foodpaths

FOODPathS Data Management Plan (DMP)

DELIVERABLE 8.2: DATA MANAGEMENT PLAN

SUSANNE HANSEN

foodpaths

Data Management Plan

Document ID	D8.2
Due date	30th November 2022
Submission date	16th December 2022
Dissemination level	Public
Work package	WP8
Author(s)	Susanne Hansen (AU)
Document version	1

Grant agreement	101059497	**Duration**	42 months
Start Date	June 2022	**End date**	Novembre 2025

Contributors

Name	Organisation
Susanne Hansen	AU
Hugo De Vries	INRAE
Damien Guimond	INRAE Transfert
Emmanuelle Lagendijk	INRAE Transfert

Revision history

Version	Date	Reviewer	Modifications
V0.1	07/10/22	/	Input from WP leads and deputy leads in Forms
V0.2	15/11/22	/	Initial version done
V0.3	25/11/22	/	Feedback from Project lead and deputy lead
V0.4	07/12/22	/	Feedback from WP leads and deputy leads
V.1	13/12/22	/	Deliverable updated for finalisation

Table of content

1. Data summary

The project will not generate research data *per se*, however as part of the implementation project data from various stakeholders will be gathered and shared within and between various work packages. Data will be used in the formulation of R&I needs for the SRIA (WP6), in educational needs (WP5), in the co-creation of a Hub of Food Systems Living Labs (WP4 & 7), in the Food Systems Observatory (WP2) and in the definition of funding strategies (WP3).

The data will be generated and re-used across the different WPs.

The project is expected to produce both qualitative and quantitative data, including survey, interview, and workshop/focus group data (Table 1).

Table 1: 'Research' Data Produced in the Different WPs

Work package	Data produced	Purpose
WP1	Organisational data of partner organisations	Project management
	Personal data of project participants to annual meetings	
	Documents for realisations (contracts etc.)	
	Data for reporting to the EC (financial)	
	Progression data	
WP2	Case study data of territorial food systems	Data collected to gain insight into territorial food systems and to map and analyse Co-creation concepts.
	T2.1: Literature reviews	
WP4	Literature review	
	Field visits – stakeholders and experts	
	Expert interviews – stakeholders and experts	
WP 3	Surveys of funders	Set up network of funders. Understand practices in funding to help with transnational funding and calls
	Interviews and focus groups of funders	
	Events	
WP5	Data re. educational programs, exemplary university-driven food ecosystems, etc.	To create a repository of knowledge regarding educational programs.
WP6	Public consultation	Ensure SRIA is fit for purpose
	SRIA Workshops with stakeholders	
	Survey and focus group interviews with stakeholders within science based policy advice	Overview and information re. organisation of science advice processes To present overview of types of processes and needs for development
WP 7	Semi-structured interviews	Mapping existing SSFS actions
	Desk and field research.	
	Consultations with non-EU experts	
WP8	T8.2: Personal contact data	Subscription to receive information on the project via email.

The size of the data is not yet known, but it is expected to be rather limited, as it will be collected to facilitate the project development, not as part of a research project.

Data quality assurance will be limited to ensuring that the data collected is fit for purpose through the use of templates, interview guides and questionnaires, validating answers (i.e. ensuring that surveys have been completed), and data cleaning.

Outside of the project the data collected may be useful to people and organisations working within sustainability and food domains, working within funding strategies and calls, as well as people and organisations involved in setting up transnational networks. The overall aim of FoodPathS as a CSA is to support the coming HE partnership SFS (P-SFS), and the consortium behind P-SFS may wish to use some of the data for their processes.

The project will primarily makes use of Microsoft Office software for the processing and analysis of data, as the volume of data will be low. File formats will be included in the metadata descriptions.

2. FAIR data

2.1. Making data findable, including provisions for metadata

The projects chosen repository, ERDA[1], includes the possibility of ascribing a Digital Object Identifier (DOI) to chosen datasets. Once the data processing is complete, each dataset will be assigned a DOI, so that it is findable and persistently citable. No datasets have been publicly shared yet, as most have yet to be collected (M6).

The consortium has also established a naming convention for project datasets. Aside to the temporarily identified title, shared project datasets name should contain the following items:

- A prefix "FOODPathS" indicating a FoodPaths dataset
- "WPx" indicating which Work Package the data was produced in
- The title of the dataset
- Date yyyymmdd
- Versioning number

For instance, a dataset produced in WP6 would be named: FOODPathS_WP6_Workshop_20221118_v1.0.
Data versioning will follow the Major.Minor numbering rule, similar to software versioning systems (e.g. v2.1). Major data revision will indicate a change in the formation and/or content of the dataset. Minor revisions will rather involve quality improvement over existing data items. Only minor edits are expected along the project implementation although major revisions are possible beyond the end of the project.

Each dataset will also be annotated with metadata that includes the identifier of the described data and that are indexed in a searchable resource to increase data findability. ERDA follows the Data Documentation Initiative (DDI) metadata standards that includes key data documentation such as:

- Creators and affiliations
- Data location and DOI
- Chosen license for data sharing
- Contributors
- Subjects
- Dataset overview incl. variable lists/codebooks
- Methodological information
- Software and tools information
- Date of first version and dates of subsequent changes

Different levels of confidentiality are considered within FoodPathS:

- Confidential to partner. This option is applied when, regardless of the long-term value and scope for wider use, the dataset contains personal data that cannot be protected once disclosed. These include among others videos and images collected during the project tests.
- Confidential to consortium (including EC services). This option is applied for data containing confidential information (e.g. related to personal interviews) or those with no wide-scope of use and long-term value.
- Confidential to consortium and Advisory Board (AB). This option is applied for data where advice from the AB is requested. This includes data that are relevant for related Partnerships in the wider agrifood domain.
- Public. This option is applied to most datasets.

While datasets confidential to partner will be safely stored by the developing partner, all other datasets will be shared via the SharePoint platform or via ERDA.

---

[1] ERDA or Electronic Research Data Archive at University of Aarhus (AU) is meant for storing, sharing, analysing, and archiving research data. The intended audience is employees, their collaboration partners and students.

2.2. Making data interoperable

The consortium will strive to collect and document the data in a standardized way to ensure that datasets can be correctly understood, interpreted, and re-used. Aside from metadata, the datafiles will include relevant templates, surveys, interview guides and codebooks used to generate the data.

Where necessary partners will be asked to provide documentation describing the main variables included in the datasets in order to support the interpretation and re-use. Standard vocabulary will be used for all data types present in the dataset to allow inter-disciplinary interoperability.

The interoperability of data, in particular, is also relevant while guiding the future P-SFS consortium at their initial activities.

2.3. Increasing data re-use

Aside from metadata, the data files will include relevant templates, surveys, interview guides and codebooks used to generate the data. Where necessary partners will be asked to provide readme files.

Public data will be made available for re-use. To avoid any potential doubt, the consortium will attach specific licenses to the deposited data to define all conditions under which the work is provided under either open or restricted access.

Data produced in WP2 to 8 will be shared under the creative commons license CCby, to ensure that the authors of the original datasets are credited for their work. The data will not include any personal data (contact information or names of respondents).

Data collected in WP1 however, cannot be shared, as the data collected as part of the project management is explicitly intended for internal use to the project management and the consortium. Data collected in WP8 Task 2 will also not be shared externally, as this data consists of contact information of individuals who have signed up for the newsletter. Most of the data will be made available internally, to partners with confidentiality status of consortium member.

All data will be stored in ERDA as soon as possible, at the latest upon publication of the related public reports and will remain re-usable for a minimum of 10 years.

3. Other research outputs

As the project will not generate research data per se, all the outputs identified in M6 have been included in the main sections of the DMP.

4. Allocation of resources

At this preliminary stage of the project, the only costs foreseen for data management are related to:
- The working time needed to set up and perform the data collection, including synchronisation of devices, and analysis activities.
- The working time to setup local and shared data collection devices/servers.
- The working time needed to write documentation, metadata, etc.
- The working time needed to set up and perform the data collection and analysis activities.
- The working time needed to share relevant data with the future P-SFS consortium, with the Advisory Board and any other relevant actor in the development of the future P-SFS.
- Writing up of results in open access publications and reports.

Dedicated financial resources have been already allocated in each partner budget for such activities.

Since ERDA is free, no costs are predicted to come from long-term-storage in the repository, however, eventual long-term storage costs can be incurred for data confidential to the single partner.

The project Coordinator is in charge of the DMP from both the scientific and technical perspective. AUs role includes the first version release as well as the regular update.

Validation and registration of datasets and metadata, as well as backing up data for sharing through open access repositories is the responsibility of the partner that generates the data in the WPs.

Each partner will identify a specific responsible person for each dataset. Quality control of these data is the responsibility of the relevant WP leader, supported by the Project Coordinator.

Each partner should respect the policies set out in this DMP.

5. Data security

As mentioned, partners are responsible for data security and backing up datasets while collecting and processing the data. Generally, local backup system will be guaranteed during the project lifespan.

Once uploaded to ERDA, data will be backed up and stored on internal AU servers with state-of-art data protection software and processes.

6. Ethics

The project consortium fully agrees that the protection of personal data is a priority, and will ensure that all partnership activities, and particularly stakeholders' engagement, dissemination, and communication activities, are conducted in accordance with this EU's General Data Protection Guidelines (GDPR). Further, in case of publication of data, all scientific ethical principles will be respected.

As a general rule, the partners will strive to anonymise personal research data before making them openly available, thus fulfilling both the open research data and data protection rules. The partners will also follow the dissemination rules as setup in the Consortium Agreement. Nevertheless, complete anonymization is not always possible, especially in case of raw data, such as taped interviews.

Partners will request informed consent to disseminate data in public reports, communications etc. as well as for long-term storage. At the same time informed consent can never legitimise the use of data in an open access environment considering that the purposes for further use of data are unknown. In such cases data will be kept confidential.

7. Other issues

We will not be making use of other national, sectorial or departmental procedures for data management in the CSA.

foodpaths