



DIVERFARMING

Crop diversification and low-input farming across Europe: from practitioners' engagement and ecosystems services to increased revenues and value chain organisation



DATA MANAGEMENT PLAN

Deliverable D1.1

Version 1.0

Issue date: 21/10/2017

Author: Mathieu Kessler and Raul Zornoza





DIVERFARMING

This project has received funding from the *European Union's Horizon 2020 Research and Innovation Programme* under grant agreement No 728003

Copyright © DIVERFARMING Project and Consortium
www.diverfarming.eu





DIVERFARMING

Document summary	
Document title	Data Management Plan
Author	Mathieu Kessler and Raul Zornoza
E-mail of principal author	mathieu.kessler@upct.es / raul.zornoza@upct.es
Lead beneficiary	Universidad Politécnica de Cartagena
Deliverable No.	D1.1
Work Package	WP1. Project Management and Coordination
Dissemination type	Report
Dissemination level	Confidential, only for members of the consortium (including the Commission Services)
Deliverable due date	31/10/2017 (month 6)
Release date	31/10/2017
Copyright	© 2017 DIVERFARMING Project and Consortium

List of Diverfarming participants

Name	ACRONYM	COUNTRY
Universidad Politécnica de Cartagena (Coordinator)	UPCT	Spain
Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria	CREA	Italy
Agencia Estatal Consejo Superior de Investigaciones Científicas	CSIC	Spain
Università degli Studi della Tuscia	UTu	Italy
Asociación Regional de Empresas Agrícolas y Ganaderas de la Comunidad Autónoma de Murcia	ASAJ	Spain
Consorzio Casalasco del Pomodoro Società Agricola cooperativa	CCP	Italy
Arento Grupo Cooperativo Agroalimentario de Aragón	GA	Spain
Barilla G.E.R. Fratelli SPA	Bar	Italy
Disfrimur Logistica SL	DML	Spain
Universidad de Córdoba	UCO	Spain
Wageningen University	WU	Netherlands
Firma Nieuw Bromo van Tilburg	NBT	Netherlands
Industrias David S.L.U.	InDa	Spain
University of Portsmouth Higher Education Corporation	UPO	United Kingdom
Universität Trier	UT	Germany
Eidgenössische Technische Hochschule Zürich	ETH	Switzerland
Weingut Dr. Frey	WDF	Germany
University of Exeter	Exeter	United Kingdom
Pecsi Tudományegyetem - University of Pecs	UP	Hungary
AKA Kft	AKA	Hungary
Nedel-Market Kft	NMT	Hungary
Luonnonvarakeskus	Luke	Finland
Paavolan kotijuustola	PK	Finland
Polven juustola	PJ	Finland



DIVERFARMING

Executive summary

With the long-term view of increasing diversification and biodiversity in Europe (CAP objective) and fostering sustainable development of bioeconomy, the Diverfarming consortium has come together to develop and deploy innovative farming and agribusiness models. Diverfarming will increase the long-term resilience, sustainability and economic revenues of agriculture across the EU by assessing the real benefits and minimising the limitations, barriers and drawbacks of diversified cropping systems using low-input agricultural practices that are tailor-made to fit the unique characteristics of six EU pedoclimatic regions (Mediterranean South and North, Atlantic Central, Continental, Pannonian and Boreal) and by adapting and optimising the downstream value chains organization through executing 14 field case studies and 8 additional long-term experimental plots. This approach will provide: i) increased overall land productivity; ii) more rational use of farm land and farming inputs (water, energy, machinery, fertilisers, pesticides); iii) improved delivery of ecosystem services by increments in biodiversity and soil quality; iv) proper organization of downstream value chains adapted to the new diversified cropping systems with decreased use of energy; and v) access to new markets and reduced economy risks by adoption of new products in time and space. This data management plan will contribute to the management of research data collected or generated in the course of Diverfarming work and how these data is going to be stored, published, cited and made FAIR beyond the project life. Our goal is to meet the requirements of excellent scientific practice and to allow for accessibility, interoperability, reproducibility of Diverfarming research results.



DIVERFARMING



DIVERFARMING

Table of Contents

1. DATA SUMMARY	1
2. FAIR DATA	2
2.1. Making data findable, including provisions for metadata	2
2.2. Making data openly accessible	2
2.3. Making data interoperable	3
2.4. Increase data re-use (through clarifying licenses).....	3
3. ALLOCATION OF RESOURCES	4
4. DATA SECURITY	4
5. ETHICAL ASPECTS.....	4
6. SUPPORT IN DEVELOPING THIS DMP	4



DIVERFARMING

1. Data summary

Data collection is an essential part of the Diverfarming project which aims at assessing the real benefits and minimising the limitations, barriers and drawbacks of diversified cropping systems. This assessment relies on 22 field case studies (14 newly established and 8 long-term).

On the one hand, data will derive from collection of existing databases, outcomes from previous and ongoing European and national projects and literature review (sources: Zenodo, PubMed, SCOPUS, Web of Science, Sciencedirect, etc.) to identify the current situation on crop diversification, land management and value chain organization, with opportunities/limitations across pedoclimatic regions (WP2). These data will be integrated in Excel spreadsheets (.xlsx).

Data corresponding to history of the case studies and Communities of Practitioners will be also collected, such as classification of soil, climate, property, crop history, land management and experimental design of the field case studies (geolocation, plot characteristics, replications, crop diversification applied, management practices, sampling and analyses) (WP3, 4, 5). These data will be integrated in Excel spreadsheets (.xlsx).

New data will firstly derive from surveys and questionnaires and the decision making process, together with the associated output of the multicriteria decision analysis to tailor diversified cropping systems (WP2). These data will be integrated in Excel spreadsheets (.xlsx) to make them interoperable and reproducible.

The project will produce, for the field case studies, a set of experimental indicators related to the associated diversified cropping systems output: land productivity, crops quality and the delivery of ecosystem services (soil fertility, prevention of soil and water contamination, water availability, soil structure and quality, reduced greenhouse gas emissions, carbon sequestration, erosion prevention, above and belowground biodiversity and pest and disease control). These data will be integrated in Excel spreadsheets (.xlsx).

In this line, WP4 will produce new transcriptomic data of soil biological communities (taxonomic and functional), that will be stored in .bam and .fastg files, which can be operated with the open access R software.

Diverfarming will also generate observational images related to UAV-based optical methods for vegetation cover and diversity (WP5), stored as .jpg or .tiff. Spectral data derived from UAV will be stored in Excel files (.xlsx) and geotiff, .img, .asc and .shp after processing with GIS software; these data will be accessible through QGIS, ArcGIS or SAGAGIS. In this line, WP7 will generate new data related to upscaling, Geostatistical interpolation and GIS implementation of data generated in the field case studies, with extensions geotiff, .img, .asc, .shp (shapefile) or geojson, accessible through QGIS, ArcGIS or SAGAGIS, and extensions .kml, .kmz accessible through Google Earth. Simulation models will be applied in WP7 to assess long-term sustainability of crop diversification. Outputs will be finally produced in .xlsx to be fully interoperable.

WP6 will provide new data about agricultural value chain organization and optimization pathways by use of surveys/questionnaires and qualitative and quantitative analysis of primary data. These data will be integrated in Excel spreadsheets (.xlsx). Furthermore, STATA software will be used for modelling, with extension .dta, .log and .do. However, outputs will be delivered as pdf, .xls/.xlsx or .doc/.docx files.



DIVERFARMING

Data generated from the economic assessment will be gathered in Excel spreadsheets (.xlsx) to identify real economic benefits of diversified cropping systems along the value chain and the critical factor affecting the rational use of external inputs (WP8). WP8 will perform GAMS models whose inputs and outputs are .xls/.xlsx files.

Overall, the datasets generated by the project to fulfil its objectives will be diverse but mostly numeric and to a lesser extent text as a result of the surveys/questionnaires. It will mainly consist of observational data (geolocation data, experimental design data, surveys/questionnaires, value chain and economic assessment), experimental data captured by field and laboratory equipment (delivery of ecosystem services, genomics, UAV-based images for vegetation cover) and simulation data (models applied to assess long-term sustainability of crop diversification in WP7).

Regarding the size of the generated data, relatively many datasets will be produced, but none will be overly big, which will allow them to be collected on spreadsheets for easy accessibility except for the Next Generation Sequencing data (.bam and .fastg) and GIS maps (geotiff, .img, .asc, .shp, geojson, accessible through QGIS, ArcGIS or SAGAGIS, and extensions .kml, .kmz accessible through Google Earth).

Data derived from the implementation of Diverfarming will be useful for researchers in the field of sustainable agriculture and crop diversification along the entire value chain, as well as professionals in the field of crop and land management.

This DMP is not in conflict with the Finish Advisory Board on Research Integrity (TENK: www.tenk.fi/en).

2. FAIR data

2.1. Making data findable, including provisions for metadata

Due to the fact that the collected datasets will be diverse and belong to different fields, the metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository, which will be used. The depositors will create the metadata manually when uploading the datasets to Zenodo.

The repository assigns Handle/DOIs for persistent identification and citability of the dataset. The Project dataset identification follows the naming: Diverfarming_<WPno>_<serial number of dataset>_<dataset title>_versionnumber(date). Example: Diverfarming_WP2_1_datamining-woody-crops_v1(2017-10-06)

All datasets will be accompanied by a “readme file” in pdf format, which describes data factors and variables names and units. Keywords will be provided based on the expertise of the depositors who are familiar with the most commonly used keywords in their field.

2.2. Making data openly accessible

All datasets supporting publications will be made openly and publicly available on the OpenAire repository Zenodo upon acceptance of the manuscript for publication. Moreover, all datasets generated along the



DIVERFARMING

project work plan will be made public three years after completion of the Diverfarming project, regardless of manuscript publication.

Only data gathered by partners outside of the project work plan and protected by IPR, or inside the work plan but containing confidential information (e.g. related to personal interviews) will be kept closed for privacy reasons.

Data will be shared among members of work packages continuously as they are being generated, by using SharePoint online libraries associated to the O365 platform of tools for collaborative work (online platform OneDrive). Furthermore, when a dataset turns openly accessible, it will be announced on the web page of the Diverfarming project www.diverfarming.eu, where a link to the dataset on Zenodo will be available.

The vast majority of the datasets will be made available as spreadsheets (.xls/.xlsx format) and to a lesser extent as .doc/.docx or .pdf. Some transcriptomic datasets will be provided (WP4) in .bam and .fastg files, which can be operated with the open access R software. Some observational images related to UAV-based optical methods for vegetation cover and diversity (WP5) will be stored as .jpg or .tiff. Finally the GIS implementation of data generated in the field studies will generate geotiff, .img, .asc, .shp (shapefile) or geojson files accessible through QGIS, ArcGIS or SAGAGIS, and extensions .kml, .kmz accessible through Google Earth.

2.3. Making data interoperable

The depositors will strive to use metadata commonly and internationally used in their field of research, since no widely accepted specific metadata vocabularies exist. Data variables names and units commonly accepted by the international scientific community and international system of units, will be used. The vast majority of the datasets will be made available as spreadsheets (.xls/.xlsx format) and to a lesser extent as .doc/.docx or .pdf, so that they are fully used by potential users.

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

The project will make use of the CC-BY license. The particular license applicable to each dataset, will be decided on an individual basis, though the recommendation will be either CC-BY or CC-BY-SA.

The data will remain re-usable after the end of the project by anyone interested in it, with no access or time restrictions, since the Zenodo repository will be used. Each archived data set will have its own permanent repository ID and will be easily accessible. We expect most of the data generated to be made available without restrictions and only data sets subject to IPR and confidentiality issues will be restricted. Where this is going to be the case, agreements will be made based on the individual data sets. Requests for the use of the data by externals will be approved by the project consortium.

Regarding data quality, since the vast majority of datasets will support publication in peer-reviewed open access journals, the standards of quality (validation of the sample, replication and comparison with results of similar studies) will be met.



DIVERFARMING

3. Allocation of resources

There are no costs associated to the described mechanisms to make the database FAIR and long term preserved since we are using the Zenodo repository. The project coordinator has the ultimate responsibility for the data management in the project.

4. Data security

On one hand, the Zenodo repository takes care of security and backup of the deposited data. On the other hand, regarding the data-sharing process involved between partners during research, SharePoint Online libraries associated to the O365 platform will be used, and data protection services are provided by Microsoft to prevent the loss of data.

5. Ethical aspects

The ethical aspects related to the personal data collected in this dataset are addressed in the Ethics WP 11 deliverables:

1. D11.2. POPD - Requirement No. 1
2. D11.1. NEC - Requirement No. 2

6. Support in developing this DMP

This DMP has been created with the tool “Pla de Gestió de Dades de Recerca” (www.dmp.csuc.cat).