

Project¹ Number: 774124

Project Acronym: SUPER-G

Project title: Developing SUsustainable PERmanent Grassland systems and policies

DATA MANAGEMENT PLAN

¹ The term ‘project’ used in this template equates to an ‘action’ in certain other Horizon 2020 documentation

Contents

1. Data Summary.....	3
2. FAIR data	5
2.1. Making data findable, including provisions for metadata	5
2.2. Making data openly accessible	6
2.3. Making data interoperable	7
2.4. Increase data re-use.....	7
3. Allocation of resources	8
4. Data security	8
5. Ethical aspects.....	8
6. Other sources of information	8

1. Data Summary

The purpose of the data collection/generation and its relation to the objectives of the project

Data will be gathered to meet the following SUPER-G objectives:

- i. To benchmark and model the ability of PG to deliver farm productivity, profitability and key ES within a selection of biogeographic regions.
- ii. Improve understanding of trade-offs and synergies associated with the management of permanent grassland (PG) for ecosystem service (ES) delivery in a range of contexts and scenarios.
- iii. Clarify the benefits PG provides for society and generate policy options that are either tailored to or robust across multiple PG contexts to deliver the range of ES.

Data will be made findable, accessible, interoperable and reusable (FAIR).

Types and formats of data generated and collected

SUPER-G will generate and collect:

- Social, economic and environmental data and metadata from PG farm networks, field experiments and demonstration platforms.
- Knowledge, experiences and insights based on stakeholder interactions.
- Personal data, records and images.

All personal data will be collected and processed in accordance with General Data Protection Regulations (GDPR - Regulation (EU) 2016/679). Data will also be used in physical models and relational databases.

The following types of data (and associated metadata) will be generated and collected:

Task no.	Task name	Data type
T2.1	Grassland typology	• New geographical information data layers
T2.2	Farming systems on PG	• New geographical information data layers
T2.3	ES provided by PG	i. Outputs from systematic review and meta-analysis ii. Modelling outputs iii. Quantitative expert elicitation data
T2.4	Transdisciplinary workshops	• Quantitative/qualitative coded expert elicitation data
T3.1	Farm networks	• Quantitative/qualitative coded survey data
T3.3	Synergies and trade-offs	• Quantitative data from field trials, demonstrations and experiments
T3.4	Field experimentation and demonstration	• Quantitative data from field demonstrations and experiments
T3.5	Modelling and simulation	• Modelling and simulation outputs
T3.6	Road testing	• Quantitative and qualitative data from road testing and demonstration
T4.1a	Review of facilitators of and barriers to adoption and choice of PG management options to deliver ES	• Outputs from systematic review
T4.1b	Economic drivers of farmer adoption	• Outputs from systematic review
T4.1c	Review of existing policies and impacts	• Outputs from systematic review
T4.2	Farmer priorities and preferences for ES in relation to PG	• Quantitative/qualitative coded survey data
T4.3	Citizen priorities and preferences for ES in relation to PG	i. Coded focus group transcripts ii. Quantitative/qualitative coded survey data

Task no.	Task name	Data type
T4.4	Developing policy options for ES in relation to PG	<ul style="list-style-type: none"> • Metadata focused on impact pathways

In WP3, scientific data will be generated through demonstration and field experimentation on commercial farms and experimental platforms to improve understanding of trade-offs and synergies associated with the management of PG. Data will be gathered to investigate the relationship between PG management and ES delivery. The aim is that from this data and the systematic reviews carried out in WP2 (Delivering sustainable systems), a relational database could be developed with estimates of pros and cons of PG management options, that could enable extrapolation of experimentally determined option impacts to new situations, based on similarity of characteristics.

Two questionnaire surveys will be carried out on farm management practices and business profitability (at the beginning and at the end of the project).

Contact details of organisations representing farmers, advisers, researchers, retailers, policy makers and other stakeholders will also be collected and processed in accordance with General Data Protection Regulations (GDPR - Regulation (EU) 2016/679)..

Surveys will also be used to collect information, attitudes and opinions from farmers, policy makers and citizens.

Re-use of existing data

We will gather and analyse technical information and data from farm networks and existing European (e.g. Eurostat) and national databases, although it will not be possible to share all of this existing data in a project database.

Data from experimental platforms will also be collated and reviewed.

Existing data from Eurostat and national statistics databases (e.g. June census) will also be used in models to investigate the relationship between PG management and ES delivery.

Data utility

The generated and collated data will be useful for natural scientists, social scientists, economists, agricultural advisers, farmers, policy makers and other stakeholders.

2. FAIR data

2.1. Making data findable, including provisions for metadata

Metadata will be gathered and integrated into newly created or existing databases to improve accessibility and interrogation. The data will also be identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers).

No SENSITIVE personal data will be collected, but certain personal data in the form of contact data (name, address, email address, telephone number for example) will be collected. This will be stored separately from the project data with the only link being through a code which will be accessible to only the lead researcher in each country where that data has been collected. Therefore, in terms of the GDPR the project data will not be fully anonymised until the link is destroyed. All participants will be made aware of this as part of the informed consent process. As it would be misleading if the informed consent documentation said that the data collected would be anonymised, participants will be informed that the research data collected will only be linked to their personal data by a code, which will be kept securely in a separate location from the research data and only accessible to the lead researcher. They will also be told for how long their personal details will be retained, and when and if the code/link will be destroyed. All the metadata associated with new data generated by the project will be collected on laptops or hand held devices and then uploaded to a central storage area.

For activities where audio or video recordings are taken, the data will be transcribed and then the recording medium erased or destroyed. No video recordings will be used as part of formal data collection. However, separate video recordings may be used as part of dissemination activities (for example, imagery of farms and farmers). Under these circumstances, then individual consent to this will be obtained.

Metadata requirements

SUPER-G will collect the following standard information in addition to specific metadata relevant to each individual discipline and database:

- Title – how the data are known
- Description – a brief methodological overview of the data
- Keywords – related to the content of the data
- Creators – the main researchers involved in creating the data
- Access conditions – how the data can be accessed and whether there are any restrictions in place

Supporting documentation will describe the data and include:

- Code, field and label descriptions
- Software used
- Methodology used to gather the data and links to experimental protocols
- For instrumental data - precision/accuracy of the data and approximate typical random error
- Dates of collection
- Geographic location

Work package leaders will be responsible for selecting data and metadata vocabularies, standards and methodologies to make SUPER-G data interoperable within each discipline.

Organising data

The project will use file naming conventions that are consistent, concise, meaningful and findable. The most recent version of each file will be retained on the project internal communication platform (SharePoint) along with relevant older versions.

Within document and spreadsheet directories, file and folder naming will be consistent across work packages (WPs), with WP, task and deliverable folders including a SUPER-G prefix and relevant suffix, e.g.:

- SUPER-G_WP1
- SUPER-G_T2.1
- SUPER-G_D3.1

File names will not include spaces as some software can't read them. Alternatives include underscores, hyphen and camel case, e.g.:

- SUPER-G-data-mgt-plan_en(1)
- SUPER-G_dissemination_plan_20181123_clean
- SUPER-G_T3.4_UNEW_Interview-audio_20181014(1)

File names should not exceed 50 characters.

Files and samples will be numbered using an appropriate scale, e.g. if 15 samples are to be taken, the first one will be numbered 01 rather than 1.

Version control tables will be used for important documents, including version number, date, author, date and whether the document is 'clean' or includes 'tracked changes' changes.

If required, documents will be dated at the end of the file/ folder name using a YYYYMMDD suffix.

File structure

SUPER-G will use a hierarchical file structure organised by WPs, tasks and deliverables. The aim is to achieve a balance between breadth and depth, with a limited number of broad topics (WPs) at the highest level.

Folders will be backed up on SharePoint. Relevant files should be transferred from organisation shared drives (for which back up is automatic) and individual hard drives on a regular basis.

Folders and files will be reviewed on a bi-monthly basis to ensure that they are not retained needlessly and to separate current and completed activity. An archive folder within the hierarchical structure will be used to avoid a cluttered workspace.

File formats

File formats will be determined by how data is to be collected and analysed. For long term storage, we will use formats that do not have restrictions on their use, increasing the likelihood that they will be accessible in the future. The following table presents some popular data formats and file format options.

Format	Ideal for preservation	Acceptable for preservation
Textual data	.rtf; .txt; .xml	doc; .docx; .html
Tabular data	.por; .csv; .tab	xls; .xlsx; .sav; .mdb; .txt; .dbf; .dta; .ods
Image	.tif (version 6)	.jpeg; .jpg; .pdf; .raw; .psd
Video	.mj2; .mp4	
Audio	.flac; .wav	.mp3; aif
Geospatial data	.shp; .shx; .dbf; .tiff; .tfw;	.mdb; .mif; .kml; .dxf; .svg

2.2. Making data openly accessible

All data produced in the project, apart from personal data, will be made openly available, where possible, by deposition in a certified repository supporting open access. Data will be made available in formats that are acceptable for preservation (see Table above). Documentation describing the metadata will also be provided. Access to the data will be facilitated by registration. We will provide research data repositories, designed with FAIR as their guiding principle, to provide open platforms for the preservation and sharing of research datasets.

The ethics advisory committee will ensure rigorous application of ethical standards, and data privacy standards, and compliance with the General Data Protection Regulation (Regulation (EU) 2016/679), especially in research where human participants are involved. Membership will comprise representatives of Wageningen Research, University Of Newcastle upon Tyne, ADAS and Sveriges Lantbruksuniversitet.

The committee will be tasked with ensuring ethical compliance with the GDPR, through establishing best practice, monitoring ethical procedures, including that associated with human participants, overseeing the collation and archiving of ethical documentation, and delivering annual reports regarding ethical compliance within activities within the project. The committee will meet annually in person at plenary meetings, and remotely in the intermediary period between meetings. If needed, ad hoc virtual meetings will be called by the chair to deal with specific ethical issues arising.

Publications will be made open access via the 'green' (no-cost) open access route; copies will be deposited within appropriate subject repositories and local institutional repositories. Wherever possible research publications/articles related to the project will also be linked to openAIRE: a European Open Science network.

In some instances 'green' open access may not be the most appropriate route, either due to publisher restrictions or because 'gold' (paid) open access will improve the citation / discoverability. In these cases funds to pay the Article Processing Charge (APC) will be included in the grant as a project-specific cost.

2.3. Making data interoperable

The data produced in the project will be interoperable, allowing data exchange and re-use between researchers, institutions, organisations and countries. SUPER-G will adhere to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins.

Some data provided by human participants will be transferred between EU and non EU countries to enable analysis. This specifically involves activities in:

- Task 2.4. Transdisciplinary workshops to discuss technical feasibility of ES delivery with farmers and experts (qualitative/quantitative coded data about technical aspects of PG management). This will involve transfer between EU member states with the exception of Switzerland, where expert elicitation data will be transferred to the UK.
- Task 4.2 Farmer priorities and preferences for ES in relation to PG: qualitative coded data about farmer priorities for policy. This will involve transfer of data between EU member states with the exception of Switzerland, where data will be transferred to the UK.
- Task 4.3. Citizen priorities and preferences for ES in relation to PG: qualitative (coded focus group transcripts) and quantitative (coded survey data) survey. This will involve transfer of data between EU member states with the exception of Switzerland, where data will be transferred to the UK.
- Task 4.4. Developing policy options for ES in relation to PG. Stakeholder workshops. Meta data focused on impact pathways will be transferred between EU member states, with the exception of Switzerland, and potentially Montenegro, where data will be transferred to Sweden.
- If authorisations for any transfers are required these will be obtained and kept on file to be provided to the Agency upon request. Details about the data to be imported /exported to/from the EU will be submitted as a deliverable.

Other information or documentation that must be either obtained and kept on file or provided to the Agency before the relevant research commences within each research task will be provided at the appropriate time, e.g.:

1. Detailed information on data collection, storage, protection etc. (To be sent to the Agency)
2. Details of the procedures for recruitment of participants (To be sent to Agency)
3. Details of informed consent procedures (To be sent to Agency)
4. Templates of informed consent documentation (To be sent to Agency)
5. Details of Incidental findings policy (Retained on file)
6. Details about the data to be imported /exported to/from the EU will be submitted as a deliverable
7. Any national declarations on compliance (Retained on file)
8. Copies of Opinions of DPOs (Retained on file)
9. Copies of any necessary import/export authorisations (Retained on file)

We will investigate the use of standard vocabularies for all data types to allow inter-disciplinary interoperability.

Where uncommon or project specific ontologies or vocabularies are used or generated, SUPER-G will provide mappings to more commonly used ontologies.

2.4. Increase data re-use

Data produced and/or used in the project will be useable by third parties, and will be made available for re-use after publication in a peer-reviewed journal by SUPER-G project partners. Publications will include a data access statement, describing where supporting data can be found and under what conditions they can be accessed.

Filr could be used for collaborators outside the SUPER-G network who require regular access to data. Filr is a Newcastle University service that allows user access to data from anywhere in the world. It can be used to synchronise data on the University's servers with data on a remote computer or mobile device, as well as to access files through a web interface from a device that is not already synchronised. The data will not leave the repository and will be automatically backed up. Access control will be delegated and permissions allocated on a folder basis.

3. Allocation of resources

Costs for making data FAIR will be covered by the organisations involved in the gathering and analysis of data, and therefore ultimately in the publication of related research papers. Costs related to open access to research data are eligible as part of the Horizon 2020 grant.

Data management will be the responsibility of all SUPER-G participants involved in data collation, collection, transfer and storage. Task leaders, WP leaders and the scientific coordinator will be responsible for the development of data management protocols and it is the responsibility of other participants to implement these protocols.

Long term preservation of data, in terms of how long the data needs to be retained, will be discussed with the Research Executive Agency (REA or the Agency).

4. Data security

All non-personal data produced in the project will be made available by deposition in a certified repositories supporting open access.

5. Ethical aspects

The SUPER-G Ethics Summary Report raised issues regarding the protection of personal data, humans and third countries. These issues are addressed in WP7 on Ethics requirements. SUPER-G will ensure rigorous application of ethical standards, and data privacy standards, and compliance with the General Data Protection Regulation (GDPR - Regulation (EU) 2016/679), especially in research where human participants are involved. All requirements detailed in the Ethics Summary Report provided by the Agency will be dealt with within this WP.

See section 5.1 of the SUPER-D Description of Action (DoA) for more information.

6. Other sources of information

The following sources will be used to finalise the Data Management Plan through the course of the project:

- The Research Data Alliance provides a [Metadata Standards Directory](#) that can be searched for discipline-specific standards and associated tools.
- The [EUDAT B2SHARE](#) tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.
- [Registry of Research Data Repositories](#)
- [DMP online](#)
- Platforms for making individual scientific observations available such as [ScienceMatters](#).

SUMMARY TABLE 1**FAIR Data Management at a glance: issues covered in the SUPER-G DMP**

DMP component	Issues addressed
1. Data summary	<ul style="list-style-type: none"> • Purpose of the data collection/generation • Relation to the objectives of the project • The types and formats of data generated/collected • How existing data will be used • Origin of the data • Data utility: to whom will it be useful
2. FAIR Data 2.1. Making data findable, including provisions for metadata	<ul style="list-style-type: none"> • 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 for clear versioning • Standards for metadata creation.
2.2 Making data openly accessible	<ul style="list-style-type: none"> • Data to be made openly available • How the data will be made available • Methods or software tools needed to access data • 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	<ul style="list-style-type: none"> • Interoperability of the data. • Data and metadata vocabularies, standards or methodologies to be used. • Use of standard vocabulary for all data types to allow inter-disciplinary interoperability
2.4. Increase data re-use (through clarifying licences)	<ul style="list-style-type: none"> • How the data will be licenced. • When the data will be made available for re-use. • Whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project.

DMP component	Issues addressed
	<ul style="list-style-type: none"> • Data quality assurance processes • Length of time for which the data will remain re-usable
3. Allocation of resources	<ul style="list-style-type: none"> • Estimated costs for making data FAIR • Responsibilities for data management • Costs and potential value of long term preservation
4. Data security	<ul style="list-style-type: none"> • Data recovery, secure storage and transfer of sensitive data
5. Ethical aspects	<ul style="list-style-type: none"> • See ethics section of DoA and ethics deliverables.
6. Other	<ul style="list-style-type: none"> • Other national/funder/sectorial/departmental procedures for data management.

HISTORY OF CHANGES		
Version	Publication date	Change
1.0	26.11.18	▪ Initial version
2.0	30.11.20	▪ Version 2