Del 8.2
Data Management Plan & Support Pack
Document Summary

Deliverable Title: Data Management Plan & Support Pack

Version: 1.0

Deliverable Lead: Agricultural University of Athens (AUA)

Related Work package: WP8

Author(s): Nikos Mylonas

Contributor(s): All partners

Communication level:
- PU Public

Project Number: 773718

Grant Agreement Number: 773718

Programme: OPTIMA

Start date of Project: September, 2018

Duration: 40 months

Project coordinator: Agricultural University of Athens (AUA)

Abstract

This report will specify how data will be collected, processed, monitored, classified, and disseminated during the project’s lifetime. It will also include a Support Pack with guidelines during their research activities, which software tools and services should be used, and how they can align the project requirements with their institutions’ standard practices and systems.
Table of contents

1. Introduction ........................................................................................................................................ 4
   1.1 Data management in Horizon 2020 program .................................................................................. 5
2. Methodology ......................................................................................................................................... 7
   2.1 Data management strategy in OPTIMA ....................................................................................... 7
   2.2 DMP template ............................................................................................................................... 8
   2.3 Data summary ............................................................................................................................... 8
   2.4 FAIR data ....................................................................................................................................... 9
       2.4.1 Making data findable, including provisions for metadata ..................................................... 9
       2.4.2 Making data openly accessible ............................................................................................. 9
       2.4.3 Making data interoperable .................................................................................................. 10
       2.4.4 Increase data re-use (through clarifying licenses) ............................................................... 11
   2.5 Allocation of resources ............................................................................................................... 11
   2.6 Data security .................................................................................................................................. 12
   2.7 Ethical aspects .............................................................................................................................. 12
   2.8 Other issues ................................................................................................................................... 12
3. Mapping the OPTIMA data management landscape ........................................................................ 13
4. OPTIMA Support Package ............................................................................................................... 25
5. Conclusions .......................................................................................................................................... 26
   .......................................................................................................................................................... 27
1. Introduction

This document provides the plan for the management of research outcomes (and more specifically, the research publications and datasets) that will be produced during the OPTIMA project lifetime, as well as those that will be collected from the OPTIMA partners for the respective Work Package activities. It aims to ensure that the research activities of the project are compliant with the H2020 Open Access policy and the recommendations of the Open Research Data pilot. In this context, the project’s Data Management Plan (DMP) described in this document outlines how research data and metadata will be collected, processed or generated within the project; what methodology and standards will be adopted; whether and how this data will be shared and/or made open; and how this data will be curated and preserved during and after the project. 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 OPTIMA project duration. In this respect, the 2nd version (D8.2) will be submitted on Month 20, while the 3rd and final version (is due on Month 40 at the end of the project. As required, the upcoming versions of the OPTIMA DMP will have a clear version number and include a timetable for any occurring data updates.

Free and open access to scientific publications and research data is nowadays critically important for researchers, in order to base their work on them and make the next step in their research fields, instead of having to duplicate existing experiments and research work. However, scientific publications are usually accessible only through commercial publishers and accompanied by an access fee, which needs to be paid either by the researcher’s institutional library (as an annual subscription fee or on a request basis) or by the researcher himself (in case the institutional library does not have an agreement with the specific publisher). At the same time, research data are not always accessible or at least easily discoverable, as data publishing is not a common practice yet even for institutional repositories. As a result, such data remain stored in offline locations, such as the hard disks and other storage solutions used by the researchers. This issue is not only due to the fact that researchers are not aware of common practices or specific solutions available for the storage and preservation of research data, but also due to the (usually) huge volume of research data which renders commercial data sharing solutions often inappropriate for the specific purpose.
This situation was noticed by the European Commission (EC)\(^1\) and it was decided that actions should be taken for ensuring that at least research publications and relevant datasets that have been funded through programmes of the EC have to be publicly available to all stakeholders. The first steps were taken in the context of the Open Access Pilot of the FP7 funding programme, where the design and implementation of an Open Access Plan by projects funded through the FP7 programme was optional, followed by the Horizon 2020 programme in which the Open Access and Data Management Plan is a mandatory part of the proposals.

1.1 Data management in Horizon 2020 program

In the context of the Horizon 2020 programme, the European Commission published a document titled “Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020”\(^2\). The document clearly describes the need that led to the mandate for open access to scientific publications, research data and their associated metadata that have been produced under the Horizon 2020 programme. At the same time, the document states the European Commission’s view on the important aspect of data re-use: “information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full”.

According to the latest Guidelines on FAIR Data Management in Horizon 2020 released by the EC Directorate-General for Research & Innovation\(^3\) on the 30\(^{th}\) of July 2016 “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”\(^4\) originally broadcasted on 7 July, 2016 and currently available online.


FAIR data management is part of the Open Research Data (ORD) Pilot\(^5\) 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.

\(^5\) [https://www.openaire.eu/what-is-the-open-research-data-pilot](https://www.openaire.eu/what-is-the-open-research-data-pilot)
2. Methodology

The first step towards the implementation of the data management plan for the OPTIMA project is the identification and analysis of the collected and generated data, referring to the data that needs to be covered by the specific data management plan. A data analysis needs to take place, focusing on the data types and formats, as well as the existing licensing options used, in order to allow the data management plan meet any specific requirement that exists due to the nature or the license applied on data.

As the OPTIMA project is in its early stage while compiling the Data Management Plan (DMP), this report will not provide deep elaboration of the entire data spectrum to be handled, instead establishing the methodology of the data management and the foreseen data activities that have been already identified by the project partners. Updated versions will follow later that will include crucial changes about important aspects such as access, exploitation and preservation.

2.1 Data management strategy in OPTIMA

An approach tailored to the specific OPTIMA requirements was followed, in order to identify the data type, format, access, storage etc. Intention is to map the landscape of data in the specific context of the OPTIMA project and to obtain a better understanding on the framework in which the DMP would function. Hence, the initial analysis was focused on the WP level (Table 1) and all the data related requirements.

<table>
<thead>
<tr>
<th>WP No</th>
<th>Work Package Title</th>
<th>Lead Participant</th>
<th>Start Month</th>
<th>End month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Community Building and User Requirement</td>
<td>UNITO</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>Advanced methods for early detection methods of plant diseases</td>
<td>WR</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>Combined synthetic and biological PPPs application and assessment of host and PPP resistance mechanisms</td>
<td>AUA</td>
<td>1</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 1: List of Work Packages
For each individual WP, the guidelines that constitute the DMP and are presented below, are respected to clearly identify the specific data characteristics. Partners from every WP have provided their inputs regarding the expected data outputs, while strictly aligning with the DMP prerequisites for the correct data characterization. Essential to note that at the early stage of the project, data management aspects could not be accurately described within a forty-month projection, but define the framework and data organization aspects.

### 2.2 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 OPTIMA reflects to these sections.

### 2.3 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.
2.4 FAIR data

2.4.1 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 OPTIMA 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_Questionnaire_answers_Spain. 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. As far as file version numbering is concerned, a clear demarcation will be followed at the end of the file name (<file name_v3>).

The use of a standard identification mechanism in for the datasets of OPTIMA will be decided by the project consortium in collaboration with the project coordinator. 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\(^6\) will be considered.

2.4.2 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 possible that some beneficiaries have decided to keep their data closed. For example, during the implementation of pilots, participants or beneficiaries may be unwilling to disclose sensitive operational information regarding cultivation methods or environmental impacts of their activities. The DMP will effectively handle those issues, including the option for Non-Disclosure Agreements to be signed between OPTIMA consortium and actors so as to ensure that all information given will be used only for the project purposes and no sensitive data will be made available to third parties.

\(^6\) https://doi.org
In the context of OPTIMA, the following options for open repositories of data, metadata, documentation or code will be considered: (a) The Registry of Research Data Repositories\(^7\), (b) Zenodo\(^8\), (c) OpenAIRE\(^9\). Finally, the project website will be also used as a repository for project deliverables, results or product demonstrators and trial versions, as well as a widely used commercial repository (e.g. Dropbox) that will be used as the storage destination for all the necessary files, that are shared within the consortium on a daily basis. In the latter case, access will be restricted just for the OPTIMA partners.

Specific care will accompany the access policy of the personal data, derived from the end-user questionnaires. Personal data will be accessible only from the IP’s of each partner premises, using different security credentials and server-side encryption for each staff member. The data will be handled with appropriate confidentiality and technical security, as required by law in the individual countries and EU laws and recommendations. All activities will be carried out ensuring the ethical principles in accordance with the GDPR 2016/679\(^10\) of the European Parliament.

### 2.4.3 Making data interoperable

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

Data will be interoperable, adhering to standards for data annotation, data exchange, compliant with available software applications, and allowing re-combinations with different datasets from different origins. Standards like Dublin Core\(^11\) and ISO/IEC 11179 Metadata Registry\(^12\) which addresses issues in the metadata and data modelling space, will be considered. Even though the project is in its early stage, based on the Grant Agreement and

---

\(^7\) [http://www.re3data.org/](http://www.re3data.org/)
\(^8\) [https://zenodo.org/](https://zenodo.org/)
\(^9\) [https://www.openaire.eu/](https://www.openaire.eu/)
\(^12\) [https://www.iso.org/standard/61932.html](https://www.iso.org/standard/61932.html)
the partners’ feedback, OPTIMA output data will come in standard formats (.shp, .docx, .xls), allowing easy interpretation and adaptability from common software.

In the context of the OPTIMA project for data interoperability among different WPs, there is such need to allow connectivity and data exchange. For example, the data streaming of the field information (e.g., disease information) towards the DSS portal will be facilitated via APIs, allowing only authorized users (farmers, operators, and researchers) to carry out this task.

### 2.4.4 Increase data re-use (through clarifying licenses)

This section describes the licenses, if any, under which data will be re-used in OPTIMA. 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.

The collected anonymized data in the project derived from questionnaires will be available to third parties in contexts such as scientific scrutiny and peer review. Private data (e.g., name) will be restricted only to consortium members, by providing access credentials to the secure repository, which uses encryption content tools. Deliverables marked as publicly available will be uploaded to the OPTIMA website and no specific licensing is required for the users to access these files.

Information that are described as confidential will be identified during the course of the project, and the licensing policy to be followed will be established through non-disclosure agreements.

### 2.5 Allocation of resources

FAIR data management in OPTIMA project is under WP8 – Project management lead by the project coordinator Agricultural University of Athens. Within the project budget, a specific amount of person months has been dedicated for that activity. Project management team (project coordinator and project manager) will be responsible for the continuous monitoring of all data-related activities. Furthermore, they will supervise the smooth adaptation of the data management plan guidelines and practices inside the project. Their observations and the feedback received from the partners will be the product for the upcoming DMP updates.
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.

2.6 Data security

Data security is of major importance in the OPTIMA project. Special attention will be given to the security of personal 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 and appropriate European and Internet security standards from ISO, ITU, W3C, IETF and ETSI.

2.7 Ethical aspects

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

- Personal data collection during the questionnaires distribution with the participation of end-users,
- Ethical approvals,
- Ownership of field data,
- User testimonies and evaluation of IPM from end-users demonstrated in OPTIMA website,
- Procedures for data collection, storage, protection, retention and destruction,
- Application of the right to opt out and to erasure in personal data collection,
- Procedures regarding the recruitment of users during the survey phase (WP1) and evaluation groups (WP5).

Details related to these issues are provided in the Ethics requirements chapters of the OPTIMA project.

2.8 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.
3. Mapping the OPTIMA data management landscape

As described in Section 2.1, a first assessment of the project’s data management characteristics will be conducted, in order to better understand the status quo. For the first version of the DMP, the elaboration will reach only the WP level of the project, without diving deeper in each individual Task. Later versions of this living document will ensure the full coverage of all the data management requirements.

WP1 - Community Building and User Requirement analysis

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP1 - Community Building and User Requirement analysis</th>
</tr>
</thead>
</table>
| 1. Data summary | Objectives and purpose of data creation:  
- set up protocols for assessing end-users requirements and create brochures for explaining the features of OPTIMA IPM system and questionnaires for evaluating the proposed IPM approach;  
- identify, analyze and prioritize the end-users requirements in the pilot countries; This requires the creation of data derived from the analysis of the requirements.  
- provide users participation in the design (improvement) of the IPM system based on a co-creation concept; Datasets to describe the user advices will be created.  
- assess the quality of the proposed (developed) IPM system by the end-users. Datasets to document the evaluation outcomes will be created.  
Data format: (*.xlsx), (*.docx), (*.pdf)  
Excel files for collecting and processing the questionnaire answers. Word and .pdf files for documenting the results of each task and creating the questionnaires and brochures.  
No significant storage requirements, as the files are expected to be small (≤ 300 MB each) |
| 2. FAIR Data | Considering the data nature, no need for metadata is required. Anonymized questionnaire outputs along with community’s guidelines will be available on the website. For the questionnaire naming conventions while uploading the data either in the OPTIMA repository or in the website, a clear reference of the origin country is required:  
e.g WP1_1_<Questionnaire name>_country.<extension> |
| 2.1 Making data findable, including provisions for metadata |  |
| 2.2 Making data openly accessible | Private data that can be attributed to survey participants will not be openly accessible, but only designated partners will be granted |
access through the OPTIMA repository. Access will be given via specific credentials from the data manager.

Interesting research outputs from the WP1 that do not conflict with protection of private data and user testimonies will be accessible in the OPTIMA website. Only web browser and Internet access are needed to access the data.

### 2.3 Making data interoperable

The deliverable files will be “as is” in a .pdf format only to be used for reference. On the other hand, the questionnaire data accumulation files 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). Hence, several software will be able to understand and read this data (e.g Adobe Reader, Excel).

### 2.4 Increase data re-use (through clarifying licenses)

Deliverables 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 WP1 before upload and since no one else but WP1 can import them into the website, no specific requirements for data insurance is needed.

On the contrary, questionnaires will remain available for a specific period (60 months). After extracting the useful information for the analysis, retention of such private information is considered unnecessary.

### 3. Allocation of resources

N/A

### 4. Data security

As for the private data security, the OPTIMA repository and the encryption tools guarantee that access is limited to OPTIMA partners, while securing the repository from external threats.

### 5. Ethical aspects

A consent form has been designed for the collection, retention and exploitation rights of private data derived from end-users’ questionnaires.

### 6. Other

N/A

---

**WP2 - Advanced methods for prediction and early detection of plant diseases**

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP2 - Advanced methods for prediction and early detection of plant diseases</th>
</tr>
</thead>
</table>
1. Data summary

<table>
<thead>
<tr>
<th>Objectives and purpose of data creation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- develop a Decision Support System (DSS) for disease control scheduling; This will yield to advisory data creation (prescription maps, meteorological info)</td>
</tr>
<tr>
<td>- develop advanced detection systems for in-field localization and monitoring of the selected diseases in the use-case crops; Hyperspectral imagery of the pilot crops will be created.</td>
</tr>
<tr>
<td>- use pattern recognition through artificial intelligence/deep learning to detect, segment and quantify plant diseases. Hyperspectral imagery of the pilot crops will be created.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data format:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(<em>.docx), (</em>.csv), (<em>.tiff), (</em>.shp), (<em>.hdr), (</em>.bin)</td>
</tr>
<tr>
<td>Prescription maps (DSS) will be stored in a shapefile format, while images captured from the hyperspectral camera will be in a (<em>.tiff) or (</em>.hdr) and (*.bin) format. It might be that different data formats will be generated from the hyperspectral camera. More information will be provided in the updated version.</td>
</tr>
<tr>
<td>The expected size of the hyperspectral imagery is not yet defined (~ 50 MB), but the image stream will not challenge the storage capacity, because we will not store every single image.</td>
</tr>
</tbody>
</table>

2. FAIR Data

<table>
<thead>
<tr>
<th>2.1 Making data findable, including provisions for metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata such as date, bounding box, confidence score and other parameters will be useful for proving information for the disease data.</td>
</tr>
<tr>
<td>In addition, each dataset produced will be associated with a unique ID corresponding to the pilot of interest.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2.2 Making data openly accessible</th>
</tr>
</thead>
<tbody>
<tr>
<td>All deliverables are classified as ‘confidential’ so none of them will be under Public Domain Dedication and License (PD DL). Only consortium members will be able to access this information.</td>
</tr>
<tr>
<td>Field images and prescription maps will be obtained after farmers consent and will be used to evaluate the capacity of the vision based system by the project members. Access restrictions will be imposed in this data, by safekeeping in the OPTIMA repository.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2.3 Making data interoperable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data handled in the DSS portal, will be provided to other nodes through APIs. Based on that, there is possibility for combining with other datasets and using in various open software applications.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2.4 Increase data re-use (through clarifying licences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is mainly composed of spatial-temporal crop information (images, weather conditions). Re-usability is limited to the needs of this specific research effort.</td>
</tr>
<tr>
<td>Prescription maps record will provide a sufficient application history for the pilot sights. Project partners and farmers will be able to retrieve information for analysis and decision support.</td>
</tr>
</tbody>
</table>

3. Allocation of resources

| Funding dedicated for the preservation of prescription maps and other useful application records, have been considered for the DSS portal administrators (AGENSO). |
4. Data security
The algorithms that will be created for the purposes of OPTIMA, will be included in the deliverables of DSS, disease detection system and pattern recognition. These deliverables will be confidential among the OPTIMA project partners. To ensure confidentiality, the partners, will gain access to the file by authenticating themselves into the website.

5. Ethical aspects
Data to be used from the pilot areas are farmers’ property and consent forms will safeguard their data protection rights.

6. Other
N/A

WP3 - Combined synthetic and biological PPPs application and assessment of host and PPP resistance mechanisms

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP3 - Combined synthetic and biological PPPs application and assessment of host and PPP resistance mechanisms</th>
</tr>
</thead>
</table>
| 1. Data summary | Objectives and purpose of data creation:  
- Evaluate a collection of different categories of commercial and under development bio-PPPs; Datasets with experimental procedures and findings will be created.  
- Screen and select new generation synthetic PPPs for the disease/crops cases; Datasets with experimental procedures and findings will be created.  
- Understand the genetic, evolutionary and agronomic drivers of host resistance and evolution of PPP pathogen resistance; Datasets with experimental procedures and findings will be created.  
- Assess the dosage, frequency and timing of the applications of both bio- and synthetic PPPs for selected genotype; Datasets with experimental procedures and findings will be created.  
Data format:  
(*.xlsx), (*.docx), (*.pdf)  
Excel files for collecting and processing the lab test results. Word and .pdf files for documenting the results of each task.  
No significant storage requirements, as the files are expected to be small (≤ 50 MB each) |
| 2. FAIR Data | Meaningful metadata will be produced as a result of every Task in WP3 (time and date of data creation or data amendments, owners of actions that took place, lab and field experiments set-up). Metadata will assist the discoverability of the data and related information. |
| 2.1 Making data findable, including provisions for metadata | Meaningsfull metadata will be produced as a result of every Task in WP3 (time and date of data creation or data amendments, owners of actions that took place, lab and field experiments set-up). Metadata will assist the discoverability of the data and related information. |
Name conversions will be determined from the WP3 team as soon as research will start.

2.2 Making data openly accessible
All WP3 data will be openly available. The data will be downloadable through OPTIMA website or scientific repositories complying with Open Access policy.

2.3 Making data interoperable
WP3 data will come as guidelines and research results, all reported in standard word (*.docx) or (*.pdf) format. Hence, several software will be able to understand and read this data.

2.4 Increase data re-use (through clarifying licenses)
The WP3 data will be freely available on the OPTIMA website. In addition, significant scientific findings from the research activities will be published in free-access peer-reviewed journals.

3. Allocation of resources
N/A

4. Data security
N/A

5. Ethical aspects
N/A

6. Other
N/A

WP4 - Development and optimization of innovative spraying technologies

DMP component | WP4 - Development and optimization of innovative spraying technologies
--- | ---
1. Data summary | Objectives and purpose of data creation:
- define optimal spray configuration and parameters for the different crop-disease combinations, considering on every case the specifications of the selected PPPs (WP3) in the selected crops; Datasets will be created out of this experiment effort to document the findings.
- develop drift reduction technologies for the different crop-disease-PPP combinations; Datasets will be created out of this experiment effort to document the findings.
- develop dedicated smart sprayers for the Alternaria/carrots case; the downy mildew/vineyard case and the apple scab/apple case; Datasets will be created out of this integration effort to document the methodology.

Data format:
(*.xlsx), (*.docx), (*.pdf)
The *.xlsx format files will be used for collecting the experiment findings for drift reduction and bio-PPP efficacy tests. The guidelines and methodology to integrate smart components will be documented in .docx and .pdf format.

No significant storage requirements, as the files are expected to be small (≤ 50 MB each)

This data will be useful to provide the optimal spraying operation settings, as well as the experimental results of drift and bio-PPP efficacy to the operators and researchers.

2. FAIR Data

2.1 Making data findable, including provisions for metadata

Meaningful metadata will be produced as a result of every Task in WP4 (time and date of data creation or data amendments, owners of actions that took place, lab and field experiments set-up). Metadata will assist the discoverability of the data and related information.

Name conversions will be determined from the WP3 team as soon as research will start.

2.2 Making data openly accessible

Data related with drift reduction and bio-PPP efficacy will be openly available. The data will be downloadable through OPTIMA website or scientific repositories complying with Open Access policy.

On the contrary, deliverables and reports that contain confidential information about the sprayers’ capabilities and working principles are marked as confidential and access shall be restricted among the consortium members. Therefore, no public release of this data is foreseen.

2.3 Making data interoperable

The deliverables and reports file will be “as is” in a pdf or .docx format only to be used for reference. On the other hand, the experiment data accumulation files 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 licenses)

Publicly available data will be released without restrictions for further re-usability.

No quality assurance measures will be imposed after the initial release, since the validation procedures and the project partners guarantee the quality of the original data.

3. Allocation of resources

N/A

4. Data security

Confidential data regarding the sprayer information will be secured in the OPTIMA private repository, with limited access only from the partners involved.

5. Ethical aspects

N/A
WP5 - Pilot testing of developed technologies

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP5 - Pilot testing of developed technologies</th>
</tr>
</thead>
</table>
| 1. Data summary | Objectives and purpose of data creation:  
- evaluate the disease detection system and the DSS in 3 areas in Spain (apple orchards), Italy (vineyard), France (carrots); Evaluation sheets and assessment forms will follow the implementation of this effort.  
- evaluate and quantify the efficacy of selected synthetic and bio-PPPs and the defined best IPM practices for the three diseases under commercial field conditions; Evaluation sheets and assessment forms will follow the implementation of this effort.  
- evaluate and quantify the improvements on deposition, coverage and drift reduction obtained with the developed smart sprayers; Evaluation sheets and assessment forms will follow the implementation of this effort.  
- Globally evaluate and quantify the improvements generated by the OPTIMA holistic IPM system through field trials. Evaluation sheets and assessment forms will follow the implementation of this effort.  

Data format:  
(*.xlsx), (*.docx), (*.pdf)  
The *.xlsx format files will be used for collecting the evaluation results from the pilot testing. Evaluation reports will be documented in .docx and .pdf format.  
No significant storage requirements, as the files are expected to be small (≤ 50 MB each)  
This data will be useful for the researchers and end-users because they provide with evaluation results of the IPM system and all the OPTIMA related developments. |
| 2. FAIR Data |  |
| 2.1 Making data findable, including provisions for metadata | Meaningful metadata will be produced as a result of every Task in WP5 (time and date of data creation or data amendments, owners of actions that took place, evaluation protocols). Metadata will assist the discoverability of the data and related information.  
Name conversions will be determined from the WP5 team as soon as research will start. |
| 2.2 Making data openly accessible | All WP5 data will be openly available. The data will be downloadable through OPTIMA website or scientific repositories complying with Open Access policy. |
2.3 Making data interoperable

The deliverables and reports file will be “as is” in a pdf or .docx format only to be used for reference. On the other hand, the evaluation data files 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 WP3 data will be freely available on the OPTIMA website. In addition, significant scientific findings from the evaluation activities will be published in free-access peer-reviewed journals.

3. Allocation of resources

N/A

4. Data security

N/A

5. Ethical aspects

N/A

6. Other

N/A

WP6 - Human Health, Environmental and Socioeconomic Life-Cycle Analysis & Risk Assessment

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP6 - Human Health, Environmental and Socioeconomic Life-Cycle Analysis &amp; Risk Assessment</th>
</tr>
</thead>
</table>
| 1. Data summary | **Objectives and purpose of data creation:**  
- perform an extended Life-Cycle Assessment (LCA) combined with Human and Environmental Risk Assessment (HERA) using a Multi-Criteria Decision Analysis (MCDA); Questionnaires and reports will be produced.  
- assess human health, environmental and socio-economic impacts and risks of treatments of the selected crops, comparing conventional and the proposed IPM crop protection system; Questionnaires and reports will be produced.  
- conduct comparative life-cycle costing of the new IPM system and the conventional practices using sensitivity scenarios. Questionnaires and reports will be produced.  
  
**Data format:**  
(*.xlsx), (*.docx), (*.pdf)  
The *.xlsx format files will be used for collecting all the necessary data under WP6. Assessment reports and deliverables will be documented in *.docx and *.pdf format.  
No significant storage requirements, as the files are expected to be small (≤ 50 MB each) |
This data will be useful for the researchers and end-users because they provide with socio-economic impact analysis, LCA and life-cycle costing reports of the IPM system.

| 2. FAIR Data | Meaningful metadata will be produced as a result of every Task in WP6 (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.
| 2.1 Making data findable, including provisions for metadata | Name conversions will be determined from the WP6 team as soon as research will start.
| 2.2 Making data openly accessible | All WP6 data will be openly available. The data will be downloadable through OPTIMA website or scientific repositories complying with Open Access policy.
| 2.3 Making data interoperable | The deliverables and reports file will be “as is” in a pdf or .docx format only to be used for reference. On the other hand, data related with assessment efforts in ever Task 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 WP6 data will be freely available on the OPTIMA website. In addition, significant scientific findings from the evaluation activities will be published in free-access peer-reviewed journals.

| 3. Allocation of resources | N/A |
| 4. Data security | N/A |
| 5. Ethical aspects | N/A |
| 6. Other | N/A |

### WP7 - Dissemination, Communication & Exploitation

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP7 - Dissemination, Communication &amp; Exploitation</th>
</tr>
</thead>
</table>
| 1. Data summary | Objectives and purpose of data creation:
- Widely disseminate the project results at regional, national and European level; Data created will be the communication channel with OPTIMA audience.
- ensure as widespread as possible the communication of the new products and services generated within the consortia; Data created will be the communication channel with OPTIMA audience. |
- develop the exploitation channels and commercialization strategy for the novel products; Data created will be the communication channel with OPTIMA audience.

**Data format:**
(*.jpg), (*.docx), (*.pdf), (*.ai), hard copy
The above mentioned formats will include digital data and hard copies of brochures, leaflets, banners, press releases, posters etc. Furthermore, information from the visitors of the OPTIMA website acting as newsletter subscribers, will be collected.

No significant storage requirements, as the files are expected to be small (≤ 100 MB each)

This data will be useful for promoting the OPTIMA dissemination activities, while hearing from the end-users and communities about the impact and acceptance of the project. Moreover, this data will be a useful tool to maximize the effectiveness and penetration of OPTIMA activities towards the outer world.

<table>
<thead>
<tr>
<th>2. FAIR Data</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Making data findable, including provisions for metadata</td>
<td></td>
</tr>
<tr>
<td>2.2 Making data openly accessible</td>
<td>Most of WP7 data will be openly available. The data will be downloadable through OPTIMA website or scientific repositories complying with Open Access policy. Only personal information of the newsletter subscribers and the Exploitation Plan of the project will have a restricted access only to consortium members.</td>
</tr>
<tr>
<td>2.3 Making data interoperable</td>
<td>The deliverables and dissemination material files will be “as is” in a digital format (.pdf, .docx etc.) or in hard copies, so the interoperability with other means of processing information is easy.</td>
</tr>
<tr>
<td>2.4 Increase data re-use (through clarifying licences)</td>
<td>N/A</td>
</tr>
<tr>
<td>3. Allocation of resources</td>
<td>No additional budget is required to keep the data FAIR.</td>
</tr>
<tr>
<td>4. Data security</td>
<td>Personal information obtained from newsletter subscribers will be safely stored in the website’s administrator premises (AGENSO), using encryption tools to prevent any security breach event.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>5. Ethical aspects</td>
<td>Newsletter subscribers will first accept the terms and conditions, that are aligned with the GDPR, before providing their personal data.</td>
</tr>
<tr>
<td>6. Other</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**WP8 - Project Management**

<table>
<thead>
<tr>
<th>DMP component</th>
<th>WP8 - Project Management</th>
</tr>
</thead>
</table>
| **1. Data summary** | **Objectives and purpose of data creation:**
To manage and document all the data that will be produced from the Project Management related activities during the OPTIMA project. |
<p>| | <strong>Data formats:</strong> *.xlsx , <em>.doc, <em>.pdf files</em></em> |
| | The data includes doc and pdf files that will contain the transcript of the teleconferences, webinars, and discussions held among the consortium members and the Project Coordination Team, in order to make it accessible and readable to the hearing impaired. Reports related to the PCT will also be in doc format. |
| | No existing data will be used or reused. Data will be produced only within the time duration and in the context of OPTIMA project as described in the GA. |
| | 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 or as advices for a better project implementation when risks are identified. In addition, it would be useful to anyone who would like to reflect on the impact that the PCT would have to the evolution of the project itself. |
| <strong>2. FAIR Data</strong> | <strong>The inclusion of metadata for the current dataset has not been yet decided.</strong> |
| <strong>2.1 Making data findable, including provisions for metadata</strong> | Unique and persistent identifiers will not be used for this dataset. |
| <strong>2.2 Making data openly accessible</strong> | Doc transcripts of the conferences will not be openly available since they represent an initial and unstructured reflection of the internal project meetings. These initial data will be kept in secure repositories shared only to OPTIMA partners. Same applies for the |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3 Making data interoperable</td>
<td>Not applicable. There are no interoperability issues related to this dataset.</td>
</tr>
<tr>
<td>2.4 Increase data re-use (through clarifying licenses)</td>
<td>See point 2.1 for data access. Public data will remain visible during the entire website lifecycle.</td>
</tr>
<tr>
<td>3. Allocation of resources</td>
<td>All costs related to the data collection and processing are covered by the project budget with dedicated person months under WP8.</td>
</tr>
<tr>
<td>4. Data security</td>
<td>Audio and doc files will be deposited in OPTIMA private repository and will be protected with the server’s security protocol. Pdf files that will be deposited in the project website will follow the security protocol of the website hosting service (AGENSO).</td>
</tr>
<tr>
<td>5. Ethical aspects</td>
<td>There are ethical issues regarding the WP8 data. First, for the audio recording of the project Management derived data, the consent of the PCT participants for this recording was asked prior to the beginning of the discussion. Their consent is recorded in the same audio file.</td>
</tr>
<tr>
<td>6. Other</td>
<td>N/A</td>
</tr>
</tbody>
</table>
4. OPTIMA Support Package

The Support Package includes a list of recommendations that should be applied for the project in general. From the previous Chapters in this document the following general guidelines for data management in OPTIMA can be derived:

- To prepare the Data Management Plan within the first six months of the project and update the report during the lifetime of the project when significant changes will happen (e.g. new data types to be included, changes in the consortium)
- Research papers that are derived from the project should be published according to the open access policy
- Research data should be stored in a central repository according to the FAIR principles: findable, accessible, interoperable and reliable
- Continuously investigate what kind of data are extracted from every Task during the project implementation.
- Analyze and explore the Project Tasks in a deeper way in order to identify which data management issues potentially play a role and define plans how to deal with them.
- Each partner should align with the DMP, when managing data and comply with the Open Access policy in the event of data release.
- Especially for confidential and ethical aspects, the project Coordinator and Project Management team will give emphasis and monitor closely the data management practices followed. Partners managing such data will be alerted in advance, to ensure that no deviation from the plan will take place.
- The next DMP update will take place by sending to all OPTIMA partners a questionnaire asking about all the DMP aspects for every task. All partners will contribute according to their involvement in the data creation.
5. Conclusions

Data rapidly has become a new resource or asset in the current economy, also in the agricultural sector. This development leads to several issues that have to be addressed such as data availability, quality, access, security, responsibility, liability, ownership, privacy and costs. In Europe several initiatives and projects have already been established to work on this in a general context (e.g. the EU open data policy) and more specifically in agriculture (e.g. EIP-Agri, agINFRA).

In this document we have presented a first version of a Data Management Plan with concrete actions to be taken into consideration in order to establish open, transparent data management in OPTIMA. It includes the basic methodology and guidelines that will be followed as well as the template that will accompany all datasets.

As most of the research is on a premature level, it is expected that the next update of the DMP will be able to have all the related datasets clarified and resolved. These aspects include questions related to hosting the data (persistence), appropriately describing the data (data provenance, relevant audience for re-use, discoverability), access and sharing (rights, privacy, limitations) and information about the human and physical resources expected to carry out the data management plans per dataset.