ, JPG, PNG, KML, HTML, IMG, SIT, IS2, IR, FLS, PCD, ANA, FTS
Deliverables Dataset	Material generated or collected by the EXCELSIOR consortium for the preparation of Project Deliverables.	Documents Presentations Questionnaires	DOCX, ODT, RTF, PDF PPTX, PPSX, ODP, PDF DOCX, XLSX, ACCDB, CSV, XML, ODS
		Photos	JPG, TIFF
		Video	MP4, MOV, WMV

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Description	Type	Form	Format
Trainings Dataset	Material generated or collected by the EXCELSIOR consortium in the framework of the project training activities. The dataset includes presentations, tutorial videos, images from various applications, including training data from: - Medium and High resolution Earth observation datasets (e.g. Landsat series/ Sentinel data, TerrasSAR, Copernicus Data Warehouse), - data from Aerial and UAV, - data from in-situ measurements (e.g. meteorological stations, atmospheric LIDAR, sun-photometers, VOL sensors), - data from Ground remote sensing techniques (e.g. geophysical prospection, ground spectroscopy), - data from GIS databases - Spatial data (e.g. vector and raster data).	Presentations	PPTX, PPSX, ODP, PDF
		Numeric data	DOCX, XLSX, ACCDB, CSV, XML, ODS
		Images	JPG, TIFF
		Video	MP4, MOV, WMV
		Sensor and Satellite Data (Multispectral, optical and radar data sets, RGB, Infrared, Hyperspectral and LIDAR images) and Metadata, Binary Data of Atmospheric State, Statistics, Historical Data, Third Party Scientific Data, Modelled Data.	Sentinel Standard Archive Format for Europe (SAFE), JP2, XML (for metadata), TIFF, NC, GEOTIFF, CCR, HDF, DOQQ, CCM, ASCII, Fast-L7A, EOSAT Fast Format, EEF, NetCDF, TIFF, TFW, RPC, HDR, SHP, SHX, DBF, GEOTIFF, XML, N1, E1, E2, NTIF, NetCDF,
		Vector data	shp, gdb,dwg, dxf
		Aerial and UAV RBG, infrared, Hyperspectral and LIDAR images and data	PLY, FBX, DXF, OBJ, PDF, SHP, GEOTIFF, LAS, LAZ, XYZ, JPG, PNG, KML, HTML, IMG, SIT, IS2, IR, FLS, PCD, ANA, FTS
		Data for Aerial optical thickness, atmospheric profile, wind speed and direction, air temperature, humidity, surface temperature, incoming solar radiation, CO, CO2, NO.	NetCDF

2.3 Re-use of existing data

The Project will use existing data which is openly available through international data portals, such as Copernicus Data Hubs for Sentinel satellite images, Copernicus Services portals, NASA for Landsat images, ESA for third party missions, ACTRIS and PollyNET networks, Cloudnet, Vector data for GIS applications from EU portals, GEOSS & GEO-CRADLE portal, Inspire portal, as well as data that will be provided by Project Partners (i.e. DEC-MTCW, DLR, NOA, TROPOS), stakeholders and governmental organisations that have provided commitment letters to the EXCELSIOR Project (e.g. Ben-Gurion University of the Negev, the Land Surveys Department and the Water Development Department). The

data will be used primarily for training and capacity building activities of the project (in the framework of WP6). The data will not be further managed or distributed by the project, since the current DMP focuses on the data that will be generated or collected by EXCELSIOR consortium and not the data collected, generated or transferred to and from the ECoE.

2.4 Data origin

The data will originate mainly from activities and tasks of the EXCELSIOR Project, as described in the Grant Agreement of the Project. Remote Sensing and satellite EO data, provided openly by European and International organisations, as well as Earth Observation data granted by commercial companies will also be utilised for training purposes, as well as for the preparation of the scientific publications of the EXCELSIOR project.

2.5 Data size

The expected size of the EXCELSIOR documentary datasets will approximately be 80TB; however, this is an estimated size of the datasets, considering that raw and processed satellite images will not be archived by the Project. In this case their origin and source will be referenced.

2.6 Data utility

The data of the EXCELSIOR Project will be useful to the Project partners, personnel of the newly established ECoE, as well as researchers and scientists (operating in the fields of RS and satellite-based EO and Big Earth Data analytics), stakeholders from industry, local and national authorities in the EMMENA region. Most of the used satellite data will be available and transferrable on a free and open basis. For some commercial and copyrighted Earth observation data, rules and restrictions of the originator may apply.

3 FAIR data

3.1 Making data findable, including provisions for metadata

The datasets collected and generated by the EXCELSIOR project will be discoverable with descriptive metadata, and the template of the KTISIS (certified repository of the Cyprus University of Technology) has been adopted for the four datasets of the Project. This template (Table 2) is based on a Qualified Dublin Core based metadata schema (DuraSpace, 2020). Furthermore, to comply with the Inspire Directive, the ISO 19115-1 metadata standard (ISO 19115-1, 2014) will be used to develop a metadata template for the EO and RS data collected and generated by the Project. Each data will be provided with at least five keywords and a version number will be provided for each dataset.

Table 2: KTISIS metadata template utilised for the EXCELSIOR Datasets

Title	Publisher	ISSN
Authors	Source	URL
Keywords	Journal	DOI
Category	Conference	Rights
Field	Abstract	Type (e.g. article, poster, leaflet, deliverable report)
Issue Date	Appears in collection	

Each dataset will be made identifiable and locatable through a persistent URL, which will be based on the Handle System (CNRI, 2020), administered by the Corporation for National Research Initiatives. However, the DOI of Scientific publications will also be included in the metadata of this type of dataset.

All files irrespective of the data type will be named in accordance with the following file naming convention, aiming to have short (less than 35 characters excluding separators) but descriptive file names.

DATE_PROJECT_WORKPACKAGE_TITLE_VERSION_DISSEMINATIONCLASS_ARCHIVE_FILETYPE

Where:

DATE is the date of file creation. The ISO 8601 format is adopted (YYYYMMDD).

PROJECT is the acronym of the relevant project (e.g. EXCELSIOR for EXCELSIOR Phase 2 Project documents)

WORKPACKAGE is the relevant project work package number, with WP as a prefix.

TITLE represents the description of the data item contents excluding capitalisation and punctuation characters (mandatory).

VERSION is the version number consisting of integer numbers only without leading zeros, prefixed with V (mandatory).

DISSEMINATIONCLASS is the dissemination classification allocated to a document type that define the data access (especially after archiving). Classification options include:

PU (for Public),

RE (for restricted to a group specified by the consortium) and

CO (for Confidential, only for members of the consortium; Commission services always included).

ARCHIVE this is a single character defining the allocation of the data item for future archiving and is represented by a Y or N.

FILETYPE indicates the format of the file

3.2 Making data openly accessible

The four datasets produced during the life-time of the EXCELSIOR Phase 2 Teaming Project will include public, restricted and confidential data, saved in open source file formats. The public data will also include processed open data, originating from European and international infrastructures through related links (such as Sentinel mirror sites) and will be openly accessible without restrictions. While, restricted and confidential data (including confidential deliverables and periodic management reports submitted to the European Commission as well as IPR related documents) will not be shared outside of the consortium, in order to comply with the provisions of the CA and the Grant Agreement of the Project. Furthermore, personal data (e.g. those collected for the purpose of undertaking communication events) will not be shared, but will be strictly kept confidential, to comply with national and European legislation (i.e. GDPR) as described in the CA of the Project as well as section 1.1 of this report.

The provisions of Articles 29.1 to 29.3 of the EXCELSIOR Project GA will be applied to make openly accessible the Scientific Publications Dataset. The aim will be to ensure open-access within six-months of publication; and this will include both the scientific articles as well as the research data generated for each scientific article. To ensure appropriate protection of the IPR (Intellectual Property Rights) generated by the Project, the DMP will follow the provisions of the CA of the project, as described in section 3.4 below.

Depending on the dissemination status of the data collected and generated by the Project, the four datasets and/or the associated metadata will be accessible (and archived) through KTISIS, the certified repository of CUT (Ktisis, 2020), where a dedicated section has been created[1] for the EXCELSIOR Phase 2 Project (https://ktisis.cut.ac.cy/cris/project/pj00347). KTISIS is compatible with OpenAIRE and will provide direct open-access to Project data with public dissemination status (e.g. scientific publications, presentations and public deliverables). Furthermore, KTISIS will provide indirect restricted access to Project data with restricted/confidential dissemination status; this restricted access will be provided only to members of the EXCELSIOR consortium by using persistent digital identifiers (i.e. ORCID iD as shown in Figure 1). If necessary, an embargo period will be applied to sensitive information, such as personal data, which should only be accessible to the designated controller, as described in the CA of the Project.

[1] As stated in the GA of the EXCELSIOR Project, the CUT Library will provide technical support to the EXCELSIOR Project as well as access to related infrastructures and facilities of the Library (e.g. KTISIS repository). It is noted that the Library Director is a member of the CUT team participating in the EXCELSIOR project. Furthermore, CUT has been committed to provide access to the state-of-the-art Data Centre and E-learning facility of the University.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

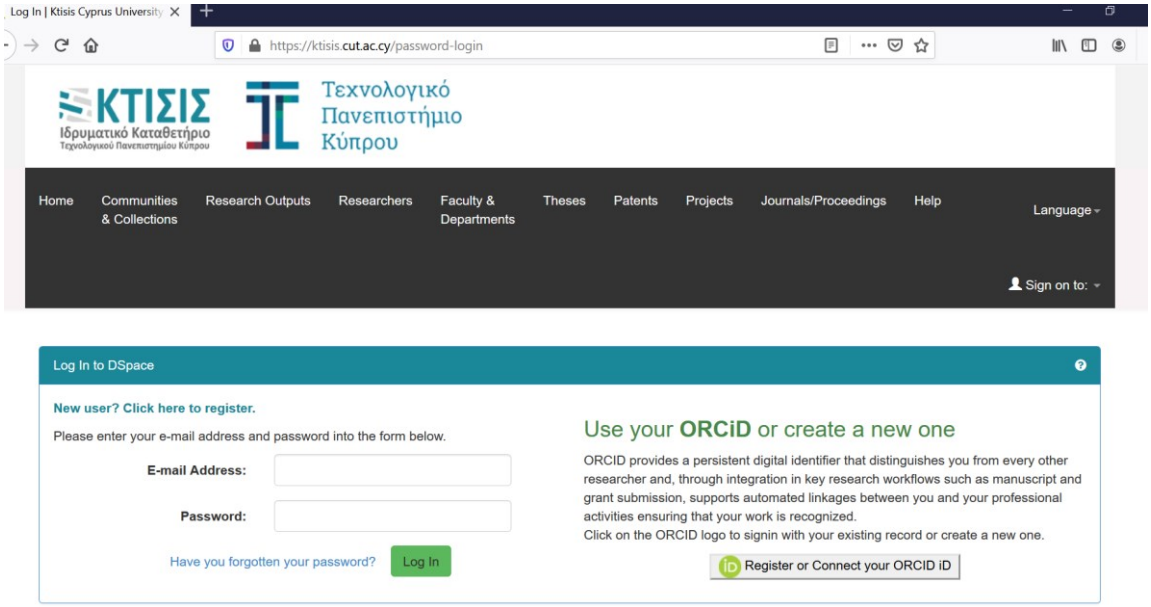

Figure 1: Access to restricted data of the EXCELSIOR Project in KTISIS by using ORCID ID

Ready-to-use data cube software (LARDA[3]) will be provided by TROPOS for display and analysis of data collected with the Ground-based remote sensing station (GBS). This software provides online access for remote-sensing data and an API for detailed data analysis. A similar data-cube software will also be implemented for the collected satellite datasets.

The following software will be required to access directly the data of the EXCELSIOR Project deposited on KTISIS:

- Web-browser to access the KTISIS repository (https://ktisis.cut.ac.cy).
- Microsoft Office software or any other OpenOffice software to read documents, spreadsheets and presentations.
- PDF reader to read PDF files.
- Geographical Image viewer and/or web-based applications.

The EXCELSIOR project will also provide instructions for using specialised software, required to read the EO and RS data.

3.3 Making data interoperable

The data produced by the EXCELSIOR Project will be interoperable; Table 3 shows the standard file formats that will be used for the EXCELSIOR datasets, deposited on the KTISIS repository.

Table 3: File formats to be used for EXCELSIOR data available on the KTISIS repository

Data form	Standard format
Documents	PDF, ODT, RTF
Presentations	PDF and ODP
Spreadsheets	ODS
Numeric data	CSV, XML
Vector data	DXF
Photos and Images	Tiff
EO / geographical images	GeoTiff
Movie	MP4

Furthermore, to make the data interoperable, the KTISIS repository applies the Schema.org vocabulary and Qualified Dublin Core based metadata schema, as well as the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

3.4 Increase data re-use (through clarifying licenses)

The EXCELSIOR Project will promote data re-use, which will be achieved through open-access and licensing provisions.

Open-access will be provided for the Communication Dataset (excluding personal data covered by GDPR) and the Scientific Publications Dataset, as well as for Deliverables and Training data with Public dissemination status, by following the procedures described in section 3.2 of the DMP.

While the provisions of Section 8 "Results" and Section 9 "Access Rights" of the CA of the EXCELSIOR Project will be applied for licensing, to achieve the widest reuse possible of the EXCELSIOR Project data. These provisions focus on a range of aspects relating to licensing, including the following.

- Ownership and Joint Ownership of Results.
- Transfer of Results.
- Dissemination.
- Background included to perform the Project activities.
- General Principles for Access Rights.
- Access Rights for implementation of the Project.
- Access Rights for exploitation.
- Access Rights for software.

Free and open satellite data, such as Copernicus and Landsat, can be reused and re-distributed immediately. However, negotiations with commercial data suppliers of "restricted" datasets tends to take additional time. Commercial data suppliers are open to negotiate a free access to their data provided that: (a) it is used for academic purposes, (b) it has a limited geographic coverage and (c) the data is over a year old, at minimum. Commercial data suppliers will provide data, provided that it does not interfere with their business interests and that they also benefit when granting and distributing data for free.

Therefore, utilizing archival data from both free and open satellite data as well as from commercial sources would provide the ability to perform longitudinal studies for applications. With its limited geographic extent, a long time series data deposit would also not occupy a large amount of storage space. The provisions of the Quality Assurance manual of the EXCELSIOR project (Deliverable D1.11, section 4.3) will be applied for the purpose of data quality assurance.

3.5 Data Storage

In accordance with the GDPR policy at the Cyprus University of Technology (CUT), the following table outlines the type of data and the duration that it is retained by CUT.

Table 4: CUT GDPR policy for retaining data

Data type	Retaining Period
Job applications for administrative, academic and other academic staff	20 years
All necessary documents for employment of administrative and academic staff	62 years
Outsourcing or Employment: A Research Associate involved in University Research Programs either by contract of employment or by outsourcing	30 years
New employee: By hiring a new employee, Human Resources opens up a portion of the employee within the accounting software. **Records for employees who will soon retire.** **Records for other categories of staff who will not retire.**	 100 years 62 years
Contact list for events - Contacts for event updates	Deleted upon request
Web site photos and videos - Photographs and videos are not deleted for historical and research purposes	Unlimited
Social network data - No data is deleted from Social Media unless deemed inappropriate by the Media	Unlimited
Cameras Used for Video Assessment and / or Intervention Sessions – retention of visualized files from cameras	7 years in external disk 1-2 years on server depending on capacity
Voice recording of evaluation and / or intervention sessions - audio files on tape recorders	6-9 months
Building Rental Contracts -	20 years from termination of contracts
Closed surveillance system at the Rector's building	1 month

It should be noted that other organizations have different retainment times for data. For example, research data and related material should be retained for a minimum of 5 years or any longer period required by an approving body, such as the research funder or under legislation. The periods commence on the date at which the final report was sent to the research funder, or the date on which the output was published. The term 'data' here includes all unpublished evidence, whether numerical or otherwise, on which the publication is based and from which results can be replicated or reproduced. Hard copy, such as laboratory notes, field notes, questionnaire responses, signed consent

forms, the research protocol for the project, photographic records, and subsequent electronic files should all be retained.

In terms of compliance with the Data Protection Act, the research exemption allows researchers to keep data indefinitely as long as it is still useful and there is an intention to use them further for research. If research data / records are no longer useful, then disposal should be considered. Data which requires retention periods beyond 20 years are usually related to studies where there is a greater likelihood that scientists will be asked about reproducibility (i.e. because their research findings are high profile/high impact or in some way contentious).

It may be necessary to consider the impact of the data at several stages during its lifecycle, particularly for studies which have long-term goals, which may be running for many years. Once decisions have been made with regards to what will be kept and in what format, it is important to review what is held on an ongoing basis. In cases where data is archived for long term preservation or longitudinal studies the data will be made available in the most open manner appropriate.

Research data relating to studies which directly inform national policymaking should be considered for permanent preservation in an appropriate archive. In some cases, the potential impact on policy may be a clear aim of the study, while in others the significance may only come to light later.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

4 Allocation of resources

4.1 Cost breakdown

Resources are required for the costs involved for making FAIR the data of EXCELSIOR as well as for producing, analysing and preserving data in the long-term.

The Communication and Deliverables Datasets will be hosted, free of charge, on the data repository of CUT (Ktisis, 2020); while the Scientific Publications dataset will be provided in open-access, either through the Gold or the Green route. Within the 7-year period of the EXCELSIOR project, the consortium is expected to produce up to 200 peer-reviewed scientific publications, and it is estimated that a budget of up to 200,000 EUR will be required to make these publications openly accessible.

4.2 Budget summary

The following table outlines the EXCELSIOR budget allocated for making the Project data FAIR.

Table 5: Outline of budget for making FAIR the EXCELSIOR Project data

Amount	Funding Source	Duration	Purpose
€117000	EXCELSIOR Project funding	Year 1-7	Open-access (Gold route) and conference registrations for Scientific Publications
€100000	Cash contribution of the government of the Republic of Cyprus	Year 1-7	Open-access (Gold route) and conference registrations for Scientific Publications
€130000	Cash contribution of the Cyprus University of Technology	Year 1-7	Open-access (Gold route) and conference registrations for Scientific Publications
€100000	In-kind contribution of the Cyprus University of Technology	Year 1-7	Use of the University data centre, E-learning facility and KTISIS repository

4.3 Responsibility for Data Management

The Leader of WP10 (Dissemination and Exploitation) will be responsible for the data management of EXCELSIOR and will be supported by the Quality Assurance Manager of the Project.

4.4 Long-term preservation of data

The EXCELSIOR Project datasets will be preserved in the KTISIS repository and the associated costs will be covered by the in-kind financial contribution of the Cyprus University of Technology.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

5 Data security,

The EXCELSIOR datasets will be securely stored in KTISIS, CUT's certified repository, to ensure long term preservation and curation after the end of the project. The KTISIS servers are hosted by the state-of-the-art Data Centre of CUT and, hence, there are no issues with data storage solutions recoverability and security. The team of the EXCELSIOR Project will collaborate with the Library and the IT services of CUT to tackle security issues by considering the newest technical solutions and standards. It is noted that when the new entity ECoE accedes to the EXCELSIOR Project, the data security policy of the current DMP will be revised accordingly to cover the activities of the ECoE.

Furthermore, all Project deliverables will be available on the H2020 Participant Portal of the European Commission and, this will ensure the long-term preservation and recoverability of the deliverable reports.

References

CNRI, 2020. *Handle.Net Registry.* [Online] Available at: https://handle.net/

DuraSpace, 2020. *DSpace - The Softwarwe of choice for academic, non-profit, and commercial organizations building open digital repositories..* [Online] Available at: https://duraspace.org/dspace/resources/ [Accessed 06/03/2020 March 2020].

European Parliament and the Council, 2016. *REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.* Brussels: European Union.

ISO 19115-1, 2014. *Geographic information — Metadata — Part 1: Fundamentals,* Geneva: International Organization for Standardization.

Ktisis, 2020. *Open access institutional repository of the Cyprus University of Technology.* [Online] Available at: https://ktisis.cut.ac.cy/ [Accessed 3 2 2020].

.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

Appendix A –Template for Data Management Plan

Data Set	**DATA MANAGEMENT PLAN**			
No.	**Partner Legal Name**	**Project/GA No/WP**	**Researcher's name**	**email**
Please choose from available lists or add options if applicable.				**NOTES**
1. DATA SUMMARY				
	Purpose of the Data	**Purpose**	**Project Objectives**	
			1	*State the purpose of the data collection/generation, indicating the relation with the objectives of the project. Add additional objectives if necessary.*
			2	
			3	
	Type and Format of Data	**Form**	**Format**	**Describe the type of data used or generated within the project, specifying the form and format of the data.**
	Text	*type_of_text*	*type of format*	*Form: Field or laboratory notes, survey responses .. Format: in plain text, (txt), HTML, XLM, PDF/A …*
		type_of_text	*type of format*	
		type_of_text	*type of format*	
	Numeric	*tables*		*Tables, row counts, measurements - in .XLSX, .CSV …*
		tables		
	Audiovisual	*image*	*jpg*	

		movie	*MP4*	*Images, sound recordings, video - in .JPEG, .JPG, .PNG, .TIFF, AIFF, WAVE, .MP3, .MP4…*
	Simulated			*Please state the model, model type and computer code - and specify output data type and format.*
	model			
	model type			
	computer code			
	data type			
	format			
	Discipline specific information	*discipline*	*format*	*Specify discipline and format)*
		discipline		
	Instrument specific	*equipment*	*format*	*Equipment output (specify equipment and format).*
	Reused-Data (rd)	*yes_rd*	*write explanation*	*Indicate if you re-use existing data (generated outside the EXCELSIOR project). If so, explain how.*
	Data Origin	**Define and describe the origin/source of your data. Data can be gathered from different sources.**		
	Observational			*Data captured in real time - often not reproducible i.e. sensor readings, images, telemetries, sample data…*

	Experimental			*Data from lab equipment, often reproducible*
	Simulation			Data generated by computational models where model and metadata are equally important to output data - i.e. climate models, economic models, materials models,…
	Derived/Compiled			Data coming from analysis or compilation. Reproducible but with high costs - i.e. the results of text and data mining, compiled databases..
	Reference or Canonical (links)		*write the reference or canonical*	*Collection or conglomeration of smaller (peer-reviewed) datasets published and curated - i.e. chemical structures, gene sequence databanks, spatial data portals..*
	Dataset is:			***Fixed** : never change after being collected or generated. **Growing** : new data may be added, but the old data is never changed or deleted. **Revisable** : new data may be added, and old data may be changed or deleted.*
	Quantity		in MB/GB	*of each experiment*
			in MB/GB	*overall*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

				In case not just digital archiving is required, indicated quantities of other form of storage.
	Data Security & Storage	*select or add: type of storage*	*write data security policy*	*(i.e. Office computer, Hard Drive, Tape back-up system, Institute network drive, Institute Central Data storage, private Cloud storage …), briefly describing the data security policy applied.*
	Data Value (long term)			*Describe to whom the data could be useful.*
				Estimate potential value of long-term re-use of the data.
2 FAIR DATA				
2.1 FAIR DATA - Making data findable				
	Discoverability of data (metadata provision)	*write explanation*	*write information*	*Explain how data are documented and if metadata are provided, listing the information made available/discoverable.*
	Identifiability of data (refer to standard id mechanisms)	*choose mechanism*	*write how are made identifiable*	*Outline the identifiability of data and refer to standard identification mechanism.* *Do you make use of persistent and unique identifiers such as Digital Object Identifiers*
	Naming conventions used	*describe*	*refer*	*Describe the system used to name and structure electronic files and folders. Refer also to any file renaming procedure or tools used.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

	Search keywords approach	*indicate*		*Indicate the approach to keywords generation, indexing and tagging. (For materials modelling the MODA provide this answer.)*
	Clear versioning approach	**Versioning**	**Traceability**	*Describe the versioning and traceability approach used (especially if the dataset is growing or revisable).*
	Standards or procedures for metadata creation applied			*Specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how*
2.2 Fair data Making data openly accessible				
	Data openly available	*indicate ownership*		*Indicate ownership of the data, if it is openly available or can be made openly available.*
	Data kept closed	*users*	*reasons*	*Indicate if data access is restricted, to what users, and explain the reasons.*
	How data will be made available	*indicate*		*Indicate how you intend to make data available.*
	Methods or software (SW) tools for data access	*write methods and tools*		*Indicate methods and SW tools needed to access the data. Clarify if the relevant software (e.g. in open source code) is included in the data set.*

	SW documentation and other information needed	*indicate*		*Indicate any specific SW documentation that is needed to access the data, or additional information that is needed to understand the data (i.e. abbreviations, supplementary notes).*
	Repository for deposit of data, metadata, documentation and code	*indicate open or private*		*Indicate the (open or private) repositories in which the data, metadata, documentation and code are stored and/or those in which they will be stored in the future.*
	Access restrictions	*indicate*	*explain*	*Indicate if there are limitations and restrictions to access the data, and if they are linked to a specific timeframe. Explain how access will be provided after these restrictions are lifted.*
2.3 Fair data Making data interoperable				
	Data interoperability assessment	*indicate* *select*	*range of utilization*	*Assess the level of interoperability of the dataset.* *Indicate data and metadata vocabularies, standards and methodologies followed to facilitate interoperability.* *Indicate if open standards are used, and (if you know) the range of utilization of proprietary SW and methodologies used to generate and manage the data.*

	Standard vocabulary or mapping to commonly used ontologies	*refer*		*Refer to commonly used ontologies to map the dataset, considering also the use of existing common platforms and tools*
2.4 FAIR DATA – Increase data re-use (through clarifying licenses)				
	Data licensing for wide reuse	*define*	*indicate*	*If applicable, define data licensing approach for the dataset wide reuse. Indicate the chosen licenses tools.*
	Timing of data availability for re-use (incl. indications on embargo)	*define*	*indicate*	*If applicable, define the timeframe for making data available for re-use. Indicate any embargo period if required.*
	Data usability by Third Parties (after the end of the project)	*indicate*		*Indicate any limitation to the use of the data by Third Parties, after the end of the project.*
	Restrictions to data re-use	*indicate*	*explain*	*Indicate and explain any restriction to the re-use of data (i.e. confidentiality agreements, other issues).*
	Quality assurance process	*how assured*	*how controlled & documented*	*Explain how quality of the data is assured, how the consistency and quality of data collection is controlled and documented.*
	Length of time of data re-usability	*indicate*		*Indicate the time limit for the data re-usability, if any.*
3 ALLOCATION OF RESOURCES				

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510.

This project has received funding from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development.

	Costs estimates for making data FAIR	*estimate*	*describe*	*Estimate the costs for making your data FAIR (findable, accessible, interoperable and reusable) and describe how you intend to cover these costs (i.e. institute dedicated resources, dedicated part of the project budget …).*
	Data Management Responsibilities	*identify*		*Identify responsibilities for data management of this dataset (within your research group and institute, and within the project if applicable).*