Data Management Plan

MATS Deliverable 7.4

Summary

The data management plan (DMP) covers the principles and protocols for data collection, processing, sharing and storage. The DMP will ensure that data management is compliant with the EU's Guidelines on Findable, Accessible, Interoperable and Reusable (FAIR) Data Management in Horizon 2020, and with relevant national data protection laws and institutional data management policies.

Deliverable title: Data Management Plan
Deliverable number: D7.4
Authors: Mila Hyytinen, Qiuzhen Chen, Bodo Steiner, Karlheinz Knickel
Due date: 31 December 2021
Submission date: 20 December 2021
Nature[1]: R
Dissemination Level[2]: PU
Work Package: WP7
Lead Beneficiary: University of Helsinki (UH)
Contributing Beneficiaries: all partners

---

[1] R = Report, P = Prototype, D = Demonstrator, O = Other

[2] PU = Public, CO = Confidential, only for members of the consortium (including the Commission Services)

Table of contents

1. Introduction

MATS is a 3.5-year project implemented from July 2021 to December 2024. The consortium consists of 14 partners in Europe and Africa, including six universities (in Finland, Spain, Greece, Netherlands, South Africa, and Switzerland), four research institutions (in Italy, Netherlands, Tanzania, and Germany), two civil society organizations (in Uganda and Belgium), and two small and medium-sized enterprises (in Italy and Greece).

MATS aims to identify key leverage points for changes in agricultural trade policy that foster the positive and reduce the negative impacts of trade on sustainable development and human rights. The focus is on improving the governance, design, and implementation of trade practices, regimes, and policies at national, EU, African, and global levels. The MATS project develops and pilots new tools for systemic analysis and assessment of the interactions between agricultural trade, investments, sustainability, and development.

The research data collected and generated will be made available to the research community and public to the extent possible while considering ethical, personal data protection, and IPR matters.

MATS is participating in the Pilot on Open Research Data in H2020. All partners are committed to the open dissemination of knowledge and have the shared objective to maximize open access to project data and results. The project is compliant with EU principles and standards, relevant national data protection laws, and institutional data management policies. The management of qualitative and quantitative datasets follows established principles of good data management at every stage of the data lifecycle, ensuring that data is findable, accessible, interoperable, and reusable (FAIR).

2. Data summary

2.1 Purpose of data collection and generation

The MATS project consists of 7 work packages (WPs) presented in **Figure 1**, and of WP8 'Ethics Requirements' in which three ethics issues (Humans, Protection of Personal Data, Non-EU countries) have been identified according to the Research Executive Agency's ethics screening of the MATS proposal.

The data types collected and processed in MATS cover surveys, interviews, workshops, online questionnaires, photographs and videos for online dissemination, conferences, exhibitions, and publications.
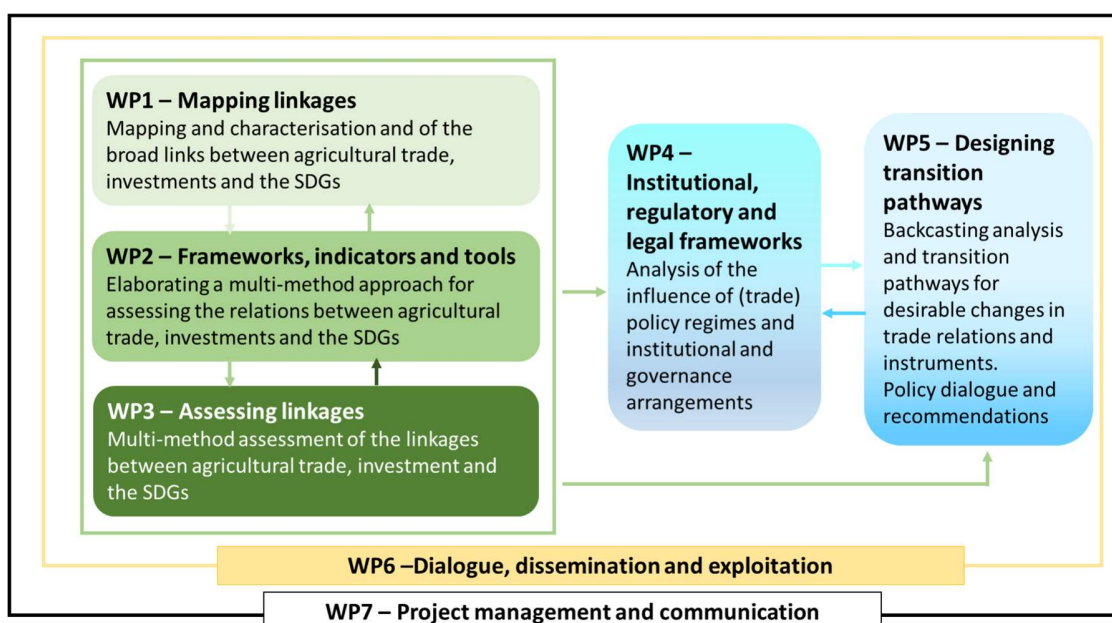

Figure 1. Work packages (WPs) 1-7

WP1: Mapping linkages

The overall goal of WP1 is to provide a review for mapping and characterization of the broad links between agricultural trade, agricultural and rural investments, environmental sustainability, and human well-being. The tasks in WP1 are implemented in M1-8 (i.e., July 2021 - February 2022), the initial

phase of project implementation. Existing statistical data related to agricultural trade, market, investment, and environmental indicators are used.

More specifically, the following databases and data sources are utilized:

- World Integrated Trade Solutions (WITS): https://wits.worldbank.org/
- United Nations Conference on Trade and Development (UNCTAD): https://unctad.org/
- Climate Change Knowledge Portal (World Bank Group): https://climateknowledgeportal.worldbank.org/
- World Trade Organization (WTO): https://rtais.wto.org/UI/PublicMaintainRTAHome.aspx
- WTO data source: https://data.wto.org/en https://agims.wto.org/
- WTO 'Trade and Environment': https://edb.wto.org/charts
- CEPII, Research and Expertise on the World Economy: http://www.cepii.fr/cepii/en/cepii/cepii.asp
- Global Preferential Trade Agreements Database (GPTAD): https://wits.worldbank.org/gptad/database_landing.aspx
- United Nation's Department of Economic and Social Affairs – Sustainable Development: https://sdgs.un.org/goals
- ITC Standards Map: https://standardsmap.org/
- ICT data source: https://www.intracen.org/itc/market-info-tools/
- FAO data source: https://www.fao.org/faostat/en/#data

WP2: Frameworks, tools, and indicators

The overall goal of WP2 is to identify relevant frameworks, indicators and tools for assessing the relations between trade and the sustainability of human well-being. There is no primary data collection in this work package. The secondary data used are from literature reviews and published data sources, including official statistics.

The Sustainable Development Goals (SDG) indicators are identified, and their suitability assessed for analyses of business and sustainability relationships. A complete indicator list will be created from the SDG indicators Metadata repository (https://unstats.un.org/sdgs/metadata/), scientific articles, and other official information from the United Nations for identifying trade-related

indicators. The data from the case studies and consultations of the case study leaders will also be used. WP2 provides an enhanced analytical framework for WP3.

WP3: Assessing linkages

WP3 includes a set of 15 case studies to provide a deeper understanding of trade impacts at the local market and the societal level and how these shape national, regional, and global socio-economic and environmental trends. Data comes from interviews, focus groups, workshops, structured questionnaires, and surveys. Some interviews will be recorded. The data collection for the 15 case studies will start approx. M13 (i.e., July 2022). More detailed information about data collection and processing will be obtained from the case study leaders during 2022.

Table 1 provides an overview of the key products, main locus of country, and Sustainable Development Goals per case study product and area are shown in.

Table 1: Overview of the key aspects for the 15 case studies

Case study			Sustainable Development Goals*						
Product	**Main locus of FVC**	**Lead**	SDG1	SDG2	SDG3	SDG 6	SDG 13	SDG 15	Others
1. Coffee	Uganda, Tanzania	UH	�(shaded)	▢	▢	▢	▢	▢	▢
2. Oats	Finland, Sweden, Norway	UH					▢	▢	▢
3. Dairy	Finland	UH					▢	▢	▢
4. Food products	Sub-Saharan Africa	ESRF	▢	▢			▢		
5. Poultry	Ghana	UPM	▢						▢
6. Cacao	EU, Côte d'Ivoire	OWW	▢	▢	▢			▢	
7. Milk powder	EU, Africa	OSOL	▢	▢				▢	
8. Biofuels	EU, America, Africa, Asia	OSOL	▢	▢			▢	▢	
9. Coffee	Tanzania, DR Congo, Burundi, Uganda, Ethiopia	OWW			▢	▢	▢	▢	
10. Beef	Germany, France, Italy, USA, Brazil, Argentina, Morocco, South Africa and Namibia	CRPA	▢						
11. Beef, milk	Angola, Vietnam	CRPA	▢						
12. Wine	South Africa	NWU	▢						▢
13. Dairy	Germany, France, Italy, USA, Brazil, Argentina, South Africa, Zimbabwe, Algeria	CRPA	▢						
14. Pork	EU, Brazil	UPM	▢	▢					
15. Food, etc.	EU, Tunisia	TNI	▢	▢					
Number of case studies per SDG			**12**	**12**	**9**	**8**	**10**	**12**	**8**

* SDG1 No poverty; SDG2 Zero Hunger; SDG3 Health, Well-being; SDG 6 Clean Water; SDG 13 Climate Action; SDG 15 Life on Land

WP4: Institutional, regulatory and legal frameworks

The overall goal of WP4 is to gain a better understanding of the role of institutional, regulatory, and legal frameworks given the impacts of agricultural trade on the SDGs and in respect of global agreements on environmental and climate challenges.

There is no primary data collection and processing foreseen in this work package. WP4 will use information and materials that are collected from the 15 case studies in WP3. In addition, secondary data will be used from a review of academic literature on the role of institutional, regulatory and legal frameworks, and international governance failure, corroborated by recent case studies and counterfactual information made available by WTI's African and European networks.

WP5: Transition pathways and policy recommendations

The overall goal of WP5 is to derive transition pathways for desirable changes in trade relations and instruments and to formulate policy recommendations in respect of global agreements on environmental and climate challenges.

The two tasks of this WP involve a visioning process and road mapping and will use interviews and questionnaires with relevant stakeholders.

WP6: Communications, dialogue, and dissemination

The overall goal of WP6 is to contribute to an enhanced civil society dialogue by providing improved data, analysis, and methods. The project communication plan has been submitted as Deliverable D7.2, and communication toolkits for all project communication are elaborated in Deliverable D7.3.

In WP6, a questionnaire is used for gathering basic information on key stakeholders and decision-makers. Also, their preferences for the sustainable trade platform, their information needs, the type of information required, the way they access it, and information on their availability for collaboration is collected. The targeted respondents include non-governmental organisations, government officials, policymakers, and key actors from international platforms, academia, the private sector, agri-food trade, producer associations, and consumer organisations. Respondents are from subnational, national, European and other macro-regions, and global levels. The online survey is carried out in M6 (i.e., prior to November 2021). In WP6, a semantic framework for resource annotation is elaborated and it will be used for characterizing and categorizing all MATS assets (data, reports, publications) that are made available through the sustainable trade platform.

The core instrument in MATS for connecting the project with relevant decision-makers, stakeholders, and initiatives is the evidence-based communication platform 'Sustainable Trade Hub' (https://sustainable-agri-trade.eu/). The platform provides data, information, and insights at a global scale to relevant stakeholders about agricultural trade, sustainability and investments, data from the MATS project, and external resources. It offers analytics and visualization functionalities and allows users to build their data stories and share

them within the platform and to broader communities enhancing the discourse and communication between stakeholders.

WP7: Project management

The overall goal of the WP7 is to ensure a successful implementation of the project. The project documents will be shared with the project members via the Microsoft TEAMS platform which is also used for internal communication.

WP8: Ethics and Security

All personal data collected and processed in the project will be handled according to EU General Data Protection Rules (GDPR) with pseudonyms. All published data will be anonymous, and any personal data or sensitive information will not be exposed or retained. The visualized data will be exclusively aggregated, with anonymization techniques applied to prevent any ability of back-tracing to the original content.

Ethical requirements will be described in more detail in Deliverables D8.1, D8.2 and D8.3. The data security issues will be described in Deliverable 8.2, which is also containing the Data Privacy Notice.

2.2 Types and formats of research data collected and generated

Table 2 summarises the file formats for data sharing, reuse and preservation are recommended/accepted by the UK Data Service as well as by the University of Helsinki. The MATS project will follow these file formats.

Appendix 1 presents a table with basic information that MATS project partners will use for managing the data they collect and process. The table will be updated and finalised in conjunction with forthcoming updates of the Data Management Plan.

Table 2: Planned/recommended data formats used at MATS project

	Recommended formats for data collection and processing
Audio data	.mp3/.mp4/ .flac
Video data	.mp3/.mp4/.avchd
Textual data	.rtf/txt/.doc/.docs/.html
Tabular data	.csv/.tab
Image data	.tf/.jpeg/.jpg/.jp2/.gif/.raw./.psd/.png/.psd

Source: File formats recommended by UK Days Service
https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/

Description of data types and formats for data collected and processed in MATS:

- **Workshop (collected):** Workshops may be recorded with the appropriate authorization by the participants. Notes may also be taken.

- **Survey (collected):** The information will be collected through suitable online platforms (e.g. Google forms) and the data will be stored in a MS Excel file.

- **Interview (collected)**: Interviews will be recorded with the appropriate authorization of the interviewee. Notes may also be taken.

- **Gray literature (collected):** Reports and general information on the Web or internal to each case will be used for the collection of additional information. The sources will be stored in their original format.

- **Case studies template (generated):** A template will be generated for each case study according to the guidelines elaborated in WP3, Task 3.1 in M9-13 (i.e., March-July 2022).

- **Set of indicators (generated):** The set of indicators will be generated into an MS Excel file.

- **Toolbox (generated):** An electronic document (or platform, space), in which different tools (policy tools, research tools) are presented. This implies a variety of information that will be integrated into a format that

will allow user-friendly navigation and search. The Sustainable Trade Toolbox includes frameworks, indicators and tools for a systemic analysis and assessment of the linkages between agri-food trade, investments, environmental sustainability, and human well-being. The final format is not yet defined but information will as soon as available be included in an update of the DMP.

- **Database (generated)**: From the responses of the questionnaire, a database will be generated. All outcomes (data, discussion papers, policy briefs, scientific publications, videos, research and innovation protocols etc.) generated in the project are made accessible in a very user-friendly way on the interactive Sustainable Trade Hub.

- **Simulation models (generated)**: The models are stored using Vensim Software (the file format is *.mdl).*

Data size

The number of people interviewed/participating in workshops and questionnaires/case studies is not defined yet. The size of the data files, however, will be small and therefore it does not require special planning for storage.

2.3 Data utilization and re-use of the existing data

Planning which data formats will be used in the project is essential for data sharing arrangements as well as during the project. The precise plans for the formats to be used will be confirmed in conjunction with the planning of empirical analyses in WP3. The related information will be provided in the update of this DMP.

3. Findable, Accessible, Interoperable and Re-usable (FAIR) data

The description of the data will be provided for each published dataset in metadata repositories such as the SDG indicators metadata repository (https://unstats.un.org/sdgs/metadata/).

The format of the metadata will follow the Data Documentation Initiative (DDI) recommendations (https://www.fsd.tuni.fi/en/data/misc/ddi-records/) which is an international standard for describing survey, questionnaire, statistical data files, and social sciences study information. Formats, keywords, and coding practices will be decided before data collection. The keywords will be the same as those included in scientific articles wherever applicable.

Regarding the grey literature, suitable keywords will be specified. For the storage of case study data, naming conventions will be agreed.

Project outputs will be made available on the Sustainable Trade Hub platform. It includes presentations, recorded webinars and publications that are relevant for wider audiences, as well as case study abstracts, factsheets, Policy Briefs, etc.

A consortium partner SCIO.P.C, a private company, and developer of the hub, anticipates maintaining it after the project has ended. The plan is to offer basic data analytics services for the end-users free of charge, whereas more sophisticated data analytics will be offered via a subscription fee. The overall data management and visualization solution are based on the SCIO.P.C's Quantum platform that can seamlessly incorporate data ingestion, processing, analysis and visualization components under an extensible architecture.

The more academic research results are disseminated through scientific publications, presentations at international conferences, and workshops. Access to the publications will be provided via open access publishing and/or using repositories such as OpenAire. Publications by the University of Helsinki are openly accessed via Helda (https://helda.helsinki.fi/).

For data storage, Zenodo will be considered as a possible repository (https://zenodo.org/). Zenodo is an open digital repository integrated with OpenAIRE and GitHub, which is a software for version control and data management (https://github.com/).

4. Allocation of resources

Open-access fees and the costs of cloud services, data processing and other dissemination activities have been budgeted in each project partner's budget. The costs of cloud services for the Sustainable Trade Hub are about 7,500€. Costs for anonymizing data can be covered through the respective partners budgets.

5. Ethical aspects and data security

All personal data will be treated as strictly confidential and processed in compliance with General Data Protection Regulation 2016/679. Published data will be anonymous, and no personal data will be made available to third parties.

Consents will be asked from the participants to record meetings and interviews, either by recording the audio or in writing.

The Project Steering Group (PSG) defines access rights to the data in consideration of ethical and data protection issues.

The Consortium Agreement (CA) defines the ownership and access to key knowledge (IPR, research data, etc.). A non-disclosure agreement (NDA) will be used if confidential information is shared within the MATS project. The Project Advisory Group (PAG) is responsible for monitoring matters identified in the ethical assessment and submitting them to the General Assembly when necessary.

The implementation of data management for each WP is the responsibility of the leader of the relevant WP. Within the framework of the DMP, the respective consortium partners are responsible for collecting and transferring the data needed for their research activities. All partners are fully responsible for a

proper understanding and practices regarding the collection, storage, preservation, protection, security and backup of their respective data.

Pseudonymized data from surveys and interviews will be securely stored in a safe location at the project coordinators or partners' facilities. Internally, pseudonymized data may be shared via Teams/Dropbox/OneDrive/WeTransfer/Google Docs.

Contact information for research subject will be stored separately at partner´s facilities and personal data will be handled and shared between the project members only as pseudonyms.

For data sharing, methods for ensuring data security will be used. The Sustainable Trade Hub database is encrypted, and access is given via secure protocols like HTTPS and FPTS. In the case that user-derived documents are uploaded into the platform, a personally identifiable information (PII) identification mechanism will be enacted to examine the document, notify the user on the possible existence of PII and ask for their consent before making the document available.

All data collected and processed in the project, will be internally stored during the project lifetime. Anonymization of the data will be done when the project has been ended.

Appendix 1.

Data type	Collected / produced /reused	Source	File format	Personal / sensitive data	Ownership	Size estimate	Metadata types	Storage during project	Opening