This project is part of the PRIMA Programme supported by the European Union

INNOVATIVE GREENHOUSE SUPPORT SYSTEM IN THE MEDITERRANEAN REGION: EFFICIENT FERTIGATION AND PEST MANAGEMENT THROUGH IOT BASED CLIMATE CONTROL — IGUESSMED

o•••o•••o

Deliverable 1.9
Data Management Plan

Version 1.0

Due date:	01/10/20
Submission date:	14/04/21
Deliverable leader:	CREA
Author list:	Alejandra Navarro Garcia (CREA), Luca Incrocci (UNIPI), Aliona Lupu (Subcontractor Iniziativa Cube S.r.l.)

Dissemination Level

☒	**PU:**	Public
☐	**PP:**	Restricted to other programme participants (including the Commission Services)
☐	**RE:**	Restricted to a group specified by the consortium (including the Commission Services)
☐	**CO:**	Confidential, only for members of the consortium (including the Commission Services)

Abstract

The aim of the Data Management Plan (DMP) is to provide the plan for managing the data generated and collected during the project. The DMP describes the data management life cycle for all datasets to be collected, processed and/or generated by a research project.

It covers: a) the handling of research data during and after the project, b) what data will be collected, processed or generated, c) what methodology and standards will be applied, d) whether data will be shared/made open and how and e) how data will be curated and preserved.

The deliverable will be updated during the project lifetime.

Table of Contents

Table Summary

1 Introduction

The iGUESS-MED project aims to develop a Decision Support System (DSS) able to effectively manage fertigation and prevent plant diseases and pests in tomato crops grown in soil and soilless in commercial greenhouses of the Mediterranean region. This innovative greenhouse DSS will be developed to (i) help greenhouse farmers to improve the management of fertigation in areas with low (saline) quality waters (ii) to reduce the use of chemicals by a sustainable and integrated pest and disease control and (iii) to improve the climatic efficiency in the existent greenhouse by low-cost climate actions. The DSS will allow obtaining healthier and higher quality productions and higher yields, while will reduce the use of water and the losses of nutrients and chemicals to the environment. iGUESS-MED will be able to manage efficient fertigation, to forecast diseases and pests, and to improve the climatic efficiency in tomato greenhouses, using only climate data acquisition and basic information on cropping system. The DSS will provide feedbacks and alerts about crop needs and real time recommendations to the farmers through friendly portable real time data visualization tools as PC, tablets or smartphones. To achieve this objective, new models for calculating crop evapotranspiration will be performed by integrating sensor data from plant, soil and climate, and forecasting models for assessing disease and pest risks will be developed by using the Integrated Pest Management.

The project consortium (research centres, SMEs and end-users of EU and non-EU countries belonging to the Mediterranean basin) will collaborate from the beginning to make the DSS marketable involving, end-users and stakeholders to validate the system in own greenhouses, reducing gaps between research, application developers and farmers. The application of DSS will benefit the workers and the consumers, providing better working conditions, crop healthiness and reduction of environmental impact.

1.1 Purpose and scope of this document

The Deliverable D1.9 Data Management Plan (DMP) represents the first version of the DMP of the iGUESS-MED project. The current deliverable has on purpose to ensure proper and sound management of the data that will be collected, processed and generated within iGUESS-MED.

The concrete objective of the document is to ensure that the work is performed according to clear and appropriate rules and methods, at all level and all times, to ensure a satisfactory completion of the Grant Agreement (GA). The Data Management Plan (DMP) covers principles and protocols for data collection, data transfer, data processing, data storage, and data retention and sharing, and will be regularly reviewed and updated throughout the period of the grant. The DMP will ensure that data management and protection in iGUESS-MED is compliant with EU principles and standards, including standards outlined in COE108 and 95/46/EC, and with relevant national data protection laws and institutional data management policies.

The DMP describes the data management life cycle for all datasets to be collected, processed and/or generated by the project. It covers:

- the handling of 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.

The DMP is not a fixed document, but it is likely to evolve during the whole lifespan of the project serving as a working document.

1.2 Data management

According to the latest Guidelines on FAIR Data Management in Horizon 2020 released by the EC Directorate-General for Research & Innovation on the 30th of July 2016 “beneficiaries must make their research data findable, accessible, interoperable and reusable (FAIR) ensuring it is soundly managed”. FAIR data management is also part of the ORD Pilot promoted by the European Commission. The purpose of the ORD is to improve and maximize access to and re-use of research data generated by H2020 projects and to take into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security, as well as data management and preservation issues.

iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection. Notwithstanding this option, the Consortium members, as beneficiaries, adopt a FAIR approach in data management.

New data generated by iGUESS-MED will include quantitative (and/or qualitative) data from the application of sensor technology, IoT, advanced agronomic management, simulation models and self-calibrating mathematical algorithms, etc. New datasets will also be produced by revisiting previous studies, reviewing national and regional secondary data. iGUESS-MED will additionally use and analyse secondary data sourced from public archives such as Eurostat and similar national and regional archives.

Besides, as part of the iGUESS-MED work plan, interviews, and workshops in the context of Living Labs (LL) will take place. The results of these interaction with people who are not members of the consortium will be used to provide data for the project. Therefore, it is important for the iGUESS-MED partners to ensure that all forms of engagement will be conducted in an ethical way. The Consortium will also engage with a large number of human participants who are involved in farming and agricultural sector or experts in these areas. The processing of personal data and on the free movement of such data will include personal data management, privacy and security, informed consent forms and information sheets.

Individual partners will be responsible for the collection, transfer and storage of data from their own project activities, within the framework of the DMP. The DMP outlines access rights for consortium participants to data held by other consortium members, with raw and/or personal data only shared where necessary for analysis in order to meet project objectives. Data to be shared across the consortium will be stored on secure server maintained by CREA and other partners involved in data management, with daily back-up on a secure external mirror server, and transferred using secure cloud system.

Security protocols for data will be tailored according to risk assessments, with encryption of any sensitive personal data or commercially sensitive data. Where relevant and appropriate, personal data will be anonymised as soon as practical after collection.

1.3 Document structure

The deliverable is structured in the following chapters:

- Chapter 1 includes an introduction of the deliverable and a brief description on how Data Management is approached in Horizon 2020 program;
- Chapter 2 includes a summary of the data management under iGUESS-MED;
- Chapter 3 describes the methodology used for data management;
- Chapter 4 describes the allocation of the resources;
- Chapter 5 includes data security issues;
- Chapter 6 describes ethical aspects;
- Chapter 7 includes a description of the datasets to be used in iGUESS-MED reflected on the template provided by the EC;
- Chapter 8 includes a summary table of the iGUESS-MED datasets;
- Chapter 9 includes a conclusion to the deliverable.

2 Data Summary

2.1 Purpose of the data collection/generation and its relation to the objectives of the project

The general objective of iGUESS-MED is to support a transition toward innovative, sustainable and competitive Mediterranean horticultural greenhouses by developing, validating and transferring a pioneering Decision Support System (DSS) for the MED greenhouses, which is able to: (i) reduce the nutrient leakages into sub-surface and groundwaters by optimizing the fertigation management (both irrigation and fertilization) under low quality water conditions (i.e., saline water); (ii) reduce the use of chemicals by a sustainable and integrated pests and diseases control; (iii) increase the productivity by an improved and cost-effective efficiency of climatic control procedures, introducing specific low-cost solutions to apply to pre-existent greenhouse structures (i.e., use of diffusing cover films, improved natural ventilation). The work will be carried out on tomato as reference crop, in soil and soilless culture in low-tech greenhouses typical of the Mediterranean region, by applying participatory and integrated interdisciplinary toolkit of novel and emerging technologies such as sensor technology, IoT, advanced agronomic management, simulation models and self-calibrating mathematical algorithms.

iGUESS-MED will follow an implementation based on the following steps:

1. A scientific approach based on experimental results to design the operational framework and their analysis to obtain a whole and complete feedback to improve the support system and, to publish, disseminate and communicate the evaluation results, discussing the best practices based on the gained experience. In this first step, accurate studies (first 18 months) to monitor and record **water, soil, crop and climate parameters** will be conducted, integrating **sensor data from plant (transpiration – sap flow sensor), soil (evaporation – lysimeters, or soil water volume sensors) and climate (continuous micro-climate and environmental data - extendable meteorological mast) in different cropping systems and latitudes**. These preliminary studies will allow, on the one hand, the development of improved models to estimate ET_c considering especially saline waters to achieve a high precision in the fertigation management and to reduce chemical fertilizers supply, and, on the other hand, the development of an IPM approach based on climate monitoring to predict favourable conditions for pests or diseases occurrence, reducing insecticides and fungicides supply.
2. Development of an intermediate DSS prototype by processing and analysing the collected data, to implement selected algorithms for parameters and processes estimation. Thus, an easy alarm system based on climate, soil and crop system data, like the OPI system hardware/software bundle, entirely developed in-house by EVJA and already available on the

market, will be produced for recommendation/decision making to the farmers. The intermediate DSS will be tested and validated by the demo-pilot commercial greenhouses involved in the project and located in different Mediterranean countries (Italy, Spain, Tunisia, and Turkey) and presented to other stakeholders.

3. Creation of LL in each site for actively involving local stakeholders in the development, validation and testing of the new technology and the associated new greenhouse management system, in real production facilities and contexts. Knowledge gaps, end-user needs, working conditions, gender issues as well as the usefulness and social acceptability of the innovation will be discussed in the LL: qualitative and quantitative information will be collected that will feed into the socio-economic impact assessment. Preliminary assessment results will be discussed during the labs to create reinforcing feedback loops and allow technology and management system adaptation to site-specific needs and issues.

Within the above mentioned phases different environmental and greenhouse activity related data will be collected, such as humidity, temperature, solar radiation, leaf wetness, soil water status, plant transpiration, tomato pests and disease occurrence, biocontrol data, fertilization related data, etc.. Besides, due to the creation of LL, personal data of human participants will be also collected.

2.2 Types and formats of data generated/collected

Project partners will collect and manage data sets related to the following domains (according to the Classification of Statistical Activities (CSA)[1]):

- Agriculture, forestry, fisheries
- Environment
- Business and agricultural surveys
- Dissemination, data warehousing
- Management and development of technological resources

Collected data will be represented by qualitative (binary and nominal) and quantitative (both discrete and continuous) data. Among the main format, the following will be considered: text, spreadsheets, software, maps and audio-visuals.

[1] http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=DSP_GEN_DESC_VIEW_NOHDR&StrNom=CSA&StrLanguageCode=EN

2.3 Data Management Programme Responsible and Data Set Responsible

Each iGUESS-MED beneficiary has to respect the policies set out in the DMP. Datasets have to be created, managed and stored appropriately.

At the overall project level, the role of Data Management Programme Responsible (DMPR) was assigned to Mr. Luca Incrocci (UNIPI) and partners handling and processing data have been asked to nominate a Data Set Responsible (DSR).

The DMPR is responsible for management of data, he must ensure compliance with the safety rules in physical and logical terms by applying local and EU regulations and according to best practices in respect of security, personal data and ethics issues.

Validation and registration of datasets and metadata is the responsibility of the beneficiary that generates the data (DSR). The DSR is responsible for the management of data coming from his/her organization or processed provided in relation to all aspects, ranging from collection and conservation to anonymization, and aggregation, when required. He/she must ensure compliance with the safety rules in physical and logical terms by applying the rules laid down by local and EU regulations and by best practice in respect of security, personal data and ethics issues.

Backing up data for sharing through open access repositories is the responsibility of the member processing the data. Updates and management of the different versions are in charge of the DSR.

Quality control is the responsibility of the DMPR.

3 Methodology

3.1 Fair Data

The Data Management Plan regards all the data sets that will be collected, processed and/or generated within the project. The methodology the consortium follows to create the DMP is as follows:

- create a data management policy according to the EC DMP template. The elements of the defined data management policy were used to create a DMP template which was sent to the partners of the consortium in order to fill it in with information for each relative data set;
- analyse the completed by the project's partners DMP templates.

According to the EC guidelines regarding the DMP, this document should address for each data set collected, processed and/or generated in the project the following elements:

- data set reference and name, standards and metadata;

- data set description;
- data sharing;
- archiving and preservation.

In this section, we provide a detailed description of these elements in order to ensure their understanding by the partners of the consortium. For each element, we also describe the strategy that will be used to address it.

3.2 Data set reference and name, standards and metadata: making data findable, including provisions for metadata[2]

Data will be formalised in structured database for the purposes of the project, and they will be **findable, accessible, interoperable and reusable (FAIR) by Consortium members**.

The **data will be stored** on data custodian servers and must be accompanied by appropriate metadata to allow users to search them.

In order to be able to distinguish and easily identify data sets, each **data set is assigned with a unique name**. This name can also be used as the identifier of the data sets.

In order to design the data set names, we use the following practice:

- Each data set name consists of four different parts separated with a dot: *CSACode.CountryCode.PartnerName.DatasetName*, where:
    - The *CSACode* part describes the theme that the dataset fits. We use the Classification of Statistical Activities (CSA)[3] in order to represent different themes. The Classification of Statistical Activities (CSA) is used to classify the statistical activities undertaken by national and international statistical organizations. It is used as the basis for the Database of International Statistical Activities, and for the list of subject matter domains in the Content-oriented Guidelines, produced by the SDMX (Statistical Data and Metadata eXchange) initiative. The classification has three levels. The five "domains" form the first level, and relate to the broad type of statistical activities: (a) Demographic and social statistics, Economic statistics, (b) Economic statistics, (c) Environment and multi-domain statistics, (d) Methodology for data collection, processing dissemination and analysis, and (e) Strategic managerial issues of official statistics. The second level specifies "activities" within these domains and the third level covers more detailed "subject areas". To represent the theme in the name of the

[2] Metadata: the details about what, where, when, why, and how the data were collected, processed, and interpreted. Metadata include descriptions of how data and files are named, physically structured, and stored as well as details about the experi-ments, analytical methods, and research context. It is generally the case that the utility and longevity of data relate directly to how complete and comprehensive the metadata are. The amount of effort devoted to creating comprehensive metadata may vary substantially based on the complexity, types, and volume of data."
(Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4619636/)

[3] http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=DSP_GEN_DESC_VIEW_NOHDR&StrNom=CSA&StrLanguageCode=EN

data sets, we use the number that corresponds to each second or third level "activities" of CSA. For example, a data set that comes from the environment domain will have 3.1 as theme which corresponds to "Environment" in CSA.

- The CountryCode part represents the country associated with the dataset using ISO Alpha-2 country codes:
    - IT for Italy
    - ES for Spain
    - TR for Turkey
    - TN for Tunisia

- The *PartnerName* part represents the name of the organization (e.g. data owner, data custodian) associated with the dataset:

Participant organisation short name
CREA
EVJA
UNIPI
BIOPLANET
UAL
CAJAMAR
La Caña
Akdeniz University
CRRHAB

- The *DatasetName* represens the full name of the dataset. An example of a data set's name could be the following:

3.1.IT.CREA.Environment

The above name indicates that the data set describes Environment from the CREA in Italy and that the data set regards the 3.1 (Environment) domain.

Search keywords will be identified for each data set to optimize possibilities for open data re-use.

When appropriate, depending on the nature of single data set, the **version** will be used in form of progressive number, with indication of major and minor release or in form of date.

3.3 Data sharing: making data openly accessible, interoperable and data re-use

This section includes a description of the **data that will be made accessible and how**. It also explains why some datasets cannot be made open due to possible, legal, contractual or ethical issues. It is possible that some beneficiaries have decided to keep their data closed. A description of the potential data repositories is also included along with the potential software tools required to access the data.

For iGUESS-MED project, the following options for open repositories of data, metadata, documentation or code will be considered: (a) The Registry of Research Data Repositories[4], (b) Zenodo[5], (c) OpenAIRE[6]. In addition, the project website will be used as a repository for public project deliverables and results. Finally, additional specific repositories may be used for the iGUESS-MED purposes hosted by the project's partners.

Data interoperability is detailed for every dataset of iGUESS-MED. Issues such as the allowing of data exchange between researchers, institutions or even countries are covered along with all the technicalities including standards for formats, metadata vocabularies or ontologies of vocabularies. The issue of interoperability will be discussed among the consortium members in the upcoming project plenary meeting.

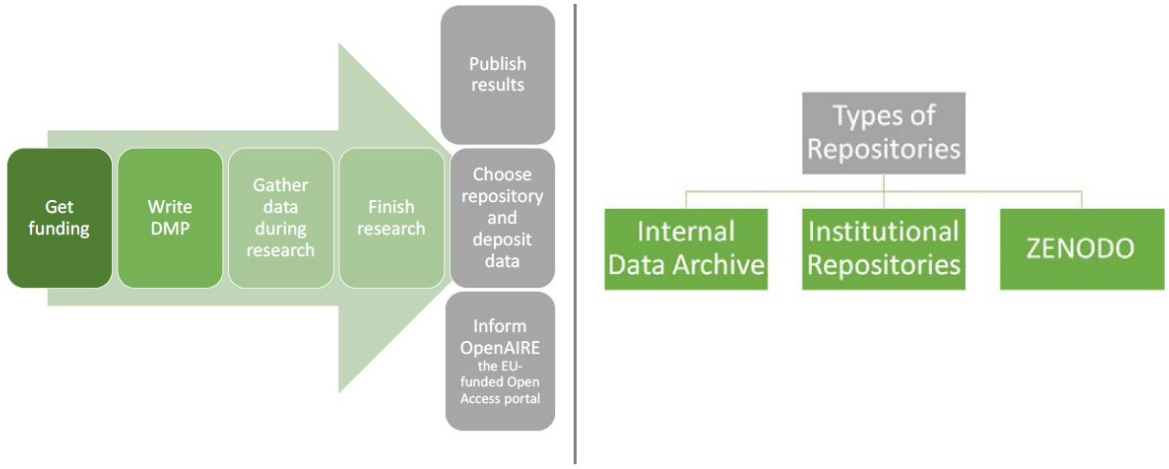
This section will also describe in the future versions the licenses, if any, under which data will be re-used in iGUESS-MED. It includes provisions regarding the period when data will be available for reuse and if third parties will have the option to use the data and when. Quality assurance processes will be also described.

3.3.1 Data Repository - insights

The publication of the data requires:

- the choice of a repository to deposit data
- Inform OpenAIRE

[4] http://www.re3data.org/
[5] https://zenodo.org/
[6] https://www.openaire.eu/



It is mandatory to store and track data that has been opened.

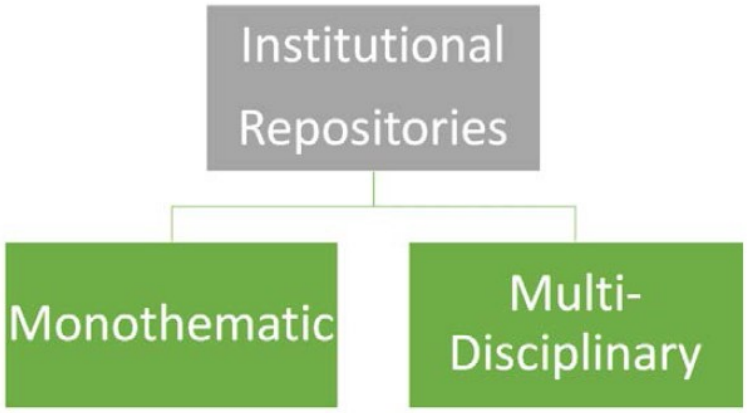
There are pro and cons about the choice of one of these repository typologies.

1) Repositories – Self-Archiving

Example: Organization website, Project website with a reserved area for the datasets, hard disk, pen drive.

PRO	CONS
Easy to use (support by internal IT personnel)	Need to have a persistent identifier
Easy to update data	Need to have a standardized communications protocol
Easy to data maintenance	Need to have an authentication and authorization procedure
	Need to have a formal, accessible, shared, and broadly applicable data usage license
	Ensure access after the closure of the project
	Link the own web to OpenAIRE

2) Repositories – Institutional



Nowadays, most of European universities have an institutional repository where papers, books and data can be archived.

Example: University website

PRO	CONS
Easy to update data	Data upload only for members
Persistent identifier	Created for academic scopes
Easy to data maintenance	Data quality in charge of the author
Standardized communications protocol	Some costs
License to access	
Accessible, shared, and broadly applicable language	
Clear and accessible data usage license	
Data preservation	
Compliant with OpenAIRE	

3) Repositories – Zenodo (https://help.zenodo.org/)

ZENODO builds and operates a simple and innovative service that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results (data and publications) that are not part of the existing institutional or subject-based repositories of the research communities. ZENODO enables researchers, scientists, EU projects and institutions to: easily share the long tail of small research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science. Display their research results and get credited by making the research results citable and integrate them into existing reporting lines to funding agencies like the European Commission. Easily access and reuse shared research results.

Main characteristics of a multidisciplinary repository:

- Multiformat;
- License to choose (CC, GNU, BSD…);
- Free of charge;
- Simple to search;
- Filter by document type;
- List of results: filter by document type, author, access, FP-7 and Horizon2020 projects;
- Ranking of results.

PRO	CONS
Persistent identifier	Search tool to be completed
Altmertrics	
Personal collections	
Login with ORCID and Github	
DOI versioning	

The selected repositories for iGUESS-MED project will be:

- The self-archiving repository represented by:
  a) project website for public data, e.g. public project deliverables and results
  b) reserved/dedicated repository on the CREA organisation website for closed dataset with a share point intranet access only to Consortium members
- The institutional repositories on the public research centres and universities websites for publications.

3.4 Archiving and preservation

The partners will decide and describe the procedures that will be used in order to ensure long-term preservation of the data sets. This field will provide information regarding the duration of the data preservation, the approximate end volume, the associated costs and the plans of the consortium to cover the costs.

4 Allocation of Resources

FAIR data management in iGUESS-MED project is under WP1 – Project coordination lead by the project coordinator CREA. Within the project budget, a specific amount of person months has been dedicated for that activity. The coordinator has already assigned the task to a dedicated data manager, appointing Mr. Luca Incrocci from UNIPI as DM. All costs related to FAIR data management, if any, that will occur during project implementation will be covered by the project budget. Any other cost that may relate to long-term data preservation will be discussed among consortium members.

5 Data Security

Data security is of major importance in the iGUESS-MED project. Special attention will be given to the security of sensitive data. The protection of personal data will be ensured through procedures and appropriate technologies, like the use of HTTPS protocol for the encryption of all internet transactions (https://creagov.sharepoint.com/sites/iGUESSMED).

iGUESS-MED project data files will be managed, processed, and stored in a secure environment (lockable computer systems with passwords, firewall system in place, power surge protection, virus/malicious intruder protection) and protecting by controlling access to digital files with encryption and/or password protection.

6 Datasets Analysis

The DMP currently includes 25 datasets proposed by the partners of the Consortium. This section presents the results of the analysis of these data sets.

6.1 Dataset reference and name

Table 1 presents the datasets recorded by the partners of the consortium along with the CSA domain they belong to. Data sets come from 5 different domains.

#	DATASET	CSA DOMAIN
	CREA	
1	3.1.IT.CREA.Environment	Environment
2	4.5.IT.CREA.Dissemination, data warehousing	Dissemination, data warehousing
	EVJA	
3	3.1.IT.EVJA.Environment	Environment
4	4.5.IT.EVJA.Dissemination, data warehousing	Dissemination, data warehousing
5	5.5.IT.EVJA.Management and development of technological resources	Management and development of technological resources
	UNIPI	
6	2.4.1.IT.UNIPI.Agriculture, forestry, fisheries	Agriculture, forestry, fisheries
7	3.1.IT.UNIPI.Environment	Environment
8	4.3.4.IT.UNIPI.Business and agricultural surveys	Business and agricultural surveys
9	4.5.IT.UNIPI.Dissemination, data warehousing	Dissemination, data warehousing
	UAL	
11	3.1.ES.UAL.Environment	Environment
12	4.5.ES.UAL.Dissemination, data warehousing	Dissemination, data warehousing
	CAJAMAR	
13	2.4.1.ES.CAJAMAR.Agriculture, forestry, fisheries	Agriculture, forestry, fisheries
14	3.1.ES.CAJAMAR.Environment	Environment
15	4.3.4.ES.CAJAMAR.Business and agricultural surveys	Business and agricultural surveys

16	4.5.ES.CAJAMAR.Dissemination, data warehousing	Dissemination, data warehousing
Akdeniz University		
18	2.4.1.TR.Akdeniz University.Agriculture, forestry, fisheries	Agriculture, forestry, fisheries
19	3.1.TR.Akdeniz University.Environment	Environment
20	4.3.4.TR.Akdeniz University.Business and agricultural surveys	Business and agricultural surveys
21	4.5.TR.Akdeniz University.Dissemination, data warehousing	Dissemination, data warehousing
CRRHAB		
22	2.4.1.TN.CRRHAB.Agriculture, forestry, fisheries	Agriculture, forestry, fisheries
23	3.1.TN.CRRHAB.Environment	Environment
24	4.3.4.TN.CRRHAB.Business and agricultural surveys	Business and agricultural surveys
25	4.5.TN.CRRHAB.Dissemination, data warehousing	Dissemination, data warehousing

Table 1 – Datasets and CSA domains

6.2 iGUESS-MED project data sets

The tables below present the datasets recorded by the partners of the consortium along with the DMP Methodology presented above.

6.2.1 CREA- iGUESS-MED project data sets

CSA Domain	Dataset accessibility
Agriculture, forestry, fisheries	Public and open
Environment	Private and closed
Business and agricultural surveys	Public, but the processing of personal data and on the free movement of such data will include personal data management, privacy and security, informed consent forms and information sheets.
Dissemination, data warehousing	Dissemination data is public and open, while data warehousing is private and closed.
Management and development of technological resources	Private and closed

Table 2 – Dataset accessibility

DMP component	3.1.IT.CREA.Environment 4.5.IT.CREA.Dissemination, data warehousing
Data summary	• CREA will collect data necessary for the calibration and validation of the sub-routine of DSS; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
Making data findable, including provisions for metadata	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be developed
Making data openly accessible	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables and dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee • The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
Making data interoperable	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_CREAenvironment_Navarro_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
Increase data re-use	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.

DMP component	3.1.IT.CREA.Environment 4.5.IT.CREA.Dissemination, data warehousing
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.2 EVJA- iGUESS-MED project data sets

DMP component	3.1.IT.EVJA.Environment 4.5.IT.EVJA.Dissemination, data warehousing 5.5.IT.EVJA.Management and development of technological resources
Data summary	• EVJA will collect data necessary for the calibration and validation of the sub-routine of DSS; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, sensors, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
Making data findable, including provisions for metadata	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be developed
Making data openly accessible	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables and dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee • The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
Making data interoperable	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example:

DMP component	3.1.IT.EVJA.Environment 4.5.IT.EVJA.Dissemination, data warehousing 5.5.IT.EVJA.Management and development of technological resources
	20210410_iGUESSMED_EVJAenvironment_Parisi_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
Increase data re-use	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.3 UNIPI- iGUESS-MED project data sets

DMP component	2.4.1.IT.UNIPI.Agriculture, forestry, fisheries 3.1.IT.UNIPI.Environment 4.3.4.IT.UNIPI.Business and agricultural surveys 4.5.IT.UNIPI.Dissemination, data warehousing
Data summary	• UNIPI will collect data necessary for the calibration and validation of the sub-routine of DSS and personal data in LL and agricultural surveys; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, sensors, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
Making data findable, including provisions for metadata	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be developed
Making data openly accessible	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables, public dataset dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website

DMP component	2.4.1.IT.UNIPI.Agriculture, forestry, fisheries 3.1.IT.UNIPI.Environment 4.3.4.IT.UNIPI.Business and agricultural surveys 4.5.IT.UNIPI.Dissemination, data warehousing
	• Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
Making data interoperable	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_UNIPIenvironment_Incrocci_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
Increase data re-use	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.4 UAL- iGUESS-MED project data sets

DMP component	3.1.ES.UAL.Environment 4.5.ES.UAL.Dissemination, data warehousing
Data summary	• UAL will collect data necessary for the calibration and validation of the sub-routine of DSS; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, workshops, and

DMP component	3.1.ES.UAL.Environment 4.5.ES.UAL.Dissemination, data warehousing
	compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
Making data findable, including provisions for metadata	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be developed
Making data openly accessible	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables and dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
Making data interoperable	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_UALenvironment_Gallardo_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
Increase data re-use	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.5 CAJAMAR- iGUESS-MED project data sets

DMP component	2.4.1.ES.CAJAMAR.Agriculture, forestry, fisheries 3.1.ES.CAJAMAR.Environment 4.3.4.ES.CAJAMAR.Business and agricultural surveys 4.5.ES.CAJAMAR.Dissemination, data warehousing
Data summary	• CAJAMAR will collect data necessary for the calibration and validation of the sub-routine of DSS and personal data in LL and agricultural surveys; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, sensors, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
• **Making data findable, including provisions for metadata**	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections Metadata, standard identification mechanisms, search keywords will be developed
• **Making data openly accessible**	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables, public dataset dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee • The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
• **Making data interoperable**	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_CAJAMARenvironment_Fernandez_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
• **Increase data re-use**	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all.

DMP component	2.4.1.ES.CAJAMAR.Agriculture, forestry, fisheries 3.1.ES.CAJAMAR.Environment 4.3.4.ES.CAJAMAR.Business and agricultural surveys 4.5.ES.CAJAMAR.Dissemination, data warehousing
	• The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.6 Akdeniz University - iGUESS-MED project data sets

DMP component	2.4.1.TR.Akdeniz University.Agriculture, forestry, fisheries 3.1.TR.Akdeniz University.Environment 4.3.4.TR.Akdeniz University.Business and agricultural surveys 4.5.TR.Akdeniz University.Dissemination, data warehousing
Data summary	• AU will collect data necessary for the calibration and validation of the sub-routine of DSS and personal data in LL and agricultural surveys; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, sensors, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
• **Making data findable, including provisions for metadata**	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be developed
• **Making data openly accessible**	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables, public dataset dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee • The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
• **Making data**	• Only some elaborated data will be available in outreach and scientific

DMP component	2.4.1.TR.Akdeniz University.Agriculture, forestry, fisheries 3.1.TR.Akdeniz University.Environment 4.3.4.TR.Akdeniz University.Business and agricultural surveys 4.5.TR.Akdeniz University.Dissemination, data warehousing
interoperable	publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_AUenvironment_Buyuktas_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
• **Increase data re-use**	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

6.2.7 CRRHAB - iGUESS-MED project data sets

DMP component	2.4.1.TN.CRRHAB.Agriculture, forestry, fisheries 3.1.TN.CRRHAB.Environment 4.3.4.TN.CRRHAB.Business and agricultural surveys 4.5.TN.CRRHAB.Dissemination, data warehousing
Data summary	• CRRHAB will collect data necessary for the calibration and validation of the sub-routine of DSS and personal data in LL and agricultural surveys; • Generated and collected data will be both qualitative and quantitative data (text, spreadsheets, software), maps and audiovisuals. • New data will be used for these CSA domains. • The origin of the data will be direct observation, field work, sensors, workshops, and compilations of data from other studies, with and expected size of 1 terabyte • Data will be useful for decision-makers, planning department, investors, entrepreneurs, academia, local community
FAIR data	
• **Making data findable, including provisions for**	• We follow naming conventions and clear version numbers as mentioned and agreed in the previous sections • Metadata, standard identification mechanisms, search keywords will be

DMP component	2.4.1.TN.CRRHAB.Agriculture, forestry, fisheries 3.1.TN.CRRHAB.Environment 4.3.4.TN.CRRHAB.Business and agricultural surveys 4.5.TN.CRRHAB.Dissemination, data warehousing
metadata	developed
• **Making data openly accessible**	• Some elaborated data will be openly available in outreach and scientific publications. Public deliverables, public dataset dissemination data will be openly available too. Datasets belonging to private domains cannot be openly accessible and will be kept accessible only for Consortium project members, in line with the reasons for opting out. (View table 2). • Private data will be made accessible on an on-line repository and public data on the project website • Access to private dataset will be via share point cloud, while public data via webpage • An open source code will not be included, since many domains are private dataset and will be only opened for Consortium members • Appropriate arrangements with the identified repository will be developed • Access to closed dataset will be provided via login and authentication system • There is no need for a data access committee • The identity of the person accessing the data will be ascertained via microsoft/ CREA authentication system
• **Making data interoperable**	• Only some elaborated data will be available in outreach and scientific publications. Public deliverables and some dissemination data will be openly available too through the project webpage. • Private Datasets cannot be openly accessible and are in line with the reasons for opting out. (View table 2). • The files will provide a preview of the content, and will be organized in a logical way (yyyy-mm-dd) identify the responsible party and convey the work history. For example: 20210410_iGUESSMED_CRRHABenvironment_Laarif_v1.02.docx • The use of standard vocabularies for all data types present in the data set to allow inter-disciplinary interoperability among consortium members will be developed. • Project will be used common or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Yes
• **Increase data re-use**	• PRIVATE datasets will not be licensed in Open source CC. • Data produced and/or used in the project may be usable by third parties only if are public, for private data, time to publish or seek patent will be discussed by consortium members in a later stage, and afterwards the data becomes available to all. • The private data produced and/or used in the project will not be reusable by third parties. iGUESS-MED project Consortium has opted out of the ORD Pilot to allow the protection of project results at the end for commercial exploitation and for the reason of incompatibility with privacy/data protection.
Archiving and preservation	CREA cloud
Allocation of resources	N/A
Data security	Yes
Ethical aspects	Yes
Other	N/A

7 Conclusions

The purpose of this document was to provide the plan for managing the data generated and collected during the project. Specifically, the DMP described the data management life cycle for all datasets to be collected, processed and/or generated within iGUESS-MED project.

It covered:

- the handling of 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.

Following the EU's guidelines regarding the DMP, this document will be updated during the project lifetime.