D1.7: 2nd Data Management Plan

WP1 – Project Management

Authors: Polimachi Simeonidou, Ioanna Pavlou

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687412.

Disclaimer

Copyright message

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	Research and innovation Action		
Start Date	1st May 2016	**Duration**	34 months
Project URL	-		
Document URL	-		
EU Project Officer	Kristiina URPALAINEN-MENON		
Project Coordinator	Polimachi Simeonidou		
Deliverable	D1.7: 2nd Data Management Plan		
Work Package	WP1 – Project Management		
Date of Delivery	**Contractual** M12	**Actual**	M12
Nature	R – Report	**Dissemination Level**	PU - Public
Lead Beneficiary	DRAXIS		
Lead Author	Polimachi Simeonidou	**Email**	msimeonidou@draxis.gr
	DRAXIS	**Phone**	+30 2310 274566
Other authors	Ioanna Pavlou (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	01/04/2017	Draft	Draft for review	DRAXIS
2.0	21/04/2017	Draft	Review Feedback	Starlab
3.0	24/04/2017	Final	Final version	DRAXIS

This project is co-funded by the European Union

Table of Contents

This project is co-funded by the European Union

Executive summary

This deliverable represents an update of the first version of APOLLO Data Management Plan (D1.6). The deliverable outlines the updates on the APOLLO project data management life cycle for all datasets that are collected, processed or generated and describes how this data will be handled during the project.

The deliverable is structured in the following chapters:

- Chapter 1 includes an introduction to the deliverable.
- Chapter 2 includes the updated description of the datasets along with the documented changes and additional information.

This project is co-funded by the European Union

1 Introduction

The current deliverable D1.7 "2nd Data Management Plan" represents the second version of the APOLLO Data Management Plan (DMP) delivered in M6. This deliverable documents and explains the types of information it contains and the data collected during the project in detail. The DMP will be updated one more time and documented in deliverable D1.8 (M34).

The "Guidelines on Data Management in Horizon 2020"[1] released by the European Commission Directorate – General for Research & Innovation mentions:

"A Data Management Plan should be updated – if appropriate – during the project lifetime (in the form of deliverables). Please create and submit a new version of the DMP whenever significant changes arise in your project such as:

- *new datasets*
- *changes in consortium policies*
- *external factors.*

Deadline of updates:

- *as a minimum, update the DMP in time to be evaluated as part of the mid-term and final project reviews for projects that have planned reviews*
- *DMP updates needed at least at the end of projects, for projects where reviews are not planned*
- *and at other moments if it is defined by project consortia."*

Although the DMP is being developed by DRAXIS, its implementation involves all project partners' contribution. The next version of the DMP, to be published at M34, will describe more in detail the practical data management procedures implemented by the APOLLO project.

The Work Packages that have not occurred any changes are not included in this deliverable.

[1] European Commission, (26 July 2016*), Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0

2 DMP Components in APOLLO

2.1 DMP Components in WP3 – Earth Observation data products (UBFCE)

2.1.1 Satellite data and Pre-processing

DMP Component	Issues to be addressed
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: • 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 file 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[2] • Sentinel-1 CGR calibrated and geo referenced (GeoTiff format) Size of the input data: • 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 output data: • Sentinel-2 L2A BOA: 1200MB *73days= 87600MB~85GB; Test areas - Serbia=1 tile, Greece=1 tile, Spain=4 (not whole) tiles • 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 algorithms, including soil moisture, Crop condition, and Biomass.

[2] http://step.esa.int/main/third-party-plugins-2/sen2cor/

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 processing requested. Naming conventions for the data will be: • S2_BOA_[ID]_[YYYYMMDD].tiff • S1_CGR_[ID]_[YYYYMMDD].tiff
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.
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.
Increase data re-use	The data will be accessible for use to all APOLLO project partners through REST API from the Rasdaman database.
Allocation of resources	N/A
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.
Ethical aspects	N/A
Other issues	N/A

2.1.2 Soil Moisture Data

DMP Component	Issues to be addressed
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.

This project is co-funded by the European Union

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 processing requested. Naming conventions for the data will be: • sm_[ID]_[YYYYMMDD].tiff
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.
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.
Increase data re-use	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).
Allocation of resources	N/A
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.
Ethical aspects	N/A
Other issues	N/A

2.1.3 Data validation

DMP Component	Issues to be addressed

Data Summary	The purpose of the data sets used in this task is the validation 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. The EO data sets produced within APOLLO are: soil moisture, vegetation indices, leaf area index (LAI) and biomass, and will all be subject to validation. The data sets against which validation will be performed include: modelled data (e.g. GLDAS-Noah land surface model), other EO-based products (e.g. soil moisture from produced within H SAF) and in-situ data (collected from the pilot areas and, where available, soil moisture from the International Soil Moisture Network (ISMN)). 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.
Making data findable, including provisions for metadata	Within Task 3.6 results are presented as specific statistical analysis for validation of EO data (e.g. triple collocation, mean error, mean absolute error, bias corrected mean, absolute error).
Making data openly accessible	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 published in peer-reviewed scientific journals. There are two types of data that will be used for validation, both openly available: external data sets (e.g. modelled data) and internally produced data sets which is the in-situ data collected in the pilot study areas. The external datasets which will be used include: data produced with the GLDAS-Noah land surface model, H SAF soil moisture, and depending on availability in-situ data from the International Soil Moisture Network (ISMN). A detailed description of the data based on in-situ measurements collected in the three pilot areas is given in Sub-section 3.6.1 of this document. The validation analysis for the soil moisture product 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[3] along with the relevant documentation[4].
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.

[3] https://pypi.python.org/pypi/pytesmo

[4] http://rs.geo.tuwien.ac.at/validation_tool/pytesmo/docs/index.html

Increase data re-use	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.
Allocation of resources	The budget includes funds that will cover any fees associated with publication of any relevant validation results in open-access peer-reviewed journals.
Data security	N/A
Ethical aspects	N/A
Other issues	N/A

2.2 DMP Components in WP4 – Development of added value services (AUA)

DMP Component	Issues to be addressed
Data Summary	We don't collect or generate any kind of data. 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
Making data findable, including provisions for metadata	N/A
Making data openly accessible	N/A
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A
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.
Ethical aspects	N/A
Other issues	N/A

2.3 DMP Components in WP5 – Platform development and integration (DRAXIS)

2.3.1 Website content

DMP Component	Issues to be addressed

Data Summary	APOLLO users will generate data via the platform. These data contain users' personal information, farm information and shapefiles containing farm location 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. 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. Shapefiles containing farm location information will be produced by the system based on the users' input and will be available for the users (farm owners and consultants) to download. No existing data will be reused.
Making data findable, including provisions for metadata	Meaningful metadata are 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 app will be able to discover all the data generated by the platform. The database is not 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) are able to discover the database.
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. The database is only accessible by the authorized technical team.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A

This project is co-funded by the European Union

Data security	All data generated by the platform are saved on the APOLLO database server. Sensitive user data such as passwords are encrypted using strong algorithms (bcrypt). All data are 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.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

2.3.2 Communication tool

DMP Component	Issues to be addressed
Data Summary	APOLLO's communication tool is a place for APOLLO users to 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.
Making data findable, including provisions for metadata	All registered users will be able to discover posts, comments and other users' username. 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.
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.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A

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.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

2.3.3 Maps produced by the agricultural models

DMP Component	Issues to be addressed
Data Summary	One of the main offerings of APOLLO is the generation of maps, based on agricultural models, that can help 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[5] are being used as input to our models.
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.

[5] https://soilgrids.org/#/?layer=geonode:taxnwrb_250m

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 have created. The hard drive that will host all the produced .tif files, will only be accessible by the authorized technical team.
Making data interoperable	N/A
Increase data re-use	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.
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.
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.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

2.3.4 Statistical Reports

DMP Component	Issues to be addressed
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.

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.
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.
Making data interoperable	Statistical reports will be downloadable in xml format so that it is easy to be further used for other purposes.
Increase data re-use	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.
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.
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.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

2.3.5 Mockups

DMP Component	Issues to be addressed
Data Summary	Mock ups have been created by UI/UX experts in order to help the development team design an intuitive and easy to use application. The type of the files produced is psd. The expected size of each document is 1,5 KB.
Making data findable, including provisions for metadata	The mock ups might become both discoverable and accessible to the public once it is delivered to the EU and the consortium decides to do so. In any case given that the mock ups depict the look and feel of the final application, once the APOLLO product is out on the market, everyone who uses it will be able to view them.
Making data openly accessible	The mock ups will not be openly available before the completion of the APOLLO platform and if the consortium decides to do so.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A
Data security	All data will be securely saved in the DRAXIS premises.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

2.4 DMP Components in WP6 – Pilot operation and evaluation (UBFCE)

2.4.1 Pilot in situ data

DMP Component	Issues to be addressed
Data Summary	In situ data will be collected during WP6 in order to improve and further validate models of estimation of biophysical parameters of crops and soil moisture. The data will include (LAI, biomass, chlorophyll content, 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 and Agriculture cooperative of Pella) for all the crops involved in APOLLO project. 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.

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.
Making data openly accessible	All APOLLO in situ data will be openly available. The data will be downloadable through APOLLO website. The in-situ soil moisture data collected within the project will be made openly available for scientific purposes through the International Soil Moisture Network (ISMN). The ISMN is an initiative by the Global Energy and Water Cycle Experiment (GEWEX) and the European Space Agency (ESA) to promote scientific studies on calibration and validation of satellite based and modeled soil moisture products.
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.
Increase data re-use	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.
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. The in-situ soil moisture measurements collected throughout the project are hosted free of charge on the ISMN network.
Data security	N/A
Ethical aspects	No personal data will be distributed within the data.
Other issues	N/A

2.4.2 Evaluation of APOLLO platform services data

DMP Component	Issues to be addressed

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.
Making data findable, including provisions for metadata	N/A. Task 6.3 will produce results of specific statistical analysis for validation and evaluation of APOLLO's services.
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.
Making data interoperable	N/A. It will be decided only if we make the raw data public.
Increase data re-use	N/A
Allocation of resources	N/A
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.
Ethical aspects	N/A
Other issues	N/A

2.5 DMP Components in WP7 – Exploitation and communication (EVF)

DMP Component	Issues to be addressed
Data Summary	Analysis of the market of competing solutions for the purpose of positioning the APOLLO solution within this framework. This database 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.). The format is Microsoft Excel, and the data is collected through publicly open sources (available online). Estimated size: >1 Mb.

Making data findable, including provisions for metadata	Given the open character of the data and the fact that the analysis and recommendations that feed into APOLLO's business plan are considered confidential, there will be no external interface to the database.
Making data openly accessible	Both the deliverables and the excel files will be openly accessible by anyone through the APOLLO website.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	This dataset does not require specific allocation of resources for its maintenance.
Data security	The database is stored using industry-standard cloud storage tools.
Ethical aspects	All data have been collected through publicly accessible on-line sources. The validation data collected through interviews are not a structured dataset and will not be made publicly available.
Other issues	N/A

3 Conclusion

This second deliverable reflects the updated procedures to be implemented by APOLLO project to efficiently manage the data it will produce. In particular, the 2nd DMP anticipates the data management strategy regarding the collection, management, sharing, archiving and preservation of data. The DMP is not a fixed document but it will be updated one more during project lifespan (M34).

This project is co-funded by the European Union

Abbreviations

API	Application Programming Interface
BBHC	Biologische Bundesastalt, Bundessortenamt und Chemische Industrie
CEOS	Committee on Earth Observation Satellites
DMP	Data Management Plan
EO	Earth Observation
ESA	European Space Agency
EU	European Union
GEWEX	Global Energy and Water Cycle Experiment
IP	Internet Provider
ISMN	International Soil Moisture Network
LAI	Leaf Area Index
LAN	Local Area Network
LPV	Land Product Validation
REST	Representational State Transfer
SAR	Synthetic Aperture Radar
SQL	Structured Query Language
SSL	Secure Sockets Layer
UI	User Interface
UX	User Experience
VLAN	Virtual LAN
WMS	Web Mapping Services

This project is co-funded by the European Union