D3.2 Data Requirements and Interoperability - Final version

quantifarm.eu

Acronym	QuantiFarm
Full title	Assessing the impact of digital technology solutions in agriculture in real-life conditions
G.A. No.	101059700
Topic	HORIZON-CL6-2021-GOVERNANCE-01
Type of action	Research and Innovation Actions
Start date of project	1st July 2022
Duration	45 months
Author(s)/Organisation(s)	Jack Verhoosel/TNO, Nikos Kalatzis/NP, Daan di Scala/TNO, David de Best/TNO, Christopher Brewster/TNO
Contributor(s)	All partners
Work Package	WP3 – Digital Tools
Delivery Date (DoA)	31/12/2024 (M24)
Actual Delivery Date	31/12/2024 (M24)
Dissemination Level	PU
Abstract:	This deliverable describes the background, requirements and design of the QuantiFarm Common Semantic Model (QCSM) and the QuantiFarm Data Platform which are used to collect DATS generic and usage information to be fed into the various tools of the QuantiFarm Toolkit. This is the updated and final version of this deliverable that includes all the required updates derived from the use of the QuantiFarm Toolkit.

Document Revision History			
Date	**Version**	**Author/Contributor/ Reviewer**	**Summary of main changes**
09/09/2024	v0.5	Jack Verhoosel, Nikos Kalatzis, Daan di Scala, David de Best, Christopher Brewster	First version based on D3.1 plus structure of new sections to be filled
05/10/2024	v0.6	Jack Verhoosel, Daan di Scala	Updates of the QCSM integrated.
08/12/2024	v0.8	Nikos Kalatzis, Christopher Brewster, Jack Verhoosel	Updates to main sections, addition of introduction and conclusions.
20/12/2024	v0.9	Dionysios Solomos, Tania Demirtzioglou, Alessandro Iachetti	Internal review
24/12/2024	v1.0	Jack Verhoosel	Final version

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

QuantiFarm Consortium			
Participant Number	**Participant organisation name**	**Short name**	**Country**
1	GAIA EPICHEIREIN ANONYMI ETAIREIA PSIFIAKON YPIRESION (GAIA)	GAIA	GR
2	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO	TNO	NL
3	POLITECNICO DI MILANO	POLIMI	IT
4	NEUROPUBLIC AE PLIROFORIKIS & EPIKOINONION	NP	GR
5	CONSULAI, CONSULTORIA AGROINDUSTRIAL LDA	CONSULAI	PT
6	CONFEDERAZIONE GENERALE DELL AGRICOLTURA ITALIANA	CONFAGRICOLTURA	IT
7	FOODSCALE HUB GREECE ASSOCIATION FOR ENTREPREUNERSHIP AND INNOVATION ASTIKI MI KERDOSKOPIKI ETAIREIA	FSH	GR
8	PETERSON PROJECTS BV	PETERSON	NL
9	LUONNONVARAKESKUS	LUKE	FI
10	GEOPONIKO PANEPISTIMION ATHINON	AUA	GR
11	OKYS LTD	OKYS	BG
12	COMITE DES ORGANISATIONS PROFESSIONNELLES AGRICOLE DE L'UNION EUROPEENNE COPA ASSOCIATION DE FAIT	COPACOGECA	BE
13	COMITE EUROPEEN DES GROUPEMENTS DE CONSTRUCTEURS DU MACHINISME AGRICOLE	CEMA	BE
14	. TEAGASC - AGRICULTURE AND FOOD DEVELOPMENT AUTHORITY	TEAGASC	IE
15	INSTITUTO TECNOLOGICO AGRARIO DE CASTILLA Y LEON	ITACyL	ES
16	HORTA SRL	HORTA	IT
17	KATHOLIEKE UNIVERSITEIT LEUVEN	KUL	BE
18	DELPHY BV	DELPHY	NL
19	INSTITUT DE L'ELEVAGE	IDELE	FR
20	AUGMENTA AGRICULTURE TECHNOLOGIES MONOPROSOPI IDIOTIKI KEFALAIOUCHIKI ETAIREIA	AUGMENTA	GR
21	ASOCIATIA NATIONALA A INDUSTRIILORDE MORARIT SI PANIFICATIE DIN ROMANIA	ANAMOB	RO
22	UAB ART21	ART21	LT
23	AGROSMART SIA	AgroSmart	LV
24	BENCO BALTIC DOO ZA SAVJETOVANJE IUSLUGE	BENCO	HR
25	FARM FRITES POLAND DWA SPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA	FFP2	PL
26	AGROMAIS PLUS COMERCIO E SERVICOSAGRICOLAS S.A.	AGROMAIS	PT
27	KMETIJSKO GOZDARSKA ZBORNICA SLOVENIJE KMETIJSKO GOZDARSKI ZAVOD MURSKA SOBOT	KGZS	SI
28	TERRA LITTERA DO0	Terra	RS
29	ANYSOLUTION SL	AnySol	ES
30	A.M. FILAGROTIKI SYMVOULEFTIKI LTD	Filagro	CY
31	AGRIDEA SCHWEIZERISCHE VEREINIGUNG	AGRIDEA	CH
32	FLOX limited	FLOX	UK

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

LEGAL NOTICE

Funded by the European Union under Grant Agreement No 101059700. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Funded by
the European Union

Table of Contents

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

List of Figures

List of Tables

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

LIST OF ABBREVIATIONS AND ACRONYMS	
API	Application Programming Interface
AKIS	Agriculture Knowledge and Innovation Systems
CAP	Common Agricultural Policy
DATs	Digital Agriculture Technologies
DB	Database
DIH	Digital Innovation Hub
DMP	Data Management Plan
EDIH	European Digital Innovation Hubs
EIP-AGRI	European Innovation Partnership for Agricultural productivity and Sustainability
EC	European Commission
EO	Earth Observation
EU	European Union
EUPL	European Union Public Licence
FMIS	Farm Management Information System
GDPR	General Data Protection Regulation
ICT	Information and Communications Technology
IPR	Intellectual Property Rights
JRC	Joint Research Centre
QA	Quality Assurance
QC	Quality Control
RDF	Resource Description Framework
TC	Test Case
USGS	United States Geological Survey
VRA	Variable Rate Application
WP	Work Package

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

1 Introduction

1.1 Project Summary

The QuantiFarm project focuses on supporting the further development of Digital Agriculture Technology Solutions (DATSs) in order to improve the sustainability (economic, environmental and social) and competitiveness of the EU's agricultural sector. To this end, QuantiFarm develops a comprehensive Assessment Framework for independent qualitative and quantitative assessments of the costs and benefits of digital agriculture technologies. The project will support the uptake of digital technologies by providing innovative tools and services, which will provide practical recommendations of relevance and practical utility to farmers, advisors, and policy makers across Europe. QuantiFarm is building its assessment and recommendation tools using data derived from 30 Test Cases (TCs) which span over 20 countries in 10 Biogeographical regions across Europe, thereby capturing multiple social, environmental and economic parameters. More than 100 farms of different types, sizes, ownership and operating conditions, have committed to participate in the project, both directly and through cooperatives and large umbrella organisations. In addition, the QuantiFarm Digital Innovation Academy will provide capacity building capabilities for advisors and other Agriculture Knowledge and Innovation Systems (AKIS) actors, by providing data on the various types of digital technologies available, their costs, benefits and impact on sustainability. QuantiFarm involves 32 partners, representing a variety of stakeholders, including 8 scientific organisations and 12 farmer representatives and consultants.

1.2 Document Scope

This document is part of the set of deliverables of WP3 of the QuantiFarm project and is the result of T3.1 on Data Requirements and Interoperability. This task has identified the relevant data sources for the assessment of DATSs and worked on a data platform that allows to clean, integrate, link, harmonise, store and make available data coming from various heterogeneous data sources. Data formats have been standardised with the addition of suitable metadata, using semantic interoperability mappers and reusing and extending existing relevant ontologies.

The goal of this document is to describe the results of T3.1 activities, specifically focusing on an update of the QuantiFarm Common Semantic Model (QCSM) and the QuantiFarm Data Platform that have been developed. Both the model and the platform are defined to collect data from various sources, including the 30 test cases of the project and other open data sources that are important for supporting farmers, advisors and policy makers about the use of Digital Agricultural Technologies. The Data Platform will provide collected data from all the sources in QuantiFarm Common Semantic Model format to the QuantiFarm Toolkit developed in T3.2 and T3.3 to be used by farmers, advisors and policy makers.

The initial version of this document is entitled "D3.1 Data Requirements and Interoperability" and was created in June 2023 (M12). This is the second version of this document and contains all the realised updates that were implemented during these months aiming to address data interoperability needs and information richness.

The main updates that this document includes are:

- The introduction of additional parameters within the QCSM addressing data modelling needs of "Farms", "DATS" and "DATS Assessment parameters". The "DATS Assessment parameters" are derived due to the implemented updates of the QuantiFarm Assessment Framework of WP2 (section 3.1).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

- Updates of the QCSM for addressing the needs of the updated versions of the QuantiFarm Toolkit's services. (section 3.3)
- Updates of the QCSM addressing the needs fro improved descriptions of the DATSs managed/hosted by the QuantiFarm Toolkit (section 4.2).
- Update of the assessment concepts and properties in the QCSM (section 4.3)
- Update of the QuantiFarm data platform (part of the QuantiFarm Toolkit) based on the introduced functionalities (section 5.2)

The outcomes of this deliverable are accompanied with the actual specification of the QCSM which is available in QuantiFarm project code repository. The most up to date version of the ontology is available here: https://gitlab.com/QuantiFarm/qcsm/-/blob/main/qcsm.ttl

Even though the QuantiFarm activities on data modelling are officially completed in case there is a need of further updates these will be realised by updating the QCSM: https://gitlab.com/QuantiFarm/.

1.3 Document Structure

This report is structured as follows.

- Chapter 1 provides a summary of the project, the document's scope and its overall structure.
- Chapter 2 introduces the current landscape and background on agricultural interoperability standards and ontologies and existing data reference architectures to be reused.
- Chapter 3 discusses the goal of the QuantiFarm data platform and the requirements from various sources on this platform and the common semantic model that is to be used by the data platform.
- Chapter 4 presents the QuantiFarm Common Semantic Model (QCSM), the guidelines that have been used to define it, the existing ontologies that are being reused and examples of how DATS, sustainability and behavioural parameters are being modelled.
- Chapter 5 describes the high-level design of the QuantiFarm data platform, and its components based on use case definitions it needs to satisfy.
- Chapter 6 gives a brief overview of the main ingredients of the document and indicates next steps to be taken.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

2 State of the art analysis on relevant tech sectors

This chapter briefly describes the basic semantic technologies that are being used in the QuantiFarm Common Semantic Model and the QuantiFarm Data Platform. In addition, a brief summary is given of the existing agrifood interoperability standards and ontologies. Finally, a short overview is presented of existing data management architectures.

2.1 Background

In the last decades various methodologies for describing information in a semantically unambiguous way have been defined. The World Wide Web consortium (W3C) has been very active in this field, which has led to a number of standard mechanisms to define semantics, i.e., the Resource Description Framework (RDF), RDF Schema (RDFS) and the Web Ontology Language (OWL). See Figure 1for the various semantic standards and how they relate.

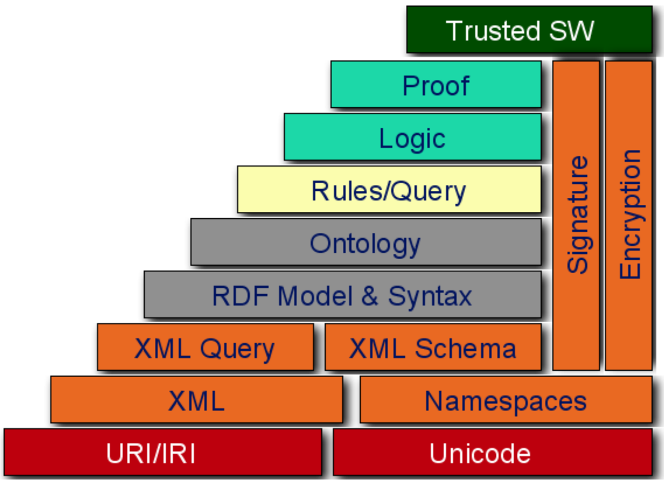

Figure 1: Semantic standards and their relationships

The Resource Description Framework (RDF) is a World Wide Web Consortium (W3C) standard originally designed as a data model for metadata. It has come to be used as a general method for description and exchange of graph data. RDF provides a variety of syntax notations and data serialisation formats, with Turtle (Terse RDF Triple Language) currently being the most widely used notation. RDF is a directed graph composed of triple statements. An RDF graph statement is represented by: 1) a node for the subject, 2) an arc that goes from a subject to an object for the predicate, and 3) a node for the object. Each of the three parts of the statement can be identified by a URI. An object can also be a literal value. This simple, flexible data model has a lot of expressive power to represent complex situations and relationships, while also being appropriately abstract. RDF was adopted as a W3C recommendation in 1999. The RDF 1.0 specification was published in 2004, the RDF 1.1 specification in 2014. SPARQL is a standard query language for RDF graphs. RDFS, OWL and SHACL are ontology languages that are used to describe RDF data.

The RDF Schema (Resource Description Framework Schema, abbreviated as RDFS) is a set of classes with certain properties using the RDF extensible knowledge representation data model, providing basic elements for the description of ontologies. It uses various forms of RDF vocabularies, intended to structure RDF resources. RDF and RDFS can be saved in a triplestore, then one can extract some knowledge from them using a query language, like SPARQL. The first version [1,2] was published by the World-Wide Web Consortium (W3C) in April 1998, and the final W3C recommendation was

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

released in February 2014 [3]. Many RDFS components are included in the more expressive Web Ontology Language (OWL).

The Web Ontology Language (OWL) is a family of knowledge representation languages for authoring ontologies. Ontologies are a formal way to describe taxonomies and classification networks, essentially defining the structure of knowledge for various domains: the nouns representing classes of objects and the verbs representing relations between the objects. The OWL languages are characterised by formal semantics and a first version was published in 2004. They are built using the Resource Description Framework (RDF). W3C announced the version 1.1 of OWL on 27 October 2009 [5]. This version, called OWL2, soon found its way into semantic editors such as Protégé and semantic reasoners such as Pellet, RacerPro, FaCT++ and HermiT. The OWL family contains many species, serialisations, syntaxes and specifications with similar names. OWL and OWL2 are used to refer to the 2004 and 2009 specifications, respectively.

2.2 Interoperability and data standards

There is a considerable body of work building ontologies for the food and agriculture domain which has gone hand in hand with the development of Linked Data (and "Linked Open Data") in the agri-food domain. The major effort here has been AGROVOC developed by the FAO and maintained by a network of institutes around the world [6]. It is nowadays the most comprehensive multilingual thesaurus and vocabulary for agriculture. AGROVOC has now been partially mapped onto the US National Agricultural Library of the USDA and the CABI thesaurus in the form of the GACS ontology which has mapped and integrated the top 15,000 concepts [7]. Other recent work in this area has also focused on developing ontologies for sharing of research data including the Crop ontology initiative, the Agronomy Ontology (AgrO), and the Plant Trait Ontology (TO) supported by CGIAR [8]. FOODON integrates a number of existing ontologies, but its focus seems to be again on research data although its ambition is to provide a mechanism for data integration across the food system. Considerable efforts have been put into extending and integrating the FOODON ontology with various other ontologies extending its utility to areas such as nutrition, and integrating it with the Foodex2 standard from EFSA [9, 10] As mentioned, most work on ontologies for the agrifood domain has up to now mostly been targeted towards the clear definition of domain concepts and terms in the form of a vocabulary for the annotation of research publications or research data sets [8]. As a result, there is little or no use yet of ontologies for supporting the actual sharing and exchange of data across the agrifood chain. Only a few papers about the use of ontologies for traceability and data analysis in the dairy sector [11, 12]. In addition, a few innovation projects have dealt with the use of ontologies in the horticultural supply chain and greenhouse domain [13]. However, the use of semantic standards for information exchange by standardisation organisations, like GS1, ISO or AgGateway is not yet common practice or even beginning to be picked up. More generally under the auspices of the EC funded projects like ATLAS (https://www.atlas-h2020.eu/) and especially DEMETER (https://h2020-demeter.eu/), there is support for greater use of semantic standards and technologies where the data and data models are explicitly specified, where URIs are widely used, and where data integration is consequently made far easier [14].

2.3 Architectures for data management, sharing, querying

The use of semantics for data sharing or data exchange in general has a long history epitomised by the Linked Data initiative and more recently the FAIR data movement. There has been a tendency for this to be largely a focus for researchers rather than commercial applications with the exception of the life sciences. Within the food and agriculture sector, the majority of efforts at data sharing have focussed on ontologies for research data and assumed relatively centralised approaches to repositories (e.g. the CGIAR platform for Big Data [8]), or else have been based on XML. Three major data standardisation efforts in the agri-food areas are ICAR (for livestock and dairy production), ISOBUS (for machine-to-machine data sharing) and AgGateway (for FMIS applications). These are currently available as XML standards (with codebooks) but are slowly moving towards JSON-LD versions of their standards. [15] make the case for using Linked Data principles and a variety of ontologies so that data can be integrated for farmer decision support. Other literature emphasises the importance of semantics for integrating IoT derived data in agriculture, while [16] similarly uses ontologies to ensure data integration for supply chain data. These papers, as do most others, assume centralised architectures. The Linked Pedigree architecture was proposed to enable data sharing across supply chains by formalising the GS1 EPCIS standards as ontologies and enabling SPARQL queries across distributed triple stores, and this work partly inspires the technological approach described below [17, 18].

2.3.1 Interoperability and legacy systems

The issue of making existing legacy systems interoperable is not new and has been researched as part of the data integration topic [19]. The goal of data integration is focused largely on the need for answering user queries over distributed relational data sources and using SQL as the query language. The data integration landscape has changed with the increase in (types of) data sources and applications that can benefit from it. Other approaches that use semantic technology propose query-rewriting of the overall query to integrate distributed heterogeneous data sources [20]. This is different from our approach in that the data sources are integrated using semantic domain knowledge instead of a user query. Our focus lies on composing web service APIs of the distributed legacy data sources. Applying semantic technologies to web services has been the topic of extensive research. Two topics can be distinguished, although they are often investigated together.

On the one hand, research has focused on giving semantic descriptions of the inputs/outputs and pre/post conditions of web services. to be able to automatically discover or select relevant services [21]. On the other hand, research focused on how to automatically compose multiple web services together [22]. Despite the many proposals to include semantics into the descriptions of web services [23–25], many of today’s web service APIs do not describe their formal invocation, pre- and post-conditions, and semantic input/output. The research on semantic web service composition in general has investigated several techniques to achieve this, and the approach is similar to rule-based ‘planning’ [26–28]. However, in contrast to these approaches, our approach uses graph patterns of a common semantic model to be able to automatically compose distributed heterogeneous data sources.

In the area of Internet of Things, with the growth of IoT devices efforts have been devoted to overcoming interoperability challenges [29, 30]. In this context the W3C has launched the Web of Things Working Group (https://www.w3.org/WoT/) to counter the fragmentation arising from the use of IoT devices and support the use of semantic standards. This work largely builds on the Semantic Sensor Network (SSN) ontology (https://www.w3.org/TR/vocab-ssn/). The Open Mobile Alliance and a variety of telecommunication standards bodies have supported the Next Generation Service Interface (NGSI) [31]. The NGSI API and the associated NGSI context model has been adopted by key organisations in the IoT standardisation efforts.

In the work described below, we have used extensively the concepts and principles of semantic web technologies. This enables us to achieve a number of objectives: a) using semantics for the data model allows flexibility as the full range of parameters to be captured in assessment model; b) it ensure flexibility and adaptability when interacting with external data sources; c) it allows for easy construction of complex queries which may not be foreseeable at data model design stage; d) by representing all data collected in a knowledge graph, we can use a growing set of standardised tools for mining the knowledge graph.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101059700

3 Technical requirements for the QuantiFarm data platform

This section specifies the technical and data modelling requirements that will guide the implementation of the QuantiFarm data platform. Prior to the specification of these requirements, it is necessary to identify the involved data sources, the data properties that need to be modelled and the data needs of the QuantiFarm tools to be developed. Figure 2 illustrates in a conceptual and high-level manner the overall flow of data starting from the "Data sources" to be integrated, the "Assessment Framework" that will perform the Economic, social and Environmental assessment of the DATSs and finally the set of "Tools" that will consume these data aiming to provide more advanced decision-making services. As it is evident, the heterogeneity of data derived by different sources needs to be harmonised in order to ensure the maximal reuse and applicability of the QuantiFarm Assessment Framework and Tools. For this reason the QCSM is introduced that will be utilised for enforcing interoperability at a semantic level.

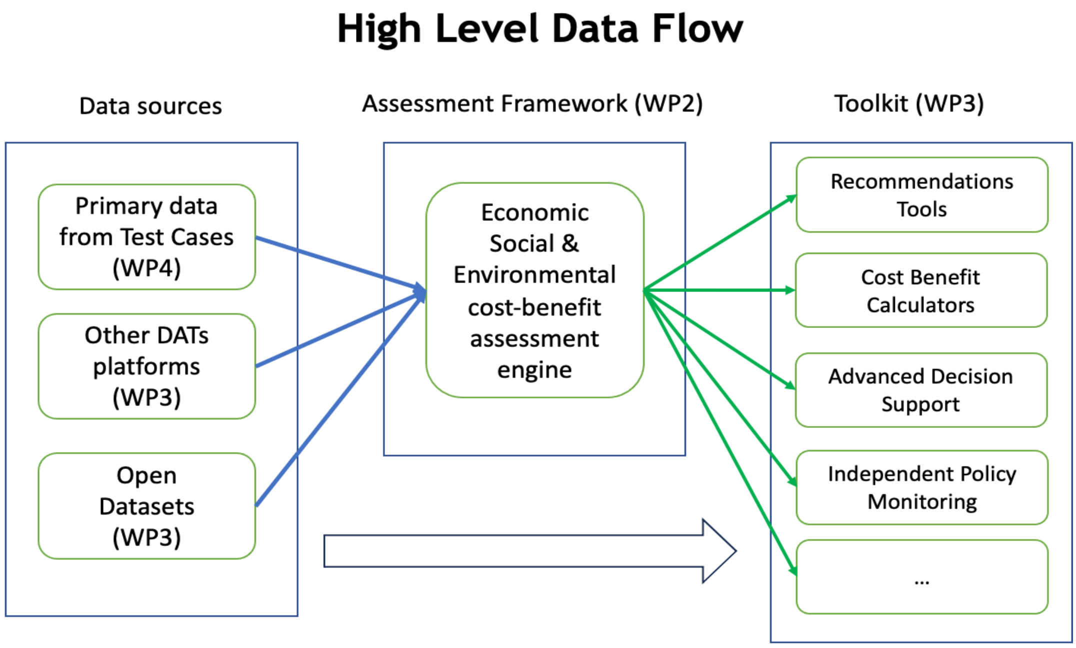

Figure 2: High level data flow from sources via assessment framework to toolkit

3.1 Data provision by Test Cases and other platforms

3.1.1 Test cases

The project will collect data from 30 Test Cases (TCs) across multiple countries and biogeographical regions. Full details of the TCs are provided in "D4.1 Testing and Assessment Guidelines". The TCs will enable the evaluation of different DATSs in real life conditions, across 3 iterations/growing seasons. Each TC includes one or more farms using a specific DATS, and a set of equivalent and comparable farms not using any DATS. Care is being taken to find 1-to-1 matching so as to ensure consistent, contrasting yet comparable participants in each TC. An overview is provided in Table 1.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

No	Sector	Crop/Animal	Digital Solution	Country	Biogeographical Region
1	Arable	Potatoes	SF DSS/ App	Greece	Mediterranean
2	Arable	Corn	Precision irrigation	Portugal	Mediterranean
3	Arable	Wheat	DSS/Agri-environmental Monitoring	Spain	Mediterranean
4	Arable	Catton	VRA add-on for old tractors	Greece	Mediterranean
5	Arable	Wheat	SF DSS/ App	Turkey	Anatolian
6	Arable	Wheat, onion, potato	Machinery with VRA, data analytics	Netherlands	Continental
7	Arable	Potatoes	SF DSS/ App	Poland	Continental
8	Arable	Wheat, Rapeseed, Rye, Barley	Drones, sensors, silo management, AI	Latvia	Boreal
9	Arable	Barley, Corn, Wheat	FMIS/ Financial Modelling	Slovenia	Alpine
10	Arable	Wheat	FMIS/ app	Romania	Steppe
11	Fruit	Olives	SF DSS/ App	Greece	Mediterranean
12	Fruit	Apples	Drones and soil sensors	Poland	Continental
13	Fruit	Grapes	SF DSS/ App	Italy	Mediterranean
14	Fruit	Strawberry, Blueberry	Precision Irrigation / Variable root pruning	Serbia	Pannonian
15	Fruit	Olives	SF DSS/ App	Cyprus	Mediterranean
16	Fruit	Apples	Drones and soil sensors	Netherlands	Continental
17	Fruit	Grapes	Harvesting robotic and SF DSS	Romania	Black Sea
18	Vegetables	Tomatoes	SF DSS/ App	Italy	Mediterranean
19	Vegetables	Tomatoes	Automated greenhouses	Netherlands	Continental
20	Fruit	Bananas	Precision Irrigation, Monitoring	Spain	Macaronesian
21	Vegetables	Tomatoes	Automated greenhouses	Finland	Boreal
22	Meat	Poultry	Cleaning robot, AI	UK	Atlantic
23	Meat	Cows	Heat box collar	France	Continental
24	Meat	Pigs	Automated monitoring, AI	Belgium	Continental
25	Dairy	Cows	Feeding robot	France	Continental
26	Dairy	Cows	Milking robot	Ireland	Atlantic
27	Dairy	Cows	Automated monitoring	Germany	Continental
28	Dairy	Cows	Livestock feeding DSS	Romania	Steppe
29	Apiculture	Bees	Automated monitoring	Lithuania	Boreal
30	Aquaculture	Oyster	Sensors for quality assessment	Croatia	Mediterranean

Table 1: Test Cases summarization

TCs will act as data providers for the Assessment Framework, collecting data for the assessment of the digital agriculture solutions they are using on their farms, and equivalently to assess the impact of not using DATSs. To properly assess the impact of using a digital technology in agriculture in real conditions, the assessment needs to take into consideration the outcomes of the whole growing season. Given the fact that agricultural production is affected by many unpredictable factors such as extreme weather conditions, which can alter the results drastically, there needs to be enough growing seasons. Therefore, testing has started in January 2023 (M7) and will run for three (3) annual cycles, covering an equal number of growing seasons. The data collected by all 30 TCs will be integrated into the QuantiFarm data platform and toolkit to support the respective assessment and decision support tools for farmers, advisors and policy makers with the necessary input.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

All TCs include farms that are already using specific DATSs. The DATSs to be assessed in the context of the TCs are the DATSs already being used by the farmers. No additional DATSs will be developed or selected during the project. There is direct farmer involvement in all QuantiFarm TCs. The DATS is being applied for a particular part of the farm (a parcel or a group of animals) and its performance is being monitored and compared with a similar part of the farm (or another farm) for which the DATS is not applied. The first cultivation cycle in 2023 of all the TCs have resulted in lessons learned on how to collect which data and how to interpret/process it. Moreover, additional generic and specific assessment parameters have been defined.

Table 2 provides more details on the farm parameters that have been collected per testcase.

FARM PARAMETERS		
Type of output/product	-	Milk
Crop	-	Apple
Animal		
Start of the growing/production season	(dd/mm/yyyy)	
End of the growing/production season	(dd/mm/yyyy)	
Biogeographical region	-	
Total farm size	ha	
Number of comparison(s)	n	
Number of employees	n	
Number of full-time employees	n	
Number of seasonal employees	n	

Table 2: Farm parameters collected per test case

Table 3 gives a list of generic DATS parameters that have been collected per test case.

DATSs PARAMETERS		
Initial investment for DATSs (the total for all the solutions assessed in the project)	€	
Expected life span of each DATSs	y	
Time to install and make DATSs operational	h	
DATs maintenance costs	€/y	
DATS annual fees	€/y	
Training hours for the use of the DATSs in a year	h	
Number of people involved in the training	n	
Cost of training service by external people (advisor, DATS providers, ...)	€/y	
DATSs name	-	

Table 3: Generic DATS parameters collected per test case

Apart from the more static parameters of the farm and DATS, the test cases provide measured data about the parcels being used. For each parcel with or without DATS, a large number of parameters are being monitored or calculated. Table 4 gives a list of the main information of a parcel that is assessed in a test case either with or without DATS.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

Parcel assessment information			
#	data	u.m.	value
1	Parcel dimension	ha	
2	Number of plants/trees	n	
3	Sand content in the soil	%	
4	Clay content in the soil	%	
5	Silt content in the soil	%	
6	Soil Bulk Density	g/cm^3	
7	Soil Organic Carbon	%	
8	Number of animals	n	
9	Livestock units per hectare	LSU/ha	
10	Number of beehives	n	
11	Area under aquaculture	m^2	

Table 4: Main information of a parcel with/without DATS in a test case

In addition, various parameters around usage of water, fuel, electricity, pesticides, fertilisers, feed and drugs and labour are monitored or calculated. Finally, economic and output data about the performance of the parcel is collected as well.

3.1.2 External DATSs platforms

One of the key objectives of the QuantiFarm project is to build on the results and lessons learnt from existing platforms that already maintain a large number of DATSs. To avoid reinventing the wheel, the QuantiFarm Data Platform and Toolkit will be built on top of the databases of the Smart-AKIS and FAIRshare platforms. Although these two platforms maintain the descriptions of a large number of DATSs, access to these is only available through web interfaces. To this end and in the context of the QuantiFarm project both platforms have been extended to facilitate data exchange through an API. This will create a unified, comprehensive data platform that brings together important information and tools from different sources. APIs will allow QuantiFarm to interact with various external platforms and consume their data and functionalities. At the same time, QuantiFarm's own data and functionality can also be exposed and made accessible to these external platforms.

3.1.2.1 Integration with FAIRshare

One of the key platforms that QuantiFarm integrates with is FAIRshare (https://fairshare-pnf.eu/). FAIRshare's mission is to encourage farm advisors to utilise Digital Advisory Tools and Services[1] that contribute to more productive and sustainable agriculture. Figure 3 provides a screenshot of FAIRShare's inventory of DATSs. As it is illustrated, the inventory currently maintains about 300 DATSs and it also provides a set of filters that uses as input user defined preferences in order to refine and personalise the rendered results.

[1] In the context of FAIRshare a DATS is a "Digital Advisory Tool or Service" as opposed to the QuantiFarm definition as "digital technologies in agriculture".

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

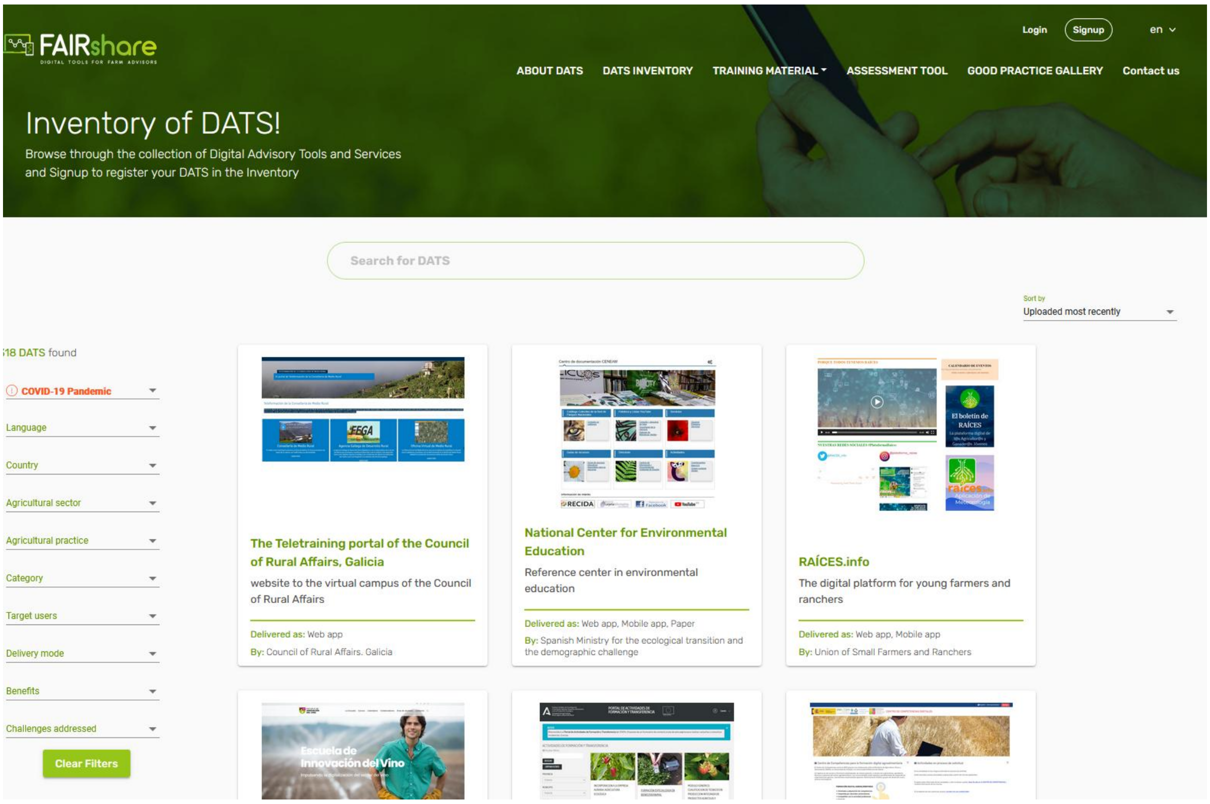

Figure 3: The FAIRshare inventory of DATSs

The FAIRShare inventory provides the necessary functionality through the website for registered users to add, update, and remove DATSs. The platform already had an internal API to allow communication between a front end (in Angular) and the backend (in Node.js and MongoDB).

To enable integration with QuantiFarm, FAIRshare's API underwent an extension of functionalities. The process involved addressing minor pre-existing issues, such as character encoding, and optimising functionalities to prevent critical errors. The result was a more efficient and reliable API that can provide good results based on the keywords provided. The expanded API now includes critical calls like 'searchTools', which retrieves DATS that includes at least one of the provided keywords, and 'Tools', which enables the retrieval of the entire collection of DATS available on the FAIRshare database. To aid in this integration and for better understanding, a detailed documentation of the Fairshare API, presenting technical information and examples of the responses for each of the available requests can be accessed here: https://documenter.getpostman.com/view/26161901/2s93JnW7Sx.

In Annex B the properties used for describing a DATS as present in the FAIRshare system are presented. As it will be described in section "4.2 Modelling digital agricultural technologies", these properties are also incorporated within the QCSM ontology.

3.1.2.2 Integration with Smart-AKIS

Smart-AKIS was another external platform from a European funded project that collected smart farming technologies. It was a free platform providing tools for the dissemination and application of Smart Farming technologies. As illustrated in Figure 4, the Smart-AKIS dashboard provided a platform that allows users to browse, search and assess a broad database of smart farming technologies available on the market and coming from research projects, getting useful information on how they work, their economic and environmental benefits and demonstration materials. Unfortunately, the Smart-AKIS dashboard is at the time of writing this deliverable (end 2024) not up and running anymore.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

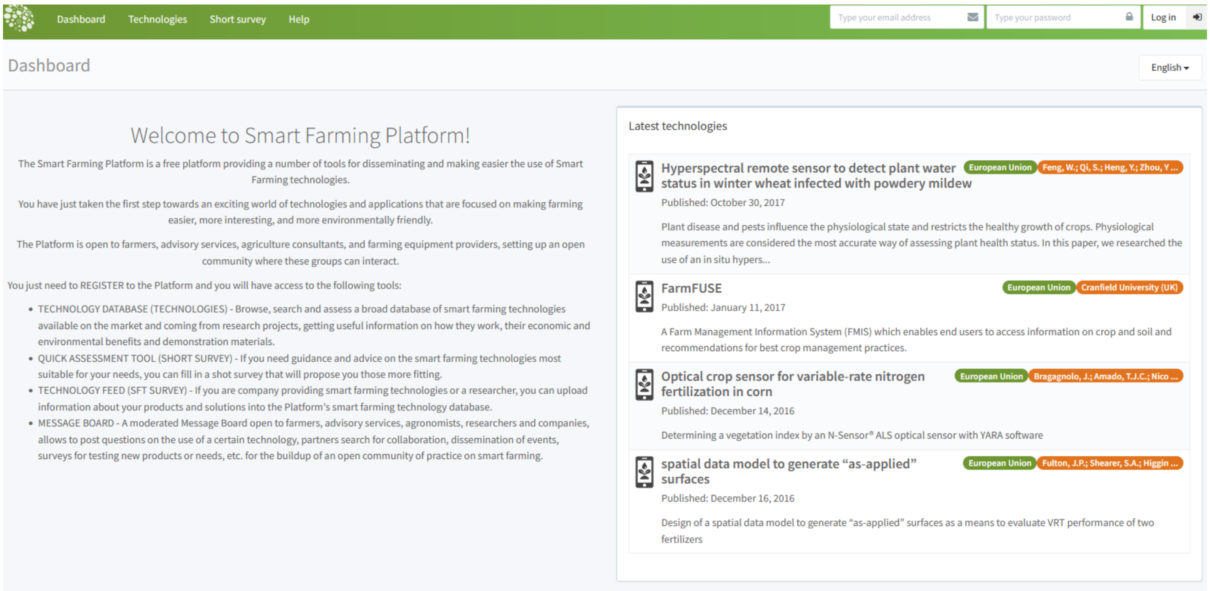

Figure 4: The Smart-AKIS inventory of Smart Farming technologies

Nevertheless, in order to facilitate integration with QuantiFarm's data platform, an API for Smart-AKIS platform was developed using the Laravel framework (https://laravel.com/). The Smart-AKIS API supports queries to the underlying database for retrieving smart farming technologies with the use of filters such as: 'Category options', 'Countries', 'Crop Systems', 'SFT Types', 'SFT Effects', 'Crop Types', and 'Technologies'. These calls enable the retrieval of different subsets of information.

The detailed documentation for the Smart-AKIS API is accessible here:
https://documenter.getpostman.com/view/2626884/2s93m33iXy.

The API is still available and gives a long list of technologies. However, it must be noted that according to our research the registration of new technologies to the platform is limited while the most recent update was in 2019. This presents a significant challenge as the digital agriculture sector is very fast moving.

In conclusion, the integration with external platforms such as FAIRshare and Smart-AKIS has provided a core dataset for QuantiFarm. It extends the range of available data and functionalities, thereby enhancing the value offered to end-users. Such integration efforts also underscore QuantiFarm's commitment to promoting interoperability and seamless data exchange in the agricultural domain. As will be described in the following sections, "Data Loaders" will periodically connect to SmartAKIS and FAIRShare databases in order to retrieve any available updates of digital agricultural technologies and store them within the QuantiFarm data platform.

3.1.3 Open Datasets

In some cases it will be necessary to use datasets that are provided by third parties. For example, in order to capture the context that a DATS is operating in, it will be important in some cases to record weather parameters that have occurred during the cultivation period. In addition and especially for the operation of the Independent Monitoring Policy tool (more on this will be presented in Section 3.3) the following candidate open data sources are being evaluated for integration:

- Copernicus Climate Data store

This dataset provides daily surface meteorological data for the period from 1979 to present as input for agriculture and agro-ecological studies. This dataset is based on the hourly ECMWF ERA5 data at surface level and is referred to as AgERA5. DOI: 10.24381/cds.6c68c9bb

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

- Common Agricultural Policy datasets

Data on national and European agriculture and common agricultural policy (CAP), provided by the European Commission's agricultural and rural development department. Of interest for the QuantiFarm project are dataset and calculated indicators derived by the performance monitoring and evaluation framework (PMEF) assessing CAP 2023-27 national strategic plans. In a similar manner, the data from the Farm Accountancy Data Network (FADN) and its successor scheme Farm Sustainability Data Network (FSDN) will be assessed as to whether it is useful. Context indicators provide information on the agricultural sector and rural areas as well as general economic and environmental trends. Some of this information drills down to regional level (NUTS 2-3).

Other datasets that are likely to be useful and relevant to policy implementation include data sets on the boundaries of environmental sensitive/important regions such as: Natura2000 areas, Wildlife sanctuaries, Waterbodies, Biodiversity indicators, etc.

It should be noted that given the complexity and the large volume of these datasets, we will avoid duplicating such data within the QuantiFarm Data Platform. A more optimal approach will be followed by retrieving and integrating them directly according to the decision-making needs.

3.2 Assessment Framework

Among the main objectives of the QuantiFarm project is to develop an Assessment Framework for evaluating the added value provided by DATSs in agriculture. A mixed-method approach will be used including the farm as well as the wider society for the co-creation of the framework. The Assessment Framework will combine different quantitative and qualitative methodologies to conduct a comprehensive assessment of the economic, social and environmental benefits and costs of DATSs. The current specification of the Assessment Framework is documented in D2.2 “Assessment Framework and Governance Mechanisms – first updated version” of January 2024. According to these specifications, the outcome of the assessment process will be a composite cost-benefit monetary index which will be complemented with a set of additional descriptive indicators on the impact of DATSs. These are intended to avoid the risk of drawing simplistic conclusions when relying only on a single value to capture the complexity of multi-dimensional aspects

The parameters that are being used to measure the performance of DATSs are defined in D2.2. The Assessment Framework also defines sustainability categories and indicators on environmental, economic and social aspects. Each of the test cases will report repeatedly on all of these categories and indicators, in addition to some general information about the DATS. Seven categories have been identified on environmental aspects, four on economic aspects and two on social aspects. Various subcategories of performance measurement have been identified, for which there are a large set of indicators (48 environmental, 16 economic and 16 social). The final set of indicators to be included in the assessment framework has taken into account the particularities of the test cases in terms of crop/product, geographical location, harvesting season and DATS used, if any. The various cost-benefit and sustainability aspects, categories and indicators of the Assessment Framework are summarised in Table 3.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101059700

COST-BENEFIT ANALYSIS

DATSs COSTS

- **INITIAL INVESTMENT**
  - DATS cost (hardware, software)
  - set-up
  - initial training
- **RUNNING COSTS**
  - annual fee
  - periodical training
  - maintenance

BENEFITS

COSTS SAVINGS

- technical input (seeds, fertilizers, pesticides, drugs, feed, ...)
- water
- fuel
- utilities (energy, gas, ...)
- labour
- other direct costs of activities impacted by DATSs

INCREASE OF REVENUES

- increase of sales (> yield)
- higher price (> quality)
- additional revenue (e.g. data valorisation)

NET BENEFIT

PBP, NPV, ROI

SUSTAINABILITY IMPACTS

ENVIRONMENTAL

- Total greenhouse gas (GHG) emissions
- Water Consumption
- Water productivity
- Fuel oil/diesel/propane consumption
- Gas consumption
- Wood consumption
- Electricity consumption
- Share of renewable energy
- Nitrogen use & efficiency
- Phosphorus use & efficiency
- Potassium use & efficiency
- Herbicides use
- Insecticides use
- Fungicides use
- Animal welfare: ease of movement
- Animal welfare: total indoor area
- Animal health: Mortality rate
- Animal health: Mortality rate at birth
- Animal health: Cows with high SCC
- Animal health: Quantity of drugs used

ECONOMIC

- Input productivity (land, labour, milk, meat, bees, oyster)
- Feed conversion ratio
- Efficiency indicators (Rate of time (to complete an activity, precision & accuracy,...)
- Intrinsic product quality (e.g. BRIX, Protein %, ...)

SOCIAL

- Training hours (for the use of DATSs)
- Working time
- Frequency rate of rates of occupational injuries
- Incidence of occupational injuries
- Contribution to local employment

+ a set of additional descriptive / qualitative indicators referring to social impact of DATSs

Table 5: QuantiFarm *Assessment Framework with cost-benefit and sustainability aspects, categories and indicators*

The Assessment Framework and the respective algorithms that calculate the indicators reported in Table 3 is implemented as a reusable software component that supports the dynamic provisioning of different input datasets referring to different DATSs and respective contexts. The dataset to be utilised as input and the calculated performance indicators of each category have been further detailed into measurable parameters that are being modelled in our QuantiFarm Common Semantic Model (QCSM). This is described in more detail in Chapter 4.

3.3 Data needs for the QuantiFarm Tools

Besides the "Assessment Framework", QuantiFarm will design and develop a set of software tools that will offer a variety of services. These services will be bundled in the "QuantiFarm Toolkit" and will be offered as **interactive and user-friendly** web-based tools. These tools are intended to fulfil the needs of specific user groups, viz. Farmers, Advisors and Policy Makers. The overall objective is to provide access to the assessment results and support decision-making across the various contexts that are analysed within the QuantiFarm's project Test Cases. Through a personalised interface, the Toolkit will present information on the DATSs' costs, benefits and sustainability impact, and thereby support both individual decision making and broader policy design. The detailed specifications and implementation details of the tools will be documented in D3.2 "Tools for DATS Assessment and Policy Monitoring" (to be delivered in December 2023). Nonetheless at this stage it is important to identify core requirements, especially regarding data. These requirements guide the specification of the data modelling properties of the QuantiFarm data model.

a) Cost and benefit calculator

A financial cost and benefit calculator for different sectors and different types of DATS will be available to farmers and advisors. After entering the field size, crop or livestock system and the DATS, the

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

calculator will be able to provide the potential benefits and costs (both running and investment costs) when using a specific DATS under certain conditions, compared to not using this DATS. These include different calculations for DATSs in livestock (e.g. robotic milking, virtual fencing, automatic feeding, livestock grazing, etc), arable crops (e.g. VRA for fertilisers, robotic field operations for seeding, spraying, weeding, etc) and fruits and vineyards (e.g. precision irrigation, precision spraying, selecting harvesting, etc). In addition, the calculator will also calculate the risk in the investment or the increased risk in the production due to climate change adaptation. These calculations are described in "D2.5 Cost and Benefit Calculators Design". According to these specifications the following two tables contain the set of required data types, each focusing in a different application domain.

Application Domain: Crop Farming Systems Calculator		
Calculator's Modules	**Data types required**	
	Provided by the end-user	**Provided by the system**
Investment Cost Calculator	• Years of Usage, • Total Area (ha), • Number of Units (sensors etc.)	• DATS's Initial Cost of Investment (€), • Subscription Cost for 1ha (€), • Monthly Subscription Cost (€), • Annual Subscription Cost (€), • Total Cost of DATS Purchase (€)
Yield Increase Calculator	• Current Yield (tons/ha), • Market Price (€/1 Ton)	• Yield increase (%), • Current Revenue (€ in 1 year), • Increased Yield (tons/ha), • Price of Increased Yield (€/ ha).
Revenue Increase Calculator	• Total Area (ha)	• Price of Increased Yield (€/ha), • Current Revenue (€/year)
Fertiliser Use Calculator	• Current fertiliser usage (Kg/ ha), • Fertiliser Cost (€/Kg)	• Fertilisation saving (%), • Current Fertiliser Cost (€/ 1 Kg), • Current Fertilisation Cost (€/ ha), • Reduced Fertiliser usage (kg/ha), • Fertiliser cost savings (€/ ha), • Fertilisation cost saving (€ in 1 year)
Water Use Calculator	• Current water usage (m3/ha), • Current water cost (€/ m3)	• Water saving (%), • Current Irrigation Cost (€/ ha), • Reduced water usage (m3/ha), • Water cost savings (€/ ha), • Water cost savings (€ in 1 year)
Pesticide Use Calculator	• Current pesticide usage (kg or lt/ ha), • Current pesticide cost (€/ kg or lt)	• Pesticide saving (%), • Current pesticide cost (€/ ha), • Reduced pesticide usage (kg or lt /ha), • Pesticide cost saving (€/ ha), • Pesticide cost savings (€ in 1 year)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

Application Domain: Crop Farming Systems Calculator		
Calculator's Modules	**Data types required**	
	Provided by the end-user	**Provided by the system**
Labour Cost Calculator	• Current Labour Cost (€ in 1 year)	• Labour saving (%), • Labour cost savings (€ in 1 year)
FUEL Cost Reduction Calculator	• Current Fuel Cost (€ in 1 year)	• Fuel Saving (%), • Fuel cost savings (€ in 1 year)

Application Domain: Livestock Farming Systems Calculator		
Calculator's Modules	**Data types required**	
	Provided by the end-user	**Provided by the system**
Investment Cost Calculator	• Number of animals, • Years of Usage, • Number of Units (sensors etc.)	• Initial Cost of Investment (€), • Start-up Fee per Animal (€), • Monthly Cost per Animal (€)
Milk Yield Increase Calculator	• Average Milk Price per Litre (€/Litre), • Average Litres of Milk Produced per animal per Day (Litre)	• Milk Yield Increase (%), • The Price of Milk Yield in One Year (€)
LABOR Cost Reduction Calculator	• Current Labour Cost (€ in 1 year)	• Labour saving (%), • Labour cost savings (€ in 1 year
Energy Cost Calculator	• Current Energy Consumption (in kWh), • Current Cost of Energy (€/year)	• Energy Saving (%), • Cost of Energy per kWh (€), • Energy Cost Savings (€/year)
WATER Use Reduction Calculator	• Current Water Cost (€/ ha)	• Water saving (%) • Water Cost Saving (€ in 1 year) • Increase Profit per animal per year (€) • Profit per animal Increase in 1 year (€)
Feed Cost & Waste Calculator	• Current Feed Cost (€/year), • Current Feed Waste Cost (€/year)	• Feed Saving (%), • Feed Cost Saving (€/year), • Feed Waste Saving (%), • Feed Waste Cost Saving (€/year)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

Application Domain: Livestock Farming Systems Calculator		
Calculator's Modules	**Data types required**	
	Provided by the end-user	**Provided by the system**
Antibiotics Cost Calculator	• Current Antibiotics Cost (€/year)	• Antibiotics Saving (%), • Antibiotics Cost Saving (€/year)
Mortality Cost Calculator	• Current Mortality Cost (€/year)	• Mortality Rate Decrease (%), • Mortality Cost Saving (€/year)

b) Criteria-based interactive recommender tool

The Recommendation Tool is a criteria-based interactive tool that allows farmers (and farm advisors) to select the most appropriate solutions for their identified needs under a given set of conditions. Based on the criteria specified by the users, the tool proposes specific DATSs that cover their needs and provides specific quantified information on the costs, benefits and impact on sustainability of each solution. The user will be able to see the overall costs and benefits of each DATS in the form of a monetary value (in Euros) and with details provided on how this was calculated. The tool will be easy to use and the farmer will not need special training or the help of the advisor. Technically the tool will be based on a knowledge-based recommender system (building on existing systems and expertise) using a combination of reasoning and constraint satisfaction to provide a ranked list of the most suitable DATSs for any given farm/crop/biogeographical region combination. The Recommender tool takes various data elements as input.

The first main category will be the characteristics of the requesting farmer and his/her context:

- Name of the farm
- Country of the farm
- Biogeographical region (flat land, mountains)
- Agricultural sector (arable, horticulture, meat, dairy, apiculture, aquaculture)
- Agronomic model (Organic or traditional/industrial/greenhouse etc.)
- Type of agronomic activity (crop type/livestock/dairy)
- Location (GPS coordinates: latitude, longitude, altitude – or derived from postal address)
- Size of parcels/greenhouse/animals (ha or number)
- Soil type (sand, clay, silt, loam, peat, chalk)

These characteristics are used to search for similar farms and contexts that have used or are currently using DATSs which might be of interest to the requesting farmer.

The second category of input data are the parameters of a DATS the farmer is searching for:

- Functionality
- Target agricultural sector
- Target crop type
- Benefits/challenge to be solved
- Language/country
- Cost structure
- Digital form

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

The Recommendation tool will take values for these parameters indicated by the farmer and search for those DATSs with fully matching or closely similar values. This list can then be presented to the farmer.

The third category of input data are the aspects that the farmer wants to know about the selected DATSs:

- Initial investment costs
- Maintenance costs
- Expected ROI timeline
- Installation time
- Lifespan
- Amount of training needed
- Usability level
- Interoperability with other DATSs
- Regional suitability
- Environmental sustainability

The output of the Recommendation Tool is a list of DATSs and their characteristics that match the requested input criteria. In addition, for each DATS, a calculation is made of the costs and benefits for the specific farm and its context. This is expressed in terms of monetary value. Upon request, the farmer can get insight into the specific sub-categories that are used to calculate this monetary value. In the current version of the Recommendation Tool only a subset of the parameter lists above are incorporated. See the online version of the tool for further details: https://quantifarmtoolkit.eu/tool1.html.

c) Policy Monitoring tool

The "Policy Monitoring" tool will integrate and render -on a parcel and/or regional level- EO data products (e.g. crop type, NDVI, NDWI, land use), in-situ information (e.g. farm calendar exports describing applied farming practices, digital logs by utilised DATSs) along with qualitative and quantitative results from questionnaires on evaluated DATSs performance. The Policy Monitoring tool will provide policy makers with a visual policy monitoring dashboard, allowing the generation of analytical reports based on analytical queries, including summary tables as well as graphical charts. The Policy Monitoring tool will also facilitate data exchange through standardised API. A number of QuantiFarm test cases have been selected to evaluate their integration based on the technologies that will be assessed. An indicative list of TCs is: 1, 3, 4, 7, 8, 11, 17, 27, 29, 30 while the respective technologies that will be evaluated as ground truth data providers are Farm Management Information Systems, remote sensing with the use of UAVs, Variable Rate Application components applied on tractors, livestock automated monitoring, milking robots, automated monitoring for bees, robotic harvesting.

Data needs:

- Parcels location expressed in polygons with the use of coordinates,
- Farm/cultivation properties, applied farming practices and access to digital farm calendars,
- Digital data logs generated by the use of DATSs.
- Earth Observation data products (e.g. parcels shapes, NDVI, crop type estimation)
- Other relevant open datasets/repositories for the areas of the test cases: climate baseline values, biodiversity indicators, polygons-coordinates for environmental sensitive areas (Natura2000, water bodies, wildlife sanctuaries.

The data outputs will be selected policy performance indicators. Examples: Fertilisers (N/P/K kg/ha), Pesticides (list of active substances, quantity per substance of product, frequency of applications),

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

Irrigation (m3/hectare/year, total m3/year, frequency of irrigations), Land management (frequency of applied practices- ploughing, mowing), Harvested yield (number of harvests, quantity, quality).

As a final note, all tools are currently under specification and their design details including their data needs and calculated outputs will be documented in the respective deliverables. The flexible data modelling approach that the QCSM follows supports the necessary adaptations e.g. incorporation of additional domain specific ontologies. Hence, final specification of QuantiFarm Tools will dictate the final form of the QCSM.

3.4 Technical Requirements specification

The main purpose of the QuantiFarm data platform is to provide data about DATSs coming from various, heterogeneous sources in a uniform and unambiguous form to a set of QuantiFarm tools to be used by farmers, advisors and policy makers. This includes static data about DATSs, such as title, description, provider, functionality etc., but also dynamic data about its usage in the context of a specific setting, such as the QuantiFarm test cases.

Functional requirements

- A QuantiFarm database with DATS information including static and usage DATS data forms the core of the QuantiFarm data platform.
- The QuantiFarm database should provide data that is compliant to a common semantic model.
- The QuantiFarm tools can access the data in the QuantiFarm database upon request using a question-response mechanism.
- The QuantiFarm database collects information about DATSs from other existing DATSs repositories, such as FairShare and SmartAKIS.
- The QuantiFarm database collects all information of the QuantiFarm test cases about the usage of DATSs in practical situations.
- The QuantiFarm database maintains a mandate scheme that is used for secure access to the QuantiFarm test case data.
- The QuantiFarm database provides an identification, authentication and authorization mechanism that supports secure access control to its data based on the mandate scheme.

Non-functional requirements

- Performance: a query on the QuantiFarm database must respond within a reasonable time from the end-user's perspective. The exact acceptable response time depends on the complexity of the query. If needed intermediate results need to be returned iteratively to satisfy the user's request.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

4 QuantiFarm Common Semantic Model

Based on the requirements for the QuantiFarm data platform, a semantic model with common concepts and relations has been developed, the QuantiFarm Common Semantic Model (QCSM). In the next section, the design of the QCSM is described, its guiding principles and existing ontologies that have been reused and extended for modelling the required concepts and relations in the QCSM. Then, in the next subsections, the modelling of digital agricultural technologies and sustainability parameters in the QCSM is briefly identified. The complete QCSM and some of the reused and extended ontologies can be found on gitlab webpage: https://gitlab.com/QuantiFarm/qcsm.

4.1 Design of the QCSM

The design of the QCSM has been guided by a number of principles described in the next subsection. Afterwards, the existing ontologies that have been reused and extended are being listed.

4.1.1 Guiding principles

For the design of the QCSM, the following guidelines have been used:

1. The QCSM is based on semantic technologies, like RDF and OWL, because it is currently the best way of intuitively defining formal semantics (OWL) and provides the flexibility for modular reuse of existing data models or extending them.
2. The QCSM should be a small, core model that covers the main common concepts in the agrifood domain with a focus on the farm, its digital agricultural technologies and sustainability performance.
3. The concepts and relations for the QCSM are selected from general properties of digital agricultural technologies on the farm and sustainability characteristics as identified by the assessment framework for the QuantiFarm test cases.
4. Existing ontologies that already define required concepts are reused by the QCSM as much as possible. In addition, proposals for extension of existing ontologies are being made as long as this fits into the scope of these existing ontologies.
5. Required concepts that cannot be modelled in or added to an existing ontology will be defined within the QuantiFarm namespace for the QCSM, namely https://quantifarm.eu/ontology/ prefixed as `qcsm`.
6. Existing ontologies are only reused when they have a clear formal OWL structure that is publicly available and accessible or downloadable in a .owl, .ttl or .rdf format, for instance at OBO foundry website, W3C website or the AgroPortal (http://agroportal.lirmm.fr). Consequently, no reuse of proprietary ontologies of different projects will be done.
7. Vocabularies and thesauri/taxonomies that simply define and list a large set of hierarchical terms will not be reused in the QCSM other than using the `rdf:isDefinedBy` property to point to the definition of the concept in a vocabulary.

A consequence of these guiding principles is that existing data models that are not proper ontologies are not explicitly inherited and extended by the QCSM. Unfortunately, extensive data models like NGSI-LD and the Demeter AIM fall into this category.

4.1.2 Reuse and extension of existing ontologies

As a consequence of guiding principle 4 on reusing and extending existing ontologies as much as possible, the QCSM imports various existing ontologies. We therefore have built further on the work that has been done in the Ploutos project (https://ploutos-h2020.eu) in which already various ontologies have been reused and extension proposals for some of them have been generated. Table 4 below sketches the reused ontologies and their scope.

Prefix	Name	Base URI	Scope
ENVO	Environment Ontology	http://purl.obolibrary.org/obo/envo.owl#	Ecosystems, environment, geographical components and processes
S4AGRI	SAREF4AGRI	https://saref.etsi.org/saref4agri/	Smart appliances and devices for agricultural purposes
SSN	Semantic Sensor Network	http://www.w3.org/ns/ssn/	Sensor networks and processes
SOSA	Sensor Observation Sample Actuator	http://www.w3.org/ns/saso/	Observations, actuations and samples from sensors and actuators
OM	Ontology of units of Measure	http://www.ontology-of-units-of-measure.org/resource/om-2/	Units hierarchy for measurements
Weather	BIMERR Weather Ontology	https://bimerr.oit.linkeddata.es/def/weather#	Weather phenomena and parameters
GS1	GS1 web vocabulary	https://gs1.org/voc/	Supply chain product information including certification
Ploutos	Ploutos Common Semantic Model	https://www.tno.nl/agrifood/ontology/ploutos/common#	Supply chain actors and operations

Table 6: Existing ontologies to be reused by the QCSM

As part of the Ploutos' use cases on traceability and certification, proposals have been made for the extension of the Saref4Agri, GS1 and Weather ontologies with new concepts and/or properties. In the context of the QuantiFarm project, reuse and extension proposals have been made again to Saref4Agri as will be discussed in the next sections.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

4.2 Modelling digital agricultural technologies

Information about digital agricultural technologies is at the moment being collected and made available by a few platforms, specifically FairShare and SmartAkis. In addition, the QuantiFarm test case assessment framework defines another few properties of DATSs. The collection of the most important DATS properties is defined in Table 5.

Name	Description	QCSM property
Title	Short title of the DATS	qcsm:DATS -> qcsm:hasTitle -> xsd:string
Description	Short description of the DATS and its purpose	qcsm:DATS -> qcsm:hasDecription -> xsd:string
Keywords	Keywords that characterise the DATS and its usage	qcsm:DATS -> qcsm:hasKeyword -> xsd:string
Provider name	Supplier of the DATS where it can be purchased	qcsm:DATS -> qcsm:hasProvider -> xsd:string
Functionality	Main services the DATS supports	qcsm:DATS -> qcsm:hasFunction -> xsd:string
Agricultural sector	Sector in the agriculture in which the DATS is meant to be used	qcsm:DATS -> qcsm:isTargetedtoSector -> xsd:string
Benefits	Main advantages of using the DATS	qcsm:DATS -> qcsm:hasBenefits -> xsd:string
Initial investment	Initial capital expenditures to purchase and start using the DATS	qcsm:DATS -> qcsm:hasInitialInvestment -> xsd:decimal
Maintenance costs	Operating expenditures for using the DATS	qcsm:DATS -> qcsm:hasMaintenanceCost -> xsd:decimal
Lifespan	Duration in which the DATS can be used normally functioning	qcsm:DATS -> qcsm:hasLifespan -> time:Duration
Installation time	Time interval in which the DATS can be made operational	qcsm:DATS -> qcsm:hasInstallationTime -> time:Interval

Table 7: Main DATS properties modelled in the QCSM

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

In the QCSM, the class of DATSs is modelled as the concept *qcsm:DigitalAgriculturalTechnology* with various datatype properties of basic types like *xsd:string*, *xsd:date* and *xsd:decimal*. In addition, relations with other concepts are defined as object properties to for instance the *qcsm:TechnologyProvider*, *saref4agri:Farm* and *gs1:Country* classes. An overview of the DATS concept, its properties and relations with other concepts is visualised in Figure 5.

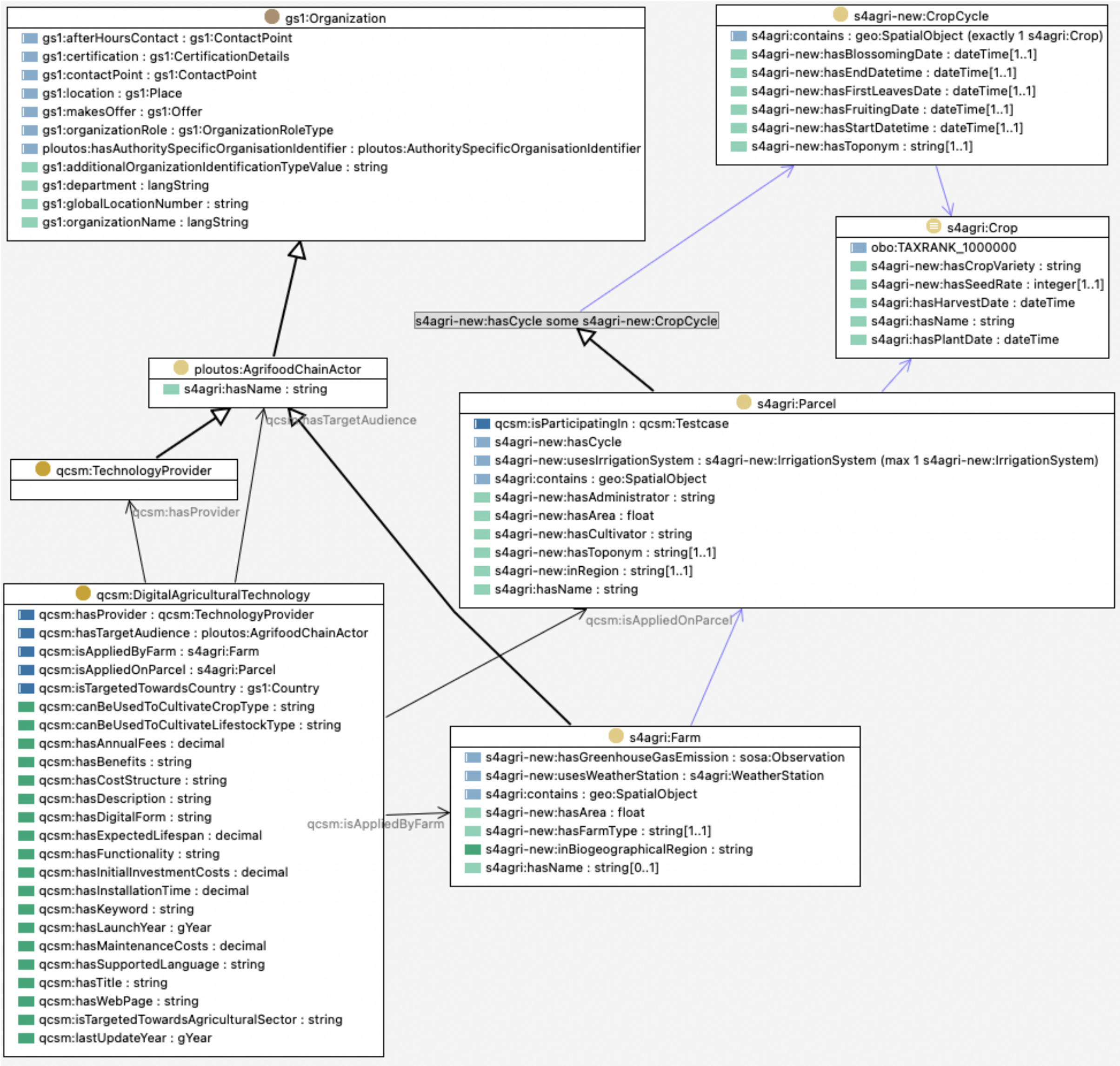

Figure 5: QCSM model with DATS concept and properties

4.3 Modelling sustainability parameters

Within the QuantiFarm assessment framework three main pillars of sustainability parameters are distinguished:

- Environmental sustainability, which includes parameters that indicate how sustainable the processing and usage of resources is.
- Economic sustainability, which includes parameters that indicate whether a farm is viable in the long term.
- Social sustainability, which includes parameters that indicate how socially responsible the farm operates to keep its personnel vital for the long term.

These parameters are being measured and monitored in each test case on a yearly basis when using one or more specific DATSs. In more detail, every test case has DATSs information, and every assessment year of a test case has market information and one or more comparisons of two parcels. The comparison contains values for the sustainability parameters of the two parcels, one with the usage of DATSs and the other without. This is modelled in the QCSM using the class *qcsm:AssessmentInformation* depicted in Figure 6.

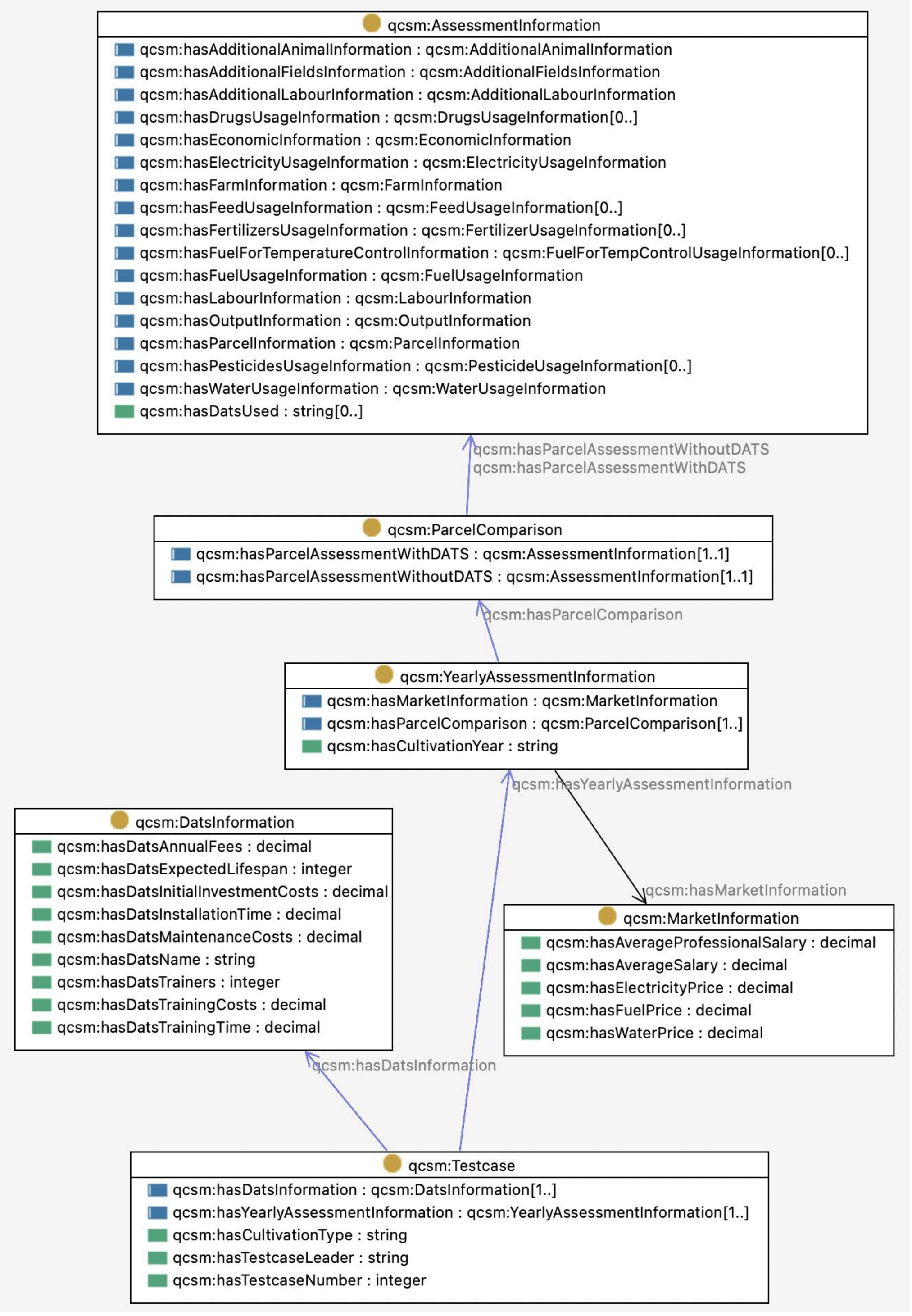

Figure 6: QCSM classes and properties for a test case, with DATSs and yearly assessment information

More details about the exact sustainability parameters are presented in the next subsections.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

4.3.1 Environmental sustainability parameters

These parameters include aspects such as usage of water, fuel and electricity, the health of the animals or the crop in the field, and the type and quantity of pesticides, fertilizers, feed and drugs used at a farm. Table 6 contains the environmental sustainability indicators per category defined in the QuantiFarm assessment framework.

WATER		
#	**data**	**u.m.**
1	Irrigation	m^3
2	Fertigation	m^3
3	Pest dilution	m^3
4	Cleaning	m^3
5	Drinking	m^3
6	**Total water used**	m^3

FUEL		
#	**data**	**u.m.**
7	Fertilisation	liter
8	Pesticides	liter
9	Irrigation	liter
10	Sowing / Planting	liter
11	Temperature and Humidity Control	liter
12	Pruning	liter
13	Beehives visits	liter
14	Transporting oysters to the laboratory	liter
15	Milking	liter
16	Feeding	liter
17	Quantity of greenhouse fuel X (multiple times)	liter
18	**Total fuel used**	liter

ELECTRICITY		
#	**data**	**u.m.**
19	Temperature and Humidity Control	kWh
20	Irrigation	kWh
21	Milking	kWh
22	Heat detection	kWh
23	Calving detection	kWh
24	Feeding	kWh
25	**Total electricity used**	kWh

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

PESTICIDES		
#	**data**	**u.m.**
26	Treatments with pesticide X (multiple times)	number
27	Average quantity per treatment with X	kg/ha

FERTILIZERS		
#	**data**	**u.m.**
28	Treatments with fertilizer X (multiple times)	number
29	Mineral or organic fertilizer	string
30	Average quantity per treatment with X	kg/ha

FEED		
#	**data**	**u.m.**
31	Quantity of feed X (multiple times)	ton

DRUG		
#	**data**	**u.m.**
32	Quantity of drug X (multiple times)	gram

HEALTH		
#	**data**	**u.m.**
33	Number of death of animals/oyster/beehives	number
34	Calves died in the first 24h	number
35	Tot. number of calves born	number
36	Number of days per year with access to pasture	number
37	Outdoor loafing area	ha
38	Number of cows producing high SCC milk	number

Table 8: Main environmental sustainability indicators of the assessment framework.

These parameters are modelled in the QCSM by additional classes with an *owl:DatatypeProperty* for each parameter with a basic type as *rdfs:range*, i.e. *xsd:string* or *xsd:decimal*. An overview of these classes and properties is given in Figure 7.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

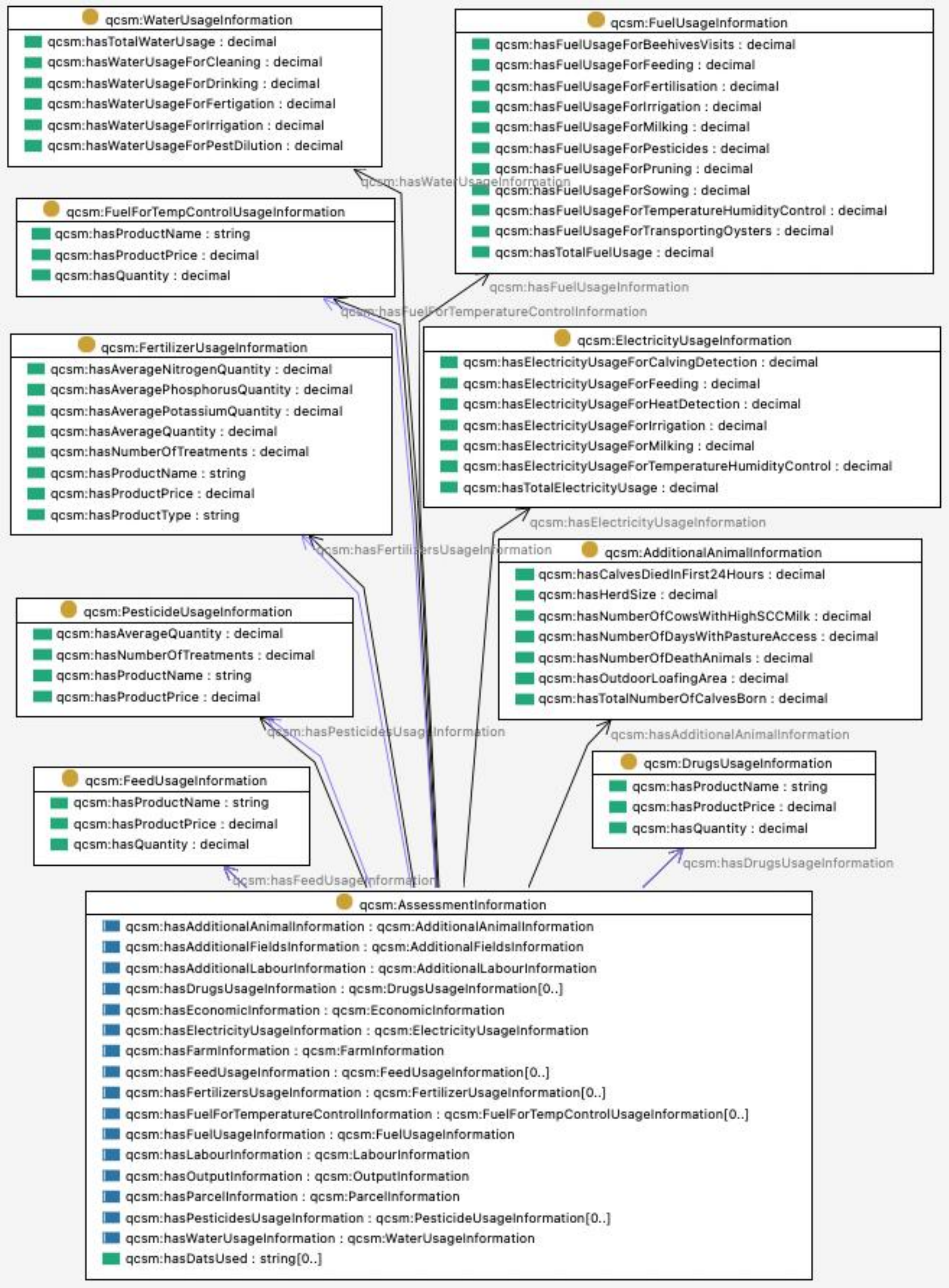

Figure 7: QCSM classes and properties for modelling environmental sustainability parameters

In addition, a specific part of the QCSM models all the possible operations on a parcel, including irrigation operations, fertilisation operations and chemical control operations in which water, fertiliser

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

and pesticides respectively are added to the soil of a parcel. An overview of the *saref4agri:Parcel* concept, its *ploutos:Operations* and the use of water, fertiliser and pesticides is given in Figure 8.

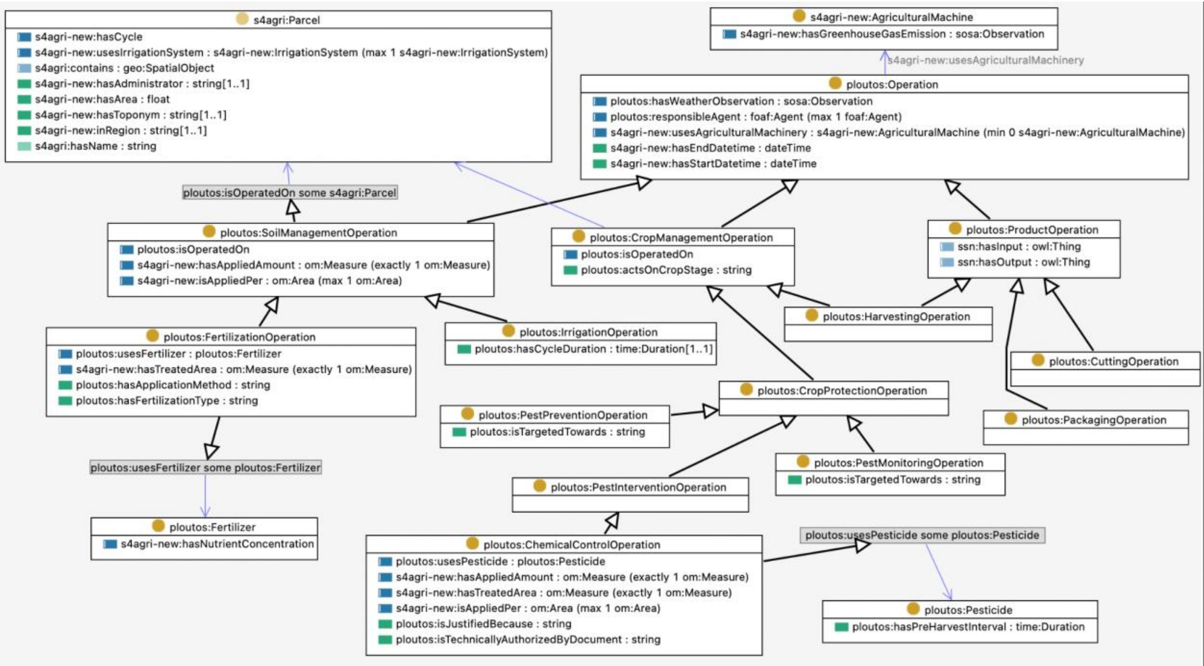

Figure 8: QCSM operations and the use of water, fertiliser and pesticides

4.3.2 Economic sustainability parameters

These parameters consider aspects about costs and benefits that influence the economic viability of the farm, but also aspects about the quality of the product. They include market costs data, such as prices of resources, product market prices, salary costs etc., revenue and quality related data and output data. Table 7 contains the economic sustainability indicators per category defined in the QuantiFarm assessment framework.

MARKET COSTS AND PRICES		
#	**data**	**u.m.**
39	Water price	€/m^3
40	Fuel price	€/liter
41	Number of fuel for greenhouse temperature control price	number
42	Fuel for greenhouse temperature control price	€/m^3
43	Electricity price	€/kWh
44	Average salary (farmer and employees)	€/hour
45	Average professional salary (agronomist, veterinary, ...)	€/hour
46	Sold product market price	€/ton
47	Pesticide price (multiple possible)	€/kg
48	Fertiliser price (multiple possible)	€/ton
49	Feed price (multiple possible)	€/ton
50	Drug price (multiple possible)	€/gram

Table 9: Main economic sustainability indicators of the assessment framework

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

REVENUE AND QUALITY		
#	**data**	**u.m.**
51	Number of on-time fulfilled orders in a year	number
52	Number of orders received in a year	number
53	Number of wrong orders in a year	number
54	Additional revenues	€/l milk
55	Protein % on dry matter	%
56	Humidity	%
57	Test Weight	kg/hl

OUTPUT		
#	**data**	**u.m.**
58	Crop production	ton
59	Milk	liter
60	Meat	kg
61	Animal	number
62	Oyster	pieces
63	Honey	kg
64	Silos product management	ton

These parameters are modelled in the QCSM by additional classes with an *owl:DatatypeProperty* for each parameter with a basic type as *rdfs:range*, i.e. *xsd:string*, *xsd:integer* or *xsd:decimal*. An overview of these classes and properties is given in Figure 8.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

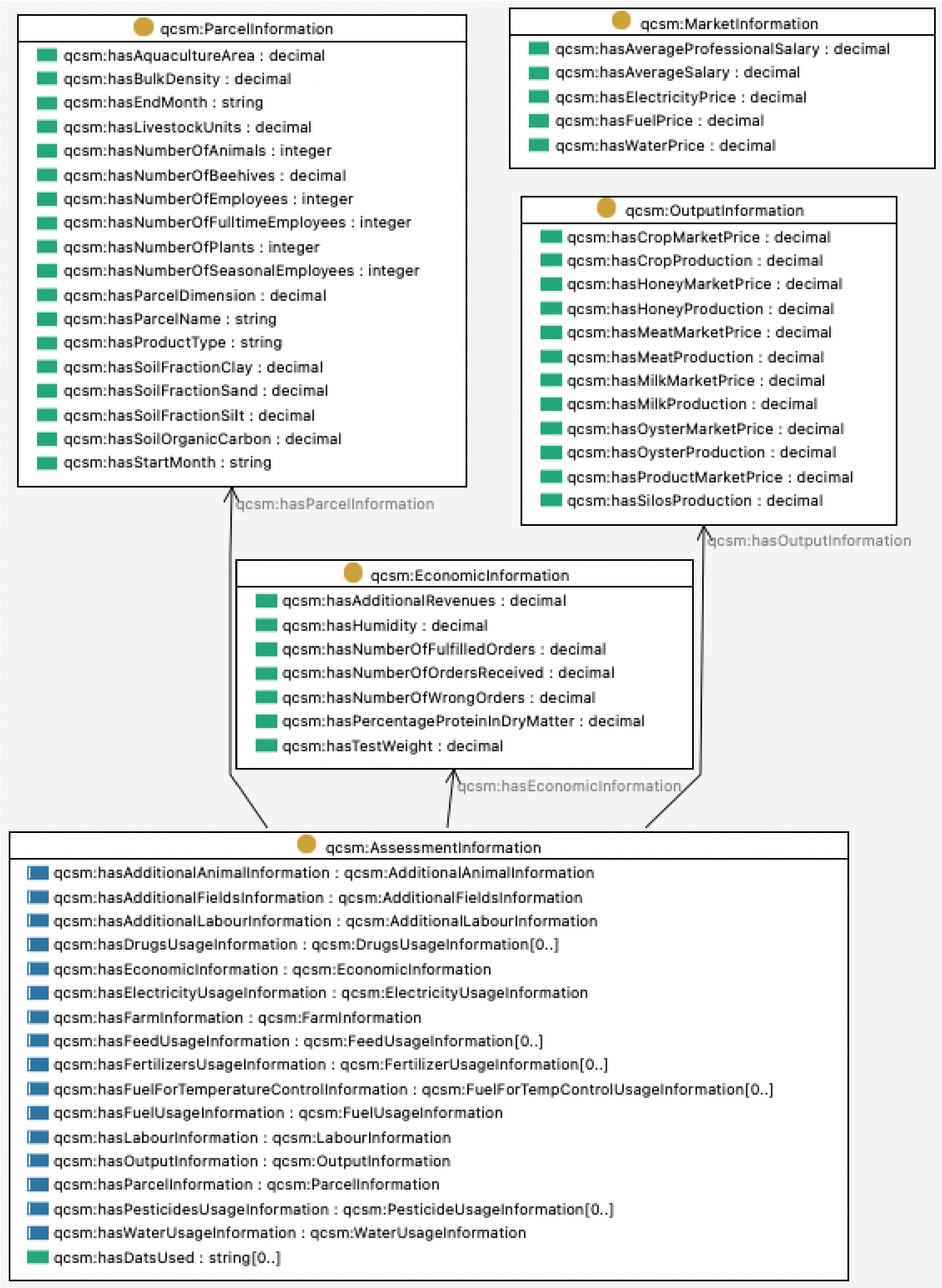

Figure 9: QCSM classes and properties for modelling economic sustainability parameters

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

In addition, a specific part of the QCSM models the operations on a parcel in which fertilisers and pesticides are added to the soil of a parcel and is shown in Figure 9. It also includes the cost in Euros of a fertiliser or pesticide per unit of volume, e.g. litre or kilograms.

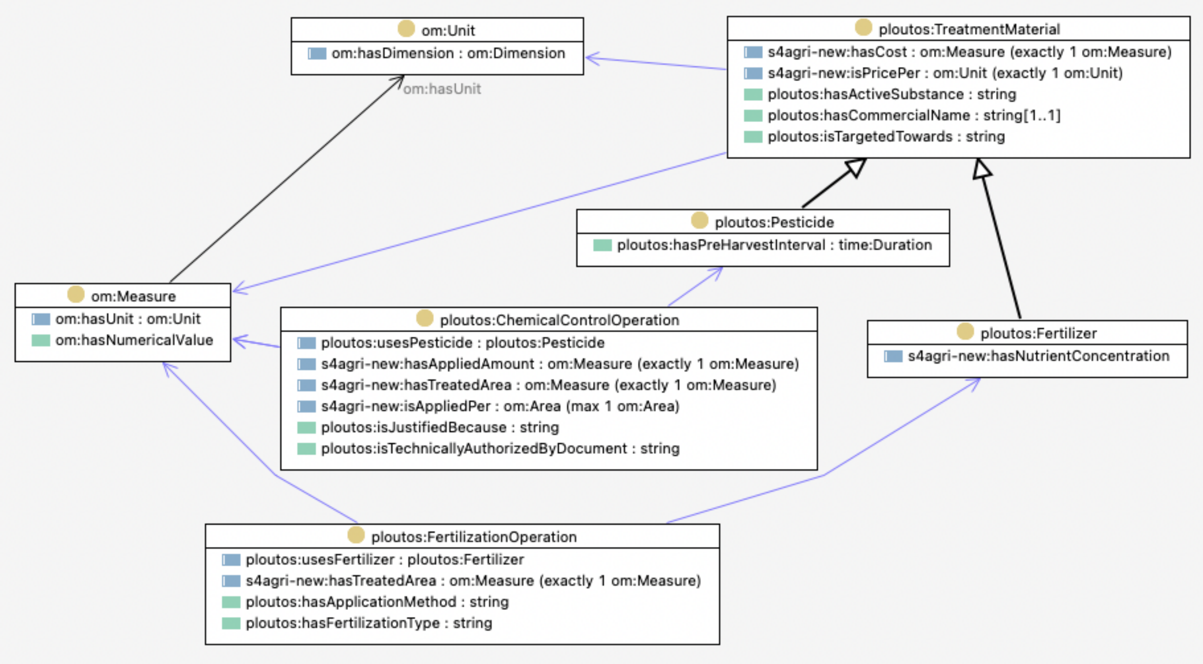

Figure 10: QCSM fertiliser and pesticide operations

4.3.3 Social sustainability parameters

These parameters consider aspects like the mandatory number of working hours/weeks, the actual amount of hours spent on various activities, the number of employees hired in total and the safety of the working environment. Table 8 contains the social sustainability indicators per category defined in the QuantiFarm assessment framework.

LABOUR		
#	**data**	**u.m.**
65	Irrigation	hours
66	Pesticides	hours
67	Fertilisation	hours
68	Sowing / Planting	hours
69	Pruning	hours
70	Field visits	hours
71	Greenhouse Management	hours
72	Tracking the quantity of grain available in the silo	hours
73	Assessing product quality (grain, vegetables, fruit...)	hours
74	Controlling purchase and sale prices and quantities	hours
75	Organise logistics	hours
76	Cleaning	hours

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

LABOUR		
#	**data**	**u.m.**
77	Stable visits	hours
78	Heat detection	hours
79	Calving detection	hours
80	Pigsty management	hours
81	Milking	hours
82	(Milk) Quality assessment	hours
83	Feeding	hours
84	Beehives visits	hours
85	Sampling of oysters	hours
86	Transporting oysters to the laboratory	hours
87	Harvesting	hours
88	Storage	hours
89	Administrative activities, data collection, etc.	hours

SAFETY AND CAPACITY		
#	**data**	**u.m.**
90	New injury cases	number
91	Total number of lost working hours due to occupational injuries	number
92	Working hours	hours
93	Number of working weeks in a year	number
94	Number of new local employees hired	number
95	Number of new employees hired	number

Table 10: Main social sustainability indicators of the assessment framework

These parameters are modelled in the QCSM by additional classes with an *owl:DatatypeProperty* for each parameter with a basic type as *rdfs:range*, i.e. *xsd:string*, *xsd:integer* or *xsd:decimal*. An overview of these classes and properties is given in Figure 10.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

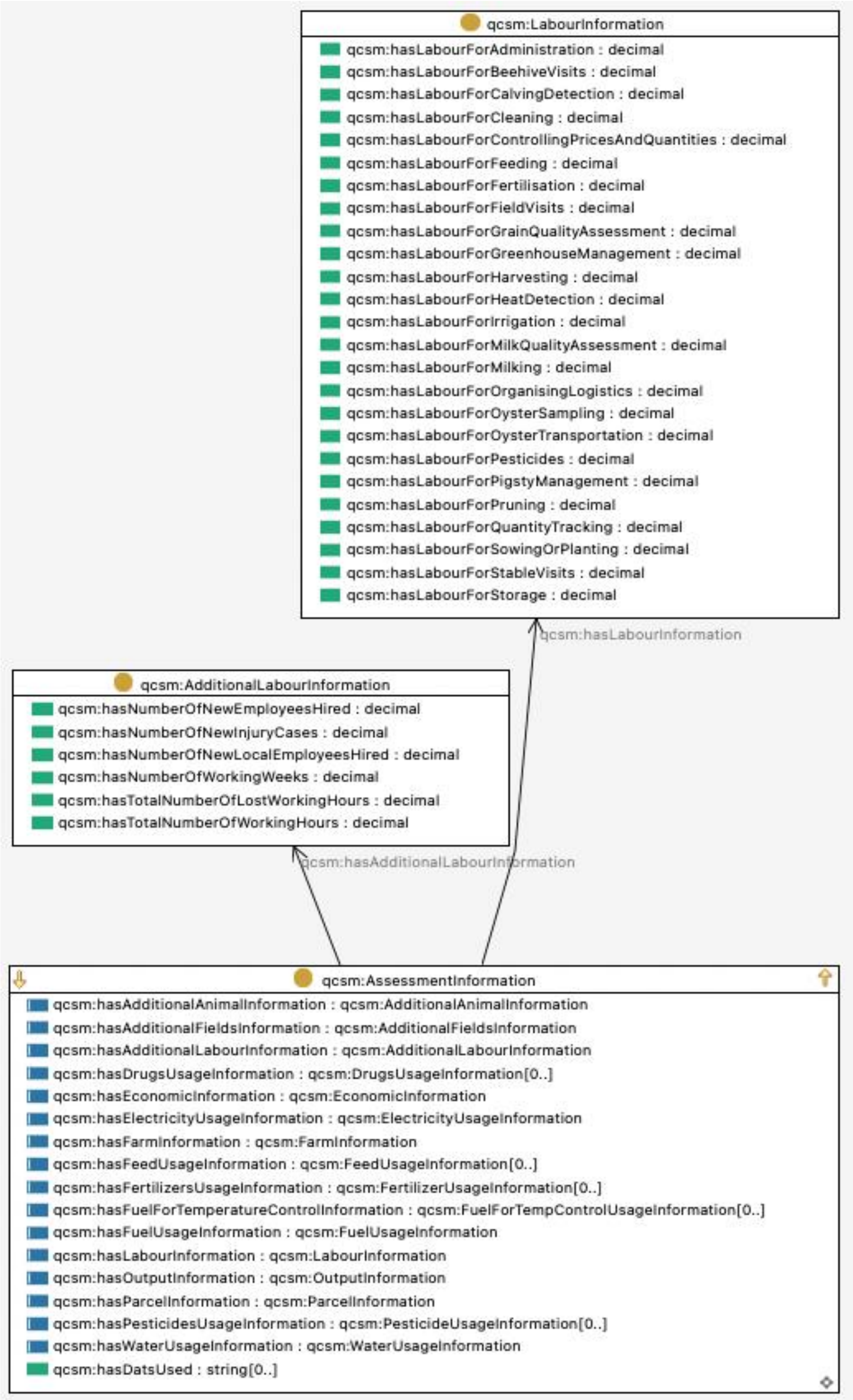

Figure 11: QCSM classes and properties for modeling social sustainability parameters

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

In addition, we added new concepts to cover injuries and workers. That specific part of the QCSM is depicted in Figure 11 and models the *s4agri-new:FarmWorker* that works at a *s4agri:Farm*, executes various *ploutos:Operations* and can get hurt by some *s4agri-new:Injuries*.

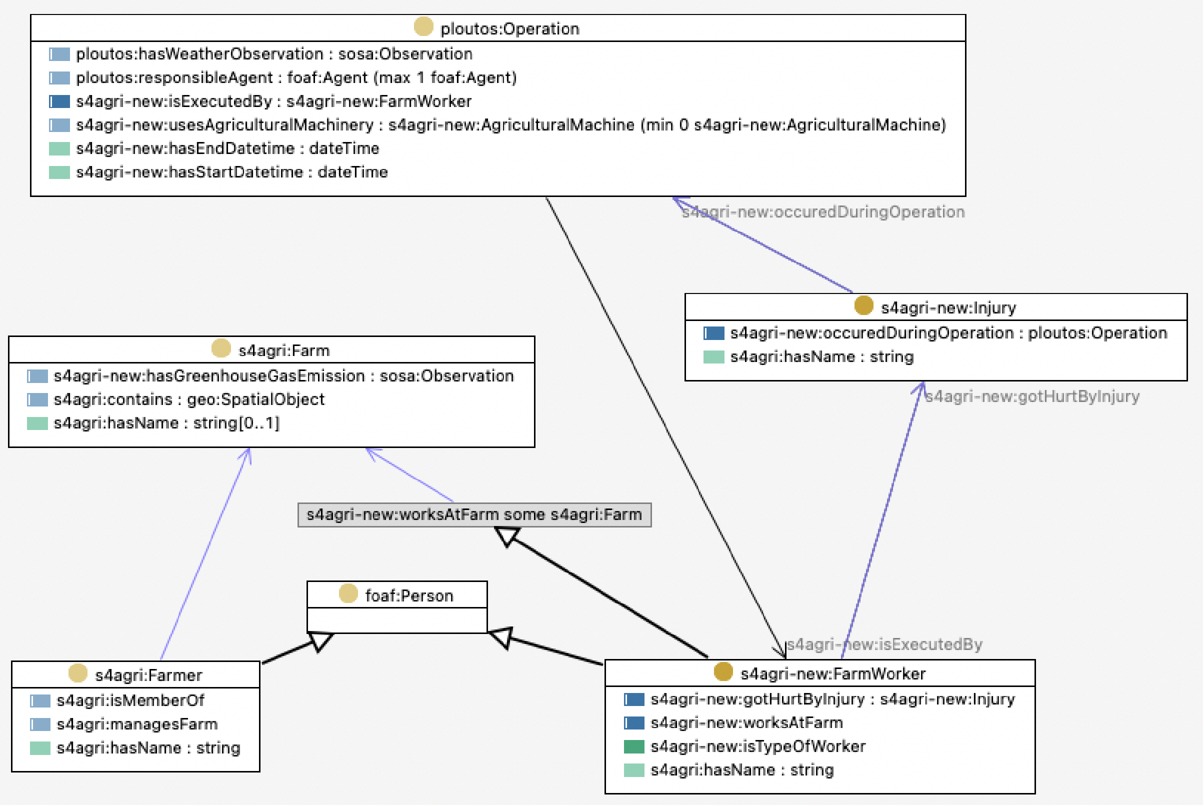

Figure 12: QCSM part of the FarmWorker and possible injuries

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

4.4 Modelling behavioural intentions

Besides DATS parameters and environmental, economic and social sustainability indicators, also behavioural intentions influence the consideration, uptake and usage of DATSs. These behavioural intentions and the factors that determine them are being studied by work package 1 of the QuantiFarm project and are currently being described in D1.1. The infographic that is made for an overview of these factors is presented in Figure 12.

When studying this infographic, it can be seen that some of the factors can be determined objectively, such as age, gender, size of business or type of farm. However, other factors are more subjective and denote the opinion of the DATS user, e.g., ease of use, prediction of risks or difficulty of learning. Nevertheless, both types of factors are important to measure during the process of encountering, consideration, implementation and usage of a DATS.

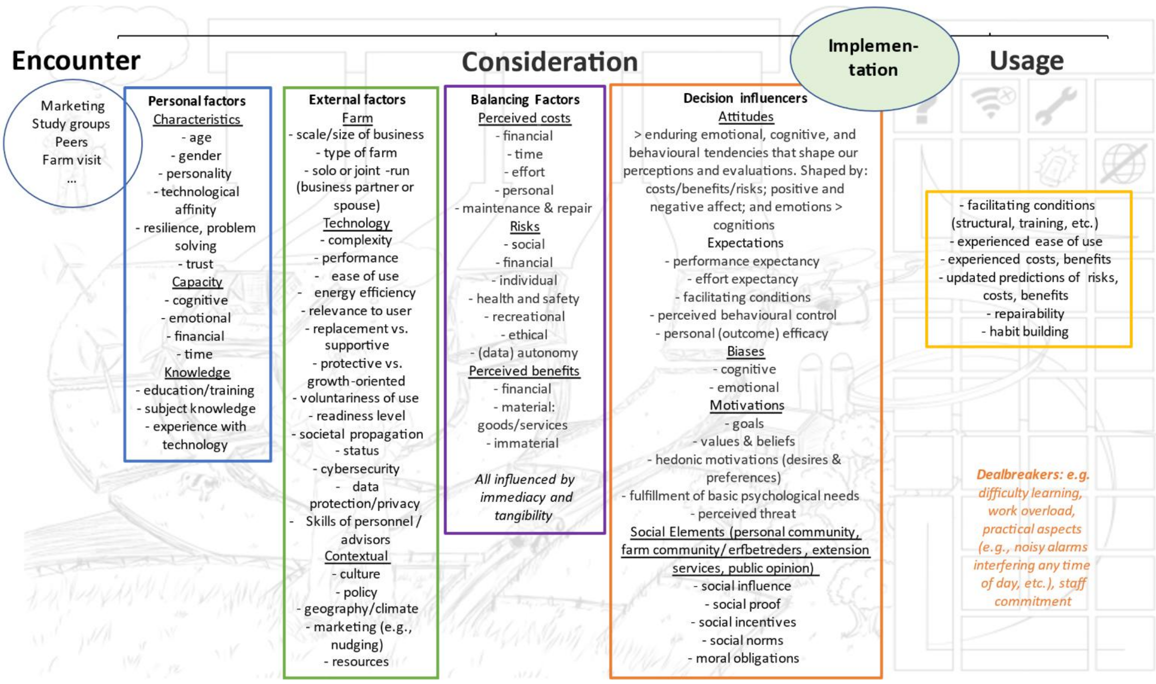

Figure 13: Infographic of the behavioural intention factors

With respect to adding these parameters to the QCSM, a subset is already present or can easily be added to existing concepts. The Personal Factors can be added as properties to the Farmer concept, the External Factors of the farm can be added as properties to the Farm concept and technology parameters as properties to the DATS concept. Most of the Balancing factors about costs, risks and benefits have overlap with the economic sustainability indicators and can be added to these parts of the QCSM. During the second half of 2023 and year 2024, we have further discussed the use of these parameters in the light of giving a recommendation for a DATS. This has resulted in a small subset of parameters that were added to the QCSM to model farm and farmer characteristics that could be used to match with the DATS that are available in the QuantiFarm toolkit.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

5 QuantiFarm Data Platform Design

5.1 Use cases for QuantiFarm Data platform

A crucial step for the design of a software system - the QuantiFarm data platform in our case- is the specification of the "Use Cases[2]" in Unified Modelling Language (UML) format. UML is a standardised modelling language consisting of an integrated set of diagrams, developed to help system and software developers for specifying, visualising, constructing, and documenting the artefacts of software systems, as well as for business modelling and other non-software systems. A UML use-case diagram models the behaviour of a system and helps to capture the requirements of the system. Use-case diagrams describe the high-level functions and scope of a system. These diagrams also identify the interactions between the system and its actors. The use cases and actors in use-case diagrams describe what the system does and how the actors use it, but not how the system operates internally. Use-case diagrams illustrate and define the context and requirements of either an entire system or the important parts of the system. To this end, Figure 13 illustrates the external entities expected to interact with the QuantiFarm data platform and most important interactions.

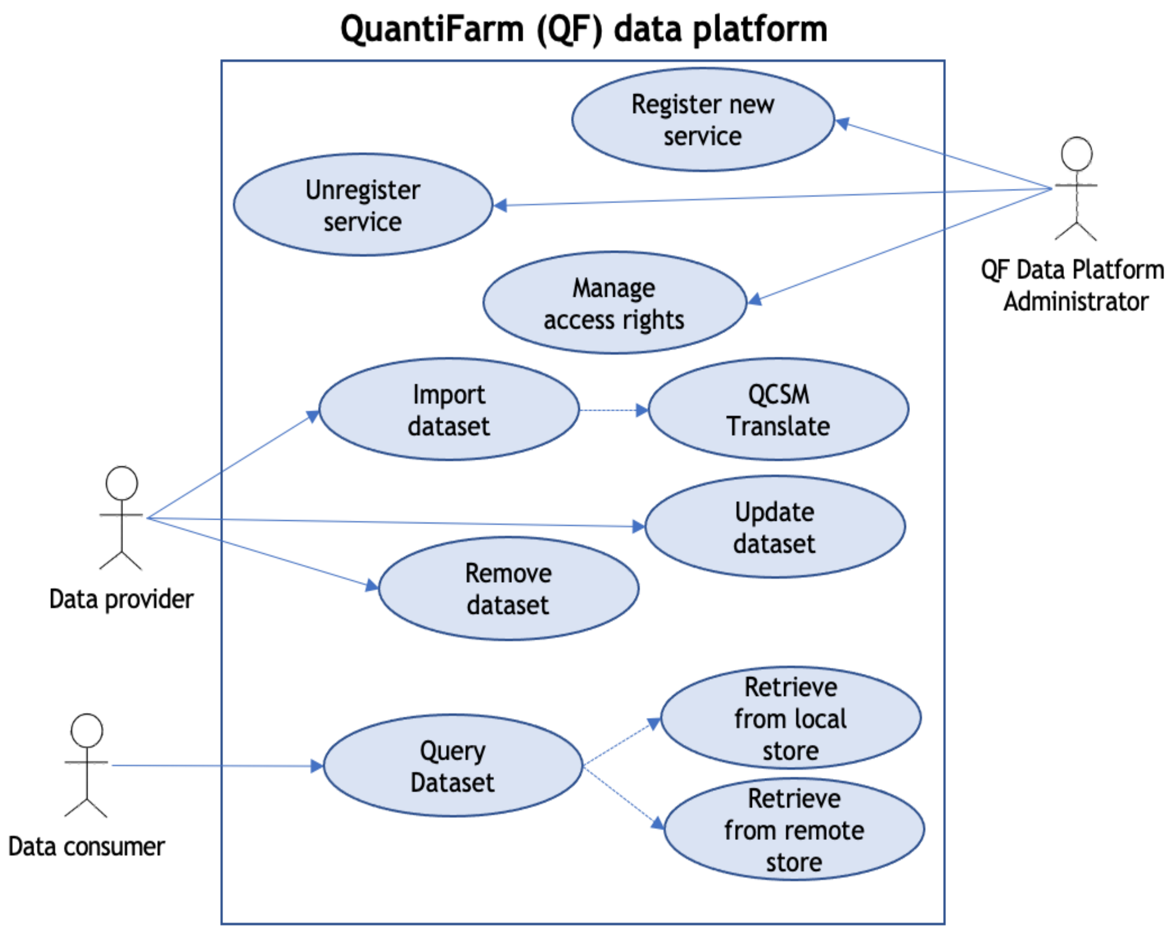

Figure 14: Use case diagram for QuantiFarm data platform

[2] "IBM Rational Software Architect documentation" https://www.ibm.com/docs/en/rational-soft-arch/9.6.1?topic=diagrams-use-case

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

At this stage three main actors of the QuantiFarm data platform are identified:

- QuantiFarm data platform administrator: This actor represents an entity (actual person or a service) allowed to register a new data provider, remove an existing one and to update/configure the access rights of a registered data provider.
- Data provider: This actor represents data providers within the data platform. A data provider can be a registered service that utilises QF data platforms APIs for storing data within the platform. As it is illustrated in Figure x it is expected that a data provider will be able to import, update and remove datasets according to the access rights that is assigned with. During dataset import a data translation process needs to be realised to ensure data harmonisation with the use of the QSCM.
- Data consumer: This actor represents entities that have the right to perform queries to the database and retrieve selected dataset. According to the issued request parameters the response can be routed either to the local data store or to remote data repositories.

5.2 QuantiFarm Data platform functional architecture

Figure 11 illustrates the QuantiFarm data platform in a diagram. The individual components are described below, but broadly speaking, we rely on a *central data storage* where most of the relevant data for the tools is stored. *Mappers* are components that can retrieve specific data from external APIs that are either too big to be loaded into the central data storage entirely or have the requirement to be available in real-time. The *generic engines* are used to calculate values of aggregated parameters for test case assessment and cost-benefit assessment. The *tool services* (corresponding to the tools defined in the project proposal) retrieve data from the central data storage and mappers to fulfil their data needs. The *user interface* gives a human friendly interface with all tool functionality in a single consistent interface.

Design principles:

- Large datasets–such as granular weather data–are not stored in our triple store, as it would be impractical to synchronise them.
- External datasets that need to be available for the tools in realtime are not stored in our triple store, but disclosed via a web service that translates data from the external data format to the QuantiFarm Common Semantic Model (QCSM).
- When data is ingested into the QuantiFarm Data Platform, it should be mapped into the QCSM at the moment of ingestion.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

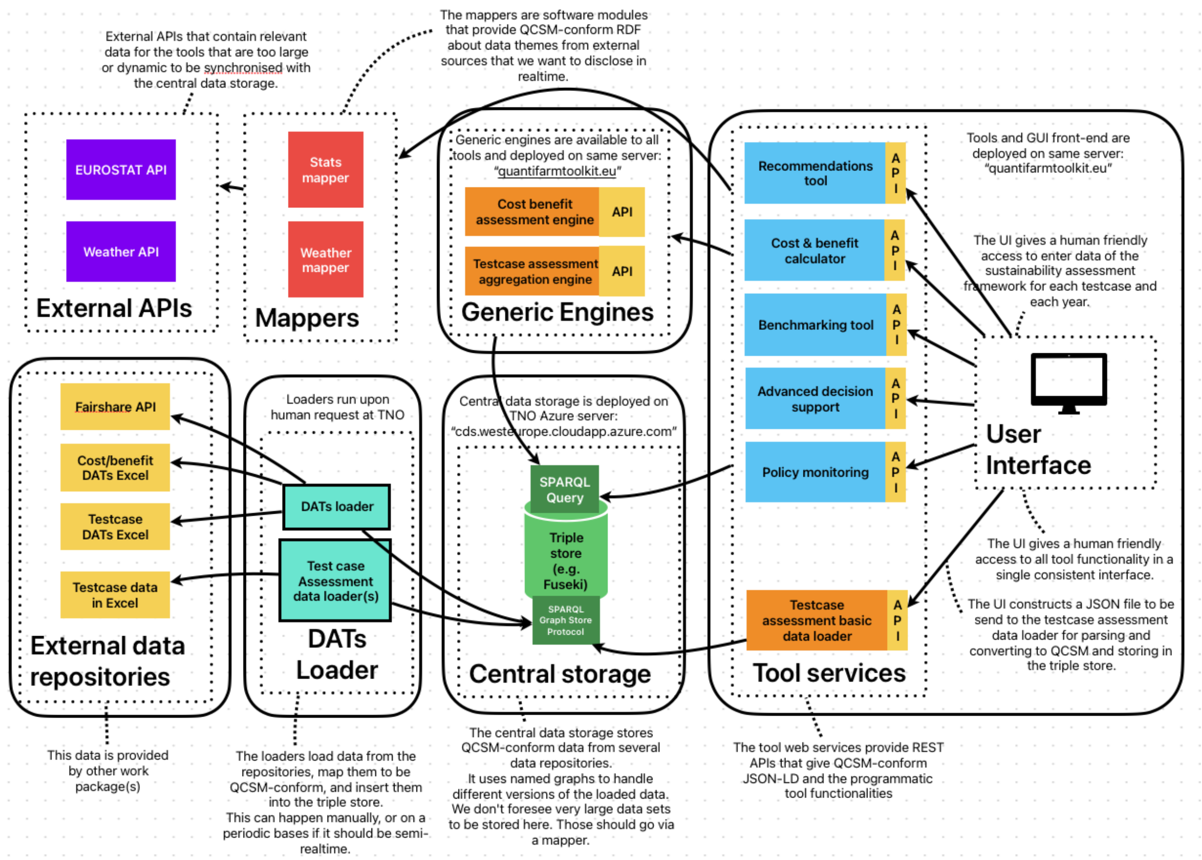

Figure 15: Functional architecture of QuantiFarm data platform

5.2.1 Central data storage

The *central data storage* (green components in the bottom middle of the diagram) is responsible for storing RDF data that need to be available for the tool services. All data that is ingested into the data storage should be conformant with the ontologies in the QCSM. For the tool services to retrieve the necessary data, SPARQL queries can be used, which is a flexible and standardised query language for RDF data. In the current prototype, the Apache 2.0-licensed triple store *Apache Jena Fuseki* is used. It provides a multitude of ways to load a retrieve data from the triple store, using standardised protocols such as SPARQL.

5.2.2 Loaders

The *loaders* (cyan components in the bottom of the diagram) load data from several data sources in an *extract, transform load (ETL)* manner. The data concerned does not involve real-time use cases, so it can be loaded into the data storage whenever it is deemed necessary. The task of a loader is to access its corresponding data repository, map it into QCSM-conform RDF, and load it into the central data storage. In our data platform, this concerns the repository of DATSs in FAIRshare, KPIs of farms that participate in test cases (and which DATSs they use, if any), among others. An important requirement is that the loader can be executed again after a while (e.g., when the data changed, or when the loader is improved), and then it should replace its previous data. This is implemented with the SPARQL Graph Store HTTP protocol, specifically with the HTTP PUT method[3] on a named graph for the loader. To produce RDF, the well known rdflib library in Python is used. It provides utilities for generating statements in a graph, and serialising the graph into formats that the central data storage understands.

[3] https://www.w3.org/TR/sparql11-http-rdf-update/#http-put

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

5.2.3 Mappers

Some external data sources are simply too big to be loaded into the central data storage, or they need to be made available in a real-time fashion instead of ETL. For these data sources, we use special *mappers* which are software modules that can retrieve specific data from these sources and map them to QCSM-conform RDF for the tool services to use. These mappers were originally envisioned to be separate microservices, but it is more practical to simply implement them as reusable code modules that the tool services can import.

5.2.4 Data repositories

There are several *data repositories* that we want to leverage in the data sharing platform. In general, these are not available in a homogeneous way and should thus be transformed into QCSM-conform RDF before they can be conveniently consumed by the tools. The FAIRshare project has produced a list of Digital Advisory Tools and Services which roughly translate into what we call Digital Agricultural Technology Solutions in QuantiFarm. These are enriched with relevant metadata about the technology so that it can be transformed into RDF statements about the technology. WP2 produced KPIs and metadata about farms participating in the test cases in the form of Excel sheets. The data from these sheets are loaded and transformed by the loaders using the pandas tool in Python. Subsequently, they are loaded into the central data storage.

5.2.5 External APIs

We envision several *external APIs* to be consumed directly as they are needed, and so bypass the central data storage because they're either too big to be loaded entirely, or are needed in real-time. Examples include (historical) weather APIs, EUROSTAT and CAP indicators.

5.2.6 User Interface

The *user interface* is the frontend that the user interacts with. There will be a single consistent user interface that gives personalised access to all the tools using a single thematic style. We have a non-functional prototype, but other than that, this is planned work, and more will be written about it in planned deliverables (D3.6). The user interface uses the tool services to populate its views with the content and functionality that are required.

5.2.7 Tool services

Each tool is implemented as an independent *tool service*, which comprises one or more web services with well defined APIs. With these APIs, the user interface can retrieve the data that it needs to present the content and provide the functionality.

5.2.8 Generic engines

Generic engines are available to all tools and have the functionality to calculate aggregated values of parameters. In particular, the cost-benefit assessment engine calculates costs and benefits for using a specific DATS and the testcase assessment engine calculates the sustainability parameters for a testcase and the DATSs used in that testcase. The engines have a well-defined API via which the calculated values can be retrieved. In addition, the results of the engines are stored in the central data storage as well.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

5.3 Implementation

The prototype of QuantiFarm data platform is implemented and an end-to-end working example of data collection from an external source, data harmonisation with the use of QCSM, data storage to a central triplestore and data retrieval through the webservice is demonstrated. The implementation process and code development follow open source principles even from an early stage and hence a public gitlab platform is available for code development in a collaborative manner.

https://gitlab.com/QuantiFarm/data-platform

Directions on how to deploy the software components and to initiate the working examples are available in the README file (https://gitlab.com/QuantiFarm/data-platform/-/blob/main/README.md) and also available in the Annex C of this Deliverable.The source code of the data sharing prototype is available as a Docker Compose project. The docker-compose.yml file is also included in Annex D.

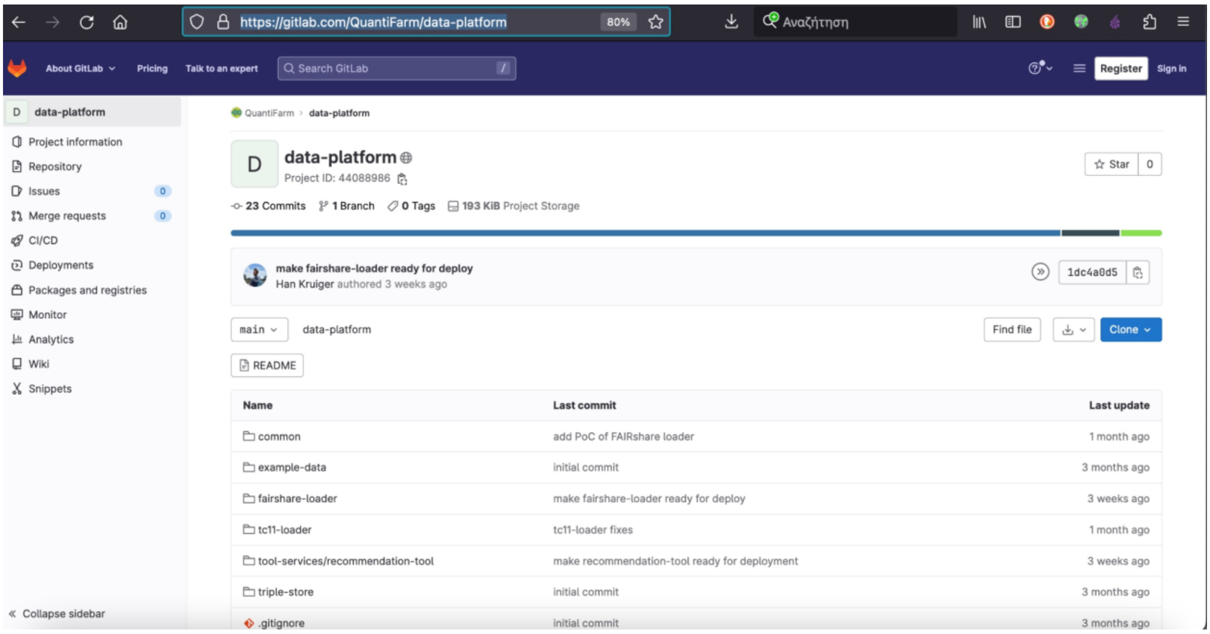

Figure 16: Screenshot of QuantiFarm's gitlab project

6 Conclusions

In this deliverable, an updated version of the QuantiFarm Common Semantic Model (QCSM) and the QuantiFarm Data Platform are described. This model and data platform is being used to collect data from various sources that is needed for the QuantiFarm toolkit for farmers, advisors and policy makers. The collected data is stored in a central datastore and made accessible in terms of the QCSM. The data requirements for the model and platform are collected from various angles, such as the QuantiFarm assessment framework from work package 2 that is used to assess the 30 QuantiFarm test cases and the behavioural intentions framework from work package 1 that describes motivations for using DATSs. In addition, existing sources of DATS information, such as FairShare and SmartAKIS, are being reused.

The current version of the QCSM might be considered as complete with respect to the QuantiFarm requirements. Nevertheless, in 2025 further development of the QuantiFarm toolkit and its assessment engine might cause the QCSM to be extended with additional concepts that cover the full assessment calculation of DATSs usage.

The most up to date version of the QCSM is available here: https://gitlab.com/QuantiFarm/qcsm/-/blob/main/qcsm.ttl

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

7 References

[1] Brickley, Dan; Guha, Ramanathan V.; Layman, Andrew, eds. (1998-04-09). "Resource Description Framework (RDF) Schemas". W3C. W3C Working Draft. RDF Schema Working Group. Retrieved 2021-04-23.

[2] Bikakis N.; Tsinaraki C.; Gioldasis N.; Stavrakantonakis I.; Christodoulakis S., eds. (2012-03-21). "XML and Semantic Web W3C Standards Timeline-History" (PDF). Archived from the original (PDF)on 2013-04-24. Retrieved 2021-04-23.

[3] Brickley, Dan; Guha, Ramanathan V., eds. (2014-02-25). "RDF Schema 1.1". W3C. 1.1. RDF Working Group. Retrieved 2021-04-23.

[4] Knublauch, Holger; Oberle, Daniel; Tetlow, Phil; Wallace, Evan (9 March 2006). "A Semantic Web Primer for Object-Oriented Software Developers". W3C. Retrieved 19 November 2017.

[5] "W3C Standard Facilitates Data Management and Integration". W3.org. 27 October 2009. Retrieved 15 October 2013.

[6] S. Rajbhandari and J. Keizer, The AGROVOC Concept Scheme – A Walkthrough, Journal of Integrative Agriculture 11(5) (2012), 694–699. 35 doi:10.1016/S2095-3119(12)60058-6. https://www.sciencedirect.com/science/article/pii/S2095311912600586.

[7] T. Baker, B. Whitehead, R. Musker and J. Keizer, Global Agricultural Concept Space: Lightweight Semantics for Pragmatic Interoperability, npj Science of Food 3(1) (2019), 16. doi:10.1038/s41538-019-0048-6. https://www.nature.com/articles/s41538- 019- 0048- 6.

[8] E. Arnaud, M.-A. Laporte, S. Kim, C. Aubert, S. Leonelli, B. Miro, L. Cooper, P. Jaiswal, G. Kruseman, R. Shrestha, P.L. Buttigieg, C.J. Mungall, J. Pietragalla, A. Agbona, J. Muliro, J. Detras, V. Hualla, A. Rathore, R.R. Das, I. Dieng, G. Bauchet, N. Menda, C. Pommier, F. Shaw, D. Lyon, L. Mwanzia, H. Juarez, E. Bonaiuti, B. Chiputwa, O. Obileye, S. Auzoux, E.D. Yeumo, L.A. Mueller, K. Silverstein, A. Lafargue, E. Antezana, M. Devare and B. King, The Ontologies Community of Practice: A CGIAR Initiative for Big Data 41 in Agrifood Systems, Patterns 1(7) (2020), 100105. doi:10.1016/j.patter.2020.100105. https://www.sciencedirect.com/science/article/pii/S2666389920301392.

[9] EFSA, The Food Classification and Description System FoodEx2 (Revision 2), Technical Report, European Food Standards Agency, 2015. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/sp.efsa.2015.EN- 804.

[10] D.M. Dooley, E.J. Griffiths, G.S. Gosal, P.L. Buttigieg, R. Hoehndorf, M.C. Lange, L.M. Schriml, F.S.L. Brinkman and W.W.L. Hsiao, FoodOn: A Harmonized Food Ontology to Increase Global Food Traceability, Quality Control and Data Integration, npj Science of Food 2(1) (2018), 23. doi:10.1038/s41538-018-0032-6.

[11] L. Magliulo, L. Genovese, V. Peretti and N. Murru, Application of Ontologies to Traceability in the Dairy Supply Chain, Agricultural Sciences 04(05) (2013), 41–45. doi:10.4236/as.2013.45B008.

[12] J.P.C. Verhoosel and J. Spek, Applying Ontologies in the Dairy Farming Domain for Big Data Analysis, in: Proceedings of the Joint Proceedings of the 3rd Stream Reasoning (SR 2016) and the 1st Semantic Web Technologies for the Internet of Thing (SWIT 2016) Workshops (SR+SWIT), Kobe, Japan, 2016, pp. 91–100. http://ceur- ws.org/Vol- 1783/paper- 09.pdf.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

[13] J. Verhoosel, M. Van Bekkum and T. Verwaart, Semantic Interoperability for Data Analysis in the Food Supply Chain, International Journal on Food System Dynamics Vol 9 (2018), 101–111 Pages. doi:10.18461/IJFSD.V9I1.917. http://centmapress.ilb.uni-bonn.de/ojs/index.php/fsd/article/view/917

[14] D. Project, D2.3 Common Data Models and Semantic Interoperability Mechanisms - Release 2, Deliverable, Deliverable D2.3, Demeter Project, 2021.

[15] F.K. van Evert, S. Fountas, D. Jakovetic, V. Crnojevic, I. Travlos and C. Kempenaar, Big Data for Weed Control and Crop Protection, Weed Research 57(4) (2017), 218–233. doi:10.1111/wre.12255.

[16] J. Verhoosel, M. van Bekkum and T. Verwaart, Semantic Interoperability for Data Analysis in the Food Supply Chain, International Journal on Food System Dynamics 9(1) (2018), 101–111. doi:10.18461/ijfsd.v9i1.917. http://centmapress.ilb.uni-bonn.de/ojs/index.php/fsd/article/view/917.

[17] M. Solanki and C. Brewster, Enhancing Visibility in EPCIS Governing Agri-Food Supply Chains via Linked Pedigrees:, International Journal on Semantic Web and Information Systems 10(3) (2014), 45–73. doi:10.4018/IJSWIS.2014070102. http://cbrewster.com/papers/ Solanki_IJSWIS15.pdf.

[18] M. Solanki and C. Brewster, OntoPedigree: Modelling Pedigrees for Traceability in Supply Chains, Semantic Web 7(5) (2016), 483–491. doi:10.3233/SW-150179. http://www.semantic- web-journal.net/system/files/swj980.pdf.

[19] A.Y. Halevy, A. Rajaraman and J.J. Ordille, Data Integration: The Teenage Years, in: Proceedings of the 32nd International Conference on Very Large Data Bases, Seoul, Korea, September 12-15, 2006, U. Dayal, K.-Y. Whang, D.B. Lomet, G. Alonso, G.M. Lohman, M.L. Kersten, S.K. Cha and Y.-K. Kim, eds, ACM, 2006, pp. 9–16. http://dl.acm.org/citation.cfm?id=1164130.

[20] I. Horrocks, M. Giese, E. Kharlamov and A. Waaler, Using Semantic Technology to Tame the Data Variety Challenge, IEEE Internet Computing 20(6) (2016), 62–66. doi:10.1109/MIC.2016.121. http://ieeexplore.ieee.org/document/7781544/.

[21] M. Klusch, P. Kapahnke, S. Schulte, F. Lecue and A. Bernstein, Semantic Web Service Search: A Brief Survey, KI - Künstliche Intelligenz 30(2) (2016), 139–147. doi:10.1007/s13218-015-0415-7.

[22] A.L. Lemos, F. Daniel and B. Benatallah, Web Service Composition: A Survey of Techniques and Tools, ACM Computing Surveys 48(3) 20 (2015), 33:1–33:41. doi:10.1145/2831270.

[23] J. Domingue, D. Roman and M. Stollberg, Web Service Modeling Ontology (WSMO): An Ontology for Semantic Web Services, in: W3C Workshop on Frameworks for Semantics in Web Services, Innsbruck, Austria, 2005. https://www.w3.org/2005/04/FSWS/Submissions/1/wsmo_position_paper.html.

[24] D. Martin, M. Paolucci, S. McIlraith, M. Burstein, D. McDermott, D. McGuinness, B. Parsia, T. Payne, M. Sabou, M. Solanki, N. Srinivasan and K. Sycara, Bringing Semantics to Web Services: The OWL-S Approach, in: Semantic Web Services and Web Process Composition, J. Cardoso and A. Sheth, eds, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2005, pp. 26–42. ISBN 978-3-540-30581-1.

[25] J. Miller, K. Verma, P. Rajasekaran, A. Sheth, R. Aggarwal and K. Sivashanmugam, WSDL-S: Adding Semantics to WSDL, Kno.e.sis Publications (2005). https://corescholar.libraries.wright.edu/knoesis/199.

[26] S.A. McIlraith and T.C. Son, Adapting Golog for Composition of Semantic Web Services, in: Proceedings of the Eights International Conference on Principles of Knowledge Representation and Reasoning, KR’02, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002, pp. 482–496. ISBN 978-1-55860-554-1.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101059700

[27] D. Redavid, L. Iannone, T.R. Payne and G. Semeraro, OWL-S Atomic Services Composition with SWRL Rules, in: Foundations of Intelligent Systems, 17th International Symposium, ISMIS 2008, Toronto, Canada, May 20-23, 2008, Proceedings, A. An, S. Matwin, Z.W. Ras and D. Slezak, eds, Lecture Notes in Computer Science, Vol. 4994, Springer, 2008, pp. 605–611. https://eprints.soton.ac.uk/265132/1/ ISMIS08.pdf.

[28] E. Sirin and B. Parsia, Planning for Semantic Web Services, in: Proceedings of the ISWC 2004 Workshop on Semantic Web Services: Preparing to Meet the World of Business Applications, SWS@ISWC 2004, Hiroshima, Japan, November 8, 2004, D. Martin, R. Lara and T. Yamaguchi, eds, CEUR Workshop Proceedings, Vol. 119, CEUR-WS.org, 2004. http://ceur-ws.org/Vol-119/paper1.pdf.

[29] D. Androcˇec, M. Novak and D. Oreški, Using Semantic Web for Internet of Things Interoperability: A Systematic Review, International Journal on Semantic Web and Information Systems (IJSWIS) 14(4) (2018), 147–171. doi:10.4018/IJSWIS.2018100108. https: //www.igi-global.com/article/using- semantic- web- for- internet- of- things- interoperability/210657.

[30] M. Noura, M. Atiquzzaman and M. Gaedke, Interoperability in Internet of Things: Taxonomies and Open Challenges, Mobile Networks and Applications 24(3) (2019), 796–809. doi:10.1007/s11036-018-1089-9.

[31] M. Bauer, E. Kovacs, A. Schülke, N. Ito, C. Criminisi, L.-W. Goix and M. Valla, The Context API in the OMA Next Generation Service Interface, in: 2010 14th International Conference on Intelligence in Next Generation Networks, 2010, pp. 1–5. doi:10.1109/ICIN.2010.5640931.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

8 Annex A

Test Case and DATS data collection template. In this template questions referring to data types and data recording process were introduced that assisted the requirement specifications for the QuantiFarm data model and the QuantiFarm data platform. The relevant questions are listed here and are underlined within the template.

Technology Description

3. Referring to the DATS's software, please indicate and/ or describe:
a. The parameters or indicators that can be extracted from the software providing important information to the farmer.
b. The output delivered by software analysis and how it influences the farmers decision making process.
c. The platform and display mode used to consult the data provided by the DATS.

4. Referring to the DATS's hardware, please indicate, and describe:
The sensors and technologies used for the collection of data on the parameters referred in the answers above

8. Regarding data collection, please indicate:
a. How often do you collect data?
b. How does the data collection process work?
c. Is the data collected through the DATS enough for decision making purposes, or some other tool is used (even manually)?
d. Is there any data that you collect periodically and have to communicate/ report to any external organisation? Please detail.

Test Case (TC) Description Form		
This form will serve to better understand the Test Case (TC) details with special focus on the DATS (Digital Agricultural Technology Solution) description, the cost benefit relation, and sustainability impact. With this information we will be able to develop the most suitable analysing methods for your Test Case. The objective is to coordinate efforts among all of us and plan the methodologies that will be used to process and analyse the data provided by you. Given the multilateral aspects of this project, we will be testing and analysing data from different technologies, sectors, crop/ animal, and biogeographical regions. Also, given the importance of cycles on the farm, parts of this survey, along with new questions, will be asked again during the project. In order to facilitate this task, the analysing parameters should be set accordingly to the TC needs so we may do as little changes as possible along the project, avoiding major methodological changes and complications in the data collection process along the projects' duration. For this reason, it is important to gather as much information as possible about the technology, the parameters it provides, how it assists in the decision-making process, and the conditions in which it is operating. During the application stage you have provided some information about the TC. To normalise the information gathered from all TCs, we are now asking for more details about the DATS and the impact on agriculture activity. Please feel free to use the information provided at the application stage to give background on specific details asked in this document. We count on your expertise on the field and technology to provide a Test Case description as detailed as possible. Note - Throughout the questionnaire the word "DATS" is mentioned repeatedly. In the context of QuantiFarm and this questionnaire, a DATS is defined as a particular Digital Agriculture solution that consists of hardware (e.g.: tractor, drone, computer, robot, weather station, etc.) and/or software, and that helps the farmer in performing an individual production step (e.g.: fertilisation, weeding, irrigation, harvesting etc) or in managing the farm.		
Partners involved		
Partner name	**Role / Value Chain Actor Type**	**Involvement**

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

	(e.g.: Farmer, Cooperative, Agri-food company, Technology Provider, Contractor etc.)	*(Full partner / Third party)*

Country and Biogeographical Region

Agricultural Sector
• *Please select from the following list which agricultural objects are the focus of your Test Case* ○ *Apiculture* ○ *Aquaculture* ○ *Arable* ○ *Livestock - Dairy* ○ *Horticulture - Fruit* ○ *Livestock - Meat* ○ *Horticulture -Vegetables*

Please provide the following information about the farm and the farmers involved
• *Please fill-out the following tables (according to necessity) for each of the farms. Copy the table as many times as necessary.* • *Please copy paste as many tables as necessary to organise the information*

Exact location of the farm where the DATS is used	
Number of parcels on the farm where the DATS is applied	
Size of these parcels (ha or number of animals)	
Crop cultivated on these parcels (there might be more than one)	
Total approximate revenue of the farm (please indicate an interval of values)	
Farmers Name	
Farmers contact (email)	

Exact location of the farm where the DATS is used	
Number of parcels on the farm where the DATS is applied	
Size of these parcels (ha or number of animals)	
Crop cultivated on these parcels (there might be more than one)	
Total approximate revenue of the farm (please indicate an interval of values)	
Farmers Name	
Farmers contact (email)	

……

Description of the end user of the DATS
• Please provide some information on who will be mostly using the DATS in daily practice and in what role (e.g.: farmer, advisor, agricultural cooperative, etc).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

Technology Description
To contextualise the Test Case, please provide a general description on the DATS (250 words max)
• What are the main goals in using the technology (e.g.: input/cost reduction, yield increase, risk reduction, environmental sustainability, animal welfare, etc.) • Part of the farm operations where the technology is applied (e.g.: individual production steps or "whole-farm approach"). • In which process phases are the DATS applied and how does it help (e.g.: cultivation, irrigation, sowing, harvesting, farm management, traceability)?
1- Referring to the technology provider: a. Who is/are the technology developer(s)? Please describe the organisations that have been involved in the development of the DATS and their specific roles.
• *Please indicate the name of the organisations involved, and what roles they had (e.g.: researcher, prototype development, manufacturer, etc.)*
b. Who is responsible for the DATSs' commercialization and what is the commercialization model?
• *Please let us know who the owner of the DATS is* • *Did the farmer buy the solution or is it just paying for services? Did a contractor buy the solution and is using it to offer services to farmers?* • *Who made the decision to adopt the DATS?* • *Was there an investment needed to adopt the DATS?* If yes, how much was the total investment? • *Who were the sponsors / public incentives / investors (if there were any)?* • *If a different funding model was used, please describe it as detailed as possible (e.g.: annual subscription with no additional investment).*
2- Referring to farmers: What type of benefits have you obtained thanks to the use of the DATS? Was the investment "worth it"?
• *On what specific tasks is the DATS helpful? How is it helpful? (e.g.: time saving, money saving, worker comfort, friendlier impact on the environment, information management, etc.)*
3- Referring to the DATS's software, please indicate and/ or describe: **a.** The parameters or indicators that can be extracted from the software providing important information to the farmer.
• *Some of the information provided by the software comes in the form of parameters and indicators like - (for precision irrigation) the water evapotranspiration (for animal welfare) body temperature, (arable) growth rate, (milking) amount of milk. Please let us know what is provided by your DATS.*
b. The output delivered by software analysis and how it influences the farmers decision making process.
• *Apart from single parameters and indicators, some software compiles that information providing direct suggestions and lines of action (e.g.: when to harvest, what section needs water, what animal has a health problem etc.). Please elaborate for your DATS.*
c. The platform and display mode used to consult the data provided by the DATS.
• *Please let us know where you check the information provided by the DATS (e.g.: phone, laptop, if there is remote access, if you must be close to the DATS)* • *Please indicate how the data is made available (e.g.: via an application programming interface, email, cloud repository, paper)*
4- Referring to the DATS's hardware, please indicate, and describe: The sensors and technologies used for the collection of data on the parameters referred in the answers above.
• *In the previous question you mentioned the important information gathered from the software. Now, we would like to ask you about the sensors and technologies (hardware) that gather the same information. (e.g.: a tractor, pivot, weather tower, processing line, drone, temperature sensor, infrared sensor, pressure sensor, humidity sensor, specific substance detector, etc.)*
5- Regarding DATS installation, please describe the process

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

• *Please describe, to the best of your knowledge, the installation steps.* • *Who was responsible for installing the DATS on the farm? (e.g.: DATS provider's technical team, farmer, advisor, a combination of stakeholders, etc.)* • *The installation process went according to plan, or were there adjustments needed to accommodate specific needs?* • *The DATS installation finished on time? If not, was it longer or shorter than expected?*
6- Usually, the DATSs are implemented to address a certain need on the farm and to assist in performing the tasks that are directly related with that need. But there might be cases when the DATS use enables farmers to accomplish additional tasks beyond the focus of the DATS (e.g.: use a particular system of the DATS, like a tractor, a drone, weather tower, or a specific sensor). If that is the case, please indicate and describe: **a.** What other tasks benefit from the use of the DATS?
b. What DATSs are involved in those other tasks.
• *e.g.: weather station, tractor, drone, etc.*
c. How do these systems contribute to the accomplishment of additional tasks?
• *e.g.: a weather station may be there to support a DSS on pest management but could also support the farmer in getting an early warning about frost, even if this was not foreseen originally by the DATS provider.*
7- For how long this DATS has been used on the farms integrated in the TC (consider the first farms to systematically use the DATS)?
8- Regarding data collection, please indicate: **a.** How often do you collect data?
• *Hourly, daily, weekly, monthly, annually, etc.?*
b. How does the data collection process work?
• *Some DATSs have this process automated via IoT, but there might be the case where you must dislocate to the DATS to extract the data directly from there. Please describe your situation.* • *List of data types that the Test Case is expected to generate (e.g.: weather parameters recordings, fuel consumption, scouting outcomes about pest infestation)* • *Please indicate the unit for measuring each data type (e.g.: Celsius, Litres)* • *What data model is available that describes the data that is being measured about these agricultural objects?*
c. Is the data collected through the DATS enough for decision making purposes, or some other tool is used (even manually)?
• *Do you use excel or any other tool/app to treat the data somehow? If you do, please let us know the tools and the methods used to analyse the data.* • *Please tell us about the format that the Dataset is expressed (e.g.: JSON, XLS, CSV).* • *Is it necessary to extract data using a hard drive?*
d. Is there any data that you collect periodically and have to communicate/ report to any external organisation? Please detail.
• *What data is shared, and to what purpose?* • *Which data types are open and available to be reused in the context of QuantiFarm?* • *To which organisation do you communicate/ report it to?* • *Is this communication/reporting mandatory?*
9- The TC will consist of comparing data from a parcel using the DATS to another parcel that is not using the DATS. Referring to the parcel that will not be using the DATS, please indicate if the same tasks/ processes are done. If so, please describe how the same tasks/ processes are performed.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

• *We would like to know how the tasks were made before DATS implementation (or without DATS) and compare a "before and after" scenario.* • *If some of the tasks were not done before DATS implementation, please state it and explain why.*
10- Still referring to the parcel not using the DATS, please indicate in more detail: **a.** If and how the parameters/ indicators monitored by the DATS (when the DATS is available) are controlled when the DATS is not available.
• *Please try to be as specific as possible on how the parameters analysis was done without the DATS. If the analysis was not possible because there was no data available, please let us know the method that was used on the farm to understand how the decisions were taken (e.g.: how did you know when to water the crop, when to harvest, when to call the vet?).*
b. How many more (or less) farm workers are involved in performing the tasks without the DATS when compared to the same tasks using the DATS?
• *It is natural that DATS implementation allows for the reduction of the working force on performing certain tasks, but in some cases that might not happen. Please let us know if there was a reduction in the workload and to what point.*
c. How many additional hours do the farm workers spend performing the tasks without the DATS?
• *If there is no direct reduction on the number of workers with DATS implementation, perhaps there was a reduction in the amount of time they spent performing the tasks. Please let us know an approximation of the time saved in person-hours.*
d. Are there any tools or practices that became obsolete with DATS implementation? Does this obsoletion have any consequences to other tasks in the farm? Please describe it.
• *With DATS implementation, some previously used tools and common practices for the same purpose may have lost their practicality / usefulness (or part of it). Please let us know, in that case, if there are additional methods that have "suffered" or tasks that became harder to accomplish.*

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

9 Annex B

The schema utilised by the FAIRShare platform for modelling the properties of Digital Agriculture Technologies. These properties have been incorporated within the QCSM.

```
const toolSchema = mongoose.Schema({
  name: { type: String, required: true },                // No 1
  countries: [{ type: String, required: true }],          // No 2
  languages: [{ type: String, required: true }],          // No 3
  website: {type: String, required: true},                 // No 4
  cost: {type: String, required: true},                    // No 5
  launchYear: {type: String, required: true},              // No 6
  updateYear: {type: String, required: false},             // No 7
  usersNo: {type: String, required: false},                // No 8
  downloads: {type: String, required: false},              // No 9
  providerName: {type: String, required: true},            // No 10
  providerEmail: {type: String, required: true},           // No 11
  providerWebsite: {type: String, required: true},         // No 12
  category: [{type: String, required: true}],              // No 13
  target: [{type: String, required: true}],                // No 14
  sector: [{type: String, required: true}],                // No 15
  applied: {type: String, required: true},                 // No 16
  organic: {type: String, required:false},                 // No 17
  languageOther: {type: String, required: false},          // No 18
  title: { type: String, required: true },                 // No 19
  titleOther: { type: String, required: false },           // No 20
  desc: {type: String, required: true},                    // No 21
  descOther: {type: String, required: false},              // No 22
  keywords: {type: String, required: true},                // No 23
  keywordsOther: {type: String, required: false},          // No 24
  delivery: [{type: String, required: true}],              // No 25
  conn: {type: String, required: false},                   // No 26
  source: [{type: String, required: true}],                // No 27
```

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

```
  access: {type: String, required: true},            // No 28

  know: {type: String, required: true},             // No 29

  train: {type: String, required: true},            // No 30

  benefits: [{type: String, required: true}],         // No 31

  challenges: [{type: String, required: true}],       // No 32

  imagePath: [{type: String, required: true}],      // No 33

  documentPath: [{type: String, required: false}],   // No 34

  videos: {type: String, required: false},           // No 35

  creator: {type: mongoose.Schema.Types.ObjectId, ref: 'User', required:
true }, // No 36

  isDraft: {type: String, required: true} // No 37

});
```

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

10 Annex C

This Annex provides directions on how to deploy the data platform and initiate the software modules that load data to the platform. The directions are also available in the README.md file here: https://gitlab.com/QuantiFarm/data-platform/-/blob/main/README.md

QuantiFarm Data Plaform

This is the source code repository of the QuantiFarm Data Platform.

The main components are the following:

- `triple-store`: Triple store for storing test case and assessment data.
    - This triple store will be queried by the various *tool web services*, which are not yet included in this prototype.
- `tc11-loader`: Proof-of-concept loader (ETL) of test case data that reads data from a web service, maps it to QCSM-conform RDF, and loads it into the triple store.

It should be noted that every test case loader should write tripes to **its own graph**. This way, when source data is updated, it is trivial to replace any existing data using the SPARQL Graph Store Protocol.
Implementation status
Currently, this is a prototype, and not yet deployed anywhere.

Running the platform
To start the platform and run the loader(s), run the following:

```
docker compose up
```

Now, the services should be available at:

- http://localhost:3030/quantifarm/query
    - The triple store's endpoint for SPARQL 1.1 Query (with ARQ extensions)
    - To be used by the tool web services
- http://localhost:3030/quantifarm/update
    - The triple store's endpoint for SPARQL 1.1 Update (with ARQ extensions)
    - To be used by the loaders
- http://localhost:3030/quantifarm/data
    - The triple store's endpoint for SPARQL Graph Store Protocol and Quad extensions
    - To be used by the loaders and tool web services
- http://localhost:3030/quantifarm/get
    - The triple store's endpoint for SPARQL Graph Store Protocol and Quad extensions (read-only)
    - To be used by the tool web services

To receive data from the triple store, you can, for example, do the following:

receive and write all of the data in the TC11 graph to a Turtle file.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

```
curl
localhost:3030/quantifarm/data?graph=https://quantifarm.eu/graphs/tc
11 --header "Accept: text/turtle" > tc11-graph.ttl
```

To update the data from the loaders, they can simply be run as a Docker Compose service:

```
# this will overwrite the data that it previously wrote, if any
docker compose up tc11-loader
```

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

11 Annex D

The QuantiFarm data platform is bundled with the use of the Docker framework. The docker-compose.yaml configuration file for initiating the platform is available in QuantiFarm's gitlab (https://gitlab.com/QuantiFarm/data-platform/-/blob/main/docker-compose.yml) and also here:


```
services:
  triple-store:
      platform: linux/amd64
      image: registry.gitlab.com/quantifarm/data-platform/triple-store:0.0.2
      build:
      context: ./triple-store
      args:
      JENA_VERSION: 4.8.0
      ports:
      - "3030:3030"
      command: [ "--conf", "/fuseki/configuration/conf.ttl"]
      volumes:
      - ./fuseki-conf.ttl:/fuseki/configuration/conf.ttl:ro
      - triple-store-logs:/fuseki/logs
      - triple-store-data:/fuseki/databases
  tc11-loader:
      build:
      context: .
      dockerfile: ./tc11-loader/Dockerfile
      environment:
      GAIASENSE_URL:
      GAIASENSE_USERNAME:
      GAIASENSE_PASSWORD:
  fairshare-loader:
      build:
      context: .
      dockerfile: ./fairshare-loader/Dockerfile
      environment:
      TRIPLE_STORE_ENDPOINT:
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700

```
      volumes:
      -                ./fairshare-loader/tools-collection.json:/data/tools-
collection.json
  recommendation-tool:
      build:
      context: .
      dockerfile: ./tool-services/recommendation-tool/Dockerfile
      environment:
      SPARQL_QUERY_ENDPOINT:
      restart: unless-stopped
      ports:
      - "8000:8000"
volumes:
  triple-store-data:
  triple-store-logs:
```

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101059700