MIXED
EFFICIENT AND RESILIENT MIXED FARMING & AGROFORESTRY

MIXED

GA no:	862357
Action full title:	Multi-actor and transdisciplinary development of efficient and resilient MIXED farming and agroforestry-systems
Call/Topic:	Climate-smart and resilient farming
Type of action:	Research and Innovation action (RIA)
Starting date of action:	1 October 2020
Project duration:	48 months
Project end date:	30 September 2024
Deliverable number:	D8.1
Deliverable title:	Data Management Plan - DMP
Document version:	Ver3
WP number:	WP8
Lead beneficiary:	1-AU
Main author(s):	Tommy Dalgaard (1-AU)
Internal reviewers:	Simon Moakes (2-FiBL), Catherine Pfeifer (2-FiBL)
Nature of deliverable:	ORDP
Dissemination level:	PU
Delivery date from Annex 1:	M6
Actual delivery date:	27.09.2023 (M36)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 862357

Executive Summary

This document describes the MIXED project data management plan (task 8.3). The data management plan is based on the H2020 guidelines. The data will be systematically managed and stored for the use in the project in general as well as for the future of the EU. In accordance with the data minimisation principle, only necessary data related to the implementation of the project are included.

The task will further include management of knowledge and IPR issues according to the rules and agreements stated in the Consortium Agreement. The deliverable schedule outlines how the collected and generated data will be handled during and after the project.

Abbreviations

AGROVOC	Linked Open Data set about agriculture of The United Nations www.fao.org/agrovoc/
Corine	Copernicus Land Cover Monitoring Service
CORDIS	Community Service and Research Information and Development Service
D	Deliverable
DMP	Data Management Plan
DOI	Digital Object Identifier, mainly to identify academic, professional & government information
EC	European Commission
FADN	European Commission Farm accountancy data network
FAIR	Findable, Accessible, Interoperable Reusable
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
MDEF	Meta-data entry form
MiFAS	Mixed Farming and Agroforestry Systems
MIXED	Acronym for the www.MIXED-project.eu
NT	National Team
NTL	National Team Leader
ORCID	Open Researcher and Contributor ID to uniquely identify authors and contributors of scholarly communication
ORDP	Open Research Data Pilot
PID	Persistent Identifier for Data
WP	Work Package
WT	Work Task

Contents

1. Introduction of the Data Management Plan of MIXED 5
2 Data Summary 5
3 FAIR data 10
3.1 Making data findable, including provisions for metadata 10
3.2 Making data openly accessible 11
3.3 Making data interoperable 12
3.4 Increase data re-use (through clarifying licences) 13
4 Allocation of resources 13
5 Data security 14
6 Ethical aspects 16
7 Other issues 17
8 Annex 18

1. Introduction of the Data Management Plan of MIXED

The Data Management Plan (DMP) for the MIXED-project lays out how the project will secure and manage data and in compliance with data management policies of the FAIR (Findable, Accessible, Interoperable Reusable) data management principles and the General Data Protection Regulation (GDPR). The DMP has guidelines, for the consortium members and those managing and using the data.

The DMP is a 'living document' that will be updated when final decisions on experimental design have been taken.

A DMP describes the data management life cycle for the data to be collected, processed and/or generated by MIXED. To make the research data findable, accessible, interoperable and re-usable (FAIR) the plan includes information on the handling of research data during and after the end of the project. It explains what data will be collected, processed and/or generated and which methodology and standards are applied. It explains how the data will be shared and made open access (Open Research Data Pilot (ORDP) and how data will be curated and preserved after the end of the project.

In this document, firstly a detailed data summary is provided for each work package, task and experiments conducted in each task. This is followed by sections on the FAIR data management principles and implications of the project, data security and ethical aspects (also specified in the WP9 ethical deliverables, and the privacy policy of WP7), as well as the GDPR principles' integration into the project.

2 Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

The proposed research in MIXED involves the collection and processing of primary "Farm level general data" to be collected in WP2. Case studies (WP2, 3 and WP4), "Input for feed-back on decision support tools for farmers (WP5), and "Feed-back on policy level decision support" (WP6).

The following data summary show the type of data created in the project. The activities are shown per work package and task. The task leader and activity leaders with the task are also shown. Within the work package 2, several different experimental trials are conducted per task and then the country of the trial is also show. The data summary gives the headline data used or generated here by activity. The subject of the data e.g. project data, Human/farmer related data (fx position of farmer in value chain, efficiency and resilience indicators), soil data, crop data, livestock data, environmental data, economic data, social data.

What types and formats of data will the project generate/collect?

The types of data that will be collected and stored are specified in the summary below:

Work package 1 - Developing efficient and resilient mixed farming and agroforestry systems through the multi-actor approach and transdisciplinary learning"

- Participatory workshops with groups of farmers and their key influencers: interview transcripts and detailed workshop minutes incl. schemes, photos, and summaries (potentially transcripts of workshops).

Work package 2 - Co-design and evaluation of farm systems for efficiency and resilience

Co-design approach to identify with farmers the technical and/or organizational changes furthering the integration between crops, livestock and trees and with the capacity to improve farm efficiency and resilience in the face of climate change.

- Participatory workshop will be organised in each network for co-designing MiFAS based either on the outcomes of WP1 or of previous projects
- Quantitative and qualitative analysis of the scenarios designed
- Develop indicators of resilience and efficiency at the farm level for use in T2.3 and 2.4.
- Human/farmer related data (e.g. position of farmer in value chain, efficiency and resilience indicators): economic data, social data.
- MIXED handbook which will include the following data: a) Livestock Indicators; b) Crop Indicators; c) Soil Indicators; d) Resilience Indicators; e) Efficiency Indicators and graphical representation; f) Environmental impact Indicators.
- These will include basic agronomic and environmental indicators, and indicators developed to measure integration from the point of view of the farming system (e.g. % of pastures in rotation with crops, stocking densities), of work organization (e.g. versatility of farm workers, workload) and of sales management (e.g. % of farm products based on products from at least two enterprises).
- Farm management data collection and collation. This uses semi-structured interviews to collect and collate farm management, technical and economic data required
- Field testing strategies for increased integration. North (DK), North West (UK), Eastern (PL) and Southern Europe (PL) using both research station (2.4 a, b) and on/farm experiments

Work package 3 - Landscape and regional impacts: Bottom-up and top-down approaches for the assessment of multiple farm MiFAS for increased resilience

- Review and integration of existing research, data and literature
- Graph-based modelling approach for farm interactions; data can be represented by fluxes of manure or feed, or even information, some semi-qualitative information might be asked to farmers in the networks involved in the work package.
- Mapping mixed landscapes in Europe using existing data; Using widely available datasets such as FADN and GIS mapping layers to determine relevant indicators to characterise mixed landscapes in Europe. The approach will be applied systematically to all of Europe and will build on previous work in the FP7-CANTOGETHER project, and data obtained from the Farm Accountancy Data Network database, and used in accordance with the rules of DG-AGRI.
- Data will be collected in selected case study networks in order to have a measure or an estimation of fluxes of feed and nitrogen between farms, in supplement to variables describing farm; Guidelines for data collection will be provided and data will be collected in selected farm network case studies.

Work package 4 - Value chains – environmental and economic impacts

- Climate and environmental impact of products from mixed and agroforestry systems;
- Report on impact of MiFAS on costs along value chains;
- Mapping of value chain network;
- Guideline on data collection and analysis; Data on relationships including information flow and the use of contracts will be collected with 3 to 6 key value chain actors (face-to-face

interviews) and a sample of 10 to 20 farmers (depending on the size and complexity of the network) through an online survey.
- Report on governance in the value chain;

Work package 5 - Farm level decision support for optimising efficiency, resilience and environmental benefits

- Performance of mixed and agroforestry systems, using farm data from WP2
- Test version- Prototype of the serious game available
- Farm scale impacts of mixed system adoption, using typical farms defined from FADN data collated in WP3
- Agroforestry system development tools, using farm data from WP2
- Final version-A landscape level serious game, working with farmers in multiple countries
- Labour use questionnaire, using farm data collected through WP2

Work package 6 - Multiscale integrated assessment of efficient and resilient MiFAS

- Report on multi-scale assessment framework for mixed farming systems
- Report and open access paper on efficiency and resilience assessment at farm level.
- Report and open access paper on upscaling of efficiency and resilience analysis to community, regional, national and EU-level.
- (As part of T6.4) Participatory workshops with groups of farmers and their key influencers: interview transcripts and detailed workshop minutes incl. schemes, photos, and summaries (potentially transcripts of workshops).
- Report with policy recommendations and open access paper on transition scenarios towards mixed farming and agroforestry systems including policy recommendations.

Work package 7 - Communication, dissemination and out-reach

There will be no research data generated in WP7. Data generated in WP7 will only serve the purpose of monitoring website, social media and newsletter activities

- The www.mixed-project.eu traffic is monitored via the system Site-Improve.
- The activity on SoMe channels are administered via the tool HootSuite
- Newsletter statistics on subscribers, opening rates etc. are monitored by the system Mailchimp, by which the newsletter is distributed.

Will you re-use any existing data and how?

The proposed research in MIXED also involves the further processing of secondary personal data. Permission will be obtained from the national statistical offices and other primary data holders of the case study countries and usage will be in line with the procedures described below in addition to the data safety procedures prescribed by the respective statistical offices and data holders.

The types of secondary data that will be collected and stored are specified in the summary below:

Work package 1 - Developing efficient and resilient MiFAS through the multi-actor approach and transdisciplinary learning

- Data collected in national participatory workshops with farmers and industry stakeholders from previous projects. The type of data used could include detailed meeting minutes, meeting summaries and photos.

Work package 2 - Co-design and evaluation of farm systems for efficiency and resilience

- National / International published data on land use, fertiliser use, feed quality, nutrient contents, livestock production, crop production and climate.

Work package 3 - Landscape and regional impacts: Bottom-up and top-down approaches for the assessment of multiple farm MiFAS for increased resilience

- The Cordis database will be used for a review of EU projects
- FADN, Corine, Copernicus data as well as other open geographical data selected after a review will be used for the top-down analyses.

Work package 4 - Value chains – environmental and economic impacts

- Data based on WP2 supplemented with existing local value chain information

Work package 5 - Farm level decision support for optimising efficiency, resilience and environmental benefits

- National and international published data related to farm system design and characterisation, including livestock nutrition and production, crop inputs and productivity.
- LCA inventory databases including the commercial ecoinvent LCIA database
- FADN (secondary data), will be used and stored during the project. Post-project only average results can be retained

Work package 6 - Multiscale integrated assessment of efficient and resilient MiFAS

- We will use FADN, national statistics and GIS data, and climate variables

Work package 7 – N/A

What is the origin of the data?

The proposed research in MIXED involves the collection and processing of the following primary personal data, to be collected in WPs 1, 2, 3, 4, 5 and 6

Work package 1 - Developing efficient and resilient MiFAS through the multi-actor approach and transdisciplinary learning

- Summaries, transcripts and photos from participatory workshops with farmers, academics and key industry stakeholders.

Work package 2 - Co-design and evaluation of farm systems for efficiency and resilience

- Network hold personal contact details (phone numbers, email addresses)
- Farm system, economic and labour data at the farm-level, for selected farms in selected networks, in coordinated with the WP3, WP4 and WP5 data collection.

Work package 3 - Landscape and regional impacts: Bottom-up and top-down approaches for the assessment of multiple farm MiFAS for increased resilience

- FADN data (see also WP6); only secondary data are public available
- Some semi-qualitative information asked to farmers

Work package 4 - Value chains – environmental and economic impacts

- Surveys in the regional value chain cases studied

Work package 5 - Farm level decision support for optimising efficiency, resilience and environmental benefits

- Farm system, economic and labour data at the farm-level, collected through WP2
- FADN data (farm type averages) supplied through WP3 analysis.
- Many secondary sources from literature and management handbooks

Work package 6 - Multiscale integrated assessment of efficient and resilient MiFAS

- FADN data; only secondary data which are publicly available

Work package 7 – N/A

The proposed research in MIXED involve data collection and processing in 10 countries – 8 members states (Denmark, Germany, the Netherlands, Poland, Romania, Austria, France and Portugal) as well as Switzerland and the UK. Any data transferred from the EU to a non-EU country (i.e. Switzerland) or international organisation will be in accordance with chapter V of the General Data Protection Regulation 2016/679. Each country has a National Team Leader (NTL) that will ensure all ethical and data management issues in this document are dealt with at the local level, while the coordinator of the project (AU) will centrally oversee the ethics and data security management. All primary data collection will be undertaken in all case study countries.

This table provides an overview of the research involved data collection and processing in 10 countries.

	DK	**CH**	**UK**	**NL**	**FR**	**AT**	**PL**	**DE**	**RO**	**PT**
Developing efficient and resilient mixed farming and agroforestry systems through the multi-actor approach and transdisciplinary learning (WP1)	X	X	X	X	X	X	X	X	X	X
Co-design and evaluation of farm systems for efficiency and resilience (WP2)	X	X	X	X	X	X	X	X	X	X
Landscape and regional impacts: Bottom-up and top-down approaches for the assessment of multiple farm MiFAS for increased resilience (WP3*)	X		X	X	X			X	X	X
Value chains – environmental and economic impacts (WP4)	X	X		X		X				X
Farm level decision support for optimising efficiency, resilience and environmental benefits (WP5*)	X	X	X	X	*X*	X	X	X	X	X
Multiscale integrated assessment of efficient and resilient MiFAS (WP6)	x	x	x	x	x	x	x	x	x	x
Communication, dissemination and out-reach (WP7)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

* The data is collected through WP2 for all networks, but the serious game will in particular work with partners in FR, NL, DK, UK, PT

What is the expected size of the data?

The primary data collection is likely to generate small datasets only in the order of magnitude of maximum a couple of Gigabytes in total. Secondary data collection is likely to generate a larger number of datasets. To be further detailed, if it appears during the project.

To whom might it be useful ('data utility')?

Our main target audience to disseminate the project data consists of (1) academics and (2) a broader audience of farmers' organizations, insurance companies, financial Institutions, policy makers, NGOs and consumer organizations that are interested in the sustainability and resilience of the EU agricultural and/or food sector.

3 FAIR data

3.1 Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

MIXED as on Open Research Data Pilot aims to make data open whenever possible, but as closed as necessary when taking into consideration personal data and privacy. If it affects personal confidentiality and privacy, data will not be shared publicly. Aiming for making the open data as discoverable and easily used as possible; this includes making use of machine readable, persistent and unique identification measures such as DOIs and presenting adequate metadata.

What naming conventions do you follow?

We will follow standard academic data conventions like Latin naming conventions, SI units (The International System of Units) and Harvard reference standards. Search keywords are provided that optimise possibilities for re-use and we use clear version numbers and dates for all information and documents generated. The UN FAO AGROVOC Linked Open Data set about agriculture may be used as reference when suitable (see also section 3.3). The metadata created will include WP, Task, Experiments, Lead and participating partners, other participants. But also, experimental data. The data will be made publicly available through uploading them in a data repository and made identifiable trough a PID (Persistent Identifier for Data) for each dataset and the associated metadata, as defined in the project.

Will search keywords be provided that optimize possibilities for re-use?

In order to ensure that the open data is findable, representative keywords will be provided and links to the data will be maximised in any (open access) journal publication related to the project.

Do you provide clear version numbers?

Version numbers will be used in order to refer to the latest version of the data, a version history will be appended to the metadata.

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

For each dataset described in the data summary an extended meta-data entry form (MDEF) will be created. This data form will hold metadata on the experiments and the dataset collected or hold. It will be kept concise and described with precise language to make the data finable. The meta-data

entry forms will be updated through the project life-cycle. The data will be held in a simple spreadsheet format readable by a relational database software. They will be identifiable with standard identification mechanisms.

The data collected in the MDEF include the PID, origin of the data, description of the data, purpose of the experiment, duration and location, and the size and scale (number of raw and processed/calculated parameters and number of values in each parameter including missing values). Geographical data will follow ISO 19115 standard.

Final versions of documents and deliverable to be openly available are published on the project website and on Organic e-Prints. Pervious and working document versions will be evaluated yearly (see appendix II for yearly overview of MIXED publications) Pervious and working document versions are stored on the Aarhus University SharePoint site to which all partners have secure data access.

3.2 Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Data, apart from that which includes sensitive data and general personal data affected by GDPR, or raising any ethical concerns (unless consented), will be shared.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

The farm accountancy data network (FADN) monitors EU farms' income and business activities. FADN data will be used in WP3 and WP 6 and the primary dataset will only be accessed by named individuals, however, averages can be stored publicly. MIXED will follow EU guidelines for accessing the FADN data (Farm accountancy data network, European Commission (europa.eu))

How will the data be made accessible (e.g. by deposition in a repository)? Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

A file will be provided with metadata and direct link to all those datasets during the research on SharePoint. After publication, on open access servers and repositories. Metadata will be located on a dedicated page of the MIXED SharePoint site which will remain active for a minimum of 5 years after the completion of the project.

What 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 project's data will be deposited in the secure repository of the Aarhus University SharePoint platform. This is in accordance with the GDPR-compliant regulations of the host and in accordance with the Aarhus University regulation on data management. No specific or proprietary software will be required to access the data, rendering specific access documentation or software obsolete.

Have you explored appropriate arrangements with the identified repository?

The project's publications and data will be deposited at PURE and Organic Eprint hosted at the Aarhus University platform.

If there are restrictions on use, how will access be provided? Is there a need for a data access committee? How will the identity of the person accessing the data be ascertained?

In accordance with the Grant Agreement all research related data (excluding personal and sensitive data) will be stored for at least five years after the end of the research project (in case there is a high interest in the datasets or due to different national legislation, data may be stored for a longer period, which will be transparently discussed and approved within the consortium and relevant parties). Data that are used for publication will be stored at least five years after publication.

Are there well described conditions for access (i.e. a machine readable license)?

The project coordinator, Aarhus University will use its repository called 'PURE' to provide the access point for the university open-access research repository. Once an output is deposited it will be further validated by Aarhus University Library Research team. The record will be checked to ensure that deposit in 'PURE' is in line with the policy of the publisher and that the full text version, and the information regarding the output is correct 'PURE' is hosted on secure servers, fulfilling EU regulations on data protection. The university's repository system does not require any specific software and provides open access according to national and EU regulation.

3.3 Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering 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)?

MIXED will strive to produce data in the project which are interoperable as possible by (1) focussing on facilitating data exchange and re-use between stakeholder by using open access repositories, (2) making use of standard data formats (generally flat text files or open spreadsheet formats) that can be opened and manipulated using standard open software applications (such as Open Office), and (3) facilitating re-combinations with different datasets from different origins by including appropriate the metadata and links to the original data collection methods, and (4) by using standard vocabularies in alignment with the UN FAO AGROVOC.

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

In the unlikely case that uncommon or project-unique specific ontologies or vocabularies have to be used we will provide mappings to more commonly used standard vocabularies for all uncommon data types to allow maximum interoperability. Any assumptions made, and the mapping steps to more commonly used vocabularies will be documented in a readme file and stored on SharePoint.

Data produced through MIXED consumer and other stakeholder research (e.g. semi or unstructured (narrative) interviews, focus groups and citizen jury participant information or observation accounts, photographs, videos, drawings) will use standard formats (e.g. TXT, DOC, XLSX JPEG, RAW, AVI formats) and made compliant with available (open), software applications, facilitating the recombination with various datasets from different origins. As with the field trial and control experimental data, during instances where the use of uncommon vocabularies is unavoidable, or where it is necessary to generate project specific ontologies mappings will be provided to more commonly used vocabularies/ ontologies and stored on SharePoint.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

Standard vocabulary that is commonly used in the project's research domain—sustainability and resilience of the EU agricultural and/or food sector—will be used in the dataset itself, the related metadata, and any accompanying documents. As the project is multi and trans-disciplinary in kind, we foresee no issues regarding overly specific jargon.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Yes, in case. However, this is not expected.

3.4 Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

Data will be Open Access licensed data after considering of personal data, intellectual property rights and any additional legal and ethical requirements. Based on The Open Research Data Pilot of the European Commission (what-is-the-open-research-data-pilot (openaire.eu)), we aim to allow data to be re-used by third parties, but with restrictions if IPR or other rights demanding such restrictions. Data licensing will be based on guidance provided by Aarhus University. Copyright of the data a*re based on EU H2020 guidelines*.

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Access to the research data will be dependent on any agreed 'embargo period' based on national and EU regulations. The 'embargo period' is applied to give time to publish the work or seek patents, were applicable and this will be as short as possible until the work is accepted for publication or patent, bearing in mind that research data should be made available as soon as possible.

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

In case of sharing data or restricting certain data with third parties outside of the consortium, a data sharing agreement will be set up that will detail anonymising or aggregating data, participant consent for data sharing, copyright permissions, and agreement on an embargo period. Data will be used in standard forms allowing reuse, as well as allowing searchability. Data quality assurance processes will be undertaken, including applied standards and methodologies.

How long is it intended that the data remains re-usable?

Data will be stored based on the contractual terms, until which it can be re-usable (all research related data will be stored at least for five years after the end of the research project). Because of the combined natural and social science nature of the data there is no time limit for its re- usability.

Are data quality assurance processes described?

MIXED will ensure data quality, data generated in the project will be (1) scrutinized by several experienced researchers, (2) used in peer-reviewed scientific publications and (3) discussed in-depth with a co-creation platform and/or through various stakeholder workshops.

4 Allocation of resources

What are the costs for making data FAIR in your project?

The project budget specifically includes costs for 6 open access publications (equalling 18'000 Euro, budgeted within WPs). Other cost for FAIR data management handling are included in each of the WP budgets, see specifications below.

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

A budget of € 3.000 was foreseen in the project budget for each open access publication (6 open access publications for each of WP1, 2, 3, 4, 5 and 6, and the WP corresponding lead partners of ABER, SRUC, INRAE, AU, FiBL, WU). In total € 18.000.

Who will be responsible for data management in your project?

The project coordinator (1-AU) is responsible for coordinating the data management efforts, with the main contact person Tommy Dalgaard (tommy.dalgaard@agro.au.dk) with support from the Information Specialist of AU library Anne Mette Emdal Navntoft (amen@au.dk).

In addition, the data management support hub will also provide assistance (https://medarbejdere.au.dk/research-data-management/). The work package leaders are responsible to oversee the data flows, requirements and security issues mentioned below within their respective work packages and to report back to the project coordinator (1-AU).

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

Data are stored throughout the life time of the project and at least 5 years after (see point above), and handled by the relevant partners in the project WPs, under supervision of the Project Steering Committee (PSC). Longer term storage cost are held by individual partners

5 Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Data to be stored

The aim is to ensure that data sets and records provide sufficient information to external parties to establish the scientific integrity of the information and that the scientific outcomes are reliably based on the evidence collected by researchers. The following data will be stored:

Data used in publications

- Sample copies of records used to collect information (questionnaire, fieldwork notes, samples of diaries, audiotapes, complete questionnaires, photographs, videos, drawings and films etc.)
- Analysed research data (such as datasets held in databases, spreadsheets, midterm reports, provisional findings)
- Supporting documents (date of the collected information, source of the information, place, notebooks, etc.

Data that might be used in future research

This includes interim reports, final reports and drafts for manuscripts in academic or applied farmer-facing literature. Subject to adherence to any restrictions arising from the need for compliance with GDPR and ethical consent procedures (e.g. respondent anonymity) the following types of data will be stored:

- Original data e.g. video and audio recordings, hard copy surveys, paper or digital experimental records

- Raw data, e.g. unprocessed digital data, site measurement data, plant and soil samples, digital audio, photographic or video recordings
- Processed data sets used in publications, including information to understand the processing, in particular coding steps, data cleaning logs, missing values, metadata for each data set
- Where applicable Intermediate data sets to guarantee transparency and scientific integrity of data processing
- Data sharing agreements

Data storage

Original data and paper records will be kept in lockable cabinets or offices with controlled access, when not under the direct supervision of a member of the research team. Access to electronic data and records are controlled by passwords and, where appropriate, access to individual files folder or databases will also be password protected. Passwords will be known only by authorised individuals. Access controls will regularly be reviewed and updated as individuals join, leave or change roles within the project. Computers and software will not be left logged in and unattended. Digital data will normally be stored in the format in which they were generated.

The project-specific shared repository SharePoint is hosted by Aarhus University. Access is limited to authorised individuals with password. SharePoint uses of a folder structure based on the functions and activities of projects to store records. This makes it easier to share information with other authorised personnel. Back-ups are implemented locally and checking procedures are established to ensure that the system works effectively. Small media devices such as USB sticks or CDs are not considered suitable as the long-term primary storage location for personal data or confidential information and will be encrypted if used to transport respondent identifiable data or data that would cause significant harm or distress to somebody if released.

Accessibility

Privacy and confidentiality of data are guaranteed at all times, in line with legal requirements, Aarhus University policy and best practice. Normally, data in process are only available to the researchers involved in a project. Access by third parties will have to be negotiated with the project leader unless otherwise specified. If ethical questions or concerns about the integrity of research data or data management are raised, access to data has to be granted in line with legal requirements, Aarhus University policy and guidance as well as host institutions policy, and, where applicable, European Commission policy. Established guidelines for academic good practice, as for example provided by academic associations or research councils, will be consulted where appropriate.

Is the data safely stored in certified repositories for long term preservation and curation?

Nearing the project's final stages, we will openly discuss how the project's data will be safely stored and maintained beyond the project's duration. However, the coordinating institution currently hosts both Organic Eprints and PURE on the Aarhus University platform, and that will remain after the project, so there is already an infrastructure in place.,

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

In order to safely store and share data within the consortium and potentially with third parties, we make use of our Sharepoint platform, PURE-system and organic eprint. As the server of these platforms is located and managed by project partners, the safety of the data is ensured and falls under EU legislation.

The D8.1 Data Management Plan and the present implementation of security measures have been presented during the project consortium meetings and General Assembly, whereby all partners have agreed to take all the relevant measures related to this.

6 Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review.

The ethical aspects were already described in the deliverable 9.1-9.3x, and they are referenced here again to make the DMP a stand-alone document. The following text already has the feedback from the ethics review incorporated.

The procedures and criteria that will be used to identify/recruit research participants will be outlined in (D9.1-9.3). The informed consent procedures that will be implemented for the participation of humans, will be specified before recruitment commences (D9.1). Templates of the informed consent forms and information sheets covering the voluntary participation and data protection issues (in language and terms intelligible to the participants) will be submitted as a deliverable before the relevant recruitment commences (and is available from D9.1, section 5).

Several tasks within the MIXED project will involve the recruitment of human research participants/ farmer network survey. In each case, we will ensure that the appropriate best practices are adopted in relation to recruiting research participants and that there are clear and demonstrable rationales for recruitment. In each case the procedures and criteria that will be used to identify/recruit research participants along with a clear justification for these strategies will be recorded and collated within deliverable D9.1.

In the case of developing a comprehensive database of key stakeholders interested in contentious inputs in mixed farming (as hosted in task 2.1) the rationale behind the recruitment strategy will be to draw on and develop the pre-established networks of MIXED partners to ensure that we are able to disseminate the project's results to as broad an audience as possible. As such, we will recruit a very broad range of stakeholders, including: policy makers at the EU, national and devolved nation level (e.g. federal states, autonomous regions); organic producers and producer organisations, organic certifying bodies and assurance schemes; animal, plant and soil scientists, environmental and animal welfare NGOs. Stakeholders will be recruited primarily through personal contact using our existing networks - advisory we will adhere to current ethical standards regarding informed consent and data protection.

Informed consent will be secured from all human participants when participating in any primary data collection activities within the MIXED project (for example, farmer surveys). All participants will take part in the research voluntarily and can decide at any moment to discontinue their participation. The MIXED WP leaders will finalise and approve an information sheet for all potential local and regional participants in the project, providing a clear written account of the goals of the research, the methods employed and use made of its results as well as its relevance to the respondent and the implications of participation. This document will be translated into all languages that will be associated with primary data collection. The information sheet will explain any potential risks that might be involved

(for example being identified as part of a particular group, or providing personal information) and how confidentiality will be ensured throughout the collection, analysis and dissemination of data. The information sheet will also provide information on how and who to contact regarding any further questions about the project. The information will be explained in person at the start of each new data collection activity, after which the participants will be asked to sign an informed consent form indicating their consent to participate in the research. Permission will be sought to digitally record data certain data collection activities explaining that all possible steps will be taken to ensure confidentiality and anonymity of participants at the level of the individual (unless participants request otherwise). The researchers will ensure that participants are aware of the archiving process, the accessibility of the data and ask their consent to their information being stored in this way. MIXED will follow Aarhus University's data protection rules, which are in line with EU laws on data protection (in particular EU Directive 95/46/EC, from 25/05/2018 the new General Data Protection Regulation No 2016/679, and EU Directives 2002/58/EC and 2006/24/EC), and national law applicable in the relevant case study countries. A detailed overview of all ethical committees involved as well as the relevant additional local laws and/or codes of conduct can be found in the table in Annex.

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Informed consent will be obtained from all participants in the project's research by presenting them with an informed consent form that will be explained by the enumerators and then needs to be signed by the participants. The case study partners will ensure that these forms will be made available in the local language, that the research will not contain any content that could be considered offensive to the participants and that participants will be allowed to ask any remaining questions to the enumerators. The informed consent sheets will also contain information sheets with clear details on (i) the aims and methods of the research, (ii) the wider implications of the research project, (iii) the nature of participation (expectations, benefits, potential discomforts), (iv) the voluntary base of participation and (v) how data will be securely and anonymously handled and shared and preserved. This information will be explicated in its entirety by the enumerators in the local language.

7 Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

The AU data management regulations and data management plan format will serve as a minimum data management policy within the consortium.

HISTORY OF CHANGES		
Version	**Publication date**	**Change**
1.0	2 July, 2021	Initial version of DMP shared with the consortium
2.0	26 September, 2022	Revised version based on Periodic Review comments
3.0	27 September, 2023	Updated list of scientific publications in Table 8.2, as reviewed during the project meeting in Toulouse

8 Annex

Table 8.1 presents an overview of the relevant ethical committees and additional local laws and/or codes of conduct regarding data management of all National Teams

Partner	Country	Ethical commission	Contact	Relevant laws/codes
1. AU	Denmark	The Research Ethics Committee	Chief Consultant, Tove Bæk Jensen tbj@au.dk	- The AU Code of Ethics - Commission on Academic Ethics Rules
2. FiBL	Switzerland	Data manager	Lisa Haller, dataprotection@fibl.org	- Data protection at FiBL
3. Aber	United Kingdom	Project manager	Pip Nicholas-Davies, pkn@aber.ac.uk	-
4. WU	Netherlands	WASS Social Sciences Ethical Committee (SEC)	Dr. Esther Roquas, Esther.Roquas@wur.nl	- The Netherlands Code of Conduct for Scientific Practice - Integrity code of Wageningen UR
5. SRUC	Scotland, UK	Social Ethics Committee	Chair Klaus Glenk, Klaus.Glenk@sruc.ac.uk	- SRUC Ethical approval process
6. INRAE	France	INRAE Ethics Committee	General secretary, Christine Charlot, christine.charlot@inrae.fr	- INRAE code of ethics and scientific integrity
7. BOKU	Austria	Project manager	Professor Werner Zollitsch werner.zollitsch@boku.ac.at	-
8. IUNG-PIB	Poland	Disciplinary Proceedings Representative at the Scientific Council	IUNG-PIB Disciplinary Committee Professor Jerzy Grabiński Jerzy.Grabinski@iung.pulawy.pl	- The IUNG Code of Ethics, Poland

Partner	Country	Ethical commission	Contact	Relevant laws/codes
9. IEA-AR	Romania	Ethical Committee of the IEA-AR	Dr. Marioara Rusu, rusu.marioara@gmail.com	- Ethical committee Standing Rules of IEA-AR
12 CONSULAI	Portugal	Data Management Board Data Protection Officer	Pedro Santos, psantos@consulai.com rgpd@ulisboa.pt -	- National Labor Law - General Data Protection Regulation (GDPR)
15. IFLS	Germany	Not existing	Simone Sterly, sterly@ifls.de	-

MIXED Scientific Publications

Table 8.2 Overview of scientific publications (S) and data (incl. metadata) produced in MIXED. The list is evaluated yearly at project meetings.

no	type	title	authors	journal	vol.	data and metadata (incl. reference)
1	S	How does Life Cycle Assessment capture the environmental impacts of agroforestry? A systematic review	Mónica Quevedo-Cascante, Lisbeth Mogensen, Anne Grete Kongsted, Marie Trydeman Knudsen	Science of The Total Environment	Volume 890, 10 September 2023	10.1016/j.scitotenv.2023.164094
2	S	Mixed farming and agroforestry systems: A systematic review on value chain implications	Guy Low, Tobias Dalhaus, Miranda P.M. Meuwissen	Agricultural Systems	Volume 206, March 2023	10.1016/j.agsy.2023.103606
3	S	A participatory approach based on the serious game Dynamix to co-design scenarios of crop-livestock integration among farms	Julie Ryschawy, Myriam Grillot, Anaïs Charmeau, Aude Pelletier, Marc Moraine, Guillaume Martin	Agricultural Systems	Volume 201, August 2022	10.1016/j.agsy.2022.103414