Ref. Ares(2025)758822 - 31/01/2025

FoodDataQuest

Fuelling the quest for sustainable food systems with data-driven innovations

D2.1 Framework Design

Grant agreement:	No 101134138
Date:	31/1/2025
Prepared by:	Xuelin Zhang, Eva Maes, Cor Verdouw, Tom Quets, Siavash Farahbakhsh

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Project no: 101134138

Deliverable factsheet FoodDataQuest	
Document name:	D2.1 Framework Design
Responsible Partner:	EV ILVO
WP:	2
Task:	2.1
Version:	1
Version Date:	31st January 2025

Dissemination level	
×	U = Public
	PP = Restricted to other programme participants (including the EC)
	RE = Restricted to a group specified by the consortium (including the EC)
	CO = Confidential, only for members of the consortium (including the EC)

Summary

FoodDataQuest (FDQ) project aims make our food systems more inclusive and more sustainable by fostering data driven innovation throughout the food chain. To enable this, data sharing plays an important role, especially private data sharing. It helps promote collaborations among all the stakeholders by enhancing transparency and trust in the food value chain. The FDQ project also aligns and contributes to the SDGs which have been promoted by United Nations (UN) since 2015, such as SDG 2 (Zero Hunger), SDG 12 (Responsible Consumption and Production), and SDG 13 (Climate Action). FDQ aligns with the SDGs.

FoodDataQuest consists of 6 Work Packages (WPs) are moving together toward to the same objective (Figure 1): Trustworthy and secure private data sharing in the food value chain. There are 4 Use Cases (UCs), UC1 develops a DNA-driven AI system to enhance olive oil traceability and sustainability; UC2 focuses on reducing hospital food waste and improving patient nutrition by creating a dashboard; UC3 combines agricultural, health, and consumer data to promote sustainable diets, track food footprints, and support healthy choices via dashboards and apps; UC4 aims to enhance consumer awareness of sustainable choices using AI-powered personalized communication to provide targeted, impactful guidance on healthy and eco-friendly lifestyles. And to achieve these goals and help UCs, WP2 will develop a structured framework, identify principles and good practices, specify the project's Reference Architecture, and design a generic data model for private data sharing ensuring transparency and sustainability across the agri-food systems. Specifically in T2.1, it is a pivotal task within the FDQ project, focusing on developing a comprehensive and integrated **framework** to facilitate trustworthy and secure data sharing in the food value chain. T2.1 developed the FDQ framework and defined aspects and elements that are necessary for trustworthy data sharing.

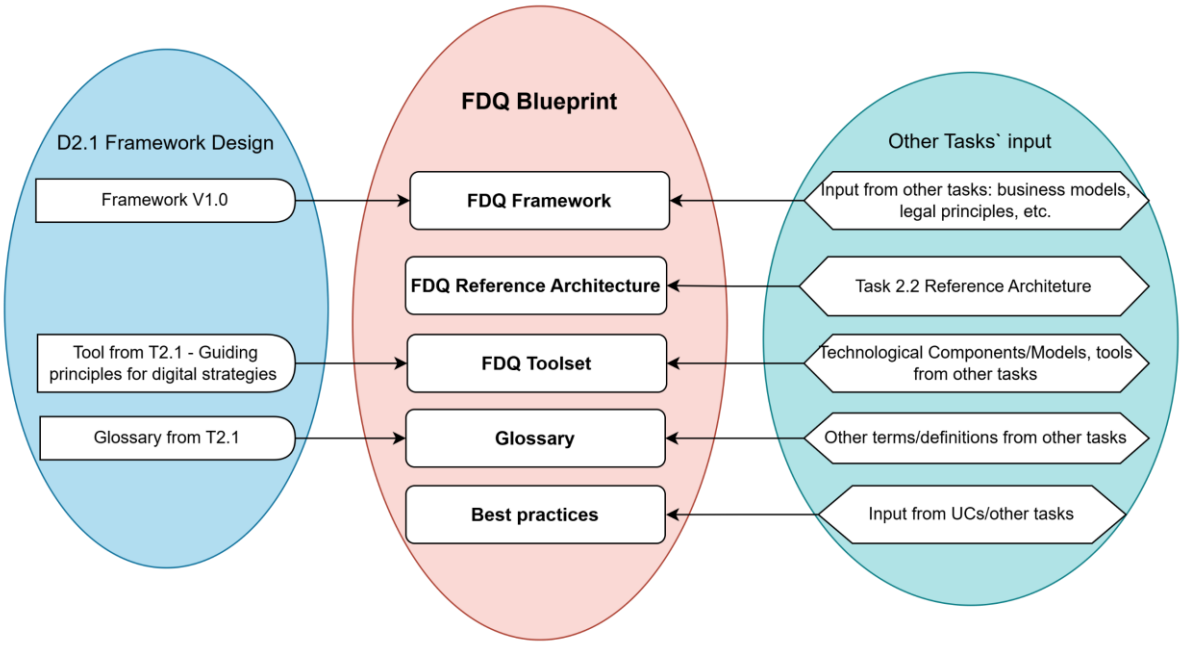

Figure 1 Overview of the FoodDataQuest project

With this context, one of the tangible results of FDQ project, will be the FoodDataQuest **Blueprint** for UCs working on collaborative data-driven innovation, using private data. It is our big ambition to provide the basic understanding of private data sharing in agri-food sector, setup framework, data-driven solutions development for multi-stakeholders throughout the food systems. Hence, the FDQ Blueprint is dynamic and iterative, which is designed to outline principles and strategies for data sharing within agri-food systems. It will be updated with the development of project. This blueprint provides a structured, stepwise approach to ensure adaptability and sustainability across the value chain. It consists of 1) **FDQ Framework**, 2) **FDQ Reference Architecture**, 3) **FDQ Toolset**, 4) **Glossary**, and 5) **Best Practices.**

With this deliverable, which is the result of Task 2.1 (Framework Design for Private Data Sharing), the first version of the FDQ Blueprint is established. The blueprint's design is both modular and adaptable, the results of other tasks will build upon this basis, improve the framework and add other tools, best practices and guiding principles to improve the blueprint.

Generally, WP1, including T1.1 (Mapping the current landscape of private data sharing), T1.2 (Identify gain creators, needs and opportunities), and T1.3 (Define strategies and digital technologies) provide basic understanding for the private data sharing in the food systems; WP2, T2.1 set up the FDQ framework and other tasks including T2.3 (Legal and policy, regulatory frameworks) and T2.4 (Business model development) contribute to the FDQ Framework; T2.2 carries out Reference Architecture and technical specifications for the services and technologies to be developed in WP3 (For 4 UCs). All the technological components, models, and tools developed by tasks will be integrated into FDQ Toolset including the tool that is created by T2.1. The glossary serves as a “mini-dictionary” containing a list of terms or words, along with their definitions or explanations collected from all the tasks coming from the food value chain. At last, the input, feedback, and applications from UCs and tasks will showcase the best practices in the real world of FDQ Framework.

The creation of the FDQ framework is initially from T2.1 Framework deliverable report. See the *Figure 2* below, it shows the basic concept and structure of FDQ Framework. The development process of the framework involved literature reviews, UCs' and stakeholders' inputs, and alignment with existing regulatory frameworks and sustainability goals. Moreover, input was collected from UCs to showcase the practices and improved the framework based on the applications and feedback.

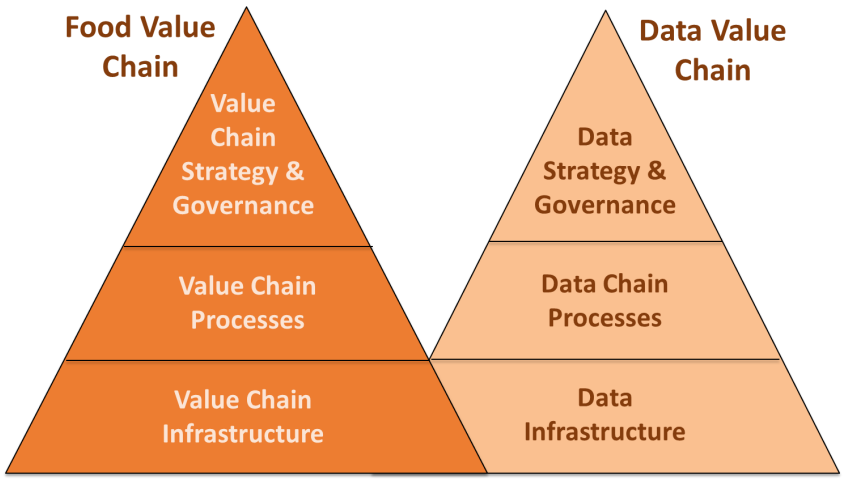

Figure 2 Framework Overview

A tool, which contains a series of questions, is created by T2.1 to provide stepwise approach and guidelines for trustworthy private data sharing and digital strategies. The tool has been piloted and used by the 4 UCs. It also validated that the identified building blocks are appropriate for trustworthy data sharing in food value chain. Additionally, this T2.1 tool is also a valuable resource for other tasks to further develop and refine their models/tools. A **glossary** was also made, which will be integrated into Glossary for the whole project.

The analysis and results of the UCs show that **trust and privacy** are key building blocks for private data sharing in food value chain. Interoperability is also important to successful data sharing among stakeholders. Common challenges include unclear data ownership, data quality issues, and differing requirements across UCs. Ethical concerns, such as consent, are critical for building trust. Due to the fact that UCs are various and at different stages of development, the framework must be flexible and adaptable. It addresses these by offering flexible and practical solutions, ensuring it meets the diverse needs of all UCs across the food value chain.

As a result, the framework remains as designed, with no changes required at this moment(Figure 2). It is structured around two main chains: the food value chain and the data value chain. Each chain is further divided into specific aspects——**The food value chain includes: i) Value Chain Strategy and Governance, ii) Value Chain Processes, and iii) Value Chain Infrastructure. The data value chain includes: i) Data Strategy and Governance, ii) Data Chain Processes, and iii) Data Infrastructure.** Each aspect comprising various elements. These aspects and elements collectively ensure the robustness and effectiveness of the framework.

TABLE OF CONTENTS

Summary **3**
TABLE OF CONTENTS **6**
Abbreviations **9**
1. Introduction **10**
2. Methodology **13**
2.1 What actually is a framework? 13
2.2 Method of framework development 15
2.3 Method of the development of a tool 15
2.3.1 Purpose of the T2.1 Tool 15
2.3.2 Development of T2.1 Tool 15
2.3.3 Input used to develop the T2.1 Tool 16
2.4 The design of T2.1 Tool 18
3. Literature Review **23**
3.1 Data sharing in agri-food sector 23
3.2 Private Data Sharing for Healthy Diets and Sustainability 23
3.3 Trust in data sharing 24
3.4 Address trust and privacy restrictions 27
4. Reference models of data sharing in agri-food **28**
4.1 Existing reference models 28
4.2 DSSC Blueprint 29
4.3 CHAIN 30
4.4 Data Value Chain Model (DATA2X) 30
4.5 Integration of reference models into FDQ 31
5. Definition of the framework **32**
5.1 Food Value Chain 32
5.1.1 Value chain strategy and governance 32
5.1.2 Value chain processes 33
5.1.3 Value chain infrastructure 34
5.2 Data Value Chain 35
5.2.1 Data strategy and governance 35
5.2.2 Data chain processes 37
5.2.3 Data infrastructure 38

6. Tool for framework 40
7. Mapping of the Use Cases 42
7.1 Why are we mapping the UCs 42
7.2 Brief description of Use Cases 42
7.3 Application of the framework by the UCs 44
7.3.1 Value chain process 44
7.3.2 Value chain strategy and governance 45
7.3.3 Value Chain infrastructure 48
7.3.4 Data chain process 48
7.3.5 Data infrastructure 52
7.3.6 Data strategy and governance 54
8. Analysis of the building blocks 58
9. Blueprint and the contribution of T2.1 to the blueprint 60
10. Conclusions 61
11. Glossary 62
12. Reference 65
13. Annex 69
T2.1 Tool 69
UC 1 input 75
UC2 input 82
UC3 input 113
UC4 input 127

List of Tables

Table 1 Process Matrix....18
Table 2 T2.1 Tool....18
Table 3 List of existing relevant reference model, and relation to FDQ....28
Table 4 Questions from Value chain process....44
Table 5 Questions from Value chain strategy and governance....46
Table 6 Questions from Data chain process....48
Table 7 Questions from Data infrastructure....52
Table 8 Questions from Data strategy and governance....54
Table 9 Overview of the UCs applications....58
Table 10 Glossary....62

List of Figures

Figure 1 Overview of the FoodDataQuest project....3
Figure 2 Framework Overview....4
Figure 3 Overview of the framework development methodology....15
Figure 4 Development of T2.1 Tool....16
Figure 5 DSSC Building Block Overview(Version 1.5)....30
Figure 6 Data Value Chain Overview....31
Figure 7 Key data sharing approaches....39
Figure 8 Workflow of the tool....40

Abbreviations

GA ----------Grant Agreement

FDQ ----------FoodDataQuest

UCs ---------- Use Cases

DSSC ----------Data Spaces Support Centre

WPs ----------Work Packages

CTE's ----------Critical Tracking Events

KDE's----------Key Data Elements

GTIN ----------Global Trade Item Number

GLN ----------Global Location Number

CHAIN ---------- Connecting humans in the agriculture industry

DATA2X ----------Data Value Chain Model

SDG---------- Sustainable Development Goals

F2F ---------- Farm to Fork

UKCM---------- Univerzitetni klinicni center Maribor

UCPH---------- University of Copenhagen

NDAs---------- Non-Disclosure Agreements

LCA ---------- Life Cycle Assessment

TAU ---------- Tampere University

WUR ----------Wageningen University & Research

AI ---------- Artificial intelligence

GDPR---------- General Data Protection Regulation

IP(s) ---------- Intellectual Property(Properties)

API----------Application programming interface

GXDCH----------Gaia-X Digital Clearing House

SaaS ----------Software as a service

UN ----------The United Nations

1. Introduction

Project aim

The FoodDataQuest (FDQ) project aims to enhance the efficiency, transparency, and sustainability of food systems by leveraging data-driven innovations. However, significant concerns about privacy, data sovereignty, and the compatibility of diverse systems restrict the adoption of digital technologies in agriculture (Gabriel et al., 2021). To address these challenges, the project's main ambition is to transform agri-food value chain by fostering collaboration among stakeholders while ensuring data sharing aligns with the protection of private data privacy.

More precisely, the FDQ project focuses on developing data-driven solutions in the food value chain, with a focus on the collaboration between stakeholders to improve the efficiency of food production and distribution while promoting environmentally sustainable practices. The project will develop a Blueprint for this type of data-driven solutions, based on a practical, integrated methodological framework. By integrating data from various stakeholders along the entire food value chain—from producers to consumers, and even food waste —FDQ seeks to improve quality and sustainability in the food value chain, reduce waste, improve collaboration and transparency between multiple actors, and promote healthy diets for consumers.

In doing so, the project is organized in multiple work-packages (WPs). Within the first WP, the available (un)conventional data, stakeholders' perceptions of data sharing and data-driven technologies will be investigated. Moreover, in this WP, there will be a closer look at how consumers approach privacy and how they can guided towards sustainable and healthy diets. In WP2, the focus will be on the structure of how to combine all the potential data and data-driven technologies from technical, business and legal point of views. In WP3, data-driven innovations and solutions will be made relying on 4 Use Cases (UCs).

Task 2.1

Within WP2, Task 2.1 is aimed at developing an integrated framework (i.e. FDQ framework) that would primarily comprehend data provenance/collection mechanisms, describing and analyzing available data relying on WP1 and UCs, trust and privacy concerns, considering data types, related actors, etc. Moreover, this task will set the basis for other tasks in the project, that will further build on this framework and dive deeper into certain aspects, or propose guidelines, good practices and tools to be used when working on a data-driven solution. All the research from Task 2.1 is documented in Deliverable 2.1, which is this document.

Use Cases

As anticipated, the project, as well as Task 2.1 specifically, relies on four UCs that will (further) develop advanced data-driven solutions to address critical challenges in agri-food supply chains and consumer

behaviour. An overview of the UCs[1] is provided below, and their characteristics are further elaborated later in this document.

1. **DNA-Based Tools and AI for Food Sustainability and Traceability**: This Use Case targets the olive oil supply chain, aiming to improve sustainability and traceability. By combining DNA-based technologies with AI, the project will create a sustainability protocol that values local species and producers, ensuring transparency from farm to fork. The integration of DNA data enables real time monitoring of risks like climate-induced mutations and mislabelling, while AI will optimize supply chain management, reducing food waste, energy use, and costs.

2. **Reimagining Food Services in Hospitals and Elderly Care Homes**: This Use Case focuses on enhancing patient health and reducing food waste by developing a data-driven system to monitor and improve food services. By providing insights on nutritious and sustainable food options, the system will transform food services in hospitals and elderly care homes, emphasizing the use of locally sourced and organic produce. This use case aims to reduce the carbon footprint of healthcare facilities while supporting a "green" approach to operations.

3. **Farm-to-Fork (F2F) Sustainability Insights for a Healthy, Climate-Neutral Food Chain**: This Use Case seeks to boost sustainability within the food supply chain by integrating farm-level data into decision-making processes. The goal is to create a consumer app that offers detailed information about product nutrition, environmental impact, and sustainability scores. AI tools will help optimize sourcing and processing methods, aligning with climate-neutral goals and enhancing transparency and trust throughout the food chain.

4. **Enhancing Consumer Awareness of Eco-Friendly Choices**: The focus of this Use Case is on empowering consumers to make eco-friendly lifestyle choices through personalized, AI-driven communication. By tailoring complex environmental information to individual preferences across different platforms, this Use Case will simplify decision-making for consumers, promoting sustainable habits without overwhelming them.

The above-mentioned UC not only help to understand the practices, opportunities and needs of different actors in the food value chain but also help test the FDQ framework and the building blocks for good data sharing practices. The input and feedback from the UCs are valuable for further refinement and improvement of the framework design.

Specific Objectives of T2.1

As described earlier, this deliverable, as the main outcome of Task 2.1, focusses on defining the framework, and elaborating on two important elements to take into account for data-driven innovation, i.e. 1) understanding data provenance and collection mechanisms, and 2) defining guiding principles for trustworthy data sharing, as further explained below. The objectives of this task are:

Framework design: Develop a structured and comprehensive framework for trustworthy and secure data sharing across the agri-food sector. The framework consists of a few essential building blocks. By setting up 4 collaborative UCs using data-driven strategies, the building blocks of the framework conclude foundational elements and important aspects. Use a

[1] More information on the UCs, including their detailed aims, objectives, key performance indicators, etc., is provided in D3.1, which is a public deliverable.

stepwise approach and create a T2.1 Tool to support and foster healthy and sustainable choices throughout the food value chain.

Understand Data provenance and collection mechanisms: Data is being collected in many different ways, by different devices, using various technologies and standards. To develop collaborative UCs, there needs to be a good overview on how data is collected at all the stages of the food value chain from the different stakeholders that want to work together, and some guidelines can be agreed upon to make interoperability easier.

Defining Guiding principles for trustworthy data sharing: Trust is one of the basic conditions to develop common data-driven UCs. There are different ways to approach and enhance trust to share data. Based on the current policy and legal aspects regarding the private data sharing across the agri-food system in the EU on farm, sectoral, regional, and national level, define guiding principles to foster collaboration among stakeholders, maintaining data sovereignty and respecting privacy.

Moreover, this task contributes to the WP3, which is focused on UCs. Hence, the produced framework will serve as a bridge between the technical design and real-world applications, guiding stakeholders in understanding how data can drive technology innovation while addressing legal, technical, trust and governance aspects.

Last but not least, T1.1 and T1.2, complementing T2.1, will further examine identifying and analysing private data sharing practices and exploring opportunities for data from different sources. T2.1 examines data structures, integration, as well as different protocols and use of standards of the data produced across the value chain. By addressing these aspects, T2.1 not only supports secure and effective private data sharing but also fosters trust among stakeholders, ensuring that shared data is used responsibly and transparently. Together, these tasks (T1.1, T1.2, T2.1) contribute to building a trustworthy framework for data sharing in the agri-food sector. Furthermore, the data collected within these tasks will provide input for the Reference Architecture (Task 2.2).

The rest of this document is organized as follows: First, a methodological section (Section 2) outlines how to develop the framework (i.e., T2.1 tool). In this section, an initial tool is also presented, using current information from the literature study and the UCs. Second, the document elaborates on trust and data-sharing, and how these concepts can be integrated into the framework (Sections 3 and 4). Third, the framework is explored in depth, with its definition detailed and its areas of application covered (Section 5). Fourth, the developed framework (i.e., T2.1 tool) is presented in Section 6, following the logic of the previous sections. Fifth, the developed framework is applied to the UCs as a validation and testing ground (Sections 7 and 8), based on which conclusions are drawn (Section 11). Furthermore, the overall contribution of the resulting tool to the project blueprint is discussed in Section 9. Last but not least, a glossary (Section 10) is provided to clarify the terminology used and developed throughout the document.

2. Methodology

This section introduces the methodology that was used to develop the "framework" and a "tool". To start with, a literate study is done about data sharing in agri-food, and some reference models that have been used in this context are summarised. But the focus in this chapter is on the definition of the framework and the development of some important aspects. After the set-up of the framework and tool, the data was collected from UCs by using the T2.1 tool. Task 2.1 reviews and assesses the data collected in the agri-food system value chain.

Since the word 'framework' is used in different contexts and with different meanings, we first define what a framework means in this project. Next, the method of developing the framework is introduced.

Last but not least, a glossary is provided at the end of this document to clarify used terminologies and avoid potential confusion arising from different backgrounds and disciplines. This glossary will help guide constructive discussions among stakeholders, not only with regard to the "framework" but also to the project as a whole.

2.1 What actually is a framework?

Generally speaking, a framework defines a set of ideas, conditions, or assumptions that determine how something will be approached, perceived or understood (Merriam-Webster, n.d.). As such, a framework can range from very basic conceptual structures to specific principles and strategies for technical implementations. In the context of data and digitalization, different types of frameworks can be distinguished:

Theoretical framework: "A theoretical framework consists of concepts and, together with their definitions and reference to relevant scholarly literature, existing theory that is used for your particular study. The theoretical framework must demonstrate an understanding of theories and concepts that are relevant to the topic of your research paper and that relate to the broader areas of knowledge being considered." (Labaree, 2024)

Conceptual framework: "A conceptual framework is a structured approach to organizing and understanding complex ideas, theories, or concepts. It provides a systematic and coherent way of thinking about a problem or topic,and helps to guide research or analysis in a particular field."(Muhammad Hassan, 2024)

Architecture framework: "Conventions, principles and practices for the description of architectures established within a specific domain of application and/or community of stakeholders." (ISO/IEC/IEEE, 2022)

In the FDQ project, the framework has a central role as the 'glue' that relates the different activities and results into a coherence and well-balanced set of results. The project proposal underlines the importance of the framework concerning:

Methodological basis: that defines the key concepts and relations of data sharing in a food value chain context.

A guide: a concrete and practical help that consists of guidelines, general principles, data strategies, as well as good practices that can be implemented in UCs.

The second perspective is usually not included in a framework but in a reference model or blueprint. In order to avoid conceptual confusion, in the remainder of this report we will distinguish these two concepts, which are defined in this project as follows:

"**The *FDQ Framework*** is a systematic classification of concepts for analyzing, designing and implementing data sharing systems in food value chains."

Explanatory notes:

A ***data sharing system*** is defined as a digital system to share data between multiple actors of a value chain. Data sharing is related to a wide variety of data for every stage of the value chain, including: structured and unstructured data, public and private data, master and process data, and different functional domains (for example: financial data, field data, food quality data, compliance data, traceability data, consumption data.

A ***value chain*** is a series of interlinked processes, in which every step adds value to the product or service to the end customer (Michael E. Porter, 1985).

A ***consumer-driven value chain*** is a value chain in which all involved firms are focused on end consumer demand, which implies a chain-wide generation and dissemination of market Intelligence, and a supply chain-wide responsiveness to it.

In a **food value chain** the end customers are food consumers and the added value is related to multiple key parameters including nutritional value, taste, price, health, and sustainability. Food value chains have network structures of diverse actors including farm input suppliers, farmers, producer cooperatives, food processors, retailers, food packaging firms, certification and inspection organizations, food labs, advisors, traders and food service companies.

"**The FDQ Blueprint** is a toolbox of practical building blocks for analyzing, designing and implementing data sharing systems in food value chains.

Explanatory notes:

A **building block** is a concrete method, guideline, tool, template, reference architecture, or a technical component that operationalizes a particular concept of the framework.

The **blueprint** is based on the framework, which means that it follows the structure of the framework and its building blocks are operationalizing concepts as defined in the framework.

The FDQ Blueprint will include the practical results of the entire FDQ project (so beyond this task and beyond WP2). It will be continuously enriched over the course of the project.

2.2 Method of framework development

Figure 3 summarizes the envisioned process for developing the project framework. In this regard, as anticipated, before initializing a framework development process, a literature review was organized, highlighting general approaches to frameworks and how they could look. This step can be supported by existing information on UCs and tool in the project. This step can result in a very early version of the framework (i.e., v0.1), containing the current knowledge. Later on, further developments and refinements can be made by elaborating more on existing building blocks and investigating the UCs in more depth, resulting in a newer version of the framework (i.e. v1.0) that is empirically validated and can inform the overall project "blueprint" approach and guidance principles. The framework v1.0 is described in section 4, 5 and 6.

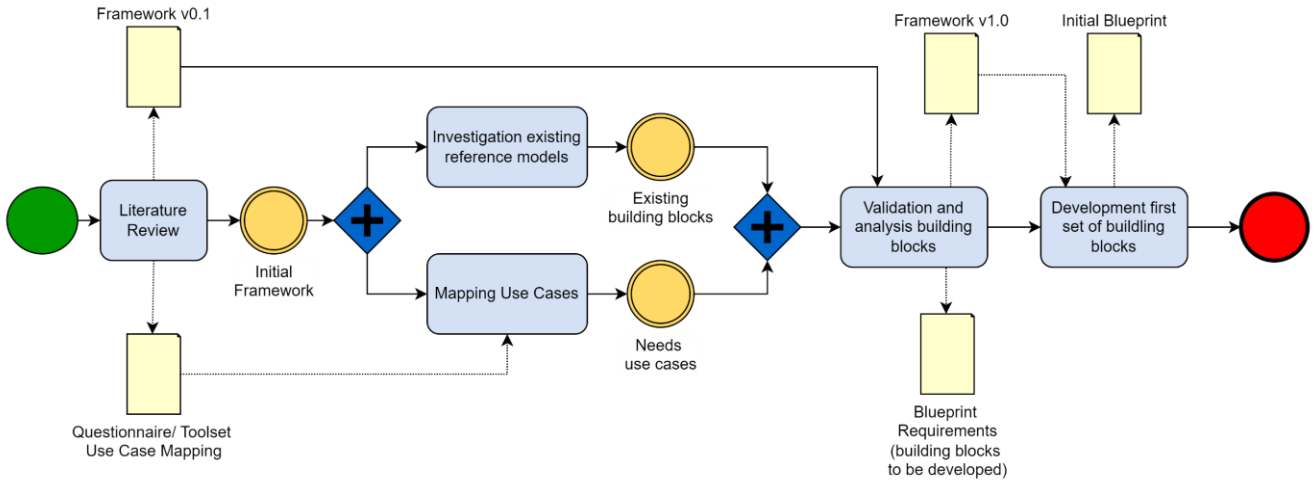

Figure 3 Overview of the framework development methodology

2.3 Method of the development of a tool

2.3.1 Purpose of the T2.1 Tool

To make the framework and the Blueprint as practical as possible, a "tool" will be developed. We designed a questionnaire to get input for the different aspects of the framework. This list has been made, using the input of different reference models, and from the perspective of a UC. This questionnaire provides guidelines to the different UC stakeholders in a logical way to discuss and unravel all elements of the UC, to come to a collaborative way of working, reaching individual goals but also a common UC objective. With the feedback on this question list, we will work further on developing this practical tool to be used by future data-driven UCs that want to implement the FDQ Framework.

The primary purpose of the T2.1 tool is to consolidate various data elements for analysis, enabling a comprehensive understanding of data structure and datasets from different channels. By integrating diverse data inputs, the tool supports investigating data structure & interrelationships, governance, and management in the data chain, as well as the roles of stakeholders, activities and critical events in the Value chain. It ensures alignment with framework objectives and provides actionable insights for further refinement.

2.3.2 Development of T2.1 Tool

The tool was developed for T2.1 to collect input from UCs and provide guidelines for stakeholders. This tool is also for other actors and other UCs across the food value chain to cover all the elements and aspects when sharing data across the food value chain and making data-driven strategies based on the framework. The diagram below(*Figure 4*) shows how the T2.1 Tool has been developed.

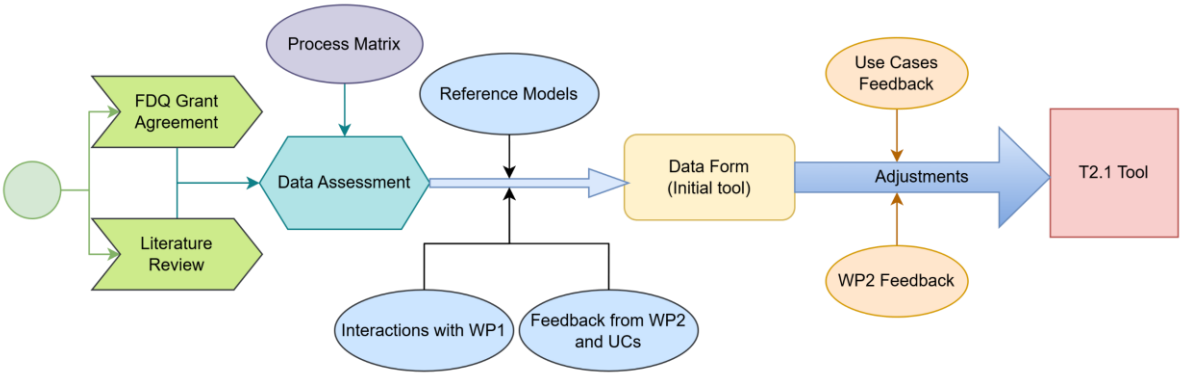

Figure 4 Development of T2.1 Tool

The T2.1 Tool starts with an idea of reviewing and analysing the raw data from different channels across the value chain to understand the content, data structure and interrelationships in data sets. While reviewing FDQ Grand Agreement(GA), we also conducted the literature review on data assessment in agri-food sector. This led to the creation of an initial survey called "Data Assessment" to collect data from UCs and get input. However, we found that it had some overlap with "Process Matrix" which was developed from GS1 as it was developed for UC2. After discussions with WP2 stakeholders and UCs leaders, we decided to combine them together to provide a broader landscape of datasets. By referencing existing models from other EU Horizon projects related to the agri-food sector, and after several discussions with UC leaders and meetings with WP2 and WP1, we developed a final draft of the "Data Form." This served as the initial **tool** for the framework. In order to gather further input and feedback, we asked WP2 consortium and UCs to review it and evaluate its practicality. Finally, after a few months of work and incorporating feedback, we completed "**T2.1 Tool**".

2.3.3 Input used to develop the T2.1 Tool

The first version of the tool was constructed using components and building blocks derived from the Framework Overview which is shown in the Chapter of "Definition of Framework". Additionally, it was consolidated with outputs from related tasks to ensure coherence, completeness and applicability to the UCs. Key contributions in the development of the T2.1 Tool include **Data Assessment** excel file and **Process Matrix**.

Data Assessment

Data assessment was developed to document how different UCs handle their data in alignment with the project scope and framework. This document encompasses multiple dimensions, including:

1. **General information**: This section provides a broad description of the UC, including its technology and data aspects.
2. **Relationship with the project reference architecture**: This includes details on inputs, outputs, and interactions with other partners, particularly when data exchanges are involved. Special considerations, such as privacy, sensitivity, GDPR compliance, etc., are outlined here.
3. **Data sharing willingness**: The extent to which the UC is willing to share its data.
4. **Data processing approach**: This covers the methods and software used for data processing.
5. **Internal UC dynamics**: This section addresses how data sharing occurs among the UC's stakeholders and partners, as well as the expected changes in these dynamics in the future.

Accordingly, this Excel file was filled out by UCs and is intended to be updated in the future to reflect any changes in UC dynamics. In this manner, UCs contribute to the overall framework development process. Hence, the aforementioned dimensions serve as inputs for framework development and inform the high-level reference architecture. The Excel file is available in the FDQ shared folder.

Process Matrix

The "Process Matrix" was – as also mentioned to above – developed at the kick-off of UC2 is based on and inspired by the **GS1 Global Traceability Standard**[2], which focusses on enabling interoperability through multi-actor supply chains through standardization of the identification, capture and sharing of supply chain data. This international standard was developed through the **GS1 Global Standards Management Process**[3]**,** which is a transparent, consensus-driven, collaborative mechanism through which standards are created, improved, and managed.

The "Process Matrix" uses key concepts for interoperable traceability from the above-mentioned standard, being:

1. *Critical Tracking Events* (CTE's), which are the actual events that occur during a supply chain lifecycle, such as harvesting, producing, receiving, packing, shipping, transporting, (waste) disposing, etc…
2. *Key Data Elements* (KDE's) are the elements of data that describe the actual instances of the CTEs, such as harvest date, lot number, farmer ID, etc...
3. *Identification Keys*[4] for enabling unique identification, capture and share of KDE's, such as products (GTIN, Global Trade Item Number) and locations (GLN, Global Location Number)

In addition to the above concepts, the "Process Matrix" has been enriched with additional data components considered relevant and important to FDQ UC2, such as process stakeholders (actors, decision makers), systems and data sources.

[2] https://www.gs1.org/standards/traceability/traceability/2-0
[3] https://www.gs1.org/standards/development
[4] https://www.gs1.org/standards/id-keys

Within the context of FDQ, the "Process Matrix" can be used across the UCs to map all CTE's in the respective supply or value chain, and to highlight the most important data elements (KDE's) that – by means of Identification Keys – can be linked to the relevant conventional and unconventional data sources. In this way, through the "Process Matrix", a connection is made between the data sources and the supply chain, clearly demonstrating the incremental value the specific data sources bring to the supply and value chain.

The table below(*Table 1*) illustrates the above and visualizes the format of the "Process Matrix".

Table 1 Process Matrix

USE CASE				***Example***
Value/Supply Chain Process	Process 1	Process 2	Process n	*Food Waste Collection*
Data Chain Process	Process 1	Process 2	Process n	*Food Waste Data*
Process description				*Registration of food waste during waste collection*
Process stakeholders				*kitchen staff*
Critical Tracking Events (CTE's)				*Transport of waste, disposal of waste*
Key Data Elements (KDE's)				*Product, Batch, Date/Time, Location*
Identification Keys				*GTIN (product), GPC (product classification), GLN (location), SSCC (pallet)*
(IT) Systems				*Transactional (ERP) system), AI image capture system, …*
Conventional Data Sources				…
Unconventional Data Sources				…

2.4 The design of T2.1 Tool

Combining, referencing and adjusting the "Data assessment" and "Process matrix", we also discussed it with UCs and WP2 consortium, after that we finalized the T2.1 tool based on the comments and feedback(*Table 2*). This tool includes the key elements and aspects that UCs need to consider when sharing data securely and implementing data-driven strategies. The tool's content is shown in the table below.

For more details, please refer to the annex (T2.1 Tool), which contains the full questions along with explanations.

Table 2 T2.1 Tool

T2.1 Tool	
Elements/aspects	Questions
Supply Chain process	
	What are the activities of the supply chain involved in your UC?

	Who are the main types of suppliers, service providers involved in supply chain process of the Use Case?
	Who are the key process actors (partners) involved in the supply chain process of the Use Case? Name the partners and other essential organisations.
Supply chain strategy and governance	
	What roles do each of these partners play in this supply chain?
	Who are the decision makers involved in each stage of this supply chain?
	Which data-driven innovation do you want to work on?
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain?
	What is required to verify that each critical event meets regulatory or sustainability standards?
Data chain process	
Data capture	
	Which data is available in each part of chain?
	What type of data is collected in each part of chain?
	How is the data collected?
	What is the estimated volume of data collected?
	What data standards are used (e.g., GS1, ISO, proprietary standards)?
	Does the dataset contain data considered personal in the GDPR?
	Does the dataset contain data considered sensitive in the GDPR?
	How often is the data updated?
Data processing	
	Is there data processing? If yes, how?
	What data needs to be standarized? To which format?
Data analytics	
	Is there data analysis? If yes, how?

Data sharing	
	What data are (can be/willing to be) shared in the process? Under which conditions?
	What data are not (cannot be/not willing to be) shared in the process? Why?
	What systems/technology is used to share the data?
	Are there third-party stakeholders involved in data sharing who are not part of the project?
Data network	
	Who are the data providers?
	Who are the data consumers?
	Who are the data users?
	If there is a technical failure or data problem, how can stakeholders in the data network work together to restore or repair the data flow?
Data infrastructure	
Platforms and tools	What systems/technology/platforms/tools/software is (could) used for data identification?
	What systems/technology/platforms/tools/software are used for data capture?
	What technologies/platforms/tools/software are used to ensure the accuracy of data flow?
Data storage	What platforms and tools are currently in use or in development to store data in the Use Case?
	Where will data (preferably) be stored if shared with other parties?
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?
Data strategy & governance	

Value proposition	
	What is the common goal for all partners?
	What is the win for every partner in the supply chain?
	How will the collaboration take place, what action does every partner take?
	Are there any risks in the data chain process that need to be mitigated?
(Cyber)security	
	What specific security measures are in place to ensure the integrity and confidentiality of the data?
	Are there specific solutions to ensure data security?
Ethics	
	What ethical guidelines/principles are followed throughout the chain?
	What measures are in place to ensure compliance with data privacy regulations?
Data management	
	Who owns the data?
	Who has access to the data?
	What specific data elements need to be traced through the system?
	Which identification standards will be used to ensure traceability across the supply chain?
Legal and policy, regulatory frameworks	
	What European policies/protocols are followed when data is shared between supply chain actors?
	Are there any special industry compliance requirements or regulations involved in the data sharing process?
	If the supply chain involves cross-border data sharing, are there different legal or compliance requirements to comply with?
Business model development	

	How do you come up/adapt your business model to this data driven innovation?
Expected changes, potential needs	
	Are there any gaps or missing data that need addressing?

3. Literature Review

——Current state of data sharing and trust in agri-food sector

Before, we define the methodology of framework and T2.1 tool. The design of framework is an important part of D2.1 for secured and trustworthy data sharing of food value chain. The tool serves as a practical guide for UCs and stakeholders to identify, collect, analyze, manage, store, and share data within the framework. In order to align with FDQ project objectives as well as T2.1, we need to define the content and focus including aspects and elements within the framework. So, literature review has been conducted for general data sharing and private data sharing in agricultural food sector.

In the Europe context, the European data strategy aims to make more data available and flow freely (European Union, 2020). Creating a single market for data will benefit the multi-stakeholders across the value chain within the EU. To align with European data strategy, we recognize that understanding the importance of trust in data sharing is essential for fostering collaboration among stakeholders.

3.1 Data sharing in agri-food sector

The agri-food system faces a wide range of challenges, including societal concerns, public health issues, individual nutrition needs, environmental impacts, food waste, and overall sustainability (Durrant et al., 2021; Griffiths et al., 2023). Data sharing can help address a series of negative problems in the agri-food Value chain, such as lack of traceability, interoperability, transparency, food safety, and provide the EU food chain stakeholders with deep insights and enhance the transition towards sustainable healthy diets (Durrant et al., 2021; Top et al., 2022).

In Europe, several important regulations focus on data sharing, but challenges remain. In 2018, the EU introduced the EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement and the General Data Protection Regulation (GDPR) (Copa et al., 2018; Hoofnagle et al., 2019). These frameworks provide clear guidelines for sharing agricultural data, including non-personal and private data, aiming to ensure transparency and trust. However, practical issues persist, such as the complexity of managing diverse datasets and combining personal with non-personal data (van der Burg et al., 2021a; Wiseman et al., 2018).

3.2 Private Data Sharing for Healthy Diets and Sustainability

Private data sharing is valuable, because it can benefit the public good such as understanding the main society problems for better scientific research in medicine, migration, and environmental areas as well as generating the economics of scope by decreasing the cost using the integration of data (Carol Tenopir, 2011; Susha et al., 2019; Teece, 1980).

The European Commission recognizes the benefits of data sharing and made one of its priorities to create strategies for digital innovations. As digital technology transforms, the EU's digital strategy seeks to utilize this change for the benefit of individuals and businesses while supporting the goal of a climate-neutral Europe by 2050. The European data strategy is one of the pillars of the digital strategy that aims to create a single market for data. This single data market allows data to flow freely within the EU and across sectors. The common European data spaces not only benefit businesses,

researchers and public administrations but also keep companies and individuals who generate the data in control (European Commission, 2020).

At the same time, people acknowledge that a lack of trust between stakeholders is a key factor contributing to the low levels of data sharing. A study (Gabriel et al., 2021) on the reasons and barriers for the acceptance of digital technologies in agriculture shows that initial investment and economic uncertainties in particular limit the acceptance of digital technologies in agriculture. The uncertainty about privacy, data protection and data sovereignty or the ownership of data are also rated as rather inhibiting. And also, the lack of interoperability of different solutions or systems is an inhibiting aspect.

The Data Governance act seeks to increase this trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data. It provides a framework to enhance trust in voluntary data sharing for the benefit of businesses and citizens and will also support the setup and development of Common European Data Spaces in strategic domains, involving both private and public players, in sectors such as health, environment, energy, agriculture, mobility, finance, manufacturing, public administration and skills (European Commission, 2020).

The European commission also has strategies to trustworthy Artificial Intelligence (AI), as it has the potential to improve healthcare outcomes, more environmentally friendly and more efficient manufacturing processes etc(European Commission, 2020).

In the context of private data sharing in agri-food systems, Task 1.1 makes more efforts to provide critical insights of current practices, actors, roles and the possible influencing drivers by mapping the current landscape of private data sharing within food value chain. It also identifies bottlenecks, barriers, and challenges in private data sharing. The outcomes of Task 1.1 serve as input for improving framework for private data sharing under WP2, particularly Task 2.1.

3.3 Trust in data sharing

Trust is the core of data sharing, especially for private data sharing. Private data is opposed to public or open data. In the FDQ project private data refers to information that is confidential and intended for restricted access. According to the GDPR, personal data are any information which are related to an identified or identifiable natural person (individual) (Finck & Pallas, 2019). However, sensitive data concerns special categories of personal data, which are subject to a higher level of protection under privacy laws and regulations due to its highly personal and potentially sensitive nature (Quinn & Malgieri, 2020). These data include genetic, biometric and health data, as well as personal data concerning a person's sex life or sexual orientation or data revealing racial and ethnic origin, political opinions, religious or ideological convictions or trade union membership (European Commission, 2016). Based on the definitions, private data include sensitive or personal data that is not meant to be shared publicly. This covers personal information, consumer food preference, and data held by a hospital or doctor etc.

Without trust, collaboration among multiple stakeholders becomes difficult and impossible. Standardized worldwide private data collection, processing, analysing and sharing might optimize global dietary habits to improve population health (Micha et al., 2018). A national data sharing initiative in Sweden for prevention and treatment of obesity has been investigated, the results showed that private data sharing was important to all participants and it might be used for precision healthcare

(Christenson et al., 2024). Data sharing can also be beneficial to the sustainability by enabling more efficient resource management, reducing waste, and supporting the transition to environmentally friendly practices. For example, sharing data across the value chain improves traceability and transparency, it ensures that products meet sustainability standards.

All the examples emphasize that data sharing is important from many perspectives and building trust is essential. This means that a structured and comprehensive framework is essential to build, maintain and secure trust.

Trust **in data sharing is not monolithic, there are different ways to start building trust.**

To enable data-driven innovation, farmers, companies as well as consumers need to be willing to share their data. To foster trust in data sharing among all stakeholders in the value chain, the EU Code of Conduct on Agricultural data sharing by contractual agreement has been developed in 2018 by CEMA and Copacogeca. This was a big step forward in the sector, because the roles and responsibilities of sharing non-personal data were clarified, and data providers agreed to only share data with consent of the famers. By giving more transparency on the data use – who will use which data for which objective – and giving farmers the option to consent, some control is handed back to them, which increases trust in the system. Van der Burg, et al. (2021) conclude that a contract, based on the principles of this Code of Conduct, can enhance trust relationships, but that this is only the case if all information is well understood by the more vulnerable party and that the more powerful partner takes a responsibility in this (Van der Burg et al., 2021b). Van der Burg also states that to really foster trust, the code of conduct should not only focus on consent, but also on engaging all partners in a data sharing case, to realise fairer and more equitable outcomes for all involved partners.

Moreover, Farahbakhsh et al. (2023), using a simulation model, highlight the importance of economic gains and the imbalance between farmers and market actors in shaping trust related to data sharing. They demonstrate that when economic benefits are unevenly distributed between farmers and other actors, the rate of data sharing can decline. The underlying mechanism is that if farmers perceive their gains as lower compared to those of market actors, they feel less satisfied and may leave the data-sharing platform. This negative experience could discourage them from rejoining the platform in the future. Therefore, to improve trust in data sharing, addressing economic aspects and ensuring equitable benefits are crucial. Moreover, this economic perspective on data sharing arguably contrasts with other approaches, such as viewing data sharing as an altruistic behaviour or emphasizing data ownership, legal frameworks, or institutional initiatives like the EU Code of Conduct or the European Data Strategy. By focusing on economic gains and imbalances, this perspective underscores the importance of viewing data-sharing platforms not merely as data highways supported by trust-based mechanisms but as environments where fair economic benefits among actors can be facilitated.

In this line of reasoning, new data governance models need to be designed to enhance trust as well as ensure fair economic gains while considering the voluntary aspects of data sharing (Ingram et al., 2022; Sjaak Wolfert et al., 2021).

Also, consumers are more and more aware of the sensitivity of sharing personal data, and many don't agree with many data sharing requests (Aiello et al., 2020). However, trust can be fostered through the implementation of trust protection-related policies, frameworks, and regulations. Based on the findings from Task 1.3, which examines consumer attitudes toward data sharing and AI transparency

through a comprehensive survey across five European countries (n=1330), trust emerges as a critical factor in consumers' willingness to share personal data. The research reveals a moderate overall willingness to share data (mean score 3.64 out of 7), with significant variations across different types of data. Consumers show highest willingness to share food preferences and recipe data (5.42/7) but are more hesitant with physiological (4.59/7) and location data (4.87/7). Trust levels vary notably across organizations, with healthcare providers garnering the highest trust (60-65%), while local organizations and industry associations receive substantially lower trust ratings (11-20%). Direct personal benefits, particularly financial incentives and discounts, prove most effective in motivating data sharing. The analysis identified three distinct consumer segments: Skeptics (30%) who show low trust and willingness to share, Optimists (45%) who demonstrate high trust and recognition of benefits, and Ambivalents (25%) who display moderate levels across all measures. These findings highlight the importance of building trust through transparent data practices and clear value propositions to facilitate effective data sharing in food systems.

In this regard, the recent Data Act and Data Governance Act, called into life in the framework of the European data strategy, that envisages a single European data market, also work strongly on this trust component. Transparency in data use and management are strongly promoted, by have a clear framework on how data should be collected, shared and used. This goes a step further then the code of conduct that was already established in agriculture. With the Data Governance Act, the focus is on strengthening mechanisms to increase data availability in an ethical way and overcome technical obstacles to the reuse of data. Giving more control to users of connected devices on their personal data – on access, rectifications, and the right to be deleted – helps also to build trust. The Data Governance Act supports the deployment of Common European Data Spaces in strategic domains as agriculture, involving both private and public players. These Data Spaces will implement a governance framework according to previous principles.

Another aspect of trust depends on which roles in the chain actors fulfil, how the connection between different actors is and what their interests are. There are different types of benefits to be gained out of data sharing. In the Fairshare Horizon Europe research project, the value of data has been described by determining 8 types of benefits of data sharing in agriculture: knowledge gathering, monetary benefits, increasing market share, administrative simplification, develop extra services, get better advice, get discounts and the use of technology (Eva Maes & Ella Deroo, 2023).

At last, technological components and infrastructure can also ensure trust. If there are robust data protection mechanisms, individuals and companies are reassured that their data is safe. Ensuring privacy, security and control is key in this context. Several initiatives are working on providing on frameworks or components to ensure trust.

In the book "**Designing Data Spaces,** The Ecosystem Approach to Competitive Advantage" (Boris Otto et al., 2022), the chapter 'building trust in data spaces' describes a security concept for building trust, enabling data sovereignty and usage control enforcement. Key elements are a certification process for components, such as connectors, and the architecture and system security of the connectors themselves.

Gaia-X Digital Clearing House (GXDCH) is one of data spaces as part of an open cloud network to support secure, transparent and sovereign data exchange within the sectors. The main building blocks of Gaia-X framework are the management of trust, identity, and governance (Sołtysiński et al., 2024).

IDSA, the International Data Space Association, also provides IDS certification for data space components. If e.g. a connector is certified, it show that it is technically capable of meeting the highest standards in terms of security and functionality.

3.4 Address trust and privacy restrictions

This aspect of trust will be addressed later on in the framework by enhancing collaboration among actors. The privacy restrictions have been carefully considered and integrated into the design of the tool. The inclusion of privacy restrictions is based on an analysis of existing practices, stakeholder inputs, and alignment with the FDQ project objectives. This ensures that the tool addresses critical privacy concerns while aligning with the European data strategy.

Other tasks, such as T1.3 and T2.2, further investigate and make sure data flow in practice and cooperation efficiently in a trusted and secured way. T1.3 aims to provide information for the consumers at the point of purchase and help consumers engage their behavior changes by effectively utilizing digital technologies and private data. It will further investigate how privacy concern and trust impact personal data sharing by conducting a survey on the privacy calculus. To be more specific, T1.3 will do research about the level of transparency of private data sharing and what personal data they feel ready to share. T1.3 has finished the data collection part and part of the data analysis results has been mentioned previously. Further details will be available on Deliverable 1.3 report at 18th month of the project. T2.2 will design Reference Architecture and Data Model to support Data Spaces for interoperability. This will enhance the trust among stakeholders for data sharing in the value chain.

Moreover, T2.3 and T2.4 will provide additional insights into privacy restrictions in data sharing. T2.3 focuses on reviewing of current policy and legal aspects regarding the data sharing across the agri-food system in the EU. This includes identifying key principles on data rights, transparency, and data sovereignty, which are critical to addressing privacy restrictions. And T2.4 will carry out a business assessment and create business model based on UCs. To identify the governance mechanisms in different stakeholders, T2.4 will help identify privacy-related challenges in business practices and propose solutions to data interoperability and improve performance.

4. Reference models of data sharing in agri-food

4.1 Existing reference models

To contextualize the set-up of the framework for the agri-food sector, it is crucial to examine existing blueprints, frameworks, and reference models already established in other projects or collaborations (*Table 3*). These existing resources provide valuable insights and best practices that have proven effective in addressing challenges such as data sharing, sustainability, and traceability. By analyzing these examples, we identify key elements to adapt or refine in creating our framework. Here lists relevant framework, blueprint, and reference models in European projects below.

Table 3 List of existing relevant reference model, and relation to FDQ

Number	Name	Website
1	DSSC Blueprint (Data Spaces Support Center)	https://dssc.eu/space/bv15e/766061169/Data+Spaces+Blueprint+v1.5+-+Home
	The DSSC has made a blueprint for data spaces, describing the different elements to consider when building a data space. This blueprint is made on a version basis, each time including more information and updating the framework based on expert and user feedback. The blueprint also includes tools, a glossary, checklists, etc. and served as a conceptual idea for the FoodDataQuest blueprint.	
2	CHAIN (Connecting humans in the agriculture industry)	https://djustconnect.be/nl/aan-de-slag-stap-3
	This project focussed on how digital data sharing can support cooperation between different stakeholders, and how data sharing can improve workable work in agrifood. The steps developed to develop a common Use Case in the chain, have been used as a starting point for the T2.1 tool questions.	
3	Data Value Chain Model	https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/
	The data value chain model has been used to define the different components of this chain, as a basis for the framework of FoodDataQuest.	
4	ScaleAgData	https://scaleagdata.eu/en
	ScaleAgData made data sharing architecture and data governance using innovative technologies. It enables the assessment of agricultural innovations and data governance framework.	
5	AgriDataSpace	https://agridataspace-csa.eu/
	AgriDataSpace incorporates elements of DSSC. It provides the necessary data spaces and specifications for UCs. It is one of the actionable platforms for data sharing in agri-food sector within the Common European Agricultural data space (CEADS) framework. For more information, please check the deliverable: 20240611 AgriDataSpace_D3.2_Arch_Revised_afterFinalReviewclean.docx	
6	Demeter and AIM initiative	https://h2020-demeter.eu/
	DEMESTER aims to help farmers and farmer cooperatives gain new knowledge and evolute their techonologies by using loT techonologies, data science and smart farming. Agriculture Information Model (AIM) initiative is a data model that shows all data needed by DEMETER applications.	

7	GREAT (European Green Deal)	https://www.egi.eu/project/great-project/
	GREAT did research on building blocks, which referenced the DSSC content, and made a roadmap for implementing and deploying the Green Deal Data Space.	
8	The Internet of Food and Farm 2020 (IoF2020)	https://www.smartagrihubs.eu/iof2020
	IoF2020 showcased 34 UCs about the practical applications of IoT in the real world. It demonstrated how smart farming techniques help improve and benefit the value chain.	
9	Smartagrihubs	https://www.smartagrihubs.eu/Deliverables/pdfs/D3.6-Common-challenges-analysis-and-technology-reusability-exploitation_final_PU.pdf
	SmartAgriHubs develops a multi-stakeholder approach and covers a value-chain network across Europe. It drives the agriculture to a digital transformation by connecting all the dots in the value chain.	

Below, we introduce one blueprint(DSSC) and two important models for the development of FDQ framework. And we also explain what we can learn from them and how we integrate them into the framework.

4.2 DSSC Blueprint

The Data Spaces Support Centre (DSSC) is an important initiative within the European Data Strategy. It aims to support the interoperable and trustworthy data spaces. Its objective is to accelerate digital transformation across sectors while aligning to EU values of data protection, transparency, and innovation. DSSC create a blueprint, which is a comprehensive set of guidelines to support the development of data spaces. In the blueprint, participants can find the conceptual model of data space, building blocks, and specifications and co-creation method. This blueprint is a practical resource for participants to establish, develop, and operate data spaces effectively.

Key DSSC Blueprint Assets

Glossary: Glossary contains a list of definitions to explain concepts and detailed terms in the building blocks. In this way, the glossary makes it easier to guarantee coherence and consistent use of terms and avoid misunderstanding for people who read and use data spaces.

Data Space building blocks: It aims to make stakeholders understand the fundamental components and establish and operate data spaces. It is essential that comply the law and meet stakeholders' interests and needs that make their data space initiatives successful and sustainable. DSSC created, adopted, and updated building blocks every time updating the blueprint. Below *Figure 5* is the latest version(Version 1.5) of DSSC building blocks.

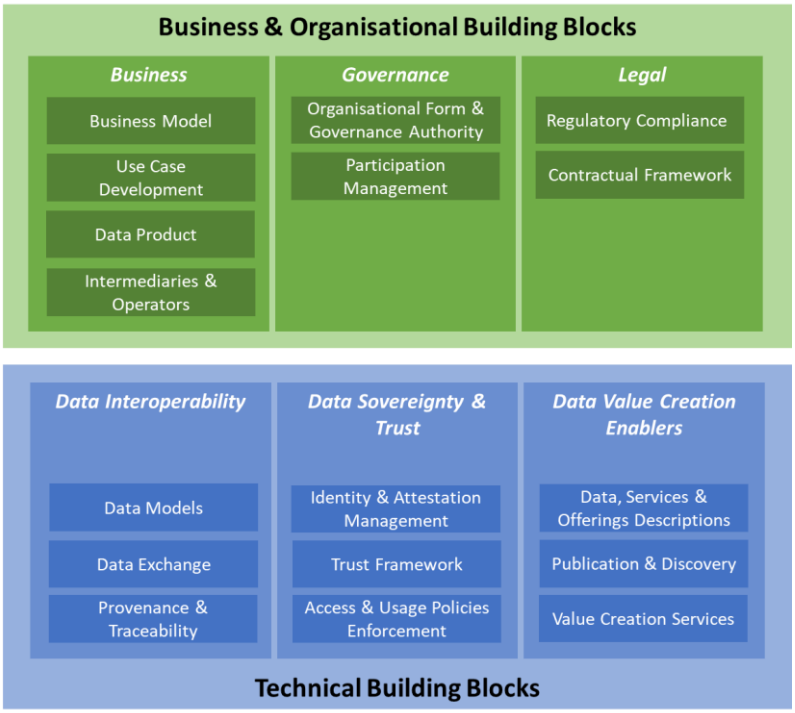

Figure 5 DSSC Building Block Overview(Version 1.5)

4.3 CHAIN

CHAIN is called “Connecting humans in the agriculture industry”, a research project funded by the European Social Fund (ESF). It investigates how digital data sharing can support cooperation between different stakeholders. This project develops a website that functions as a script and bring people together, efficiently organizing the various partners to work together in the agri-food sector: "Connecting Humans in the Agrifood INdustry (CHAIN)".

Four steps of matrix offer an opportunity for multi-actors to decide whether to further collaborate and prioritize which topics.

→Step 1: Map your chain/network

→ Step 2: Determine chain priorities and information flow

→ Step 3: Impact on workability

→ Step 4: To start or not to start?

4.4 Data Value Chain Model (DATA2X)

Data2X is a civil society organization established under the United Nations Foundation. Its focus is on improving the quality, availability, and use of gender data to advance gender equality. The "Data Value Chain Model" (*Figure 6*) is part of Data2X as a research tool aiming at demonstrating the process of data creation, use and re-use. This model applies to development data more broadly not only just collecting and documenting gender data.

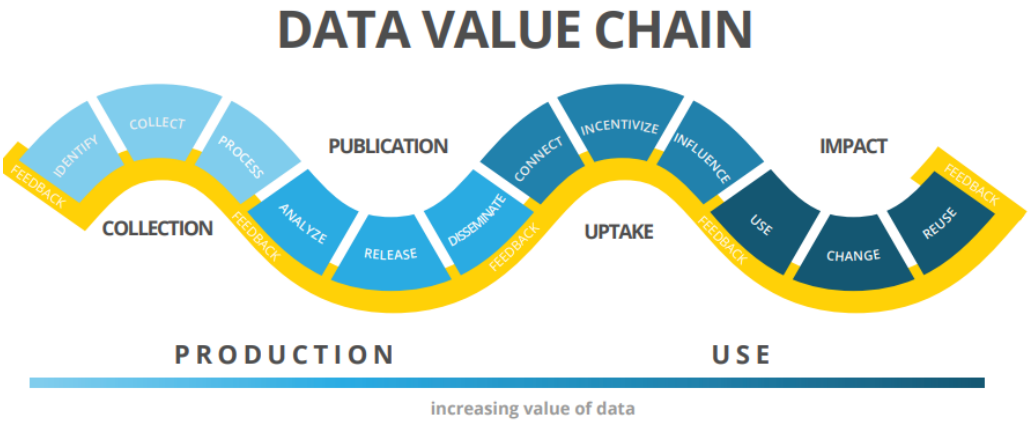

Figure 6 Data Value Chain Overview

4.5 Integration of reference models into FDQ

It is a crucial step to integrate existing reference models into the FDQ framework. These models, found from European projects, provide proven methodologies, best practices, and valuable insights of enabling data sharing, ensuring traceability, and fostering healthy diets and sustainability.

Key Elements to Integrate

To develop a structured and comprehensive framework, we have integrated some key elements from existing methodologies and projects. These elements provide structured approaches for the framework of making data driven strategies and data sharing in the food value chain.

DSSC Blueprint:

Building Blocks: We take building blocks as an important concept into FDQ project. It serves as an essential foundation in the framework for ensuring trust and secure data sharing.

Glossary: We created a similar glossary document as a reference to standardize terminology and facilitate consistent communication across stakeholders.

CHAIN Project:

Four-Step Matrix: We integrate this structured approach to mapping networks and prioritizing collaborations in T2.1 tool to help stakeholders create suitable data strategies.

Data Value Chain Model:

Some parts of the data value chain model have been considered into the framework such as data identity, collect, process, analyze, use and connect. It offers valuable understanding of datasets. We put these important elements into the tool that T2.1 created to help stakeholders better understand and implement the framework.

5. Definition of the framework

The FDQ framework is about data-sharing between multiple actors, directly as well as indirectly involved in food systems, farm-to-fork, to maximize interoperability between these actors. A major complexity of such applications is that they require collaboration between many different stakeholders having different roles in the data value chain. For this reason, the framework draws upon literature review on value chain management and data management. It combines two important chains: **Food Value Chain** and **Data Chain**. The framework is designed to be both conceptual and actionable, enabling interoperability across diverse systems and actors. A visualization of this framework illustrates its components and building blocks (*Figure 2*).

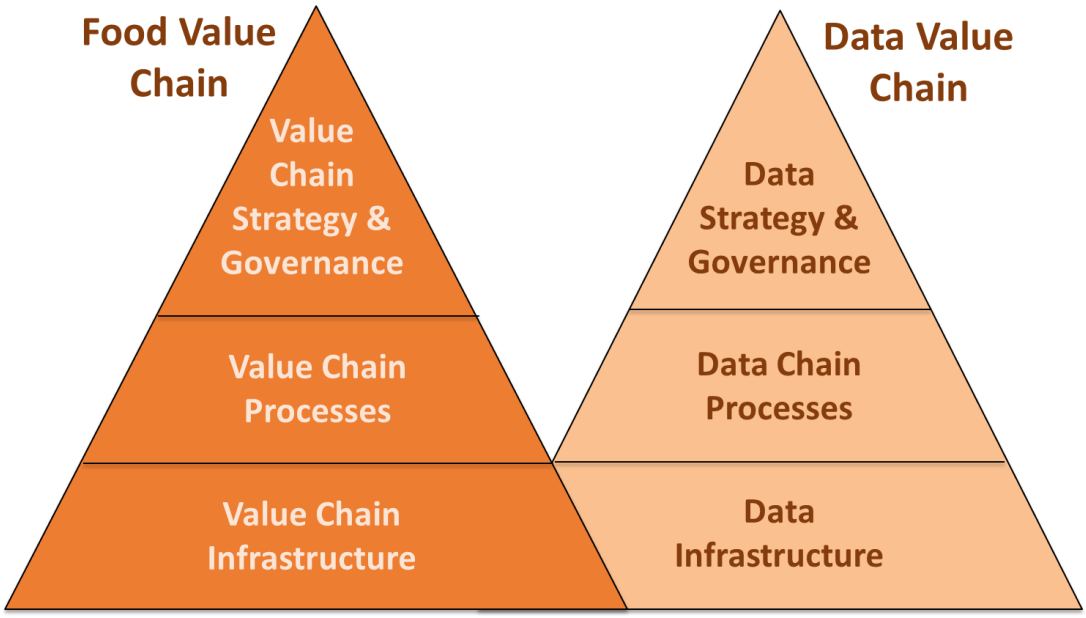

Figure 2 Framework Overview

5.1 Food Value Chain

Value chains are composed of value-adding processes that are performed by different organizations in a network. The participating organization are independent and need aligned strategies, shared governance mechanisms and common infrastructures to coordinate the processes along the value chain. As a consequence, three main categories of framework are defined: **i) Value Chain Strategy and Governance**, **ii) Value Chain Processes** and **iii) Value Chain Infrastructure**.

5.1.1 Value chain strategy and governance

The strategic dimension of a value chain system looks outwards to remain viable in fulfilling its mission, also known as its 'raison d'etre' (Beer, 1984). It involves making strategic choices about how to deliver and enhance distinctive value within specific (ultimate) product-market combinations. Subsequently, it develops strategies and policies to guide its implementation, including ethical guidelines. These strategies may include a vision and mission statement, long-term objectives and plans and business models. Having a value chain strategy implies that the strategies of involved actors are aligned based on common strategic objectives and collaborative business models, including agreements about benefit sharing.

The governance dimension defines how the business processes and the collaboration between actors in the value chain are organized and coordinated, both formally and informally (i.e. trust, power relations). Generally speaking, it determines who has power, who makes decisions, how other stakeholders make their voice heard and how account is rendered (Termeer et al., 2010). The formal governance of value chains is about the allocation of property and decision rights amongst the different involved actors and the interorganizational incentive mechanisms aiming at rewarding desirable behaviour and preventing undesirable behaviour. When studying hybrid forms of organization such as a value chains, three basic forms of network governance can be distinguished (Lazzarini et al., 2001):

- **Managerial Discretion:** discretionary actions by a coordinating agent, who centrally plans the flow of products and information;
- **Standardization:** standardized rules and shared mechanisms to orchestrate transactions;
- **Mutual Adjustment:** alignment of plans through mutual feedback processes and joint problem solving and decision making.

Last but not least, the informal organization of a value chain system is about the unwritten rules that influence the social interaction and trust between the stakeholders, including cultural and ethical considerations. In particular, trust is a critical enabler of effective governance and cooperation in food value chains, which are characterized by high interdependence among diverse stakeholders, including farmers, processors, distributors, retailers, and consumers (de Vreis et al., 2022; Fischer, 2013). It influences information sharing, resource allocation, and joint problem-solving, which are critical for addressing challenges such as food safety, traceability, and sustainability.

5.1.2 Value chain processes

After having defined objectives based on strategic choices, the intended value must be realized by executing business processes. A business process is a set of logically related tasks performed to achieve a defined business outcome (Davenport & Short, 2003). Business processes can be subdivided into primary and supporting business processes (Porter, 2008).

- **Primary Business Processes**: contribute directly to value creation by development and production of the product, its marketing and delivery to the buyer. Examples are R&D, production, logistics, distribution, sales and procurement;
- **Supporting Business Processes**: take care for the development, deployment and maintenance of resources that are required in primary processes. In this way they contribute indirectly to the value adding process. Examples are Human Resource Management, maintenance of machines and equipment, financial administration and IT service management.

The FDQ framework is focusing on integration of especially primary **processes in food value chains**, including:

- **Demand Management:** a planning methodology used to manage the demand for products and services
- **Sales and Marketing:** refers to selling goods and services of satisfying and retaining customers
- **Procurement:** refers to the process of acquiring goods, services, or works from an external source

- **Inbound Logistics:** refers to receiving and storing raw materials, finished or semi-finished food products.
- **Farming**: cultivating crops and raising animals for food and non-food products, including arable farming, livestock farming, horticulture, fishery/aquaculture, etc.
- **Food Production**: transformation of raw farm outputs (such as crops, livestock, fish, etc.) into finished or semi-finished food products.
- **Distribution/ Outbound Logistics**: refers to shipping and moving of raw materials, finished or semi-finished food products to customers or end users.
- **Transportation**: moving food products within the supply chain by road, rail, air, or water freight.
- **Retailing**: selling and distribution of food products to consumers through various channels such as supermarkets, grocery stores, farmers' markets, or web shops.
- **Food service**: preparing and serving food to consumers in restaurants, catering, hospitals, etc.
- **Consumption**: the final stage where food products are consumed.
- **Waste Management**: the collection, recycling, or disposal of food waste generated during production, processing, distribution, and consumption.

The value chain processes are executed by a dynamic network of different organizations. Food supply chains have diverse network structures where many small and medium enterprises trade with huge multinationals. **Important actors include**:

- Suppliers of agricultural inputs, including genetics, seeds, feed, pesticides, medicines, farm equipment, packaging, etc.
- Primary producers, including farmers, growers, and fishermen, and producer cooperatives;
- Food processors
- Certification and inspection organizations
- Authorities
- Traders
- Transporters and logistic service providers
- Retailers
- Food service organizations: restaurants, catering companies, (prepared) food delivery services, and institutional food providers such as schools, hospitals and work places.
- Consumers
- Others who are (in)directly involved in value chain activities: R&D etc.

5.1.3 Value chain infrastructure

Value chain infrastructure refers to the shared resources that support the effective flow of goods, information, and finance across all stages of production, distribution, and consumption. Infrastructure in value chains includes physical assets such as transport systems (such as roads, airports, port facilities, and rail), energy utilities and storage facilities, as well as digital infrastructures, such as (wireless) communication networks and digital platforms. Such infrastructures are prerequisite for providing organizations access to value chains.

5.2 Data Value Chain

Food value chains in the context of the FDQ are embedded in a data economy and therefore characterized by the use of data as a primary asset for the production, distribution, and utilization of value. Data can only be used to create value if they can be shared smoothly and trustworthy among various stakeholders and efficiently processed to actionable insights. This implies that a data value chain is integral part of the food value chain. Following the definition of Porter (1985), a data value chain is a series of interlinked processes from from its origin to its final use, in which every step adds value to the data product or service to the end user. It comprises the sequence of activities from data capture, data storage, data transfer, data transformation, data analytics and data utilization (Curry, 2016; Faroukhi et al., 2020; Wolfert et al., 2017). It is composed of similar elements as the food value chain itself: i**) Data Strategy and Governance**, ii) **Data Chain Processes** and iii) **Data Infrastructure**.

5.2.1 Data strategy and governance

A key starting point of any data-driven value chains is a robust data strategy (Marr, 2017). It outlines how an organization will collect, manage, use, and analyze data to create of distinctive business value (DalleMule & Davenport, 2017). In essence, a data strategy describes the reason-of-existence and how data align with value chain objectives, but it also defines long term policies and plans of how these objectives are to be achieved. To do so, important components of a data strategy are (Analytics8, 2024; Goethals, 2008; Wolfert et al., 2017):

- **Data vision and mission statement**: including the alignment with business objectives.
- **Assessment of the data readiness/maturity**: evaluating the starting point concerning required data management capabilities is needed to be able to define realistic plans to become data-driven.
- **Data architecture**: creating a common understanding on the data value chain by mapping the viewpoints of different stakeholders involved.
- **Data-driven business models**
- **Plans and policies for data-driven management**, including a roadmap, human resource plan, etc.

The AgriDataSpace projects identifies the following types of data-driven business models (Agri Data Space 2024):

- **Data monetisation**: Unilateral approach under which companies make additional revenues from the data they share with other companies. Data can also be monetised through the provision of services.
- **Data marketplaces**: Trusted intermediaries that bring data suppliers and data users together to exchange data in a secure online platform. These businesses generate revenue from the data transactions that take place on the platform.
- **Software as a service (SaaS)**: In this model, the data providers and analytics providers get financially compensated for their data and services and the consumers pay a subscription for accessing data and platform.
- **Industrial data platforms**: Collaborative and strategic approach to exchange data among a restricted group of companies. They voluntarily join these closed, secure and exclusive

environments to foster the development of new products/services and/or to improve their internal efficiency. Data may be shared for free, but fees may also be considered.

- **Technical enablers**: Businesses specialised in and specifically dedicated to enabling data sharing through a technical solution. Revenues are obtained from setting up, using, and/or maintaining the solution (not from the data exchanged).
- **Open data policy**: Companies that opt to share data for free to foster the development of new products and/or services.

The following are the most commonly used revenue or cooperation models (Agri Data Space business model deliverable):

- **Free-to-all**: Also known as open data (connected to the Open Data Policy business model). Data Sharing Initiatives can leverage open-source models to create communities where data contributors freely share their datasets, tools, and algorithms. This approach promotes transparency, collaboration, and rapid development, driving collective progress and enabling the democratisation of data-driven solutions.
- **Freemium**: The freemium model, popularised by companies like Spotify and LinkedIn, combines free and premium offerings to attract and retain users. In the context of Data Sharing Initiatives, this model allows organisations to offer basic data access and services for free, while offering advanced features, premium data sets, or enhanced support through paid subscriptions. Freemium models enable broader data accessibility, driving user adoption and fostering long-term customer relationships.
- **Licensing**: Data is shared based on licensing agreements.
- **Sponsorship/branded advertisement**: Free access to data. The cost of the data is subsidised by sponsors or advertisers.
- **Demand-oriented**: Cost of data depends on availability and complexity.
- **Barter system**: Data consumer gain access to data based on the exchange of data with data provider (i.e. pay with data).

Data governance concerns the organizational rules of the 'data game'. Similar to value chain governance, it defines how the processes from data capture to it utilization are organized and how collaboration between the actors involved in this chain are coordinated, both formally and informally (i.e. trust, power relations). Important formal arrangements for data governance include agreements on data availability, data quality, access to data, security, responsibility, liability, data ownership, privacy and distribution of costs (Wolfert et al. 2017). The informal organization of a data value chain is about the unwritten rules that influence the social interaction and trust between the stakeholders, including cultural and ethical considerations (Wolfert et al. 2017).

More specifically, **three key data governance challenges** can be addressed: **the investment challenge, the participation challenge**, and **the data challenge**.

The primary challenge is the investment challenge, encompassing the following key issues:

- Legal Formalization: a data-sharing initiative requires a robust legal framework to address critical matters, such as:
    - Financial obligations for IT products and services, e.g., hardware leasing, hosting fees and software licenses
    - Intellectual property concerns, e.g., ownership and modification rights for software

    - Liability for risks like security breaches and privacy violations.
- Governance mode: the ownership model of the shared data infrastructures (e.g. platforms or data spaces) must be determined. Options include:
    - Collective Ownership: Shared by all members of the "data cooperative," including farmers.
    - Investor Ownership: Owned by investors with a clear profit-driven focus.
    - Hybrid models are also viable, such as investor ownership with farmer participation in decision-making, or investor ownership paired with nonprofit goals. These structures help balance the dual challenges of securing investment and ensuring stakeholder participation.

The second major challenge is participation: determining who controls the shared data infrastructure and forms its governing body. Typically, only large organizations have the means to build such a platform. This dominance can raise concerns among smaller participants, like farmers, about the power imbalance, such as the potential for increased service costs or misuse of datasets. These fears may discourage their participation. For a collective, member-owned data-sharing platform, effective governance is critical. Allowing members to influence key decisions fosters trust and engagement, particularly from farmers as data providers. Broad farmer participation increases the platform's value, as the platform thrives on extensive use and farmer contributions.

And in the third main challenge is the data challenge, it is important that the governance of the platform must take special provisions or measures in the management of all data. This involves : who owns the data, how exactly the data is processed, who has access to the data, security of the data (security). Do all "members" of the initiative gain access to the data if the farmers give their consent? But when it comes to "confidential" data, do the other parties have to pay for the data? All agreements about the aspects must be laid down in agreements and rules (Markus and Bui, 2011).

5.2.2 Data chain processes

The data value chain process comprises the sequence of activities from data capture to its utilization:

- **Data Capture**: the process of collecting raw data from various sources, including sensors, databases, applications, or human input. Data capture focuses on ensuring accuracy and completeness while minimizing loss or distortion of information. This step is crucial for providing the foundational inputs for subsequent data processing.
- **Data Storage**: refers to the methods and technologies used to securely retain captured data for future access and use. This includes databases, data lakes, and cloud-based storage solutions. Effective data storage ensures scalability, accessibility, and compliance with data governance standards.
- **Data Transfer**: the process of sharing data between systems, devices, or locations, ensuring seamless integration across the data ecosystem. Data transfer requires robust protocols and networks to maintain data integrity, speed, and security during transmission, particularly in distributed systems.
- **Data Transformation**: the process of cleaning, structuring, and enriching raw data into a usable format. This includes operations like normalization, aggregation, and feature

engineering. Transformation ensures that data aligns with analytical requirements and is suitable for deriving insights or powering applications.

- **Data Analytics**: involves analyzing processed data to identify patterns, trends, and actionable insights. This step uses statistical, machine learning, and AI techniques to extract value from data, driving evidence-based decision-making and predictions.
- **Data Utilization**: refers to applying insights gained from analytics to inform strategies, optimize processes, and generate value. Data utilization occurs across industries, powering innovation, enhancing customer experiences, and enabling automation while ensuring alignment with organizational goals.

5.2.3 Data infrastructure

The data infrastructure includes all computers, networks, peripherals, systems software, application packages (application software), procedures, technical, information and communication standards that are used and necessary for adequate data management in the inter-organizational control of farming processes (Van der Vorst et al., 2005). Components to be mentioned here encompass (Wolfert et al., 2017):

- **Data resources** stored in shared databases and a shared understanding of its content (shared data model of the database).
- **Information systems and services** that allow us to use and maintain these databases. An information system is used to process information necessary to perform useful activities using activities, facilities, methods and procedures.
- **Standards and Communication**: the whole set of formalised data standards (both technical, syntax and semantic) with associated procedures for use, connected to shared databases, which are necessary to allow seamless and error-free automated communication between business partners in a food value chain.
- **The necessary hardware infrastructure**: this includes workstations for individual users, communication and application servers, networking equipment, and peripheral devices. This interconnected system ensures secure, reliable access and supports real-time operations across the network.

Concerning the infrastructure for four key approaches can be distinguished:

As visualized in *Figure 7*, there are 4 key approaches to share data in a supply chain:

- **Point-to-point integration** connects two separate software applications or systems directly to exchange data and communicate without intermediaries.

- **Centralized chain systems** that store and manage data from various sources across a supply chain in single, unified database.

- **Distributed ledgers** in which supply chain data are managed collaboratively by multiple participants. In this approach blockchain technology enables multiple participants to have an identical copy of the data and as a consequence, there is no need for a centralized database. Validation of the transactions is performed by the network based on a tamper-proof consensus mechanism, and not through a central intermediary. In a blockchain, each

transaction is encrypted and cannot be modified after approval, making this solution secure and tamper-proof.

- **Federated data sharing (in dataspaces)**: refers to a decentralized approach to managing and exchanging data across multiple organizations while maintaining data sovereignty. Participants retain control over their data, which remains stored in their systems, but agree to share it through standardized interfaces and governance frameworks. This model enables secure, interoperable data access and collaboration without requiring central storage, supporting trust, scalability, and compliance with data privacy regulations. Federated data sharing is commonly used in domains where sensitive data must be shared responsibly among multiple stakeholders. For this reason, it is a key approach for the Food Data Quest project.

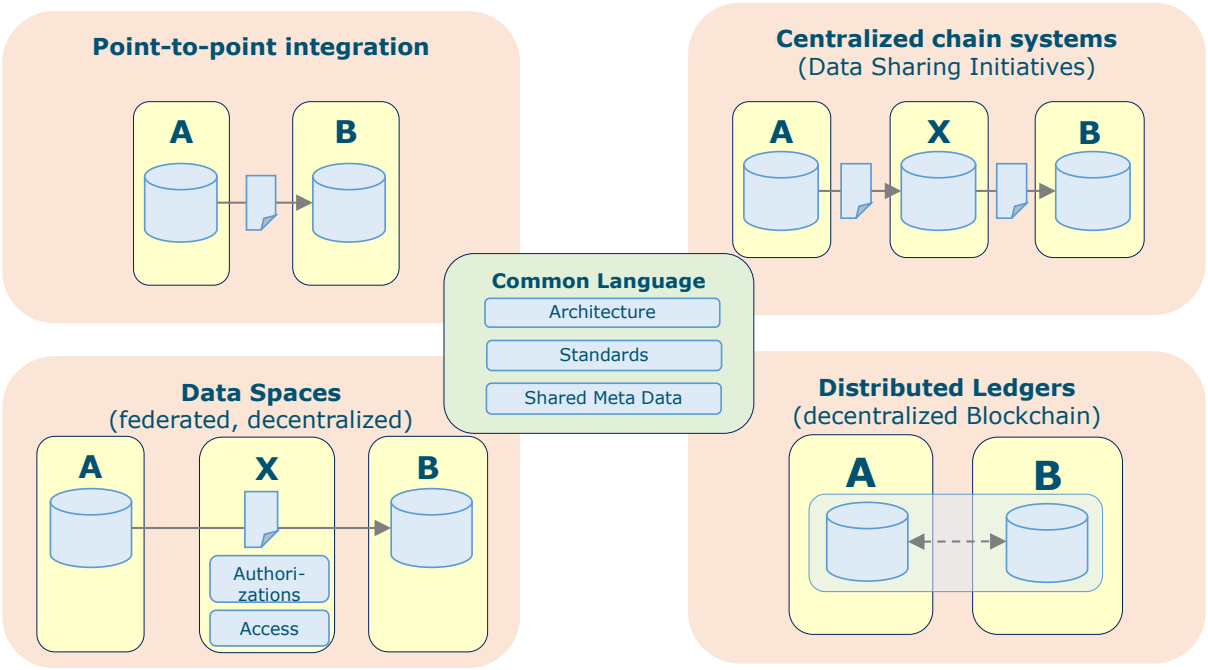

Figure 7 Key data sharing approaches

These approaches are between software information systems (A,B) in some cases centralized or decentralized (7). In all approaches agreements on common language is crucial(Cor Verdouw, 2024).

6. Tool for framework

We hope to help UCs apply the framework effectively by providing a tool. This tool is designed to be a questionnaire, which guides UCs step by step in investigating all the important elements necessary to do trustworthy data-sharing and develop their data-driven strategies. Other ongoing tasks will provide other tools such as technological specifications or business models, etc.

This tool focuses on specific elements rather than covering everything at once. It concentrates on value chain process, value chain strategy and governance, data chain process, data infrastructure, and part of the data strategy and governance.

Below you will find the workflow that we use in the tool to gather input from UCs (*Figure 8*).

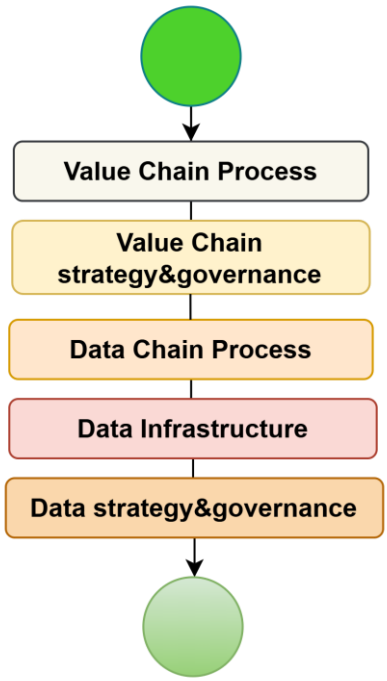

Figure 8 Workflow of the tool

The elements were chosen to ask input from UCs because it is important to build foundation for trustworthy data sharing principles. Other aspects such as data network, the rest of the data strategies and governance, and other expected changes, potential needs will be added later as the UCs evolves to ensure secure and trustworthy data sharing.

The sequence of the questions in each section of the tool was carefully designed to ensure a logical progress for developing data-driven strategies. It starts with mapping the value chain process to identify all the important and necessary stakeholders to establish a clear understanding of objectives and value chain process. Secondly, the questions in data strategies and governance aim to address the roles of the partners and identify the critical events and verifications. Then the data chain process aims to address the key data points and flow — data capture, process, analytics, and data sharing. It is the foundation for effective data chain management. Next, the technical requirements of data chain

infrastructure are explored by determining the tools, and platforms necessary to support the data sharing and storage. Finally, data governance and strategies section focus on data management like ownership, accessibility, and traceability. It is necessary to ensure compliance with ethical and legal standards for trust data sharing.

This step wise approach allows UCs to map their most appropriate data strategies for data sharing in agri-food sector and not overlook critical elements. The input of the tool from UCs can be found in the Annex 2 UCs tool input.

7. Mapping of the Use Cases

7.1 Why are we mapping the UCs

We are mapping the UCs to ensure that the framework we develop is applicable for all the UCs, not only collect input from each UC, but it also helps improve the framework to be comprehensive and inclusive of the diverse. By analyzing the UCs, we can identify commonalities and differences, which help us design the building blocks that are both general and targeted. Additionally, mapping the UCs allows us to collect input for the framework, which is critical for validating and refining the framework and tool, making it practical and relevant for its intended applications.

7.2 Brief description of Use Cases

Use Case 1 DNA-Based Tools and AI for Food Sustainability and Traceability

The Use Case 1 is to improve quality, sustainability and transparency in the olive oil supply chain. By integrating DNA-based technologies (e.g., sequencing and RT-PCR) and AI, it aims to verify the origin, authenticity, and quality of olive oil in real-time. It helps establish a sustainability protocol to mitigate risks like mislabelling and climate-induced mutations while optimizing supply chain efficiency and reducing food waste. The UC1 also aligns and contributes to multiple Sustainable Development Goals (SDGs) such as SDG2(Zero Hunger), SDG3(Good Health and Well-being), SDG9(Industry, Innovation and Infrastructure), and SDG12(Responsible Consumption and Production).

UC1 was officially kicked off on 10th May 2024 in Athens. Now Agroknow is defining and collecting publicly available datasheets to support data-driven, especially AI, DNA-based technologies. So far, Agroknow has assessed olive oil varieties to identify the most suitable ones for sustainability and traceability. The stakeholders were also actively involved in the olive oil sector to ensure their needs are addressed in the project's design. A key milestone has been identifying, which is novel DNA primers. BioCoS will use them to build a DNA traceability system for accurate identification and monitoring of olive oil throughout the value chain. Agroknow is contributing by managing public data collection to enhance transparency in the supply chain.

Looking ahead, the project aims to strengthen traceability and authenticity, ensuring olive oil's origin is verified and preventing fraud by integrating DNA-based technologies. A sustainability framework will also be developed to promote eco-friendly practices, optimize resource use, and protect high-quality local varieties.

Use Case 2 Reimagining food services in hospitals and elderly care homes promoting health and sustainability

Use Case 2 focuses on hospitals and elderly care facilities, introducing a data-driven approach to improve and monitor food services. By combining nutritional information with sustainability insights, it delivers recommendations to encourage healthier eating and reduce food waste. This Case prioritizes locally sourced ingredients and environmentally conscious practices, aiming to transform food services into health-oriented, sustainable systems that lower carbon emissions. UC2 contributes to many sustainable development goals such as SDG9(Industry Innovation & infrastructure), SDG12(Responsible consumption & production), SDG13(Climate action), SDG 15(Life on land).

UC2 officially launched with a successful kick-off meeting held on May 30, 2024, in Maribor. This event marked the beginning of several key activities, including informative sessions, collaborative discussions, and an insightful tour of the UKCM hospital. These initiatives were designed to gain a comprehensive understanding of the hospital's food distribution processes and data usage, aiming to enhance efficiency, promote sustainability, and support healthier diets.

As part of its efforts, the UC2 team mapped out the primary data processes and identified key datasets by completing the assessment Excel files provided by other tasks. To maintain consistent progress, monthly online meetings have been scheduled, ensuring that UC2 activities align with overall project goals and Work Package (WP) requirements.

Looking ahead, UC2 plans to focus on developing the business processes and reference architecture in close collaboration with relevant stakeholders. This collaborative approach will enable UC2 to contribute valuable input to other tasks. A detailed description of food chain processes for both patients and hospital staff will also be prepared to ensure a thorough understanding of the system.

In addition, the team will work on drafting terms and conditions, and ethics guidelines applicable to UC2 activities. This effort will be carried out with support from UCPH, ensuring adherence to ethical and legal standards. These next steps will further solidify UC2's contributions toward improving food systems in healthcare environments.

Use Case 3 F2F Food Sustainability insights for a healthy and climate neutral food chain

This Use Case supports climate-neutral goals by integrating farm-level data into supply chain decision-making. This means that UC3 is also in line with SDGs (3,9,12,13,15). A proof of concept (POC) of a B2B dashboard for retailers and a B2C app for consumers will be developed and they will provide insights into product nutrition, environmental impact, and sustainability. It initially focuses on pork and dairy industry. The dashboard and application aim to promote healthier and more sustainable purchasing decisions by incorporating social and environmental factors into consumer and business choices.

Currently, the B2B dashboard and B2C consumer app are in the planning stage. The POC dashboard will be designed to help retailers and food producers make informed choices by integrating diverse datasets, including non-conventional social and environmental indicators. The consumer app, based on the existing PALAU platform, will introduce new features to guide users toward sustainable habits by adding new elements. Over time, UC3 will achieve these goals throughout the project, ensuring it meets the needs of users and aligns with the project's goals.

Use Case 4 Enhancing Consumer Awareness of Eco-Friendly Choices through High-Density Data-Driven Communication

UC4 focuses on leveraging AI to enhance consumer awareness and promote eco-friendly and healthy food choices by developing and implementing AI-powered behaviour intervention pipelines. It integrates diverse data sources, such as Life Cycle Assessment (LCA) data, consumer consumption patterns, and food choice motivations, to create effective behavioral intervention pipelines.

The objectives of UC4 are to bridge the gap between the availability of environmental information about food choices and consumers' understanding and awareness of it. Current communication strategies often lack personalization and fail to engage consumers effectively, limiting their impact.

There is a clear opportunity to utilize diverse data sources and harness AI to create more meaningful and personalized messages. By doing so, the project aims to increase consumer awareness and encourage more sustainable food choices through innovative, data-driven communication methods.

In the long term, the UC4 plans to develop AI-driven behavior intervention systems that can deliver tailored and impactful messages to individual consumers. These personalized messages are expected to enhance consumer engagement with environmentally friendly food options, promoting measurable changes in dietary habits toward sustainability. The project will also establish a scalable solution adaptable to multiple communication platforms, along with a scientifically validated framework for using generative AI responsibly in food behavior research. The long-term goal is to empower consumers to make more informed and sustainable choices by data-driven communication strategies, ultimately supporting a healthier planet and a more equitable food system.

More information on the description and evolution of the Use Cases will be provided in the WP3 deliverables.

7.3 Application of the framework by the UCs

7.3.1 Value chain process

The Value Chain Process examines the stages involved in the supply chain of each Use Case, providing a structured overview of inputs, processes, and outputs. These stages may be very different depending on the specific focus and scope of each UC. The analysis is tailored to reflect the unique aspects and real-world operations of each case. Main questions that were asked the UCs for input are shown below (*Table 4*):

Table 4 Questions from Value chain process

Questions from Value chain process	
Elements/Aspects	Questions
Food Value Chain Activities	What are the activities of the supply chain involved in your UC?
Food Value Chain Actors	Who are the main types of suppliers, service providers involved in supply chain process of the Use Case?
	Who are the key process actors (partners) involved in supply chain process of the Use Case? Name the partners and other essential organisations.

UC1

In UC1, the input of the value chain includes farming and production of olives, while the process consists of the collection and transport of olives, milling, storage and bottlling. The output is bottled olive oil, but the post-ouput has not been identified yet at this early stage. The main types of suppliers and service providers involved are olive farmers, mill owners, bottlers, and olive oil producers. Key

actors include olive oil companies, BIOCOS for DNA traceability, and machine learning providers for data analysis.

UC2

In UC2, the input involves sourcing and procurement of food supplies, which are provided by local farms, specialty food producers, and wholesalers. It also involves storage and inventory monitoring input, which are provided by UKCM (logistics department). In the process of value chain, UKCM (Nutritionists, dietitians and cooks) and UKCM (Hospital staff and canteen staff) are responsible for food preparation, and meals transport and distribution separately. The output consists of meals consumed by patients, the key actors are mainly patients and nurses and hospital staff responsible for meal distribution/collection in UKCM. Furthermore, the post-output activities include food waste management and monitoring/data analysis. Key actors are UKCM as well as external waste management companies to take care of food waste management while INTRA will create the dashboard for monitoring food usage and waste.

UC3

The supply chain process in UC3 encompasses six key stages: farming input, agricultural production, storage, transport, and trade, processing and transformation, retail, and consumption.

Input (farming input and agricultural production): This stage includes farming activities such as providing animal feed, crop protection products, veterinary medicines, fertilizers, and the production of livestock and dairy products. Key actors involved are Nijsen company, EV ILVO, dairy farmers, and livestock farmers.

Process (storage, transport, and trade; processing and transformation): Activities at this stage involve the storage, transport, and trade of livestock and milk, as well as their transformation in slaughterhouses and dairy processing facilities. B2B meat and dairy transport are also included. The main partners are Colruyt Group, slaughterhouses, and transport service providers.

Output (retail and consumption): This stage includes distribution and provisioning through supermarkets and online points of sale (POS), as well as the final consumption of meat and dairy products. It also considers food waste, packaging waste, and consumer applications, which could be potential post-output of the value chain. Key actors are Colruyt Group, GS1, Palau BV, Inex, and Sport watch partner.

UC4

The value chain process in UC4 is divided into three main stages: retail (input), consumer (process), and R&D (output). In the input stage, loyalty card data is collected from retailers. During the process stage, consumer behavior data is gathered and analyzed to understand preferences and motivations. The output stage involves research institutions (TAU and WUR) utilizing this data to derive insights into consumer preferences and motivations. The main actors at each stage include retailers for the input, consumers during the process stage, and research institutions during the output phase.

7.3.2 Value chain strategy and governance

This element examines the roles and responsibilities of various partners involved in the value chain. It also helps identify the decision makers at each stage. Furthermore, it explores the critical events in

the value chain to ensure quality, compliance, or traceability at each stage of the value chain. These critical events must be monitored and managed to meet regulatory or sustainability standards, which requires the implementation of mechanisms to verify that they adhere to these criteria. *Table 5* shows the questions were asked for input of value chain strategy&governance.

Table 5 Questions from Value chain strategy and governance

Questions from Value chain strategy and governance	
Elements/Aspects	Questions
Roles and Decision-Making	What roles do each of these partners play in this supply chain?
	Who are the decision-makers involved in each stage of this supply chain?
	Which data-driven innovation do you want to work on?
Critical Events and Standards	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain?
	What is required to verify that each critical event meets regulatory or sustainability standards?

UC1

In UC1, the value chain has a few important processes: Farmers provide the raw olives, and then it is milled into olive oil by Milling companies. The final product is bottled olive oil, that will be stored and sold by olive oil producers and retailers. BioCoS uses DNA traceability and machine learning post-processing analysis to ensure the authenticity and traceability of the olive oil. The main decision-maker in this process has been identified as Olive oil company, which makes sure the whole process aligns with data-driven innovation they want to work on: the DNA traceability and sustainability framework. The critical events are monitored at every stage, including pre- and post-milling, storage conditions. Analytical outcomes and production documentation from the field are stored. The transport certificates, quality analysis reports, lab test results, batch numbers, and LoT numbers are used to verify compliance with regulatory and sustainability standards and ensure full transparency from the field to the store.

UC2

In UC2, the value chain strategy and governance encompass various roles, actors, processes, and critical events across multiple stages. UKCM plays a vital role and is the decision maker throughout the whole value chain while different departments collaborating together with other external actors function differently. Like mentioned before, in UC2 there are multi-stages of each value chain process. Briefly, for input, the UKCM Procurement Department manages sourcing strategies, evaluates sustainable suppliers, and Logistics Department provides transportation solutions and oversees

storage conditions, inventory tracking, and efficient resource utilization. Other actors, like food producers and wholesalers supply fresh, organic, and nutrient-rich ingredients.

During the process of food preparation and distribution, UKCM nutritionists, dietitians, and cooks develop patient-specific menus aligned with dietary and health needs. And hospital staff and canteen staff coordinate the timely transport of meals to patients and manage distribution points within the hospital. The output is consumption and feedback. Patients and healthcare personnel enjoy their meals and share feedback regarding the quality and their level of satisfaction. Meanwhile, nurses and hospital staff gather any leftover food to assess preferences and analyze waste patterns. Lastly, UKCM hospital staff is responsible for collecting the food wastes. They also implement recommendations to refine sourcing, preparation, and distribution practices. INTRA will develop and maintain dashboards to track food usage, waste, and patient feedback, offering real-time data insights for decision-making and process optimization.

Each stage identifies critical events which are necessary for quality, compliance, or traceability. These include supplier documentation and audits, traceability checkpoints, inventory inspections, nutritional quality assessments, patients feedback collection, leftover analysis, waste management, and data validation, etc. A series of critical events are verified at each stage. The integration of data-driven innovation across the value chain is the dashboard created by INTRA. The dashboard will centralize data from all stages of the food chain, providing real-time insights and performance metrics.

UC3

The value chain strategy and governance in UC3 involves multiple partners with different roles across different value chain stages. ILVO provides the KLIMREK tool, offering LCA scans at the farm level to enhance sustainability. Nijsen company provides circular feed to pork farmers, and in turn, pork farmers provide products to Colruyt. Retailers collaborate with dairy farmers and milk processors to produce dairy products and leverage platforms like the one from GS1 to establish data standards. GS1 can make a connection with the dashboard and the found data, and the possibility to synchronize their databases with PALAU will be explored. The consumer app, provided by PALAU, guides consumers toward sustainable and healthier choices.

Decision-makers are different across the value chain: KLIMREK coordinators and farmers must consent to share data. Retailers, feed companies, processing companies, and consumers influence downstream processes.

The data-driven innovations being pursued include API-based climate scans, additional attributes in data model, and advanced machine learning algorithms. Additionally, UC3 is interested in making a connection with the health data space of a data space with consumers info.

Critical events include obtaining consent for data collection, ensuring secure data transfer, and maintaining data quality control. At each stage, compliance with data standards, anonymization of sensitive data, and adherence to sustainability metrics are prioritized. To verify these events, clear legal contracts and regulations such as Data Act and Data Governance Act are essential. Data-sharing platforms and eID are used to ensure secure, transparent, and controlled data.

UC4

UC4 involves WUR, TAU, retailers and consumers who play different roles in the value chain. Consumer preferences and loyalty card data provided by TAU university. Consumption database provided by WUR. However, in the retail part, the retailers are the decision makers while consumers make their own choices in the process. R&D mainly is determined by WUR and TAU. For UC4, the AI-models are used as innovative technologies in the value chain to analyze consumer data to provide personalized dietary recommendations.

For UC4, the data-driven innovation for all the partners is to get insights from data and transform them into more personalized choices-recommendations. And it is important to test the effectiveness of these recommendations in driving shifts toward more sustainable consumption patterns.

The critical events include ensuring data anonymization, data quality checks, test recommendations, evaluate the behavioural change and environmental impacts, and compliance with GDPR and other data protection regulations. They need to be verified and validated by reviewing the data, technical audits of data transfer, analyzing the pilot data, getting ethical approval, performing impact assessment, validate models output etc.

These critical events can be verified by providing agreements, audit results, certificates, assessment results and other documentation to prove.

7.3.3 Value Chain infrastructure

For this aspect, we acknowledge that the necessary infrastructure for each Use Case is an essential part of the existing value chain. The value chain infrastructure in this context includes transportation systems (such as roads, airports, port facilities, and rail), public utilities (such as energy, water supply, sewer systems, and irrigation), communication networks (such as telecommunication and broadband), and social infrastructure (such as schools and hospitals) (Luo & Xu, 2018). These elements are fundamental enablers of the value chain processes and are typically fixed. Given their established nature, this infrastructure is not something that can be directly modified or identified as part of the analysis. Therefore, this section does not include additional data collection for value chain infrastructure. Instead, we focus on identifying and addressing the dynamic elements of the value chain, such as the processes and strategy/governance.

7.3.4 Data chain process

The data chain process outlines the journey of data throughout the value chain. It encompasses four key elements: data capture, processing, analytics, and sharing. Each element is defined by taking different questions into account to ensure data availability, quality, compliance, and usability. Example questions that were asked for input are shown below (*Table 6*).

Table 6 Questions from Data chain process

Questions from Data chain process	
Elements/Aspects	Questions
Data capture:	

	Which data is available in each part of chain?
	What type of data is collected in each part of chain?
	How is the data collected?
	What is the estimated volume of data collected?
	What data standards are used(e.g., GS1, ISO, proprietary standards)?
	Does the dataset contain data considered personal in the GDPR?
	Does the dataset contain data considered sensitive in the GDPR?
	How often is the data updated?
Data processing	
	Is there data processing? If yes, how?
	What data needs to be standardized? To which format?
Data analytics	
	Is there data analysis? If yes, how?
Data sharing	
	What data are (can be/willing to be) shared in the process? Under which conditions?
	What data are not (cannot be/not willing to be) shared in the process? Why?
	What systems/technology is used to share the data ?
	Are there third-party stakeholders involved in data sharing who are not part of the project?

UC1

Data Capture

Private data is available from farmers and growers, including transport documentation. In the milling process, sensitive and private data such as olive quality indexes, olive volume (weight), and the amount of olive oil produced is collected. At the storage stage, it involves sensitive data, such as the volume of olive oil and storage tank, volume of tank, as well as the date of storage. During bottling stage, data such as the number of bottles and the bottling date (including lot number, which is public) is recorded. Data is mainly collected through manual inputs in documents or government software. Additionally, BioCoS DNA traceability outcomes are provided in the form of confidential reports.

The collected data primarily is considered as private and sensitive. The estimated volume of data is expected to be small, generally only a few megabytes. The data standards used in this process are proprietary. The data is updated annually, coinciding with each harvest season.

Data Processing

In UC1, the DNA traceability data is primarily processed combining DNA and ML. No standardization of data is required at this stage.

Data Analytics

UC1 focuses on DNA testing, using molecular techniques and machine learning.

Data Sharing

Cultivation, production, and DNA traceability data will be/can be shared. However, certain data, such as experimental protocols, primer sequences, and ML methodology, cannot be shared due to intellectual property (IPs) concerns. Data is typically shared in CSV format or through cloud systems. Need to mention that: Olive oil producing companies and associations, are involved in data sharing as third parties but are not part of the project team.

UC2

Data Capture

The data capture of UC2 involves collecting diverse data from input, process, output, and post-output.

Input data includes ingredient data (quality grade, nutritional content, batch numbers) and quantity data, real-time inventory levels and movement logs are collected for storage and inventory monitoring. The process of the value chain collects data for food preparation and meal transportation and distribution. The data involves recipes, quantity of usage, and meals preparation time. As output, patients provide feedback on meal quality, portion sizes and consumption data. Dietary compliance and leftovers data are recorded and collected at the output stage. For the post-output, waste management data is recorded. The integrated data including KPI throughout the whole value chain provides predictive insights and recommendations using dashboard.

The data of the whole value chain is collected by diverse ways including documentation and certifications, management systems/database, and survey to investigate. As for the data standards, the AACI 5.0 standard will be followed which is designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations. Furthermore, there is only a few data containing personal and sensitive data. To further clarify, consumption data and feedback from patients, integrated data, health-related data are personal and sensitive. However, the volume and update frequency of data are not yet clear as UC2 is focusing on the business process and reference architecture design.

Data Process

The data will be processed accordingly by the requirements and purposes, and standardized using standardization format with different units, specific timestamp format and CSV.

Data Analytics

Data analytics in UC2 focuses on optimizing the food value chain including cost optimization, analyzing demand patterns for stock optimization, and refining recipes for cost-efficiency. Patient feedback and dietary preference are analyzed to improve service and meal planning. Integrated data supports KPI tracking, predictive insights, and real-time decision-making through dashboards.

Data Sharing

Throughout the value chain in UC2, most of the data that is mentioned can be shared while aligning with the regulations and requirements. However, personal and sensitive data such as feedback from consumers, raw data from integrated systems including unprocessed data that might contain sensitive or identifiable information cannot be shared. Sharing these data will be further discussed during the next steps of the UC2 in order to be aligned with GDPR and ethical requirements. The shared data will be stored and shared through the combination of available systems and databases, the MEDIS and EHR systems are under the consideration.

UC3

Data Capture

There is lots of useful data available from Nijsen Company, Colruyt Group, PALAU, and ILVO. Nutritional data, dashboard sustainable feed, energy use, CVB, LCA, logistics data, etc, are available from Nijsen Company, and storage/logistics data, suppliers' information, environmental impact information are available from Colruyt Group. PALAU offers data on scanned products, consumer diet data, eco-scores, nutri-scores, and product details. ILVO's data on sustainability, animal welfare, and Klimrek results will be collected. Most of the data, such as primary farm data from climate scans, is structured, but some, like feed/impact logistics data, consist of unconventional datasets. While most of the data is collected through applications or devices, some are entered manually, but the volume of the data is not yet clear. The data standards used are CSV or JSON. Certain datasets, such as unconventional datasets, are considered personal or sensitive data but will be updated regularly.

Data Process

The data will be processed on different platforms depending on different data provenance. The platform DjustConnect, Sustell tool, GS1, and the PALAU application will process the data and help standardize the data into CSV or Json if necessary.

Data Analytics

The data collected is analyzed by using calculation model of KLIMREK and Sustell calculation tool. The eco-score and nutri-score will be calculated based on the data available.

Data Sharing

In the value chain of UC3, all collected data from farming input will be shared with the consent of farmers and under legal contracts with clear rules in accordance with the legislation and the Code of Conduct on agricultural data sharing /Data Act/Data Governance Act. However, there are agreements that are needed to be achieved for the stakeholders through the rest of the value chain to share data. The stakeholders have recognized that the datasets will only be shared through excel/document/CSV/ API for the purpose and to the partners who will get consent of both data providers and the data subject. Due to the personal and sensitive data included, the datasets will not be public available nor open.

Right now, no third party stakeholders are involved in data sharing who are not part of the project. If there is any, will update this information soon.

UC4

Data Capture

In UC4, loyalty card data, demographics, consumer preferences are available at the input stage. The recommendation data and consumer choice data are available during the consumer process. After R&D, machine learning model output data will be available. The types of data are private and structured, public and unstructured, semi-structured and personalized respectively. Over 35,000 loyalty card data and over 500 consumers data will be collected using retail applications, experiments, and surveys. The volume of machine learning models' data is not yet clear. There will be personal data but no sensitive data, and the dataset is in CSV format. All the data is exported at a certain moment, and no update is foreseen.

Data Process

The data collected will be processed, and data cleaning is necessary based on the inclusion criteria. Loyalty card data will be standardized into semi-structured for time-series analysis.

Data Analytics

The data will be analyzed through Rstudio, SPSS, and data modelling such as loading, cleaning, merging, time-series analysis, dimension reduction techniques, and regression/ANOVA analysis.

Data Sharing

The purchase data, demographic data, preference/motivation and consumption data will be shared based on eliciting consent. However, sensitive and personal data will be shared using pseudonymization. The systems and technology that are used to share the data are still under the discussion, but retail app is being used at this moment. Additionally, the Finish retailer is involved in data sharing as a third-party by sharing the data with TAU University.

7.3.5 Data infrastructure

A robust data infrastructure is essential for trustworthy data sharing. We recognize 2 elements of this aspect: platforms and tools, data storage. This aspect explores the platforms, tools, and technologies currently in use or under consideration for these purposes. Relevant questions asked for UCs are shown below (*Table 7*):

Table 7 Questions from Data infrastructure

Questions from Data infrastructure	
Elements/Aspects	Questions
Platforms and tools	
	What are the activities of the supply chain involved in your UC?

	Who are the main types of suppliers, service providers involved in supply-chain process of the Use Case?
	Who are the key process actors (partners) involved in supply-chain process of the Use Case? Name the partners and other essential organisations.
Data storage	
	What platforms and tools currently in use or in development to store data in the Use Case?
	Where will data (preferably) be stored if shared with other parties?
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?

UC1

Platforms and tools

UC1 utilizes manual collection for data capture and the accuracy of data will be ensured by Internal relational databases and ML pipelines.

Data storage

For data storage, internal database is currently in use in UC1, and preferably to be shared with other parties by internal platform. The security and GDPR are expected to be aligned with when storing and sharing the data.

UC2

Platforms and tools

In UC2, it is still to be decided what systems/technologies/platforms will be used for data identification and data capture. But APIs will be used to ensure the accuracy of data flow.

Data storage

Relational data base will be used to store data and if sharing with other parties, the data will be transferred and stored in Cloud infrastructure. The specific requirements or needs for data storage and secure storage for personal data will be discussed and decided soon.

UC3

Platforms and tools

For the input side, the Data sharing platform DjustConnect, with its different API will be used for data identification. The data is captured through DjustConnect and API based climate scan application, but

also existing datasets will be looked into. A POC of a retail application will be built to capture data. Consumer data will be available on the PALAU application. These platforms such as DjustConnect are also used to ensure the accuracy of data flow.

Data storage

The own data base is the first choice for most of the data storage. But data vault webserver and external APIs are also possible. When sharing with other parties, the data will be stored in relational data base. The structured data in the data base of the organization is shared through API.

UC4

Platforms and tools

In UC4, they rely on Confocus EU and NEVO online to identify data and survey, existing datasets in applications to capture data. This is not applicable for UC4 to ask the technologies/platforms for the accuracy of data flow because UC4 exports the data at a certain point.

Data storage

During the output (R&D stage), internal project servers or a secure cloud platform are used to store data. Possible, machine learning will be used along with diffusion models and large language models for data storage. When sharing with other parties, a GDPR-compliant cloud storage platform is planning to be used. Need to make an agreement with all the parties when storing and sharing data, personal data is stored in a secure platform with compliance with GDPR, and support for large-scale analytics.

7.3.6 Data strategy and governance

Data strategy and governance is the foundation for effective data sharing. There are Value proposition, (Cyber) security, ethics, data management, Legal and Policy & Regulatory Frameworks, and Business model development to be considered. It makes sure that partners balance the risks and mutual benefits while aligning with established guidelines and compliance with data privacy regulations.

At the current stage of the project, the UCs were asked to answer part of the questions in this element. This is reasonable to do because UCs are in different situations and progresses, and it allows for capturing the relevant information from different UCs' specific situations. They may further develop and update the detailed information over time.

There are some questions that were asked to the UCs below(*Table 8*).

Table 8 Questions from Data strategy and governance

Questions from Data strategy and governance	
Elements/Aspects	Questions
Value proposition	What is the common goal for all the partners?
	What is the win for every partner in the supply chain?

	How will the collaboration take place, what action does every partner do?
	Are there any risks in the data chain process that need to be mitigated?
(Cyber)security	What specific security measures are in place to ensure the integrity and confidentiality of the data?
	Are there specific solutions to ensure data security?
Ethics	What ethical guidelines/principles are followed throughout the chain?
	What measures are in place to ensure compliance with data privacy regulations?
Data management	Who owns the data?
	Who has access to the data?
	What specific data elements need to be traced through the system?
	Which identification standards will be used to ensure traceability across the supply chain?
Legal and Policy, Regulatory Frameworks	What European policies/protocols are followed when data is shared between supply chain actors?
	Are there any special industry compliance requirements or regulations involved in the data sharing process?
	If the supply chain involves cross-border data sharing, are there different legal or compliance requirements to comply with?
Business model development	How do you come up/adapt your business model to this data driven innovation?

UC1

Value proposition

In UC1, BIOCOS will invite relevant stakeholders and gather necessary information to scale the DNA traceability system, while Agroknow will provide public environmental data to integrate with the DNA traceability outcomes for developing a sustainability framework. The main risk to address is about intellectual property rights and background knowledge.

Data management

The ownership of olive oil production data is the companies involved, and BioCoS owns the DNA traceability outcomes and testing protocols. The access to the data has not been fully defined yet, However, access will be possible to relevant partners and companies. Olive oil stakeholders to share data and receive outputs. BioCoS has access to perform DNA traceability. There is some specific data

elements that need to be traced including production information, DNA outcomes, and environmental data, within five main categories identified: Field & Harvest Data, Milling, Storage, Bottling, and DNA data. Traceability across the value chain will be ensured through identification standards such as geolocation, legal documents, olive batches, and olive oil lot numbers.

UC2

Value proposition

The specific responsibilities and actions of the partners have yet to be fully outlined. However, it is important to acknowledge potential risks that could lead to adverse outcomes. These include issues such as data traceability challenges, stock mismanagement, inconsistencies in recipes, delays in routing, biases in feedback data and analysis, and the presence of data silos.

Data management

UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through third-party sources. The access of the data is yet to be decided. The production details, inventory levels, recipe ingredients, meal records, consumption data, waste management data, and KPIs will be traced through the value chain. HACCP standard will be used to ensure the traceability of the value chain data which requires the traceability of raw materials, storage practices, food preparation processes, routes distribution, food waste and disposal etc.

UC3

Value proposition

The stakeholders that are involved will collaborate together across the value chain. Colruyt Group will focus on addressing the Scope 3 climate impact, with particular attention to dairy farmers, supporting LCA at several farms to help reduce their climate impact and increase resilience. Additionally, Colruyt will support the creation of a Retail dashboard (B2B) to aid in sustainable supply chain management. Nijsen Company, on the other hand, will focus on the pig sector by providing primary agricultural data sets and contributing to the development of the Retail dashboard. They will also explore social aspects and investigate collaboration with Colruyt to close the cycle within the pig sector. PALAU will enhance the B2C consumer app by adding new features, incorporating more specific and unconventional datasets, and exploring rewards systems. Furthermore, PALAU will investigate AI decision tools to support sustainable product sourcing and processing methods. GS1 will focus on establishing standards for the dashboards and exploring synchronizing PALAU with their database, potentially allowing new information to be pushed directly to consumers. ILVO is hosting a robust data-sharing platform, aiming to connect a wider range of diverse data sources to enhance the platform's operations. They will also provide the API-based climate scan application, Klimrek, which supports climate impact assessments and decision-making. Additionally, ILVO is exploring the integration of more unconventional datasets in the dashboards and to expand the scope of data analysis and improve insights into climate-related social and health challenges.

Data management

At this stage of UC3, farmers give consent if the data can be shared, so famers have the rights to the data such as farming input and agricultural production information they provide. And retailers and PALAU have the ownership of the data they provide. Through the value chain, the farm ID, production

identification, purchase transaction information needs to be traced using GS1 standards to ensure traceability of the value chain.

UC4

Value proposition

Different partners TAU and WUR will collaborate closely but focus on different tasks. TAU will analyze the provided loyalty card data, applying AI techniques and behavioral science methods. They will lead on the integration of cognitive models with food consumption data to generate personalized interventions. They will develop and test the AI pipelines for behavior change based on the results they get. At same time. WUR will focus on designing AI-powered behavioral intervention frameworks and communication strategies. WUR team leads the ethical and data governance aspects, ensuring compliance with GDPR and ethical standards. After the generation of personalized interventions, WUR also responsible for managing the practical application of findings, including pilot testing and reporting. There are some risks such as data breaches, misalignment on data formats, and non-compliance with GDPR to be aware of and be mitigated during the whole project.

Data management

From data management aspect, the data ownership of UC4 depends on the data provenance. The loyalty card data belongs to retailer and TAU, and they have the access to it. The consumption data is owned by WUR and can only be accessed by WUR. The research institutions TAU and WUR share the ownership of the results of machine learning modelling. Due to the nature of research, it is not applicable/required to ensure the traceability of the data across the value chain.

To conclude, this chapter provides an overview of the UCs and their status in applying the FDQ framework. It demonstrates the practices of the framework in the real world. This mapping ensures a structured approach for private data sharing and making data-driven strategies in agri-food sector.

8. Analysis of the building blocks

In this section, we aim to analyze and validate the building blocks of the framework we developed to ensure their relevance, completeness, and adaptability for all the UCs. This process is essential for adjusting and improving the framework. Ensuring the framework aligns with real-world applications.

Each UC provides valuable inputs into the food value chain and data chain. They offer diverse insights on the challenges and opportunities in applying the framework. Below shows the overview of how UCs applied framework by the tool that T2.1 created (*Table 9*).

Table 9 Overview of the UCs applications

Elements	UC 1	UC 2	UC 3	UC 4
Value Chain Process	Farming and production of olives	From food sourcing to food waste in hospitals and elderly care homes	Animal production chain focusing on pork and dairy products	AI-powered behaviour intervention at retail and consume sections
Value chain strategy and governance	Documentation compliance check	Data validation and Compliance reporting	Legal contracts with clear rules in accordance with regulations/laws	Data Validation and legal agreements
Data chain process	Documentations are collected and shared in CSV through cloud	Collect and share data from MAOP and PREHRANA	Data collected from documents, devices and app is shared in excel, document, CSV or API	Data will be collected and shared by Retail app
Data infrastructure	Internal platform	Relational data base and Cloud infrastructure	Relational data base and APIs	Cloud platform
Data Strategy and governance	Information of production, DNA outcomes and environmental data will be traced.	Ingredient batch numbers, Recipe ingredients, Route information, Meal records, Waste segregation records will be traced.	Farm ID, product identification number, purchase transactions will be traced.	Purchase transactions, demographic profiles, and AI-generated recommendations will be traced.

Key findings from the analysis include:

1. **Trust and privacy are core building blocks**: Trust and privacy are found as the most critical building blocks across the UCs. All the stakeholders need to have consent and agreements on it to further collaborate. Addressing privacy concerns and building trust are critical to share data in food value chain. Effective data strategies and governance mechanisms, such as clear ownership rules and data access policies, are one of the solutions for ensuring secure and trustworthy data sharing and storage.
2. **Interconnection and interoperability**: Smooth in data sharing is vital for achieving the objectives of the FDQ project. The results of UCs emphasize the need for interconnection and interoperability of data infrastructure. Establishing common standards and guidelines with technical specifications of data chain ensures data can be shared effectively and smoothly.
3. **Other Common Challenges**: Most UCs face challenges related to ownership clarity, and data quality management. These challenges reveal the need for a comprehensive framework that offers practical solutions to data governance and infrastructure.
4. **Different requirements and ethical aspects**: Due to specific and different context, UCs have different value chain scale, data type and quality, and technological maturity. Ethical issues, such as consent for data sharing, are critical for building and maintaining stakeholder trust. This means that to address these, the framework need to cover not only technical support and tools but also actionable regulations and guidelines.

The framework and the building blocks are currently designed to be flexible and inclusive. The building blocks remain the same as designed after the analysis of the results of UCs' input. Because the framework incorporates the diverse needs of UCs even UCs vary because of the different processes and development stages. By addressing shared challenges and recognizing unique requirements, the framework serves as a solid foundation for guiding UCs toward sustainable and data-driven strategies.

The analysis results of the applications from UCs provide valuable insights into how the FDQ framework supports agri-food sector private data sharing in different contexts. It helps ensure that the framework remains practical and adaptable. While the findings highlight valuable guidance, there are some limitations for further improvement. First, the current UCs may not cover all possible scenarios. Second, we need to do further validation of the building blocks because of differences in different countries, UCs stages could also impact the application of the framework.

Future steps could include refining the framework based on the feedback, improving building blocks, and updating the glossary. These could be done by FDQ blueprint.

9. Blueprint and the contribution of T2.1 to the blueprint

The FDQ Blueprint serves as a practical guide for organizations to work collectively toward shared goals, ensuring that their systems and processes are aligned across technical, business, legal, governance, and policy dimensions. By addressing challenges such as data sovereignty, interoperability, and trust, this framework empowers stakeholders to leverage data as a catalyst for sustainability and innovation.

The blueprint aims to provide an integrated framework for data sharing within the agri-food sector, focusing on interoperability and secure, trusted data sharing/storage. This reference framework will provide input for T2.2, which is Reference Architecture Design, to better integrate data space and develop data model. This reference framework also aligns with the broader FDQ objectives, which include leveraging data-driven innovation to enhance sustainability, traceability, and efficiency throughout the food Value chain and transform the food systems into more inclusive and sustainable.

The concept of the blueprint framework in this project was inspired by the DSSC blueprint, a well-regarded resource for consistent and comprehensive guidelines to support data spaces. Building on this idea, the FDQ blueprint is envisioned as a dynamic and evolving deliverable. It is not limited to Task 2.1 but represents a key outcome for the whole project. The blueprint will be hosted on the FDQ website, serving as a living document that can be continuously updated with new insights, updated information, and additional data. This iterative approach ensures the framework remains flexible, adaptable, and aligned with evolving needs, fostering broader adoption and integration across various UCs.

This iterative document will evolve through collaboration with WP3 input(UCs) and input from other future/new UCs, addressing general challenges for realistic UCs while contributing to the project's broader goals, such as:

Preserving data sovereignty.

Supporting policy and legal aspects of data sharing.

Identifying scalable business models for data-driven innovation.

Through its structured approach, the blueprint ensures that the framework not only meets the immediate needs of UCs but also aligns with the long-term objectives of fostering innovation and interoperability within the EU food system.

10.Conclusions

In this deliverable 2.1, we propose a framework, which contains two chains: food value chain and data chain. In each chain, they consist of different aspects and elements for different circumstances considering realistic agriculture-food sector. This framework takes into account the complexities of the agri-food sector, makes sure it is adaptable to various contexts for data sharing and data-driven strategies.

We also created an inclusive "tool" for UCs. This tool aims to collect input from UCs to better understand different datasets and provide different aspects for UCs to consider for better data-sharing. Each UC is unique and different, so UCs filled in the tool we provided for collecting input in various ways. These UCs are at early stages of development, and it's a good starting point to think about the framework. We need to stay flexible and ready to adapt to gather more insights and refine our approach as we go.

Furthermore, the glossary included here is just an initial version. People working on different parts inside and outside the project might use different words or terms, so it’s essential that we remain proactive in defining and updating these terms as the project evolves. We hope that terminology from the glossary spreads to people, and we hope to get feedback and input during the practices.

Lastly, other tasks will help building the framework in a more actionable and practical way, and the practical tools will be integrated and T2.1 tool will be further improved.

11. Glossary

To support a clear and consistent understanding of key terms within different initiatives, people involved in FDQ and future users, a glossary is created to define essential terminology used throughout the framework. This glossary serves as a reference, ensuring that all stakeholders, from industry experts to project participants, share a common understanding of specialized concepts. It includes definitions for technical terms related to data infrastructure, data chain processes, value chain, and other foundational concepts within the agri-food sector's data chain. The glossary can be found below(*Table 10*).

Table 10 Glossary

Glossary of terminology – FDQ T2.1
FDQ Framework
The FDQ Framework is a systematic classification of concepts for analyzing, designing and implementing data sharing systems in food value chains.
Framework Toolset
The Framework Toolset is a set of tools from FDQ tasks, which are designed for UCs and other participants. It serves as a practical guide to address key elements and aspects within the framework while aligning with EU regulations.
Blueprint
The blueprint is based on the framework, which means that it follows the structure of the framework and its building blocks are operationalizing concepts as defined in the framework.
FDQ Blueprint
The FDQ Blueprint is a toolbox of practical building blocks for analyzing, designing and implementing data sharing systems in food value chains.
Building block
A building block is a concrete method, guideline, tool, template, reference architecture, or a technical component that operationalizes a particular concept of the framework.
Public or Open data
Public data is information that can be shared, used, reused and redistributed without restrictions on access, usage, or redistribution. This covers any data that is publicly accessible and free to use, with the intention of promoting transparency, innovation, and collaboration. Usually shared under a creative commons or similar license that state the conditions governing re-use of this data (see licensed data below).
Private data

As opposed to public or open data. In the FDQ project, private data refers to information that is confidential and intended for restricted access. Private data include sensitive or personal data that is not meant to be shared publicly. This covers data and datasets that are owned by companies, organizations or individuals. There is clear ownership of these datasets, and access to private data is usually restricted to authorized individuals or entities. Often there are legal and regulatory requirements governing its handling and protection to ensure privacy and to prevent unauthorized access or disclosure. This data can be provided for internal use within the consortium under the provisions of the Grant Agreement.
Licensed data
Data and datasets that are shared shared under a licensing agreement, where the owner of the data grants specific permissions and restrictions on its use (Alamoudi et al., 2020). The terms of this agreement may include restrictions on usage, redistribution, modification, and commercial exploitation of the data. Licensed data includes open data but also entails any datasets shared under more restrictive licenses. A data licence is a legal agreement between the data owner or licensor of the data and the end-user, or the place where the data will be deposited, specifying what users can do with the data. The terms of this agreement may include restrictions on usage, redistribution, modification, and commercial exploitation of the data.
Personal data
In the FDQ project we follow the definition of personal data according to the GDPR, not to be confused with private data. Personal data are any information which are related to an identified or identifiable natural person (individual) (Finck & Pallas, 2019). Identifiable means that data subjects can be directly or indirectly identified, through identifiers such as name, identification numbers, location data, or other characteristics. In practice, this includes all data that can be assigned and traced back to a person in any kind of way. Personal data is often sensitive and subject to privacy regulations and protections, as its misuse or unauthorized access can pose risks to individuals' privacy and security. Given that FDQ will proceed with collecting, processing, or storing personal data, it is required to adhere to legal and ethical standards to ensure the protection and proper handling of this information.
Sensitive personal data
This concerns special categories of personal data, which are subject to a higher level of protection under privacy laws and regulations due to its highly personal and potentially sensitive nature (Quinn & Malgieri, 2020). These data include genetic, biometric and health data, as well as personal data concerning a person's sex life or sexual orientation or data revealing racial and ethnic origin, political opinions, religious or ideological convictions or trade union membership (European Commission, 2016). Given that in FDQ processes related to collecting, processing, or storing sensitive personal data will take place, this means that stricter requirements and obligations to ensure the security, confidentiality, and lawful processing of this information are required.
Data sharing initiative
Any existing collaborative initiative where committed partners initiate, develop, maintain a data space where they exchange data (Antti "Jogi" Poikola (SITRA) et al., 2024). With a number of committed

participants, a shared plan, infrastructure and governance framework to enable data sharing between the parties involved. Most (or all) UCs will fit this definition.
Data space
A decentralised infrastructure for trustworthy data sharing and exchange in data ecosystems, based on commonly agreed principles (Antti “Jogi” Poikola (SITRA) et al., 2024; Curry et al., 2022). These follow an agreed-upon set of technical specifications, standards, and governance mechanisms that allow organisations to share data without the need for a central data store. Data spaces support trust and data sovereignty.

12. Reference

Aiello, G., Donvito, R., Acuti, D., Grazzini, L., Mazzoli, V., Vannucci, V., & Viglia, G. (2020). Customers' willingness to disclose personal information throughout the customer purchase journey in retailing: The role of perceived warmth. *Journal of Retailing*, *96*(4), 490–506.

Alamoudi, E., Mehmood, R., Aljudaibi, W., Albeshri, A., & Hasan, S. (2020). *Open Source and Open Data Licenses in the Smart Infrastructure Era: Review and License Selection Frameworks* (pp. 537–559). https://doi.org/10.1007/978-3-030-13705-2_22

Analytics8. (2024, December 18). *7 Elements of a Data Strategy | Analytics8*. Analytics8 Website. https://www.analytics8.com/blog/7-elements-of-a-data-strategy/

Antti "Jogi" Poikola (SITRA), Bert Verdonck (Fraunhofer), Rieks Joosten (TNO), Tobias Guggenberger (Fraunhofer), & Satu Salminen (SITRA). (2024). *DSSC Report*. https://docs.google.com/document/d/15x6WHHGSoG4ZuXQw8u3AinpJrgbydriL/edit

Beer, S. (1984). The viable system model: Its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, *35*(1), 7–25.

Boris Otto, Michael ten Hompel, & Stefan Wrobel. (2022). Designing Data Spaces. In *Designing Data Spaces* (1st ed.). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-030-93975-5

Carol Tenopir, S. A. D. U. A. W. R. M. F. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLOS ONE*, *6*(6), 1–21. https://doi.org/10.1371/journal.pone.0021101

Christenson, A., Hemmingsson, J. U., & Lagerros, Y. T. (2024). A national data sharing solution for the prevention and treatment of obesity—a qualitative study of stakeholders' needs. *DIGITAL HEALTH*, *10*, 20552076241297740. https://doi.org/10.1177/20552076241297740

Copa, C., CEMA, F. E., Ceettar, C., Ecpa, E., & Fefac, E. S. A. (2018). *EU Code of conduct on agricultural data sharing by contractual agreement*. Academic Press. https://fefac.eu/wp-content/uploads/2020/07/eu_code_of_conduct_on_agricultural_data_sharing-1.pdf

Cor Verdouw. (2024). *Towards digital ecosystems for agrifood supply chains. 2nd Workshop on Digital Innovation Ecosystems in Agri-Food Systems*.

Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In J. M. Cavanillas, E. Curry, & W. Wahlster (Eds.), *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe* (pp. 29–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-21569-3_3

Curry, E., Scerri, S., & Tuikka, T. (2022). Data Spaces: Design, Deployment, and Future Directions. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data Spaces : Design, Deployment and Future Directions* (pp. 1–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_1

DalleMule, L., & Davenport, T. H. (2017). What's your data strategy. *Harvard Business Review*, *95*(3), 112–121.

Davenport, T. H., & Short, J. E. (2003). The new industrial engineering: Information technology and business process redesign. *Operations Management: Critical Perspectives on Business and Management*, 97–123.

de Vreis, J., Turner, J., Finlay-Smits, S., Ryan, A., & Klerkx, L. (2022). Trust in agri-food value chains: a systematic review. *International Food and Agribusiness Management Association*. https://doi.org/10.22434/IFAMR2022.0032

Durrant, A., Markovic, M., Matthews, D., May, D., Leontidis, G., & Enright, J. (2021). How might technology rise to the challenge of data sharing in agri-food? *Global Food Security*, *28*, 100493. https://doi.org/https://doi.org/10.1016/j.gfs.2021.100493

European Commission. (2016, April 27). *What personal data is considered sensitive?* . Eur-Lex.Europa.Eu. https://commission.europa.eu/law/law-topic/data-protection/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

European Commission. (2020). *Communication from the commission to the European parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "A European strategy for data," Brussels*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066

European Union. (2020, February). *European data strategy*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en.

Eva Maes, & Ella Deroo. (2023, August 8). *Good practice "Determine the value of data."*

Faroukhi, A. Z., El Alaoui, I., Youssef, G., & Amine, A. (2020). Big data monetization throughout Big Data Value Chain: a comprehensive review. *Journal of Big Data*, *7*, 3. https://doi.org/10.1186/s40537-019-0281-5

Finck, M., & Pallas, F. (2019). They Who Must Not Be Identified - Distinguishing Personal from Non-Personal Data Under the GDPR. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3462948

Fischer, C. (2013). Trust and communication in European agri-food chains. *Supply Chain Management: An International Journal*, *18*. https://doi.org/10.1108/13598541311318836

Gabriel, A., Gandorfer, M., & Spykman, O. (2021). *Nutzung und Hemmnisse digitaler Technologien in der Landwirtschaft Sichtweisen aus der Praxis und in den Fachmedien*. https://doi.org/https://doi.org/10.12767/buel.v99i1.328

Goethals, F. G. (2008). Important Issues for Evaluating Inter-Organizational Data Integration Configurations. *Electronic Journal of Information Systems Evaluation*, *11*(3), pp127-138.

Griffiths, E., Hitzler, P., Janowicz, K., Solanki, M., Sharda, A., Shimizu, C., Wallace, E., Brewster, C., Kalatzis, N., Nouwt, B., Kruiger, H., & Verhoosel, J. (2023). Data sharing in agricultural supply chains: Using semantics to enable sustainable food systems. *Semantic Web*, *15*(4), 1207–1237. https://doi.org/10.3233/SW-233287

Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: what it is and what it means*. *Information & Communications Technology Law*, *28*(1), 65–98. https://doi.org/10.1080/13600834.2019.1573501

Ingram, J., Maye, D., Bailye, C., Barnes, A., Bear, C., Bell, M., Cutress, D., Davies, L., de Boon, A., Dinnie, L., Gairdner, J., Hafferty, C., Holloway, L., Kindred, D., Kirby, D., Leake, B., Manning, L., Marchant, B., Morse, A., … Wilson, L. (2022). What are the priority research questions for digital agriculture? *Land Use Policy*, *114*, 105962. https://doi.org/10.1016/J.LANDUSEPOL.2021.105962

ISO/IEC/IEEE. (2022). *ISO/IEC/IEEE 42010:2011(en), Systems and software engineering — Architecture description*. https://www.iso.org/obp/ui/#iso:std:iso-iec-ieee:42010:ed-1:v1:en

Labaree, R. V. (2024). *Research Guides: Organizing Your Social Sciences Research Paper: Theoretical Framework*. https://libguides.usc.edu/writingguide/theoreticalframework

Lazzarini, S., Chaddad, F. R., & Cook, M. (2001). Integrating supply chain and network analysis: The study of netchains. *Journal of Chains and Network Science*, *1*.

Luo, X., & Xu, X. (2018). Infrastructure, value chains, and economic upgrades. *Journal of Infrastructure, Policy and Development*, *2*, 1. https://doi.org/10.24294/jipd.v2i2.691

Marr, B. (2017). *Data strategy: How to profit from a world of big data, analytics and the internet of things*. Kogan Page Publishers.

Merriam-Webster. (n.d.). *Framework*. In Merriam-Webster.Com Dictionary. Retrieved January 3, 2025, from https://www.merriam-webster.com/dictionary/framework

Micha, R., Coates, J., Leclercq, C., Charrondiere, U. R., & Mozaffarian, D. (2018). Global Dietary Surveillance: Data Gaps and Challenges. *Food and Nutrition Bulletin*, *39*(2), 175–205. https://doi.org/10.1177/0379572117752986

Michael E. Porter. (1985). in Competitive Advantage: Creating and Sustaining Superior Performance. In *The value chain and competitive advantage* (1st ed., pp. 33–61). Free Press.

Muhammad Hassan. (2024, March 25). *Conceptual Framework - Types, Methodology and Examples*. https://researchmethod.net/conceptual-framework/

Porter, M. E. (2008). *Competitive advantage: Creating and sustaining superior performance*. simon and schuster.

Quinn, P., & Malgieri, G. (2020). The Difficulty of Defining Sensitive Data – The Concept of Sensitive Data in the EU Data Protection Framework. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3713134

Sjaak Wolfert, Lan van Wassenaer, Simone van der Burg, Mark Ryan, Laurens Klerkx, Kelly Rijswijk, Mariette McCampbell, Ioannis Athanasiadis, & George Beers. (2021). *Navigating the Twilight Zone: Pathways towards digital transformation of food systems*. https://doi.org/https://doi.org/10.18174/552346

Sołtysiński, T., Niederhausen, J., & Eichstädt, S. (2024). Digital Calibration Certificate in a trusted quality infrastructure federated data space: A proof of concept. *Measurement: Sensors*, 101484. https://doi.org/10.1016/J.MEASEN.2024.101484

Susha, I., D. Rukanova, B., Ramon Gil-Garcia, J., Hua Tan, Y., & Gasco, M. (2019). Identifying mechanisms for achieving voluntary data sharing in cross-sector partnerships for public good*. *Proceedings of the 20th Annual International Conference on Digital Government Research*, 227–236. https://doi.org/10.1145/3325112.3325265

Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization*, *1*(3), 223–247. https://doi.org/10.1016/0167-2681(80)90002-5

Termeer, C. J. A. M., Dewulf, A., & Lieshout, M. (2010). Disentangling Scale Approaches in Governance Research: Comparing Monocentric, Multilevel, and Adaptive Governance. *Ecology and Society 15 (2010) 4*, *15*. https://doi.org/10.5751/ES-03798-150429

Top, J., Janssen, S., Boogaard, H., Knapen, R., & Şimşek-Şenel, G. (2022). Cultivating FAIR principles for agri-food data. *Computers and Electronics in Agriculture*, *196*, 106909. https://doi.org/https://doi.org/10.1016/j.compag.2022.106909

van der Burg, S., Wiseman, L., & Krkeljas, J. (2021a). Trust in farm data sharing: reflections on the EU code of conduct for agricultural data sharing. *Ethics and Information Technology*, *23*(3), 185–198. https://doi.org/10.1007/s10676-020-09543-1

van der Burg, S., Wiseman, L., & Krkeljas, J. (2021b). Trust in farm data sharing: reflections on the EU code of conduct for agricultural data sharing. *Ethics and Information Technology*, *23*. https://doi.org/10.1007/s10676-020-09543-1

Van der Vorst, J., Beulens, A., & van Beek, P. (2005). Innovations in logistics and ICT in food supply chain networks. In *Innovation in agri-food systems* (pp. 245–292). Wageningen Academic.

Wiseman, L., Sanderson, J., & Robb, L. (2018). Rethinking Ag data ownership. *Farm Policy Journal*, *15*(1), 71–77.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming–a review. *Agricultural Systems*, *153*, 69–80.

13.Annex

T2.1 Tool

Supply Chain process		Input	Process	Output	Post-output	Examples and clarification:
	What are the activities of the supply chain involved in your UC?					Farming input, production, transport, trade, procssing, retail, consumers, etc.
	Who are the the main types of suppliers, service providers involved in supply chain process of the use case?					suppliers, logistics, supermarkets, factories, etc.
	Who are the key process actors (partners) involved in supply chain process of the use case? Name the partners and other essential organisations.					e.g. Company X doing Y, famers, consumers, etc.. DjustConnect is data sharing platform
Supply chain strategy and governance						
	What roles do each of these partners play in this supply chain?					Nijsen provides circular feeds for pork farmers who also supply for the Colruyt
	Who are the decision makers involved in each stage of this supply chain?					e.g. Company X, etc..
	Which data-driven innovation do you want to work on?					
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain ?					Critical events or checkpoints are specific points within the supply chain where essential activities or checks are performed to ensure quality, compliance, and traceability. Such as: inspections, audits, certifications, or specific actions to meet standards like ISO, HACCP, or other industry-specific regulations.
	What is required to verify that each critical event meets regulatory or sustainability standards?					At each critical event, specific data or documents are collected to verify compliance with standards. Includes: quality certificates, regulatory forms, temperature logs, inspection

						records, or sustainability documentation such as Lab test reports, GPS tracking data, batch tracking data etc...
Data chain process						
Data capture						
	Which data is available in each part of chain?					Indicate what kind of data is available in each supply chain(for different partners) and indicate through what methods or applications etc. Example: In the Food retail&provision -- Retailers: data of consumer preference, products, sales are available on Application logs, mobile apps, sensors, application databases, other records or documents, manual inputs etc... e.g. Private data from retailers -- Product preference data, purchasing frequency, customer demographics, and loyalty program activity etc...
	What type of data is collected in each part of chain?					Indicate: ①public, private, licensed data, open data. (Definition is in Glossary document)， ②structured data, unstructured data, and semi-structured data, audio, video, text etc.
	How is the data collected?					e.g. Application logs, mobile apps, sensors, application databases, other records or documents, manual inputs
	What is the estimated volume of data collected?					A rough indication of the size of the data-set
	What data standards are used(e.g., GS1, ISO, proprietary standards)?					Data standards and formats of the data-set, e.g. CSV, XML, JSON, other
	Does the dataset contain data considered personal in the GDPR?					What is considered personal data by GDPR: see https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_en
	Does the dataset contain data considered sensitive in the GDPR?					Does your data include data that is considered sensitive data: genetic data, biometric data processed solely to identify a human being; health-related data; data concerning a person's sex life or sexual orientation.
	How often is the data updated?					e.g. continuous, daily, yearly, irregular

Data processing						
	Is there data processing? If yes, how?					Data processing refers to the systematic handling and transformation of raw data into valuable insights or outputs. Please briefly explain how the data is processed through what system/platform/software etc...
	What data needs to be standarized? To which format?					E.g. Unstructured Data (images, audio, video) or semi-structured data into CSV, XML, JSON, other
Data analytics						
	Is there data analysis? If yes, how?					Briefly explain statistical analysis methods
Data sharing						
	What data are(can be/willing to be) shared in the process? Under which conditions?					Conditions: license, anonymisation, bilateral agreemenet, etc.;
	What data are not (cannot be/not willing to be) shared in the process? Why?					e.g. Private data - unsure about who can access this data and how we can ensure that my company retains the rights & ownership over this data. OR second example: Licensed data - not sure if we can share this data with the reference architecture under the current license
	What systems/technology is used to share the data ?					e.g. data cloud, csv, data sharing platform, etc...
	Are there third-party stakeholders involved in data sharing who are not part of the project?					E.g. company x
Data network						
	Who are the data providers?					Data providers are entities or individuals that generate, collect, or supply data to others such as farmers or companies.
	Who are the data consumers?					Data consumers are entities that utilize the data provided by others to make informed decisions or to conduct further analysis such as researchers or retailers
	Who are the data users?					Data users are individuals or systems that directly interact with data for specific purposes such as consumers or companies/retailers

	If there is a technical failure or data problem, how can stakeholders in the data network work together to restore or repair the data flow?					e.g. Contingency Plans? But what exact action/systems/tech?
Data infrastructure						
Platforms and tools						
	What systems/technology/platforms/tools/softwares is (could) used for data identification?					e.g. DOIs,UUIDs (Universally Unique Identifiers), Dublin Core etc...
	What systems/technology/platforms/tools/softwares is used for data capture?					e.g. Surveys and Forms, IoT etc...
	What technologies/platforms/tools/softwares are used to ensure the accuracy of data flow?					Examples of Systems/tech: Blockchain, AI/ML models, APIs, IoT sensors, Data validation and ETL tools, Encryption and secure protocols
Data storage						
	What platforms and tools currently in use or in development to store data in the use case?					Tools and platforms you are currently using to store your data, e.g. specific software that you are using for data storage, specific technologies (e.g. machine learning algorithms) that you use, specific platforms where you store your data
	Where will data (preferably) be stored if shared with other parties?					With you as the partner, at an external organisation/company, in a data space, anywhere, etc
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?					
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?					
Data Strategy&governance						
Value proposition						
	What is the common goal for all the partners?					
	What is the win for every partner in the supply chain?					

	How will the collaboration take place, what action does every partner do?					Company X starts sharing info Y…
	Are there any risks in the data chain process that need to be mitigated?					Risk: Data Inconsistency and Misalignment, Data Loss During Transformation, Data Duplication, Security risk, Delays in syncing data etc. How do UCs mitigate the risks?
(Cyber)security						
	What specific security measures are in place to ensure the integrity and confidentiality of the data?					Such as encryption, authentication, etc.
	Are there specific solutions to ensure data security?					Such as multi-factor authentication, permission management, log auditing and other measures
Ethics						
	What ethical guidelines/principles are followed throughout the chain?					Please specify the guidlines/principles, examples are GDPR Compliance, Confidentiality Agreements,
	What measures are in place to ensure compliance with data privacy regulations?					E.g. Customer Data Protection, Compliance Monitoring
Data management						
	Who owns the data?					Define the ownership of data. It is possible to be multi-owners. Data ownership typically lies with the entity that generates, collects, or originally holds the data including controlling access, usage, and sharing of the data. Ownership might be shared or partially transferred based on project agreements, often detailed in contracts or grant agreements. However, primary ownership usually remains with the original data provider, who may allow restricted access to specific partners for defined uses. Even if multiple partners have access to data, ownership remains with the original provider unless otherwise specified.

	Who has access to the data?					Access is possible at different stages for different project partners, companies, and organisations. For example, company X needs to be able to upload data and receive outputs, project partner Y needs to be able to access one of the software tools listed above and be able to collect outputs. Company Z + project partner B need to be able to access and share personal sensitive data
	What specific data elements need to be traced through the system?					Provide traceability information relevant to each product in the supply chain, such as batch numbers, product identification numbers, or other critical IDs.
	Which identification standards will be used to ensure traceability across the supply chain?					List any identification types (e.g., GTIN, GLN, batch or lot numbers) and specify the standards (e.g., GS1) governing them. This information ensures that traceability requirements align with established standards throughout the supply chain.
Legal and Policy, Regulatory Frameworks						
	What European policies/protocols are followed when data is shared between supply chain actors?					Example: GDPR, DGA, EU Directive 2011/24/EU on cross-border healthcare
	Are there any special industry compliance requirements or regulations involved in the data sharing process?					Example: HACCP in the food industry, ISO 22000
	If the supply chain involves cross-border data sharing, are there different legal or compliance requirements to comply with?					Cross-border data sharing in the EU, particularly in the context of your case, involves additional legal and compliance requirements to ensure that data is handled in accordance with both EU-wide regulations and country-specific laws.
Business model development						
	How do you come up/adapt your business model to this data driven innovation?					Exaplain it based on your case
Expected changes, potential needs						
	Are there any gaps or missing data that need addressing?					

UC 1 input

Supply Chain process		Input		Process		Output	Post-output
	What are the activities of the supply chain involved in your UC?	Farming and production input	COLLECTION & TRANSPORT (olives)	MILLING & STORAGE	BOTTLING	OLIVE OIL BOTTLE	Example: Food loss/waste
	Who are the the main types of suppliers, service providers involved in supply chain process of the use case?	Farmers (Olive Growers), owners of mills, bottlers, olive oil producers (can include all the above)	Farmers	Milling companies	Olive oil producers	Olive oil producers, or retailers	Example: Waste collection and recycling centers…
	Who are the key process actors (partners) involved in supply chain process of the use case? Name the partners and other essential organisations.	Company X producing Olive oil, Farmers Y providing olive fruits), BIOCOS DNA traceability and Machine Learning post-processing analysis					
Supply chain strategy and governance							
	What roles do each of these partners play in this supply chain?	Farmers provide raw material, Milling companies mill olives to olive oil, Olive oil companies store bottle and sale, BioCoS is analyzing via DNA data driven ML teachniques data.					

	Who are the decision makers involved in each stage of this supply chain?	Olive oil company					
	Which data-driven innovation do you want to work on?	DNA traceability and Sustainability framework					
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain ?	Pre- and post-milling stage, storage conditions to ensure quality of the olive oil. Analytical outcomes and documents to ensure compliance and traceability of the label. Production documentation from field to store and DNA traceability					
	What is required to verify that each critical event meets regulatory or sustainability standards?	At each critical event, specific data or documents are collected to verify compliance Transport certificates, quality analysis, lab test reports, batch annotation data, lot numbers.					
Data chain process							
Data capture							

	Which data is available in each part of chain?	Growers/Farmers: private data (transport documentation); Milling process (sensitive and private): olive quality indexes, olives volume (weight), olive oil produced. Storage (sensitive and private data): olive oil volume, storage tank, volume of tank, date of storage; Bottling: number of bottles (sensitive), date of bottling (lot number-public). The records are by manual input in documents or govermental softwares. BioCoS DNA traceability outcomes are provided as report (document, confidential data).					
	What type of data is collected in each part of chain?	Private, Sensitive and documents					
	How is the data collected?	Manual					
	What is the estimated volume of data collected?	Few Mb					
	What data standards are used(e.g., GS1, ISO, proprietary standards)?	Proprietary standards					
	Does the dataset contain data considered personal in the GDPR?	Yes					

	Does the dataset contain data considered sensitive in the GDPR?	yes					
	How often is the data updated?	every harvesting season					
Data processing							
	Is there data processing? If yes, how?	DNA traceability data are processed combining DNA and ML (BioCOS)					
	What data needs to be standarized? To which format?	No					
Data analytics							
	Is there data analysis? If yes, how?	DNA testing using molecular techniques and ML					
Data sharing							
	What data are(can be/willing to be) shared in the process? Under which conditions?	Cultivation, production, DNA traceability.					
	What data are not (cannot be/not willing to be) shared in the process? Why?	Experimental protocols, primers sequences, ML methodology due to IPs					
	What systems/technology is used to share the data ?	csv or cloud					
	Are there third-party stakeholders involved in data sharing who are not part of the project?	Olive oil producing companies and associations					
Data infrastructure							
Platforms and tools							
	What systems/technology/platforms/tools/softwares is (could) used for data identification?						

	What systems/technology/platforms/tools/softwares is used for data capture?	Manual collection					
	What technologies/platforms/tools/softwares are used to ensure the accuracy of data flow?	Internal relational databases and ML pipelines					
Data storage							
	What platforms and tools currently in use or in development to store data in the use case?	Internal database					
	Where will data (preferably) be stored if shared with other parties?	Cultivation and production data and DNA outcomes					
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?	Internal platform					
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?	security and GDPR					
Data Strategy&governance							
	How will the collaboration take place, what action does every partner do?	BIOCOS will invite relevant stakeholders and collect the relavant information to scale the DNA traceability system. Agroknow will bring public data (environmental) to be connected with the DNA traceability					

		outcomes for establishing a sustainability framework					
	Are there any risks in the data chain process that need to be mitigated?	intellectual property rights and background knowledge					
Data management							
	Who owns the data?	Olive Oil production data are ownership of the companies (stakeholders), DNA traceabilty outcomes and DNA testing protocols are of BioCoS.					
	Who has access to the data?	Not defined yet - However, access will be possible to relevant partners and companies. Olive oil stakholders to share data and receive outputs. BioCoS to have access to perform DNA traceability					

	What specific data elements need to be traced through the system?	Information of production, DNA outcomes and environmental data. Up to now identified 5 main categories of the first two: 1) Field & Harvest Data, 2) Milling, 3) Storage, 4) Bottling and 5) DNA data					
	Which identification standards will be used to ensure traceability across the supply chain?	Geolocus, Legal documents, Olive batches, olive oil lot number					

UC2 input

Supply Chain process		Input		Process		Output	Post-output	
	What are the activities of the supply chain involved in your UC?	**Sourcing and Procurement**	**Storage and inventory monitoring**	**Food preparation**	**Meals transport and distribution**	**Consumption and feedback**	**Food waste management**	**Monitoring and Data analysis**
	Who are the the main types of suppliers, service providers involved in supply chain process of the use case?	Hospital's Proccurement department that will contact and select the suppliers:	i. Person responsible for checking the cold and other storage facilities	i. Nutritionists & dietitians	i. Hospital staff that are responsible for meal distribution and meal delibery sercices	i. Hospital staff responsible for tracking patient health realted to dietary intake	i. Hospital staff responsible for seperating waste	i. Provider of the dashboard for analyzing food usage and waste
		i. Local farms	ii. Logistics firms	ii. Cooks	ii. Canteen staff	ii. Hoospital staff managing patient interactions	ii. Outsource company for collecting and transporting waste to appropriate facilities	ii.Hospital staff interpreting data to improve supply chain efficiency
		ii. Specialthy food producers						

		iii. Wholesalers						
	Who are the key process actors (partners) involved in supply chain process of the use case? Name the partners and other essential organisations.	UKCM (Procurement department), Food producers and wholesalers	Logistic firms, UKCM (logistics department)	UKCM (Nutritionists, dietitians and cooks)	UKCM (Hospital staff and canteen staff)	Patients and UKCM (nurses and hospital staff responsible for meal distribution/collection)	UKCM (Hospital staff responsible for seperating waste), Waste outsource company	INTRA (Dashboard creation and maintenance), UKCM (Hospital staff responsible for data analysis)
Supply chain strategy and governance								
	What roles do each of these partners play in this supply chain?	**UKCM (Procurement Department):**	**Logistics firms:**	**UKCM (Nutritionists, dietitians, and cooks):**	**UKCM (Hospital staff and canteen staff):**	**Patients:**	**UKCM (Hospital staff responsible for separating waste):**	**INTRA (Dashboard creation and maintenance):**
		Manages sourcing strategies and partnerships with sustainable suppliers.	**Provide** transportation solutions to move ingredients to storage facilities.	Develop patient-specific menus aligned with dietary and health needs.	Coordinate the transport of prepared meals to patients in a timely manner.	Act as end-users, consuming meals and providing feedback on quality and satisfaction.	Seperate food waste into appropriate categories for recycling or composting.	Develop and maintain digital tools for tracking food usage, waste, and patient feedback.
		Ensures procureme	Ensure timely and	Prepare meals	Manage meal	UKCM (Nurses and Hospital Staff	Ensure compliance	Provide real-time data for

		nt aligns with sustainability and health goals.	efficient delivery with minimal spoilage.	using sustainable practices to minimize waste and optimize resources.	distribution points within the hospital.	Responsible for Meal	with waste management policies.	decision-making and process optimization.
		Handles supplier evaluation and contract management.						
			UKCM (Logistics Department):			**Distribution/collection):**	**Waste outsource Company:**	**UKCM (Hospital staff responsible for Data Analysis):**
		Food producers and Wholesalers:	Oversees storage conditions, inventory tracking, and efficient resource utilization.			Deliver meals to patients, ensuring dietary compliance and portion accuracy.	Collect and process waste for composting or conversion to bioenergy.	Analyze data to identify inefficiencies and areas for improvement.
		Supply fresh, organic, and	Implements processes to minimize food loss			Collect leftovers for feedback on meal preferences	Handle recycling of non-organic materials,	Implement recommendations to refine

		nutrient-rich ingredients.	and ensure timely restocking.			and waste analysis.	ensuring sustainability goals are met.	sourcing, preparation, and distribution practices.
		Ensure consistent delivery of quality products to meet hospital requirements.						
	Who are the decision makers involved in each stage of this supply chain?	UKCM (Procurement department)	UKCM (Logistics team)	UKCM (cooks)	UKCM (Hospital staff responsible for meals transport and distribution)	UKCM (Hospital staff - nurses)	UKCM (Hospital staff responsible for separating waste)	UKCM (Hospital staff responsible for Data Analysis)
	Which data-driven innovation do you want to work on?	Data-Driven Innovation	Data-Driven Innovation	Data-Driven Innovation	We won't focus on this step of the food chain.	Data-Driven Innovation	Data-Driven Innovation	Data-Driven Innovation
		The dashboard will help to predictive analytics to forecast	The dashboard will help to track inventory levels,	The dashboard will stand as a recipe optimization tool, that		The platform will support the patients' feedback analytics, through collection and	The platform will support food waste recording generated at various	The dashboard will centralize data from all stages of the

		ingredient needs and optimize procurement schedules.	predict stock depletion, and suggest reordering timelines.	balance patient dietary requirements with available resources. Also it will provide insights regarding the waste reduction based on historical data from meal preparation processes.		analysis of feedback on meals using satisfaction surveys.	stages (preparation, distribution, consumption)	food chain, providing real-time insights and performance metrics.
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain ?	**(Traceability checkpoints**: Ensuring every batch of ingredients is traceable back to its source.)	**Inventory inspections :** Routine audits to ensure stock integrity and expiration date tracking.	**Nutritional quality**	(**Delivery time monitoring :** Tracking delivery schedules to ensure meals are fresh upon arrival.)	**Patient feedback collection:** Regularly gathering feedback from patients on meal quality and satisfaction.	**Waste segregation audits:** Ensuring proper separation of organic and non-organic waste.	**Data validation:** Ensuring accuracy and reliability of data inputs from all supply chain stages.
				assessments: Routine analysis to		**Dietary compliance checks**: Ensuring	**Waste disposal records:**	**System updates:** Regular

				confirm meals meet prescribed dietary standards.		meals align with prescribed dietary requirements for each patient.	Documenting the quantity and type of waste sent for composting or bioenergy.	updates and maintenance of dashboards and analytics tools.
						Leftovers analysis: Monitoring uneaten portions to assess meal suitability and reduce waste.	**Environmental impact assessments:** Periodic evaluations of waste management processes for sustainability.	**Compliance reporting:** Generating reports for regulatory bodies on quality, safety, and sustainability metrics.
						Traceability logs: Keeping records of meal distribution for accountability		
	What is required to verify that each critical event meets regulatory or sustainability standards?	i. Supplier documentation, such as certificates for organic, fair trade, or sustainabili	i. Regular inventory logs documenting product condition, shelf life, and quantity.	i. Laboratory testing of meals to confirm nutritional content matches prescribed standards.	Logs of transportation times and temperature conditions during transit.	**Patient feedback collection**	**Waste segregation audits**	**Data validation**

		ty practices.						
		ii. Periodic audits of suppliers to ensure compliance with agreed standards.	ii. Compliance checks with food storage standards, such as HACCP (Hazard Analysis and Critical Control Points).	ii. Recipe documentation showing adherence to dietary guidelines.		i. Use of standardized patient satisfaction surveys.	i. Visual inspections and audit reports verifying correct segregation of organic and inorganic waste.	i. Cross-checking data against source records to confirm accuracy.
		iii. Alignment with local and international traceability regulations (e.g., ISO 22005 for food traceability).	iii. Reports from storage audits conducted by internal or external inspectors.			ii. Documentation of feedback and integration into improvement plans.	ii. Compliance with local waste management laws (e.g., EU Waste Framework Directive).	**System updates**
						Dietary compliance checks	**Waste disposal records**	i. Maintenance logs documenting updates to

								dashboards and analytics tools.
						i. Patient dietary records and compliance logs verified by nutritionists or medical staff.	i. Logs documenting waste type, quantity, and destination for composting or recycling.	ii. Compliance with cybersecurity and data protection regulations (e.g., GDPR in the EU).
						Leftovers analysis	ii. Certifications from waste processing companies ensuring responsible handling.	**Compliance reporting**
						i. Quantitative records of uneaten portions to identify patterns and areas for improvement.	**Environmental impact assessments**	i. Generating reports adhering to regulatory requirements for food quality and safety.
						Traceability logs	i. Reports analyzing the carbon footprint and	ii. Audits to ensure accuracy and

							sustainability of waste processes.	completeness of submitted reports.
						i. Detailed logs of meal distribution, documenting time, location, and recipient.	ii. Periodic evaluations against environmental standards, such as ISO 14001.	
Data chain process								
Data capture								
	Which data is available in each part of chain?	i. Ingredient specifications (type, quality, organic certification, batch numbers).	i. Real-time inventory levels (quantity, location, and status of items).	i. Recipes, seasonal menus (SS/AW) and nutritional profiles for each meal.	i. Delivery schedules and routes (planned and actual).	i. Patient feedback on meal quality, portion size, and satisfaction.	i. Waste segregation logs (type and quantity of organic and inorganic waste).	i. Centralized data from all food chain stages (sourcing, storage, preparation, distribution, consumption, and waste).
		ii. Traceability records linking ingredients to their sources.	ii. Inventory movement logs (inbound and outbound).	ii. Quantity and type of ingredients used for meal preparation.	ii. Records of distributed meals, including time and	ii. Records of dietary compliance for individual patients.	ii. Records of waste collection schedules.	ii. Historical trends and predictive insights.

					recipient details			
			iii. Historical trends of inventory usage and replenishment cycles.	iii. Data on preparation time and efficiency.		iii. Data on leftovers and uneaten portions.		iii. Reports for regulatory compliance and internal audits.
				iv. Meal customization data (e.g., dietary requirements and patient preferences).		iv. Trends in meal preferences and dietary adjustments over time.		
						v. Log of meals served per day by type and quantity.		
	What type of data is collected in each part of chain?	**Ingredient data:** Batch numbers, quality grades, nutritional content.	**Inventory levels:** Quantities of raw materials and finished goods.	**Recipe data**: Ingredients used, nutritional values, and dietary requirements.	**Route data**: Planned and actual delivery routes.	**Feedback data:** Patient satisfaction scores, comments, and preferences.	**Segregation data**: Types and quantities of organic and inorganic waste.	**Integrated data:** Aggregated data from sourcing, storage, preparation, and distribution.
		Quantity data: About	**Movement data**:	**Input usage**	**Recipient data:**	**Consumption data:** Records of	**Disposal records:**	**KPI data**: Metrics on

		the products delivered to the hospital facilities.	Records of stock received and dispatched.	**data:** Quantities of ingredients used in each recipe.	Records of meal distribution to patients.	meals consumed vs. returned.	Waste collected, processed, or recycled (e.g., composted, bioenergy).	waste reduction, cost efficiency, and dietary compliance.
				Time data: Preparation time for each meal type.		**Dietary compliance data**: Alignment of meals with patient dietary needs.		**Compliance data:** Reports for internal audits and regulatory submissions .
						Portion data: Leftover quantities for assessing portion accuracy.		
						Trend data: Patterns in meal preferences over time.		
	How is the data collected?	**Ingredient data (Nutritional Content):**	**Inventory levels:**	**Recipe data:**	Record through a badge with GSRN	Use of bracelets with 2D/QR around patients wrist (GSRN + SRIN)	**Segregation data:**	**Integrated data:**
		-Collected through supplier-	Regular updates logged into	-Stored in a recipe manageme	**Recipient data:**	**Feedback data:**	-Collected through visual	-Aggregated from multiple

		provided documentation and certifications.	inventory management software.	nt database linked to nutritional analysis tools.			inspection sorting systems.	systems into a centralized dashboard.
		-Nutritional data sourced from supplier specifications and verified through lab testing.	**Movement data:**	-Collected during meal planning by nutritionists and dietitians.	Collected via distribution systems, with manual confirmation by hospital staff.	Gathered through surveys conducted via digital forms or apps.	-Quantities recorded via bins.	-Collected via APIs or data integration platforms.
		Quantity data (Products Delivered):	Recorded during check-in/out processes at storage facilities.	**Input usage data:**		**Consumption data:**	**Disposal records:**	**KPI data:**
		-Input into inventory management systems upon delivery.		-Logged into preparation records through kitchen staff.		Recorded manually by nursing staff.	-Logged by waste collection companies.	-Generated from analytics software using real-time data inputs.
		-Verified by comparing delivery		**Time data:**		**Dietary compliance data:**	-Verified through receipts and	**Compliance data:**

		notes with purchase orders.					processing certificates.	
				Tracked through time logs maintained by kitchen staff.		Collected through health records and cross-referenced with meal plans.		-Generated through compliance audit logs.
						Portion data:		-Compiled into reports for review by internal or external regulators.
						Logged through weighing uneaten portions and recording manually or digitally.		
						Trend data :		
						Analyzed from historical patient feedback and consumption records.		
	What is the estimated volume of data collected?	TBD	TBD	TBD	TBD	TBD	TBD	TBD

	What data standards are used(e.g., GS1, ISO, proprietary standards)?	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.	We follow the AACI 5.0 standard, an internationally recognized framework developed by the American Accreditation Commission International (AACI), designed to uphold the highest levels of quality, safety, and patient-centered care in healthcare organizations.

	Does the dataset contain data considered personal in the GDPR?	No	No	No	Yes	Yes	No	Yes
	Does the dataset contain data considered sensitive in the GDPR?	No	No	No	No	Yes	No	Yes
	How often is the data updated?	TBD	TBD	TBD	TBD	TBD	TBD	TBD
Data processing								
	Is there data processing? If yes, how?	Collection of ingredient data (nutritional content) from suppliers.	i. Transformation of movement data (received and dispatched stock) into actionable insights for stock optimization.	i. Conversion of recipe data into portion sizes and tailored nutritional profiles.	i. Recording and validation of recipient data (meal distribution logs) for traceability.	i. Collection and analysis of patient feedback (satisfaction scores, preferences).	i. Segregation and recording of waste types and quantities.	i. Aggregation of data from all stages of the supply chain into a centralized system.
			ii.Generation of inventory usage reports for operational planning.	ii. Aggregation of meal preparation data for reporting and improveme	ii. Integration of delivery data into distribution management systems for	ii. Cross-referencing dietary compliance data with meal plans and health records.	ii.Analysis of trends in waste generation to identify inefficiencies and areas for	ii. Generation of KPI reports (e.g., waste reduction, cost efficiency,

				nt initiatives.	compliance reporting.		improvement .	dietary compliance).
						iii. Aggregation of consumption data (meals consumed vs. returned) to identify patterns.	iii.Generation of reports for regulatory and internal audits.	iii. Use of machine learning and AI models for predictive analysis and optimization.
	What data needs to be standarized? To which format?	**Data to standardize:**	**Data to standardize:**	**Data to standardize:**	**Data to standardize:**	**Data to standardize:**	**Data to standardize:**	**Data to standardize:**
		1. Ingredient data: Batch numbers, quality grades, nutritional content.	1. Inventory levels: Quantities of raw materials and finished goods.	1. Recipe data: Ingredients, nutritional values, dietary requirements.		1. Feedback Data: Satisfaction scores, comments, preferences.	1. Segregation Data: Types and quantities of organic/inorganic waste.	1. Integrated data: Data from sourcing, storage, preparation, and distribution.
		2. Quantity data: Product volumes delivered.	2. Movement data: Records of stock received and dispatched.	2. Input usage data: Quantities of ingredients used.	1. Route data: Planned and actual delivery routes.	2. Consumption data: Records of meals consumed vs. returned.	2. Disposal Records: Waste collected and processed.	2. KPI data: Metrics on waste reduction, cost efficiency, dietary compliance.
				3. Time data: Preparation times.	2. Recipient data: Meal distribution logs.	3. Dietary Compliance data: Alignment with patient needs.		

		Standardization format:	**Standardization format:**			4. Portion data: Leftover quantities.	**Standardization format:**	**Standardization format:**
		Nutritional content: Standardized unit format (e.g., grams, kcal, mg).	1. Inventory levels: Numerical format with units (e.g., kilograms, liters).	**Standardization format:**	**Standardization format:**	5.Trend data: Meal preferences over time.	1. Waste data: Use weight in kilograms (e.g., Organic: 25kg, Inorganic: 10kg).	1. Integrated data: Use a centralized schema with uniform field names and data types.
			2. Movement data: Standardized logs with timestamps in ISO 8601 format.	1. Recipe data: Structured text format with fields for each ingredient (e.g., { "ingredient": "rice", "quantity": "200g" }).	1. Route data: Geographic coordinate format		2. Disposal records: Structured log format (e.g., { "date": "2024-01-01", "type": "Composted", "quantity": "25kg" }).	2. KPI data: Use percentage or numerical formats (e.g., Waste Reduction: 15%).
				2. Nutritional values: Use per 100g or per serving format for standardization (e.g.,	2. Delivery time: specific timestamp format.	**Standardization format:**		

				Energy: 200 kcal/100g).				
				Time data: ISO 8601 duration format (e.g., PT45M for 45 minutes).		1. Feedback data: Use Likert scale (e.g., 1-5 for satisfaction)		
						2. Consumption data: Structured CSV or database format with fields (e.g., { "meal_id": "56789", "consumed": "80%", "returned": "20%" }).		
						3. Portion data: Weights in grams (e.g., 150g uneaten).		
						4. Trend data: Time series format (e.g., { "date": "2024-01-01", "preference": "Vegetarian" }).		

Data analytics								
	Is there data analysis? If yes, how?	1. Cost optimization: Identifying trends in ingredient costs to optimize procurement decisions.	1. Inventory optimization: Analyzing stock levels to reduce overstocking or shortages.	1. Recipe optimization: Analyzing ingredient usage to ensure nutritional and dietary compliance while minimizing waste.	Delivery timeliness: Assessing data on delivery times to ensure meals reach patients on schedule.	1. Feedback analysis: Sentiment analysis and scoring of patient feedback to understand satisfaction and preferences.	Waste quantification: Analyzing segregation data to understand waste patterns and volumes.	1. KPI tracking: Analyzing aggregated data to monitor performance metrics like waste reduction, cost efficiency, and compliance.
		2. Demand forecasting: Using historical procurement data to predict future needs.	2. Demand patterns: Identifying usage trends to align inventory levels with preparation needs.	2. Efficiency tracking: Measuring preparation times and processes to improve kitchen efficiency.		2. Compliance Checking: Comparing dietary data with meal distribution to ensure adherence to patient-specific requirements.		2. Predictive analytics: Using models to predict future trends, such as ingredient demand or waste generation.
				3. Waste analysis: Identifying sources of waste		3. Waste Insights: Analyzing leftovers to adjust portion sizes and		3. Operational insights: Real-time dashboards

				during preparation to refine processes and reduce leftovers.		improve meal planning.		providing insights across the supply chain for decision-making.
						4. Preference trends: Analyzing historical data to identify popular meal options and inform menu planning.		4. Compliance Reporting: Generating detailed reports to demonstrate adherence to regulatory standards.
Data sharing								
	What data are(can be/willing to be) shared in the process? Under which conditions?	Ingredient data: batch numbers, quality grades, nutritional content.	Inventory levels: quantities of raw materials and finished goods.	Recipe data: ingredients used, nutritional values, dietary requirements.	Delivery timeliness:	Feedback data: patient satisfaction scores, comments, and preferences.	Waste segregation data: types and quantities of organic and inorganic waste.	Aggregated data from sourcing, storage, preparation, and distribution.
		Quantity data: volumes delivered.	Movement data: records of stock received	Input usage data: quantities of	a. Delivery schedules, expected times, and actual	Consumption data: records of meals consumed vs. returned.		KPI data: waste reduction, cost efficiency,

			and dispatched.	ingredients used.	completion times.			dietary compliance.
				Waste generated during preparation.	b. Patient ID or anonymous identifiers to link meals with recipients (for traceability and feedback purposes).	Dietary compliance data: alignment of meals with patient needs.	Regulatory reporting: Share data with authorities for compliance with waste disposal laws.	Compliance data: audit logs and regulatory reports.
		Operational purpose: Shared only with relevant stakeholders (e.g., hospitals, suppliers) for sourcing efficiency.	Internal access only: Limited to hospital staff and supply chain managers.					
			Real-time access: Share real-time data only for operational purposes	Internal sharing: Accessible only to hospital nutritionists,	Operational conditions:	Anonymization: Patient feedback must be anonymized unless specific consent is given.		Confidentiality: Only aggregated data should be shared to protect sensitive

			like restocking or auditing.	dietitians, and kitchen staff.				operational details.
				Ethical use: Data on dietary requirements must exclude personal patient identifiers unless required by law.	a.Data sharing should occur in real-time or at predefined intervals to maintain logistical efficiency.	Compliance: Shared only under frameworks like GDPR for sensitive data handling.		Regulatory compliance: Data shared must meet legal and audit requirements.
				Data protection: Ensure that preparation records are stored and transmitted securely.	b. Sharing must align with operational priorities, such as timely route optimization or feedback integration.	Purpose limitation: Data shared only for improving meal services or auditing purposes.		Stakeholder access: Limit access to relevant internal stakeholders and approved external parties (e.g., consultants).
					Consent and ethical conditions:	Restricted access: Limit sharing to internal teams like dietitians or		

						quality assurance personnel.		
					a. Patient-related data (even anonymized) must be shared only with explicit consent if legally required.			
					b. Ethical considerations must be addressed to ensure data use aligns with patient and organizational expectations.			
	What data are not (cannot be/not willing to be) shared in the process? Why?	N/A	N/A	N/A	N/A	**Data not shared:** Detailed feedback linked to individuals: Comments or	N/A	**Data not shared:**

						ratings that can identify specific patients.		
						Sensitive dietary information: Specific medical or dietary conditions tied to individuals.		Raw data from integrated systems: Unprocessed data that might contain sensitive or identifiable information.
						Why?		**Why?**
						Data Protection laws: Ensures compliance with regulations like GDPR.		Operational sensitivity: Raw data may expose internal inefficiencies or vulnerabilities.
						Ethical considerations: Respects patient confidentiality and trust.		

	What systems/technology is used to share the data ?	MAOP: Food Orders and Procurement Platform	Currently, there is no dedicated system for storage and inventory monitoring. However, some relevant data can be extracted from MAOP.	PREHRANA serves as the primary entry point for diet and meal planning and sources its data from MEDIS, an electronic health records (EHR) platform.	Transport and distribution processes are managed dynamically, leveraging data partially sourced from MEDIS and PREHRANA.	Consumption feedback is occasionally collected manually, typically recorded in Excel, and consolidated from various data sources.	Food waste management is not performed continuously. Waste is collected and weighed post-festum by a garbage disposal service, but it is reported only as a total, without segregation per wards or departments.	Monitoring and analysis are conducted using a combination of available systems and data sources.
	Are there third-party stakeholders involved in data sharing who are not part of the project?	No	No	No	No	No	No	No
Data infrastructure								
Platforms and tools	What systems/technology/platforms/tools/softwares is (could) used for data identification?	TBD	TBD	TBD	TBD	TBD	TBD	TBD
	What systems/technology/platforms/tools/softwares is used for data capture?	TBD	TBD	TBD	TBD	TBD	TBD	TBD

	What technologies/platforms/tools/softwares are used to ensure the accuracy of data flow?	APIs	APIs	APIs	APIs	APIs	APIs	APIs
Data storage								
	What platforms and tools currently in use or in development to store data in the use case?	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base
	Where will data (preferably) be stored if shared with other parties?	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure	Cloud infrastructure
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?	TBD	TBD	TBD	TBD	TBD	TBD	TBD
Data Strategy&governance								
Value proposition								
	How will the collaboration take place, what action does every partner do?	TBD	TBD	TBD	TBD	TBD	TBD	TBD

	Are there any risks in the data chain process that need to be mitigated?	Traceability gaps: Missing or incomplete records linking ingredients to their sources.	Stock mismanagement: Errors in inventory levels leading to overstocking or shortages.	Recipe inconsistencies: Errors in ingredient data leading to incorrect nutritional values.	Route celays: Unplanned disruptions causing late meal deliveries.	Bias in feedback data: Incomplete or inaccurate patient feedback skewing results.	Improper waste tracking: Missing or incomplete records of waste segregation and disposal.	Data silos: Lack of integration across data sources leading to incomplete insights.
				Data loss: Loss of preparation logs or recipe details.		Inadequate analysis: Poorly processed data leading to incorrect insights.	Environmental impact: Lack of accurate data leading to unoptimized sustainability efforts.	Inaccurate predictions: Errors in predictive models or data analytics.
Data management								
	Who owns the data?	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through third-party sources.	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through third-	UKC Maribor holds the exclusive ownership rights to all data generated internally or obtained through

		third-party sources.	third-party sources.	third-party sources.	third-party sources.		party sources.	third-party sources.
	Who has access to the data?	TBD	TBD	TBD	TBD	TBD	TBD	TBD
	What specific data elements need to be traced through the system?	Ingredient batch numbers: For identifying the origin and production details of ingredients.	Inventory levels: Quantities of raw materials and finished goods.	Recipe ingredients: Type, quantity, and nutritional values of ingredients used.	Route information: Planned and actual delivery routes, including timestamps.	Meal records: Type of meal served, portion size, and dietary customization.	Waste segregation records: Types (organic/inorganic) and quantities of waste.	Integrated ddata: Aggregated data from sourcing, storage, preparation, and distribution.
			Movement logs: Records of stock received and dispatched, with timestamps.	Preparation logs: Records of when and how meals were prepared.	Recipient data: Meal delivery logs, including patient or location IDs.	Patient feedback: Satisfaction scores, preferences, and comments.	Disposal logs: Records of waste collection, processing, or recycling.	Performance metrics (KPIs): Waste reduction, cost efficiency, and dietary compliance rates.
				Waste records: Quantities and types of waste generated during preparation.		Consumption logs: Records of meals consumed versus leftovers.	Compliance data: Adherence to waste management regulations and sustainability goals.	Compliance records: Logs and reports for regulatory audits.

						Dietary compliance: Alignment of meals with prescribed dietary needs.		Predictive insights: Trends and forecasts generated from historical data.
	Which identification standards will be used to ensure traceability across the supply chain?	The HACCP standard ensures traceability of raw materials to their sources and obligates us to maintain records of approved suppliers and monitor the quality of food ingredients received.	HACCP mandates proper storage practices, ensuring that inventory records include details like batch numbers, dates, and storage conditions.	HACCP requires the monitoring of food preparation processes, such as cooking temperatures, storage conditions, and hygiene practices.	I context of HACCP: Meals are sorted and labelled by wards, patients, or dietary categories. This ensures meals are distributed correctly and reduces mix-ups. Routes are clearly defined and schedules minimize delays and reduce the	We use various tools and methods to monitor feedback on consumption.	Traceability extends to food waste, documenting the disposal process to ensure no contamination risks are posed by leftover or discarded food. (HACCP)	According to HACCP hospitals have to monitor CCPs (Critical Control Points) such as refrigeration, cooking, and cooling, with logs detailing time and temperature. Unfortunately we do not keeping this kind of logs.

					risk of food sitting outside safe temperature ranges. Hygiene practices are strict for transport staff, including sanitizing trolleys and wearing appropriate personal protective equipment. The use of insulated containers or temperature-controlled trolleys ensures meals remain at safe			

					temperatures etc.			

UC3 input

Supply Chain process		Farming input	Agricultural production	Storage, transport, trade	Processing & transformation	Retail	Consumption
	What are the activities of the supply chain involved in your UC?	FARMING INPUT	AGRICULTURAL PRODUCTION	STORAGE, TRANSPORT & TRADE	PROCESSING & TRANSFORMATION	RETAIL & PROVISIONING	CONSUMPTION
	Who are the the main types of suppliers, service providers involved in supply chain process of the use case?	Animal feed, crop protection products, veterinary medicines, fertilisers	porc farming dailry farming	Livestock sales/purchase, milk sales/purchase, livestock transport, milk transport	Slaughterhouse, dairy processing, B2B Meat transport, B2B dairy transport	Distribution, Food wasten retail supermarkets POS, online POS	meat and dairy products consumption, Food watse, Packaging waste, Consumer application
	Who are the key process actors (partners) involved in supply chain process of the use case? Name the partners and other essential organisations.	Nijsen company, EV ILVO	Dairy farmers Livestock farmers	Colruyt Group	Colruyt Group Slaughterhouse? Transport?	Colruyt Group, GS1	Palau BV, GS1, Inex, and Sportwatch partner
Supply chain strategy and governance							
	What roles do each of these partners play in this supply chain?	ILVO provides KLIMREK tool: LCA scans on	Nijsen provides circular feed	Retailer that works together with dairy	Retailer who is connected to	GS1:	PALAU has a consumer app guiding

		farm level to improve sustainability	for pork farmers who supply for Colruyt	farmers and milk processor which is used for 12 dairy products	the GS1 platform		consumers toward healthier and more sustainable choices
						- Establishing standards for the dashboards	experience in CSRD reporting
						and explore	
						- make connection between the dashboard and the found data	
						- look into synchronizing PALAU with their database	
	Who are the decision makers involved in each stage of this supply chain?	KLIMREK coordinator	Farmers	processing company	processing companuy	retail	consumers
		Farmers need to give consent if their data can be shared	Feed company	retail	retail	consumers	
		Retail also plays an	retail				

		important role in this					
	Which data-driven innovation do you want to work on?	Interest in Data space: Involved in DjustConnect and the AgriDataSpace project. Interest in Health data space or data spaces with consumers info (if available)	API based climate scan	API based climate scan		additional attributes in data model	new machine learning algorithms
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain ?	Data collection: consent management compliance with the Code of Conduct	secure data transfer and data quality control	secure data transfer and data quality control	secure data transfer and data quality control	use of data standards	ensure anonymization of data
	What is required to verify that each critical event meets regulatory or sustainability standards?	legal contracts with clear rules in accordance with the legislation and the Code of Conduct on agricultural data sharing, Data Act, Data Governance Act	Safe, transparent and controlled transport of data through data sharing platform	Safe, transparent and controlled transport of data through data sharing platform	Safe, transparent and controlled transport of data through data sharing platform	Interoperability: REST-API, taking standards into account	

		identoty management through e-ID					
Data chain process							
Data capture							
	Which data is available in each part of chain?	Nijsen Company: Nutritional data, dashboard sustainable feed, energy use, Agrimatie, CVB, LCA data (PEF - GFLI), commodity prices BASF/Sustel/Trouw: impact porc feed	Nijsen Company: stable climate/manure data, logistics data, cycle guide ILVO: results sustainability monitor, animal welfare scan, results Klimrek	Colruyt Group: avarage throughput/storage time	Colruyt Group: data on which farmers to which supplier, generic impact of 1 supplier, scope 3 impact reporting CO2 equivalent of 12 dairy products	Colruyt Group: generic CO2 impact of total sales	PALAU: data on scanned products, consumer diets data and evolutions, eco-score and nutri-score (general and per consumer), publicly available data per product (unconventional data)
	What type of data is collected in each part of chain?	Primary farm data (from climate scans); structured data unconventional datasets	impact of circular feed/impact logisctics (Nijsen):		CO2 equivalent of 12 dairy products		Data on CSRD advice (PALAU): unconventional datasets

			unconventional datasets				
	How is the data collected?	Structured data: Climate Scans for dairy and pig farmers (ILVO- Klimrek) (application, manual input) ; Nijsen (application);	Structured data: Nijsen (application) ;				Structured data: Consumers (IOT devices, applications, manual input); Palau advice service (application, manual input);
	What is the estimated volume of data collected?	No idea yet, but limited volume	No idea yet, but limited volume	No idea yet, but limited volume	No idea yet, but limited volume	No idea yet, but limited volume	No idea yet, but limited volume
	What data standards are used(e.g., GS1, ISO, proprietary standards)?	csv or json	csv or json	csv or json	csv or json	csv or json	csv or json
	Does the dataset contain data considered personal in the GDPR?	Yes. unconventional datasets; Primary farm data (from climate scans)					
	Does the dataset contain data considered sensitive in the GDPR?	Yes. unconventional datasets					unconventional datasets
	How often is the data updated?	Primary farm data (from climate scans)- irregular;	impact of circular feed/impact logisctics				(PALAU) - yearly; unconventional datasets

		impact of circular feed/impact logisctics (Nijsen) yearly	(Nijsen) yearly				- continious and irregular
Data processing							
	Is there data processing? If yes, how?	through data sharing platform DjustConnect	yes, looking into Sustell tool			GS1 platform	yes, within PALAU app
	What data needs to be standarized? To which format?	into csv or json	into csv or json	into csv or json	into csv or json	into csv or json	into csv or json
Data analytics							
	Is there data analysis? If yes, how?	yes, calculation model of KLIMREK	Sustell calculation tool			Yes, eco-score and nutri-score calculation	
Data sharing							
	What data are(can be/willing to be) shared in the process? Under which conditions?	All data, with the consent of the farmers and under legal contracts with clear rules in accordance with the legislation and the Code of	agreements are needed	agreements are needed	agreements are needed	agreements are needed	anonymisation is needed

		Conduct on agricultural data sharing, taking into account Data Act, Data Governance Act					
	What data are not (cannot be/not willing to be) shared in the process? Why?	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non personal, but sensitive data.	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non personal,	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non personal, but sensitive data.	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non personal, but sensitive data.	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non	Not at this moment; however, data sets will only be shared for the purpose and to the partners that will get consents of both the data provider and the data subject. datasets will not be public available nor open; This includes both personal and non personal, but

			but sensitive data.			personal, but sensitive data.	sensitive data.
	What systems/technology is used to share the data ?	excel, document, csv or api	excel, document, csv or api	excel, document, csv or api	excel, document, csv or api	excel, document, csv or api	excel, document, csv or api
	Are there third-party stakeholders involved in data sharing who are not part of the project?	not now, but maybe later in the project					
Data infrastructure							
Platforms and tools	What systems/technology/platforms/tools/softwares is (could) used for data identification?	Data sharing platform DjustConnect and API based climate scan application					
	What systems/technology/platforms/tools/softwares is used for data capture?	Data sharing platform DjustConnect and API based climate scan application existing data sets				retail application (needs to be build: POC)	application used by consumers
	What technologies/platforms/tools/softwares are used to ensure the accuracy of data flow?	Data sharing platform DjustConnect and API based climate scan application					

Data storage							
	What platforms and tools currently in use or in development to store data in the use case?	own database or data vault	own data base	own data base	own data base	own data base	PALAU: consumer app
		webserver and external APIs					own data base
	Where will data (preferably) be stored if shared with other parties?	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization , shared through API	Structured data in data base of organization, shared through API
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization, shared through API	Structured data in data base of organization , shared through API	Structured data in data base of organization, shared through API
		Relational data base	Relational data base	Relational data base	Relational data base	Relational data base	Relational data base
		Specific data formats: JSON	Specific data formats: JSON	Specific data formats: JSON	Specific data formats: JSON	Specific data formats: JSON	Specific data formats: JSON
Data Strategy&governance							
Value proposition							

	How will the collaboration take place, what action does every partner do?	Colruyt Group will focus on addressing the Scope 3 climate impact, with particular attention to dairy farmers, supporting Life Cycle Assessments (LCA) at several farms to help reduce their climate impact and increase resilience. Additionally, Colruyt will support the creation of a Retail dashboard (B2B) to aid in sustainable supply chain management. Nijsen Company, on the other hand, will focus on the pig sector					

		by providing primary agricultural data sets and contributing to the development of the Retail dashboard. They will also explore social aspects and investigate collaboration with Colruyt to close the cycle within the pig sector. PALAU will enhance the B2C consumer app by adding new features, incorporating more specific and unconventional datasets, and exploring rewards systems. Furthermore,					

		PALAU will investigate AI decision tools to support sustainable product sourcing and processing methods. GS1 will focus on establishing standards for the dashboards and explore synchronizing PALAU with their database, potentially allowing new information to be pushed directly to consumers. ILVO is hosting a robust data-sharing platform, aiming to connect a wider range of diverse data sources to enhance the					

		platform's operations. They will also provide the API-based climate scan application, Klimrek, which supports climate impact assessments and decision-making. Additionally, ILVO is exploring the integration of more unconventional datasets in the dashboards and to expand the scope of data analysis and improve insights into climate-related social and health challenges.					
	Are there any risks in the data chain process that need to be mitigated?						

Data management							
	Who owns the data?	farmers give consent if the data can be shared (LCA scans)	farmers give consent if the data can be shared (LCA scans)			retailer	data in applications only available on aggregated level
	Who has access to the data?	farmers	farmers			retailer	Palau
	What specific data elements need to be traced through the system?	farm ID	Farm ID	product identification number	product identification number	Purchase transactions	
	Which identification standards will be used to ensure traceability across the supply chain?					looking into working with GS1 standards	

UC4 input

Supply Chain process		Input	Process	Output
		Retail	Consumer	R&D
	What are the activities of the supply chain involved in your UC?	Loyalty card data input	Consumption data	Consumer preferences and motivations
	Who are the the main types of suppliers, service providers involved in supply chain process of the use case?	Retailer	Consumers	Reseacher Institutions
	Who are the key process actors (partners) involved in supply chain process of the use case? Name the partners and other essential organisations.	TAU university	Wageninngen University and Research (WUR)	TAU and WUR
Supply chain strategy and governance				
	What roles do each of these partners play in this supply chain?	Consumer preferences and loyalty card data provided by TAU university. Consumption database provided by WUR	Data flow and model designed by TAU and WUR	Large language models, diffusion models, Personalized communication
	Who are the decision makers involved in each stage of this supply chain?	Retail management	Consumers	WUR, TAU
	Which data-driven innovation do you want to work on?	Insight into consumption and communication preferences to allign with more personalised	More sustainable or less unsustainable intentions and choices	Data Analysis: Utilizing historical loyalty card data to uncover insights into consumer

		choices/recommendations.		behavior related to food purchases.
				AI-Powered Personalization: Building AI models to generate personalized, sustainability-focused dietary recommendations.
				Behavioral Impact: Testing the effectiveness of these recommendations in driving shifts toward more sustainable consumption patterns.
	What are the critical events necessary for quality, compliance, or traceability at each stage of the supply chain ?	Data Collection (Retailer):	Recommendation Deployment:	Data Transfer (Retailer to Project Team):
		Critical Event: Ensure data anonymization and compliance with GDPR and other data protection regulations.	Critical Event: Test recommendations with a pilot group to ensure they align with sustainability goals and user expectations.Impact Assessment	Critical Event: Secure data transfer using encrypted channels. Data Validation (Project Team):

			(Post-Recommendation):	
			Critical Event: Evaluate the behavioral change and environmental impact of the personalized recommendations.	Critical Event: Perform data quality checks (e.g., completeness, accuracy, and consistency).AI Model Development:
				Critical Event: Validate model outputs against ethical AI principles and fairness standards.
				Purpose: Ensure the personalized recommendations do not introduce biases or ethical concerns.
	What is required to verify that each critical event meets regulatory or sustainability standards?	Data Collection (Retailer): Review of legal agreements and data handling policies.	Recommendation Deployment: Analyze pilot data and ensure recommendations comply with project-defined sustainability metrics. Impact Assessment (Post-	Data Transfer (Retailer to Project Team): Validate integrity and security of the data transfer using technical audits. AI Model Development: Peer review of model
		Ethical commission will be obtained from university ethical commission		

			Recommendation): Review and validate impact assessment methodology and findings	outputs and compliance audit of development processes. Data Validation (Project Team): Audit of data cleaning processes and validation scripts to ensure the dataset meets project standards.
Data chain process				
Data capture				
	Which data is available in each part of chain?	Loyalty card data, demographics, consumer preferences	Recommendation data, Consumer choice data	Machine learning model output data
	What type of data is collected in each part of chain?	Private and structured data (loyalty card data, demographics, preferences)	Public and unstructured data (recommendation system, behavioural change techniques) and private and structured data (consumer choice data)	Semi-structured data and personalised images (machine learning model output)
	How is the data collected?	Retail app	Survey, experiment	Machine learning model

	What is the estimated volume of data collected?	Over 35000 loyalty card holders data	>500 consumers	
	What data standards are used(e.g., GS1, ISO, proprietary standards)?	GDPR for personal data, dataset in CSV format		
	Does the dataset contain data considered personal in the GDPR?	Yes		
	Does the dataset contain data considered sensitive in the GDPR?	No		
	How often is the data updated?	Consumer app: the dataset is exported at a certain moment		
Data processing				
	Is there data processing? If yes, how?	Yes: data cleaning on inclusion criteria		
	What data needs to be standarized? To which format?	Loyalty card data is semi-structured for e.g. time-series analysis		
Data analytics				
	Is there data analysis? If yes, how?	Rstudio, SPSS: data loading, cleaning and merging. Possible time-series analysis, dimension reduction techniques and regression/ANOVA analysis.	Data modelling, Rstudio, SPSS: data loading, cleaning and merging. Possible time-series analysis, dimension reduction techniques and regression/ANOVA analysis.	Applying large language models to data model

Data sharing				
	What data are(can be/willing to be) shared in the process? Under which conditions?	Purchase data and demographic data based on eliciting consent. Futhermore, preference/motivation and consumption data.		
	What data are not (cannot be/not willing to be) shared in the process? Why?	Sensitive personal data that can be used to trace individuals. Data is received pseudonymised.		
	What systems/technology is used to share the data ?	Retail app but this is not a continues process		
	Are there third-party stakeholders involved in data sharing who are not part of the project?	Yes, the Finish retailer that shares the data with TAU university. Retailer is the largest in Finland.		
Data infrastructure				
Platforms and tools				
	What systems/technology/platforms/tools/softwares is (could) used for data identification?	COMFOCUS EU, NEVO online		
	What systems/technology/platforms/tools/softwares is used for data capture?	Survey, existing data sets (e.g. consumption data, LCA data) , retail app used by consumers.		
	What technologies/platforms/tools/softwares are used to ensure the accuracy of data flow?	Not applicable		

Data storage				
	What platforms and tools currently in use or in development to store data in the use case?			Internal project servers or a secure cloud platform. Possible, machine learning will be used along with diffusion models and large language models.
	Where will data (preferably) be stored if shared with other parties?			In a GDPR-compliant cloud storage platform.
	What solution(s) are you currently using or planning to use for storing your data (e.g., internal server, cloud, third-party platform)?			A GDPR-compliant cloud storage platform.
	What are your specific requirements or needs for data storage (e.g., secure storage for personal data, real-time data analysis, GDPR compliance)?			Secure storage for personal data, compliance with GDPR, and support for large-scale analytics.
Data Strategy&governance				
Value proposition				

	How will the collaboration take place, what action does every partner do?			**TAU:**
				Analyze the provided loyalty card data, applying AI techniques and behavioral science methods.
				Lead on the integration of cognitive models with food consumption data to generate personalized interventions.
				Develop and test the AI pipelines for behavior change.
				WUR:

				Focus on designing AI-powered behavioral intervention frameworks and communication strategies.
				Lead the ethical and data governance aspects, ensuring compliance with GDPR and ethical standards.
				Manage the practical application of findings, including pilot testing and reporting.
	Are there any risks in the data chain process that need to be mitigated?			Data breaches, misalignment on data formats, non-compliance with GDPR.
Data management				

	Who owns the data?	Loyalty card data: finish retailer/TAU university	Consumption data: WUR.	Research institutions
	Who has access to the data?	Loyalty card data: finish retailer/TAU university	Consumption data: WUR.	Research institutions
	What specific data elements need to be traced through the system?			Purchase transactions, demographic profiles, and AI-generated recommendations.
	Which identification standards will be used to ensure traceability across the supply chain?			Not applicable