D1.6: 1ˢᵗ Data Management Plan

WP1 – Project Management

Authors: Polimachi Simeonidou, Ioanna Pavlou

Disclaimer

Any dissemination of results reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

Copyright message

APOLLO

This project is co-funded by the European Union

Document Information

Grant Agreement Number	687412	Acronym		APOLLO	
Full Title	Advisory platform for small farms based on earth observation				
Horizon 2020 Call	EO-1-2015: Bringing EO applications to the market				
Type of Action	Innovation Action				
Start Date	1st May 2016	**Duration**		34 months	
Project URL	-				
Document URL	-				
EU Project Officer	Iulia SIMION				
Project Coordinator	Polimachi Simeonidou				
Deliverable	D1.6: 1st Data management plan				
Work Package	WP1 – Project Management				
Date of Delivery	Contractual	M6		Actual	M6
Nature	R – Report	**Dissemination Level**		PU - Public	
Lead Beneficiary	DRAXIS				
Lead Author	Machi Simeonidou	**Email**		msimeonidou@draxis.gr	
	DRAXIS	**Phone**		+302310274566	
Other authors	Lazaros Xenidis (DRAXIS), Ioanna Pavlou (DRAXIS), Ioulia Anastasiadou (DRAXIS), Camille Pelloquin (Starlab), Andrés Cuesta García (AgriSat), Anna Osann (AgriSat), Andreea Bucur (TUW), Alexander Gruber (TUW), Dragutin Protic (UBFCE), Aleksandar Sekulić (UBFCE), Ognjen Antonijević (UBFCE), Stefan Milutinović (UBFCE), Lefteris Mamais (EVF), Dimitri Papadakis (EVF), Zisis Tsiropoulos, (AUA), Evangelos Anastasiou (AUA), Spyros Fountas (AUA)				
Reviewer(s)	Camille Pelloquin (Starlab)				
Keywords	Data management, Horizon 2020, FAIR, Open Research Data Pilot				

Document History

Version	Issue Date	Stage	Changes	Contributor
1.0	12/10/2016	Draft	Initial version for review	DRAXIS
2.0	26/10/2016	Draft	Comments received	Starlab
3.0	28/10/2016	Draft	Comments included	DRAXIS
4.0	29/10/2016	Final	Submission	DRAXIS

Table of Contents

Acronyms and Abbreviations 6

Executive summary 7

1   Introduction 8

    1.1   Data management in Horizon 2020 program 8

2   Methodology 9

    2.1   DMP template 9

        2.1.1   Data summary 9

        2.1.2   FAIR data 9

            2.1.2.1   Making data findable, including provisions for metadata 9

            2.1.2.2   Making data openly accessible 9

            2.1.2.3   Making data interoperable 10

            2.1.2.4   Increase data re-use (through clarifying licenses) 10

        2.1.3   Allocation of resources 10

        2.1.4   Data security 10

        2.1.5   Ethical aspects 11

        2.1.6   Other issues 11

    2.2   Methodological steps in APOLLO 11

3   Datasets 12

    3.1   WP1 – Project Management 12

        3.1.1   Advisory board data 12

    3.2   WP2 – Users' needs analysis and specifications 13

        3.2.1   User requirements data 13

        3.2.2   Co-creation data 14

    3.3   WP3 – Earth Observation data products 15

        3.3.1   Satellite data and Pre-processing 15

        3.3.2   Soil Moisture Data 17

        3.3.3   Local meteorological data 18

        3.3.4   Crop condition data 19

        3.3.5   Biomass data 20

        3.3.6   Validation results 21

    3.4   WP4 – Development of added value services 22

        3.4.1   Added value service dat 22

    3.5   WP5 – Platform development and integration 23

        3.5.1   System architecture and design 23

        3.5.2   Specification of APOLLO service requirements 24

Acronyms and Abbreviations

AB	Advisory board
API	Application programming interface
DMP	Data Management Plan
DEM	Digital elevation model
DOI	Digital Object Identifier
EO	Earth Observation
EC	European Commission
XML	Extensible Markup Language
H2020	Horizon 2020 program
IDF	International DOI Foundation
ISMN	International Soil moisture network
LAI	Leaf Area Index
LAN	Local Area network
ORD Pilot	Open Research Data Pilot
REST	Representational state transfer
RMSE	Root mean square error
SSL	Secure Sockets Layer
SAR	Synthetic aperture radar
TIFF	Tagged Image File Format
VIs	Vegetation Indices
VLAN	Virtual LAN
WMS	Web mapping service

Executive summary

The purpose of the current deliverable is to present the 1st Data Management plan of the APOLLO project. The deliverable has been compiled with the collaborative work among the coordinator and the consortium partners who were involved in data collection, production and processing. It includes detailed descriptions of all datasets that will be collected, processed or generated in all Work packages during the course of the 34 months of APOLLO project. The deliverable is submitted six months after project start as required by the European Commission (EC) through the latest guidelines: The Open Research Data Pilot (ORD Pilot). For the methodological part, the latest EC guidelines have been adopted for the current deliverable.

The deliverable is structured in the following chapters:

Chapter 1 includes an introduction to the deliverable and a brief description on how Data Management is approached in Horizon 2020 (H2020) program along with the importance of it.

Chapter 2 includes a description of the methodology used, an analysis of the chapters of the provided template and last the methodological steps followed in APOLLO.

Chapter 3 includes a description of the datasets to be used in APOLLO reflected on the template provided by the EC.

Chapter 4 includes a conclusion to the deliverable.

Chapter 5 includes a summary table with all the datasets included in 1st APOLLO DMP.

1 Introduction

The Deliverable D1.6 1st Data Management Plan (DMP), represents the first version of the DMP of the APOLLO project. APOLLO is an Innovation Action project funded under the Horizon 2020 program of the EC that will last 34 months. As such, APOLLO participates in ORD Pilot, and, therefore, is providing, as requested, the current deliverable six months after the beginning of the project (M6, October 2016).

The current deliverable has on purpose to ensure proper and sound management of the research data that will be collected, processed and generated within APOLLO. The concrete objectives of the document are to (a) detail the handling of research data during and after the project, (b) describe the methodology and standards required, (c) identify whether and how data will be shared, exploited or made accessible for verification, and re-used, and (d) identify how they 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. This document is the first of the three versions to be produced throughout APOLLO project duration. In this respect, the 2nd version (D1.7) will be submitted on Month 12 (April 2017), while the 3rd and final version (D1.8) is due on Month 34 at the end of the project (February 2019). As required, the upcoming versions of the APOLLO DMP will have a clear version number and include a timetable for any occurring data updates.

1.1   Data management in Horizon 2020 program

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[1] "*beneficiaries must make their research data findable, accessible, interoperable and reusable (FAIR) ensuring it is soundly managed*". Further information on how to approach the writing of a DMP has been provided in the Joint EUDAT-OpenAIRE webinar "*How to write a Data Management Plan*"[2] originally broadcasted on 7 July, 2016 and currently available online.

FAIR data management is 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. For the 2017 work programme, the Open Research Data pilot has been extended to cover all the thematic areas of H2020.

The inclusion of a DMP is a key element for FAIR data management in a H2020 project. In a DMP, the data management life cycle for the data to be collected, processed and/or generated by a H2020 project is described and analysed. DMP should also include information on (a) the handling of research data during & after the end of the project, (b) what data will be collected, processed and/or

---

[1] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] https://eudat.eu/events/webinar/joint-eudat-openaire-webinar-%E2%80%9Chow-to-write-a-data-management-plan%E2%80%9D

generated, (c) which methodology & standards will be applied, (d) whether data will be shared/made open access and (e) how data will be curated & preserved (including after the end of the project)[3].

2 Methodology

2.1 DMP template

In order to assist the beneficiaries with the completion of the DMP, the EC produced and provided a template that act as a basis for data description. The template contains a set of questions that beneficiaries should answer with a level of detail appropriate to the project. If no related information is available for a given dataset, then the phrase "*Non-applicable*" or N/A will be used. In the following paragraphs, the main sections and proposed contents of the template are listed and presented, along with the way APOLLO reflects to these sections.

2.1.1 Data summary

In this section, beneficiaries are asked to describe (a) the purpose of the data collection or generation and how this purpose reflects to the objectives set in the project as a whole, (b) the types and formats of data that will be generated or collected, (c) the origin of the data, (d) the expected size of the data, and also (e) whether existing data will be reused and (f) the usefulness of the described datasets.

2.1.2 FAIR data

2.1.2.1 Making data findable, including provisions for metadata

This section includes a description of metadata and related standards, the naming and keywords to be used and if a standard identification mechanism such as Digital Object Identifier (DOI) will be used.

In the context of APOLLO the following naming convention will be used for all the datasets of the project. First the work package number will be placed, then the serial number of the dataset within this work package and last the dataset title, all separated with underscore (Data_<WPno>_<serial number of dataset>_<dataset title>). An example can be the following Data_WP1_1_Advisory_board_data. However, it has to be noted that this naming convention describes only the general dataset that can contain files of different size and format. The naming of each separate file follows a different naming convention that is proposed by the partners who creates the files.

The use of a standard identification mechanism in for the datasets of APOLLO will be decided by the project consortium in collaboration with the data manager. If it turns out to be necessary, the use of the Guidelines and standards provided by the International DOI Foundation (IDF) and the DOI system and ISO 26324[4] will be considered.

2.1.2.2 Making data openly accessible

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

---

[3] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

[4] https://www.doi.org

This project is co-funded by the European Union

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.

In the context of APOLLO, the following options for open repositories of data, metadata, documentation or code will be considered: (a) The Registry of Research Data Repositories[5], (b) Zenodo[6], (c) OpenAIRE[7]. An addicional specific repository may serve the APOLLO purposes include is the International Soil moisture network[8], hosted by the Technical University of Vienna. That repository is a data hosting facility for in-situ soil moisture networks, where data are harmonized, quality-checked and made available to registered users on a cost-free basis. Finally, the project website will be used as a repository for project deliverables, results or product demonstrators and trial versions.

In the context of the 1st APOLLO DMP, not any arrangements have been made with an identified repository. The need for a data access committee will be decided by the consortium during the upcoming plenary meeting. However, a dedicated data manager is working together with the coordinator DRAXIS on APOLLO data management.

2.1.2.3 Making data interoperable

In this section, data interoperability is detailed for every dataset of APOLLO. 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.

2.1.2.4 Increase data re-use (through clarifying licenses)

This section describes the licenses, if any, under which data will be re-used in APOLLO. 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 are also described.

2.1.3   Allocation of resources

FAIR data management in APOLLO project is under WP1 – Project management lead by the project coordinator DRAXIS. 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. All costs related to FAIR data management 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.

2.1.4   Data security

Data security is of major importance in the APOLLO 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

---

[5] http://www.re3data.org/

[6] https://zenodo.org/

[7] https://www.openaire.eu/

[8] https://ismn.geo.tuwien.ac.at/

and appropriate European and Internet security standards from ISO, ITU, W3C, IETF and ETSI. If data will be kept in a certified repository, then the security standards of that repository will apply.

2.1.5   Ethical aspects

In APOLLO, there are several ethical issues that can have an impact on data sharing. These issues include the following:

- Personal data collection during the pilot implementation with the participation of real actual service users,
- Ethical approvals,
- Procedures for data collection, storage, protection, retention and destruction,
- Participation and data imports and exports from an EU candidate country Serbia,
- Application of the *right to be forgotten and to erasure* in personal data collection,
- Application of the *Need to know* principle,
- Procedures regarding the recruitment of users during the pilot phase.

Details related to these issues are provided in the Ethics and Security requirements chapters of the APOLLO DoA[9]. The Ethics section in Part B has been updated according to the requirements of the Ethics Screening Report.

2.1.6   Other issues

In this section, other issues can be covered not included above such as the use of other national/funder/sectorial/departmental procedures for data management.

2.2   Methodological steps in APOLLO

For the first version of APOLLO DMP, the following methodological steps were followed:

(1) The coordinator and responsible for the implementation WP1 - Project Management DRAXIS sent to all partners, well in advance, an email notifying them about the upcoming deliverable. Contribution was asked from all partners that were involved in any data collection in each task of the WPs.
(2) In parallel, the latest guidelines from the EC regarding data management were sent to all partners to be informed. Sufficient time was given to send their input.
(3) The project team collaborated efficiently and contributed with the needed information.
(4) The last version (v3.0) of the deliverable, compiled by DRAXIS, was reviewed by Starlab.

The first version of the APOLLO DMP is intended to provide an initial screening off the data to be collected, processed and produced within APOLLO. It is also the first attempt to collect the vision and input from all the partners involved in any data management option. During the upcoming plenary meeting in Athens on 10-11 November 2016, special attention and a separate session will be given to data management in order to provide further clarifications and conclusions on data management. The upcoming deliverables D1.7 and D1.8 will be further refined using the input from the plenary meetings.

---

[9] APOLLO project Description of Action, Chapter 1.4 Ethics requirements pp41/120, Part B Section 5.1 Ethics pp 119-120/120

3 Datasets

3.1 WP1 – Project Management

3.1.1 Advisory board data

Contributions for DRAXIS: Lazaros Xenidis

DMP component	WP1_1_Advisory_board_data
1. Data summary	Purpose: To manage and document all the data that will be produced from the Advisory Board (AB) related activities during the APOLLO project. Advices and quality control, from the members, ensure the development and application of state-of-the-art methodologies, algorithms and generated services.
	Data formats: *.xlsx , *.mp3, *.doc, *.pdf files
	The data includes mp3 files that will contain the recording of the teleconferences, webinars, and discussions held among the consortium members and the AB. The transcript of the discussion will be converted to text in doc format for archiving purposes but also to make it accessible and readable to the hearing impaired. Reports related to the AB will also be in doc format. In order to make the results of the AB more visible, a serie of *.xlsx format tables will be produced mainly with coded recommendations and their relation to specific WP, deliverable or responsible partner. An xls table will also be created to include the list of the contact details of the Members.
	No existing data will be used or reused. Data will be produced only within the time duration and in the context of APOLLO project as described in the GA and DoA.
	A 20-minutes recording has the size of 40 Mb.
	The total file of this dataset will be approximately 1 Gb. This will include audio files, text and xlsx files.These data would be useful for research purposes since most of the AB members come from the academic community. In addition, it would be useful to anyone who would like to reflect on the impact that the AB would have to the evolution of the project itself.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	The inclusion of metadata for the current dataset has not been yet decided. The initial AB data audio files and doc transcripts will not be findable. Unique and persistent identifiers will not be used for this dataset. The following naming for the dataset will be WP1_1_Advisory_board_data.
2.2 Making data openly accessible	Audio files and doc transcripts of the interviews will not be available openly since they represent an initial and unstructured reflection of the AB members to the project. These initial data will be kept in secure servers in DRAXIS premises. However, the final tables from the excel files will be included in the four AB meetings minutes (D1.2 – D1.5) that will be publicly available through the project's website. Pdf files that will include the tables with the member's advice will be accessible through any pdf viewer

	software. Final deliverables will be deposited in the servers used for the project website.
2.3. Making data interoperable	Not applicable. There are no interoperability issues related to this dataset.
2.4. Increase data re-use (through clarifying licences)	The dataset cannot be accessed and there is no permission related with this dataset. Final Data will be publicly available within the related deliverables deposited in the project website. Any individual or third party interested to access and reuse the AB data can download the deliverable from the project website in pdf format. There is no restriction for the use of this final data since the dissemination level of the AB Meetings minutes is public.
3. Allocation of resources	All costs related to the AB data collection and processing are covered by the project budget with dedicated person months under WP1 Project management.
4. Data security	Audio, doc and xls files will be deposited in DRAXIS servers and will be protected with the DRAXIS server's security protocol. Pdf files that will be deposited in the project website will follow the security protocol of the website hosting service.
5. Ethical aspects	There are ethical issues regarding the AB data. First, for the audio recording of the AB, the consent of the members for this recording was asked prior to the beginning of the discussion. Their consent is recorded in the same audio file. It has been made clear to the AB members that their advice will be used only in the context of APOLLO project. Moreover, the xls file related to personal AB member's details contains information that is already publicly available from their institutions' webpage such as their email, affiliation and field of expertise.
6. Other	N/A

3.2   WP2 – Users' needs analysis and specifications

3.2.1   User requirements data

Contributions for AUA from Zisis Spyropoulos, Evangelos Anastasiou and Spyros Fountas

DMP component	Data_WP2_1 User_requirements_data
1. Data summary	Questionnaires were designed and shared during the tests pilots that were held in Greece, Serbia and Spain. These questionnaires were filled by farmers and crop consultants which provided us with data about: • Demographics • General information • APOLLO services development • APOLLO platform development These data were processed and used to extract the user requirements for the APOLLO platform. The user requirements will then enable us to develop

	user friendly and more important efficient farm management/advisory services. The information extracted from the processing of the data will be available in the deliverable 2.1 and a collection of the data from the filled questionnaires will be made available in an excel file (.xlsx) of ~70KB.
2. FAIR Data **2.1. Making data findable, including provisions for metadata**	The deliverable along with the excel file (collected data accumulation) will be uploaded to the APOLLO website were any user will be able to download. There is no need for specific metadata or versioning since those data won't change in the future. Naming convention will be used where the files will be named: • WP2_1_Deliberable.pdf • WP2_1_Questionnaire_results.xlsx
2.2 Making data openly accessible	Both the deliverable and the excel files will be openly accessible by anyone though the APOLLO website.
2.3. Making data interoperable	The deliverable file will be "as is" in a pdf format only to be used for reference. On the other hand, the questionnaire data accumulation file will be in excel format (.xlsx) so any user will be able to operate on it and even export the data in formats like csv (commonly used by the majority of platforms).
2.4. Increase data re-use (through clarifying licences)	The data will be under Public Domain Dedication and License (PDDL) for any user to obtain them and use them as they wish. These data are not to be updated/altered in the upcoming future. The quality of the data is already ensured by us before upload and since no one else but us can import them into the website, no automatic insurance is needed. As long as the website is up and running, the data will be as well.
3. Allocation of resources	N/A
4. Data security	The data is openly accessible so no specific security measurements are needed.
5. Ethical aspects	N/A
6. Other	N/A

3.2.2   Co-creation data

For Task 2.3 contributions from Andres Questa and Anna Osann (AgriSat).

DMP component	WP2_2_Co-creation_Data
1. Data summary	The purpose of the data collection is to document the process of co-creation of APOLLO tools; the development & evaluation is one of the central objectives of the project. The following data formats will be produced:

	a) Excel database of extended stakeholder mapping results; b) Audio and video documentation of some events (2-3 files per pilot area). Data will be collected from events (mostly during meetings, but also online) of interaction with stakeholders. The data will be used in the ongoing multi-actor stakeholder process and will be confidential, to remain within each pilot area stakeholder community (with sharing to be provided only with the consensusby the information providers, i.e. each stakeholder).
2. FAIR Data 2.1. Making data findable, including provisions for metadata	N/A (Confidential Data)
2.2 Making data openly accessible	N/A (Confidential Data)
2.3. Making data interoperable	N/A (Confidential Data)
2.4. Increase data re-use (through clarifying licences)	N/A (Confidential Data)
3. Allocation of resources	N/A (Confidential Data)
4. Data security	N/A (Confidential Data)
5. Ethical aspects	N/A (Confidential Data)
6. Other	N/A (Confidential Data)

3.3   WP3 – Earth Observation data products

3.3.1   Satellite data and Pre-processing

Contributions from Camille Pelloquin (Starlab)

DMP component	WP3_1_Satellite_data
1. Data summary	The objective of the collection of the data is to produce pre-processed datasets for the next step of the processing, as inputs for soil moisture, Crop condition, and Biomass estimation models/algorithms. The data collected in the T3.1 are satellite datasets from both optical and SAR instruments. On one side, the optical datasets are coming from three missions: • Landsat-8 (Landsat8 OLI L1T) from USGS Earth Explorer • Sentinel-2 (S2 MSI L1C) from the ESA Scientific Hub of Copernicus in ESA SAFE format. On the other side the SAR datasets are coming from:

- Sentinel 1 mission (S1 L1 IW GRDH mode) from the ESA Scientific Hub of Copernicus in ESA SAFE format (GeoTiff for the measurement format).

The task will produce pre-processed data:

- Sentinel-2 BOA(GeoTiff format) will be generated using Sen2Cor processor for atmospheric, terrain and cirrus correction http://step.esa.int/main/third-party-plugins-2/sen2cor/
- Landsat-8 BOA (GeoTiff format) from USGS Earth Explorer or using ARCS processor
- Sentinel-1 CGR calibrated and geo referenced (GeoTiff format)

Size of the input data:
- Landsat8 OLI L1T: Raw zipped image = 1GB, unzipped ~1.6GB,
- S2 MSI L1C: Per tile (all bands, raw data)- L1C = 600MB *73days= 43800MB~42GB; (test areas - Serbia=1 tile, Greece=1 tile, Spain=4 tiles.
- S1 L1 IW GRDH: Per tile (both polarization), two satellites (S1A, S1B) = 1GB *365/3 days ~ 120GB; Serbia test area=1 tile, Greece test area=1 tile, Spain test area=1 tile.

Size of the input data:
- Sentinel-2 L2A BOA: 1200MB *73days= 87600MB~85GB (test areas - Serbia=1 tile, Greece=1 tile, Spain=4 (not whole) tiles Landsat-8 BOA: ~1.6 GB
- Sentinel-1 CGR: Per tile (both polarization), two satellites (S1A, S1B) = 60MB *365/3 days ~ 7.2GB; Test areas: Serbia=1 tile, Greece=1 tile, Spain=1 tile.

The data will be useful for partners of the project to ingest it in their physical parameters estimation algorithm, including soil moisture, Crop condition, and Biomass.

2. FAIR Data 2.1. Making data findable, including provisions for metadata	Metadata stored in the Rasdaman database together with the raster file such as date, bounding box, projection, mission will be useful for discoverability of the data. The data produced in T3.1 will not be available for farmers, only to the next steps of processing, including models and parameters estimation. In addition, each dataset produced will be associated to a unique ID corresponding to the area of interest requested. Naming conventions for the data will be: ● S2_BOA_[ID]_[YYYYMMDD].tiff ● L8_BOA_[ID]_[YYYYMMDD].tiff ● S1_CGR_[ID]_[YYYYMMDD].tiff
2.2 Making data openly accessible	Only project partners will have permissions to access the data. All the data, associated metadata and documentation will be deposited into the official Apollo web server and available through Rasdaman's REST API and Geoserver's web mapping service (WMS). Only web browser and Internet access are needed to access the data.
2.3. Making data interoperable	The input data will be in JPEG2000 or GeoTiff format, whereas output data will be in GeoTiff format with associated metadata and provided through REST API. Based on that, there is possibility for combining with other datasets and using in various open software applications.

2.4. Increase data re-use (through clarifying licences)	The data will be accessible for use to all APOLLO project partners through RESTAPI from the Rasdaman database.
3. Allocation of resources	N/A
4. Data security	All the data will be stored in Rasdaman server for the purpose of serving data and also on separate storage server, both with backup procedures.
5. Ethical aspects	N/A
6. Other	N/A

3.3.2   Soil Moisture Data

Contributions from Camille Pelloquin (Starlab)

DMP component	WP3_2_Soil_moisture_data
1. Data summary	In this task, the algorithms will use the Sentinel-1 datasets (calibrated and georeferenced) produced in task T3.1. The output data of task 3.2 are soil moisture maps for farmers, and for project partners serving as input for modelling such as tillage scheduling. The soil moisture maps will be provided in GeoTiff format and made available through the Rasdaman database. Size of the output data: • Sentinel-1 Soil moisture maps: Per tile, two satellites (S1A, S1B) = 10MB *365/3 days ~ 1220MB; Test areas: Serbia=1 tile, Greece=1 tile, Spain=1 tile.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Metadata such as date, bounding box, projection, and parameter will be useful for discoverability of the data. The data produced in T3.2 will be available for farmers, and project partners to be used at the next steps of processing/modelling. In addition, each dataset produced will be associated with a unique ID corresponding to the area of interest requested. Naming conventions for the data will be: ● sm_[ID]_[YYYYMMDD].tiff
2.2 Making data openly accessible	Only farmers participating in pilots and project partners will have permissions to access the data. All the data, associated metadata and documentation will be deposited into the official Apollo web server and available through Rasdaman's REST API and Geoserver's web mapping service (WMS). Only web browser and Internet access are needed to access the data.
2.3. Making data interoperable	The input data will be in GeoTiff format and provided through REST API. Based on that, there is possibility for combining with other datasets and using in various open software applications.
2.4. Increase data re-use (through clarifying licences)	The data will be usable by third parties through REST API, but only for those parties who paid for services. This will be clearly defined through business model. No quality check is applied automatically. The data will be validated comparing with ground or station measurements and deriving statistical indicators (e.g. root mean square error - RMSE, mean, median, dynamic range).

3. Allocation of resources	N/A
4. Data security	All the data will be stored in Rasdaman server for the purpose of serving data and also on a separate storage server, both with backup procedures.
5. Ethical aspects	N/A
6. Other	N/A

3.3.3   Local meteorological data

Contributions for UBFCE Dragutin Protic, Aleksandar Sekulić, Ognjen Antonijević, Stefan Milutinović

DMP component	Data_WP3_3_Local_meteorological_data
1. Data summary	Spatio-temporal models created from daily data from Meteorological stations (GHCN-Daily, GSOD, ECA&D) and additional predictors (e.g. MODIS, DEM, WI, geometric temperature trend, IMERG, etc.) will provide daily temperature and precipitation at 1km spatial resolution. This data is going to be used for final agricultural models in order to provide agronomic services. Additionally, the data should be commercially available to third parties. Generated and collected data are in next types and formats: ● Daily temperature data, Daily precipitation data, DEM – DEMSRE3, TWI – TWISRE3 and GPM L3 - IMERG 1 Day multi-satellite precipitation estimates are in GeoTiff format, data type - small signed integer (INT2S, from -32767 to 32767) ● Meteorological stations are in CSV format ● MODIS/Terra Land Surface Temperature LST/E Daily L3 MOD11A1 are in HDF-EOS format Size of the data: ● Temperature and precipitation: 2img * 150MB * 365days =109500MB ~ 107GB per year (whole World) ● Meteorological stations: ~2GB per year ● MODIS/Terra: ~300MB per year
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Metadata such as creation date, version, bounding box, projection, quality of the data will be useful for discoverability of the data. Only registered farmers will be able to find temperature and precipitation data for their parcels and for specific date. Metadata will be available as data info on Apollo web site. Naming conventions for the data will be: ● temp_*<date>* (e.g. temp_20160530) ● prcp_*<date>* (e.g. prcp_20160530)
2.2 Making data openly accessible	Only farmers participating in pilots and project partners will have permissions to access the data. All the data, associated metadata and documentation will be deposited on official APOLLO web server and available through Rasdaman's REST API and Geoserver's web mapping service (WMS). Only web browser and Internet access are needed to access the data.
2.3. Making data interoperable	The data will be in GeoTiff, CSV or GML format with associated metadata and provided through REST API. Based on that, there is possibility for combining with other datasets and using in various open software applications.

2.4. Increase data re-use (through clarifying licences)	Licence will be defined through business model that will be defined during the project implementation. The data will be usable by third parties through REST API, but only for those parties who paid for services. This will be defined clearly through business model. No quality check is applied automatically. The data will be validated comparing with ground or station measurements and also using statistical measures (e.g. root mean square error - RMSE, mean, median, maximum, minimum, etc.).
3. Allocation of resources	N/A
4. Data security	All the data will be stored in Rasdaman server for the purpose of serving data and also on separate storage server, both with backup procedures.
5. Ethical aspects	N/A
6. Other	N/A

3.3.4   Crop condition data

Contributions for UBFCE Dragutin Protic, Aleksandar Sekulić, Ognjen Antonijević, Stefan Milutinović

DMP component	Data_WP3_4_Crop_condition_data
1. Data summary	Crop condition data (Vegetation indices, biophysical parameters -LAI, Chl content) is combined data from different optical sensors (Sentinel-2, Landsat 8). Beside the data will be useful for agronomic services (agronomic models), it will be useful for farmers and third parties. These data will be in next types and formats: ● Vegetation Indices (Vis), Leaf Area Index (LAI) are in GeoTiff format, data type - floating point (FLT4S) ● Sentinel-2 L1C are in JPEG2000 format (.jp2) ● Sentinel-2 L2A are in GeoTiff format ● Landsat 8 Level 1T - GeoTiff format Size of the data: ● Crop condition data for one pilot: 5img * 250MB * 73days = 91250MB~ 90GB ● Sentinel-2: Per Tile(all bands, raw data)- L1C = 600MB *73days= 43800MB~42GB; Per Tile(all bands, processed data)- L2A = 1200MB *73days= 87600MB~85GB (test areas - Serbia=1 tile, Greece=1 tile, Spain=4 (not whole) tiles. These estimates are very rough, this is worst case scenario (image for tile each 5 days, without clouds, Sentinel API fully functioning etc..). In reality on many occasions the images will be full of clouds, not whole tiles cover the interest area etc. - the clouds will be cropped out together with unnecessary areas, so the real size will probably be smaller. ● LANDSAT 8: Raw zipped image = 1GB, unzipped ~1.6GB, processed ~1.6 GB

2. FAIR Data 2.1. Making data findable, including provisions for metadata	Metadata will be date, version, bounding box, projection, quality of the data, etc. This metadata will help finding the data. Metadata will be available in data info on Apollo web site. Only registered users and administrators will be able to access the data. The data will have following names: ● ndvi_<*sentinelObservationTime*> (e.g. ndvi_20160808T094139 for tile observed on 2016/08/08 at 09:41:39) ● lai_<*sentinelObservationTime*> (e.g. lai_20160808T094139)
2.2 Making data openly accessible	Only farmers participating in pilots and project partners will have permissions to access the data. All the data, associated metadata and documentation will be deposited on official APOLLO web server and available through Rasdaman's REST API and Geoserver's web mapping service (WMS). Only web browser and Internet access are needed to access the data.
2.3. Making data interoperable	The data will be in GeoTiff or GML format with associated metadata and provided through REST API. Based on that, there is possibility for combining with other datasets and using in various open software applications.
2.4. Increase data re-use (through clarifying licences)	Licence will be defined through business model that will be defined during the project implementation. The data will be usable by third parties through REST API, but only for those parties who paid for services. This will be defined clearly through business model. No quality check is applied automatically. The data will be validated comparing with ground or station measurements and also using statistical measures (e.g. root mean square error - RMSE, mean, median, maximum, minimum, etc.).
3. Allocation of resources	N/A
4. Data security	All the data will be stored in Rasdaman server for the purpose of serving data and also on separate storage server with backup procedures.
5. Ethical aspects	N/A
6. Other	N/A

3.3.5   Biomass data

Contributions for UBFCE Dragutin Protic, Aleksandar Sekulić, Ognjen Antonijević, Stefan Milutinović

DMP component	Data_WP3_5_Biomass_data
1. Data summary	Biomass data are generated from well-known, established models based on VIs. These models are obtained using VIs and in situ data from Apollo and other European projects. It will be used for final agricultural models in order to provide agronomic services and for commercial availability to third parties. Generated and collected data are in next types and formats: ● Biomass data are in GeoTiff format, data type - floating point (FLT4S) ● VIs data are explained in Data_WP3_2_Crop_condition_data Size of the biomass data for one pilot: ● 1img * 250MB * 73days = 18250MB~ 18GB

2. FAIR Data 2.1. Making data findable, including provisions for metadata	Metadata like date, version, bounding box, projection, quality of the data, etc. will help farmers finding biomass data for its parcels and for specific time. Metadata will be available in data info on Apollo web site. This data will be named: ● biomass_<*sentinelObservationTime*> (e.g. biomass_20160808T094139 for tile observed on 2016/08/08 at 09:41:39)
2.2 Making data openly accessible	Only farmers participating in pilots and project partners will have permissions to access the data. All the data, associated metadata and documentation will be deposited on official Apollo web server and available through Rasdaman's REST API and Geoserver's WMS. Only web browser and Internet access are needed to access the data.
2.3. Making data interoperable	The data will be in GeoTiff or GML format with associated metadata and provided through REST API. Based on that, there is possibility for combining with other datasets and using in various open software applications.
2.4. Increase data re-use (through clarifying licences)	Licence will be defined through business model that will be defined during the project implementation. The data will be usable by third parties through REST API, but only for those parties who paid for services. This will be defined clearly through business model. No quality check is applied automatically. The data will be validated comparing with ground or station measurements and also using statistical measures (e.g. root mean square error |(RMSE), mean, median, maximum, minimum, etc.).
3. Allocation of resources	N/A
4. Data security	All the data will be stored in Rasdaman server for the purpose of serving data and also on separate storage server with backup procedures.
5. Ethical aspects	N/A
6. Other	N/A

3.3.6   Validation results

Contributions from Alexander Gruper and Andreea Bucur (TUW)

DMP component	WP3_6_ Validation_results
1. Data summary	The purpose of the dataset is thevalidation of EO data products. Validation of datasets produced within APOLLO project is required in terms of precision and accuracy. The EO data products developed within APOLLO will be validated against already well established products (in-situ, other EO data, modelled data). The results of the validation will be presented as a report and possibly also as peer-reviewed scientific paper(s). The data sets have various origins such as field sensors, handheld sensors, model outcomes and drone imagery. It is not yet determined which datasets will be used for validations and it will be determined by the consortium and the pilot partners. The data size is unknown at the moment and not yet determined. Validation results can be useful to project partners, research community, users of the APOLLO platform since the validation results will refer to the "quality" of the data in terms of precision and accuracy. The validation of data is also consequently important for the Apollo services which are exclusively based on EO data.

2. FAIR Data 2.1. Making data findable, including provisions for metadata	N/A. Task 3.6 will produce results of specific statistical analysis for validation of EO data.
2.2 Making data openly accessible	Data used for validation purposes is openly available. The validation results will be presented as a report which will be made available on the APOLLO website; the results of the validation will also (possibly) be publish in peer-reviewed scientific journals. The validation analysis will mainly be performed with Pytesmo 0.6.0 (Python Toolbox for the Evaluation of Soil Moisture Observations) package developed by the Remote Sensing group at TU Wien; Pytesmo is freely available (https://pypi.python.org/pypi/pytesmo) along with the relevant documentation (http://rs.geo.tuwien.ac.at/validation_tool/pytesmo/docs/index.html).
2.3. Making data interoperable	The validation process will follow the current state-of-the-art methodologies. One example for this is the "Global Leaf Area Index Product Validation Good Practices" guideline produced by the Land Product Validation (LPV) subgroup of the Committee on Earth Observation Satellites (CEOS) which will be used as reference for the validation of (LAI) product produce by the APOLLO project.
2.4. Increase data re-use (through clarifying licences)	The Validation Report will be freely available on the APOLLO website. In addition, significant scientific findings from the validation activities will be published in free-access peer-reviewed journals.
3. Allocation of resources	The budget includes funds that will cover any fees associated with publication in open-access peer-reviewed journals.
4. Data security	N/A
5. Ethical aspects	N/A
6. Other	N/A

3.4   WP4 – Development of added value services

3.4.1   Added value service dat

Contributions for AUA Zisis Tsiropoulos, Evangelos Anastasiou and Spyros Fountas

DMP component	Data_WP4_1_Development of added value services
1. Data summary	In this task no data will be collected or generated. As our work package describes, we research literature on available algorithms regarding manipulation of the available data and services development on: • Tillage scheduling • Irrigation scheduling • Crop growth monitoring • Crop yield estimation

2. FAIR Data 2.1. Making data findable, including provisions for metadata	N/A
2.2 Making data openly accessible	N/A
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	N/A
3. Allocation of resources	N/A
4. Data security	The algorithms that will be created/used for the purposes of APOLLO, will be included in the deliverable D4.1 Suite of agricultural services. This deliverable will be confidential among the APOLLO project partners. To ensure confidentiality, the partners, will gain access to the file by authenticating themselves into the website.
5. Ethical aspects	N/A
6. Other	N/A

3.5   WP5 – Platform development and integration

Contributions from DRAXIS: Ioulia Anastasiadou

3.5.1   System architecture and design

DMP component	Data_WP5_1_System_architecture_and_design
1. Data summary	Component descriptions, component dependencies, API descriptions, information flow diagram, internal and external interfaces, hardware requirements and testing procedures. All technical partners were asked to answer a set of questions, based on which further online and face to face discussions took place in order to form the final architecture document. This will be the basis upon which the system will be built.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	It will become both discoverable and accessible to the public once it is delivered to the EU and the consortium decides to do so. A table containing all versions of the document, along with who contributed to each version, what the changes where as well as the date a new version was created will be part of the architecture document.
2.2 Making data openly accessible	The data will be available in D5.1: System architecture and design. The dissemination level of D5.1 is confidential. It will be available through the APOLLO wiki for the members of the consortium.

2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	Could be used as example for software engineers who want to build similar systems.
3. Allocation of resources	N/A
4. Data security	All data will be securely saved in the DRAXIS premises and will be shared with the rest of the partners using the APOLLO wiki.
5. Ethical aspects	There are no ethical aspects related to the described dataset.
6. Other	N/A

3.5.2   Specification of APOLLO service requirements

DMP component	Data_WP5_2_Specification_of_APOLLO_service_requirements
1. Data summary	Component descriptions and dependencies, information flow diagram, user stories describing the system's functionality, non-functional requirements and wireframes are included in this report. This will be the basis of the functional requirements upon which the system will be built.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	It will become both discoverable and accessible to the public once it is delivered to the EU and the consortium decides to do so. A table containing all versions of the document, along with who contributed to each version, what the changes where as well as the date a new version was created will be part of the architecture document.
2.2 Making data openly accessible	The data will be available in D2.2: Specifications of Apollo service requirements. The dissemination level of D2.2 is confidential. It will be available through the APOLLO wiki for the members of the consortium.
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licenses)	N/A
3. Allocation of resources	N/A
4. Data security	All data will be securely saved in the DRAXIS premises and will be shared only with the rest of the partners using the APOLLO wiki.
5. Ethical aspects	There are no ethical aspects related to the described dataset.
6. Other	N/A

3.5.3   APOLLO platform content

DMP component	Data_WP5_3_APOLLO_platform_content
1. Data summary	APOLLO users will generate data via the APOLLO platform both in its web and mobile formats. These data will contain users' personal information, farm information and files containing farm location (*.shp or *.kml) all of which will be useful in order for the APOLLO models to function properly (e.g. the models that produce information about tillage need to have as input the location of the farm and the crop type). The data described above will be saved in the APOLLO central database.
	Detailed log of user actions (login, logout, account creation, visits on specific parts of the app) will be kept in the form of a text file (*.txt). This log will be useful for debugging purposes.
	Reports containing information on user devices (which browsers and mobile phones) as well as number of mobile downloads (taken from play store for android downloads and app store for mac downloads) will be useful for marketing and exploitation purposes, as well as decisions regarding the supported browsers and operating systems.
	*.shp or *.kml files containing farm location information will be produced by the system based on the users' input and will be available for the users to download.
	No existing data will be reused.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Meaningful metadata will be produced as a result of every action (time and date of data creation or data amendments, owners of actions that took place). Metadata will assist the discoverability of the data and related information.
	Only the administrator of the platform will be able to discover all the data generated by the platform.
	The database will not be discoverable to other network machines operating on the same LAN, VLAN with the DB server or other networks. Therefore only users with access to the server (APOLLO technical team members) will be able to discover the database.
2.2 Making data openly accessible	Data will only be available to registered users and administrators. The data produced by the platform are sensitive private data and cannot be shared with others without the user's permission. No open data will be created as part of APOLLO platform functionality.
	The database will only be accessible by the authorized technical team.
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	N/A
3. Allocation of resources	N/A
4. Data security	All data generated by the platform will be saved in the APOLLO database server. Sensitive user data (emails, passwords) will be encrypted using

	strong algorithms. All data will be transferred via SSL connections to ensure secure exchange of information.
	In case of necessary updates, the old data will be overwritten and all actions will be audited in detail and a log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios.
	All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
5. Ethical aspects	The dataset will contain personal data therefore a consent for data provision will be asked to the user prior to data submission. The user will be provided with the option to agree or disagree to the submission of his/her data to the APOLLO platform prior to the data submission. Users who submit personal data will be notified that their data will be confidential.
	Another important ethical issue to be covered is that data will be imported from a non-EU country, Serbia, into the EU, therefore an authorization from the EU or the country where the data will be transferred to may be required. This issue is still under investigation in collaboration with the Serbian partners of APOLLO project UBFC and UPOR.
6. Other	N/A

3.5.4   Communication tool

DMP component	Data_WP5_4_Communication_tool
1. Data summary	APOLLO's communication tool is a virtual scape where APOLLO users exchange ideas and help each other. As such it will produce posts and comments by any user who wishes to participate.
	Detailed log of user actions (postid, post content, owner, date and time) will be kept in the form of a text file.
	No existing data will be reused.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	All registered users will be able to discover posts, comments and other users' usernames. All other data will only be discoverable by the technical team.
	Meaningful metadata will be produced as a result of every action (time and date of data creation or data amendments, owners of actions that took place). Metadata will assist the discoverability of the data and related information.
	The database will not be discoverable to other network machines operating on the same LAN, VLAN with the DB server or other networks. Therefore, only users with access to the server (APOLLO technical team members) will be able to discover the database.

2.2 Making data openly accessible	No data generated from the communication tool will be openly available. As a private tool, the users will have to pay in order to access the communication tool.
	Registered users will be able to access data produced by the communication tool by signing to the platform. The database that will contain all data described above, will only be accessible by the authorized technical team.
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	N/A
3. Allocation of resources	N/A
4. Data security	All data generated by the platform will be saved on the APOLLO database server. Sensitive user data will be encrypted using strong algorithms. All data will be transferred via SSL connections to ensure secure exchange of information.
	In case of necessary updates, the old data will be overwritten and all actions will be audited in detail and a log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios.
	All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
5. Ethical aspects	The dataset will contain user comments and discussions therefore a consent for data provision will be asked to the user prior to data submission. Moreover, no confidential information will be posted to the APOLLO communication tool. Comments and communications between users containing the following shall not be allowed and will be removed by the APOLLO platform administrators. • Comments not related to the functioning and purpose of the APOLLO platform. • Profane language or content. • Content that promotes, fosters or perpetuates discrimination on the basis of race, creed, color, age, religion, gender, marital status, status with regard to public assistance, national origin, physical or mental disability, or sexual orientation. • Sexual content or links to sexual content. • Conduct or encouragement of illegal activity. • Content related to advertising or promotions. • Information that may tend to compromise the safety or security of the public or public systems. • Content that violates a legal ownership interest of any other party.
6. Other	N/A

3.5.5   Maps produced by the agricultural models

DMP component	Data_WP5_5_Maps_produced_by_the_agricultural_models
1. Data summary	One of the main types of servicesof APOLLO is the generation of maps, based on agricultural models, that the enable farmers and consultants increase their efficiency. A *.tif file, bearing geolocation metadata, for each type of map produced by each service (tillage scheduling, irrigation scheduling, crop growth monitoring, yield estimation) will be produced. As an indication for the size; for a farmer owning 405 hectares of land and who has signed up for all the four services we will daily produce 100-150kb of information. Existing data from the soil database (https://soilgrids.org/#/?layer=geonode:taxnwrb_250m) are being used as input to the models.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Registered farmers will be able to discover maps corresponding to their farms. Consultants will be able to discover maps corresponding to their client's farms. Admins will be able to discover all maps that have been produced by the APOLLO platform. Meaningful metadata will be produced as a result of every action (time and date of data creation or data amendments, owners of actions that took place, service that produced the map, crop type of depicted farm, farm owner). Metadata will assist the discoverability of the data and related information.
2.2 Making data openly accessible	Maps produced by the agricultural models will not be openly accessible. Users must sign in in order to access the produced maps. As a private product, users will have to pay in order to access the maps produced. The maps and the metadata will be made available for use by the APOLLO applications through the secure API that we will create. The hard drive that will host all the produced tif files, will only be accessible by the authorized technical team.
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licences)	Maps produced during the course of the project, will be offered to anyone who asks for it. After the end of the project, these data will only be available to users who will buy the product. Paying users will be able to download such data and use it for their own purposes.
3. Allocation of resources	Preserving map tif files for long time will offer the users the opportunity to go back in time and compare current farm conditions with those of the past.
4. Data security	All data generated by the platform will be saved on the APOLLO server. All data will be transferred via SSL connections to ensure secure exchange of information. In case of necessary updates, the old data will be overwritten and all actions will be audited in detail and a log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios. All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and

	cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
5. Ethical aspects	There are no ethical aspects related to the described dataset.
6. Other	N/A

3.5.6   Statistical Reports

DMP component	Data_WP5_6_Statistical_Reports
1. Data summary	Further processing on the models' outputs will lead to the creation of statistical reports bearing useful information for field management. Plots to graphically represent this information will also be produced. Farmers and consultants will use this information to better treat their/their clients' farms.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Registered farmers will be able to discover reports and plots corresponding to their farms. Consultants will be able to discover reports and plots corresponding to their client's farms. Admins will be able to discover all reports and plots that have been produced by the APOLLO platform. Farmers will be able to download the reports should they want it. Metadata containing the dates the reports/plots where created, the user for whom they were produced and number of downloads will also be produced.
2.2 Making data openly accessible	Reports and plots will not be openly accessible. Users must sign in in order to access the produced reports and plots. As a private product, users will have to pay in order to access the maps produced. The reports and the metadata will be made available for use by the APOLLO applications through the secure API that we will create. The database that will host the above data and any related metadata will not be discoverable to other network machines operating on the same LAN, VLAN with the DB server or other networks. Therefore only users with access to the server (APOLLO technical team members) will be able to discover the database.
2.3. Making data interoperable	Statistical reports will be downloadable in xml format so that it is easy to be further used for other purposes.
2.4. Increase data re-use (through clarifying licences)	Reports produced during the course of the project, will be offered to anyone who asks for it. After the end of the project, these data will only be available to users who will buy the product. Paying users will be able to download such data and use it for their own purposes.
3. Allocation of resources	Preserving these data for long time will offer the users the opportunity to go back in time and compare current farm conditions with those of the past. Indicative cost per month is $0.024.
4. Data security	All data generated by the platform will be saved on the APOLLO server. All data will be transferred via SSL connections to ensure secure exchange of information. In case of necessary updates, the old data will be overwritten and all actions will be audited in detail and a log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios.

APOLLO

	All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
5. Ethical aspects	To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables.
6. Other	N/A

3.6 WP6 – Pilot operation and evaluation

Contributions for UBFCE: Dr. Dragutin Protic

3.6.1 Pilot in situ data

DMP component	Data_WP6_2_Pilot_in_situ_data
1. Data summary	In situ data will be collected during WP6 implementation in order to improve and further validate models of estimation of crop biophysical parameters and soil moisture. The data will include (LAI, biomass, yield, height of crops, phenology phase, pest and disease presence and soil moisture). The data will include coordinates, date of collection, measured parameter, crop type and will be structured by the APOLLO pilots. The data will be collected in the fields by pilot partners (AgriSat, Association of farmers of the Municipality of Ruma, UPOR and Agriculture cooperative of Pella, ACP). The data will potentially be useful for other scientific and innovation projects and research initiatives which are aimed at creating models for estimation of crop characteristics from EO data. The data could be also used for validating the existing models. The data will be available in *.xls and *.csv formats.
2. FAIR Data 2.1. Making data findable, including provisions for metadata	The data will be freely available through APOLLO website along with the related metadata. For each dataset, related metadata will describe data structure and the methodology used to collect the data. The data will be named: APOLLO_INSITU_LAI, APOLLO_INSITU_BIOMASS, APOLLO_INSITU_SOILMOISTURE, etc.
2.2 Making data openly accessible	All APOLLO in situ data will be openly available. The data will be downloadable through APOLLO website.
2.3. Making data interoperable	The data will be provided in commonly used physical units, BBHC scale of phenological development stages of a plant and commonly used descriptions of pest and diseases presence.
2.4. Increase data re-use (through clarifying licences)	The data will be licensed by Creative Commons Attribution CC BY license. This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. The data will be available for re-use as soon as the quality is approved. No time limits for re-use of the data will be imposed.
3. Allocation of resources	There are no significant costs for making the data FAIR. The data will be accessible through APOLLO website as long as it is maintained.
4. Data security	N/A

5. Ethical aspects	No personal data will be distributed within the described datasets.
6. Other	No.

3.6.2   Evaluation of APOLLO platform services

Contributions from AUA from Zisis Tsiropoulos, Evangelos Anastasiou and Spyros Fountas

DMP component	Data_WP6_3 Evaluation of APOLLO platform services_data

1. Data summary	Purpose of the data collection is to evaluate the accuracy of APOLLO platform's added value services. The accuracy of the APOLLO platform's services will be compared with ground truth data and each service will be validated based on its own validation methodology.
	
	Several sensor types will be used for the collection process, each of whom exports a specific file type. Ground truth data will be also included from laboratory analysis results.
	
	Finally, the data will be available in files such as:
	
	   • CSV
	
	   • XLSX
	
	   • GEOTIFF
	
	   • JPEG
	
	These data files will be useful mainly to the APOLLO project partners for evaluation purposes.

2. FAIR Data 2.1. Making data findable, including provisions for metadata	N/A. Task 6.3 will produce results of specific statistical analysis for validation of APOLLO's services.
2.2 Making data openly accessible	The validation results will be presented as reports which will be made publicly available on the APOLLO website; the results of the validation will be possibly published in peer-reviewed scientific journals. Public access of the raw data files will be determined later in the projects life according to the business model which will be decided. The project partners of APOLLO will have full access over the raw data by authenticating themselves into the platform.
2.3. Making data interoperable	N/A. It will be decided only if we make the raw data public.
2.4. Increase data re-use (through clarifying licences)	N/A

3. Allocation of resources	N/A
4. Data security	The private data (raw data) will be placed in a password protected area on the website, so only the project members can have access to them. The reports will be public, so no security measurements have to be taken.
5. Ethical aspects	N/A
6. Other	N/A

3.7 WP7 – Exploitation and communication

Contributions from EVF: Lefteris Mamais, Dimitris Papadakis.

Regarding WP7 – Exploitation and communication, in the context of the 1st DMP, no structured dataset will be generated since the data use, re-use concept does not fully apply in the purpose of WP7. The presented Market dataset is meant to be confidential and not be made publicly available since it will feed the APOLLO business model.

3.7.1 Market Data

DMP component	Data_WP7_Market Data
1. Data summary	Database of competing solutions that includes the following details: service description, technologies used, specific agricultural applications supported, targeted customers, marketing and business aspects (e.g. pricing, service provision scheme, etc.).
2. FAIR Data 2.1. Making data findable, including provisions for metadata	Given that (a) this data is collected through publicly open sources (available online) (b) the analysis and recommendations that feed into APOLLO's business plan is considered confidential, there will be no external interface to the database.
2.2 Making data openly accessible	See 2.1 above.
2.3. Making data interoperable	N/A
2.4. Increase data re-use (through clarifying licenses)	N/A
3. Allocation of resources	This dataset does not require specific allocation of resources for its maintenance.
4. Data security	The database is stored using industry-standard cloud storage tools.
5. Ethical aspects	All data have been collected through publicly accessible on-line sources. The validation data that have been collected through interviews are not a structured dataset and will not be made publicly available.
6. Other	N/A

4 Conclusion

The deliverable covered all the aspects related to data management within the H2020 project APOLLO. Using the latest guidelines and template but most importantly the collaboration among the project partners the 1st version of the APOLLO DMP is submitted six months after project start. The collaborative work identified issues and missing information related to data management that can be further clarified during the upcoming plenary meeting and during the next six months of the project till the submission of the 2nd DMP. Last, a summary table in the next chapter attempts to provide an overview of data management in APOLLO project.

5 Summary table

WP / Task	Responsible partner	Dataset name	File types	Findable	Accessible	Interoperable	Reusable	Size	Security	Ethics
1 / 1.2	DRAXIS	WP1_1_Advisory_board_data	*.xlsx , *.mp3, *.doc, *.pdf	Not yet decided	Within deliverables in project's website	N/A	In pdf format in website	1 GB	Kept in DRAXIS servers	Yes
2 / 2.1	AUA	WP2_1 User_requirements_data	*.xlsx ,*.pdf	No metadata	In project website	Yes	Public Domain Dedication and License	. xlxs file ~70KB	No	N/A
2 / 2.3	AgriSat	WP2_2_Co-creation_Data	Xlxs, audio, video	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3 / 3.1	Starlab	WP3_1_Satellite_data	GeoTIFF, ESA Safe format, Jpeg 2000	Metadata in raster files	Only for project partners	Yes	Yes	~260GB	Rasdaman server security	N/A
3 / 3.2	Starlab	WP3_2_Soil_moisture_data	GeoTIFF	Yes	Only for pilot farmers	Yes	Yes, for users under subscription	~1.2 GB	Rasdaman server security	N/A
3 / 3.3	UBFCE	WP3_3_Local_meteorological_data	GeoTIFF, gml csv, HDF-EOS format	Metadata available	Only for pilot farmers	Yes	Yes for subscribers	~110GB	Rasdaman server security	N/A

WP / Task	Responsible partner	Dataset name	File types	Findable	Accessible	Interoperable	Reusable	Size	Security	Ethics
3 / 3.4	UBFCE	WP3_4_Crop_condition_data	GeoTIFF, JPEG2000	Metadata available	Only for pilot farmers and project partners	Yes	Yes for subscribers	~650GB	Rasdaman server security	N/A
3 / 3.5	UBFCE	WP3_5_Biomass_data	GeoTIFF	Metadata available	Only for pilot farmers and project partners	Yes	Yes for service subscribers	~54GB	Rasdaman server security	N/A
3 / 3.6	TUW	WP3_6_Validation_results	Pdf	N/A	In project website, in dedicated repository	Yes	Report in project website	N/A	N/A	N/A
4 / 4.1	AUA	WP4_1_Development of added value services	Pdf	Report deposited in project website	N/A	N/A	From project website	N/A	N/A	N/A
5	DRAXIS	WP5_1_System_architecture_and_design	Pdf	Confidential, only from project partners	Only from project partners	N/A	From software engineers	N/A	DRAXIS servers security	N/A
5	DRAXIS	WP5_2_Specification_of_APOLLO_service_requirements	Pdf	Confidential, only from project partners	Only from project partners	N/A	N/A	N/A	DRAXIS servers security	N/A

APOLLO

WP / Task	Responsible partner	Dataset name	File types	Findable	Accessible	Interoperable	Reusable	Size	Security	Ethics
5	DRAXIS	WP5_3_APOLLO_platform_content	txt, kml, shp	Metadata available	Only available to registered users and administrators	N/A	N/A	N/A	APOLLO database server security	Yes
5	DRAXIS	WP5_4_Communication_tool	txt	From registered users and technical team members	Only available to registered users and administrators	N/A	N/A	N/A	APOLLO database server security	Yes
5	DRAXIS	WP5_5_Maps_produced_by_the_agricultural_models	GeoTIFF	From registered users and technical team members	Not openly accessible. Paid service	N/A	Upon request during the project, only to subscribes after the end	100-150kb daily	APOLLO server security	No
5	DRAXIS	WP5_6_Statistical_Reports	xml	From registered users and technical team members	Not openly accessible. Paid service	Yes, xml format	Upon request during the project	-	APOLLO server security	Covered by ethics section of DoA

WP / Task	Responsible partner	Dataset name	File types	Findable	Accessible	Interoperable	Reusable	Size	Security	Ethics
6	UBFCE	WP6_2_Pilot_in_situ_data	xlxs, csv	Metadata available, deposited in project website	Openly accessible	Yes	Creative Commons Attribution CC BY license.	-	N/A	No
6	AUA	WP6_3 Evaluation of APOLLO platform services_data	xlxs, csv, jpeg, GeoTIFF	N/A	In reports deposited in project website	N/A	N/A	-	Password protected in project website	N/A
7	EVF	WP7_Market Data	-	N/A	N/A	N/A	N/A	-	Stored in cloud	No

This project is co-funded by the European Union