



Ref. Ares(2019)7371622 - 29/11/2019
Funded by the Horizon 2020
Framework Programme of the
European Union



**Exploiting the multifunctional potential of belowground
biodiversity in horticultural farming**

Deliverable 7.4


DATA MANAGEMENT PLAN

Grant Agreement: 817946

Horizon 2020 Programme

Authors: Stefano Mocali (CREA)

Version: 29/11/2019

	D7.4 – Data Management Plan	Page	2/16
		Rev	3
		Date	29/11/2019

Title: Exploiting the multifunctional potential of belowground biodiversity in horticultural farming


Project acronym: EXCALIBUR

Work Package	WP7
Deliverable number	7.4
Due date	30-11-2019
Submission date	29-11-2019
Dissemination level	ORDP: Open Research Data Pilot; public
Deliverable lead	CREA
Authors	Stefano Mocali, Antonio Gerardo Pepe, Loredana Canfora
Abstract	<p>The aim of this Data Management Plan (DMP) is to describe the types of data collected or generated in the course of Excalibur project and how they will be stored, published, cited and made FAIR also beyond the project life. In fact, our goal is to meet the requirements of excellent scientific practice and to allow for accessibility, interoperability, reproducibility of Excalibur research results.</p> <p>Since the DMP is expected to mature during the project, more detailed versions of the DMP will be submitted as additional deliverables at later stages of the project.</p>
Keywords	Data, DMP, FAIR, dataset, open access, management



List of participants

NAME	ACRONYM	COUNTRY
CONSIGLIO PER LA RICERCA IN AGRICOLTURA E L'ANALISI DELL'ECONOMIA AGRARIA	CREA	Italy
INSTYTUT OGRODNICTWA	INHORT	Poland
C.R.P.V. CENTRO RICERCHE PRODUZIONI VEGETALI SOCIETA COOPERATIVA A RESPONSABILITA LIMITATA	C.R.P.V	Italy
NATURAL HISTORY MUSEUM	NHM	United Kingdom
NIAB EMR	NIAB EMR	United Kingdom
KMETIJSKI INSTITUT SLOVENIJE - AGRICULTURAL INSTITUTE OF SLOVENIA	KIS	Slovenia
UNIVERSITA DEGLI STUDI DI TORINO	UNITO	Italy
KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN - KNAW	KNAW-NIOO	Netherlands
KOBENHAVNS UNIVERSITET	UCPH	Denmark
TECHNISCHE UNIVERSITAET GRAZ	TU GRAZ	Austria
INOCULUMPLUS SAS	INplus	France
UNIVERSIDAD DE GRANADA	UGR	Spain
INTERMAG SPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA	Intermag	Poland
NSF EURO CONSULTANTS	NSF EC	Belgium
KOMPETENZENTRUM OBSTBAU BODENSEE	KOB	Germany
FORDERGEMEINSCHAFT OKOLOGISCHER OBSTBAU EV (FOKO)	FÖKO	Germany

	D7.4 – Data Management Plan	Page	4/16
		Rev	3
		Date	29/11/2019

Version history

Version	Date	Author(s)	Description
1.0	06-08-2019	Antonio Pepe (CREA)	Initial draft version, creation of the “table data overview”
2.0	05-09-2019	Loredana Canfora (CREA), Stefano Mocali (CREA)	Initial reviewed version
2.1	17-09-2019	Antonio Pepe (CREA)	Creation of new sections and FAIR guidelines
2.2	18-10-2019	Antonio Pepe (CREA), Loredana Canfora (CREA)	Reviewed version and completed “table data overview”
2.3	22-11-2019	Antonio Pepe (CREA)	Final draft version
3.0	27-11-2019	Stefano Mocali (CREA)	Final reviewed version




	D7.4 – Data Management Plan	Page	5/16
		Rev	3
		Date	29/11/2019

TABLE OF CONTENTS

LIST OF PARTECIPANTS.....	3
VERSION HISTORY	4
TABLE OF CONTENTS	5
EXECUTIVE SUMMARY	6
1. DATA SUMMARY.....	6
1.1 PURPOSE OF THE DATA COLLECTION AND TYPES/FORMATS OF DATA	6
1.2 DATA STORAGE	9
2 FAIR DATA	13
2.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA	13
2.2 MAKING DATA OPENLY ACCESSIBLE	13
2.3 MAKING DATA INTEROPERABLE	15
2.4 INCREASE DATA RE-USE (THROUGH CLARIFYING LICENSES)	15
2.5 ALLOCATION OF RESOURCES.....	15
2.6 DATA SECURITY	16
2.7 ETHICAL ASPECTS.....	16



	D7.4 – Data Management Plan	Page	6/16
		Rev	3
		Date	29/11/2019

EXECUTIVE SUMMARY

This document describes the first version of the EXCALIBUR Data Management Plan (DMP) describing how data collected or generated in the course of the project will be handled during the EXCALIBUR action and how will be stored, published, cited and made FAIR beyond the project life. Our goal is to meet the requirements of excellent scientific practice and to allow for accessibility, interoperability, reproducibility of Excalibur research results.

This document has been prepared by following the template provided by the European Commission in Participant Portal¹.

This deliverable is a “living” document. In fact, at this stage of the project a lot of questions concerning the data are still open for discussion within the Consortium. Thus, according to the EU’s guidelines, since the DMP is expected to mature during the project, more detailed versions of the DMP will be submitted - if appropriate - as additional deliverables at later stages of the project.

It covers:

- the handling of research data during and after the project
- what data will be collected, processed or generated
- what methodology and standards will be applied
- whether data will be shared/made open and how
- how data will be curated and preserved

1. DATA SUMMARY

1.1 Purpose of the data collection and types/formats of data

The present document is the Deliverable 7.4 “Data Management Plant” of the EXCALIBUR project which main objective is to provide the plan for managing the data generated and collected during the project.

EXCALIBUR will collect existing data from partners and will produce new data within the project. In order to focus on real problems and opportunities that farmers and stakeholders are facing in horticulture, EXCALIBUR applied the multi-actor approach. Thus, Excalibur will generate data of different nature, including high-throughput sequencing data of bacteria, fungi, oomycetes, nematodes and microbial eukaryotes, transcriptomic (WP1, WP4, WP5), data of mesofauna, arthropods and earthworms biodiversity (WP1, WP4), metabolic profile of microbial bioinoculants (WP2), soil mineralogical and chemical-physical parameters (WP1, WP4, WP5), bioassay data of various formulations of bioinoculants and bioeffector’s performance (WP2), field measurements of bioindicators of the soil and plant health status from experimental and on-farm trials (WP3, WP4, WP5), biomarkers and molecules driving

¹ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm



	D7.4 – Data Management Plan	Page	7/16
		Rev	3
		Date	29/11/2019

plant-soil-microorganism interactions (WP5), detailed description of the current and innovative strategy designed for each case study including agronomical (crop yield and quality), economical (input and product prices, average farm size, etc.), and environmental (soil quality, functions and biodiversity, tracking inocula persistence in soil) data (WP3, WP4), personal data collected from stakeholder expectations (WP1, WP2 and WP6), georeferenced data as input of the DSSs (WP3, WP5), pedo-climatic and weather data (all WPs).

Despite their heterogeneity, the overall data collected and produced may refer to three primary categories: environment, soil and plant;

- Environment - from meteorological unit, regional meteorological systems, soil sensors
 - Climatic
 - Geographic
- Soil – from laboratory analysis
 - physical properties
 - chemical properties
 - biological properties (microflora, microfauna, mesofauna and macrofauna)
- Plant – from laboratory analysis
 - Agronomic parameters
 - Physiological parameters
 - Phenological parameters

A schematic table questionnaire was submitted to the participants to ask what kind of data will be produced and who will produce it within Excalibur. The complete list of data is reported as Table 1 (“Table Data Overview”) and each item will be briefly explained in the following sections for each WP:


WP1 - Selection of field trials and definition of native biodiversity

WP1 will define the field trials to be carried out with the three selected crops (apple, strawberry, tomato) under contrasting conditions across EU, which will be the basis for planning the biodiversity-driven field management strategy. The activities of this WP will provide also a number of preliminary data of the native soil biodiversity of each field trial. Therefore, both initial soil and environmental data will be collected and inserted in the database. The biological data will be mainly represented by biodiversity data, which values will be considered as “*baseline*” for the subsequent analyses. All the data (metadata) regarding soil properties of experimental fields will be included in a Database on soil mapping and physico-chemical characteristics (D1.4). Furthermore, in some cases (i.e. apple) plants might be present on the field at the beginning of the trial. Thus, in those cases plant parameters will be included as well.

WP2 - Development of novel bio-products and practices under controlled conditions

This WP will be mainly dedicated to the selection of the microbial inoculants (single strains or consortia) and bio-effectors provided by the Partner’s collections to be developed and setup as multifunctional products able to improve plant nutrition, growth promotion and health. Moreover, further improvements in formulation technology will be carried out to form two product models – liquid and granular. A detailed metabolic analysis and characterization of



	D7.4 – Data Management Plan	Page	8/16
		Rev	3
		Date	29/11/2019

their plant beneficial profile will be performed on selected single microorganisms and on their combination with the examined plant growth promoting active substances (bio-effectors) and formulation components. Therefore, genetic and metabolic data of those microbial strains will be provided as output of this WP. Data of the formulation processes of the products and their shelf-life under different storage conditions will be also provided.

The overall data will be collected in Excel spreadsheets (.xlsx), csv, .txt.

WP3 - Agronomic performance and economic feasibility throughout value chains

This WP will assess the benefit of the new proposed strategy for management of microbial-based plant nutrition and protection under real field conditions. For this purpose, a set of 22 large-scale field trials have been set up (9 on apple and strawberry and 4 on tomato) in 7 EU countries and will be monitored during the next 2-3 growing seasons (depending on the crop). Each field trial will be set with plots (1/treatment) of a reasonable size to simulate real field conditions and dimensioned also according to bio-inocula availability and will be developed comparing the new proposed strategy/ies with the (local) standard cultivation system, under both organic and conventional management systems.

Soil/plant samples will be collected in each site and supplied to partners for analyses and tested in WP4 and, eventually, in WP5. Depending on the kind of analysis the samples will be collected, stored and shipped according to specific methodologies provided by partners of these WPs (see also deliverable D1.2). Therefore, the main output of WP3 is mainly represented by environmental (climatic) and plant (agronomic and phenological) data, but also by economic data. Data will be collected in Excel spreadsheets (.xlsx), csv, .txt.

Data achieved from the different trials on each crop will be used to summarize general results in the different areas and conditions, useful to support activities of WP4, 5 and 6. For example, data collected from the trials will be used to test, validate and refine the model for biodiversity management developed in WP5, fine-tuning it for maximally precise strategic application to the focal cropping systems. This work will begin after the second growing season, to allow adjustments of the trials and/or models considering seasonal climatic fluctuations.


WP4 - Environmental impacts and delivery of soil ecosystem services

The main goal of this WP is to characterize the effects of introduced products on soil biodiversity, quality and ecosystem services in samples collected in WP3. Specifically, plant-associated microbial community (endophytes and rhizosphere), will be assessed over time (3 years). WP4 will also assess the main physical-chemical properties of soil and will carry out a life cycle impact assessment (LCIA) to quantify the impact of these new bio-products on the horticulture, comparing with those systems without them. Thus, most of the data obtained by this WP will be represented by DNA sequencing data but also diversity indices, soil chemical-physical and functional activity and ecosystem assessment will be provided and collected in Excel spreadsheets (.xlsx), csv, .txt.

The outputs from this WP will be used as input to assess the overall environmental benefits of applying microbial-based products in commercial agriculture, and to formulate guidance to farmers/growers and manufacturers on using these microbial products in practice.

Data collected in this WP will include metagenomics, functional traits variance,



	D7.4 – Data Management Plan	Page	9/16
		Rev	3
		Date	29/11/2019

WP5 - Plant-soil-biodiversity interactions and modelling

This WP will deeper investigate on plant-soil-microbe interactions, allowing a closer look at mechanisms underlying effects of bio-inocula on plant responses to biotic and abiotic stress observed in WP2. Specifically, this WP will expand the agro-ecological knowledge on the links and dynamics between biodiversity and crop production, assessing the effects of the bio-inocula on the plant-associated microbiome (beneficial and pathogenic microorganisms) as well as on aboveground pests and their insect biocontrol agents, quantifying the damage of above ground pests on plants, identifying bio-active compounds that are activated in the rhizosphere in the presence of bio-inoculants and assessing effects of bio-inocula on leaching of mineral nutrients. This kind of data will be collected as sequencing data and metadata, collected in different formats such as FASTA, FASTQ, SAM, etc., according to the methodology used and the purpose, and usually renormalized in tabular formats such as Excel spreadsheets .xlsx, .csv, .txt. The same formats will be used for collecting and storing data on minerals and bioavailability of nutrients.

Data from Scanning Electron Microscopy and Energy-dispersive X-ray microanalysis (SEM-EDX) will be provided as high-resolution images and stored in the database as TIFF, JPEG or GIF files. Finally, this WP will also develop and validate methodologies and diagnostic tools for modeling soil biodiversity and create the basis for a decision supporting system (DSS) for a biodiversity-focused soil management, as well as assessing the fate of bio-inocula in soil and the bioprofiling of soil biodiversity. The interoperability and compatibility among the EXCALIBUR database and other DBs and tools required for the development of the provisional model and DSS will be evaluated and discussed in the next months with the involved Partners, after the first data will be provided.

On the other hand, **WP6** (“Dissemination, integration of social sciences and technology transfer”), **WP7** (“Consortium coordination and project management”) and **WP8** (“Ethic issues”) will provide outputs that are not directly included in the EXCALIBUR database. Eventually, depending on the kind of collected data, it will be handled and stored with a case by case approach.

1.2 Data storage

Data will be initially stored in the “Project Collaborative Workspace”, the project intranet already available and described in the deliverable D7.3. Then, data provided by WPs will be migrated in a unified Relational Database Management System, MS Azure SQL based, developed in WP7. It could be queryable by models, DSS and tools developed in WP5.

Data storage and RDBMS services will be provided and managed by CREA on Microsoft cloud platform.




Table 1 – Table Data Overview

Organisation Acronym	category	subcategory	kind of data, unit of measure, eventually broad method, which will be later defined	format
CREA	SOIL	PHYSICAL PROPERTIES	Soil texture, soil bulk density, macroporosity by image analysis of soil thin sections, aggregate stability by wet sieving and/or readily dispersible clay, water retention curve and infiltration test	excel
CREA	SOIL	CHEMICAL PROPERTIES	Soil hydrolytic enzymes (fluorometric method); CO ₂ and N ₂ O potential production (lab incubation + FTIR analyzer and/or gas chromatography)	xls
CREA	SOIL	BIOLOGICAL PROPERTIES - microflora	bacteria, fungi, archaea, protists: culture-dependent and independent approach (DNA- based). Traceability and barcoding of inocula in soil.	xls
CREA	SOIL	BIOLOGICAL PROPERTIES - microfauna	nematodes population, microscopic method	
CREA	SOIL	BIOLOGICAL PROPERTIES - mesofauna	Collembola population, monitoring	
INHORT	ENVIRONMENT	CLIMATIC	metereological data (Temp, rain, etc.) using meteorological unit or soil sensors	
INHORT	ENVIRONMENT	GEOGRAPHIC	GPS positioning and location description	
INHORT	SOIL	PHYSICAL PROPERTIES	structure and texture - common methods	
INHORT	SOIL	CHEMICAL PROPERTIES	Macro and micro elements, common methods of soil analysis	
INHORT	SOIL	BIOLOGICAL PROPERTIES - microflora	bateria and fungi population, classic medium-based method, DNA-based method (PCR), Biolog	
INHORT	SOIL	BIOLOGICAL PROPERTIES - microfauna	nematodes population, microscopic method	
INHORT	SOIL	BIOLOGICAL PROPERTIES - mesofauna	Collembola population, monitoring	
INHORT	SOIL	BIOLOGICAL PROPERTIES - macrofauna	insects and mites populations (pests and beneficials), monitoring	
INHORT	PLANT	AGRONOMIC PARAMETERS	growth and yield	
INHORT	PLANT	PHYSIOLOGICAL PARAMETERS	VOCs, others, biochemical methods	
INHORT	PLANT	PHENOLOGICAL PARAMETERS	classic phenological phases, BBCH scale	
CRPV	ENVIRONMENT	CLIMATIC	metereological data (Temp, rain, etc.) using a regional meteorological survey system (1); plus data logger in the filed (Temp.)(2)	excel
CRPV	ENVIRONMENT	GEOGRAPHIC	GPS positioning and location description	
CRPV	SOIL	PHYSICAL PROPERTIES	structure and texture - in case CREA will not do it in our fields, we can do with common methods	
CRPV	SOIL	CHEMICAL PROPERTIES	in case CREA will not do it in our fields, we can accomplish for Macro and micro elements involving a specialized external laboratory for soil analysis	
CRPV	SOIL	BIOLOGICAL PROPERTIES - macrofauna	insects and mites populations (pests and beneficials), monitoring on plants	
CRPV	PLANT	AGRONOMIC PARAMETERS	growth and yield	
CRPV	PLANT	PHENOLOGICAL PARAMETERS	classic phenological phases, BBCH scale	



Organisation Acronym	category	subcategory	kind of data, unit of measure, eventually broad method, which will be later defined	format
NHM	SOIL	PHYSICAL PROPERTIES	mineralogical composition of soils - X-ray diffraction	
NHM	SOIL	CHEMICAL PROPERTIES	Chemical composition and distribution of micronutrients in several soil reservoirs - common methods of soil analysis	
NHM	SOIL	BIOLOGICAL PROPERTIES - microflora	Sequence data on prokaryotes and eukaryotes	
NHM	SOIL	BIOLOGICAL PROPERTIES - microfauna	eukaryote data based on sequencing	
NHM	SOIL	BIOLOGICAL PROPERTIES - mesofauna	earth worms based on microscopy, eukaryote data based on sequencing	
NHM	SOIL	BIOLOGICAL PROPERTIES - macrofauna	Eukaryotes based on sequencing data	
KIS	ENVIRONMENT	CLIMATIC	metereological data (Temp, rain, etc.) using meteorological unit or soil sensors	excel
KIS	ENVIRONMENT	GEOGRAPHIC	GPS positioning and location description	excel
KIS	SOIL	PHYSICAL PROPERTIES	soil texture - common methods	excel
KIS	SOIL	CHEMICAL PROPERTIES	Macro and micro elements, common methods of soil analysis	excel
KIS	SOIL	BIOLOGICAL PROPERTIES - microflora	total DNA and cDNA samples will be provided to partners doing the general microbial activities, microbial diversity studies, metagenomics (NHM, NIAB EMR, TUG)	sample description sheet in doc or excel format
KIS	SOIL	BIOLOGICAL PROPERTIES - microfauna	soil samples will be provided to NHM and CREA-AA	sample description sheet in doc or excel format
KIS	SOIL	BIOLOGICAL PROPERTIES - mesofauna	soil samples will be provided to NHM and CREA-AA	sample description sheet in doc or excel format
KIS	SOIL	BIOLOGICAL PROPERTIES - macrofauna	soil samples for soil-macrofaunda will be provided to NHM and CREA-AA; insect and mite populations (pests and beneficials), monitoring data and template for pests and beneficials monitoring in strawberries will be provided (WP5.2)	sample description sheet in doc or excel format; template for strawberry entomofauna monitoring in doc format; data on pests and beneficials in strawberry field trials in excel format
KIS	PLANT	AGRONOMIC PARAMETERS	growth and yield; costs of material, operations and labour;	excel
KIS	PLANT	PHYSIOLOGICAL PARAMETERS	pulse amplitude modulation chlorophyll a fluorescence and gas-exchange photosynthesis parameters	excel
KIS	PLANT	PHENOLOGICAL PARAMETERS	classic phenological phases, BBCH scale	excel
UNITO	ENVIRONMENT	CLIMATIC	metereological data (Temp, rain, etc.) using meteorological stations available at local level	csv
UNITO	ENVIRONMENT	GEOGRAPHIC	GPS positioning and location description	txt
UNITO	SOIL	PHYSICAL PROPERTIES	structure and texture - common methods	csv or jpg
UNITO	SOIL	CHEMICAL PROPERTIES	Macro and micro elements, common methods of soil analysis	csv or jpg

Organisation Acronym	category	subcategory	kind of data, unit of measure, eventually broad method, which will be later defined	format
UNITO	SOIL	BIOLOGICAL PROPERTIES - macrofauna	bacteria and fungi population, classic medium-based method, DNA-based method (PCR)	csv and/or jpg
UNITO	PLANT	AGRONOMIC PARAMETERS	growth and yield, disease incidence and severity	csv
UCPH	ENVIRONMENT	CLIMATIC	Local meteorological station; data-loggers in field experiment plots	excel
UCPH	ENVIRONMENT	GEOGRAPHIC	GPS position	
UCPH	SOIL	PHYSICAL PROPERTIES	initial characterization using methods defined in consortium	
UCPH	SOIL	CHEMICAL PROPERTIES	initial characterization using methods defined in consortium	
UCPH	SOIL	BIOLOGICAL PROPERTIES - microflora	Cultivation-dependent methods in field experiment; detection of applied strains	
UCPH	SOIL	BIOLOGICAL PROPERTIES - microfauna	Initial characterization; extracted samples to CREA for ID	
UCPH	SOIL	BIOLOGICAL PROPERTIES - mesofauna	Initial characterization; extracted samples to CREA for ID	
UCPH	SOIL	BIOLOGICAL PROPERTIES - macrofauna	Monitoring in field experiment; coordinated within WP3	
UCPH	PLANT	AGRONOMIC PARAMETERS	Plant growth responses; coordinated within WP3	
TUGRAZ	ENVIRONMENT	CLIMATIC	meteorological data (Temp, rain, etc.) using meteorological unit or soil sensors	
TUGRAZ	ENVIRONMENT	GEOGRAPHIC	GPS positioning and location description	
TUGRAZ	SOIL	BIOLOGICAL PROPERTIES - microflora	bacteria and fungi population, classic medium-based method, DNA-based method (PCR), NGS analysis	
TUGRAZ	PLANT	AGRONOMIC PARAMETERS	growth and yield	
IN+	SOIL	BIOLOGICAL PROPERTIES - microflora	AMF (MPN test, mycorrhization of roots, presence/absence by PCR)	
IN+	PLANT	AGRONOMIC PARAMETERS	growth and yield	
FOEKO	ENVIRONMENT	CLIMATIC	data available by meteorological stations near the trial locations	excel
FOEKO	ENVIRONMENT	GEOGRAPHIC	location description	word
FOEKO	PLANT	AGRONOMIC PARAMETERS	Measure vigour of the shoots and harvest in comparison to control. Measure efficacy of the treatment (if scab, control of infestation rate of 50 shoots per treatment)	excel
FOEKO	PLANT	PHENOLOGICAL PARAMETERS	classic phenological phases, BBCH scale	
NIOO-KNAW	SOIL	BIOLOGICAL PROPERTIES - microfauna	AMF root colonization (microscope, staining)	excel
NIOO-KNAW	PLANT	AGRONOMIC PARAMETERS	growth and yield	Excel
NIOO-KNAW	PLANT	PHYSIOLOGICAL PARAMETERS	Plant tissue macronutrients (common AA/EA methods), chlorophyll, VOC (QTOF GC-MS), growth, biomass allocation, gene expression	Excel
NIOO-KNAW	PLANT	PHENOLOGICAL PARAMETERS	Transition vegetative/flowering phase	Excel

	D7.4 – Data Management Plan	Page	13/16
		Rev	3
		Date	29/11/2019

2 FAIR DATA

Within EXCALIBUR we adopted the Commission's approach which described collected data as *"as open as possible, as closed as necessary"*. In fact, we are committed to protect the privacy of the Partners and the confidentiality of their specific results. In these cases the data will not be made available for public use without a specific permission. However, according with Art. 29.1 of the GA, *"each beneficiary must – as soon as possible – disseminate its results by disclosing them to the public by appropriate means (other than those resulting from protecting or exploiting the results), including in scientific publications (in any medium). This does not change the obligation to protect results in Article 27, the confidentiality obligations in Art. 36, the security obligations in Art.37 or the obligations to protect personal data in Art.39, all of which still apply"*.

Thus, only those data (or part of the data) that refer to practices, technologies covered by any secrecy clauses in the Consortium Agreement or in the exploitation agreements reached within the consortium (or with external parties) will be excluded. In all other cases we will do our best to make the EXCALIBUR data as broadly available as possible. This means the FAIR principles will be held.

2.1 Making data findable, including provisions for metadata

Due to the fact that the collected datasets will be diverse and belong to different fields, the metadata standard used to describe the dataset will be the standard adopted by the European **OpenAIRE** repository (<https://www.openaire.eu/>), which will be used. The depositors will create the metadata manually when uploading the datasets to **Zenodo** (<https://zenodo.org/>). With respect to the open access for the peer reviewed publications, the bibliographic metadata must be in a standard format and must include all the following terms: „European Union (EU)“, „Horizon 2020“, acronym and grant number of the project, the publication date and length of embargo (if applicable), and a persistent identifier, e.g. a DOI (Digital Object Identifier), to make them easily and uniquely citeable.

The repository assigns Handle/DOIs for persistent identification and citability of the dataset. All datasets will be accompanied by a “readme file” in pdf format, which describes data factors and variables names and units. Keywords will be provided based on the expertise of the depositors who are familiar with the most commonly used keywords in their field.

2.2 Making data openly accessible

As mentioned in the previous section, all datasets supporting publications will be made openly and publicly available on the OpenAire repository Zenodo upon acceptance of the manuscript for publication.

Only data gathered by partners outside of the project work plan and protected by IPR, or inside the work plan but containing confidential information (e.g. related to personal interviews) will be kept closed for privacy reasons. Furthermore, when a dataset turns openly accessible, it



	D7.4 – Data Management Plan	Page	14/16
		Rev	3
		Date	29/11/2019

will be announced on the web page of the EXCALIBUR project (<https://www.excaliburproject.eu/>), where a link to the dataset on Zenodo will be available.

The vast majority of the datasets will be made available as text (.txt, .csv), spreadsheets (.xls/.xlsx) and to a lesser extent as documents (.doc,.docx or .pdf). Some observational images as .jpg. Finally the GIS implementation of data generated in the field studies will generate geotiff, .img, .asc, .shp (shapefile) or geojson files accessible through QGIS, ArcGIS or SAGAGIS, and extensions .kml, .kmz accessible through Google Earth.

EXCALIBUR adhere to the pilot for open access to research data (**Open Research Pilot, ORD**) adopting an open access policy of all project results, guidelines and reports, providing on-line access to scientific information that is free of charge to the reader. Open access will be provided in two categories:

1) **Scientific publications:** according to Art.29.2 of the GA, stating that under Horizon 2020 *“each beneficiary must ensure open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results”*. The EXCALIBUR Consortium adheres to the EU open access to publications policy, choosing as most appropriate route towards open access self-archiving (also known as „Green Open Access”), namely „a published article or the final peer-reviewed manuscript is deposited in an online repository before, alongside or after its publication. Repository softwares usually allow authors to delay access to the article (the so-called „embargo period”). The Consortium will ensure open access to the publication within a maximum of six months.

The dissemination of EXCALIBUR results will occur according with the initial plan for exploitation and dissemination of results (PEDR), such as creation of the web page of the project, public workshops, press releases, participation in international events, etc. In compliance with the GA, free-online access will be privileged for scientific publication, following the above-mentioned rules of „green” open access. All relevant information and dissemination material (leaflets, brochures, papers, public deliverables, etc.) will be also freely available on the EXCALIBUR website.

In some specific cases and according to the rules of open access, the dissemination of research results will be managed by adopting precautionary IPR protection protocols, in order to not to obstacle the possibility of protecting the achieved foreground with preventive disclosures (see also deliverable 7.2 – Management and IP guidelines).

2) **Research data:** according to Art. 29.3 of the GA, regarding the digital research data generated in the action, the EXCALIBUR beneficiaries will:

(a) *Deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate – free of charge for any user – the following:*

1. *The data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible;*
2. *Othe data, including associated metadata, as specified and within the deadlines laid down in the „data Management Plan*



	D7.4 – Data Management Plan	Page	15/16
		Rev	3
		Date	29/11/2019

(b) *Provide information - via the repository – about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and – where possible – provide the tools and instruments themselves).*

The EXCALIBUR Consortium will make a great effort, where possible, to make this research data available as open data or through open services. However, it is important to note that because of the low maturity of this document and some existing uncertainties about the data collected in the project, additional details are going to be inserted in here as the project progresses.

2.3 Making data interoperable

The depositors will strive to use metadata commonly and internationally used in their field of research, since no widely accepted specific metadata vocabularies exist. Data variables names and units commonly accepted by the international scientific community and international system of units, will be used. The vast majority of the datasets will be made available as spreadsheets (.xls/.xlsx format) and to a lesser extent as .doc/.docx or .pdf, so that they are fully used by potential users.

2.4 Increase data re-use (through clarifying licenses)

Our intention is to make as much data as possible re-useable for third parties. Restriction will apply only when privacy, IPR or other exploitations ground are in play.

When possible, the data set will be licensed under an Open Access license. Specifically, the project will make use of the CC-BY 4.0, free cultural works license. The particular license applicable to each dataset will be decided on an individual basis, though the recommendation provided by either CC-BY or CC-BY-SA.


The data set will remain re-usable after the end of the project by anyone interested in it, with no access or time restrictions, since the Zenodo repository will be used. Each archived data set will have its own permanent repository ID and will be easily accessible. We expect most of the data generated to be made available without restrictions and, as mentioned before, only data sets subject to IPR and confidentiality issues will be restricted. Where this is going to be the case, agreements will be made based on the individual data sets. Requests for the use of the data by externals will be approved by the project Consortium.

Regarding data quality, since the vast majority of datasets will support publication in peer-reviewed open access journals, the standards of quality (validation of the sample, replication and comparison with results of similar studies) will be met.

2.5 Allocation of resources

There are no costs associated to the described mechanisms to make the database FAIR and long term preserved since we are using the Zenodo repository. The project coordinator has the ultimate responsibility for the data management in the project.



	D7.4 – Data Management Plan	Page	16/16
		Rev	3
		Date	29/11/2019

2.6 Data security

On one hand, the Zenodo repository takes care of security and backup of the deposited data.

Data protection services are provided by Microsoft to prevent the loss of data.

For more details regarding data protection and other ethics aspects related to data, please, refer to deliverable D8.1 H – Requirement No.1, where the informed consent template for data sharing is included as annex.

2.7 Ethical aspects

The ethical aspects related to the personal data processing are addressed in the following Ethics WP 8 deliverables:

- D8.1. H - Requirement No. 1
- D8.2. POPD - Requirement No. 2

