# D9.1 Data management plan

robs4crops.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016807

| | |
|---|---|
| **Author(s)/Organisation(s)** | Jochen Hemming (WR) |
| **Contributor(s)** | All RobsForCrops partners contributed with their own specific DMP section |
| **Work Package** | 9 |
| **Delivery Date (DoA)** | M6 |
| **Actual Delivery Date** | M6 |
| **Abstract:** | This deliverable defines the data management plan of the Robs4Crops project. It specifies how research publications and data generated by the project will be collected, processed, monitored, catalogued, and disseminated. Next to the description of the general project principles of the project the individual data management plans of each consortium partner is included. |

| **Document Revision History** | | | |
|---|---|---|---|
| **Date** | **Version** | **Author/Contributor/ Reviewer** | **Summary of main changes** |
| 30/04/2021 | 1.0 | J. Hemming | First Draft |
| 10/06/2021 | 1.1 | J. Hemming; N. Kim<br>all Robs4Crops partners | Second draft, including individual partner contri-butions. |
| 14/06/2021 | 1.2 | Internal review by Grigoris Chatzikostas (FSH), Maria Navarro Abellán (EUT) and Ferrán Roure García (EUT) | Suggested revisions. |
| 24/06/2021 | 1.3 | J. Hemming | Suggested revisions by FSH and EUT processed. Ready for submission. |

| **Dissemination Level** | | |
|---|---|---|
| **PU** | Public | X |
| **PP** | Restricted to other programme participants (including the EC Services) | |
| **RE** | Restricted to a group specified by the consortium (including the EC Services) | |
| **CO** | Confidential, only for members of the consortium (including the EC) | |

| ROBS4CROPS Consortium | | | |
|---|---|---|---|
| Participant Number | Participant organisation name | Short name | Country |
| 1 | STICHTING WAGENINGEN RESEARCH | WR | NL |
| 2 | GIROPOMA COSTA BRAVA SL | GIR | ES |
| 3 | AGROTIKOS SYNETAIRISMOS POLISEOS XIRON KAI NOPON STAFYLION KIATOY KORINTHIAS PIGASOS | PEG | GR |
| 4 | SERRATER SL | SER | ES |
| 5 | SMART AGRI TECHNOLOGY BV | SAT | NL |
| 6 | TERRENA SOCIETE COOPERATIVE AGRICOLE | TER | FR |
| 7 | ABEMEC BV | ABE | NL |
| 8 | AGREENCULTURE | AGC | FR |
| 9 | AGRO INTELLIGENCE APS | AGI | DK |
| 10 | FOODSCALE HUB ENTREPRENEURSHIP ANDINNOVATION ASSOCIATION | FSH | SR |
| 11 | TEYME TECHNOLOGIE AGRICOLA SL | TEY | ES |
| 12 | GEOPONIKO PANEPISTIMION ATHINON | AUA | GR |
| 13 | FUNDACIO EURECAT | EUT | ES |
| 14 | KOBENHAVNS UNIVERSITET | UCHP | DK |
| 15 | UNIVERSITAET HOHENHEIM | UHOH | DE |
| 16 | PANEPISTIMIO PATRON | LMS | GR |



Funding Scheme: Innovation Action (IA) • Topic: H2020-ICT-46-2020

Start date of project: 01 January 2021 • Duration: 48 months

# Table of Contents

# List of tables



# List of figures

No table of figures entries found.

| List of Abbreviations and Acronyms | |
|---|---|
| CA | Consortium Agreement |
| DMP | Data Management Plan |
| DPO | Data Protection Officer |
| DOI | Digital Object Identifier |
| FAIR | Findable, Accessible, Interoperable and Reusable |
| N/A | Not Applicable |
| ORD pilot | Open Research Data Pilot |
| TBD | To Be Defined |

# 1 Introduction

This deliverable defines the data management plan (DMP) of the Robs4Crops project. Robs4Crops is a H2020 Innovation Actions in which research results are implemented for industry benefit. The Robs4Crops project takes part in the Open Research Data Pilot (ORD pilot) under Horizon 2020. The ORD pilot aims to improve and maximize access to and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness and protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions. The ORD pilot applies primarily to the data needed to validate the results presented in scientific publications. Other data can also be provided by the beneficiaries on a voluntary basis. As further detailed in the individual consortium partner DMPs (section 7) the shared data will be supported by documentation and long-term preservation.

The DMP will ensure that data management and protection is compliant with EU principles and standards and with relevant national data protection laws and institutional data management policies.

The DMP includes information on:

- the handling of research data
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared, how data will be shared, with open access as one option, and
- how data will be curated and preserved (including after the end of the project).

In a broader sense also items such as documents and reports can be considered as data. The DMP however targets specifically the research data collected and created during the research activities of the Robs4Crops project. Thus, the management of the administrative documents of the project, such as presentations and meeting minutes and legal and financial documents are out of the scope of the DMP described here.

This document is organized as follows: Section 0 gives an overview of the background data brought into the project and an overview of data that will be collected, generated, and processed during the project. This section also elaborates on how and where project data is stored and how it can be shared. Section 3 is dedicated to data security. Section 4 describes the ethical aspects of the DMP. Section 5 explains the general FAIR principles (Findable, Accessible, Interoperable and Reusable) and data management guidelines applied in the Robs4Crops project. The general principles described in the previous sections are then refined by the consortium partner's practices. The individual partner DMPs are provided in section 7 of this document.

# 2 Summary of the data

## 2.1 Background data of the partners as in the CA

"Background" is defined as intellectual property assets that a partner of the consortium is bringing to the project. These can be for example inventions, know-how, secret knowledge, methods etc. "Background data" is data and/or data bases already existing before the project, that are brought to the project. Partners must give each other access to background that is needed to implement their own tasks and exploit their own results. There is however no obligation to give access if there are restrictions or limits (legal or otherwise) and the partner has informed the other partners about. Background is specified in the Consortium Agreement (CA). Only AGC has defined background data in the Robs4Crops CA:

| Ownership | Description and origin | Use |
|---|---|---|
| AGC | All remote management and monitoring data from AGC is owned by AGC. | Data may be shared for experiments with other partners |

*Table 1: Background data identified by the partners and stated in the Consortium Agreement.*

## 2.2 Data that will be collected, processed and/or generated during the project

A general overview on the data generated in each WP is shown in Table 2.

| Workpackage | Data to be collected |
|---|---|
| WP1 Ecosystem building | Answers to questionnaires. Responses from stakeholders collected during workshops, interviews, demonstrations, and other events. |
| WP2 Smart implements | Data collected by sensors attached to implements. Data on the operation performance of the smart implements (e.g. counts of crop plants damaged, or percentage of weeds removed). |
| WP3 Autonomous vehicles | Data collected by sensors for autonomous navigation. Data on the telemetry and operation performance of the autonomous vehicles (e.g. mean deviation from path, number of manual overrides needed). |
| WP4 Farming controller | Data retrieved from FMIS owned by partners. Simulation data originating from the digital twins. |
| WP5 Tests | Data resulting from testing combinations of implements and vehicles, such as communication logs. |
| WP6 Large-scale pilots | In the 4 large-scale pilots large volume of data is expected to be collected:<br>• Data on cropping practices (crop type, sowing or planting date, fertilization, irrigation, weed control measures),<br>• Data on cropping results (crop yield, crop quality)<br>• Data on weeds (type, size, density) before and after weed control. |

| | |
|---|---|
| | • The characteristics of the farm (number, size, and location of fields; number and types of implements and tractors).<br>• Application techniques used.<br>• Weather data<br>• History of cultivation seasons during and before the project.<br>• Imaging data of fields and crops, including RGB, thermal, hyperspectral images.<br>• Data on soils (soil samples, EC scan, water content)<br>• Remote supervision/monitoring data from the implements and vehicles consisting out of telemetry and video data. |
| WP7 Socio-economics and ethics | Data resulting from interviews with stakeholders during the social and ethical impact assessment.<br>Economic performance data of robotic systems at farm level. |
| WP8 Exploitation, communication & dissemination | Personal data such as names, affiliations and e-mail addresses used for mailing lists and newsletters and alike. |
| WP9 Management | Agreements, project reports and project deliverables, minutes, videos and photos of project related meetings. |

*Table 2: Types of data generated in each work Package*

By default, data collected or created within the Robs4Crops project is first of all stored on the data storage systems of the partner that has collected/created and thus owns the data. Starting point is, that this data is sufficiently sensitive, and that by default it is not shared with others. How each partner deals with the various aspects of his data is further detailed in the individual DMPs in section 7.

For the project, some data will be shared between some of the consortium partners. As an example this could be data logged by the autonomous vehicles and smart implements during field operations (WP6). This data might be needed by the technical work packages (WP2,WP3, WP4, WP5) for fulfilling their tasks. Another example would be the writing of a joint publication by several partners that may require that some data is shared. For within project data sharing, Robs4Crops makes use of the Microsoft OneDrive for Business services provided by the coordinator, WR. Access rights to that system is based on credentials (username and password). Without other agreements made, the data collector/creator/owner stays responsible for his own data. General data security, ethics and FAIR principles are specified further down in this document.

Finally, some of the data of the project might be useful for others outside the consortium. The degree to which data are openly shared is case specific and there might be good reasons to keep data confidential. The general principle of Robs4Crops however is, to make data as much as possible openly accessible. Section 5 will go into detail on how this is implemented.

# 3 Data security

For within project sharing of documents and data, Robs4Crops makes use of the Microsoft OneDrive for Business services provided by the coordinator, Wageningen Research. OneDrive for Business fulfills the highest data security standards, it has restricted access, and document versioning. Access to the data is managed by credentials (username and password) including occasional two-factor authentication. OneDrive data is stored on servers inside the EU. Microsoft is committed to taking all necessary precautions to ensure the physical safety and security of the data. This includes fire prevention and protection and physical intruder prevention of the data centers. As a general rule, the Robs4Crops project will not store data, especially sensitive data, on local systems only (such as laptops, external hard drives or USB sticks etc.). It might be needed however to temporarily do this (for the minimum time needed) to carry out a specific task of the project. Locally collected data will be copied/backed up as soon as possible.

How the individual partners have arranged the data security of their own data is depicted in section 7.

For long term storage and publication (including beyond the lifespan of the project) certified repositories that ensure data security are preferred.

# 4 Ethical aspects

The confidentiality of personal information provided to the consortium during the execution of the project needs to be safeguarded. Think for example of information provided by farmers about their businesses or their fields or crops. It is not foreseen that research in Robs4Crops will involve research on human subjects. As such no special approval from an Ethics Committee is needed. Personal data collected must be held in confidence and should be destroyed when no longer required. One option to retain data for a longer period, and even make it available to others, is to anonymize personal data. All Robs4Crops partners are aware that they must comply with the European General Data Protection Regulation (GDPR). The GDPR's primary aim is to give individuals control over their personal data and to simplify the regulatory environment for international business.

# 5 FAIR data

Robs4Crops is committed to the general FAIR principles, that is making data Findable, Accessible, Interoperable and Reusable. The general FAIR principles are described in the following and are refined by the individual partner's practices, as described in the partner DMPs in section 7.

## 5.1 Making data findable

To make data findable it must be stored and organized in a structured and defined way. Each partner has defined his own principles of doing so (see section 7). Examples of making fata findable are: using a defined way of naming folders and files on a data drive and by adding meta-data and keywords to the dataset.

**Persistent identifiers**

For data published to a wider audience persistent identifiers/references will be used. A persistent identifier (PI or PID) is a long-lasting reference to a document, file, web page. Persistent data identifiers are preferred when publishing data. Following the data repository practice, a data set will be assigned an unique and persistent identifier. Example of a PID is a digital object identifier (DOI) number from the International DOI Foundation https://www.doi.org/ assigned to a publication.

**Keywords**

Suitable search keywords are used for each data set to optimize that potential users of the data can find the data set. The keywords used will match the content of the data and the standards of the research field.

**Version numbers**

In some cases, already deposited data sets are updated, for example with new observations. The versions will be identified with the date and with the key issues that are updated (e.g. number of observations), including notes on the potential corrections for the data.

## 5.2 Making data openly accessible

### 5.2.1 General principles

The degree to which data are openly shared is case specific, determined by the involved partners owning the data. The owners of the data (IPR) are defined by the rules set in the Consortium Agreement.

The partners will ensure that adequate steps towards IPR protection are taken prior to exploitation, dissemination and communication, preventing unapproved public disclosure of data, results, tools, products and services. The Project Executive Board and WP leaders will continuously work out the plan together with the relevant partner(s). Simultaneously, Robs4Crops will ensure a) the obligation to protect results b) the confidentiality obligations, c) the security obligations and d) the obligations to protect personal data, as stated in the Grant Agreement, all of which apply. Reasons for data not being shared may include commercial exploitation, protection of IPR, securing of future data production, project agreements, privacy policy or other legislation.

Apart from the above-mentioned restrictions, the general principle of Robs4Crops is to make data openly accessible when this is both possible and useful. Robs4Crops prefers that the data, once used in a publication, will be deposited in an open-access research data repository to facilitate further use of data. Measures are taken to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the data, including associated metadata, needed to validate the results presented in scientific publications and to further utilize the data.

An additional advantage of making data accessible in an open-access research data repository is that, once published, it involves not further work for the author of the data to distribute the data. Interested users can easily be provided the persistent identifier and a reference to the data.

### 5.2.2 Data repositories and software

The data and/or metadata that are shared will be shared via open-access data repositories. These repositories guarantee long-term preservation, typically >15 years. Examples of such repositories, but not limited to, are:

- Data Archiving and Network Services DANS (https://easy.dans.knaw.nl/ui/home)
- 4TU.Centre for Research Data (https://www.4tu.nl/datacentrum/en/) .

For software source code, repositories such as GIThub (https://github.com/) or local GITlab versions can be used. GitHub is a website which allows to host, review and manage code online. GIT repositories offer distributed version control and source code management functionality.

## 5.3 Making data interoperable

Data will be in the form that is a common standard in this field. This makes the data sets interoperable allowing data exchange and re-use between researchers, institutions, organisations, and countries.

**Metadata standards**

Metadata is data that describes other data. Meta is a prefix that in most information technology usages means "an underlying definition or description." Metadata summarizes basic information about data, which can make finding and working with particular instances of data easier. (http://whatis.techtarget.com/definition/metadata).

The metadata and documentation in Robs4Crops will follow commonly-agreed standards. One example of a metadata standard for Agricultural Sciences is the AGRIS Application Profile for the International Information System on Agricultural Sciences: http://www.fao.org/3/ae909e/ae909e00.htm.

## 5.4 Increase data re-use

The project will produce a number of re-usable assets, including knowledge (scientific and engineering papers), software, business models, and datasets. In general, data re-use is encouraged by assuring that easy findable and accessible and that data is supported by documentation and long-term preservation.

The Robs4Crops approach is to establish and use the connections to the pan-European Digital Innovation Hubs (DIHs) to share reusable assets, including data. Good examples are the agrifood related EU-funded projects SmartAgriHubs and agROBOfood in which some Robs4Crops partners are already integrated or have an instrumental role. By participating or co-organizing in DIH's activities, such as workshops, demos and events the re-use of data will be stimulated.

**Licensing**

For data published on open-access repositories, the data will be free of charge for any user. Still, specific licenses may apply to this data, such as one of the Creative Commons (CC) licenses. There are six different CC license types that range from "Credit must be given to the creator" to "Only noncommercial uses of the work are permitted". Details on the different CC licenses can be found here: https://creativecommons.org/about/cclicenses/.

For the data with restricted access, specific contracts can be generated between the owner and the third party, to safeguard the owner and to agree on general rules for the data use.

**Embargo**

Robs4Crops has many company partners and IPR protection is of highest priority. In some cases, embargo for opening the data may be needed to give time to publish or seek patents, yet keeping in mind that the aim is to make data available as soon as possible. Embargo may be needed to protect company business, product development and its IPR, patent applications, or repeated use of the same data by the partners for additional output. Again, the need for embargo and its length is case specific, determined by the involved partners owning the data. The procedures to follow concerning the approval for publications and to protect results have been explained during the Robs4Crops IP Workshop no. 2. Details on this can be found in the appendix in section 8.2.

## 5.5 Allocation of resources

### 5.5.1 Covering the cost of open data

The work to process and document the data is part of the other direct costs of each partner. Open data costs will be paid from the respective partner's budget. Some public open access data repositories are free of charge to use.

### 5.5.2 Responsible person

Individual partners will be responsible for the collection, transfer, storage and documentation of data from their own research activities, within the framework of the DMP. Each consortium partner has detailed in his own DMP in section 7 who will be responsible for the data.

# 6 The next steps of the Data Management Plan

This is the first released version of the DMP of the Robs4Crops project, delivered early in the project, in M6. In the current version there are still some white spots and open questions and discussions. Especially between the individual partner DMPs there are some discrepancies that will be further aligned in future revisions of this document.

The DMP will be updated during the course of the project whenever significant changes arise, such as:

- New data
- Changes in consortium policies (e.g. new innovation potential, decision to file for a patent)
- Changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

The DMP will be updated as a minimum in time with the periodic evaluation moments of the project. These evaluation moments have been earlier defined at M12, M36 and M48. As such, updated versions of the DMP will also be submitted at these moments.

# 7 Data Management Plan of each consortium partner

## 7.1 Data Management Plan of WR

Author: Jochen Hemming (WR)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | As coordinator WR will host data-sets relevant to the execution of the project. As WP5 leader, data resulting from testing combinations of implements and vehicles will be collected. Also, as responsible partner for some tasks in other WPs, such as WP2, data will be generated during experiments. |
| What types of data will be generated/collected? | WP2:<br>• Sensor data from monitoring smart weeders, such as camera images and other machine data (location, speed, application maps, etc.). Software code for applications monitoring the smart weeders.<br>WP5:<br>• Data resulting from testing combinations of implements and vehicles, such as communication logs and experimental results.<br>WP9:<br>• Datasets generated in other WPs that needs to be shared within the consortium during the execution of the project. Type of data can vary. |
| What formats of data will the task generate/collect? | WP2: Images, videos, Text files containing for example position data, other sensor data.<br><br>WP5: Text files containing log data. By preference using csv format but maybe using Excel (*.xlsx) or ROSbag format. |
| Will any existing data be used and how? If yes, what is the origin of the data? | Not expected. |
| What is the expected size of the data? (in Gb) | The expected size of data for WR is between 10Gb and 100 Gb. |
| To whom might it be useful ('data utility')? | Consortium partners, Researchers, Developers/Engineers of field robots. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification | Data is produced by conducting experiments with equipment (WP2 and WP5) and derived data from this. This data will be stored together with metadata describing the structure and content. |

| | |
|---|---|
| mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | Selected data and/or reports for a broader audience will be published in journals or data repositories that provide DOI numbers. |
| What naming conventions do you follow? | On the central OneDrive file storage of the project a separate folder for each WP will be maintained. For the structure inside this WP folder the WP leader is responsible. Where appropriate file or (sub-)folder names contain a date tag in the form of (YYY-MM-DD). |
| How are search keywords provided that optimize possibilities for re-use? | Documents /publications and released data -sets will be tagged with appropriate keywords. |
| How do you provide clear version numbers? | Documents will contain a document revision section/header that specifies date and version number. Source code will be stored on the GIT server of Wageningen University & Research (git.wur.nl). The GIT mechanism ensures that for each new commit a version number together with a description of changes is available. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | Datasets will be accompanied by documentation, for example readme (Ascii-format or MS-Word) files. This will contain for example origin and location of recordings and general data about settings and equipment used.<br>Some data types, such as image data, contain metadata in each image (EXIF tags such as timestamp, resolution, image format). |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[1]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | Data collected and/or generated is not by default shared. Only data that has the potential to be useful for others will be considered for publication. Publication of data will need to take into account IP, and privacy rules before publishing. |
| How will the data be made accessible (e.g. by deposition in a repository)? | Data will be copied/uploaded to the network or web-based/cloud based storage. |
| What methods or software tools are needed to access the data? | This depends on the type of data. The accompanying metadata and documentation (readme) will inform the user on the tools needed. |

[1] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| • Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | |
| Where will the data and associated metadata, documentation and code be deposited? | Published data will be put in a public and certified repository. By preference, the publication of data is linked to a (scientific) publication to support better findability.<br>To share data within the consortium only, data will be hosted at the network storage facilities (OneDrive or GITlab) of WR. |
| Have you explored appropriate arrangements with the identified repository? | Yes. |
| If there are restrictions on use, how will access be provided? | Access will be arranged based on user credentials. |
| How will the identity of the person accessing the data be ascertained? | For restricted data: by the credentials of the user. In case of public accessible repositories and open access platforms: WR will not have information on who has accessed the data. |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Data produced will be as much as possible compliant with available (open) software applications. The project consortium will agree on the data type to be used for the different tasks. |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | None. |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Per dataset it will be decided on the type of license to be used. For data shared open access, by preference a Creative Commons license will be applied. |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | This depends on the task the data was created. Inside the consortium data can be re-used as soon as the data is collected. Some data might only be made available after the acceptance of a scientific publication or a graduation of a student. |

| | |
|---|---|
| How long is it intended that the data remains re-usable? | In line with the data policy of WR all data will be stay at least 10 years available after the project is finished (see also below "Allocation of Resources") |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | Some data might be useful also after the project. That is the reason why we strive to publish selected data on certified long-term repositories. |
| How is the data quality assured? | The quality of the data will be primarily checked by the team that collects and analyses the data. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Costs for storing the data falls within the institutional budget (overhead) of WR.<br><br>The costs for depositing data for example at the 4TU.Cente for Research Data repository (see below) is free of charge for up to 10 GB per year. There is an institutional agreement that research datasets up to 1 TB are automatically reimbursed by WR. |
| Who will be responsible for data management in the task? | Each WP leader is responsible for the data collected in his WP. Ultimate responsible for the data in the Robs4Crops project is the project coordinator, Frits van Evert. He will be supported for the day-to-day data management tasks by Nicole Kim. |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | WR will comply the existing data policy of WR (https://www.wur.nl/en/Value-Creation-Cooperation/Collaborating-with-WUR-1/WDCC/Data-Management-WDCC/Data-policy.htm). All research data and underlying publications will be kept for a minimum of 10 years (https://edepot.wur.nl/446446) . |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data is stored at the WR network facilities that is operated and supervised by the ICT department of Wageningen University & Research. The project will make use of the Microsoft OneDrive for Business and SharePoint services, as well of the WUR GITlab servers. These services include regular backups, firewall and access based on personal credentials. In most cases two-factor authentication applies to access the network. |
| Is the data safely stored in certified repositories for long term preservation and curation? | Selected data will be published in certified repositories such as DANS (https://easy.dans.knaw.nl/ui/home) or 4TU.Centre for Research Data (https://www.4tu.nl/datacentrum/en/) . These |

| | |
|---|---|
| | repositories guarantee long-term access (>15 years) and preservation. |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | The data collected by WR does not carry any ethical or legal issues. |
| Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | WR complies with the GDPR. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | WR will comply the existing data policy of WR (https://www.wur.nl/en/Value-Creation-Cooperation/Collaborating-with-WUR-1/WDCC/Data-Management-WDCC/Data-policy.htm |

## 7.2 Data Management Plan of GIR

Authors: Joan Ponsí

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | -User feedback on working with robots/smart implements<br>-Product assessment<br>-Provide this information to the technical WPs for further robot development |
| What types of data will be generated/collected? | -Climate data<br>-Treatment efficiency<br>-Treatment precision |
| What formats of data will the task generate/collect? | -Office documents |
| Will any existing data be used and how? If yes, what is the origin of the data? | Not defined yet |
| What is the expected size of the data? (in Gb) | Not defined yet |
| To whom might it be useful ('data utility')? | All fruit producers, consortium partners. |

| FINDABLE |
|---|

| | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | There is no strategy defined yet |
| What naming conventions do you follow? | Not defined yet |
| How are search keywords provided that optimize possibilities for re-use? | Not defined yet |
| How do you provide clear version numbers? | *N/A* |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | *N/A* |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[2]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | All data will be accessible to all partners |
| How will the data be made accessible (e.g. by deposition in a repository)? | Online repository |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Regular software for each type of data.<br>To be further specified. |
| Where will the data and associated metadata, documentation and code be deposited? | Not defined yet |

[2] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| Have you explored appropriate arrangements with the identified repository? | Not defined yet |
| If there are restrictions on use, how will access be provided? | Not defined yet |
| How will the identity of the person accessing the data be ascertained? | Not defined yet |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | Not defined yet |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Creative Commons license (according to the H2020 guidelines) |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Not defined yet |
| How long is it intended that the data remains re-usable? | Not defined yet |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | |
| How is the data quality assured? | |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Not defined yet |
| Who will be responsible for data management in the task? | Giropoma Costa Brava SL |
| Are the resources for long term preservation discussed (costs and | Not defined yet |

| | |
|---|---|
| potential value, who decides and how what data will be kept and for how long)? | |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data stored in Giropoma Costa Brava SL's networks with backups, firewall etc |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | There will be no personal information in our data |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.3 Data Management Plan of PEG

Authors: Fouli Doukas (PEG)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | The purposes of the data collection are:<br>- user feedback<br>- product improvements<br>- Assess status of use cases<br>- Assess status of product development<br>- reduce applications<br>- create robots to work with accuracy in cultivation |
| What types of data will be generated/collected? | Types of data that will be collected are: |

| | |
|---|---|
| | - General information on use case (use case number etc)<br>- Climate information<br>- Test farm / company information<br>- Personal information respondent (name, e-mail address)<br>- Information of the usage of the solution from the user perspective<br>- Accuracy<br>- Time to execute orders<br>- Effectiveness |
| What formats of data will the task generate/collect? | Excel (measurements)<br>Photos<br>Word |
| Will any existing data be used and how? If yes, what is the origin of the data? | Time records, spray records, photos |
| What is the expected size of the data? (in Gb) | To be answered later |
| To whom might it be useful ('data utility')? | To robot constructors. Consortium partners. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | No metadata expected. |
| What naming conventions do you follow? | Folders and files conventions with WP name and date |
| How are search keywords provided that optimize possibilities for re-use? | Documents /publications and released data -sets will be tagged with appropriate keywords. |
| How do you provide clear version numbers? | Each file named also by date |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | No metadata expected. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly | All data can be open for all partners. |

| | |
|---|---|
| available as the default[3]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | |
| How will the data be made accessible (e.g. by deposition in a repository)? | Data shared within the consortium will be copied/uploaded to the OneDrive file shares provided by WR. |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Common Microsoft Office tools such as MS-Word, Excel. |
| Where will the data and associated metadata, documentation and code be deposited? | N/A |
| Have you explored appropriate arrangements with the identified repository? | N/A |
| If there are restrictions on use, how will access be provided? | N/A |
| How will the identity of the person accessing the data be ascertained? | e.g user registration process |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | N/A |

[3] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | To be defined. |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Inside the consortium data can be re-used as soon as the data is collected. |
| How long is it intended that the data remains re-usable? | As long as it needed for program and for many years |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | After the project is finished, the generated spatial data will be useful for other researchers who are involved with robotics and farming. |
| How is the data quality assured? | Accurate measurements, photos, Responsible personnel |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Costs for storing the data falls within the institutional budget (overhead) of PEG and WR (in the case of Onedrive for data shared within the consortium). |
| Who will be responsible for data management in the task? | A person inside PEG will be assigned for this. |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Yes |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | SharePoint/OneDrive (provided by Wageningen University and Research that ensures the data security and recovery) is accessible with credentials, Basecamp as well. |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | Data that need to comply to the GDPR etc. Names, education level, etc |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning | N/A |

| | |
|---|---|
| information provisions and access to personal data? | |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.4 Data Management Plan of SER

Authors: Raul Sanchez (SER)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | • To evaluate production development<br>• Economically, environmentally and socially Improving treatments<br>• To reduce the number of applications |
| What types of data will be generated/collected? | • Climate data<br>• Treatment efficiency<br>• Treatment precision |
| What formats of data will the task generate/collect? | Excel<br>Pictures<br>Text |
| Will any existing data be used and how? If yes, what is the origin of the data? | No |
| What is the expected size of the data? (in Gb) | TBD |
| To whom might it be useful ('data utility')? | Consortium partners, fruit producers |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | It is not foreseen to use specific methods/mechanisms for this. |
| What naming conventions do you follow? | To be defined. |
| How are search keywords provided that optimize possibilities for re-use? | N/A |
| How do you provide clear version numbers? | N/A |
| What metadata will be created? In case metadata standards do not | N/A |

| | |
|---|---|
| exist in your discipline, please outline what type of metadata will be created and how. | |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[4]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | All data will be open to all partners |
| How will the data be made accessible (e.g. by deposition in a repository)? | Using the OneDrive shared folder provided by WR. |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | All software can be used that is not sensible to steal the information |
| Where will the data and associated metadata, documentation and code be deposited? | ROS depackaging and then regular software for each type of data |
| Have you explored appropriate arrangements with the identified repository? | TBD |
| If there are restrictions on use, how will access be provided? | TBD |
| How will the identity of the person accessing the data be ascertained? | TBD |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with | Yes |

[4] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | TBD |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Creative Commons license CC-BY or CC-0 (according to the H2020 guidelines) |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | TBD |
| How long is it intended that the data remains re-usable? | TBD |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | TBD |
| How is the data quality assured? | TBD |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | - |
| Who will be responsible for data management in the task? | A person of SER will be assigned for this. |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | TBD |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | TBD |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | SER complies with GDPR. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.5 Data Management Plan of SAT

Authors: Han Hilbrands (SAT) and Jeroen Wolters (SAT)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | WP2 & WP6: T2.2, Development of smart weeder control / feedback mechanisms for autonomous weeding.<br><br>WP6: During the large-scale pilots performance evaluation of autonomous vehicles/machinery. |
| What types of data will be generated/collected? | Images of the field / crop / performance of the combination robot/implement. Videos of the field / crop / performance of the combination robot/implement.<br><br>Performance data of the combination robot and implement.<br><br>Crop & field data / task maps / prescription map / routing data. |
| What formats of data will the task generate/collect? | Mp4 video<br>JPEG photo<br>Table data (CSV, TXT)<br>Documents (Word)<br>Shape files (ISO XML, KML, GeoJson) |
| Will any existing data be used and how? If yes, what is the origin of the data? | No. |
| What is the expected size of the data? (in Gb) | Large volume of photo/video data will be collected and transferred to the cloud space of |

| | |
|---|---|
| | partner Agrointelli (AGI). But this data will not be stored by SAT.<br><br>Other log data is expected to be not more than 1 GB. |
| To whom might it be useful ('data utility')? | Robs4Crops partners, other researchers, farmers, dealers |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | Currently there are no identifiers used. |
| What naming conventions do you follow? | Folders named with company name, field name, date, crop. |
| How are search keywords provided that optimize possibilities for re-use? | No keywords are currently used. |
| How do you provide clear version numbers? | Date tag in filename or document is used for versioning. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | If needed a readme file per folder/company/field will be created. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[5]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | Most of the data will be collected by other consortium partners.<br><br>Data collected by SAT about the performance of the robots will not be open as this is related to personal data (farmer) and of commercial interest of SAT and its competitors. |
| How will the data be made accessible (e.g. by deposition in a repository)? | After a year we can upload the data to the shared folder of the Robs4Crops project. |

[5] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Al the data we generate are created within common programs. |
| Where will the data and associated metadata, documentation and code be deposited? | SAT has its own OneDrive from Microsoft. |
| Have you explored appropriate arrangements with the identified repository? | N/a |
| If there are restrictions on use, how will access be provided? | We can make an shared folder in OneDrive where we can give the partners access to. Or we can shared it in the Robs4Crops shared folder offered by WR.. |
| How will the identity of the person accessing the data be ascertained? | Own username and password from Microsoft. Or easy with the existing shared folder. |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes, the data can be opened in the common programs from Microsoft office |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | N/a |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | TBD |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | After the Robs4Crops project the open data can be used. |
| How long is it intended that the data remains re-usable? | As long as it is needed |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the | The non-proprietary data van be available on request, but the proprietary will be restricted. |

| | |
|---|---|
| project? If the re-use of some data is restricted, explain why. | |
| How is the data quality assured? | We collect our data by ourselves. Collected data will be internally reviewed. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | The data will be on our existing OneDrive from Microsoft, 502 euro per year. These costs are covered by SAT |
| Who will be responsible for data management in the task? | The people who generate the data are responsible |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Yes, if necessarily |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | The data is stored in the OneDrive from Microsoft, so the Microsoft server will make back-ups. |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes. |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | SAT complies with GDPR. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.6 Data Management Plan of TER

Author: Charles Duchemin (TER)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Define the perception of farmers of robotic system ->T1.1<br>Define the needs of farmers to build specifications -> T1.2<br>Define how the robotic system is used -> T1.2<br>Define what will be measured to follow the improvement of the project – T1.4 |
| What types of data will be generated/collected? | Discussions between partners – Lists of needs – Lists of metrics |
| What formats of data will the task generate/collect? | Text documents – Table documents - Pictures |
| Will any existing data be used and how? If yes, what is the origin of the data? | Yes: data collected before the start of the project by R4C partners for business studies was used to get a first perception of farmers for D1.1 |
| What is the expected size of the data? (in Gb) | Not relevant |
| To whom might it be useful ('data utility')? | People who deal with farmers' needs of robotic systems. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | A Readme.txt file describe how data are stored. |
| What naming conventions do you follow? | File:<br>YYMMDD_[T for Task/D for Deliverable] [Number of WP].[Number of task]-[Name of the file] {V[Number of the version]}<br>As example :<br>210412_T1.1-16 partners inputs V2<br><br>Folder:<br>[T for Task/D for Deliverable] [Number of WP].[Number of task] {Name of the Task/Deliverable}<br>As example :<br>T1.1 Co-design |
| How are search keywords provided that optimize possibilities for re-use? | Proposition in the readme.txt of keyword when |
| How do you provide clear version numbers? | With the "V[Number of the version]" at the end of each file |

| | |
|---|---|
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | readme.txt file |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[6]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | None. |
| How will the data be made accessible (e.g. by deposition in a repository)? | Data will be uploaded on R4C SharePoint (= OneDrive) and be available if necessary. |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Office software (Word – Excel…) |
| Where will the data and associated metadata, documentation and code be deposited? | R4C SharePoint (OneDrive) |
| Have you explored appropriate arrangements with the identified repository? | |
| If there are restrictions on use, how will access be provided? | Data from WP1 will be accessible to all R4C Partners |
| How will the identity of the person accessing the data be ascertained? | Not relevant |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. | Yes |

[6] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | None |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | N/A |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | During or at the end of the project |
| How long is it intended that the data remains re-usable? | 3-5 years |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | To be further worked out. |
| How is the data quality assured? | No specific process, only seriousness of Partners |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | TBD. |
| Who will be responsible for data management in the task? | TBD |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Not relevant |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data is stored in the TERs networks with backups and in the project's OneDrive/SharePoint accessible with credentials only. |
| Is the data safely stored in certified repositories for long term preservation and curation? | No |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | Partners must be careful with potential individual data, coming from one farmer or one field. TER will anonymize this kind of data as much as possible before sharing them. |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | No information will be related to an identified or identifiable living individual. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.7 Data Management Plan of ABE

Authors: Mike Roskam (ABE)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Directing, controlling and correcting the implements and the work done. |
| What types of data will be generated/collected? | WP2 5 6<br>Quality of work<br>Capacity ha/h<br>Row recognition<br>Video material for the operator |
| What formats of data will the task generate/collect? | Mp4 video<br>JPEG photo<br>Table data<br>Documents |
| Will any existing data be used and how? If yes, what is the origin of the data? | Existing experiences of the employees can be used if necessary |
| What is the expected size of the data? (in Gb) | If a robot is working it can easily use 5Gb a day just on video streaming and receiving commands to our expectation n |
| To whom might it be useful ('data utility')? | Consortium partners. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification | There are no DOI's used |

| | |
|---|---|
| mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | |
| What naming conventions do you follow? | Folders and files conventions |
| How are search keywords provided that optimize possibilities for re-use? | TBD |
| How do you provide clear version numbers? | Outline the approach for clear versioning |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | TBD |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[7]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | Not much data will be produced by Abemec.<br>For the data that is produced that still needs to be defined. |
| How will the data be made accessible (e.g. by deposition in a repository)? | The data is available on the Abemec server and can be added to the R4C drive if necessary |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Microsoft office<br>Mp4 player |
| Where will the data and associated metadata, documentation and code be deposited? | The data is available in the Abemec server and can be added to the R4C drive if necessary |
| Have you explored appropriate arrangements with the identified repository? | No. |
| If there are restrictions on use, how will access be provided? | |

[7] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| How will the identity of the person accessing the data be ascertained? | e.g user registration process |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | TBD |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | TBD |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | N/A |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | After upload |
| How long is it intended that the data remains re-usable? | For the lifetime of the repository |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | TBD |
| How is the data quality assured? | N/A |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | The costs for storing on Abemec servers fall under Abemec's budget. The sharepoint costs will fall under WR budget |
| Who will be responsible for data management in the task? | Abemec considers this a responsibility for the researching parties |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Not yet decided |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | TBD |
| Is the data safely stored in certified repositories for long term preservation and curation? | Not yet decided |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Not yet decided |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | no |

## 7.8 Data Management Plan of AGC

Authors: Christophe Aube (AGC)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | User's feedback<br>Product assessment<br>Product improvement<br>Path optimisation |
| What types of data will be generated/collected? | Data Collected :<br>- User's feedback (T1.1.)<br>- Smart Tools communication protocol (T2.3.)<br>- Path planning feedback (T4.3.)<br>- Performance assessment ( T6.3.)<br><br>Data generated:<br>- Uses of the robot and demonstration – videos images (T.1.5.)<br>- Measurements (T1.4.)<br>- Robot communication protocol (T2.3.) |

| | |
|---|---|
| | - Robot CAO<br>- Users' manual for the retrofitting kit<br>- Path generation ( T4.3.)<br>- Robot specifications (T6.1.)<br>- Video of demonstrations ( T6.5) (T8.2.) |
| What formats of data will the task generate/collect? | Type of data generated:<br>- CAO robot images<br>- Excel - Measurement metrics<br>- Schemes - Pdf<br>- System architecture<br>- Photos<br>- Word documents<br><br>Type of data collected:<br>- Information on the LSP Land Survey (Jison)<br>- Path optimisation (Jison)<br>- Anonymous questionnaire: Feedback of users and pilot sites. (Word, Excel) |
| Will any existing data be used and how? If yes, what is the origin of the data? | System plans and communication protocol, created by Agreenculture.<br>Time sheets<br>Photos and Videos |
| What is the expected size of the data? (in Gb) | To be defined |
| To whom might it be useful ('data utility')? | WP (1, 2,3,4, 5,6) partners<br>To LSP |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | NO metadata.<br>A study on the use and necessity of the DOI for AGC is being conducted. |
| What naming conventions do you follow? | Folders and files conventions<br>AGC_FOLDER_YEAR_MONTH_FILE TYPE_FILE Number<br>WP Name and task can be added |
| How are search keywords provided that optimize possibilities for re-use? | TBD |
| How do you provide clear version numbers? | According to the date and version number |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | Description, ownership, date etc.<br>readme.txt file<br>To be further defined |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[8]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | For commercial reasons and intellectual property some data is not openly available.<br>All other data shared by AGC can be used by the consortium members |
| How will the data be made accessible (e.g. by deposition in a repository)? | On the Wageningen Sharepoint |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Software for reading Jison Files<br>MS Word, Excel |
| Where will the data and associated metadata, documentation and code be deposited? | N/A |
| Have you explored appropriate arrangements with the identified repository? | N/A |
| If there are restrictions on use, how will access be provided? | Sent by AGC through encrypted file if necessary |
| How will the identity of the person accessing the data be ascertained? | User registration process<br>Identification code |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with | Yes |

[8] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| different datasets from different origins)? | |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | N/A |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | TBD |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | For some commercial information the embargo depends on the launching date.<br>Seeking patent: the period needed the for patent to be sent<br>Other documents: upon upload |
| How long is it intended that the data remains re-usable? | As long as it is needed |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | To be defined |
| How is the data quality assured? | Simulations, Measurements, Quality engineer |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | To be defined. |
| Who will be responsible for data management in the task? | WP leader |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | To be defined |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data stored in the partners' networks with backups, and in the project's SharePoint |
| Is the data safely stored in certified repositories for long term preservation and curation? | TBD |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | Data that need to comply to the GDPR<br>data from Third parties or future customers |

| | |
|---|---|
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Yes |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.9 Data Management Plan of AGI

Authors: Alea Scovill (AGI)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | AGI collects data about the Robotti in order to be able to continually develop and improve Robotti. |
| What types of data will be generated/collected? | WP1: Word, PDF, online cloud programs, Power point, office, teams<br>WP2-8: Robotti data, RGB images, videos, RTK position data, sensor data, mobile data, end user data, route planning information, all information collected with Robotti, cloud database, websites, Jira, Confluence, cloud applications, Word, PDF, online cloud programs, Power point, office, teams, CAD/Solid Works, coding (python, ROS, C, mySQL, etc), etc. |
| What formats of data will the task generate/collect? | AGI will use the most common data formats for the programs/purposes mentioned for data created in WP2-8. |
| Will any existing data be used and how? If yes, what is the origin of the data? | Yes, existing data will be used. RGB images, videos, RTK position data, sensor data, mobile data, end user data, route planning information, all information collected with Robotti, cloud database, websites, Jira, Confluence, cloud applications, Word, PDF, online cloud programs, Power point, office, teams, CAD/Solid Works, coding (python, ROS, C, mySQL, etc), etc. |
| What is the expected size of the data? (in Gb) | AGI has a system that can collect all of the data from all of the sensors, camera, etc into a ROSbag that can then be unpacked and explored. It records approximately 4GB/ 10 secs.<br>The total amount will depend on how many times collecting all of the data is performed.<br>Approximately 1 mil GB data. |

| | |
|---|---|
| To whom might it be useful ('data utility')? | Researchers, engineers, farmers, dealers, |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | Dropbox is searchable by windows and also online.<br>Confluence/Jira is searchable.<br>Coding is documented to some degree so it is searchable.<br>When images are annotated, there is a program that can search and cluster images. |
| What naming conventions do you follow? | Naming conventions for Dropbox folders, GitHub, Confluence, Jira, Solid Works, etc. |
| How are search keywords provided that optimize possibilities for re-use? | Tagging items, writing keywords in text, documenting code, annotating images,etc |
| How do you provide clear version numbers? | Dropbox and Github allows for version control. Can go back in time for each document/code. Otherwise typically if there are many versions of a document v1, v2 is used or persons initials and date is used. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | Website, databases, code, Robotti |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[9]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | Technical information about Robotti that is not proprietary will be open.<br>Non-proprietary data recorded on Robotti will be open to the project partner AGI is directly working with (Wageningen, Smart Agri Technology, ABEMEC, etc).<br>Personal data, proprietary Robotti data, Robotti ROS & website code and image annotations will not be open.<br>Proprietary data is defined as data that if another company/third party acquired the data, it could potentially harm AGI's business case. |
| How will the data be made accessible (e.g. by deposition in a repository)? | Email, PDF, dropbox, links to dropbox, etc<br>Potentially APIs if access to part of Robotti's database is needed. |
| What methods or software tools are needed to access the data? | Common programs |

[9] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| • Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | |
| Where will the data and associated metadata, documentation and code be deposited? | Jira, Confluence, GitHub, Dropbox |
| Have you explored appropriate arrangements with the identified repository? | N/A |
| If there are restrictions on use, how will access be provided? | Email, PDF, dropbox, links to dropbox, etc<br>Potentially APIs if access to part of Robotti's database is needed. |
| How will the identity of the person accessing the data be ascertained? | User name and password, email |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes, data can be opened into common programs |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | TBD |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | TBD |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | N/A |
| How long is it intended that the data remains re-usable? | As long as it is needed. |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | Proprietary data will be restricted.<br>Non-proprietary data will be available upon request. |

| | |
|---|---|
| How is the data quality assured? | We don't have any systematic double check process of the data (we are a SME).<br>When we find bugs in the code, we fix them.<br>When errors are found in the data, they are corrected. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Dropbox, Email, Confluence, Jira, Github, Amazon Web Services. Estimated 40k euro per year. Costs are covered by AGI and not by the project. |
| Who will be responsible for data management in the task? | The people who produce the data will be responsible for the data. |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Yes |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Databases are stored in AWS. Data is stored in Dropbox. The cloud services provide the data security as part of the price. |
| Is the data safely stored in certified repositories for long term preservation and curation? | TBD |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Yes |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.10Data Management Plan of FSH

Authors: Maja Zikic (FSH)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | The purpose of the data collection/generation in WP8 is to support communication, dissemination ecosystem building and exploitation activities. |
| What types of data will be generated/collected? | Communication & Dissemination material (incl. press releases, leaflets, audio-visual material, posters, images/photos)<br><br>Newsletter subscription data (i.e., name and last name, professional email)<br><br>Explotation data – data regarding partners backgorund and foreground IP |
| What formats of data will the task generate/collect? | Data will be collected in a xls, csv and Word format. |
| Will any existing data be used and how? If yes, what is the origin of the data? | Not applicable. |
| What is the expected size of the data? (in Gb) | The size of data to be collected in WP8 is approximately 100 GB. |
| To whom might it be useful ('data utility')? | Data will be used by Robs4Crops Communication & Dissemination Team and the project's Innovation Manager to support communication and exploitation activities. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | Communication and dissemination material (visual material, blog posts, etc.) are produced by Robs4Crops Communication & Dissemination Team. Newsletter subscription data are generated via subscription box on Robs4Crops website. Data concerning partners' exploitation intentions are produced via interviews.<br>We do not use Digital Object Identifiers. |
| What naming conventions do you follow? | Folders and files conventions |
| How are search keywords provided that optimize possibilities for re-use? | Not applicable. |
| How do you provide clear version numbers? | TBD |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | No metadata will be created in WP8. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly | The communication and dissemination material will be made openly available with the condition of appropriate crediting when referencing. |

| | |
|---|---|
| available as the default[10]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | The newsletter subscription dataset will be kept confidential in accordance with DGPR and R4C Privacy Policy (available on robs4crops.eu)<br><br>The exploitation data, regarding partner IP will be kept confidential, in accordance with GDPR and Business ethics. |
| How will the data be made accessible (e.g. by deposition in a repository)? | The data be made accessible by deposition in a repository and publication on official R4C channels (website and social media) in case of communication and dissemination material.<br>Other data (e.g., exploitation-related data) will be safely deposited in R4C OneDrive. |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | R4C OneDrive and R4C website for communication material.<br>Mailchimp for newsletter data. |
| Where will the data and associated metadata, documentation and code be deposited? | Data are deposited on R4C OneDrive. |
| Have you explored appropriate arrangements with the identified repository? | Not applicable |
| If there are restrictions on use, how will access be provided? | For communication & dissemination data – no restrictions, provided that R4C is appropriately credited as a source.<br>Newsletter subscription data is not available to anyone outside R4C Communication and Dissemination team.<br>Partners' exploitation intentions and associated data are available to R4C Exploitation Team and Innovation Manager. |
| How will the identity of the person accessing the data be ascertained? | Communication and dissemination material is openly available.<br>Newsletter data is only available on to R4C Communication & Dissemination Team. |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use | Not applicable. |

[10] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | Not applicable. |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Communication materials will be available for re-use with the condition of appropriate crediting of the content. In case any communication material from R4C is used after the project, the appropriate R4C logo must be used, as well the EC logo and the project funding phrase.<br>The materials will not be licensed in any other way. |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Embargo period is not applicable in our case.<br>The communication data will be made available after publishing. |
| How long is it intended that the data remains re-usable? | For the lifetime of the repository. |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | Not applicable for our data. |
| How is the data quality assured? | Comm and Diss manager who is also the Data Protection Officer (DPO) at FSH, and withing the Comm and Diss team we have a hierarchy of decision making to ensure that any data is double checked for quality before being released for public. In case we are publishing some more technical content for ex. related to activities and used cases, our Communication Manager communicates with the responsible partner to ensure that the information is correct. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | As non-profit NGO, FSH has secured G – Suite access with no charge, thus there are no direct costs for storing the data. |
| Who will be responsible for data management in the task? | FSHs DPO – Maja Žikić |

| | |
|---|---|
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Secure data repository of FSH – it can support long term persevation curration. Shared one drive of R4C. The DPO decides on all data protection issues. |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | FSH has an internal G -Suite repository (Google workspace), the access to which is limited to FSH employees only.<br>For the purposes of R4C FSH also uses the Shared one drive repository. |
| Is the data safely stored in certified repositories for long term preservation and curation? | The G -Suite repository can support long term persevation and curration od data. |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | For all dissemination, communication and exploitation activities of Robs4Crops, FSH has developed a privacy policy which can be found here[11].<br>In case any sensitive or personal data is being handled, FSH always asks for specific consent and agreement with the aforementioned privacy policy. |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | FSH pays special attention to always follow and comply to the GDPR.<br>As a non-EU country Serbia does not fall under the GDPR. However , the main piece of legislation currently regulating personal data protection in the Republic of Serbia, the Law on Protection of Personal Data (Official Gazette of the Republic of Serbia, No. 87/2018 (only available in Serbian here[12]), is fully compliant with GDPR. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | FSH has a data protection officer whose role is to ensure that their organization processes the personal data of its staff, customers, providers or any other individuals in compliance with the applicable data protection rules. |

[11] https://robs4crops.eu/privacy-policy/
[12]https://www.poverenik.rs/images/stories/dokumentacija-nova/zakoni/ZZPLnovembar2018/ZZPLnovembar2018.doc

## 7.11 Data Management Plan of TEY

Authors: Josep Vidal (TEY)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Collect data from work parameters of the sprayers to check correct operation of the implements according to the setpoint data. |
| What types of data will be generated/collected? | Sprayer ECU working parameters log |
| What formats of data will the task generate/collect? | .txt, .csv, .json, .xml |
| Will any existing data be used and how? If yes, what is the origin of the data? | No |
| What is the expected size of the data? (in Gb) | 0.1 Gb |
| To whom might it be useful ('data utility')? | For anybody capable to analyse correct operation of the combination of a hardware and a software of the sprayer which implements VRT technologies. This data alone can be quite out of context if you do not have specifications of hardware and software of the sprayer. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | No |
| What naming conventions do you follow? | TBD |
| How are search keywords provided that optimize possibilities for re-use? | VRT, SPRAYING, LOG |
| How do you provide clear version numbers? | TBD |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly | ECU data logs can be shared, since they are out of context, but I think they lose their value or future utility. |

| | |
|---|---|
| available as the default[13]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | |
| How will the data be made accessible (e.g. by deposition in a repository)? | To be determined in task 5.2 and 6.3 |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | Software capable to open .txt, .csv, .json, .xml files |
| Where will the data and associated metadata, documentation and code be deposited? | To be determined in task 5.2 and 6.3 |
| Have you explored appropriate arrangements with the identified repository? | Not yet. |
| If there are restrictions on use, how will access be provided? | TBD |
| How will the identity of the person accessing the data be ascertained? | TBD |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | There is no problem on re-use of it. |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | Not specified yet |

[13] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | After upload |
| How long is it intended that the data remains re-usable? | For the lifetime of the repository |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | No restriction in re-use log data |
| How is the data quality assured? | Not specified yet |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Not specified yet |
| Who will be responsible for data management in the task? | Not specified yet |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Not specified yet |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Not specified yet |
| Is the data safely stored in certified repositories for long term preservation and curation? | Not specified yet |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Not specified yet |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.12Data Management Plan of AUA

Authors: George Papadopoulos (AUA)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Data collection will be used in order to develop AI algorithms but also to test developed vision system in real world dataset. |
| What types of data will be generated/collected? | For WP2.1 Smart Spraying, we will collect images with apples and vineyard with infected diseases; we will collect video streams with vineyard canopy and apple orchards characteristics. Finally a geo-location will be captured for every dataset collection we already described. |
| What formats of data will the task generate/collect? | The formats of the expected data that will be utilized are:<br>- raw (image)<br>- ply (3D data such as point clouds)<br>- bag (a 'rosbag' file recording, containing video and gps streams)<br>- png (image)<br>- csv (tabular data suited to databases and spreadsheets)<br>- bin (depth matrix) |
| Will any existing data be used and how? If yes, what is the origin of the data? | We will use a dataset that is open access in Kaggle and moreover we will use images that are collected from our Robs4Crops partner Serrater in Spain and Pegasus in Greece. |
| What is the expected size of the data? (in Gb) | The dataset with the images only is expected to be 10-20 Gb. For the video size we estimate roughly 100-150 Gb. |
| To whom might it be useful ('data utility')? | Anyone who is interested in developing AI algorithms for vineyard or apple orchard vision tasks. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make | Questionnaire responses are uniquely identifiable by the Qualtrics software (each row represents a respondent). Digital<br>Object Identifiers are not used.<br>Sharepoint, where downloaded data is stored, tracks usage<br>and versions. |

| | |
|---|---|
| use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | |
| What naming conventions do you follow? | The following structure will be used for naming the folders and files:<br><br>R4C_YYYYMMDD_DISEASE/Type_of_diseases/ file_name.fileextension<br>R4C_YYYYMMDD_CanopyCharacteristics/Type_of_Canopy/ /file_name.extention |
| How are search keywords provided that optimize possibilities for re-use? | This parameter is to be defined later |
| How do you provide clear version numbers? | Outline the approach for clear versioning.<br>This parameter is to be defined later |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | This parameter is to be defined later |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[14]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | Data will be open to the public. The analysis will be reported on in R4C2021 Deliverable 2.2 |
| How will the data be made accessible (e.g. by deposition in a repository)? | This parameter is to be defined later |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | This parameter is to be defined later |

[14] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| Where will the data and associated metadata, documentation and code be deposited? | Preference should be given to certified repositories which support open access where possible.<br><br>The code will be deposited in Github or Bitbutcket platform. As far for the data and the associated metadata it will be defined later. |
| Have you explored appropriate arrangements with the identified repository? | This parameter is to be defined later |
| If there are restrictions on use, how will access be provided? | This parameter is to be defined later |
| How will the identity of the person accessing the data be ascertained? | e.g user registration process<br>This parameter is to be defined later |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Data is not being exchanged at this point. The questionnaire items (questions) are shared upon request for similar projects or Digital Innovation Hubs (DIHs). |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | This parameter is to be defined later |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | This parameter is to be defined later |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | This parameter is to be defined later |
| How long is it intended that the data remains re-usable? | This parameter is to be defined later |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | After the project is finished, the generated spatial data will be useful for other researchers.<br>This parameter is to be defined later |
| How is the data quality assured? | This parameter is to be defined later |

| ALLOCATION OF RESOURCES | |
|---|---|

| | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | This parameter is to be defined later |
| Who will be responsible for data management in the task? | For W.P 2.1 Smart Spraying, all the data management is done by AUA's personnel |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | This parameter is to be defined later |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | This parameter is to be defined later |
| Is the data safely stored in certified repositories for long term preservation and curation? | This parameter is to be defined later |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | Not applicable |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Yes |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.13Data Management Plan of EUT

Authors: Ferran Roure Garcia (EUT)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Provide information for navigation and localization algorithms. Mostly WP2, WP3 and WP4. |
| What types of data will be generated/collected? | Trajectories, GNSS data, images, 3D point clouds, and all kind of data related to the robot state and behaviour. |

| | |
|---|---|
| What formats of data will the task generate/collect? | Various formats, from Image to tabular data. All can be packaged in Rosbag format files. |
| Will any existing data be used and how? If yes, what is the origin of the data? | No |
| What is the expected size of the data? (in Gb) | 500Gb or more |
| To whom might it be useful ('data utility')? | People working on autonomous navigation methods |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | There is no strategy on this right now, due to the amount of data that will be produced. If there will be interesting captures, can be published following these standards. |
| What naming conventions do you follow? | Not defined yet. For Rosbags, capture time and description is the common standard: "210512-tractorData.bag" |
| How are search keywords provided that optimize possibilities for re-use? | Not defined yet. |
| How do you provide clear version numbers? | For navigation data, there is no versions, because data is unique. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | Not defined yet. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[15]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | A priori, all rosbags can be published openly if there is no sensible information. |

[15] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| How will the data be made accessible (e.g. by deposition in a repository)? | Online repository |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | ROS for depackaging, and then regular software for each type of data. |
| Where will the data and associated metadata, documentation and code be deposited? | Not defined yet. Regarding the code, a priori this will not be published. |
| Have you explored appropriate arrangements with the identified repository? | Not defined yet. |
| If there are restrictions on use, how will access be provided? | N/A |
| How will the identity of the person accessing the data be ascertained? | N/A |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | Not defined yet |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Creative Commons license CC-BY or CC-0 (according to the H2020 guidelines) |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Not defined yet. |
| How long is it intended that the data remains re-usable? | At least 15 years, for the lifetime of the repository. |
| Are the data produced and/or used in the task useable by third parties, | Yes |

| | |
|---|---|
| in particular after the end of the project? If the re-use of some data is restricted, explain why. | |
| How is the data quality assured? | There is no standard for data quality, besides our own criterion. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | Eurecat will cover the cost for those data stored in Eurecat's repositories. |
| Who will be responsible for data management in the task? | Eurecat |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Not defined yet. |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data stored in the Eurecat's networks with backups, firewall, etc... |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | There will not be personal information in data collected/generated by EUT |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

# 7.14Data Management Plan of UCPH

Authors: Soren Marcus Pedersen (UCPH)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | **Work package 7**<br><br>**Objective:** The objective of this work package is to conduct an overall integrated social and environmental impact assessment from the implementation of autonomous systems in the four pilot cases as outlined in WP1 and WP6 and to provide a web-tool for analysing economic performance at farm level.<br><br>University of Copenhagen is involved in the following subtasks:<br><br>Detailed objectives<br>Task 1: To assess social and environmental impact of selected systems in the four pilot cases<br>Task 2: To outline ethical concerns and barriers for selected autonomous pilot systems<br>Task 3: To conduct a cost benefit analysis of the most promising autonomous systems scenarios at the four pilot cases<br>Task 4 :To develop a web-tool for calculating economic and environmental performance of selected agri-robotic systems at farm level |
| What types of data will be generated/collected? | |

<table>
<tr><td></td><td>
<table>
<tr><th>Data</th><th>Type</th><th>Format</th></tr>
<tr><td>Task 1 : Data about environmental impact of selected technologies<br><br>Data about farmers and stakeholders perception of robotic systems</td><td>Text, table data</td><td>Excel files (.xlsx)<br><br>Word files</td></tr>
<tr><td>Task 2 Data about ethical concern among stakeholders/farmers</td><td>Tekst data from interviews</td><td>Word files<br><br>Excel files (.xlsx)</td></tr>
<tr><td>Task 3 : about cost structures and benefits forme each pilot case in<br><br>Data derived analysis algorithms.<br>Production costs and revenues, harvest data on production batches to estimate economic values of different scenarios</td><td>Tabular data with minimal metadata (column headings, variable names)</td><td>Word files<br><br>Excel files (.xlsx)</td></tr>
<tr><td>Task 4: Different measures for the Web-tool.<br>Data derived analysis algorithms.<br>Production costs and revenues, harvest data on production batches to estimate economic values of different scenarios</td><td>Tabular data with minimal metadata (column headings, variable names)</td><td>Excel files (.xlsx) and possible<br><br>comma-separated values (.csv) or delimited text with SQL data definition statements (.sql)</td></tr>
</table>
</td></tr>
<tr><td>What formats of data will the task generate/collect?</td><td>See above</td></tr>
<tr><td>Will any existing data be used and how? If yes, what is the origin of the data?</td><td>Yes, we will apply data from other sources such as public available data and data from other research projects, and if needed after necessary permissions.
<table>
<tr><th>Data</th><th>Type</th><th>Format</th></tr>
<tr><td>WP7: The already existing records of costs and benefits</td><td>Text and numeric data.</td><td>Word files<br><br>Excel files (.xlsx)</td></tr>
</table>
</td></tr>
</table>

| | |
|---|---|
| | recorded in other projects \| \| \| \| \| \| |
| What is the expected size of the data? (in Gb) | The size of the data can only be estimated, based on the current technology and knowledge of scenarios from each pilot case. Here is a forecast of the expected size of the data by WP:<br><br>WP7: Task 1-3 about 10 to 100 MB of production data for derivation of economic values and information about costs of production and revenues of production.<br><br>WP7 Task 4. In particular, for the webtool in task 4 we will refer to AUA about needed data for input files and output files in WP6. |
| To whom might it be useful ('data utility')? | WP 7 Farm economics, ethics, welfare economics, environmental impact data from WP7 to the webtool might be useful for farmers, consultants and other stakeholders and decision makers for assessing the economic and environmental and ethical impact. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | The keywords used will match the content of the data and the standards of the research field.<br><br>Published data can be find with e.g. DOI. |
| What naming conventions do you follow? | Folders and files conventions |
| How are search keywords provided that optimize possibilities for re-use? | We will apply clear task/deliverable numbers, dates and names for identifying the specific files |
| How do you provide clear version numbers? | Outline the approach for clear versioning. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | Suitable search keywords are used for each data set to optimize that potential users of the data can find the data set. The keywords used will match the content of the data and the standards of the research field. These include the farm survey and economic data, and the specific research questions. Tests are done to validate the usefulness of the keywords.<br><br>The new versions of open data will be numbered. The versions will be further identified with the date and with the key issues that are |

| | |
|---|---|
| | updated (e.g. number of observations), including notes on the potential corrections for the data. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[16]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | In principle, all data will be open and available among partners in the Robs4crops project.<br><br>Data from surveys and interviews in task 1-3 about personal matters will be kept confidential in line with the rules in GDPR.<br><br>If needed, some technical and biological data (collected from partners will be only used by UCPH and partners in Robs4crops for the purposes specified in this project. The information is not intended to be used by other parties and will be treated as confidential to protect the interests of the companies/partners providing the data. |
| How will the data be made accessible (e.g. by deposition in a repository)? | The data and/or metadata that is accessible will be shared via open-access data repositories and/or with data repositories of scientific journals with the aim of long-term preservation. |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | The phenotypic data sets are in .txt format or in Excel. We prefer to use textual data in the .txt format, supported by Metadata and documentation of the data. This allows any user to read the data without any specific software and to modify the data to the user's own preferences. |
| Where will the data and associated metadata, documentation and code be deposited? | Preference should be given to certified repositories which support open access where possible.<br><br>Data available among Robs4crops partners will be deposited at shared OneDrive manged by WR.<br><br>Data for UCPH will be deposited on UCPH servers. |
| Have you explored appropriate arrangements with the identified repository? | UCPH has certain network drives with special access |
| If there are restrictions on use, how will access be provided? | By contacting UCPH project partner |
| How will the identity of the person accessing the data be ascertained? | By using a user registration process with UCPH |

[16] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | To enhance interoperability, it is essential that the trait recording scheme and practice is described in detail e.g. in the publications, metadata and data documentation. Each research field has their own vocabulary and terminology which is well mastered by the Robs4crops partners, yet they will be always spelled out in the documents. |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | TBD |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | We do not expect to make any license for the outcome in WP7 – the webtool in task 4 is expected to be open access. **(to be discussed with AUA/WP6)** |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Data about environmental impact and indicators, cost and benefits will be made available for stakeholders and others as part of the finalized webtool in WP7 task 4 for calculating economic scenarios for the 4 pilot cases. Default input data will be made available as open access and without license – Data will be made available in compliance with GDPR rules and after agreement with Robs4crops partners that have delivered data. |
| How long is it intended that the data remains re-usable? | We expect that data is available for reuse for at least 5 years after the project ends. |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | As described above - data about environmental impact and indicators and cost and benefits will be made available for third parties as part of the finalized webtool in WP7 task 4 for calculating economic scenarios for the 4 pilot cases. Default input data will be made available as open access and without license – Data will be made available in compliance with GDPR rules and after agreement with Robs4crops partners that have delivered data. |
| How is the data quality assured? | By using review procedures from partners in Robs4crops and internal at UCPH according to rules and standards |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | UCPH has access to institutional funding schemes for open-access publishing at UCPH with public access to reports, working papers and other publications. The process of making the data and storing the data is part of the normal work at UCPH and covered by overhead costs.<br><br>UCPH also has agreements with different journals about publishing data and research results. |
| Who will be responsible for data management in the task? | UCPH in WP7 **task 1-3** and for **task 4** UCPH in collaboration with AUA. |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Data set sizes are relatively small (a few GB) which will fit standard network drive storage at either UCPH or AUA . Long term storage of data is expected for these network drives at either UCPH, AUA or Robs4crops servers |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | For data stored at partner UCPH the policy is to store everything on network drives and not on local computers. The institutional ICT department assures frequent redundant back-up for data shares across multiple locations. UCPH maintain up-to-date facilities with modern equipment as well as reliable & high-speed bandwidth for local data storage. Only authorized members of the project team have access to the data. |
| Is the data safely stored in certified repositories for long term preservation and curation? | Yes – see above |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | Most of the data collected by UCPH, does not carry any ethical or legal issues. |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | In case, personal data about economic questions or other personal information is collected - then we will keep the information confidential and comply with the GDPR concerning information provisions and access to personal data? |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.15Data Management Plan of UHOH

Authors: Galibjon Sharipov (UHOH), Dimitrios Paraforos (UHOH)

| Data Summary | |
|---|---|
| What are the purpose of the data collection/generation and its relation to a task in the project? | An application map, in terms of prescription, will be generated on the farm management system (FMIS) and will be implemented on the ISOBUS task controller. This will be carried out by developing a middleware that communicates with the farm controller. The middleware will consider recommendation data from the Perception System as well to make a final decision for the smart sprayer. To visualise all the information about the as-applied data, task data (as-applied rate, geolocation, section activated, etc.) will be recorded from the smart implements and be transmitted back to the FMIS. |
| What types of data will be generated/collected? | In WP 2, the following data types will be employed:<br>Prescription maps (Shapefile/ISOXML)<br>Task data (ISOXML)<br>Process-related data (To be defined) |
| What formats of data will the task generate/collect? | Shape files<br>ISOXML<br>Kml files |
| Will any existing data be used and how? If yes, what is the origin of the data? | For developing prescription maps, existing data such as field boundaries or historical data will be used. |
| What is the expected size of the data? (in Gb) | In each application:<br>Shape files /ISOXML ca. 0.5 Gb per file<br>Camera Data (not defined yet) |
| To whom might it be useful ('data utility')? | Farmers/Advisers will use the recorded task data to evaluate the efficiency of the operation and to consider its effects on the final yield. |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | - Both prescription and as-applied data are supposed to be available in the FMIS after being developed and recorded, respectively. They will be both stored as an active layer where a user discovers them by selecting the application/data/year of the task performed. So involved partners can have access to all data obtained.<br>- If there is another option such as Data Cloud Base (internal data server) employed, then both prescription and as-applied data can be stored in a specific folder with the application/data/year that makes it easily definable. |

| | |
|---|---|
| What naming conventions do you follow? | Follows the options in the previous chart |
| How are search keywords provided that optimize possibilities for re-use? | Application number, task name, implement type |
| How do you provide clear version numbers? | Version numbers might correspond to the number of application/crop type/year. These will be decided later when the data is recorded and options for storing are clarified. |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | It depends on what meta data will be involved in developing prescription maps. As mentioned earlier, the prescription maps might be developed by considering field boundaries, historic yield, weather forecast, etc. |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[17]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | - If an FMIS is used to visualise the recorded task data, then the data will be accessible to those who have access to the FMIS.<br>- If an internal data server is used to store the data, then they will not be public.<br><br>There won't be any restriction to share the data among partners of the project. However, making it public through the FMIS should be defined. |
| How will the data be made accessible (e.g. by deposition in a repository)? | See previous |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | - The FMIS should have the capability to present geodata and an ISOXML converter.<br>- The MATLAB algorithms to convert from ISOXML to raw data will be possible to be open source and should be defined to which extent. |
| Where will the data and associated metadata, documentation and code be deposited? | To be defined |
| Have you explored appropriate arrangements with the identified repository? | To be defined |
| If there are restrictions on use, how will access be provided? | To be defined |
| How will the identity of the person accessing the data be ascertained? | To be defined |

[17] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | The task data will only be used to evaluate the performance of the implements and how accurate the task is performed. In case of re-using the task data as metadata in the future, it can be exchanged upon request for similar projects. |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | Extensible markup language for ISOXML file |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Creative Commons license |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | To be defined |
| How long is it intended that the data remains re-usable? | As long as it remains available on the FMIS or the Data Server |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | Task-related data will be useable by third parties. |
| How is the data quality assured? | Own assessment |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | To be defined |
| Who will be responsible for data management in the task? | Team of Work Package 2: A middleware, will be developed, which will be responsible to apply the prescription maps and record the task data to/from the ECU (electronic control unit) of the smart implements. Responsibility for storing the created/recorded data comes with the option (FMIS or Data server). |
| Are the resources for long term preservation discussed (costs and potential value, who decides and | The data can be stored/used while the project lasts. Archiving and using it further when the project ends will be decided later. |

| | |
|---|---|
| how what data will be kept and for how long)? | |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | If the data created and recorded will be stored and visualised in the FMIS, then the responsibility of backup and recovery goes to the FMIS managers. |
| Is the data safely stored in certified repositories for long term preservation and curation? | To be defined |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | Data will be revealing farming strategies. End users' permission should be considered. |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

## 7.16Data Management Plan of LMS

Authors: Panagiotis Karagiannis (LMS)

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | Under WP4, data from end users related to their crops and the process executed within the fields will be collected |
| What types of data will be generated/collected? | WP4 Farming controller<br>• Field data for the modelling of the real field<br>• Process data for the process supervision<br>• Simulation data for the field (weather, soil conditions etc.)<br>• Status data exchanged among the different resources |
| What formats of data will the task generate/collect? | To be defined later |

| | |
|---|---|
| Will any existing data be used and how? If yes, what is the origin of the data? | No |
| What is the expected size of the data? (in Gb) | Several MBs - To be defined later |
| To whom might it be useful ('data utility')? | • End users – Status of crops<br>• Technical/Research partners – Technical developments<br>• Marketing partners – Dissemination material |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | To be defined later |
| What naming conventions do you follow? | • Report files and deliverables are stored in the R4C OneDrive/SharePoint folders<br>• ROSBag data files will be stored in the WP4 SharePoint folder once produced |
| How are search keywords provided that optimize possibilities for re-use? | Items and files are not tagged but can be found through the SharePoint search service. |
| How do you provide clear version numbers? | SharePoint has versioning service<br>In cases of different data sets the version will be declared in the name of the file |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | N/A |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly available as the default[18]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, | To be defined later |

[18] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| | |
|---|---|
| privacy-related, security-related, etc.). | |
| How will the data be made accessible (e.g. by deposition in a repository)? | • Publicly: Through publications made by LMS and other partners<br>• Within consortium: Through the OneDrive/SharePoint platform at WR |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | • For public access: PDF reader is required<br>• Within consortium: ROS installation and simple text editors are required |
| Where will the data and associated metadata, documentation and code be deposited? | To be defined later |
| Have you explored appropriate arrangements with the identified repository? | To be defined later |
| If there are restrictions on use, how will access be provided? | No restrictions once access provided |
| How will the identity of the person accessing the data be ascertained? | SharePoint's login process |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | Yes |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | N/A |

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | Depends on the restrictions from data owners |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | Depends on the restrictions from data owners |

| | |
|---|---|
| How long is it intended that the data remains re-usable? | As long as the SharePoint folder will be active |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | To be defined later |
| How is the data quality assured? | Data owners will be responsible for its quality. |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | This is covered by WR which offers the OneDrive/SharePoint folder |
| Who will be responsible for data management in the task? | TBD |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | No |

| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | This is covered by WR which offers the OneDrive/SharePoint folder |
| Is the data safely stored in certified repositories for long term preservation and curation? | This is covered by WR which offers the OneDrive/SharePoint folder |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | No |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | N/A |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | No |

# 8 Appendices

## 8.1 Partner template for the DMP

Authors:

| Data Summary | |
|---|---|
| What is the purpose of the data collection/generation and its relation to a task in the project? | |
| What types of data will be generated/collected? | |
| What formats of data will the task generate/collect? | |
| Will any existing data be used and how? If yes, what is the origin of the data? | |
| What is the expected size of the data? (in Gb) | |
| To whom might it be useful ('data utility')? | |

| FINDABLE | |
|---|---|
| How are the data produced and/or used in the task discoverable with metadata and identifiable? Refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers (DOI)? | |
| What naming conventions do you follow? | *Folders and files conventions* |
| How are search keywords provided that optimize possibilities for re-use? | *e.g. Tagging items (i.e. datasets, documents, code, etc.) with relevant keywords, that are automatically indexed by the search* |
| How do you provide clear version numbers? | *Outline the approach for clear versioning.* |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. | *e.g. Description, ownership, date etc.*<br>*e.g. Standard e.g. Dublin Core metadata standard*<br>*e.g. readme.txt file* |

| ACCESS | |
|---|---|
| Which data produced and/or used in the task will be made openly | |

| | |
|---|---|
| available as the default[19]? If certain or parts of datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.). | |
| How will the data be made accessible (e.g. by deposition in a repository)? | |
| What methods or software tools are needed to access the data?<br>• Is documentation about the software needed to access the data included?<br>• Is it possible to include the relevant software (e.g. in open source code)? | |
| Where will the data and associated metadata, documentation and code be deposited? | *Preference should be given to certified repositories which support open access where possible.* |
| Have you explored appropriate arrangements with the identified repository? | |
| If there are restrictions on use, how will access be provided? | |
| How will the identity of the person accessing the data be ascertained? | e.g user registration process |

| INTEROPERABLE | |
|---|---|
| Are the data produced in the project interoperable., that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)? | |
| What data and metadata vocabularies, standards or methodologies will the task follow to make the data interoperable? | *e.g. metadata format is compliant with standard formats (MARCXML, Dublin Core, and DataCite Metadata Schema)* |

[19] In case that specific datasets will be associated to scientific publications (i.e. underlying data), public projects reports and other raw or curated data not directly attributable to a publication, then the open access policy of H2020 is applicable.

| REUSABILITY | |
|---|---|
| How will the data be licensed to permit the widest re-use possible? | *e.g. Creative Commons license CC-BY or CC-0 (according to the H2020 guidelines)* |
| When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed. | *e.g. after upload, after Embargo period to publish or seek patents)* |
| How long is it intended that the data remains re-usable? | *e.g. at least 15 years, for the lifetime of the repository* |
| Are the data produced and/or used in the task useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | *After the project is finished, the generated spatial data will be useful for other researchers who…* |
| How is the data quality assured? | *Describe data quality assurance processes* |

| ALLOCATION OF RESOURCES | |
|---|---|
| What are the costs for storing the data. And how will these be covered? | *e.g. Long term storage, journal open access costs etc.*<br>*e.g project's budget* |
| Who will be responsible for data management in the task? | *e.g. data access committee* |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | |

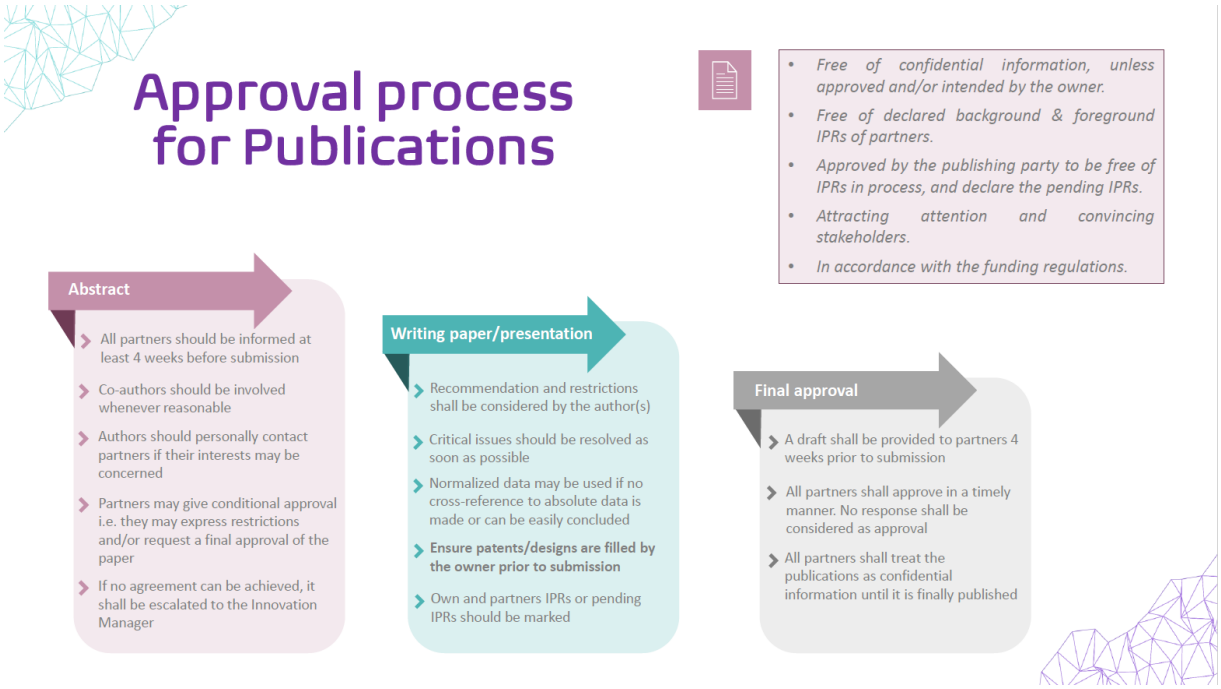
| DATA SECURITY | |
|---|---|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | *e.g, data stored in the partners' networks with backups, firewall; in the project's SharePoint accessible with credentials; Basecamp etc.* |
| Is the data safely stored in certified repositories for long term preservation and curation? | |

| ETHICS | |
|---|---|
| Are there any ethical or legal issues that can have an impact on data sharing? | *e.g. data from 3rd-party that didn't give an explicit consent, data that need to comply to the GDPR etc.* |
| Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data? Do you comply with the GDPR concerning information provisions and access to personal data? | |

| OTHER ISSUES | |
|---|---|
| Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones? | |

## 8.2 Extract of the Robs4Crops IP workshop no. 2

Below two relevant presentation slides concerning the protection of results and the approval process for publications:

## 8.3 List of all tasks per WP in Robs4Crops

WP1 Ecosystem building
T1.1 Ecosystems building
T1.2 Specfication of requirements
T1.3 Safety and regulations
T1.4 Definition of measurable metrics
T1.5 Cooperation with DIHs etc.,

WP2 Smart implements
T2.1 Smart sprayers
T2.2 Smart weeders
T2.3 Implement communication with the vehicle and the ERP
T2.4 Tractor Implement Management (TIM)

WP3 Autonomous vehicles
T3.1 Autonomous capabilities definition and coordination
T3.2 Retrofitting of conventional farm tractors into autonomous vehicles
T3.3 Sensing and mapping
T3.4 Localization and Navigation techniques to enable large-scale autonomous operations

WP4 Farming controller
T4.1 Digital world modeling for unified data representation
T4.2 Service based robot communication and interaction
T4.3 Farming simulation, optimization and forecasting
T4.4 Process execution, orchestration and control
T4.5 Cyber-security for data and system integrity

WP5 Tests
T5.1 Task define test scenario's
T5.2 Perform tests through the system architecture
T5.3 Feedback to pilot community

WP6 Large-scale pilots
T6.1 Pilot preparation and planning
T6.2 Evaluation protocol definition
T6.3 Deployment, feedback collection and performance assessment
T6.4 Scale-up demonstration activities to network of interest
T6.5 Scale-up demonstration activities to full ecosystem

WP7 Socio-economics and ethics
T7.1 Social impact assessment
T7.2 Ethical impact
T7.3 Social cost benefit analysis
T7.4 A web-tool for calculating the economic performance of robotic systems at farm level

WP8 Exploitation, communication, and dissemination
T8.1 Dissemination and Communication Strategies
T8.2 Communication and Dissemination Activities

T8.3 Exploitation model development and assessment
T8.4 Pilot Areas Business and IPR Management Plans

WP9 Management
T9.1 Manage overall direction of the project and follow-up on decisions
T9.2 Endure delivery of reports and updated implementation plan
T9.3 Financial administration
T9.4 Develop data management plan
T9.5 Plannning and preparation of project meetings
T9.6 Internal communication