Root2Res

Root phenotyping and genetic improvement for rotational crops resilient to environmental change

DATA MANAGEMENT PLAN

Timothy George (JHI)

Jean-Pierre Cohan (ARVALIS)

Funded by the European Union

UK Research and Innovation

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Root2Res has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101060124. Its work is supported by Innovate UK through the Horizon Europe Guarantee scheme Grant Agreement No. 101060124 and by the Swiss State Secretariat for Education, Research and Innovation (SERI) grant No. 23.00050.

The Root2Res Data Management Plan (DMP) describes the types of data to be produced or used in the project, how they will be documented and stored, who data will be accessible to, and how data will be shared and preserved for re-use in line with the FAIR (Findable, Accessible, Interoperable, Reusable) guidelines.

Deliverable Number	Work Package / Task
D8.3	WP8 / T8.3
Lead Beneficiary	**Deliverable Author (S)**
JHI and ARVALIS	Timothy George (JHI) Jean-Pierre Cohan (ARVALIS)
Beneficiaries	**Deliverable Co-Author (S)**
JHI, ADAS, ARVALIS, BOKU, ICARDA, UFZ, FZJ, FEUGA, FiBL	Fanny Tran (JHI) Charlotte White (WP1, ADAS) Katia Beauchene (WP2, ARVALIS) Eva Oburger (WP2, BOKU) Joanne Russell (WP3, JHI) Andrea Visioni (WP4, ICARDA) Doris Vetterlein (WP5, UFZ) Johannes Postma (WP6, FZJ) Oscar Bernardez (WP7, FEUGA) Sophie Thanner (WP7, FiBL)
Planned Delivery Date	**Actual Delivery Date**
28.02/2023	28.02/2028

Type of deliverable	R	Document, report (excluding periodic and final reports)	X
	DEM	Demonstrator, pilot, prototype, plan designs	
	DEC	Websites, patents filing, press & media actions, videos, etc.	
	DATA	Data sets, microdata, etc.	
	OTHER	Software, technical diagram, algorithms, models, etc.	

Dissemination level	PU	Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page))	
	SEN	Sensitive, limited under the conditions of the Grant Agreement	X

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, UK Research and Innovation (UKRI), European Research Executive Agency (REA) or Swiss State Secretariat for Education, Research and Innovation (SERI). Neither the European Union nor any other granting authority can be held responsible for them.

Project Number: 101060124

Project Acronym: Root2Res

Project Title:
Root phenotyping and genetic improvement for rotational crops resilient to environmental change

Data Management Plan
First version

Date: February 2023
DMP version: 1

INDEX

FIGURES

TABLES

1. The Root2Res Data Management Plan

The Root2Res Data Management Plan (DMP) describes the types of data to be produced or used in the project, how they will be documented and stored, who data will be accessible to, and how data will be shared and preserved for re-use in line with the FAIR (Findable, Accessible, Interoperable, Reusable) guidelines. The DMP has been compiled collectively in response to extensive consultation and iterative dialogue with all the WP Leads and other partners with respect to all project activities and outputs.

This DMP is a live document and will be revised during the lifetime of the project as project activities progress through the research data life-cycle. It is available to view online *via* the Project Management SharePoint, which can be accessed securely upon request to the project workspace manager: Pierre Rochepeau (p.rochepeau@arvalis.fr). The Data Manager (Root2Res.data.officer@arvalis.fr) with input from the Project Coordinators, is responsible for monitoring adherence of partners to the DMP supported by WP Leads and co-Leads, who are in turn responsible for verifying that relevant procedures are in line with the DMP. In addition, the DMP will be reviewed by the Executive Committee at the Annual General Meeting to ensure that it remains applicable to the requirements of the consortium.

2. Data Summary

2.1. Purpose of data generated within Root2Res and relation to project objectives

All data collected and generated within Root2Res will be managed securely throughout the project and made **Open Access** (with the exception of sensitive data such as personal descriptive data) following publication and appropriate exploitation according to the specific rules that are detailed in this document.

Overall, the data collected will help meet the primary objectives of the project, that are:

- **OB1**: Define, identify, and test root/rhizosphere ideotypes for a changing environment in crops common to rotational systems in Europe.
- **OB2**: Define and provide a complete set of tools to consider root traits.
- **OB3**: Identify, develop, and multiply germplasm and populations for phenotyping activities at different scales and use material to identify new candidate genes and markers connected to root traits and their plasticity, and enable novel pre-breeding germplasm, for all crops.
- **OB4**: Quantify plasticity of extended root phenotype for germplasm/populations identified in OB3 under a range of environmental conditions, including the identification of the relevant root trait, its interrelation with other characteristics (trade-offs) and the consequences for carbon sequestration.
- **OB5**: Actively engage with relevant stakeholders and disseminate new knowledge on the use of root and rhizosphere traits to develop resilience to

environmental change, while also exploiting the results to provide tools and products which improve the sustainability of agriculture with environmental change.

Given the multi-actors nature of the Root2Res consortium, the breadth of the database will benefit a wide audience ranging from farmers and farm managers, integrating farm organisations, plant breeders, advisors, and consultants, NGO/NPO, scientific/research community, policy makers industry, and the public at large.

2.2. Type and format of data generated within Root2Res

Root2Res data are being generated from several sources (Figure 1) and will be deposited in the Root2Res Data Repository and catalogued in a Data Register. These include:

- **datasets (field/controlled environment trials, genotypes, phenotypes, metagenomics, modelling)** and explanatory **metadata entry forms** (MEFs);
- methods, and **Standard Operating Procedures** (SOPs) for processes relevant to research and commercial-based activities;
- **new and historical data**, which inform WP operations and include relevant experiments carried out prior to the initiation of the Root2Res project as well as data from other EC-funded projects;
- **publications**, such as peer-reviewed papers, abstracts, practice abstracts, policy briefings, data highlights and key findings, images, and video shorts;
- **reports** written as public deliverables and milestones;
- **other resources** developed from social media-based activities, the implementation of the Modelling Toolbox, the Ideotype review being carried out by WP1, Expertise Data collected from the stakeholder groups etc...

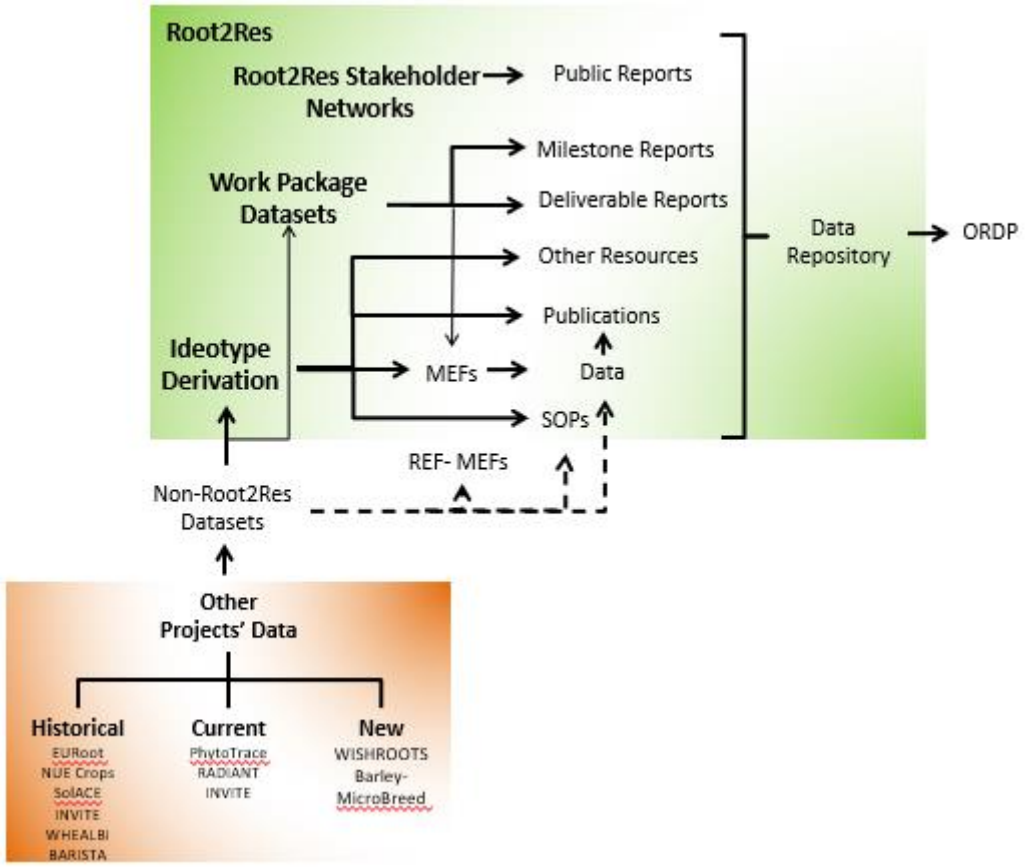

Figure 1. A graphical representation of the main types of data expected to be generated within Root2Res (green box) and other projects (orange box, only EU projects are currently listed). These include Reports Deliverables, Publications (of a wide variety including peer-reviewed, meeting, and grey literature), Datasets and explanatory MEFs, Standard Operating Procedures (SOPs, some of which will be developed into Practice Abstracts in the EIP-AGRI Common format) and Other Resources, including chosen KE outputs (e.g., abstracts, ppt presentations, trade reports etc.) and Practice Abstracts.

2.2.1. Datasets, MEFs and SOPs

Metadata Entry Form (MEF) and dataset templates will be provided to Work Package (WP) and Task Leaders to capture experimental metadata and raw data. A single MEF will accompany a single dataset for each Task or WP activity, which will be deposited in the Root2Res Data Repository once checked by the Data Manager (Root2Res.data.officer@arvalis.fr). These essential metadata explain the stored dataset to help ensure those downloading the information can interpret the data.

In addition, specific methods, or SOPs, used to gather the data are being produced and stored in the Data Repository. It is expected that some of these SOPs will be suitable for the production of Practice Abstracts.

2.2.2. Publications

Root2Res is a multi-disciplinary consortium with a variety of publishing conventions existing for each represented discipline. Hence author contributions may be less clear-cut in certain domains, which could pose difficulties in determining authorship, and consequently present potential for conflict.

When planning an experiment or designing a study, partners involved will discuss respective contributions on a case-by-case basis with full regard being paid to the guidelines provided by the targeted journal(s) (Table 1).

Table 1. List of main high impact journals targeted by the Root2Res project for publication of results.

Journals	
Agronomy	Nature
Annals of Botany	New Phytologist
Biology and Fertility of Soils	Plant and Soil
BMC Genomics	Plant Journal
Breeding Science	Plant Physiology
Environmental and Experimental Botany	Plant Stress
Environmental Microbiology	Plant, Cell Environment
European Journal of Agronomy	Field Crops Research
Frontiers in Fungal Biology	PloS
Frontiers in Plant Science	Rhizosphere
Fungal Biology	Science
ISME Journal	Science of the Total Environment
Journal of Environmental and Experimental Botany	Scientific Reports
Journal of Experimental Botany	Soil biology and biochemistry
Journal of Plant Interactions	TAG
Journal of Plant Nutrition and Soil Science	Trends in Plant Science

Other published outputs include abstracts of results, data highlights, policy briefs and key findings, images, and video shorts (see dissemination plan D7.2). All data will be uploaded annually from the start of year two, to the Data Repository alongside information and media repository for translation by WP7 and communication at stakeholder network workshops.

2.2.3. Background and Historical Data

Background or historical data are anticipated from several Root2Res partners. These data relate to relevant experiments carried out prior to the initiation of the Root2Res project and include data from other EC-funded and national projects. Only relevant historical data will be collated to be used by the consortium or will inform WP and stakeholder network objectives. The relevance of other EC-funded projects such as FP7 projects (EURoot, NUE Crops), H2020 projects (SolACE, INVITE, ADAPT, WHEALBI,

ClimBar, BARISTA, RADIANT) and ERC projects (Phytotrace) and national project (FSOV ARCHIRAC, BBSRC Rhizosphere by Design) will also be examined.

2.2.4. Deliverable Reports

Deliverables submitted to the EU portal SyGMa are considered project data and provide essential information about the overall project and its progress. These documents will be retained in the data repository and described in the Data Register (list of the data available on the Root2Res dataset).

2.2.5. Personal Data

Personal data are any information relating to an identifiable person who can be directly or indirectly identified, in particular by reference to an identifier such as a name, an email address, an identification number or a location. Personal data collected, within Root2Res, will be solely used to build a network of Stakeholders for contact by project partners. In such cases, personal data, such as name, location, workplace, and email address, will be collected by registering to attend one of our stakeholders' events. This specific type of data will be stored securely to minimise any risk of breaching the General Data Protection Regulation (GDPR, Annex 1) and the UK Data Protection Act 2018 and or the Federal Swiss Law on Data Protection. Data holders have been instructed to ensure, the data are cybersecure (*i.e.,* held in an encrypted (password protected) drive and/or PC) and avoid the use of USB drives, which can be easily misplaced. In compliance with these legislations, Root2Res will also ensure all stakeholders within the network consent to their data being added to our Data Repository and that they are informed of how their data will be used (Annex 1). In addition, mechanisms to allow access, removal and rectifications of the data held, upon request, have been put in place, e.g., personal data will be deleted from the local copy once it has been transferred into the data repository and once the local task is over.

The applicable national and EU data protection legislations, such as the GDPR, UK GDPR or the Federal Swiss Law on Data Protection, will be complied with. Specifically, Switzerland has received an adequacy decision of the European Commission (Commission Decision 2000/518/EC) and is as such considered to be a country which offers adequate protection for personal data.

2.2.6. Other Resources

Stakeholder's opinion, professional experience and other important subjective data collected at the stakeholder workshops, or via the website/social media and other social engagements will also be captured and made available with the same process as the other types of data.

2.3. Data origin

In fulfilling its purpose, Root2Res will generate new data and use existing data. The new data collected will be obtained in a variety of ways reflecting the appropriate methods and approaches listed in Table 2. In addition, Root2Res will obtain data from

a variety of sources. Where known, the origins of datasets are indicated in Table 2, this information will be updated as the data are obtained during the project.

2.4. Expected size of the data

The data collected and generated by Root2Res will be variable in size. A guide to the expected size of data files is given in Table 2.

2.5. Data utility

The primary purpose of each anticipated dataset is described on Table 2.

.

Table 2. Summary of anticipated data to be generated and collated by the project setting out the purpose, utility, origin, type and format, and estimated size of data.

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
Raw Data	Raw Data	Text, Numerical, Image, Video, Audio	Raw Data to feed into analysis in all WPs	Restricted access to Consortium members	All	.cvs	>10 TB
Imaging Data (OB1, OB2,OB4)	Microscopic studies	Numerical and Images	Studying the changes in root aerenchima after stress	Embargoed access after exploitation by Researchers	UVIGO	.xls, .jpg, .tiff	<10 GB
	CT-Data	Image	Quantify plasticity response in lab experiments	Embargoed access after exploitation by Researchers, Breeders, Modellers	UFZ	.mhd, .raw	>100 TB
	Root Nav heat maps	Image	Seedling root system heat maps of barley & faba bean		ADAS	.jpg	<1GB
	Segmented Root architecture	Image	Quantify plasticity response in lab experiments		UFZ	.mha	several TB
	Shoot phenotyping	Numerical, image	Phenotyping shoot traits under controlled and semi-controlled conditions in barley and faba bean		UVIGO	.xls. .jpg	<10 MB
	WinRhizo Scans	Image, Numerical	Quantify plasticity response in lab experiments and from field phenotyping of all crops		UFZ, ADAS, HUTTON	.tifs	<10 GB

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
Field and Phenotyping Data (OB1, OB2, OB3)	Field HTP	Numerical, Image	Quantify field response of phenotypes, genotypes, ideotypes	Embargoed access after exploitation by Researchers and Industry	ICARDA, ARVALIS	EXCEL, jpg and HTP raw data (RGB, LiDAR and multispectral)	several TB
	Physitron + HTP	Numerical, Image	Quantify field response of phenotypes, genotypes, ideotypes		ICARDA		several TB
	Field Phenotyping	Numerical	Quantify field response of phenotypes, genotypes, ideotypes	Immediate open access by Researchers, Breeders, Modellers	HUTTON, ARVALIS, ADAS, KWS, ICARDA, KIS	.xls	<1MB
	Rhizosphere Phenotyping	Numerical	Establish rhizosphere phenotyping for all crops	Embargoed access after exploitation by Root Phenotyping, Breeders, Modellers	HUTTON, BOKU	.xls	<5 MB
	Soil and Climate Data	Numerical	Establish environments for the trials hubs		ARVALIS, ADAS, KIS, ICARDA	.xls	<5 MB
	Root Phenotyping	Numerical, Image	Establish root phenotyping for potatoes, barley, durum wheat, faba bean under CE and Field conditions		HUTTON, ARVALIS, CNR, UVIGO, ADAS, ICARDA, KIS	.xls. .jpg	<10 GB
	Root and Shoot Plasticity	Numerical	Quantify plasticity response in lab experiments	Embargoed access after exploitation by Researchers, Breeders, Modellers	UFZ	.xls	<100 MB
	Root Exudates Phenotyping	Numerical	Quantify root exudates quantity and quality in crop rhizospheres		BOKU	.xls	<100 MB

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
Omics Data (OB2, OB3, OB4)	Gene expression	Numerical	Gene expression on roots focusing on genes involved in stress and symbiosis (eventually RNAseq analysis on selected samples)	Embargoed access after exploitation by Researchers, Modellers	CNR	FASTQ Sequencing and/or excel file	>10 GB
	Metabolomic data	Numerical	Finding the stress response-related metabolites in the crops		UVIGO	.xls and .raw or .mzml	>10 GB (50-100 MB per sample)
	Genotyping	Numerical	Establish crop genotypes for GWAS analysis	Embargoed access after exploitation by Breeders, Modellers	HUTTON, ARVALIS	FASTQ Sequencing;	5GB per crop
	Metagenomic	Numerical	Identify microbial component of phenotype	Open access after exploitation by Phenotyping, Breeders, Modellers	HUTTON,UNIDUN, FiBL, TEAGASC	FASTQ Sequencing; scripts (*e.g.*, .R), R objects (*i.e.*, .rds) and tab-delimited datasets (*e.g.*, .csv, .txt)	~5GB per MiSeq Library
	QTL Analysis Data	Numerical	Generation of QTL and candidate genes for phenotypes in all crops	Embargoed access after exploitation by Breeders Modellers	HUTTON, ARVALIS	.xls	1GB per crop

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
Modelling Data (OB1, OB3, OB4)	Model code/scripts	Code	Scripts and programming code to compile and run models	Open access after exploitation by Modellers and Researchers, and according to the rules set up in the description of background in the Consortium Agreement	FZJ, ARVALIS, ADAS, NEIKER	structured text files	<1000 MB
	Model input files	Code	To parameterize and run model code		FZJ, ARVALIS, ADAS, NEIKER	structured text files	<100 MB
	Model output files	Numerical	Data produced by models, tables and 3D structures, RSML, VTK		FZJ, ARVALIS, ADAS, NEIKER	structured text files, including CSV, and XML formats with schema	1 TB
	Model rendering	Numerical and Images	Visualization of models	Open access after exploitation by Researchers	FZJ	image and video formats	100 GB
Exploitation (OB4,OB5)	Exploitation of project results	Text, Image	Project result evaluation and exploitation strategy configuration	Restricted access to Consortium members	FEUGA	.doc, .xls, .pdf, .txt, .ppt	<10 GB
	Optimal IP protection	Text, Image, Numerical			FEUGA	.doc, excel, .pdf, .txt, .ppt	<10 GB
	Patentability analysis	Text, Image, Numerical			FEUGA	.doc, .xls, .pdf, .txt, .ppt	<10 GB
	Result evaluation methodology	Text, Image, Numerical			FEUGA	.doc, .xls, .pdf, .txt, .ppt	<10 GB
	State of the art	Text, Image, Numerical			FEUGA	.doc, .xls, .pdf, .txt, .ppt	<10 GB

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
Communication and Dissemination (OB5)	Conference presentations	Text	Repository of publications	Embargoed access after exploitation by Researchers and Industry	FiBL, all	raw, editable text files, .pdf	<1 GB
	Disseminatio n materials	Text, Image, Video	Material reuse and reporting for communication and dissemination		FIBL, FEUGA	raw, editable and exported image and video	1 TB
	Infographics	Image	Dissemination, communication and reporting		FiBL, all	.ia, .jpg, .pdf	<1 GB
	Podcasts	Audio	Dissemination and reporting		FiBL, all	.sesx, .omf, .mp3, .wav	<10GB
	Publications	Text	Repository of publications		FIBL, all	.pdf	<1 GB
	Tutorials	Video	Dissemination and reporting		FiBL, all	.prtl, .psd, .mp4, .mov, .avi	<10GB
	Policy briefs	Text	Repository of publications	Embargoed access after exploitation by Policy Makers Researchers and Industry	FiBL, all	raw, editable text files and .pdf	<1 GB
	Press releases	Text	Communication and reporting	Immediate open access to Publics	FEUGA	.doc, .pdf	<100 MB

Data Group (Obj)	Data/ Dataset	Data Type	Purpose	Utility	Origin	Format	Expected Size
	Stakeholder Engagement	Text, Numerical	Feedback from stakeholders	Immediate open access to Researchers, Breeders, Modellers	HUTTON, ARVALIS, ADAS, ICARDA, KIS	.xls	<1 MB
	Stakeholder profile database	Text	Engagement. Contact data from consent form and event platforms. Stratified stakeholders	Restricted access to Consortium members only	FEUGA	CSV	<1 MB

3. FAIR data

3.1. Making data findable, including provisions for metadata

The DMP is still evolving and will likely continue to do so over the life of the project. Root2Res is committed to making the data FAIR, beyond their original purpose, as described below.

3.1.1. Identifiability of data

Once data are uploaded to the project data repository, a Persistent Identifier (PID) will be assigned to each dataset and associated metadata. Once a dataset is made publicly available, a Digital Object Identifier (DOI) will be created either by the receiving journals or by uploading to a relevant open access repository such as Zenodo or Research gate to generate such an identifier (see Dissemination plan D7.2).

3.1.2. Discoverability of data

The data produced in the project will be discoverable with metadata associated with all inputs to the various data templates having a metadata sheet (MEF). Metadata will be created for the field-based, laboratory, and the simulated datasets and will include information on type of experiment, aim and hypothesis, experimental protocols, dates, and location where performed, by who with links to appropriate lab-book. Where appropriate there will be environmental information collected on the soil used (sheet with characteristics), plant species and genotypes used, environmental conditions of the experiment (temperature, daylength, humidity, soil moisture etc.) and any experimental treatments (water stress, nutrient stress etc.) and how they were imposed.

For field experimentation, any standard agronomic treatments (tillage, pesticides, fertilisation etc.) will be recorded. Any incidence of stress not imposed by the experimental treatments will also be recorded. Experimental randomisation and statistically valid experimental plans will also be presented. Within each category, the metadata describes the variable and a reference to the project's data collection protocol, which defines how the data have been collected.

3.1.3. Naming conventions, keywords, and versioning

Pre-defined "keywords" (supplied by the partners in the MEFs) will allow the databases and hence data repository to be searched. The level of complexity of the research to be made will be set up according to the need of each partner in charge of data analysis. The open-access repository chosen to publish the datasets once fully exploited, would also benefit from a search function to allow datasets to be found.

This is also aided by the naming conventions (name of files, their extensions, and their format) being standardized across all the WPs datasets. To achieve this, the naming convention from the ICASA Data Standard Dictionary will be used. This Dictionary

contains, for each variable, the short name and which unit of measure can be expressed. For example, the acronym of the wheat grain dry yield (0% moisture) is HWAM and it is expressed in kg ha^{-1}. Therefore, if in two different WPs two groups collect and report yield, they would do so using the same heading (HWAM) and same unit. In the case of variables that don't exist in the ICASA Data Standard Dictionary, project partners will agree on an exact and common nomenclature.

The Data Manager based at ARVALIS (Root2Res.data.officer@arvalis.fr) will only work with finalised versions of data files, regardless of their format. Thus, all versioning, audit trails and verification of data are subject of each partner's own internal system. However, if files deposited on the Root2Res server require updating or replacing, the data manager would issue "version iteration" with: 1) a timestamp of the change; 2) a brief description of the change; and, 3) the reason for the change. This information would then be associated with the metadata for the specific file(s).

A Data Repository versioning will be held in a table in the database and will list the version number, date and changes made since the previous version. An offline copy of each version will also be retained by the Data Manager (Root2Res.data.officer@arvalis.fr).

3.1.4. Metadata

Data collated on the Metadata Entry Forms (MEF) and to be included on the Data Repository will include:

- **Persistent Identifier (PID)** for the data set: specific dataset reference and name;
- **origin**: *i.e.,* historical or new data and any processing that may have occurred;
- **purpose**: description of the data to be generated, its purpose for the project, and its objectives;
- **complies with specific standards**: raw data and processed formats and/or types;
- **size or scale**: how many attributes or measures provided, whether these are raw or processed (calculated) parameters, and the number of 'values' received;
- **utility**: specify to whom the data will, or may be, useful.

Dataset-level metadata will be created upon deposition, allowing these resources to be found within their repositories provided they have a rich search syntax.

Table 3 shows the information/metadata that will be provided for each type of dataset.

Table 3. Metadata to be provided for each type of dataset collected by the project.

Type of Dataset	Information Required
Raw Data	Owners Address Owners Email Title Size Time (if video) Date of creation Location GPS Coordinates (if available) Main subject Text description Keywords (up to 5) Sample Type Species/Genotype Soil Treatment Environment/Weather Data
Imaging Data	
Field and Phenotyping Data	
Omics Data	
Modelling Data	Owners Address Owners Email Title Size Date of creation Location Main subject Text description Keywords (up to 5) Name and version of the model Set of parameters/calibration Input files of the model to be able to re-run a simulation Script to launch and run model Species/Genotype Soil/Environment Treatment
Exploitation	PID (file name) Title Keywords/Tags (*e.g.*, Root2Res, Horizon Europe) Author/Creator Subject Date of Publication Publisher (*e.g.*, Root2Res consortium) Version File size
Communication and Dissemination	

The final versions of all documents and deliverables that can be published open access and not deemed sensitive will be added to the project website. All previous versions will be stored on the project SharePoint platform, securely accessible by all partners.

3.2. Making data accessible

Datasets produced as part of the project will either be uploaded to an open access platform (*e.g.,* Zenodo) or published onto a specialised repository, (*e.g.,* Open Data Journal for Agricultural Research) following full exploitation. This may be that during the lifetime of the project, open-access publications will be delayed by 6 months to allow all possible routes for exploitation to be explored. All other data will be made available open-access in line with the Consortium agreement.

While most of the research data would be made freely available, some of the data may have some restrictions, due to commercially sensitive information or because it is part of a patent or because they are personal data protected by different regulation like GDPR. When this is the case, we will comply with the Exploitation management plan (see D7.2) and actions like copyrights on databases, softwares, *etc.* will be invoked.

Although different timelines for different datasets may exist, hence these timelines may need to be evaluated on a case-to-case basis. After the end of the project, partners will ensure that datasets are published within two years following project end.

3.2.1. Repository

A Root2Res Data Repository has been created. Its objective is to provide all project's partners (and agreed licensed partners) a secured database specifically designed to accommodate and manage all Root2Res data, methods and information gathered from project activities. The storage procedures and repository structure have been designed to ensure that datasets comply with FAIR principles (Findable, Accessible, Interoperable and Reusable) and GDPR of the EU (and equivalent for associated partners) for the internal needs of the project. The first version of the Date Repository has been delivered as D8.4. The next version, designed to fit some future needs of the project partners, will be included in the next version of the DMP.

The Root2Res Data Repository will store the datasets that partners will disseminate outside the project, according to the dissemination rules set up in the project. Access to these data won't be managed directly from the Root2Res Data Repository but via external open access repository adapted to each type of data (Root2Res website, Zenodo, etc.).

3.2.2. Data

In principle, all data and metadata will be made open access, following the gold model, but this will only be the case following the appropriate embargo period to allow data owners the opportunity to exploit the data (see Exploitation Plan, D7.2). However, some datasets will not be openly accessible at all due to copyright or other IP protection, and ethical concerns regarding personal data. Accessibility status will be specified in the Data Register including any reason why a specific dataset is not being made available to download *e.g.,* ethical; containing personal data; intellectual

property protection; commercially sensitive information; privacy related; and/or security-related.

Until the data and accompanying MEFs are deposited into an online repository, datasets can only be accessed by project partners. However, upon request to the Database Manager, non-Root2Res partners could be provided with data prior to the project end-date, provided express permission, and specification of the sharing term and conditions are given by the data-owner.

3.2.3. Metadata

For internal access (partners only), the metadata will be available in the data repository. For external access, they will made available via our website or other appropriate media. In such instances web browsers can be used to download .zip files. Freely available software can be used to open .zip files (*e.g.,* 7-zip*). The contents of the zip files *e.g.,* spread sheets and documents should be accessible using common software, such as Microsoft Office (proprietary), OpenOffice (free) or LibreOffice (free). Image files can be opened using Gimp*. Other software(s) required will be identified and specified once datasets have been received.

3.3. Making data interoperable

Where possible, standard vocabularies for all data types will be used to allow inter-disciplinary interoperability. Hence, the ICASA data standard (https://dssat.net/data/standards_v2/) was chosen to ensure the interoperability of the data. This data standard has been in use for more than 30 years, it has a well-established dictionary that is easy to access and read. Where standards do not already exist, Root2Res will provide a map to more commonly used ontologies (White et al. 2013 Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards. Comp. Elect. Agri. 96 (1-12).).

In addition, the Data Manager will ensure that data held on the Root2Res data repository are provided in commonly used data format and file types.

3.4. Increase data re-use

3.4.1. Licencing and Embargo Periods

The Data Register will specify any licencing requirements, embargo periods and other restrictions that may limit the re-use of the datasets and metadata to encourage the widest re-use possible. To further facilitate the process, the partners will use the dissemination and communication strategy developed in WP7 to promote the availability of the data. In addition, the Root2Res consortium includes members of focus groups national and European wide focus groups whose consortium members will further disseminate information and project findings where relevant.

3.4.2. Data Quality

Dataset will be quality assessed or 'verified' for compliance with experimental, methodological or survey data to ensure consistency and that they may be easily interpreted. In addition, checksums will be used to detect data errors introduced during data transmission and/or storage. This will help ensure data integrity (not authenticity).

All these measures will ensure that data are: **Findable**, metadata are assigned unique PIDs; **Accessible**, with clear protocols for querying or copying data; **Interoperable**, data can be interpreted and combined with other data; and, **Re-usable**, data is sufficiently well described so that they can be replicated or combined in future.

4. Other research outputs

The section of the plan will be developed as the character of other research outputs become clear.

5. Allocation of resources

5.1. Costs

The costs associated with making Root2Res data FAIR is covered within the Grant Agreement No. 101060124 funded by the European Commission, UKRI and Swiss Federal Government.

5.2. Data management within the project

- **Management of the DMP** is the responsibility of the Project Coordinator, Jean-Pierre Cohan (jp.cohan@arvalis.fr) and the Deputy Coordinator, Tim George (tim.george@hutton.ac.uk).
- **Collation of metadata from historical and new datasets to populate the Data Repository and inform the development of the DMP** is carried out by the Data Manager (Root2Res.data.officer@arvalis.fr).
- **Data Repository and management support** is provided by the Data Manager (Root2Res.data.officer@arvalis.fr), who will inform the signing-off of data for receipt into the Data Repository.

6. Data security

Data security will be administered as laid out in the Grant Agreement and Consortium Agreement. Sufficient local backups of data files exist to allow re-instatement following potential system failure. Backups will, in the case of sensitive data, be within encrypted volumes. In addition, the outward facing server at ARVALIS is held on a virtual machine, which allows rapid re-instatement following catastrophic failure. Transfer of data via webpages will be carried out through secure protocols. The transfer of sensitive data from Root2Res partners to ARVALIS is the responsibility of each individual partner.

7. Ethics

No ethical issues were identified in the building of the project proposal and therefore there are no identified impacts of ethical issues on data sharing.

Annex 1 Root2Res GDPR Guidelines

General Data Protection Regulation (GDPR) and the Data Protection Act 2018 came into force on May 25th 2018 to protect the personal information of individuals and give them greater protection and rights and is referred to as the "Data Protection Legislation".

The Data Protection Legislation applies to 'personal data', that is any information relating to an identifiable person who can be directly or indirectly identified in particular by reference to an identifier such as a name, an email address, an identification number, or a location.

The Data Protection Legislation applies to 'controllers' and 'processors'.

- A **controller** determines the purposes and means of processing personal data. If you are a controller, the Data Protection Legislation places obligations on you to ensure your contracts with processors are compliant.
- A **processor** is responsible for processing personal data on behalf of a controller. If you are a processor, the Data Protection Legislation places specific legal obligations on you. For example, you are required to maintain records of personal data and processing activities and you will have legal liability if you are responsible for a breach.

Personal Data within Root2Res

It is likely that the personal data collected within Root2Res will serve two purposes: 1. to build networks of stakeholders or farmers etc.; and 2. to conduct research, such as via focus groups and interviews. Regardless of how personal data are used, they must be stored securely. That is, access to such data should be granted on a need-to-know basis only. If it is an electronic copy, ensure it is cybersecure (i.e., held in an encrypted (password protected) drive and/or PC) and avoid the use of USB drives which can be easily misplaced. Please check with your respective institutions, that you are compliant in this regard to minimise any risks of a data breach.

More specific details are given below on how to deal with personal data depending on their uses.

Building Networks

In this specific case, personal data include name, home address, workplace, location, and email address. When collecting these data, please ensure the following to comply with the Data Protection Legislation. In particular:

- obtain explicit consent from the data subjects to add their personal data to your database and clearly explain to them how their data are going to be used;
- provide mechanisms to allow each data subject's details to be easily erased or transferred from the database or accessed or rectified upon request;
- do not contact data subjects for marketing purpose (such as advertising upcoming events) unless explicit consent was given to such marketing. Otherwise, that would constitute unsolicited marketing and may breach the Data Protection Legislation. Please note that a data subject's registration or

attendance to a specific event does not constitute or imply their consent to receiving any type of future communications.

Conducting Research

Please do not include any personal data in any report or deliverables as such documents could be made available to the public (such as via the Root2Res website and CORDIS, in the case of the deliverables). Failure to obtain consent to publish or disclose personal data which has been published or disclosed to a third party may constitute a breach of the Data Protection Legislation also.

In all outputs (whether Deliverables, Milestones, Database, SOPs, reports etc.) no personal data should be included in any manner without prior written consent from the data provider. Even when personal data is shared, it can only be shared according to the data sharing terms originally agreed by the data provider.

When carrying out interviews for a report or study, participants' personal details must be anonymised. This personal data should not be traced back to a specific individual. If the interviewee agrees to have their personal data posted online, a participant consent form will need to be signed by the data subject which clearly explains: (a) how the data are to be used by the interviewer; and (b) the legal obligations or/and objectives of the relevant project. It is extremely important that these steps are followed to minimise the risks of any data breaches or non-compliance with the Data Protection Legislation. A personal data breach means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data. This includes breaches that are the result of both accidental and deliberate causes. It also means that a breach is more than just about losing personal data. The Data Protection Legislation makes it clear that when a security incident takes place, you should quickly establish whether a personal data breach has occurred and, if so, promptly take steps to address it, including informing your national Information Commissioner's Office (ICO, in the UK)) or equivalent local regulator. Any notifiable breach should be reported to the ICO (or your local regulator) without undue delay, but no later than 72 hours after becoming aware of it.

Failure to comply with the Data Protection Legislation could result in the entire consortium being liable to fines that can be as high as 10 million Euros, or two (2%) per cent of your institution or businesses global turnover. Although it is unlikely that partners would be equally liable for one partner's data breach, it is the collective responsibility of the Root2Res project partners to ensure every partner is compliant with the Data Protection Legislation.

We strongly encourage you to contact your respective Data Protection Officer, to discuss your operations as part of Root2Res and to clarify: (a) if you are dealing with personal data; and (b) your obligations under this Data Protection Legislation.

Please remember that Data Protection Legislation compliance is a condition of our funding.

Please follow all these Guidelines whilst you are a member of Root2Res (including, but not limited to, writing a report, or organising an event).

If you are in any doubt regarding your compliance with the GDPR guidelines or adherence to the Data Protect Legislation, please do not hesitate to contact the Root2Res-Project Manager, Fanny Tran (fanny.tran@hutton.ac.uk) or the Coordinator Jean-Pierre Cohan (jp.cohan@arvalis.fr) .