Data Management Plan for the SoilCare project

Editors: Simone Verzandvoort and Rudi Hessel

Report number:	**02**
Deliverable:	**D1.2**
Report type:	**Scientific Report**
Issue date:	**30/11/2021**
Project partner:	**Wageningen Environmental Research, the Netherlands**
Version:	**3.1**

WAGENINGEN
UNIVERSITY & RESEARCH

DOCUMENT SUMMARY

Project Information

Project Title:	Soil Care for profitable and sustainable crop production in Europe
Project Acronym:	SoilCare
Call Identifier:	H2020-SFS-2015-2b
Grant agreement no.:	677407
Starting Date:	01.03.2016
End Date:	31.08.2021
Project duration	66 months
Web-Site address:	www.soilcare-project.eu
Project coordinator:	Stichting Wageningen Research (Wageningen Environmental Research)
EU project representative & coordinator of the project:	Dr. Rudi Hessel - (rudi.hessel@wur.nl) +31 317 486 530
Project manager(s):	Erik van den Elsen (erik.vandenelsen@wur.nl), Simone Verzandvoort (simone.verzandvoort@wur.nl), Falentijn Assinck (falentijn.assinck@wur.nl)

Report Information

Report Title:	Data Management Plan for the SoilCare project
Principle Author(s):	Simone Verzandvoort and Rudi Hessel
Principle Author e-mail:	simone.verzandvoort@wur.nl
Deliverable Number:	D1.2
Version:	3.1
Work Package:	WP1
WP Leader:	Wageningen Environmental Research, Netherlands
Nature:	Public
Dissemination:	Document
Editor (s):	Rudi Hessel, Simone Verzandvoort
E-Mail(s):	Rudi.hessel@wur.nl simone.verzandvoort@wur.nl
Telephone Number(s):	+31 317 486530, +31 317 486500
Report Due Date:	01-03-2020
Report publish date:	30-11-2021
Copyright	©2021 the SoilCare project and Partners Copyright notice and disclaimer: http://tinyurl.com/soilcare-disclaimer

No.	Participant organisation name	Abbreviation	Country
1	Stichting Wageningen Research	WENR	Netherlands
2	University of Newcastle upon Tyne	UNEW	United Kingdom
3	Katholieke Universiteit Leuven	KUL	Belgium
4	University of Gloucestershire	UoG	United Kingdom
5	University Hohenheim	UH	Germany
6	Research Institute for Knowledge Systems	RIKS	Netherlands
7	Technical University of Crete	TUC	Greece
8	Joint Research Centre	JRC	Italy
9	University of Bern	UNIBE	Switzerland
10	Milieu LTD	MLTD	Belgium
11	Norwegian Institute of Bioeconomy Research	NIBIO	Norway
12	Bodemkundige Dienst van België	BDB	Belgium
13	Aarhus University	AU	Denmark
14	Game & Wildlife Conservation Trust	GWCT	United Kingdom
15	Teagasc	TEAGASC	Ireland
16	Soil Cares Research	SCR	Netherlands
17	Instituto Politecnico De Coimbra	IPC/ESAC	Spain
18	National Research and Development Institute for Soil Science, Agrochemistry and Environmental Protection	ICPA	Romania
19	University of Padova	UNIPD	Italy
20	Institute of Agrophysics of the Polish Academy of	IAPAN	Poland
21	Wageningen University	WU	Netherlands
22	University of Pannonia	UP	Hungary
23	Swedish University of Agricultural Sciences	SLU	Sweden
24	Agro Intelligence Aps.	AI	Denmark
25	Crop Research Institute	VURV	Czech Republic
26	University of Almeria	UAL	Spain
27	Fédération Régionale des Agrobiologistes de Bretagne	FRAB	France
28	Scienceview Media BV	SVM	Netherlands

Data Management Plan for the SoilCare Project

With contributions from:

Marius Heinen (WENR, Database 1)

Melanie Muro (Milieu, Database 2)

Francisco Areal and Regina Hansda (UNEW, Database 2)

Abdallah Alaoui (UNIBE, Database 2)

Guido Wyseure and Ioanna Panagea (KUL, Database 3)

Hedwig van Delden (RIKS, Database 4)

Version	date	Change
1	30/08/2016	Initial draft version
2	30/08/2016	Version to be revised by SoilCare consortium
2.1	17/01/2017	There was inconsistent use of the terms 'database' and 'dataset'. Distinguished 'database' at highest level and 'dataset' at one level lower. Comments Melanie Muro (WP7) processed - Data*set* with results of interviews with stakeholders for WP7 added to Data*base* 3 as Dataset 3.2. - Accepted proposed addition to the protocol on Data sharing, use and publication in SoilCare in article 'agreement on sharing data'. Questions on Dataset 5 to be handled by the WP6-leader.
2.2	28/08/2017	- Creation of the Soil care community on the Zenodo Platform mentioned - Upload to Database for WP7 added - Questions on Dataset 5 removed (to be handled in the next version).
2.3	13/10/2017	Merge and modification of Databases for WP2
2.4	28/02/2019	Updated version for M36 report
2.5	28/10/2019	Revisions WUR Data Competence Centre processed List of uploads to the Zenodo Data Repository updated
2.6	06/03/2020	Revisions to Databases 2 and 4 List of uploads to the Zenodo Data Repository updated
3.0	31/08/2021	Revisions to all Databases List of uploads to the Zenodo Data Repository updated
3.1	30/11/2021	Reordering of items in Annex 1 according to order in SyGMa and renaming of 1 item in response to request Project Officer in review meeting 23/11/2021

Contents

Preface

The present document is the final version of the SoilCare Data Management Plan (DMP). It identifies the main databases that have been developed within SoilCare, gives a basic description of these databases and provides information on the aim and use of these databases. The data repository associated with this DMP is the community 'Soil care for profitable and sustainable crop production in Europe' on the Zenodo-platform, accessible at https://zenodo.org/communities/soilcare/.

Introduction

Introduction to SoilCare data

Data management is an important subject in research projects, and was of particular importance for SoilCare as data obtained from various sources and from the 16 different study sites had to be made available to target audiences and within the project to e.g. Work Package (WP) leaders in appropriate formats. WP leaders needed these data to perform their tasks and to compare the different study sites. Furthermore, data generated by SoilCare may also be of interest to scientists and other stakeholders that are not part of SoilCare. As the SoilCare consortium recognises the need to make data available where this is possible, SoilCare has decided to join the Open Research Data (ORD) pilot. SoilCare will thus endeavour to make its data FAIR (findable, accessible, interoperable and reusable). As part of this effort, a DMP plan has been devoloped that covers the main datasets developed in SoilCare. This DMP, for each dataset, explains the following topics:

- the handling of research data during and after the end of the project
- what data have been collected, processed and/or generated
- which methodology and standards were applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

Within Soilcare, data management was an issue for leaders of study sites and for Work Package (WP) leaders. Development of the DMP was an interactive effort initiated by the project coordinator and executed by the coordination team and WP leaders. An updated version of the DMP has been provided to the EC with each periodic technical report.

Data management, property rights and publications

Although SoilCare endeavoured to make its data accessible to others, data management is a complex issue that also has links to e.g. property rights and publications. Some data that have been used in SoilCare are not owned by the SoilCare consortium, and therefore SoilCare partners are not at liberty to make these data available to others. In other cases, the SoilCare consortium plans to write (scientific) publications based on the data that have been generated, in which case the data themselves can only be made available to others after an embargo period (in case of an Open Access publication according to the Green or Hybrid Model). There are also companies involved in SoilCare who may want to use data generated within SoilCare for commercial exploitation. Finally, privacy issues may play a role too, for example for interviews that were held with stakeholders, or for soil quality information that can be traced back to particular fields and thus land owners. All these issues require careful consideration, also in the light of what is written about these topics in the Grant Agreement (GA) and the Consortium Agreement (CA). These documents were considered when writing the DMP in order to make sure that no conflicts between documents exist, and that the interests of the different SoilCare partners were not harmed. For these reasons, data cannot always be open to others. This fact is also recognised in the Guidelines on FAIR Data Management in Horizon

2020, which state that data should be 'as open as possible, as closed as necessary'. SoilCare fully subscribes to this statement.

As specified in Article 29 of the Grant Agreement, there are three types of data that should be made O.A. Apart from the data specified in the DMP, the OA requirement also applies to scientific publications and the data underlying such publications. These latter two types are not covered in the DMP, but some of these data are made available via the Zenodo repository[1].

In Annex 1, relevant text about ownership of data, data sharing, and publications is assembled from the GA and the CA. To that an explanation is added as to how this text was applied in practice within SoilCare.

Which datasets

SoilCare generated several databases, as mentioned in the DoA (p26 of part B):

'The following types of data will be collected: Data on the classification of soil, climate and land use (including soil property, quality and productivity data) will be collected to characterise CS (cropping systems) across Europe (WP 2). Data derived from existing soil quality indicator systems and field trials will be compiled in a database to identify the most cost-effective indicators of soil threats, soil functions and land potential (WP2). The development of a database with monitoring results from the Study Sites is included in WP5, and spatial data used and generated in the project, especially in WP6, will be safeguarded to ensure that they remain available after the end of the proposed project, ...'

To this can be added a database with information obtained through interviews with stakeholders (WP3, WP4, WP7).

Hence, 4 databases are distinguished:

1. Database with data on effect sizes of several crop husbandry practices or SICS obtained from published meta-analysis studies, based on global data (WP2)

2. Database with information obtained from stakeholders (WP3, WP4, WP7)

3. Monitoring database (WP5)

4. Database of information on cropping systems at European level (WP6)

A description of these 4 databases is given in the following sections.

---

[1] https://zenodo.org/communities/soilcare

Database 1: Meta-information on cropping systems in Europe (WP2)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The purpose of this database is to assemble results from published meta-analysis studies on the effects of soil improving cropping systems (or crop husbandry practices). These were used to analyse which soil improving cropping systems (SICS) would be suitable in which parts of Europe. Earlier versions of these data formed the basis of the SoilCare deliverable D2.1 which included a SWOT analysis and from which a pre-selection of SICS for further study in SoilCare was made. • Relation to the objectives of the project The SoilCare project aims to identify soil improving cropping systems, which consist of a particular crop selection and rotation and an optimal combination of inputs, techniques and management, as function of soil type (soil threat), climate, and socio-economic conditions. The database was needed to achieve the first aim of SoilCare, namely to review which CS can be considered soil-improving, to identify current benefits and drawbacks, and to assess current and potential impact on soil quality and the environment. • Types and formats of data generated/collected The data is in the form of tables with meta-information on cropping systems. • Existing data that is being re-used (if any) All data come from the literature and refer to the main outcomes of published meta-analysis studies regarding the effectiveness of SICS. These data are based on global meta-analysis publications, and therefore not based on conditions in Europe solely. • Origin of the data Reviews published in the literature. • Expected size of the data (if known) Several MB of Excel spreadsheet(s). • Data utility: to whom will it be useful Data are useful for all WPs in SoilCare, since they provide information on existing cropping systems in Europe and their impacts on soil quality and crop productivity. It has been of particular use for WP5 (monitoring and analysis of results). The results of the analysis performed on the database has been used in WP4 of SoilCare for the assessment of SICS. The database assembles data from a range of sources and will therefore be useful to everyone interested in the effects of SICS on soil quality.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	The database is made findable through the metadata standard applied by the Zenodo data repository (the Dublin Core Metadata Standard).

	A Digital object Identifier (DOI) has been attributed to the database as the standard identification mechanism when uploaded to the Zenodo platform.
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so The database has been made available in the Zenodo Research Data Repository subject to an embargo of 6 months after finalization and licence conditions. • How the data will be made available Via the data repository. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? MS Excel or another spreadsheet program (that can import a .xlsx file) is needed. • Where the data and associated metadata, documentation and code are deposited In the original sources of the data and partly in the Database. • How access will be provided in case there are any restrictions The data that are included are without restrictions, because authors refer to the original publications.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. The database uses acknowledged names, symbols and units for target variables, domains of interest and statistical parameters in soil science and agronomy. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? Standard vocabulary is used from soil science and agronomy. In addition to the explanation of variables that are in the database, there is also a [SoilCare glossary](#) that explains how terms are used within the SoilCare project.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible The database is made available under the CC-BY licence ('Attribution'). This allows reuse, but obliges the user to credit the sources. • When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed As WP2 is writing scientific publications about its work, an embargo of 7 months is set from the date of completion of the database. • Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

		Data from the original sources are useable for third parties, since they are also available in the literature. Data resulting from the meta-analysis will be made available after publication of the results under the restrictions set by the licence selected. • Data quality assurance processes Only peer reviewed literature sources were used. The quality assurance is derived from these sources. • Length of time for which the data will remain re-usable Data will remain re-usable for the lifetime of the Zenodo repository (currently >20 years).
3.	Allocation of resources	• Costs for making your data FAIR. Describe how you intend to cover these costs Costs are low, as all data are already available, and documentation is part of the work of WP2. • Responsibilities for data management in your project This database is developed and managed by WP2 of SoilCare. • Costs and potential value of long term preservation No costs are foreseen, since archiving with Zenodo is free up until 50GB per dataset, and the size of the database will not exceed this threshold. The potential value of the database is difficult to monetize, but the database is foreseen to be a valuable asset for those who are interested in the relationships between cropping systems, soil quality, environment and yield. This potentially includes many farmers, agricultural advisors, companies and policy makers throughout Europe.
4.	Data security	• Data recovery as well as secure storage and transfer of sensitive data Provisions for data security on the Zenodo platform apply. These concern the access to the CERN Data Centre, the servers, the network, the data storage, the application and the staff from Zenodo.
5.	Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former No ethical issues foreseen for this database.
6.	Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any) The policies for data management apply of the institutes from the SoilCare partners having contributed to the database. For Wageningen Environmental Research this is the WUR data policy for storage of data during a research project (see https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Data-policy/Storage.htm).

Database 2: Information obtained from stakeholders

Dataset 2.1 Information on social factors relevant for adoption (WP3)

DMP component	Issues to be addressed

1. Data summary	• Purpose of the data collection/generation
	The purpose of this dataset is to collect information on social factors on social acceptability influencing the adoption of soil-improving innovations in the SoilCare study sites.
	
	• Relation to the objectives of the project
	The dataset informs on:
	    o How social capital could affect uptake of SICS;
	    o What stakeholders thought were the main causes and solutions to declining soil quality;
	    o How trust might affect uptake of SICS.
	
	• Types and formats of data generated/collected
	The data is in the form of interview transcripts. All data is stored in Nvivo files.
	
	• Existing data that is being re-used (if any)
	None.
	
	• Origin of the data
	The data are derived from workshops and interviews in the study sites.
	
	• Expected size of the data (if known)
	25.9 MB
	
	• Data utility: to whom will it be useful
	Other WPs .

2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset described is not accessible based on the conditions of privacy granted to stakeholders collaborating in the research. Since there is no explicit consent from participants of the workshops and interviewees, therefore, access to individual stakeholder data will not be provided. • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? The dataset described is not accessible based on the conditions of privacy granted to stakeholders collaborating in the research. • Standards for metadata creation (if any). No particular standards used.
2.2 Making data openly accessible	• The dataset described is not accessible based on the conditions of privacy granted to stakeholders collaborating in the research. However, the research output generated with the data is available in Deliverables 3.3 and 3.4 of the SoilCare project. • How the data will be made available

	Since there is no explicit consent from participants of the workshops and interviewees, and guarantees on privacy protection were given to interviewed stakeholders, individual stakeholder data will not be made available. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? The data has been analysed in NVivo software for which license is required. The licence is updated every year. • Where the data and associated metadata, documentation and code are deposited Data, metadata, documentation and code are deposited in a repository of Newcastle University. • How access will be provided in case there are any restrictions Since there is no explicit consent from participants of the workshops and interviewees, therefore, access to individual stakeholder data will not be provided.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. No vocabularies, standards and methodologies from the social sciences domain were required to facilitate interoperability with other project results. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? As stated under the previous point.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible Since there is no explicit consent from participants of the workshops and interviewees, based on the conditions of privacy granted to stakeholders collaborating in the research, access to individual stakeholder data will not be provided, and the data will not be licenced. • When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed Not applicable. • Whether the data produced and/or used in the project is 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. • Data quality assurance processes Data quality is assured through the following approach: Pilot interviews were conducted to test the clarity of the question wording and questions were subsequently refined. Interviews were conducted with trained personnel. Sampling tended to be purposeful and, for the qualitative interviews, data continued to be collected until theoretical saturation was reached. Data were analysed qualitatively in Nvivo, except for Q-methodology quantitative data, which were analysed using KenQ that scans the data before analysis to check for errors. • Length of time for which the data will remain re-usable Not applicable.

3. Allocation of resources	• Costs for making your data FAIR. Describe how you intend to cover these costs No costs are foreseen. • Responsibilities for data management in your project This dataset was developed by WP3 of SoilCare. • Costs and potential value of long-term preservation Long term preservation will contribute to the record of information on role and trust of stakeholders in the adoption of cropping systems in the study sites.
4. Data security	• Data recovery as well as secure storage and transfer of sensitive data Provisions for data security on the repositories of Partner 2 (Newcastle University) apply.

5. Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former.
	Ethical standards and guidelines have been applied to the collection, processing and storage of data about persons, as described in Part B of the DOA. These include:
	    o Identification and recruitment of participants through intermediaries
	    o Informed consent procedures
	    o Procedures for data collection, storage and protection conform national and EU legislation
	    o No existing (secondary) data will be sought that is not already in the public domain as part of this project
	    o Relevant authorisations will be sought for research permits were required to conduct workshops and interviews.

6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any) All activities for data collection, storage and protection for this dataset comply with national and EU legislation, in particular with the GDPR.

Dataset 2.2 Information on policy factors relevant for adoption (WP7)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The purpose of this dataset is to collect information on the role of policies and policy instruments in facilitating or hindering adoption of soil Improving cropping systems (policy mechanisms, outputs, outcomes, impacts, factors for success and failure) as well as adoption factors more broadly. • Relation to the objectives of the project The dataset informs on: 1. Opportunities and bottlenecks in current policy to enhance soil quality and land degradation and more specifically to facilitate the adoption of soil-improving techniques; 2. Factors hampering and enabling adoption of Soil Improving Cropping Systems as a basis for developing actions to increase the wider uptake of these, that can be facilitated by policy. • Types and formats of data generated/collected Two data sets are provided: 1. The first data set is in the form of policy inventories (excel files); 2. The second data set is an excel database containing interview data and information collected through EU and study site level stakeholder interviews and workshops. • Existing data that is being re-used (if any) None. • Origin of the data The data are derived from European- and site-level desk studies, stakeholder workshops, and interviews. • Expected size of the data (if known) 1 MB • Data utility: to whom will it be useful Other WPs and stakeholders in the study sites. Beyond the project, the database will be useful to anyone interested in EU and national policies relevant for the protection, maintenance, and improvement of soil quality.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset is made findable through the metadata standard applied by the Zenodo data repository (the Dublin Core Metadata Standard). • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? Digital object Identifiers (DOI) have been attributed to the items of the database as the standard identification mechanism after uploading to the Zenodo platform.

	• Naming conventions used SOILCARE_WP_7_[Name of deliverable]_[VERSION]_dd-mm-yy. • Approach towards search keyword None. • Approach for clear versioning Use of clear numerical identifiers. • Standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how No particular standards were applied.
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so Only data not referring to individual stakeholders have been published, based on the conditions of privacy granted to stakeholders collaborating in the research. • How the data will be made available Via the data repository and dissemination of reports aggregating/summarising the data through channels created by WP8. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? No specific software tools needed. • Where the data and associated metadata, documentation and code are deposited Partly in the Project's Data Repository. • How access will be provided in case there are any restrictions In case restrictions apply based on privacy guarantees to interviewed stakeholders, access is not provided.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. Standard vocabularies, standards and methodologies from the social sciences domain are used. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? As stated under the previous point.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible The database is made available under the CC-BY licence ('Attribution'). This allows reuse, but obliges the user to credit the sources.

		• When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed
		According to the embargo set for the Project Data Repository, i.e. 6 months after publication of the associated project deliverable (7.1 and 7.2).
		• Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why
		Data produced and/or used from this dataset are useable by stakeholders in the study sites and by other target groups of the project. Reuse of data referring to individual stakeholders is restricted.
		• Data quality assurance processes
		Data quality is assured by social science standards applied in WP7.
		• Length of time for which the data will remain re-usable
		Data will remain re-usable for the lifetime of the Zenodo repository (currently >20 years).
3.	Allocation of resources	• Costs for making your data FAIR. Describe how you intend to cover these costs
		No costs are foreseen, since archiving with Zenodo is free up until 50GB per dataset, and the size of the database will not exceed this threshold.
		• Responsibilities for data management in your project
		This dataset has been developed by WP7 of SoilCare. WP8 has disseminated part of the data.
		• Costs and potential value of long-term preservation
		Long term preservation will contribute to the record of information on the role policies, policy instruments, and practices (and combinations thereof) inhibiting or facilitating the adoption of soil improving cropping system.
4.	Data security	• Data recovery as well as secure storage and transfer of sensitive data
		Provisions for data security on the Zenodo platform apply. These concern the access to the CERN Data Centre, the servers, the network, the data storage, the application and the staff from Zenodo.
5.	Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
		Ethical standards and guidelines have been applied to the collection, processing and storage of data about persons, as described in Part B of the DOA. These include: - Identification and recruitment of participants through intermediaries - Informed consent procedures - Procedures for data collection, storage and protection conform national and EU legislation - No existing (secondary) data will be sought that is not already in the public domain as part of this project - Relevant authorisations will be sought for research permits where required to conduct workshops and interviews.

6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)
	All activities for data collection, storage and protection for this dataset comply with national and EU legislation, in particular with the GDPR. The SoilCare policy for privacy on personal data is applied to this dataset (https://www.soilcare-project.eu/glossary/all-terms/202:human-time-scale).

Dataset 2.3 Information on stakeholder perception of effects of SICS in terms of workload, risks and impact on farmer reputation (WP4)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The purpose of this dataset is to collect information in the SOILCARE study sites on how stakeholders perceive the tested soil-improving cropping systems (SICS) with regard to a) change in workload (compared to the "usual practice"), b) potential risks associated with the SICS, and c) their impact on farmer reputation. • Relation to the objectives of the project The dataset is used to identify the most important factors that determine the social / socio-cultural dimension of sustainability of the SICS. • Types and formats of data generated/collected All data is stored in a Excel spreadsheet. • Existing data that is being re-used (if any) None. • Origin of the data The data is derived from questionnaire based interviews conducted in the study sites. • Expected size of the data (if known) Small _1-2 MB). • Data utility: to whom will it be useful Other WPs and stakeholders in the study sites.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset will not be made findable for parties who are not part of the consortium because it underlies Deliverable 5.1, which has dissemination level *confidential* in the Grant Agreement of the SoilCare project. • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? No use is made of identifiers, because the dataset will not be published. • Naming conventions used No particular convention used. • Approach towards search keyword Excel search functionality can be applied • Approach for clear versioning Versioning is managed by partner 9 (UNIBE).

	• Standards for metadata creation (if any). No particular standards used.
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so Data in the dataset are part of D5.1, the database with monitoring results from the study sites generated in SoilCare. The status of this deliverable as laid down in the Grant Agreement is *confidential*, and therefore the data will not be made public. Data in the dataset has been made openly available only in aggregated form in Deliverable 5.3. After the term of confidentiality has expired, only data not referring to individual stakeholders can be published, based on the conditions of privacy granted to stakeholders collaborating in the research. • How the data will be made available The data have been made available to project partners only during the lifetime of the project. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? None. • Where the data and associated metadata, documentation and code are deposited In the archive of the SoilCare project in the institute of the project coordinator, Wageningen Environmental Research. • How access will be provided in case there are any restrictions Access to the data has only been provided to project partners due to the confidential dissemination level of the dataset. In case restrictions apply based on privacy guarantees to interviewed stakeholders, access has not be provided.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. No vocabularies, standards and methodologies from the social sciences domain were required to facilitate interoperability with other project results. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? As stated under the previous point.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), the data (or subsets of it) can

be made available under licences determined by the partner(s) who generated the data.

- When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners can make the data (or subsets of it) available for re-use and apply an embargo period.

- Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

The dataset will not be usable by third parties based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners who produced the data may decide to make the data available for re-use. In that case, reuse of data referring to individual stakeholders will not be allowed.

- Data quality assurance processes

Data quality is assured through the following approach: A semi-structured questionnaire as well as detailed guidelines on its use were developed. The questionnaire and guidelines were further refined and simplified on the basis of pilot interviews to test the clarity of the question wording, and feedback provided by study site teams. The study site teams were requested to translate the questionnaire into their local language to facilitate the interview process.

- Length of time for which the data will remain re-usable

To be determined after the term of status confidential of the dataset has expired, by the partners who generated the data, in case these decide to make the dataset re-usable.

3. Allocation of resources	- Costs for making your data FAIR. Describe how you intend to cover these costs No costs were foreseen. - Responsibilities for data management in your project This dataset was developed by WP4 of SoilCare. - Costs and potential value of long term preservation Long term preservation will contribute to the record of information on stakeholder's perception of SICS in terms of workload, associated risks, and impact on farmer reputation.
4. Data security	- Data recovery as well as secure storage and transfer of sensitive data Provisions for data security and storage in Wageningen Research apply.
5. Ethical aspects	- To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

	Ethical standards and guidelines were applied to the collection, processing and storage of data about persons, as described in Part B of the DOA. These include: - Identification and recruitment of participants through intermediaries - Informed consent procedures - Procedures for data collection, storage and protection conform national and EU legislation - No existing (secondary) data will be sought that is not already in the public domain as part of this project - Relevant authorisations will be sought for research permits were required to conduct workshops and interviews.
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any) All activities for data collection, storage and protection for this dataset comply with national and EU legislation, in particular with the GDPR.

Dataset 2.4 Information on **economic factors** relevant for adoption (WP4)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The purpose of this dataset is to collect information on cost and benefits of the SICS implementation that directly influences the adoption of SICS under different conditions of the study regions. • Relation to the objectives of the project The dataset contributes to the assessment of the overall sustainability of SICS in general, and more specifically will inform on profitability of SICS adoption. The information given in the dataset also contributes to the benefits and drawbacks of SICS assessment. • Types and formats of data generated/collected The data are in the form of a report summarizing the outcomes of the questionnaires received from the CSS teams, and survey results. Questionnaire on costs and benefits was sent to the 16 CSS with the aim to assess the profitability of 3 SICS implemented in each CSS. The questionnaire was completed by the CSS leaders together with the farmer. The information given in the questionnaire was checked and approved by the farmer using a participatory approach. The outcomes were provided in the periodic report and will be published in a scientific journal. All data are stored in Excel spreadsheets. • Existing data that is being re-used (if any) None. • Origin of the data The data are derived from interviews and survey results of multi-stakeholder panels in the study sites under supervision of the study site leaders. • Expected size of the data (if known) 100 kb • Data utility: to whom will it be useful Other WPs and stakeholders in the study sites.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset will not be made findable for parties who are not part of the consortium because it underlies Deliverable 5.1, which has dissemination level *confidential* in the Grant Agreement of the SoilCare project. After expiration of the term of confidentiality (4 years), the data (or subsets of it) can be made available under licences determined by the partner(s) who generated the data. • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? No use is made of identifiers, because the dataset will not be published. • Naming conventions used No particular convention used • Approach towards search keyword Excel search functionality can be applied

	• Approach for clear versioning Versioning is managed by partner 9 (UNIBE). • Standards for metadata creation (if any). No particular standards used.
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so Data in the dataset are part of D5.1, the database with monitoring results from the study sites generated in SoilCare. The status of this deliverable as laid down in the Grant Agreement is *confidential*, and therefore the data will not be made public. Data in the dataset has been made openly available only in aggregated form in [Deliverable 5.3](#). After the term of confidentiality has expired, only data not referring to individual stakeholders can be published, based on the conditions of privacy granted to stakeholders collaborating in the research. • How the data will be made available The data has been made available to project partners only during the lifetime of the project. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? None. • Where the data and associated metadata, documentation and code are deposited In the archive of the SoilCare project in the institute of the project coordinator, Wageningen Environmental Research. • How access will be provided in case there are any restrictions Access to the data will only be provided to project partners due to the confidential dissemination level of the dataset. In case restrictions apply based on privacy guarantees to interviewed stakeholders, access will not be provided.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. No vocabularies, standards and methodologies were required to facilitate interoperability with other project results because the data will be provided in a quantitative way. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? As stated under the previous point.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), the data (or subsets of it) can

be made available under licences determined by the partner(s) who generated the data.

- When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners can make the data (or subsets of it) available for re-use and apply an embargo period.

- Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

The dataset will not be usable by third parties based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners who produced the data may decide to make the data available for re-use. In that case, reuse of data referring to individual stakeholders will not be allowed.

- Data quality assurance processes

Data quality is assured through the following approach:
The questionnaire was completed by scientists (CSS leaders) together with a stakeholder (in most cases the farmer). Obtained data were quantified and presented using a scientific approach (statistical analysis). A comparison of all data has been done to check the coherence of the outcomes. In case of doubt, WP4 has contacted the study site leader who checked inconsistency with the stakeholder.

- Length of time for which the data will remain re-usable

To be determined after the term of status confidential of the dataset has expired, by the partners who generated the data, in case these decide to make the dataset re-usable.

3. Allocation of resources	- Costs for making your data FAIR. Describe how you intend to cover these costs
	No costs were foreseen.
	- Responsibilities for data management in your project
	This dataset was developed by WP4 of SoilCare. WP8 has disseminated part of the data but only in aggregated form.
	- Costs and potential value of long term preservation
	Long term preservation will contribute to the record of information on role and trust of stakeholders in the adoption of cropping systems in the study sites.
4. Data security	- Data recovery as well as secure storage and transfer of sensitive data
	Provisions for data security and storage in Wageningen Research apply.
5. Ethical aspects	- To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
	Ethical standards and guidelines have been applied to the collection, processing and storage of data about persons, as described in Part B of the DOA. These include:
	- Identification and recruitment of participants through intermediaries
	- Informed consent procedures
	- Procedures for data collection, storage and protection conform national and EU legislation

	- No existing (secondary) data will be sought that is not already in the public domain as part of this project - Relevant authorisations will be sought for research permits were required to conduct workshops and interviews.
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any) All activities for data collection, storage and protection for this dataset comply with national and EU legislation, in particular with the GDPR.

Database 3: Monitoring data of cropping systems (WP5)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The dataset stores monitoring data of selected cropping systems, and of already implemented cropping systems in the study sites. • Relation to the objectives of the project The dataset is used to identify the most important factors that determine profitability and sustainability, and the most important drawbacks of the selected cropping systems in the context of the study sites. • Types and formats of data generated/collected Data types collected include data about agronomic techniques, soil quality, environmental impacts and farm economy. Data formats follow the methodology for monitoring and assessing cropping systems as set out by WP4. Formats are chosen to facilitate comparison and consistency between Study Sites. • Existing data that is being re-used (if any) Existing records of agronomic techniques and farm economy of the experimental farms have been included in the database. • Origin of the data Data were derived from the monitoring activities by the study sites and compiled by WP5. The approach is described in the 'Database report' as part of deliverable D5.1, which will be shared on ZENODO. The database itself contains confidential information. • Expected size of the data (if known) For 16 SS weather-data, Study Site characteristic and experimental data have been stored in two monitoring forms. The total size of the monitoring forms for each study site is smaller than 10 MB. • Data utility: to whom will it be useful The database will be of use to scientists and research staff from authorities and companies interested in effects of agro-management on crop production.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset will not be made findable for parties who are not part of the consortium because it represents Deliverable 5.1, which has dissemination level *confidential* in the Grant Agreement of the SoilCare project. After expiration of the term of confidentiality (4 years), the data (or subsets of it) can be made available under licences determined by the partner(s) who generated the data. • Identificability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

	No use is made of identifiers, because the dataset will not be published. • Naming conventions used Naming conventions have been established by WP4 and 5, and are documented in Appendix 3 of the Database report D5.1. • Approach towards search keyword Search facilities are part of the database infrastructure. • Approach for clear versioning Versioning is established in the database infrastructure. • Standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how Metadata provision is defined in the common databases structure developed for the dataset by WP5, based on the methodology worked out by WP4.
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so Data in the dataset are part of D5.1, the database with monitoring results from the study sites generated in SoilCare. The status of this deliverable as laid down in the Grant Agreement is *confidential*, and therefore the data will not be made public. Data in the dataset has been made openly available only in aggregated form in Deliverable 5.3. • How the data will be made available For project partners through the 'cloud storage facility' operated by partner 3 (KUL). • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? The data are stored in standard SQL, which can be accessed by most software. Within WP5 opensource PostgreSQL was used for storing the database. The description of the components are available via the Database report D5.1[2]. • Where the data and associated metadata, documentation and code are deposited In the common database structure established by WP5. An empty version of the database structure has been deposited in the Zenodo Research Data Repository. • How access will be provided in case there are any restrictions Data with limited access rights due to privacy regulations have been accessible for project partners only. Public access outside the SoilCare project will only be possible after reports have been published and after the term of status 'confidential' of the deliverable has

---

[2] This report has been deposited in the Zenodo Research Data Repository.

	expired, under the condition that partners who generated the data give permission.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. Provisions for interoperability are included in the common database structure designed by WP5. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? The common database structure developed by WP5 uses standard vocabularies from agronomic, environmental and economic science. In addition to the explanation of variables that are in the dataset, there is also a [SoilCare glossary](#) that explains how terms are used within the SoilCare project.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), the data (or subsets of it) can be made available under licences determined by the partner(s) who generated the data. • When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed The dataset will be not be made available to third parties outside the consortium based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners can make the data (or subsets of it) available for re-use and apply an embargo period. • Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why The dataset will not be usable by third parties based on the confidential dissemination level of the associated deliverable. After expiration of the term of confidentiality (4 years), partners who produced the data may decide to make the data available for re-use. In that case, reuse of data referring to individual stakeholders will not be allowed. • Data quality assurance processes The data quality assurance processes were defined and controlled by WP5 and executed by the case study teams collecting the data. Data-validation and quality control took place shortly after the data-entry by study site teams. • Length of time for which the data will remain re-usable To be determined after the term of status confidential of the dataset has expired, by the partners who generated the data, in case these decide to make the dataset re-usable.

3. Allocation of resources	• Costs for making your data FAIR. Describe how you intend to cover these costs	
	Costs for setting up the common database structure and for the embedding in the project data repository were covered by the budget of WP5 in the SOILCARE project.	
	• Responsibilities for data management in your project	
	The responsibility for the data management in this dataset is with the WP5 leader, the project coordinators, and the case study leaders delivering the data to the dataset.	
	• Costs and potential value of long term preservation	
	The costs for long term preservation correspond to the costs for hosting the database in the infrastructure arranged by partner 3 (KUL) during the period in which the partner offers this service. Costs for hosting after the period of confidentiality has expired depends on the mode(s) of publication decided by the partners who generated the database and infrastructure.	
	The value for long term preservation is large, since the dataset forms the baseline for a set of monitoring results from new cropping systems in various contexts in Europe, in some case studies combined with extended records of already implemented cropping systems. Therefore, the dataset provides a basis for future long-term assessments of the performance and environmental impacts of the cropping systems.	
4. Data security	• Data recovery as well as secure storage and transfer of sensitive data	
	The dataset in the common database structure of WP5 are safeguarded in the 'cloud storage facility' operated by partner 3 (KUL) for a period determined by partner 3 (KUL).	
5. Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former	
	Ethical standards and guidelines have been applied to the collection, processing and storage of data about the farm enterprises where the new cropping systems have been implemented, and from where data were collected on already implemented cropping systems. These standards and guidelines are described in Part B of the DOA. They include a.o. procedures for data collection, storage and protection conform national and EU legislation.	
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)	
	Not applicable.	

Database 4: Dataset of information on cropping systems at European level (WP6)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation The purpose of this dataset is to support quantitative, spatially explicit impact assessment of soil-improving cropping systems in Europe and to provide information on requirements and profitability of the cropping systems. The dataset itself is not a deliverable of the project but was used to create integrated information on cropping systems and agronomic techniques at the European level under current and future conditions. • Relation to the objectives of the project The dataset supports the project objective to develop an interactive tool for selection of soil-improving cropping systems throughout Europe. • Types and formats of data generated/collected The data is in the form of maps, scenario-settings for socio-economic conditions and climate, and information on cropping systems in the form of indicators for productivity, profitability, adoption and sustainability. • Existing data that is being re-used (if any) Existing data used to create this dataset includes farm statistics, socio-economic statistical data, data on land use and land cover. • Origin of the data Existing data from outside the project were derived from a.o.Eurostat and FADN, SoilGrids, ESDAC and EEA. Newly created data is derived from the upscaling procedures applied in WP6. • Expected size of the data (if known) In the order of magnitude of GB for the applicability maps and scenario results. • Data utility: to whom will it be useful The datasets resulting from the upscaling procedures applied in WP6 are useful for policy makers and planners at the level of member states and the EU. In addition to their availability through the Zenodo data repository they were also included in the Interactive Mapping Tool.
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) The dataset will be made findable through the metadata standard applied by the Zenodo data repository (the Dublin Core Metadata Standard). • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? (A) Digital object Identifier(s) will be attributed to the dataset as the standard identification mechanism when uploaded to the Zenodo platform. • Naming conventions used The file names for the newly created data will have the following structure: Type of analysis_topic_simulated date. E.g. Scenario1_annual erosion_2030.

	• Approach towards search keyword Not applicable. • Approach for clear versioning Only final versions of the (created) data will be uploaded. • Standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how Standards of metadata provision are used as defined in the EU-wide Integrated Assessment Models developed in previous European research projects (MedAction, LUMOCAP, DeSurvey, RECARE).
2.2 Making data openly accessible	• Data that will be made openly available. If some data is kept closed provide rationale for doing so Information on cropping systems at the European level generated by the upscaling procedures will be made openly available. • How the data will be made available The data will not be made available directly, but in the form of processed results from the upscaling procedures. A conversion table in Excel will be provided as testimony of the intermediate steps in creating the SICS potential map layers for the SICS in Europe. The data will be made accessible in the Zenodo Data Repository when available. • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? The underlying applicability mapping tool and the underlying Integrated Assessment Model are required. Licensing is used. • Where the data and associated metadata, documentation and code are deposited In the Geonamica Software Environment and in the project's research data repository. • How access will be provided in case there are any restrictions Partner 6 (RIKS) has enabled access to the data, metadata, documentation and software code for all project partners for the purpose and duration of the project. Access to software might be restricted for others, as SoilCare work builds further on background IP owned by partner 6.
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. For newly created data content-related naming is used as described under 2.1. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? Standard vocabularies of the existing datasets have been used.
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible Data generated with these tools and uploaded to the Zenodo Research Data Repository will be licensed under the Creative Commons Attribution 4.0 International[3].

---

[3] https://creativecommons.org/licenses/by/4.0/legalcode

		• When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed
		The existing datasets used are already available at the original sources (FADN, Eurostat, CLC, LUCAS). Data generated through the upscaling procedures and the use of the Integrated Assessment will be made available within 6 months from the publication of the associated deliverables (by partner 6 with assistance from the project coordinator or the EC).
		• Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why
		Data generated through the upscaling procedures, including procedures applied in the Integrated Assessment Model, will be available for reuse by third parties after the expiration of embargo period of 6 months after publication of the corresponding deliverables.
		• Data quality assurance processes
		For existing data the quality assurance of the data providers is used (EC, EEA). Quality assurance of data generated by the Integrated Assessment Model and the Integrated Mapping Tool consists of the historic calibration and validation of the Integrated Assessment Model based on previous configurations for previous EU projects, complemented by expert judgement from experts within and outside the SoilCare project.
		• Length of time for which the data will remain re-usable
		Data uploaded to the Zenodo data repository will remain re-usable for the lifetime of the Zenodo repository (currently >20 years).
3. Allocation of resources		• Costs for making your data FAIR. Describe how you intend to cover these costs
		Costs are covered by the project budget from WP6.
		• Responsibilities for data management in your project
		Depositing the results of the upscaling procedures, including the IAM, is managed by WP6.
		• Costs and potential value of long term preservation
		There are no costs required for long term preservation. The value of long term preservation is the possibility to repeat the scenario analysis with new data in the future to assess the impact of actual climate and socio-economic change.
4. Data security		• Data recovery as well as secure storage and transfer of sensitive data
		Data are automatically backed-up and can thus be recovered if needed.
		For all data uploaded to the Zenodo research data repository, the provisions for data security in the repository apply. These concern the access to the CERN Data Centre, the servers, the network, the data storage, the application and the staff from Zenodo.
5. Ethical aspects		• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
		No ethical issues are associated with this dataset.
6. Other		• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)
		If parts of the dataset are published on platforms of member states of the EC, the procedures for data management of these entities apply.

Annex 1 Data items uploaded to the Zenodo Data Repository under Databases 1-3[4]

Name	Upload date	Database/dataset	Link address
Guidance and input for Study Site Teams: Participatory workshop on adoption	9-3-2020	Database 2: Information obtained from stakeholders - Dataset 2.2 Information on policy factors relevant for adoption (WP7)	https://zenodo.org/record/3702353#.YSNuDI4zaUk
Accompanying material to the Inventory of opportunities and bottlenecks in policy to facilitate the adoption of soil-improving techniques	28-3-2019	Database 2: Information obtained from stakeholders - Dataset 2.2 Information on policy factors relevant for adoption (WP7)	https://zenodo.org/record/2613625
Working with stakeholders - SOILCARE WP3 Workshop Guidelines	14-11-2016	Database 2: Information obtained from stakeholders - Dataset 2.1 Information on social factors relevant for adoption (WP3)	https://zenodo.org/record/166515
SoilCare database 3: schema (empty database) and Report 34 (D5.1): Database with monitoring data	01-10-2021	Database 3: Monitoring data of cropping systems (WP5)	https://zenodo.org/record/5541296#.YVaxtppByUk
SOILCARE_database1_WP2_SICS_aspects	20-8-2021	Database 1: Meta-information on cropping systems in Europe (WP2)	https://zenodo.org/record/5226666#.YS9HKY4zY70
SOILCARE_WP_7_D7.2_Adoption factors and policy actions	16-8-2021	Database 2: Information obtained from stakeholders - Dataset 2.2 Information on policy factors relevant for adoption (WP7)	https://zenodo.org/record/5205401#.YSNr344zaUk
Report on demonstration activities in the study sites (D5.2)	16-12-2020	Database 3: Monitoring data of cropping systems (WP5)	https://zenodo.org/record/4326767#.YO9dEugzY70
SOILCARE Study Sites	17-11-2016	Database 3: Monitoring data of cropping systems (WP5)	https://zenodo.org/record/167323#.YSNtlo4zaUk

---

[4] Data items underlying publications from SoilCare, that are published in the Zenodo RDR, are not listed in this table because they are not part of the datasets described in the DMP.

Annex 2 Data sharing, use and publication in SoilCare

Version 28 February 2019

Introduction

Data sharing, data ownership and publishing are to a large extent arranged within the SoilCare project documents, in particular in the Terms and Conditions of the Grant Agreement (GA) and in the Consortium Agreement (CA). Below these 3 subjects are discussed separately in 3 sections. Each section starts with the relevant text from the Terms and Conditions and from CA, and concludes with a summary of what this means for SoilCare, and a proposal for additional agreements where needed.

Data sharing

Terms and Conditions GA

7.1 General obligation to properly implement the action
The beneficiaries must implement the action as described in Annex 1 and in compliance with the provisions of the Agreement and all legal obligations under applicable EU, international and national law.

31.2 Access rights for other beneficiaries, for implementing their own tasks under the action
The beneficiaries must give each other access — on a royalty-free basis — to results needed for implementing their own tasks under the action.

31.3 Access rights for other beneficiaries, for exploiting their own results
The beneficiaries must give each other — under fair and reasonable conditions (see Article 25.3) — access to results needed for exploiting their own results.

CA

Section 4: Responsibilities of Parties

4.1 General principles

Each Party undertakes to take part in the efficient implementation of the Project, and to cooperate, perform and fulfil, promptly and on time, all of its obligations under the Grant

Agreement and this Consortium Agreement as may be reasonably required from it and in a manner of good faith as prescribed by Belgian law.

Each Party undertakes to notify promptly, in accordance with the governance structure of the Project, any significant information, fact, problem or delay likely to affect the Project.

Each Party shall promptly provide all information reasonably required by a Consortium Body or by the Coordinator to carry out its tasks.

Each Party shall take reasonable measures to ensure the accuracy of any information or materials it supplies to the other Parties

SoilCare

The texts above make clear that if the contract demands that data are shared, parties are obliged to deliver the data. This means, for example, that if the DOA (which is part of the contract!) states that WPX[5] will assemble data from all study sites, study sites are obliged to provide these data to WPX. This is done in a manner of good faith.

In our view, this is clear enough, so that strictly speaking no additional agreements are needed for SoilCare. However, we can add the following: WP leaders and others who receive data from study sites (or from other partners) promise to use the data they received only to fulfil their contractual obligation in SoilCare. This corresponds to clause 9.2.5. in the CA[6]. Any other use of these data (e.g. direct or indirect use of the data in a publication of any sort, or make data available to third parties outside of the consortium) needs to be discussed with the data owner(s) first, and will only be allowed after the data owner has given written consent. In case of a publication, the data owner has the possibility to become co-author of the publication, see the section on publishing.

Data ownership and use

Terms and Conditions of GA

26.1 Ownership by the beneficiary that generates the results
Results are owned by the beneficiary that generates them.

'**Results**' means any (tangible or intangible) output of the action such as data, knowledge or information — whatever its form or nature, whether it can be protected or not — that is generated in the action, as well as any rights attached to it, including intellectual property rights.

26.2 Joint ownership by several beneficiaries
Two or more beneficiaries own results jointly if:

---

[5] Work Package X, X indicating a number.
[6] 9.2.5 Results and Background shall be used only for the purposes for which Access Rights to it have been granted.

(a) they have jointly generated them and
(b) it is not possible to:
    (i) establish the respective contribution of each beneficiary, or
    (ii) separate them for the purpose of applying for, obtaining or maintaining their
       protection (see Article 27).

The joint owners must agree (in writing) on the allocation and terms of exercise of their joint ownership ('**joint ownership agreement**'), to ensure compliance with their obligations under this Agreement.

Unless otherwise agreed in the joint ownership agreement, each joint owner may grant non-exclusive licences to third parties to exploit jointly-owned results (without any right to sub-license), if the other joint owners are given:
(a) at least 45 days advance notice and
(b) fair and reasonable compensation.

Once the results have been generated, joint owners may agree (in writing) to apply another regime than joint ownership (such as, for instance, transfer to a single owner (see Article 30) with access rights for the others).

CA
8.0 Ownership of Results

Results are owned by the Party that generates them.

8.1 Joint ownership

Unless otherwise agreed:
    - each of the joint owners shall be entitled to use their jointly owned Results for non-commercial research activities on a royalty-free basis, and without requiring the prior consent of the other joint owner(s), and
    - each of the joint owners shall be entitled to otherwise Exploit the jointly owned Results and to grant non-exclusive licenses to third parties(without any right to sub-license), if the other joint owners are given:
        (a) at least 45 calendar days advance notice; and
        (b) Fair and Reasonable compensation.

The joint owners shall agree on all protection measures and the division of related cost in advance

SoilCare
Texts make clear who owns the data. It is clear that ownership rests with the partner or partners who did the work. In our view this remains the case also after data have been shared with other partners in the consortium. The focus of the texts is on exploitation of results, hence it is not about use of data within the project. Use of data within the project is taken for granted, as discussed in the section on data sharing. Of course data that are

collected in the project should be used for the purpose for which these data were intended, in accordance with the DOA. The main point here is that beneficiaries who carried out the work are the data owners. This ownership remains after the end of the project, so that in a new project it would become part of the background of the partner.

In case data are used outside the SoilCare project, both the Terms and Conditions of GA and CA mention 'fair and reasonable compensation'. The form of this compensation should be agreed upon by the involved parties, and can be material or non-material in nature.

Publishing

Terms and Conditions GA

29.1 Obligation to disseminate results
Unless it goes against their legitimate interests, each beneficiary must — as soon as possible — '**disseminate**' its results by disclosing them to the public by appropriate means (other than those resulting from protecting or exploiting the results), including in scientific publications (in any medium).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

A beneficiary that intends to disseminate its results must give advance notice to the other beneficiaries of — unless agreed otherwise — at least 45 days, together with sufficient information on the results it will disseminate.

Any other beneficiary may object within — unless agreed otherwise — 30 days of receiving notification, if it can show that its legitimate interests in relation to the results or background would be significantly harmed. In such cases, the dissemination may not take place unless appropriate steps are taken to safeguard these legitimate interests.

If a beneficiary intends not to protect its results, it may — under certain conditions (see Article 26.4.1) — need to formally notify the *Agency* before dissemination takes place.

29.2 Open access to scientific publications
Each beneficiary must ensure open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results. In particular, it must:

(a) as soon as possible and at the latest on publication, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications; Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications.

(b) ensure open access to the deposited publication — via the repository — at the latest:
    (i) on publication, if an electronic version is available for free via the publisher, or
    (ii) within six months of publication (twelve months for publications in the social sciences and humanities) in any other case.
(c) ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication.

The bibliographic metadata must be in a standard format and must include all of the following:
- the terms *"European Union (EU)" and "Horizon 2020"*;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier.

29.4 Information on EU funding — Obligation and right to use the EU emblem

Unless the *Agency* requests or agrees otherwise or unless it is impossible, any dissemination of results (in any form, including electronic) must:
(a) display the EU emblem and
(b) include the following text:

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 677407 (SoilCare project)".

When displayed together with another logo, the EU emblem must have appropriate prominence.
For the purposes of their obligations under this Article, the beneficiaries may use the EU emblem without first obtaining approval from the *Agency*.
This does not however give them the right to exclusive use.
Moreover, they may not appropriate the EU emblem or any similar trademark or logo, either by registration or by any other means.

29.5 Disclaimer excluding *Agency* responsibility

Any dissemination of results must indicate that it reflects only the author's view and that the *Agency* is not responsible for any use that may be made of the information it contains.

38.1 Communication activities by beneficiaries
38.1.1 Obligation to promote the action and its results

The beneficiaries must promote the action and its results, by providing targeted information to multiple audiences (including the media and the public) in a strategic and effective manner.
This does not change the dissemination obligations in Article 29, the confidentiality obligations in Article 36 or the security obligations in Article 37, all of which still apply.
Before engaging in a communication activity expected to have a major media impact, the beneficiaries must inform the *Agency* (see Article 52).

38.1.2 Information on EU funding — Obligation and right to use the EU emblem

Unless the *Agency* requests or agrees otherwise or unless it is impossible, any communication activity related to the action (including in electronic form, via social media, etc.) and any infrastructure, equipment and major results funded by the grant must:
(a) display the EU emblem and
(b) include the following text:

For communication activities: *"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 677407 (SoilCare project)".*

For infrastructure, equipment and major results: *"This [infrastructure][equipment][insert type of result] is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 677407 (SoilCare project)".*

When displayed together with another logo, the EU emblem must have appropriate prominence.
For the purposes of their obligations under this Article, the beneficiaries may use the EU emblem without first obtaining approval from the *Agency*.
This does not, however, give them the right to exclusive use.
Moreover, they may not appropriate the EU emblem or any similar trademark or logo, either by registration or by any other means.

38.1.3 Disclaimer excluding *Agency* responsibility
Any communication activity related to the action must indicate that it reflects only the author's view and that the *Agency* is not responsible for any use that may be made of the information it contains.

38.2 Communication activities by the *Agency*
38.2.1 Right to use beneficiaries' materials, documents or information
The *Agency* may use, for its communication and publicising activities, information relating to the action, documents notably summaries for publication and public deliverables as well as any other material, such as pictures or audio-visual material that it receives from any beneficiary (including in electronic form).
This does not change the confidentiality obligations in Article 36 and the security obligations in Article 37, all of which still apply.
However, if the *Agency's* use of these materials, documents or information would risk compromising legitimate interests, the beneficiary concerned may request the *Agency* not to use it (see Article 52).

The right to use a beneficiary's materials, documents and information includes:
(a) **use for its own purposes** (in particular, making them available to persons working for the *Agency* or any other EU institution, body, office or agency or body or institutions in EU Member States; and copying or reproducing them in whole or in part, in unlimited numbers);
(b) **distribution to the public** (in particular, publication as hard copies and in electronic or digital format, publication on the internet, as a downloadable or non-downloadable file,

broadcasting by any channel, public display or presentation, communicating through press information services, or inclusion in widely accessible databases or indexes);

(c) **editing or redrafting** for communication and publicising activities (including shortening, summarising, inserting other elements (such as meta-data, legends, other graphic, visual, audio or text elements), extracting parts (e.g. audio or video files), dividing into parts, use in a compilation);

(d) **translation**;

(e) giving **access in response to individual requests** under Regulation No 1049/2001, without the right to reproduce or exploit;

(f) **storage** in paper, electronic or other form;

(g) **archiving**, in line with applicable document-management rules, and

(h) the right to authorise **third parties** to act on its behalf or sub-license the modes of use set out in Points (b),(c),(d) and (f) to third parties if needed for the communication and publicising activities of the *Agency*.

If the right of use is subject to rights of a third party (including personnel of the beneficiary), the beneficiary must ensure that it complies with its obligations under this Agreement (in particular, by obtaining the necessary approval from the third parties concerned).

Where applicable (and if provided by the beneficiaries), the *Agency* will insert the following information:

"© – [year] – [name of the copyright owner]. All rights reserved. Licensed to the *Research Executive Agency (REA)* under conditions.

CA

8.3 Dissemination

8.3.1 Dissemination of own Results

8.3.1.1 During the Project and for a period of 1 year after the end of the Project, the dissemination of own Results by one or several Parties including but not restricted to publications and presentations, shall be governed by the procedure of Article 29.1 of the Grant Agreement subject to the following provisions.

Prior notice of any planned publication shall be given to the other Parties at least 21 calendar days before the publication. Any objection to the planned publication shall be made in accordance with the Grant Agreement in writing to the Coordinator and to the Party or Parties proposing the dissemination within 15 calendar days after receipt of the notice. If no objection is made within the time limit stated above, the publication is permitted.

8.3.1.2 An objection is justified if

(a) the protection of the objecting Party's Results or Background would be adversely affected

(b) the objecting Party's legitimate academic or commercial interests in relation to the Results or Background would be significantly harmed.

The objection has to include a precise request for necessary modifications.

8.3.1.3 If an objection has been raised the involved Parties shall discuss how to overcome the justified grounds for the objection on a timely basis (for example by amendment to the planned publication and/or by protecting information before publication) and the objecting Party shall not unreasonably continue the opposition if appropriate measures are taken following the discussion.

The objecting Party can request a publication delay of not more than 60 calendar days from the time it raises such an objection. After 60 calendar days the publication is permitted, provided that Confidential Information of the objecting Party has been removed from the Publication as indicated by the objecting Party.

8.3.2 Dissemination of another Party's unpublished Results or Background

A Party shall not include in any dissemination activity another Party's Results or Background without obtaining the owning Party's prior written approval, unless they are already published.

8.3.3 Cooperation obligations

The Parties undertake to cooperate to allow the timely submission, examination, publication and defence of any dissertation or thesis for a degree which includes their Results or Background subject to the confidentiality and publication provisions agreed in this Consortium Agreement.

8.3.4 Use of names, logos or trademarks

Nothing in this Consortium Agreement shall be construed as conferring rights to use in advertising, publicity or otherwise the name of the Parties or any of their logos or trademarks without their prior written approval.

SoilCare

The text from Terms and Conditions GA and CA above shows that intended publication should be shared within the consortium 21 days before planned publication (as this is what we stipulated in the CA)[7]. For this purpose, we will use a page on the SoilCare website with intended publications. The first author of an intended publication announces the intended publication to the project coordinator. The project coordinator places the details about the intended publication on the webpage (abstract, authors, and a link to the outline or manuscript), and informs all partners by e-mail, with reference to the webpage.

The text from the GA and CA shows that no partner can publish work from another partner unless there is written approval for that. Hence, the key is timely communication about any intended publication.
This applies to all use of Foreground, which includes:
- Someone else using data that you have collected

---

[7] Note that an exception was made for some kinds of publications, see minutes of kick-off meeting which state that: 'It is decided that the rules for sharing information about dissemination activities with SOILCARE partners will not apply to dissemination activities that do not allow 21 day notice, such as interviews, social media contributions etc. Prior notice is also not needed for dissemination that is really site specific.'

- You applying a method or model that was developed by someone else

Text from GA and CA also indicates that objections should not be continued unreasonably, which means that if an effort was made to overcome your objections, you have to accept. Hence, GA and CA in our view adequately cover how to deal with publishing results that is (partly) owned by other partners.

The GA and CA do not state rules about co-authorship. Based on the provisions on the use of Foreground above, we propose that the first author of an intended publication informs all the other partners (especially those that have shared results) as soon as an idea for a paper emerges. The first author makes a proposition for co-authors based on their role as either contributor or data provider and includes as much information about the intended paper as is available, such as key objectives, intended timeline for writing and proposed journal. We suggest that the webpage for intended publication on the SoilCare website, mentioned above, is used by first authors to propose co-authors.

Everyone who provides data, methods, models or any other form of contribution will be invited to become co-author, in accordance with general scientific practice. From co-authors it is expected that they actively contribute to the paper. If there are any objections against publication at this stage[8], these should be raised within 14 days. In that case, reasons should be provided, and if possible suggestions on how to overcome these objections should be included too. First author and partners shall then endeavour to reach a mutual conclusion about co-authorship in harmony before the paper is written. If agreement is not reached, the matter shall be discussed with the project coordinator, who shall then decide whether co-authorship is warranted or not.

Regarding Open Access to papers, it has been decided that this is the responsibility of the first author of each publication. The reasons for arranging it this way include:

- That some institutes have contracts with OA publishers, that allow them to publish in OA journals at reduced costs or even free of charge
- That several institutes have their own repositories for papers

Genarally, these facilities are only open to employees of these institutes, which precludes that OA can be arranged centrally in SoilCare. However, partners will inform coordinator and WP8 (dissemination), so thatan overview is maintained of all SoilCare papers.

---

[8] All partners will still be able to object when the paper is submitted, as described in section 8.3.1.1. of the CA. Hence, in the inception stage described here, objections should only be raised if there are objections against writing a paper on the intended subject at all.

Annex 2 Table format

Dataset x: Dataset title (WPx)

DMP component	Issues to be addressed
1. Data summary	• Purpose of the data collection/generation • Relation to the objectives of the project • Types and formats of data generated/collected • Existing data that is being re-used (if any) • Origin of the data • Expected size of the data (if known) • Data utility: to whom will it be useful
2. FAIR data	
2.1 Make data findable, including provisions for metadata	• Discoverability of data (metadata provision) • Identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? • Naming conventions used • Approach towards search keyword • Approach for clear versioning • Standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how standards for metadata from existing datasets are adopted (FADN, Eurostat, CLC, LUCAS).
2.2 Making data openly accessible	• Data that will be made openly available If some data is kept closed provide rationale for doing so • How the data will be made available • Methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

	• Where the data and associated metadata, documentation and code are deposited • How access will be provided in case there are any restrictions
2.3 Making data interoperable	• Interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. • Standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?
2.4 Increase data re-use (through clarifying licences)	• How the data will be licenced to permit the widest reuse possible • When the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed • Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why • Data quality assurance processes • Length of time for which the data will remain re-usable
3. Allocation of resources	• Costs for making your data FAIR. Describe how you intend to cover these costs • Responsibilities for data management in your project • Costs and potential value of long term preservation
4. Data security	• Data recovery as well as secure storage and transfer of sensitive data
5. Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)