SMALLDERS PROJECT

SMART **M**ODELS FOR **A**GRIFOOD **L**OCAL VA**L**UE CHAIN BASED ON **D**IGITAL TECHNOLOGIES FOR **E**NABLING COVID-19 **R**ESILIENCE AND **S**USTAINABILITY

D1.2

Data Management Plan

Project Information

Acronym: SMALLDERS

Title: Smart Models for Agrifood Local Value Chain Based on Digital

Technologies for Enabling Covid-19 Resilience and Sustainability

Coordinator: University of Calabria (Italy)

Programme: PRIMA. Thematic area Agro-food value chain. Topic 2.3.1

Start: 2022-05-01

Duration: 36 months

MEL Web: https://mel.cgiar.org/projects/smallders

Consortium:

University of Calabria, Italy (UNICAL)

University of Parma, Italy (UNIPR)

University of Extremadura, Spain (UEX)

IMT Mines Ales, France (LSR)

Faculty of Science of Tunis, Tunisia (LAPER)

DELIVERABLE

Number: D1.2

Title: Data Management Plan

Lead participant: University of Calabria, Italy (UNICAL)

Other participants: University of Parma, Italy (UNIPR)

                   University of Extremadura, Spain (UEX)

                   IMT Mines Ales, France (LSR)

                   Faculty of Science of Tunis, Tunisia (LAPER)

Work package: WP1

Dissemination level: Public (PU)

Type: Report (R)

Due date: 2022-10-31

Submission date: 2022-11-01

Main Contributors: Francesco Longo (UNICAL), Antonio Padovano (UNICAL), Vittorio Solina (UNICAL)

Reviewer(s): Giuseppe Vignali, University of Parma, Italy (UNIPR)

Document Revision

Version	Date	Editor	Overview
First Release	2022-10-10	Vittorio Solina	First version
Second Release	2022-10-28	Francesco Longo, Vittorio Solina	Second version and discussion with partners
Revision	2022-10-31	Giuseppe Vignali	Version after first review
Final Release	2022-10-31	Francesco Longo, Vittorio Solina	Final version

Table of Contents

List of Figures

List of Tables

- **Executive Summary**

This deliverable is aimed at providing the Data Management Plan (DMP) of the SMALLDERS project. The DMP is extremely useful for documenting and communicating the management aspects related to project activities. More specifically, the DMP will analyze the overall data management strategy, the various data management procedures and the approaches that the consortium will follow to be compliant with the FAIR (Findability, Accessibility, Interoperability, Reusability) Data Principles.

1   Introduction

This deliverable represents a formal document, which describes how the data will be handled during the SMALLDERS project and after its completion. It represents a significant point of reference for all partners as it aims to detail and define:

- the expected types of datasets;
- the nomenclature to be used for the project files;
- the expected data formats;
- the repositories.

Furthermore, this document briefly describes which strategies are adopted for compliance with the FAIR Data Principles and for the security and respect of ethical issues.

This document is fundamental for the entire project as in each phase well-defined rules will be needed to manage the data, in order to guarantee: effective communications between partners, avoid wasting time in managing project activities, ensure consistency with the objectives declared at the beginning of the project, to reach the correct release of the SMALLDERS platform.

2 Data summary

This section provides important details on how data will be handled throughout the project lifecycle. Specifically, the main types of data, the data formats, the nomenclature to uniquely identify the data, the way to describe the various datasets, the repositories used within the project are detailed. It is important to underline that any data present in the SMALLDERS project will be treated in accordance with current regulations on data protection, as described in more detail in the deliverable D1.3 Ethics and Privacy Plan. Table 1 provides details about the types of datasets present in the SMALLDERS project.

Table 1. SMALLDERS dataset types.

Dataset types	Description
Software (i.e., SMALLDERS Platform)	The SMALLDERS ICT platform is certainly one of the most important datasets of the project. At the end of the project. the SMALLDERS platform will be released as a prototype to demonstrate, on specific testbeds and use cases, its potentials and functionalities. The partners expect to keep working on the SMALLDERS platform even after the completion of the project to raise the Technology Readiness Level (TRL) to the highest value therefore releasing the SMALLDERS platform as open-source for public administrations. To this end, the business model will be based on the so-called "Software Re-Use Catalogs for Public Administrations". More details on this business model are provided in the section of this document concerning compliance with the FAIR Data Principles.
Scientific publications	Another important project dataset concerns scientific publications. In line with the planned communications activities, SMALLDERS will produce a number of Conference Papers, all in the form of Gold Open Access. In addition, articles on prestigious

	scientific journals will also be published, both in Gold Open Access and Green Open Access formats.
Dissemination Materials	To contribute to the dissemination of the project results, the SMALLDERS consortium will generate updates on the website and on the different social media (i.e., Facebook, LinkedIn, Twitter). Furthermore, promotional materials such as newsletters, photos, videos, flyers, etc. will be generated and distributed.
Internal Materials	A dataset will derive from the internal communications and materials, necessary to carry out all the activities of the project
Other Datasets	Considering that the SMALLDERS project is still in the early months, it is necessary to leave room for the possibility that other datasets may be needed.

2.1 Data formats

The SMALLDERS project, wherever possible, will try to use formats that are open-access and in any case well recognized. In this context, the aim is to facilitate compliance with the so-called FAIR Data Principles. Table 2, for each type of file, shows the recommended formats, based on current knowledge. It is important to underline that, during the course of the project, new formats could be added, based on the needs of the consortium.

Table 2. SMALLDERS data formats.

File type	Recommended formats
Textual document	- Plain text, ASCII (.txt) - Software-specific formats - XML (.xml) - JSON (.json)

	• MS Excel (.xls, .xlsx)
	• Adobe PDF (.pdf)
	• Hypertext Mark-up Language (.html)
	• MS Word (.doc/.docx)
	• Comma-separated values (.csv)
Source code	• Python (.py, .npy, .pyc, .pyd, .pyo, .rpy, .whl)
	• C++ (.h, .hh, .hpp, .hxx, .cpp, .c, .cc, .cxx)
	• JavaScript / Node.js (js)
	• C# (.cs, .csx, .rsp)
	• Java (.class)
	• Hypertext Preprocessor (PHP)
	• HyperText Markup Language (HTML)
Audio	• FLAC (.flac)
	• WAV (.wav)
	• MPEG-1 Audio Layer 3 (.mp3)
	• Audio Interchange File (.aif)
Video	• MPEG-4 (.mp4)
	• WAV (.wav)
	• SVO video file (StereoLabsTM) (.svo)
	• Audio Interchange File (.aif)
	• AVI video file (.avi)
	• Motion JPEG 2000 (.mj2)
	• BVH human motion file format (.bvh)
	• OGG video (.ogv, .ogg)
Image	• JPEG (.jpeg, .jpg)
	• BMP (.bmp)
	• PNG (.png)
	• TIFF (.tif, .tiff)

2.2 Files identification

All files generated and collected within the SMALLDERS project will be uniquely identifiable through the following nomenclature:

SMALLDERS_[type]_[name]_[date]_[policy]_[version]

where:

- "**type**" - describes the type of file (i.e., based on what was previously described: software (S), scientific publication (SP), dissemination material (DM), internal material (IM), other dataset (OD)).
- "**name**" - is the name, which describes the content of the file.
- "**date**" - is the date in the format "YYYYMMDD".
- "**policy**" - is the policy concerning the file (e.g., restricted, open, embargoed).
- "**version**" – is the version of the file (i.e., v01, v02, v03, etc.).
- "_" (underscore) – is used to separate the different fields.

2.3 Data description

The data generated and collected within the SMALLDERS project will be adequately described, in order to ensure order and compliance with the FAIR Data Principles. Table 3 details how the various data will be described. The partners will refer to different aspects: name of the dataset, responsibilities, data type, data format, connection with the goals of the project, data origin, data audience (i.e., expected data destination), size of the data, accessibility policy, storage.

Table 3. SMALLDERS data description template.

Field	Description
Name of the dataset	Name explaining the content of the considered dataset.
Responsibilities	It is necessary to specify which SMALLDERS WP and task the data refers to. Furthermore, it is recommended to include the responsible partner as well.
Data type	It specifies the type of data (e.g., image, textual document, etc.).
Data format	It specifies in which formats (e.g., .pdf, .docx, etc.) the dataset is available.
Connection with the goals of the project	It specifies the relationship between the data and the project objectives, hence why the data is needed.
Data origin	It specifies the origin of the data.
Data audience	It specifies the expected audience of the data.
Size of the data	It specifies the data size.
Accessibility policy	It specifies one of the following accessibility policies: • **Open access:** the data is public and available to any entity inside and outside the project. • **Restricted access**: The data is only available within the consortium. • **Embargoed access**: after a certain period of restricted access, the data will become open.
Storage	It specifies where the data is stored.

2.4 Repositories

Two different repositories will be used in the SMALLDERS project: the project website and a shared folder on Google Drive. These repositories will be extremely useful for carrying out all the project activities, therefore they will not only serve to store data, but also to encourage collaboration between partners, to communicate with the "outside world", to manage and organize data effectively and efficiently. The two repositories are detailed below.

2.4.1 SMALLDERS Website

The SMALLDERS website (https://smallders.com/) will be the official channel to keep all the stakeholders updated on the progress of the project. It will be scalable and user-friendly (see Figure 1 for the homepage) and will have different purposes, which are explained below, through the description of the main sections.
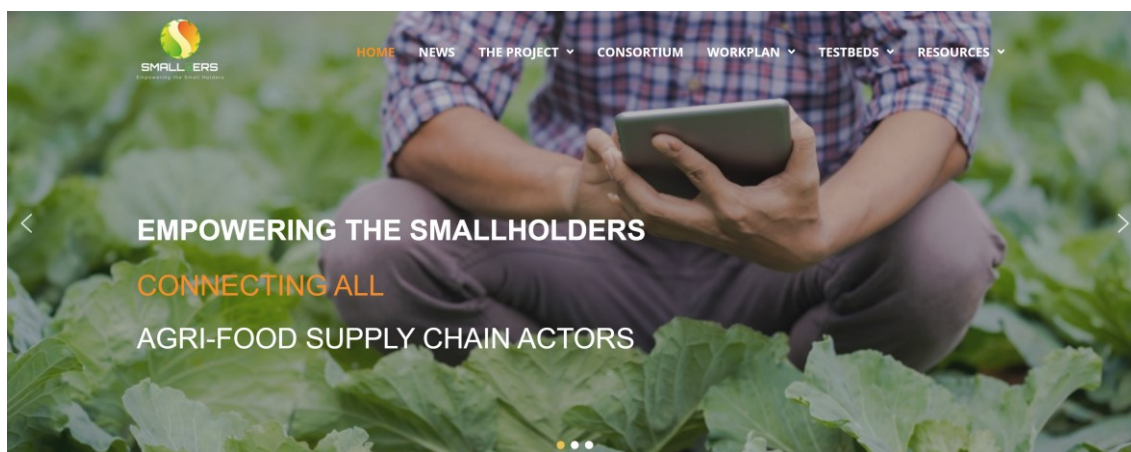
Figure 1. SMALLDERS website homepage

As shown in Figure 2, a section of the website is dedicated to reporting all the news, in order to communicate and disseminate the results gradually obtained in the project. Specifically, here you will find links to scientific publications, photos and videos about the consortium's participation in scientific conferences,

information on technical workshops, news on testbeds, photos on face-to-face meetings, and much more.

Figure 2. SMALLDERS website: news section

Another section of the website (see Figure 3) is dedicated to a general description of the project and includes: an abstract, the objectives, the expected impact and the human resources involved in the project in terms of consortium, advisory board and external partners. This section is extremely important because it provides information, which allows anyone to measure the degree of achievement of project goals over the months.

Figure 3. SMALLDERS website: section about the project

The fourth section of the website (see Figure 4) shows the main information on the partners that make up the consortium, therefore their geographical location and main competences.


Figure 4. SMALLDERS website: consortium section

Another section of the website concerns the project workplan, as shown in Figure 5. Visitors can view the Project Gantt Diagram and the content of the different work packages in terms of expected objectives and deliverables, with quite well detailed information on leaders and deadlines.
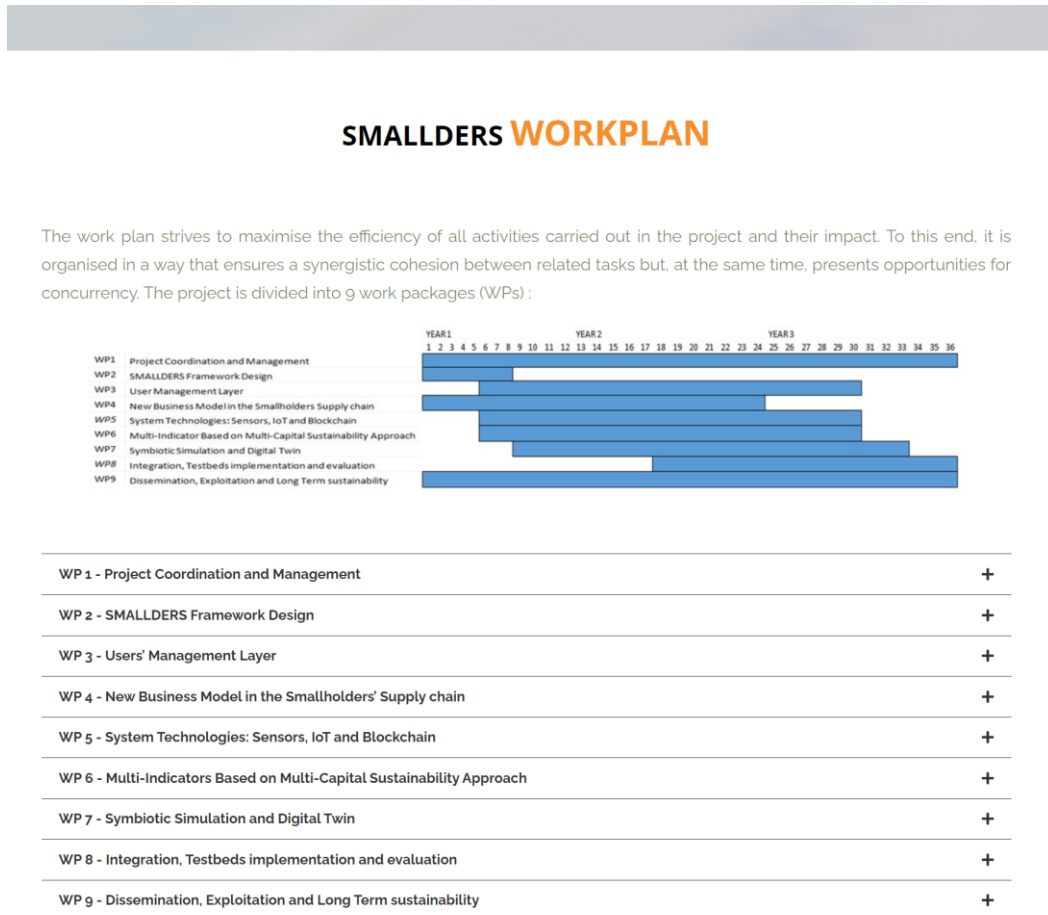
Figure 5. SMALLDERS website: workplan

Another fundamental part of the website concerns the testbeds, therefore a preliminary description of the case studies of the project, divided by geographical location (see Figure 6).
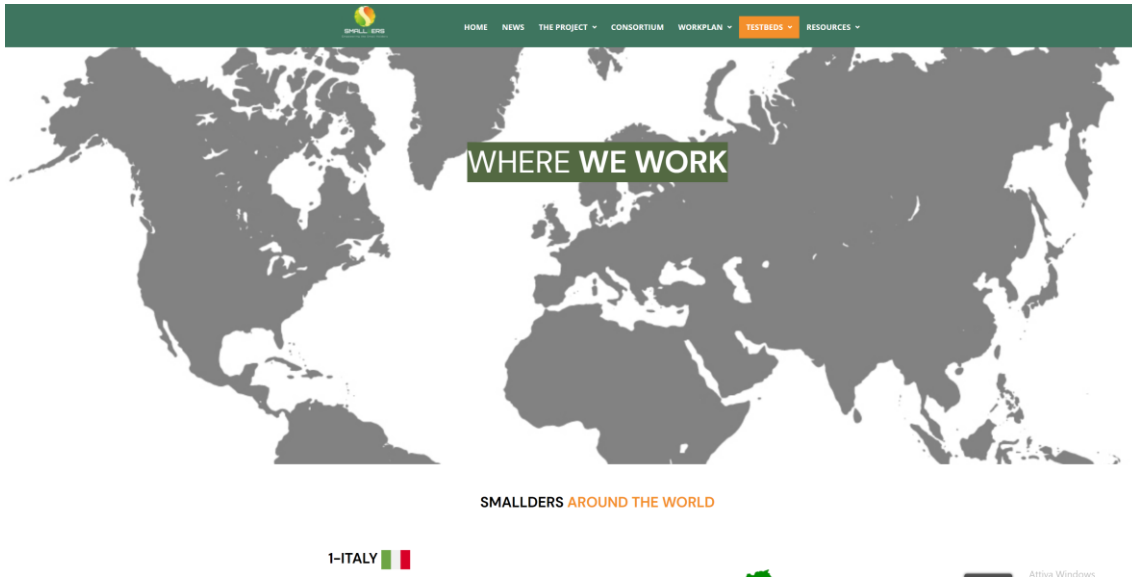
Figure 6. SMALLDERS website: testbeds

Finally, Figure 7 shows the last section of the website, which concerns the "resources", intended as scientific publications, dissemination materials and public deliverables. All these resources will be updated promptly, once they become available.
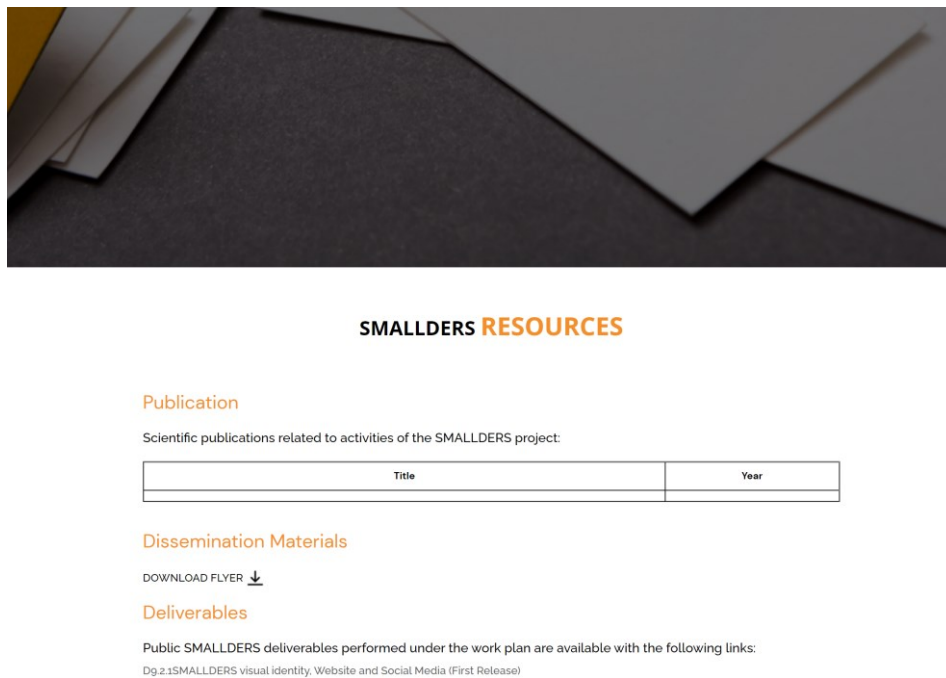
Figure 7. SMALLDERS website: resources

The SMALLDERS website is hosted by OVHcloud. A backup of the database and the pages is performed periodically. It will be maintained in accordance with modern guidelines in terms of website lifecycle management. The website will be available for the entire duration of the project and for at least 5 years after its end.

2.4.2 SMALLDERS Shared Folder on Google Drive

The second SMALLDERS repository is the shared folder on Google Drive, which will accumulate all the project data. A part of this repository will be reserved and therefore visible only to the consortium, while another part will be public. The repository sections that can only be consulted by the consortium will be safe and password protected. Basically, all SMALLDERS members, properly authenticated and authorized, will be able to use this type of repository to store, consult and manage data. The structure of the SMALLDERS Shared Folder is explored below and it is organized in a way that helps partners and external people to easily access all the data and information.

As shown in Figure 8, the first level of this repository contains 6 main paths:

- **Communications with the Project Officer**: it contains all the exchanges of emails, sorted by date, between the consortium (and in particular the Coordinator) and the Project Officer. This portion of the repository is extremely important in order not to neglect communications or deadlines and especially to keep all partners updated.

- **Consortium Agreement & Grant Agreements**: this folder contains all the official documents concerning the agreements between the consortium partners, therefore Consortium Agreement and Grant Agreements. Also in this case, it is an important portion of the repository, extremely

20

confidential, which can be consulted at any time by the partners to always be informed about rights and duties.

- **On the Press**: given the importance of the project, a part of this repository is dedicated to the presence of SMALLDERS on the press. During the course of the project, the relationship between the consortium and the press will be important in order to guarantee a valuable dissemination of results.

- **Partners Contact List**: This part of the repository contains all the basic information about the people who contribute to the project (i.e., email address, telephone number, affiliation, etc.). Basically, this is useful information to facilitate the exchange of data and the communication within the consortium.

- **Proposals, Final Version and Financial Data:** this is the part of the repository used by the consortium when writing the project proposal, which was then funded. Therefore, it contains a history of the project proposal, therefore also the final version, which can be easily consulted at any time by the partners to verify the achievement of the planned objectives. In addition, all project financial data are contained here.

- **Project Activities**: this is the part of the repository that is updated and consulted most often as it follows exactly the progress of the project. As it can be seen from Figure 9, several subfolders are available in order to guarantee a high findability of the project data:
    - **Deliverables**: it contains the final versions of the deliverables already submitted on the MEL Platform. Basically, all partners can easily consult what has been done up to the project at any time. In

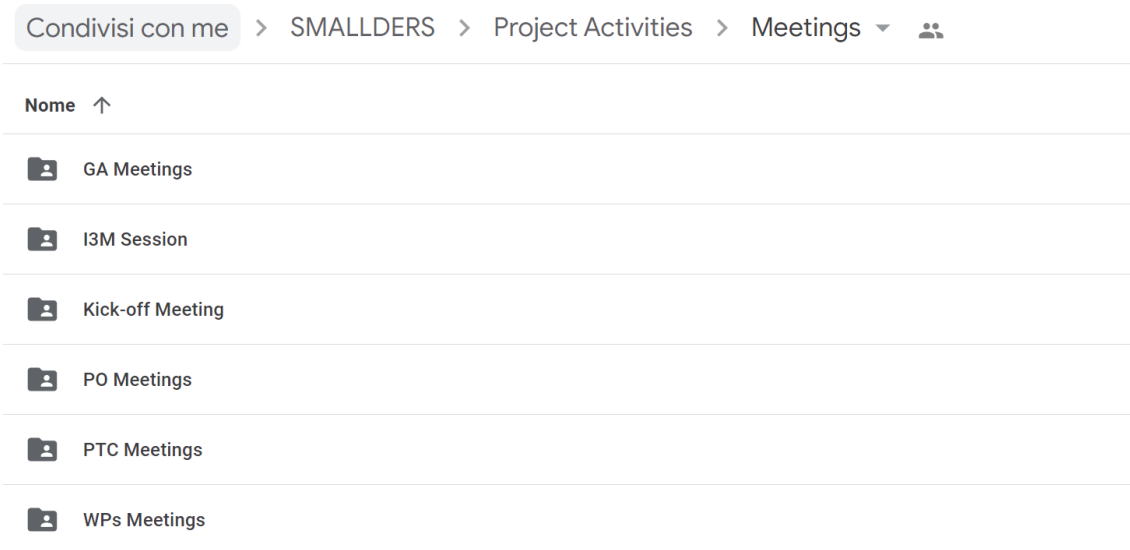this context, the portion of the folder that contains deliverables labeled as public is obviously public.

o **Meetings**: it contains agenda and minutes of all the kinds of meetings of the project (see figure 10): General Assembly, Kick-Off, meetings with the Project Officer, PTC Meetings, WPs Meetings. Moreover, here are included meetings at conferences or workshops (e.g., so far, the Special Session at the International Multidisciplinary Modeling & Simulation Multiconference(I3M)).

o **Project Presentations**: it contains all the files related to presentations (especially .ppt or .pdf files) of the project during different events.

o **Project Website**: it contains data and files in support of the SMALLDERS website (e.g., photos to be uploaded to the site, draft files for parts of the site to be designed and uploaded online, etc.).

o **Publications**: it contains all the publications of the project, divided per typology: conference papers, journal articles.

o **Reviews**: it will contain all the reviews, which will be communicated to the Consortium about deliverables submitted to the platform over time.

o **Templates & Logo**: It contains the project logo in different formats and all the templates to be used in the project documents (e.g., template to fill meeting minutes, template for deliverables, etc.).

o **WP1-WP9:** as it can be seen, a folder is used for each work package in order to ensure order and consistency. For each WP, there is a subfolder for each expected task. In this way, all the technical project documents (e.g., draft of the deliverables, numerical data, etc.) are associated with the respective tasks.
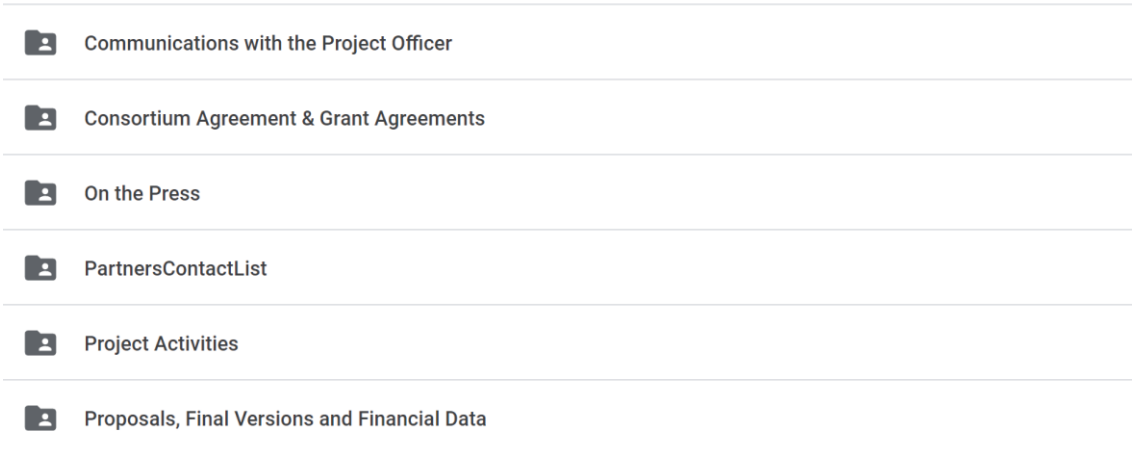
Figure 8. SMALLDERS Shared Folder on Google Drive: first level
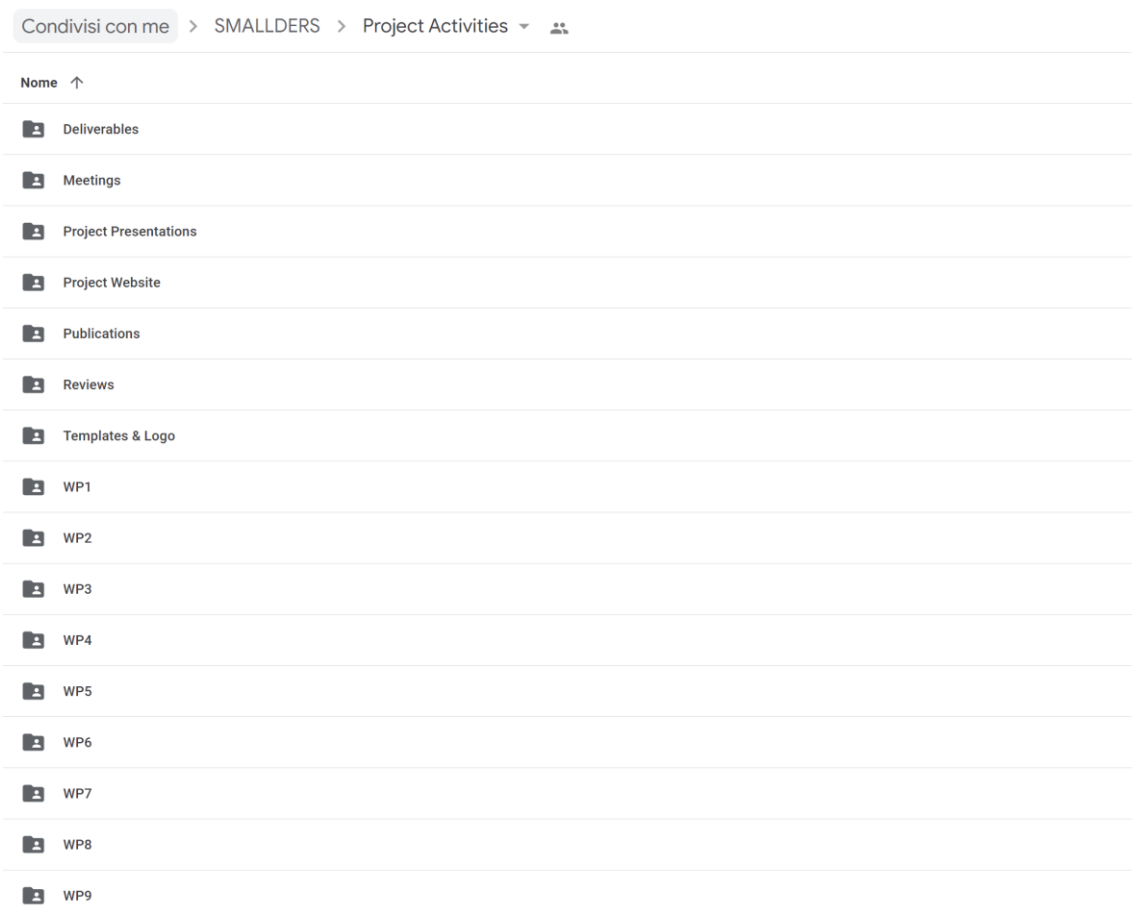
Figure 9. SMALLDERS Shared Folder on Google Drive: detail on Project Activities

Condivisi con me > SMALLDERS > Project Activities > Meetings ▾ 👥

Nome ↑

📁 GA Meetings

📁 I3M Session

📁 Kick-off Meeting

📁 PO Meetings

📁 PTC Meetings

📁 WPs Meetings

Figure 10. SMALLDERS Shared Folder on Google Drive: detail on Meetings

Overall, the tree structure of the SMALLDERS Shared Folder (on Google Drive) was designed to ensure maximum findability of data, therefore to make data management activities as effective and efficient as possible.

3 FAIR Data Management

Data management within the SMALLDERS project will be compliance with the FAIR Data Principles.

3.1 Data Findability

Findability is linked to the ease in being able to retrieve the necessary data, effectively and efficiently. As mentioned in the previous sections, the two repositories of the project are organized in such a way as to ensure that it is easy for all types of actors to retrieve data and information. The Shared Folder on Google Drive is organized according to a very precise and clear hierarchy, which allows members of the consortium and external visitors to quickly find everything they need. Similarly, the website provides for the presence of sections, which concern the main aspects of the project. Furthermore, the use of a precise and shared nomenclature to identify the files minimizes the possibility of errors in finding the data.

3.2 Data Accessibility

Accessibility is related to the ease of accessing various types of data and software documentation (referred to the SMALLDERS Platform). This principle will be guaranteed first of all by the scientific publications, which will all be Open Access, so they can be consulted by anyone. In some cases, reference will be made to the Gold Open Access Policy while in others the Green Open Access Policy. Furthermore, the deliverables of a public nature will be promptly made available on the website, so that the audience can always be informed about the progress of the project. With regard to confidential materials, contained only

in the Shared Folder on Google Drive, partners will be able to access through appropriate authentication.

3.3 Data Interoperability

Interoperability is the ability of two or more systems to exchange and use shared information. The SMALLDERS project envisages a series of actions in the idea of ensuring adequate levels of interoperability. First of all, as specified in the previous subsections, it is planned to use well-known and recognized formats for the various files in order to facilitate the exchange of information both within the consortium and with external parties. Therefore, SMALLDERS adheres to the main standards available. The main interoperability issues will be related to the communication between the various layers of the platform. In this context, well known middleware will be used; such middleware will be decided as part of the requirements specifications and final architecture definition; details will be included in the deliverables that are part of WP3.

3.4 Data Re-usability

Reusability refers to the ability to reuse data for different purposes, effectively and efficiently. The SMALLDERS project fully embraces the logic of reuse with a view to sustainability. As mentioned, the SMALLDERS platform will be released, at the end of the project, as a prototype. Nevertheless, partners expect to keep working on the SMALLDERS platform even after the completion of the project to reach the market. Once ready, the business model behind the SMALLDERS platform will be based on open source and free solution for public administrations in accordance with the concept of "Software Re-Use Catalogs for Public Administrations". The idea is to make the SMALLDERS platform available to public entities through their re-use catalogs, and, therefore, free of charge. Considering

that many public entities do not have the sufficient knowledge to install, run and maintain the software on their own, the business model is as follows: revenues do not come from the sale of the software (which, for public administrations, is free of charge, as said), but from services (e.g. , installation, training, maintenance) regulated by a "maintenance & training contract", which will be needed by the public administrations over time. Basically, this strategy allows the SMALLDERS platform to be continuously reused over time, thus adequately justifying the investments made to conduct research activities. In addition, the project data will remain available for a certain period of time, even after the end of the project to allow everyone to reuse it.

4   Data security & Ethical Issues

Data security will be guaranteed throughout the entire life cycle of the SMALLDERS project. All partners will guarantee the security and confidentiality of the data in the various collection/generation/ dissemination phases. Data that is not publicly accessible will be safely stored in the Shared Folder on Google Drive. Furthermore, for security reasons, all partners will follow strategies for a regular back-up.

Confidential data will be appropriately stored for a certain period of time. As regards users' sensitive and personal data, data anonymization techniques will be applied, in order to guarantee anonymity.

Data collection, storage, protection, retention and destruction will be managed by the SMALLDERS consortium according to General Data Protection Regulation (GDPR) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons and to all national regulations in force in each Country involved in the project. Respect for people and for human dignity will be guaranteed. Rights and interest of all research stakeholders will be protected.

SMALLDERS deliverable D1.3 "Ethics and Privacy Plan" analyses extensively the ethics and privacy aspects of the project. It specifies all the considerations taken into account within SMALLDERS regarding the research with human participants during the pilot and evaluation execution. It also specifies the security measures, guidelines and rules that will be followed by SMALLDERS, to ensure the users' privacy and safety, respecting their rights as volunteer test subjects.

5   Conclusions

In this deliverable, the data management plan of the SMALLDERS project has been detailed. Specifically, the necessary information was provided in terms of expected types of datasets, recommended data formats, nomenclature to identify the files and schema to describe the data collected and / or generated within the project. The two project repositories have been described, namely the SMALLDERS website and the Shared Folder on Google Drive. Furthermore, the approaches according to which the consortium will pursue compliance with the FAIR Data Principles were detailed. Finally, preliminary information was provided on security and privacy issues, which are dealt with in greater detail in deliverable D1.3.

All the practices described in this deliverable will be adopted by the consortium and will be fundamental for the achievement of the project results. However, it is important to point out that as the project evolves, new needs will be discovered, so it may be necessary to enrich and update some parts of the DMP. Possible new additions will be detailed in the subsequent deliverables, in particular in the D1.5 - Mid Terms progress report and D1.6 - Final progress report.