



This project has received funding from the European Union's Horizon 2022 Research & Innovation Actions – Project. 101086355 – HORIZON-CL6-2022-GOVERNANCE-01



OPEN SCIENCE AND DATA MANAGEMENT PLAN, V2.1

Updated version, 13 May 2026



Author		
Name	Organization	Draft release date
Koen Van Rossum	VITO (coordinator)	20/02/2025
Tuna Coppens	ILVO	20/02/2025
Marios Vlachos	ICCS	20/02/2025
Nuno Grosso	Deimos	20/02/2025
Panagiota Louka	NP	20/02/2025

Approval on behalf of the Executive Board		
Name	Organization	Date of approval
Maria P. Gonzalez-Dugo	IFAPA	27/02/2025
Isabelle Piccard	VITO	27/02/2025

Revision records			
Version	Date	Changes	Authors
1.0	27/09/2023	Original document	VITO
1.1	27/01/2025	Updated version of the document, outlining how feedback from the EC and external reviewers on version 1.0 of the document has been addressed.	VITO
2	28/02/2025	Updated version 2 of the document	VITO
2.1	13/05/2026	Updated version of the document, outlining how feedback from the EC and external reviewers on version 2.0 of the document has been addressed.	VITO

Acronyms and Abbreviations

Acronyms and Abbreviations	
AI	Artificial Intelligence
AgroSTAC	global repository on key agronomy observations
AIM	Agriculture Information Model
Aoi	Area of Interest
API	Application Programming Interface
BIOPAR	Biophysical Parameters
CA	Consortium Agreement
CDSE	Copernicus Data Space Ecosystem
CSA	Coordination and Support Action
CEADS	Common European Agricultural Data Space
CSV	Comma Separated Values
CSW	Catalogue Service for the Web
DA	Data Act
DES	Deimos Spain
DGA	Data Governance Act
DME	DEIMOS ENGENHARIA SA
DSS	Decision Support System
DSSC	Data Spaces Support Centre
EC	Electrical Conductivity
ECMWF	European Centre for Medium-range Weather Forecasts
EnMAP	Environmental Mapping and Analysis Program
ENVO	ENVironment Ontology
EO	Earth Observation
EOSC	European Open Science Cloud
ERP	Enterprise Resource Planning
ESA	European Space Agency
ESDAC	European Soil Data Centre
ET	EvapoTranspiration
ETa	Actual EvapoTranspiration
ETSI	European Telecommunications Standards Institute
EU	European Union
EV ILVO	Eigen Vermogen van het Instituut voor Landbouw en Visserij Onderzoek
ExBo	Executive Board
FAIR	Findable, Accessible, Interoperable, and Reusable
fAPAR	fraction of Absorbed Photosynthetically Active Radiation

Fcover	Fraction of vegetation cover
GA	Grant Agreement
GB	Giga Byte
GDPR	General Data Protection Regulation
GEOGLAM	Group on Earth Observations Global Agricultural Monitoring Initiative
GeoJSON	Geospatial Java Script Object Notation
GeoTIFF	Geographic Tagged Image File Format
GIS	Geographic Information Systems
GML	Geography Markup Language
ICCS	Institute of Communications and Computer Systems
iCM	Integrated Crop Management
IFS	Integrated Forecasting System
INSPIRE	Infrastructure for Spatial Information in Europe
IoT	Internet of Things
IPR	Intellectual Property Rights
ISO	International Organization for Standardization
JSON	Javascript Object Notation
KPI	Key Performance Indicator
LAI	Leaf Area Index
LAU	Local Administrative Unit
LUE	Light Use Efficiency
ML	Machine Learning
MSI	MultiSpectral Instrument
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NGSI	Next Generation Service Interfaces
NIR	Near Infra Red
NP	NeuroPublic
OAuth	Open Authorization protocol
OGC	Open Geospatial Consortium
OpenAIRE	Open Access Infrastructure for Research in Europe
OpenDEI	Open Digitizing European Industry
ORE	Open Research Europe
PC	Project Coordinator
PPP	Plant Protection Product
PSNC	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk
R&D	Research and Development

RC	Research Coordinator
RDF	Resource Description Framework
REST	Representational State Transfer
RIE	Research and Innovation Environment
RILab	Research and Innovation Lab
RS	Remote Sensing
SciHub	Scientific Hub
SDG	Sustainable Development Goal
SDK	Software Development Kit
SF	Smart Farming
SITRA	Finnish innovation fund
SME	Small and Medium sized Enterprise
SOC	Soil Organic Carbon
STAC	Spatio Temporal Assets Catalog
SyGMa	System for Grant Management
TRL	Technology Readiness Level
VITO	Vlaamse Instelling voor Technologische Onderzoek
VNIR	Visible and Near Infra Red
VPO	Vlaams Planbureau voor Omgeving
WP	Work Package
XML	eXtensible Markup Language

Table of Contents

1. Introduction	9
1.1. Project overview	9
1.2. Scope of the document	9
1.3. Document structure	9
1.4. Evolution of the document	10
2. Research planning, monitoring and risk management	12
2.1. Research planning	12
2.2. KPI monitoring	26
2.3. Risk management	27
3. Data summary	28
3.1. Data inputs	28
3.2. Data outputs	37
3.3. Broader context	39
4. FAIR data	42
4.1. Making data findable, including provision of metadata	43
4.2. Making data openly accessible	44
4.3. Making data interoperable	46
4.4. Increase data re-use	48
4.5. Allocation of resources	49
5. Data security	51
6. Ethical aspects	53
Annexes.....	54
Annex I: KPIs and their status	54
Annex II: Risks, mitigations and state of play	66

List of figures

Figure 1: ScaleAgData's multistep agile development approach	12
Figure 2: ScaleAgData phases in relation to the various deliverables	12
Figure 3: Functional view of the architecture of the ScaleAgData system.	17
Figure 4: DSSC coordinates development of data spaces ensuring common standards and interoperability. The ScaleAgData project, has links with the Agriculture & Green Deal data space. ...	18
Figure 5: The overall approach of the sensor data acquisition planning and sensor data collection in ScaleAgdata	19
Figure 6: Context broker and data sharing	21
Figure 7: Schematic view of model use	23
Figure 8: Overview of the workflow to package and deploy RIE developed algorithms/applications...	24
Figure 9: General Requirements of the RILabs	26
Figure 10: Sub-section of Execution Plan related to Sensor Data	26
Figure 11: First Section of Execution Plan	26
Figure 12: Taxonomy of building blocks v0.5 of the Data Spaces blueprint.....	39
Figure 13: Functional view of the architecture of a node of ScaleAgData system.....	48

List of tables

Table 1: Overview planning and monitoring of R&D activities	13
Table 2: Processes and their order of the applied co-design framework	15
Table 3: Rolling plan for the Yield Monitoring RILab (colours map technology providers to the deployment scenario/innovation area they are involved in for the RILab).....	15
Table 4: Overview of agri-environmental data products developed by the RILabs.....	22
Table 5: ScaleAgData data inputs	29
Table 6: ScaleAgData data outputs	37

1. Introduction

1.1. Project overview

ScaleAgData is a response to the call HORIZON-CL6-2022-GOVERNANCE-01-11 Upscaling (real-time) sensor data for EU-wide monitoring of production and agri-environmental conditions. The ScaleAgData project runs from January 2023 till December 2026 and consists of a consortium of twenty-six partners from fourteen countries. The vision of ScaleAgData is two-fold. On one hand, it wants to obtain insights in how the complex data streams should be governed and organized (governance call). On the other hand, it aims to develop the data technology needed to scale data collected at the farm level to regional datasets, agri-environmental monitoring, and the management of agricultural production.

To do so, ScaleAgData has five objectives:

- Developing innovative approaches for collecting in-situ data and applying data technologies.
- Enabling and promoting data sharing along the entire data value chain.
- Demonstrating how the sensor data can be scaled to agri-environmental data products at the national, regional or European level.
- Demonstrating the benefit of the improved monitoring capacities in a precision farming context.
- Demonstrating the benefit of upscaled regional datasets for the agricultural sector in general.

During its lifecycle, the project will explore seven innovation areas: innovative sensor technology, edge processing, data sharing architecture and data governance, satellite data augmentation, from data assimilation to service development, privacy-preserving technology, and data integration methodologies.

Six Research and Innovation Labs (RILabs) have been identified within the project, across various biogeographical regions of Europe, where different data upscaling and integration models or approaches will be evaluated and demonstrated. The six RILabs are: water productivity (RILab 1), crop management (RILab 2), yield monitoring (RILab 3), soil health (RILab 4), grasslands (RILab 5) and sustain dairy (RILab 6).

In the second iteration, recommendations will be formulated on how such integrated datasets can be capitalized to help national and regional policy making to strengthen both the competitiveness and sustainability of European agriculture.

1.2. Scope of the document

This Data Management Plan describes the planning and monitoring of R&D activities as well as the procedures, resources, and security rules applicable to the data assets handled by the ScaleAgData project. It also addresses the relevant aspects of making the data products FAIR (Findable, Accessible, Interoperable, and Reusable).

1.3. Document structure

This document is structured as follows:

- Chapter 1 provides a project overview and then goes on to describe the scope, responsibilities, and structure of this deliverable.

- Chapter 2 provides an overview of the research planning taking place across the various work packages
- Chapter 3 presents a summary of the data involved in the services
- Chapter 4 discusses the approach to the FAIR principles, whenever possible in the project, as well as the allocation of resources
- Chapter 5 introduces data security issues
- Chapter 6 describes the ethical aspects
- Annex I provides an overview of the KPIs and their status
- Annex II lists the project related risks, mitigations and state of play

1.4. Evolution of the document

This Open Science and Data Management Plan is intended to be a living document in which information will be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, each data management plan will have a clear version number and include a timetable for updates.

Version 1.0 of this document, submitted on 27 September 2023, showed the initial approach to data management within ScaleAgData.

In December 2024, after submission of the first interim technical report, the EC and external reviewers formulated a number of recommendations to improve the Open Science and Data Management Plan. This included, among others, adding information about the initial planning and monitoring of the research, scientific and technical progress of the project and activities related to the management of R&D aspects, including the monitoring of research objectives and KPIs, technology selection, standards adoption, scientific and research issues concerning methodologies and risk management. The requested information was provided to the extent possible in version 2.0 of this document, deliverable D1.3.

Following the review of the second periodic report, additional changes were applied to the original document D1.3, resulting in this version 2.1. An overview of the changes applied in this version 2.1 (compared to v2.0) are outlined in table below

Section	Nature of change and reason (if applicable)
2.1 – Quality assurance	For technology validation, reference is made to 2.1 Technology validation of methodological frameworks and prototypes for more details
2.1 – Co-design planning and elicitation	Updated reference to D2.2 which was delivered in M33
2.1 – Evolution requirements, validation framework and rolling plan	Updated reference to D2.2 which was delivered in M33
2.1 – Architectural design	Updated reference to D3.4 which was delivered in M41 with reference to its section 2.8 for high-level architecture per RILab
2.1 - Governance frameworks planning	Updated reference to D2.2 which was delivered in M33

Deliverable 1.3. Open science and data management plan v2.1

2.1 – Framework development and edge processing	Reference was made to deliverable D3.4 for more details on the architecture per RILab and to deliverable D3.5 chapter 5 for an overview of the edge processing enabling technologies, real time processing and privacy. Additional reference was made to upcoming deliverable “D5.4 Deployment and Evaluation Report” (due in month 46) for experiences and lessons learnt across the RILs.
2.1 – EO data management	The different types of EO data collections were described, paragraph was added on accessing the data and further reference was made to section 4 (FAIR data)
2.1 - Data Governance, Sharing Meta architecture and Integration	Updated reference to D3.4 which was delivered in M41
2.1 - Planning for development of methodological frameworks to integrate sensor data into data products	Table 4 was updated, reflecting the latest status. Reference is also made to D4.5 which was delivered in M41
2.1 - Technology validation of methodological frameworks and prototypes	The section on technology validation has been updated reflecting the project’s fit-for-purpose validation strategy, which follows a hybrid approach necessary to accommodate heterogeneous data infrastructures and domain-specific requirements, leverage existing validated environments already in use by partners.
3.1 – Data Inputs	Reference was made to where updated version of table 5 will be made available, namely D3.3 and D4.5
3.2 – Data Outputs	Reference was made to where update version of table 6 will be made available, namely in D4.5.
4 – FAIR data	RIE was further described in terms of how it operates as a flexible, hybrid infrastructure for collaborative research and prototyping, and as a cloud-based environment supporting internal development, testing, and validation within the project, providing partners with access to EO and in-situ data through a shared catalogue and Jupyter-based workflows.
4.2 – Making data accessible	Updated with descriptions on workspace organization, user access, user workflows and data access mechanisms
4.4 – Increase data re-use	Updated with information on data sharing and integration as well as public outputs
Annex 1	KPI tables were updated including measurable progress up till month 36

2. Research planning, monitoring and risk management

2.1. Research planning

The project is organized in two iterations of each two years. This multistep agile development approach of the project is illustrated in Figure 1. Each iteration has four steps:

- Phase 1: co-design the building blocks with the RILabs (WP2)
- Phase 2: development of methodological frameworks and prototypes (WP3, WP4)
- Phase 3: technology implementation and validation in a testing environment (WP3, WP4)
- Phase 4: demonstration and evaluation in the RILabs (WP5)

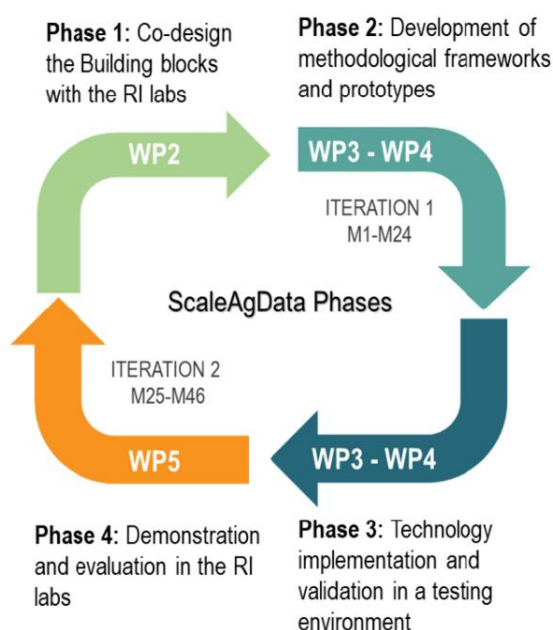


Figure 1: ScaleAgData's multistep agile development approach

The project consists of multiple work packages with tasks dedicated to design, planning and monitoring of specific components of the research within ScaleAgData.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Iteration 1	2023	KO	D1.1				D1.2, D6.1			D2.1			D3.1, D5.1
	2024		D3.2, D3.3, D4.1				D4.2, D4.3, D6.2						D5.2, D5.3, D6.4
Iteration 2	2025		D1.3				D2.2			D3.3, D3.4, D3.5			D4.4, D4.5, D6.3
	2026								D6.4, D6.5, D6.6		D5.4		D5.5

Figure 2: ScaleAgData phases in relation to the various deliverables

WP1	WP2	WP3	WP4	WP5
✓ Overall Progress monitoring (T1.2)	✓ Co-design planning and elicitation process (T2.1)	✓ Sensor selection & development & sensor data acquisition planning (T3.1)	✓ Planning for development of methodological framework to integrate sensor data into data products (T4.2)	✓ Procedures definition for monitoring, recording and progress evaluation
✓ Quality assurance (T1.1)	✓ Evolution Requirements, Validation framework & Rolling plan (T2.2)	✓ Framework development Edge processing (T3.2)	✓ Planning for development of new data assimilation methods (T4.1)	✓ Methodology development demonstrators & pilot trials (T5.1)
✓ Overall Risk management (T1.1)	✓ Architectural design (T2.3)	✓ Mapping EO data collection, processing and delivery (T3.3)	✓ Implementation of Research and Innovation Environment (RIE) with centralised data store (T4.3)	✓ RILab execution planning
	✓ Governance frameworks planning (T2.4)	✓ Data Governance, Sharing Meta architecture and Integration (T3.4)	✓ Technology validation of methodological frameworks and prototypes (T4.4)	

Table 1: Overview planning and monitoring of R&D activities

A selection of the project’s tasks related to planning and monitoring of research, scientific and technical progress is presented in Table 1, and a more detailed overview is provided here:

✓ **Overall progress monitoring**

The monitoring of the research, scientific and technical progress of the project is part of task 1.2 within WP1. This also includes activities related to the management of R&D aspects, including the monitoring of research objectives and KPIs, technology selection, standards adoption, and scientific and research issues concerning methodologies and approaches. This is supported by a series of overall consortium and WP specific meetings in which progress is monitored, challenges and risks flagged, and mitigations discussed. The regular progress monitoring meetings are:

- Bi-monthly ExBo meetings with all WP leads and RILab leads
- Monthly WP specific meetings
- Monthly RILab specific meetings
- Milestone meetings (every 6 months in first iteration, once a year in second iteration)

Progress towards the KPI targets is listed in more detail in Annex I.

✓ **Quality assurance**

Quality assurance monitoring is part of task 1.1 within WP1 and is further supported by task 2.2 within WP2 (validation framework and rolling plan), WP3, task 4.4 within WP4 (technology validation) as well as task 5.3 within WP5 (monitoring, evaluation and impact analysis). Within the ScaleAgData project, quality is secured through:

- An iterative approach that enables continuous testing, validation and refinement of methodologies and data products through pilot demonstrations. Agile development ensures multiple validation cycles, with feedback loops aligning implementations to real-world needs.
 - Data validation, ensuring the accuracy of sensors, EO, and other datasets through profiling, error detection, and deduplication before integration into RILabs. This is well described for each data product in D4.3.
 - A structured framework with KPIs, acceptance criteria, and cross-validation methods enhancing data reliability, while innovative techniques improve spatial and temporal resolution.
- ✓ Technology validation is described in more detail under “Technology validation of methodological frameworks and prototypes” later in this document

Furthermore, the project ensures interoperability and compliance with open data standards (OGC, INSPIRE) and data governance frameworks, supporting GDPR compliance. It focuses on standardized data models, data-sharing frameworks, and principles to promote seamless integration and reusability across platforms. The project adheres to structured access control to maintain data integrity and confidentiality while following best practices in data governance and interoperability. Data formats are kept consistent and aligned with industry standards, ensuring high data quality and reusability.

✓ **Overall Risk management**

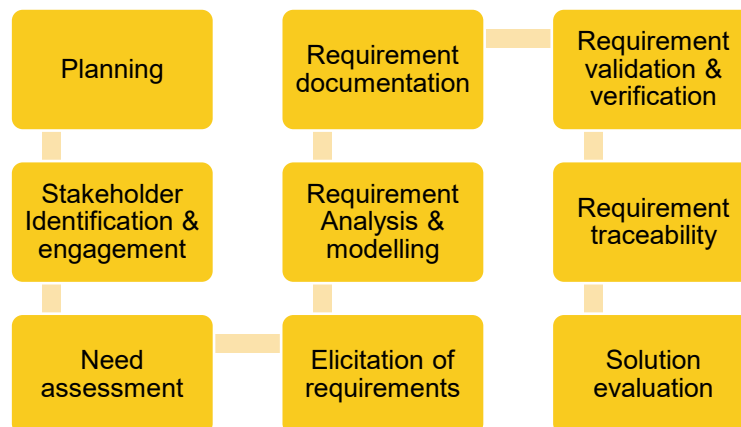
Overall risk management is part of task 1.1 within WP1, whereby regular risk assessments are conducted, addressing potential delays, technological challenges, and data integrity risks. While section 2.3 covers aspects of the risk management strategy, a detailed overview of the risks, mitigations and state of play is provided in Annex II of this document. A more regular update is provided within the continuous monitoring tool of SyGMA.

✓ **Co-design planning and elicitation process**

The co-design planning is part of task 2.1. within WP2 whereby the co-design phase is initiated by analyzing the current state of each RILab, identifying user needs, and planning and performing the elicitation process, after which findings are refined into user stories and use cases while stakeholders are engaged through workshops, interviews, and surveys. This is taking place at the beginning of each 2-year cycle iteration within ScaleAgData.

A more detailed overview of the co-design planning is available within deliverable “D2.1 Vision scenarios, requirements and innovative governance models v.1” (which was updated as D2.2 in M33).

Table 2: Processes and their order of the applied co-design framework



✓ **Evolution requirements, validation framework & rolling plan**

This task is part of WP2 whereby requirements for ScaleAgData frameworks, prototypes, and data products are defined and refined while collaborating with stakeholders. It also establishes a validation framework for monitoring results and incorporating acceptance criteria.

Additionally, a rolling plan is managed in collaboration with the Research Coordinator and WP leaders to integrate the latest advancements in sensor technology and data in the public and private domain. The rolling plan tool has been developed to identify partners, research topics and technologies impacted by changes to the intended research activities and to assess the impact of changes. Considering project objectives and available resources, proposed changes can be discussed and evaluated with all involved parties. As a dynamic document, the rolling plan continuously visualizes and reports research activities, including deployment scenarios and innovation areas, for research partners and technical providers. The work package leaders and project coordinator have, as a good practice, opted to review and evaluate the rolling plan regularly, 3-4 times per year, during the bi-monthly scheduled ExBo-meetings.

A more detailed overview is available within deliverable “D2.1 Vision scenarios, requirements and innovative governance models v.1” (which was updated as D2.2 in month 33).

Table 3: Rolling plan for the Yield Monitoring RILab (colours map technology providers to the deployment scenario/innovation area they are involved in for the RILab)

DEIMOS	EGM	KUVA	VTT	ICCS	RILab	1.1.3.1. Innovative sensor technology	1.1.3.2. Edge processing	1.1.3.3. Data sharing architecture and data governance	1.1.3.6. Privacy-preserving technology	1.1.3.7. Data integration methodologies
DEIMOS	EGM	KUVA	VTT	ICCS	Soil - AUTH	Mount hyperspectral sensors on different platforms (tractors, UAVs, robotics) to increase the mapping ability on different soil parameters	Application of edge computing to ensure high data quality while at the same time minimising the size of information transmitted to the cloud by different platforms or farmers (handheld sensors), enabling more real-time feedback	Apply innovative standardization processes in the development of soil health products that rely on satellite data as a way to increase interoperability, data sharing and reuse	Use of Federated AI to topsoil Soil Organic Carbon (SOC) model building at regional and national level	Combine sensor data with satellite images to produce an optimal estimate of soil parameters

DEIMOS	EGM	KUVA	VTT	ICCS	Soil - EV ILVO	Mount hyperspectral sensors on different platforms (tractors, UAVs, robotics) to increase the mapping ability on different soil parameters	Application of edge computing to ensure high data quality while at the same time minimising the size of information transmitted to the cloud by different platforms or farmers (handheld sensors), enabling more real time feedback	Apply innovative standardization processes in the development of soil health products that rely on satellite data as a way to increase interoperability, data sharing and reuse	Use of Federated AI to topsoil Soil Organic Carbon (SOC) model building at regional and national level	Combine sensor data with satellite images to produce an optimal estimate of soil parameters
--------	-----	------	-----	------	----------------	--	---	---	--	---

✓ Architectural design

The ScaleAgData high-level prototype architecture provides a standardized blueprint for all RILabs, enabling the development of specific applications and services. Derived from identified requirements (T2.2), vision scenarios, and use cases (T2.1), the architecture follows a structured methodology that integrates standard literature-based approaches with user-driven design from co-design workshops and partner interactions. It adopts a decentralized model where each RILab autonomously manages and processes its data, from sensor collection to transformation, while operating as an interconnected node within a larger network. This allows direct data exchange between RILabs, the Research and Innovation Environment (RIE), and external entities. Task 2.3 successfully delivered the first version of this architecture, which is being monitored in collaboration with WP5 to track implementation progress. Functional and structural diagrams further illustrate its key system layers and components, ensuring a modular, scalable, and interoperable approach to agricultural data management.

A more detailed overview is available within deliverable “D3.1 Generic architecture and data governance, sharing meta-architecture and integration of the RILabs v.1” (which was updated as D3.4 in month 41). D3.4 section 2.8 documents all the architectural implementations per RILab alongside the updated architecture of each module developed by the RILabs and technology providers.

Common European data spaces

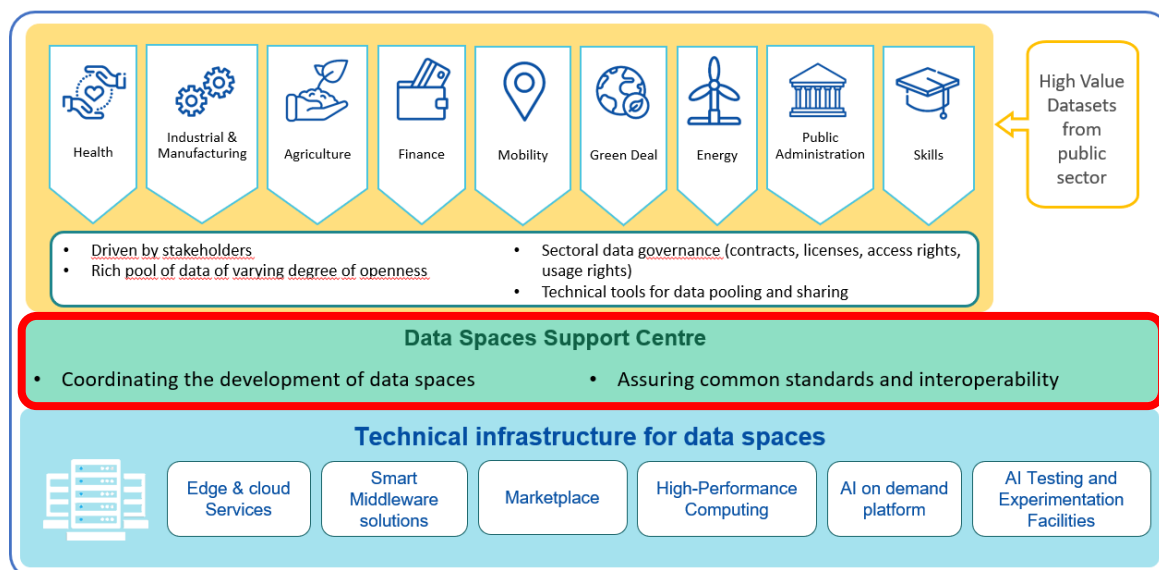


Figure 4: DSSC coordinates development of data spaces ensuring common standards and interoperability. The ScaleAgData project, has links with the Agriculture & Green Deal data space.

A more detailed overview is available within deliverable “D2.1 Vision scenarios, requirements and innovative governance models v.1” (which was updated as D2.2 in month 33).

✓ Sensor selection and development and sensor data acquisition planning

This is part of task 3.1 within WP3, whereby the set of sensors for data collection is defined, as a support to the product and service development and the RILabs. The selection of the set of sensors (provided by the consortium partners) will be based on the vision scenarios and use cases from WP2 and on the availability of the sensors and their data. This task also includes the data acquisition planning that considers the characteristics of the target services (e.g. crop development), the necessary timeliness issues of the different data sources (including EO data), and the spatial planning of the sensor data collection. A series of data quality checks on the datasets are performed aiming to preprocess and prepare the datasets for the development of the services. The acquired data and the associated metadata are directed to sensor data platforms (databases).

A more detailed overview is available within deliverable “D3.2 Sensor and edge processing selection, development, spatial planning and data collection, v.1” (which was updated as D3.5 in month 33).

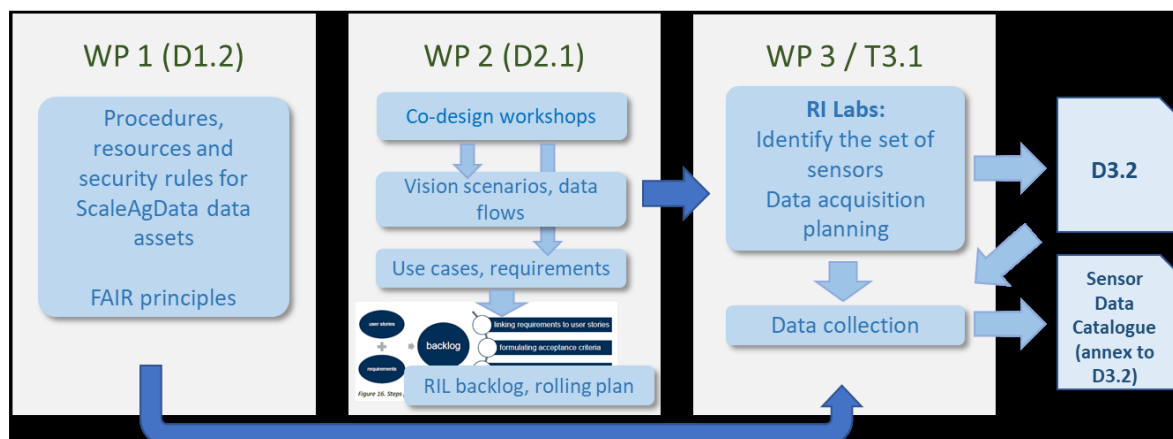


Figure 5: The overall approach of the sensor data acquisition planning and sensor data collection in ScaleAgdata

✓ Framework development edge processing

Task 3.2 under WP3 focuses on developing a framework for edge processing, enabling AI distribution and management while ensuring security and privacy through adaptive anonymization and sensitive data tagging. This includes data profiling, error detection, deduplication, cleaning, and validation, demonstrated using the EGM EdgeSpot device to connect sensors and actuators to the cloud. The EGM EdgeSpot board has been adapted into an agricultural field monitoring station called YaraSense, which features edge processing capabilities with short and long-range sensor buses (I2C and Modbus/SDI12). A second version of YaraSense is currently in development, incorporating specialized sensors such as a spectrometer, surface temperature sensor, and an eNose for soil monitoring, along with basic data quality processing. This version will be deployed at the RILab Water Productivity. Additionally, several options for firmware updates over the air are being explored, including new firmware upload mechanisms and scripting capabilities using WebAssembly. Meanwhile, the Soil Health RILab has developed an algorithm leveraging satellite images to predict topsoil organic carbon (SOC) levels in agricultural fields. To improve accuracy despite limited soil sample data, they plan to apply Federated AI, a distributed learning approach that preserves data privacy by training models across multiple local datasets instead of centralizing data. This methodology, set up using Google Cloud Platform and the Flower.ai framework, enables iterative collaborative training rounds between client-actors and a central server-actor, enhancing the SOC prediction model. Although not initially intended for Soil Health RILab experiments, this approach is highly relevant for use cases involving edge computing devices to train central algorithms while maintaining data privacy.

A more detailed overview is available in deliverable “D3.4 ScaleAgData generic architecture and data governance, sharing meta architecture and integration of RILs” concerning the architecture per RILab. A more detailed overview is also available within deliverable “D3.2 Sensor and edge processing selection, development, spatial planning and data collection, v.1” (which was updated as D3.5 in month 33). Within D3.5 chapter 5 provides an overview of the edge processing enabling technologies, real time processing and privacy. Upcoming deliverable “D5.4 Deployment and Evaluation Report” (due in month 46) will share experiences and lessons learnt across the RILs.

✓ Mapping EO data collection, processing and delivery

Providing analysis-ready EO data from Copernicus and other sources to the Research and Innovation Environment (RIE) and RILabs, tailored to specific temporal and spatial resolution needs, is part of task 3.3 within WP3. It also develops innovative methods for cloud-filling, enhancing resolution, and applying data reduction and feature selection techniques to improve information content of the EO data. The activities are driven by the needs expressed by the RILabs and their stakeholders during the co-design phase (WP2). The results are used as input for product development in WP4 and demonstration in WP5.

An overview of the EO data that the ScaleAgData partners use for the development of new data products & services in the RILabs, their sources and how the data are processed and delivered to the partners is given within deliverable “D3.3 EO data collection, spatio-temporal preparation services” (submitted in month 33 and further updated in month 41).

An overview of the EO input data used for product development can also be found in section 3.1, Table 5, of the present document.

The RIE Jupyter Notebook Environment primarily provides access to project partners to analysis-ready EO products that are in the scope of the work develop in WP3, WP4 and RI Labs so that project partners can develop their development, testing and validation activities in the project, if they don't have their own development environment.

Three main types of EO data collections are available:

- Internally hosted datasets
 - Stored in the RIE archive
 - Directly searchable and accessible through the catalogue and SDK
- Externally hosted datasets
 - Located on third-party platforms (e.g. Terrascope, Kuva Sense)
 - Accessed by users and optionally downloaded into personal RIE workspaces
 - Referenced in the RIE catalogue via metadata entries containing access links
- Externally processed datasets
 - Accessed and processed directly on platforms such as Google Earth Engine or OpenEO
 - Linked to processing scripts developed within the project
 - Referenced in the RIE catalogue via metadata and GitHub repositories

This structure allows the RIE to handle both stored data and on-demand processing workflows, depending on the use case. These datasets are only accessible internally by project partners during the development, testing and validation. Once the methodological frameworks/applications are mature a subset of the input and output datasets are made openly accessible in the ScaleAgData catalogue (<https://scaleagdata.nextgeoss.eu/>) to external users so that they can access them, understand the type of outputs generated by the project and replicate them using the project developed software components made available in the Github repository (<https://github.com/Scaleagdata>). The data is available through the CKAN powered web interface that allows users to search and discover those datasets through metadata. More details are given related to the data access in deliverables D3.3, D3.4 and D4.5 in the sections related to the RIE.

EO data management is further detailed in section 4 FAIR data.

✓ **Data Governance, Sharing Meta architecture and Integration**

Task 3.4 under WP3 focuses on developing and implementing the components of ScaleAgData’s high-level architecture, ensuring alignment with pilot needs and technical developments. The architecture integrates standardized data models, sensor data quality parameters, and governance schemes, with ICCS leading implementation, EGM handling data sharing connectors, and EV ILVO managing external API connectivity. During this period, the main building blocks for implementing the high-level architecture were identified, and some components have already been instantiated in RILabs as a common blueprint for detailed solution designs. The approach serves as an end-to-end data management system integrated in a distributed manner based on pilot requirements, incorporating defined data formats and governance schemes while ensuring semantic interoperability through standardized and extendable data models. Sensor data quality parameters such as accuracy, reliability, resolution, and timeliness are embedded as metadata. The roles of each partner in T3.4 have been clarified, and while implementing a full Dataspace is beyond the project’s scope, a roadmap for integrating relevant building blocks using Dataspace concepts has been outlined. The first implementation of the data transformation building block using the AIM ontology has been completed, and its adaptation across RILabs is supported. Additionally, the compatibility of the context broker for data sharing and the integration of data connectors using real examples were evaluated. To enable interoperable and secure data exchange, a platform based on the European open-source FIWARE ecosystem, utilizing the NGSI-LD specification from ETSI, has been developed. The functional view architecture of each RILab has been derived based on their selected components and building blocks. Data interoperability challenges are being addressed through semantic harmonization, with the AIM ontology selected as the best approach for ScaleAgData due to its agricultural sector relevance and extendibility. Initial steps toward AIM integration included translating sample RILab data into AIM and planning software translators for automated data conversion. Compliance with FAIR data governance principles has been evaluated, confirming AIM’s strong support for interoperability and reusability. The data platform is based on the Stellio solution, an open-source NGSI-LD context broker within the FIWARE Foundation, integrating mature components such as Keycloak for authentication, Grafana for visualization, and Apache NiFi for data collection. This modular, open-source approach enables flexibility and scalability, allowing for seamless adaptation to evolving needs while supporting multiple solution providers.

A more detailed overview is available within deliverable “D3.1 Generic architecture and data governance, sharing meta-architecture and integration of the RILabs v.1” (which was updated as D3.4 in month 41).

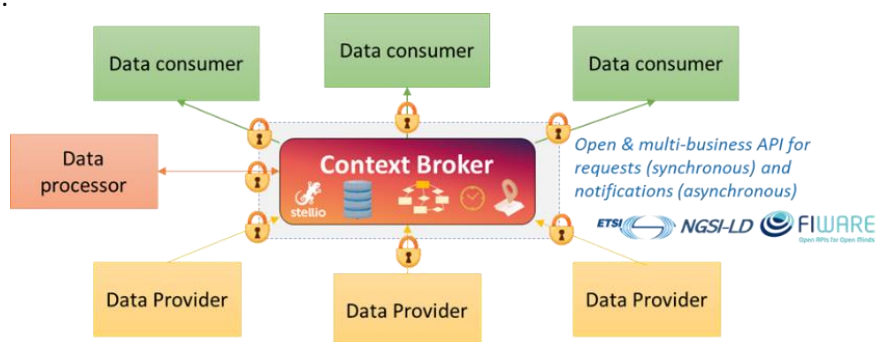


Figure 6: Context broker and data sharing

✓ **Planning for development of methodological frameworks to integrate sensor data into data products**

The development of methodological frameworks to integrate sensor data into agri-environmental data products, enhancing monitoring at local and regional scales, is part of task 4.2 within WP4. It provides strategies for RILabs and EO service providers to improve models, adjust outputs, or create new data products using limited sample data.

Within the RILabs, sensor data, EO data and other reference data that are needed for the development of new or improved agri-environmental data products have been collected. Collaborations have been set up between RILab partners and technology providers in the context of WP3, related to new sensor developments and sensor data standardization, and in WP4, to co-develop and test various methodological frameworks.

The expected agri-environmental data products are listed in Table 4 below. A detailed description of the methodological frameworks and the data products that are being developed is available within deliverable “D4.3 Sensor integrated data products v.1” (which was updated as D4.5 in month 41).

Table 4: Overview of agri-environmental data products developed by the RILabs

RIL	Data product
Water productivity	field water status and predicted yield for target crops
	satellite based field water status and predicted yield for target crops
Crop management / Agri-environmental monitoring for policy makers	soil moisture and evapotranspiration aggregation at LAU / Commune Level
	calculated indicators (aggregates) based only on the ground truth evidence
	calculated indicators based on data assimilation mechanisms along with the respective annotations
Crop management / Sustainability performance	aggregated pesticide uses for policy makers
	Calculated sustainability indicators
Crop management / Early pest detection	DSS model outputs
	statistical data on the accuracy of observations of the occurrence of agrophages
	improved predicted agrophage occurrence data based on geolocation
Yield monitoring	predicted overall level of agrophage occurrence risk for the selected region
	Potato yield estimates (subfield level)
	Improved tare yield estimates for potatoes
	Winter wheat yield estimates (LUKE)
Soil health	Winter wheat yield estimates (VITO)
	EO based regional soil organic carbon map
Grasslands	Soil health indicator estimates
	Gap-filled grasslands LAI maps at parcel level
	Estimated grassland yield at parcel level
Dairy	Improved grassland yield maps based on flux tower sensors
	Regional productivity of dairy farms
	Deviation of milk quality & quantity
	Assessment of grass yield at regional level
	Milk quality forecast on the regional level

More detailed specifications are provided in Table 6 of the present document.

The methodological frameworks and the agri-environmental data products are added to the Research & Innovation Environment (RIE) as they become available.

✓ **Planning development of new data assimilation methods**

The development of smart models to transform measurement data from various sources into actionable farming insights using digital twins is part of task 4.1 within WP4. It creates methods for automatic digital twin setup, data assimilation for process-based and ML models, and develops a framework for training intelligent agents for multi-objective decision-making.

A more detailed overview is available within deliverable “D4.2 Data-based farming services v.1” (which will be updated as D4.4 in month 36).

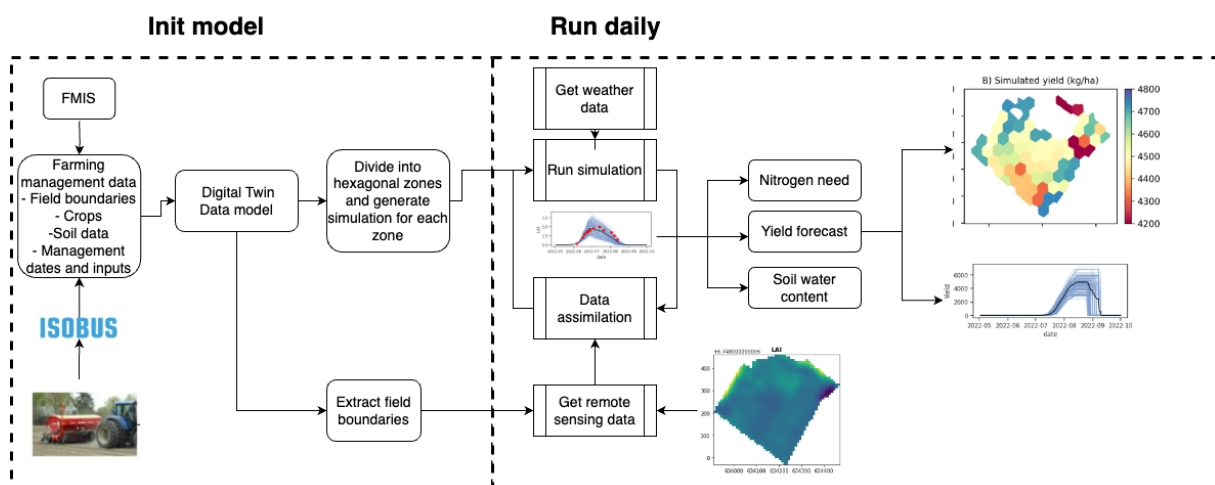


Figure 7: Schematic view of model use

✓ **Implementation of Research and Innovation Environment (RIE) with centralized data store**

This activity establishes a RIE with a centralized data store, providing input datasets to development teams via Jupyter notebooks. It uses the Deimos EO platform, services4EO, enabling collaborative EO data exploitation through standardized, customizable geospatial application modules. This is part of task 4.3 within WP4.

The overall architecture, as well as a user manual explaining how to login, how to execute Jupyter Notebooks as well as how to access the ScaleAgData catalogue is available within deliverable “D4.1 RIE”.

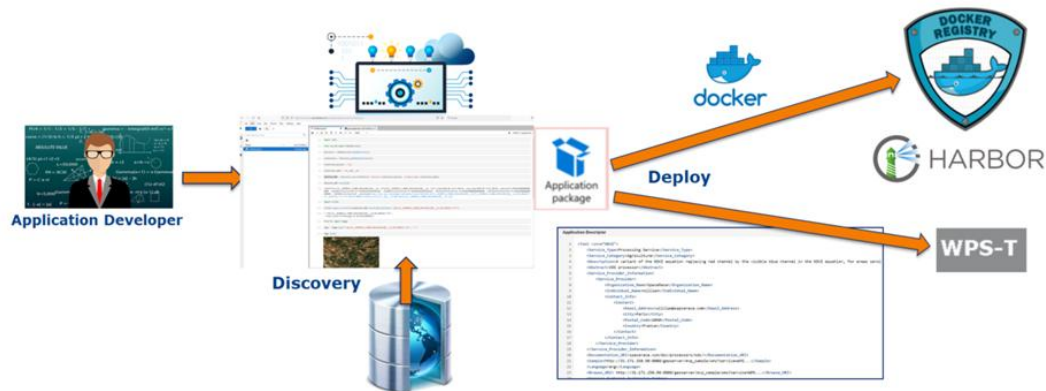


Figure 8: Overview of the workflow to package and deploy RIE developed algorithms/applications

✓ Technology validation of methodological frameworks and prototypes

Across the RILs, independent quantitative validation is systematically implemented, using domain-appropriate performance indicators depending on the specific use case. These include, among others statistical metrics such as RMSE, R^2 , bias and correlation for yield estimation and environmental variables, classification accuracy, F1-score and confusion matrices for crop mapping applications, and comparison against in-situ measurements (e.g. soil sensors, flux towers, farm records) for EO-derived products such as soil moisture, evapotranspiration, and biomass.

These validation approaches are applied independently within each RIL, using locally available datasets and ground truth data, which ensures that results are evaluated under real-world operational conditions. Given the diversity of thematic use cases, biogeographical regions and data availability across RILs, a single uniform validation protocol would not be appropriate. Instead, the project follows a fit-for-purpose validation strategy, while maintaining consistency in the use of recognised quantitative indicators.

The RIE is designed as an internal development, integration, and validation environment, supporting access to shared datasets and intermediate products, testing and validation of methodological components, EO and in-situ data workflows, and prototyping of services across partners. In practice, and in line with this design, validation activities are conducted using a combination of the RIE and partner-specific environments. This hybrid approach is necessary to accommodate heterogeneous data infrastructures and domain-specific requirements, leverage existing validated environments already in use by partners, and avoid delays that would have resulted from enforcing a single execution environment across all RILs.

The validation results are presented within deliverable “D4.3 Sensor integrated data products v.1” (which was updated as D4.5 in month 41).

✓ **Procedures definition for monitoring, recording and progress evaluation**

Effective monitoring, recording, and progress evaluation are essential components of research planning within ScaleAgData. To ensure the successful implementation of project objectives, a structured approach is adopted to define key procedures across all research and innovation activities. Regular assessments track key performance indicators (KPIs) and milestones, enabling early issue detection and corrective actions. Further details on KPI tracking and evaluation can be found in Section 2.2.

✓ **Methodology development for demonstrators & pilot trials**

This task (5.1) is part of WP5 and defines the methodology for RILabs to iteratively deploy, integrate, and evaluate demonstrators, providing guidelines and materials. It co-designs implementation plans for 10 demonstrators, detailing pilot areas, schedules, evaluation criteria, and stakeholder responsibilities.

A more detailed overview is available within deliverable “D5.1 Piloting plan & evaluation methodology v.1” for the first iteration (which was updated as D5.3 in month 24 for the second iteration).

✓ **RILab execution planning**

The execution planning, also part of WP5, consists of multiple components as outlined in D5.1:

- Areas of interest, the crops or products to be researched, and their seasonal timelines. It also includes details on pilot fields for each iteration, such as the number of fields, total area, average parcel size, and any relevant information from each RILab.
- Description of the input data for RILabs, categorized into sensor data, EO data, and other data. It covers sensor types, data provided, responsible partners, planned upgrades, sub-activities, timelines, and risks. The EO data section details sources, acquisition plans, and dependencies per iteration. The final part discusses in-situ and other input data, including their purpose, acquisition plans, and related sub-activities.
- Innovative approaches and user stories in RILabs. It details the necessary sub-activities for implementation across the first and second iterations, including dependencies, critical points, and risks. Additionally, it outlines assessment criteria and update points for each innovative approach.
- Technical providers contributing to each RILab and to the RILabs general requirements.
- Outputs of each RILab and their potential applications. It includes a brief description of each output, the responsible RILab partner, and its relevance to the deployed innovative approach and user story.
- Demonstration steps and required activities, including step descriptions, involved partners, dependencies, and risks.
- KPIs for each RILab, including a brief description, the responsible RILab partner, and the relevant innovative approach and user story. It also outlines KPI targets for the first and second iterations, along with the estimated time for achievement.

A more detailed overview is available within deliverable “D5.1 Piloting plan & evaluation methodology v.1” for the first iteration (which was updated as D5.3 in month 25 for the second iteration).

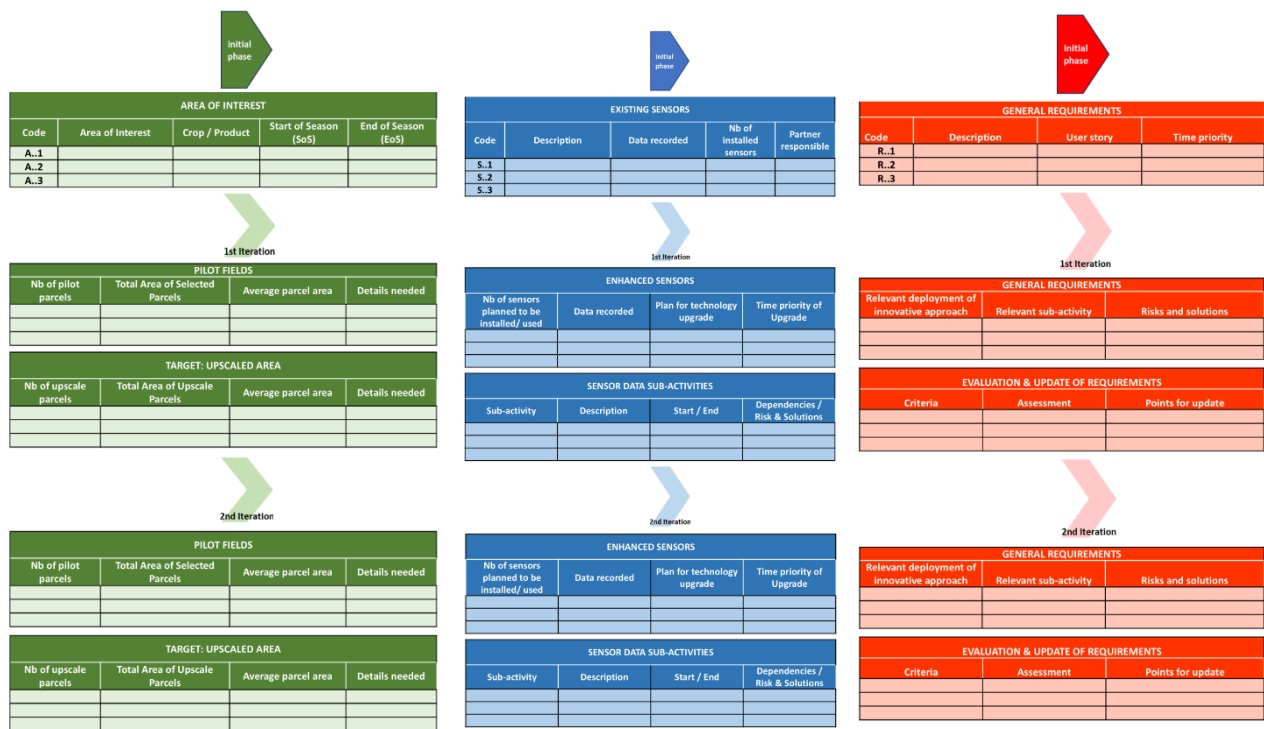


Figure 11: First Section of Execution Plan
 Figure 10: Sub-section of Execution Plan related to Sensor Data
 Figure 9: General Requirements of the RILabs

2.2. KPI monitoring

While KPI monitoring is part of WP1, discussed during bi-monthly ExBo meeting and reported within the technical reports, procedures for monitoring, KPI recording and progress evaluation within the RILabs have also been developed under WP5 (see deliverable D5.3, sections 3.3 and 3.4 as well as annex V) while progress is continuously being monitored during monthly work package meetings and RILab meetings. Regular technical reporting periods (every 18 months within ScaleAgData) help track progress, identify risks, and ensure corrective actions.

The KPIs are linked to the project’s five objectives, three outcomes and six impacts. Below you can find an overview of which WP contributes to which KPIs.

- WP2 contributes to objective indicator KPIs 01, 02, 04, 05, 06, 08, 09 and 011 as well as to impact indicator KPIs I3, I6, I9, I10, I16 and I17.
- WP3 monitors, within their coordination structure, progress towards objective indicator KPIs 03, 04, 05 and 06 and towards impact indicator KPIs I1 and I4. It further contributes to KPIs 01, 02, 07, I7, I8, I14, I15, I19 and I20.
- WP4 monitors progress with regards to objective indicator KPI 07 and impact indicator KPI I12. It further contributes to impact indicator KPIs I7, I8, I14, I15, I19 and I20.
- WP5 monitors, within their coordination structure, progress of objective indicator KPIs 01, 02, 08, 09, 010 and 011 and of impact indicator KPIs I1, I2, I3, I4, I5, I6, I7, I8, I11, I14, I15, I19, I20. It further contributes to KPI03, 04 and 07.
- WP6 contributes to objective indicator KPIs 04, 05, 06, 08 and 011, and to impact indicator KPI I3, I6, I9, I10, I13, I16 and I18. Dissemination indicators are not covered in this document.

For more details on the KPIs and their status, see Annex I.

2.3. Risk management

The ScaleAgData project incorporates a structured risk management framework to identify, assess, and mitigate potential risks that could impact its objectives. Key aspects of the risk management strategy include:

- ✓ **Risk monitoring:**
 - Regular assessments are conducted to detect potential technical, financial, and operational risks.
 - Risks are classified based on their likelihood and impact on project deliverables.
- ✓ **Risk Governance:**
 - The Project Coordinator (PC) and Research Coordinator (RC), along with Work Package (WP) leaders, oversee risk management.
 - Risks are regularly reviewed in Executive Board (ExBo) meetings to ensure proactive responses and alignment with project goals.
 - Progress is reviewed at predefined intervals (e.g., every 18 months), allowing for timely corrective actions if necessary.
- ✓ **Mitigation strategies:**
 - Project execution risks (e.g., low engagement from RILabs, unrealistic schedules, budget constraints) are addressed through continuous monitoring, adaptive work plans, and reallocation of resources.
 - Operational risks, including stakeholder coordination and data governance challenges, are addressed through clear role assignments and regular consortium meetings.
 - Technical risks are managed by implementing standardized data models, ensuring compliance with interoperability frameworks, and conducting continuous testing.
 - Technological risks (e.g., failure of key components, unexpected technical issues) are managed through iterative design, alternative solutions, and continuous assessment of emerging challenges.
 - Financial risks are mitigated through structured budget management and periodic financial reporting. As such the consortium is conducting on a 6-monthly basis internal financial reporting and on an 18-monthly basis interim financial reporting.
 - Dissemination and communication risks, and specifically the outreach to target groups, are managed through a comprehensive Dissemination, Exploitation & Communication Plan, ensuring key stakeholders are engaged and project visibility is maximized.
- ✓ **Response strategies:**
 - Financial risks due to inflation or budget misalignment are managed by reallocating resources and streamlining activities.
 - Operational risks, such as travel restrictions, are mitigated through virtual collaboration tools.
 - Technical failures, like Sentinel-1 satellite issues, are accounted for with backup solutions and flexible methodologies.
- ✓ **Data security and compliance:**
 - The project ensures adherence to GDPR and other relevant regulations to mitigate risks associated with data privacy and security.
 - Structured access control mechanisms are in place to prevent unauthorized access or data breach.

For more details on the identified risks, see Annex II.

3. Data summary

ScaleAgData is dealing with different types of data, including: EO imagery (raw, intermediate and processed data), data collected via in-situ sensors covering: soil data (i.e. soil moisture, temperature, salinity, EC, pH, clay and Soil Organic Carbon), weather data (i.e., air temperature, air humidity, wind direction and speed, solar radiation, precipitation, atmospheric pressure), machinery data (i.e., crop type, yield, dates of management practices, yield quality), key biophysical parameters (i.e. LAI, fAPAR), reflectances from cameras and spectrometers and various added value data products such as yield maps, crop phenology, crop development/field activities, growth rate, crop type, surface soil moisture, root zone soil moisture, soil maps, Eta, and methane concentration.

This chapter describes the input and output data ingested/needed by ScaleAgData as source data in the project (section 3.1 and 3.2) and discusses the broader context of data access and data sharing and related European initiatives and regulations (section 3.3).

3.1. Data inputs

Table 5: ScaleAgData data inputs provides an overview of the input data that will be collected and used by ScaleAgData to develop improved products for agri-environmental monitoring. It is a living table that will be completed / modified during the project's lifetime. An updated version will be made available in deliverables D3.3 (updated in month 41) for EO products and in deliverable D4.5 (updated in month 41) for sensor data.

Data inputs are defined in terms of:

- **Data input:** identification of the data.
- **Spatial scale:** at what geographical scale will the data be available.
- **Purpose and relation to the project:** what is the purpose of the data collection/generation and its relation to the objectives of the project?
- **Types and formats:** what types and formats of data will the project generate/collect?
- **Licensing:** licenses required and how?
- **Origin:** what is the origin of the data?
- **Size (GB):** what is the expected size of the data?
- **Data utility:** for what might it be useful?
- **Service:** service or services using the data.
- **Partner:** partner or partners involved.

Table 5: ScaleAgData data inputs

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
Sentinel-2 LAI	EU	input for crop yield estimation in RILab 3	geoTIFF, geojson	yes, available via Terrascope for all RLabs	ESA, processed by VITO	1GB/100km ²	broad range of users from agricultural sector (farmers, advisors, processors, suppliers, downstream service providers...)	Terrascope BIOPAR service	VITO
Sentinel-2 fAPAR	EU	input for crop yield estimation in RILab 3	geoTIFF, geojson	yes, available via Terrascope for all RLabs	ESA, processed by VITO	1GB/100km ²	broad range of users from agricultural sector (farmers, advisors, processors, suppliers, downstream service providers...)	Terrascope BIOPAR service	VITO
Sentinel-1 & 2 CropSAR fAPAR	EU	input for crop yield estimation in RILab 3	geojson	yes, available via Terrascope for all RLabs	ESA, processed by VITO	1GB/100km ²	broad range of users from agricultural sector (farmers, advisors, processors, suppliers, downstream service providers...)	Terrascope CropSAR service	VITO
evapotranspiration	regional	input for crop yield estimation in RILab 3	geoTIFF	yes, available via DHI	ESA, processed by DHI	<1MB/ha	broad range of users from agricultural sector (farmers, advisors, processors, suppliers, downstream service providers...)	DHI ET service	DHI
soil moisture	regional	input for crop yield estimation in RILab 3	geoTIFF	yes, available via DHI	ESA, processed by DHI	<1MB/ha	broad range of users from agricultural sector (farmers, advisors, processors, suppliers, downstream service providers...)	DHI SM service	DHI
soil texture	regional	input for crop yield estimation in RILab 3	txt	private data	CNH/UGent	10kb/field	Relationship with yield and potato tare	/	UGent
usage of pesticides and fertilizers	local	input for crop yield estimation in RILab 3	txt	private data	CNH	0,58MB/ha	Relationship with yield	/	CNH
weather data (temperature, rainfall)	EU	input for crop yield estimation in RILab 3	netCDG	ECMWF	CDS	200MB/region	Temperature, Rain, Relative Humidity, Leaf wetness data from in situ weather station	CDS	VITO

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
weather data	Local	input for DSS models, sub-lab2	txt	private data	weather stations in selected farms in North Italy	8000 kb	Temperature, Rain, Relative Humidity, Leaf wetness data from in situ weather station		Horta
potato yield sensor data	local	input for crop yield estimation in RILab 3	geojson, xls, shape file	private data	AVR harvesters	1 GB/100km2	usage in optimisation of crop rotation and field application with VRA techniques	AVR API	AVR
harvester RGB images	local	image analysis of potato and soil	JPG	private data	AVR harvesters	3GB/field		AVR API	AVR
wheat yield sensor data	local	input for crop yield estimation in RILab 3	geojson, shape file	private data	CNH harvesters	0,58MB/ha	Insight into yield variation within the field	CNH API	CNH
wheat NIR sensor data	local	input for crop yield estimation in RILab 3	geojson, shape file	private data	CNH harvesters	/	Insight into yield variation within the field	CNH API	CNH
weather station data	local	input for crop yield estimation in RILab 3	txt, csv	KMI	KMI	<1MB/ha	Temperature, Rainfall from in situ weather stations	KMI	CNH, UGent
soil scan data	local	input for crop yield estimation in RILab 3	geojson	Private data	CNH/UGent	<50MB/ha	Insight into soil variation within the field	CNH API	CNH, UGent
crop scan data	local	input for crop yield estimation in RILab 3	geojson	Private data	CNH/UGent	<50MB/ha	Insight in crop variation within the field	CNH API	CNH, UGent
soil analysis data	local	input for crop yield estimation in RILab 3	txt	private data	CNH/UGent	10kb/field	Information on soil composition, nutrients on different spots in the field	CNH API	CNH, UGent
soil data	Local	input for DSS models, sub-lab2	txt	private data	compiled in the DSS	100 kb	Texture from in situ analysis		Horta
farm log	Local	input for DSS models, sub-lab2	txt	private data	compiled in the DSS	2500 kb	Register of operations carried out in field		Horta

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
Sentinel 2	Local	input for DSS models, sub-lab2	txt	private data	external provider	100 kb	Crop monitoring by mean of remote sensing indexes (i.e. NDVI)		Horta
field characteristics	Local	input for DSS models, sub-lab2	txt	private data	compiled in the DSS	100 kb	Location (coordinates); agronomic characteristics		Horta
management of PPP	local	input for sustainability calculation, sub-lab2	txt	private data	compiled in the DSS	2500 kb	Calculate sustainability indicators, with particular reference to the reduction of risk exposure to Plant Protection Product		Horta
Sentinel-2 NDVI	Global	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (NDVI) sense-web.neuropublic.gr	ESA processed by NP's earth insight Engine	~10 KB per response	Policy Makers, Farmers, Agronomists, Scientists	NP's earth Insight service	NP
Sentinel-2 LAI	Global	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (LAI) sense-web.neuropublic.gr	ESA processed by NP's earth insight Engine	~10 KB per response	Policy Makers, Farmers, Agronomists, Scientists	NP's earth Insight service	NP
Sentinel-2 NDWI	Global	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (NDWI) sense-web.neuropublic.gr	ESA processed by NP's earth insight Engine	~10 KB per response	Policy Makers, Farmers, Agronomists, Scientists	NP's earth Insight service	NP
gaiaTron Temperature	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
gaiaTron rainfall	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
gaiaTron relative Humidity	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
gaiaTron soil moisture 10cm-60cm	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
gaiatron soil salinity 10cm-60cm	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
gaiatron windDirection	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
gaiatron windSpeed	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via gaiasense.neuropublic.gr/gaiaexmon	NP's Telemetric Station (gaiTron)	50-400 KB per parcel	Local Weather and Soil Data from farmers' parcel	NP's gaiatron Station	NP
iCM-FarmBook_Irrigation	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_Fertilization	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_sprays	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_phenological-stages	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_pests	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_harvest	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP
iCM-FarmBook_LandManagement	local	input for policy monitoring tool, sub-lab2a	geojson	yes, available via (Events) sense-web.neuropublic.gr	NP's iCM app	50-400 KB per parcel	FarmBook for each parcel	NP's iCM app (FarmBook)	NP

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
soil moisture and temperature sensor data	local	input for water productivity assessment, RILab 1	geojson	yes, available to all partners on request	Soil probes	<1 MB per region	Local IoT sensors in test fields		IES, MIGAL
weather sensor data	local	input for water productivity assessment, RILab 1	geojson	yes, available to all partners on request	Weather stations	<1 MB per region	Local IoT sensors in test fields		IES, MIGAL
yield assessment	local	input for water productivity assessment, RILab 1	geojson	yes, available to all partners on request	In situ measurements	< 1 KB per region	Measurement of pepermint and quinoa yields		IES, MIGAL
thermal data	Local	input for water productivity assessment, RILab 1	geoTIFF	yes, available to all partners on request	Airborne sensors	TBD, will be acquired during the 2nd iteration	Several airborne data acquisition campaigns using uTABI		IES, MIGAL
VNIR hyperspectral data	local	input for water productivity assessment, RILab 1	geoTIFF	yes, available to all partners on request	Airborne sensors	TBD, will be acquired during the 2 nd iteration	Several airborne data acquisition campaigns using uCASI		IES
Sentinel 2	regional	input for correlation analysis of dairy productivity and environment	geoTIFF	yes, available to all partners on request	CDSE	approx. 50 GB	analysis of productivity and precision farming at regional level		OHB
EnMAP	regional	input for correlation analysis of dairy productivity and environment	geoTIFF	yes, at moment for non-commercial use	EnMAP Portal	N.A.	analysis of productivity and precision farming at regional level		OHB
milk quantity & quality data	regional	input for correlation analysis of dairy productivity and forecasting component	json	Private data	DMK ERP system	> 5 GB per region	Laboratory analysis data and milk delivery details	Real Time Data via SFTP server	DMK, OHB

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
grass yield aggregation	regional	input for analysis of dairy productivity and environment	txt	Private data	Claas farm equipment	600 MB in test region	Harvester data		Claas, OHB
ERA-5-land reanalysis meteo parameters	regional / EU	input for correlation analysis of dairy productivity and forecasting component	netCDF/zarr	yes, available to all partners on request	CDS	27 GB for full history since 2018 Europe (hourly)	Weather model based on data from satellites, weather stations, radar, ...		OHB
ECMWF IFS forecast meteo parameters	regional / EU	input for dairy forecasting component	grib2/zarr	yes, available to all partners on request	ECMWF	10 MB per forecast	Weather model based on data from satellites, weather stations, radar, ...		OHB
grassland biomass, fPAR and LAI measurements local	local	input for grasslands RILab	geojson	yes, available to all partners on request	In situ measurements	1 Gb (2024-2025)	development of RS-based models		IFAPA
soil moisture data	local	ET models validation	geojson	yes, available to all partners on request	Soil probes	3 Mb per year	remote sensing ET model analysis		IFAPA
weather sensor data	local	input for grasslands RILab	geojson	yes, available to all partners on request	Weather stations	15 Mb per year and station	remote sensing ET and biomass model analysis		IFAPA
energy, water and carbon fluxes	local	input for grasslands RILab	geojson	yes, available to all partners on request	Flux tower measurements	65 Gb per year and tower	remote sensing ET and biomass model analysis		IFAPA
Sentinel-2 MSI L2A reflectances	regional	input for grasslands RILab	zip/SAFE, jp2	available according to the consortium agreement	ESA scihub, processed by EURAC	960 GB (2023)	estimation of grassland biophysical parameters		EURAC
Sentinel-1 backscattering	regional	input for grasslands RILab	nc	available according to the consortium agreement	ESA scihub, processed by EURAC	146 GB (2023)	development of a data fusion algorithm between Sentinel-1 and Sentinel-2		EURAC
grassland LAI and yield ground measurements		input for grasslands RILab	csv	available according to the consortium agreement	ground measurements	192.0 kB (2023-2024)	study the relationship between LAI and yield		EURAC

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
soil - bulk density content	regional/ EU	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	geotif	public available	ESDAC	250 MB	soil indicator	ESDAC-JRC request form	EV ILVO/ AUTH
soil - clay content	regional/ EU	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	geotif	public available	ESDAC	250 MB	soil indicator	ESDAC-JRC request form	EV ILVO/ AUTH
soil - coarse fragments content	regional/ EU	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	geotif	public available	ESDAC	250 MB	soil indicator	ESDAC-JRC request form	EV ILVO/ AUTH
soil - sand content	regional/ EU	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	geotif	public available	ESDAC	250 MB	soil indicator	ESDAC-JRC request form	EV ILVO/ AUTH
soil - silt content	regional/ EU	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	geotif	public available	ESDAC	250 MB	soil indicator	ESDAC-JRC request form	EV ILVO/ AUTH
Sentinel 2 and other Copernicus data	regional	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	STAC services	Open data plus the Licence for the use of the STAC services	https://stacindex.org	130 - 200 GB	satellite data		EV ILVO/ AUTH

Data input	Spatial scale	Purpose & relation to the project	Types & formats	Licencing	Origin	Size (DB)	Data utility	Service	Partner
hyperspectral EO data	regional	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	GeoTIFF	private, available to partners of the project	KUVA	~400 Mb/data cube (swath 50 km)	satellite data		EV ILVO/ AUTH
hyperspectral sensor data	local	Input of RILab Soil for soil health indicators at field level - Flemish Region and Region of Central Macedonia	npz/ tif	private, available to partners of the project	VTT	> 100 GB	spectral sensor data		EV ILVO/ AUTH
soil measurements	local	Input of RILab Soil for SOC maps- Flemish Region and Region of Central Macedonia	xls/csv	private	ILVO/AUTH	20 - 250 MB	laboratory analysis data on soil properties		EV ILVO/ AUTH
weather data	regional	Weather data collected by meteorological stations across the Poland	csv	private	Edwin	110-150MB per year (weather + rain)	temperature and humidity data to predict agrophage occurrence		PSNC/WOD R
rain data	regional	Rain data collected by meteorological stations across the Poland	csv	public	Institute of Meteorology and Water management	110-150MB per year (weather + rain)	rain data to predict agrophage occurrence		PSNC/WOD R
plants status data	local	Data from field reports about the state of plants.	json	private	Edwin	9,9 GB per year	field reports regarding plants state		PSNC/WOD R

3.2. Data outputs

In this section the data outputs are listed (see *Table 6: ScaleAgData data outputs*).

Data outputs are defined in terms of:

- **Data output:** identification of the data.
- **Types and formats:** what types and formats of data will the project generate?
- **Size (GB):** what is the expected size of the data?
- **Service:** service or services using the data.
- **Partner:** partner or partners involved.

Further updates will be provided in deliverable D4.5 (updated in month 41).

Table 6: ScaleAgData data outputs

Data output	Types & formats	Size (DB)	Service	Partner
potato yield variability maps (sub-field level)	geoTIFF, geojson	<50kb/field	Yield monitoring	VITO, AVR
potato yield estimates at field level	geojson	<10kb/field	Yield monitoring	VITO, AVR
potato yield estimates at regional level	geojson	<10kb/region	Yield monitoring	VITO, AVR
wheat yield estimates (sub-field level)	geoTIFF, geojson	<50kb/field	Yield monitoring	Luke, UGent, CNH
wheat N-uptake (sub-field level)	geoTIFF, geojson	<50kb/field	Yield monitoring	Luke, UGent, CNH
Improved tare estimates for potatoes	geoTIFF, geojson	<1MB/ha	Yield monitoring	AVR, UGent
forecasted time series of milk quality parameters	txt/json	< 10 kB/forecast & model	Dairy	OHB, DMK
regional productivity of dairy farms	txt/json	< 1MB	Dairy	OHB, DMK
gap-filled grasslands LAI maps at parcel level	geoTIFF, geojson	ca 130 KB (year 2023, 8 fields, TOT area = 23500 m ²)	Grassland	EURAC, IFAPA
estimated grassland yield at parcel level	geoTIFF,	0.25 Gb per farm per year	Grassland	IFAPA
improved grassland yield maps based on flux tower sensors	geojson	0.25 Gb per farm per year	Grassland	IFAPA, DEIMOS
regional grasslands biomass estimations	geoTIFF	0.6 Gb in test region	Grassland	IFAPA
field water status for target crops	geojson, geoTIFF	<1 MB / region	Water productivity	IES, MIGAL
satellite data-based field water status for target crops	geoTIFF	<10 MB/ region	Water productivity	IES, MIGAL
predicted yield for target crops	geojson, geoTIFF	<1 MB / region	Water productivity	IES, MIGAL
satellite data-based predicted yield for target crops	geoTIFF	<10 MB/ region	Water productivity	IES, MIGAL
EO based regional Soil Organic Carbon maps	Cloud-optimized geoTIFF	ca. 1.5 GB - 3 GB / region	Soil health	EV ILVO / AUTH
soil health indicator estimates at field level	Cloud-optimized geoTiff / geojson or equivalent vector format	ca. 300 – 700 MB / region	Soil health	EV ILVO / AUTH

soil moisture aggregation at LAU / Commune Level	geoTIFF	ca. 5 - 10 MB per region	Crop management/ Agri-environmental monitoring for policy makers	NP
calculated indicators (aggregates) based only on the ground truth evidence	json	50-400 KB per region	Crop management/ Agri-environmental monitoring for policy makers	NP
calculated indicators (LUKE) based on data assimilation mechanisms along with the respective annotations	txt	50-400 KB per region	Crop management/ Agri-environmental monitoring for policy makers	NP
aggregated pesticide uses for policy makers	txt	50-400 KB per region	Crop management/ Agri-environmental monitoring for policy makers	NP
crop classification product at LAU / Commune Level	geoTIFF	ca. 1 - 5 MB per region	Crop management/ Agri-environmental monitoring for policy makers	NP
NDVI statistics at parcel level	json	~10 KB per response	Crop management/ Agri-environmental monitoring for policy makers	NP
calculated sustainability indicators	txt	300 kb	Crop management/ Sustainability performance	Horta
DSS model outputs (fertilization rate, disease risks, crop yield and quality, soil moisture)	txt	3000 kb	Crop management/ Sustainability performance	Horta
vegetation indexes, crop nutrient status	txt	100 kb	Crop management/ Sustainability performance	Horta
statistical data on the accuracy of observations of the occurrence of agrophages	json	~13 MB/year (wielkopolska region)	Crop management/ Early pest detection	WODR/PSNC
improved predicted agrophage occurrence data based on geolocation	json	~70 MB/ year	Crop management/ Early pest detection	WODR/PSNC

predicted overall level of agrophage occurrence risk for the selected region	json	~13 MB/year	Crop management/ Early pest detection	WODR/PSNC
--	------	-------------	--	-----------

3.3. Broader context

As presented in the previous two sections, ScaleAgData partners are dealing with different types of data, including IoT data, and aiming to develop data products and services by performing transactions within loose or more organized (in terms of governance) data ecosystems. For this reason, ScaleAgData will address the issue of fair access to and use of data, accessing different elements of the technical and non-technical building blocks¹ of data spaces, in a way to facilitate data sharing and enable the creation of value from data within the vertical domains defined by the RILabs.

ScaleAgData will follow the Building Blocks Taxonomy suggested by the Data Space Support Center² (DSSC) (Figure 12), targeting the:

- The organisational governance and Regulatory Compliance within Work Package 2-Task 4 (Governance models for the vertical domains of the RILabs).
- The data models and data exchange, within Work Package 3-Task 4.
- The access and usage policy, together with the identity management within Work Package 3-Task 4 and Work Package 4-Task 3 (Research and Innovation Environment).
- Business building blocks within Work Package 6-Task 4 (IPR Management, Definition of Business Models and Policy brief).
- Data value creation building blocks (data and services within WP2, WP3, WP4 and WP5)

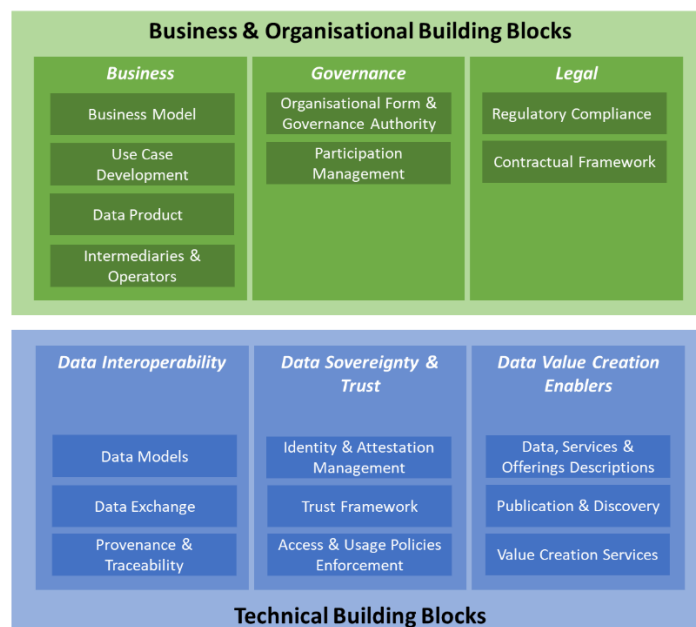


Figure 12: Taxonomy of building blocks v0.5 of the Data Spaces blueprint.

¹ Legal, Business and Governance

² Data Spaces Support Centre (dssc.eu)

ScaleAgData considers the recent EC legislation of the Data Governance Act (DGA)³, the Data Act⁴ that have as primary objectives:

- Ensure fairness in the allocation of value from data among actors.
- Foster access to and use of data.

and as sub-objectives:

- To ease the switching of providers of data processing services,
- To put in place safeguards against unlawful data transfers by cloud service providers.
- The development of interoperability standards for data to be reused between sectors.

✓ **Data governance act (DGA)**

The Data governance act creates the processes and structures to facilitate data sharing by companies, individuals and the public sector.

The Data Governance Act (DGA) aims to create a legal framework for data sharing for the benefit of the European single market, ensuring neutral access to data and interoperability and helping to avoid lock-in effects. It has three pillars. The most relevant to the ScaleAgData project is the one related to the data intermediation services, which will play a "key role in the data economy" in promoting voluntary data sharing practices between companies and facilitating the exchange of substantial amounts of data and the use of data. The data intermediation services are intended to help small and medium-sized enterprises (SMEs) and start-ups gain access to and use of the data they need. The planned data spaces aim to support this and provide data that is needed for innovation, research and the like.

✓ **Data Act (DA)**

The Data Act clarifies who can create value from data and under which conditions and introduces new rules on who can use and access data generated in the EU across all economic sectors.

The Data Act is a proposed legislation regarding harmonised rules on fair access to and use of data. The goal is to give both individuals and businesses more control over their data through reinforced portability right, copying or transferring data easily from across different services, where the data are generated through smart objects, machines, and devices (emphasis on IoT).

The Data Act specifying who, other than the manufacturer or other data holder, is entitled to access the data generated by products or related services, under which conditions and on what basis.

To align with the current developments and the convergencies related to the development of the data spaces but also with the new EC legislation, ScaleAgData will adopt, when needed, new taxonomies like the one coming from DSSC (a new effort inspired by the OpenDEI but also consider and reference the SITRA Rule book), or new terms that describe the role of each actor in the development or use of new products with data coming from sensors (Data Act). Additionally, to align with the support and coordination activities of the specific sector, ScaleAgData also monitored and followed the Common European Agricultural Data Space⁵ CSA project, AgriDataSpace, and will follow other relevant projects that aims to set up to share and access agricultural data.

³ The European regulation that aims to create a framework to facilitate European data spaces and increase trust between actors in the data market. The DGA entered into force in June 2022 and applies from Sept 2023. The DGA defines the European Data Innovation Board.

⁴ Data Act: EU institutions finalise agreement on industrial data law – EURACTIV.com

⁵ The common European data spaces is a subclass of data spaces that adheres to European rules and values. The common European data spaces were introduced in the EU data strategy and referenced in the Data Governance Act and Data Act. The Agricultural data space, it's the data space that refers to the Agricultural sector.

For the development of a Governance framework⁶ (DSSC prefers the use of framework instead of model) which is a set of principles, standards, policies (rules/regulations) and practices that apply to the governance, management, and operations within a particular scope (e.g., a data space, a data space initiative, or data spaces blueprint) as well as to the enforcement thereof, and the resolution of any conflicts, WP2 will identify, analyse, and map the current status and level of governance frameworks in the data ecosystem of the vertical domains of the RILabs.

More specific and to achieve this, T2.4 is using/will use:

- Results of T2.1 & T.2.2 (D2.1) that reveal existing legal, operational and functional agreements as well as technical standards⁷ widely adopted by RILabs actors, stakeholders and potential users of the ScaleAgData results,
- Evaluation outcomes of WP5, T5.3 (D5.3).
- Results of WP3, related to the data governance in WP3 T3.4 (D3.1, D3.4).
- Results of WP6, T6.4, as governance typically includes business aspects. (D6.4)

After the analysis of the collected material, during the second iteration of ScaleAgData, WP2, T2.4 will proceed to the initiation and tailoring of a governance framework, considering the DSSC Starter Kit⁸, DSSC Glossary and DSSC Blueprint⁹ and the results of the AgriDataSpace¹⁰ project, aiming to support the RILabs partners to realise the status of different aspects of governance within their data ecosystems¹¹, business, legal, and organisational, supporting the adoption of rules, practices and processes needed for effective data sharing and innovation within and across the vertical domains of the RILabs.

⁶ DSSC Glossary | Version 1.0 | March 2023 - Glossary - Data Spaces Support Centre

⁷ SITRA Rulebook for a fair data economy.

⁸ Starter Kit for Data Space Designers | Version 1.0 | March 2023 - Starter Kit - Data Spaces Support Centre (dssc.eu)

⁹ Public Consultation - Data Spaces Blueprint v0.5 - Blueprint (external-share.com)

¹⁰ AgriDataSpace (agridataspace-csa.eu)

¹¹ A loosely coupled set of autonomous parties engaging in data sharing

4. FAIR data

Through the RIE, the ScaleAgData project partners can have access to all data and processing infrastructure on the Deimos EO thematic exploitation solution (services4EO) during the development and testing stage of the project. It needs to be mentioned that:

- Partners already have the prerequisite infrastructure, facilities and equipment to conduct the activities envisioned by ScaleAgData.
- Partners VITO, DME/DES, ICCS, EV ILVO, PSNC, VTT, DHI and NP have access to all the cloud computing infrastructure that is needed to deploy the core datasets, services and common components of the project.
- In addition, all the partners responsible for building project demonstrators, have all the infrastructure necessary to build and deploy the ScaleAgData-powered solutions.

Therefore, the RIE functions as an optional internal development, testing and validation environment for partners who require this facility to develop their activities within the project.

Those partners that need the RIE for their work, request that the datasets that they require are ingested by the RIE catalogue/archive solution so that they can access it via the RIE Jupyter Notebook environment through a dedicated SDK. Access is exclusive for them and not shared with external users.

The analysis-ready EO products resulting from WP3, as well as data products generated by the RILabs in WP5 using the frameworks for sensor data integration developed in WP4, can be made available to other partners through the Research and Innovation Environment (RIE) catalogue established within the ScaleAgData project. The RIE is a cloud-based environment, managed by Deimos, designed to support the consortium's methodology development, integration, and validation activities.

The RIE provides a shared technical workspace where datasets, intermediate products, analytical components, and development resources are organised and made accessible to project partners, if they require. Its primary purpose is to support research and prototyping activities, rather than to operate as a production-level or public service platform.

In this context, the RIE functions as a consortium-facing development and testing environment, where EO and in-situ data can be accessed, combined, and analysed in an iterative and collaborative manner.

✓ **Cloud Infrastructure and System Architecture**

The RIE is implemented as a cloud-based infrastructure that integrates multiple components forming a flexible EO data ecosystem. The architecture can be described through the following main layers:

- Data discovery layer: a central catalogue providing metadata and access points to EO datasets
- Data storage layer: including both the internal RIE archive and external third-party platforms
- Processing layer: consisting of notebook-based environments and external processing services
- Access layer: supporting both web-based and programmatic interaction

The system supports a hybrid architecture, where data may:

- Be stored centrally within the RIE
- Remain hosted on external platforms
- Be accessed and processed remotely via scripts

This distributed setup enables flexibility in handling large EO datasets while maintaining a unified data discovery mechanism through the RIE catalogue.

✓ **Backend Software Stack and Development Environment**

The RIE provides a Jupyter Notebook-based development environment, supported by a dedicated Software Development Kit (SDK) developed by Deimos. This environment allows users to:

- Discover EO datasets via catalogue queries
- Programmatically access and download data
- Develop and test analytical workflows
- Integrate EO and in-situ data

The SDK includes structured classes such as:

- DataDiscovery for dataset search and filtering
- DataCollection and DataCollections for managing collections
- EODataset and EODatasets for dataset metadata and access

This backend setup enables reproducible, script-based data processing workflows, supporting experimentation and iterative development.

In addition to the internal tools, the RIE also interfaces with external EO processing platforms such as:

- Google Earth Engine
- OpenEO
- Terrascope

These platforms are accessed through scripts and linked via the catalogue.

Once the methodological frameworks/applications are developed and mature, a subset of both the input and output datasets will be provided in the ScaleAgData catalogue (<https://scaleagdata.nextgeoss.eu/>) to external users in order for them to understand what was developed in the scope of the project and replicate those results using the components made available in the Github repository (<https://github.com/Scaleagdata>).

More information about the RIE data access and processing architecture can be found in D4.1 and D3.4 (section RIE).

4.1. Making data findable, including provision of metadata

The findability of data is ensured through the use of Persistent Identifiers, which are crucial for unambiguously identifying datasets and facilitating data citation. These identifiers guarantee that datasets remain referenceable and accessible over time. Additionally, rich metadata descriptions, such as those following standards like Dublin Core and ISO 19115, provide essential contextual information for both human users and automated systems. By including detailed metadata, datasets become more understandable and easily discoverable, enhancing their overall usability.

The catalogue solution provided by Deimos supports multiple standard interfaces for discovering data, including CSW (Catalogue Services for the Web) and OpenSearch. These interfaces allow users to query the catalogue directly and retrieve relevant datasets efficiently. A Python library is available within the RIE, offering easy access to the discovery methods embedded in these standard interfaces. This library streamlines the interaction with the catalogue, enabling users to conduct searches more effectively. Additionally, datasets are associated with rich metadata and published within recognized repositories, making them more visible and accessible to both researchers and automated discovery tools.

Data accessibility is further enhanced by well-defined access policies and open licenses. These measures ensure that datasets can be freely used while respecting licensing terms. Access mechanisms, including APIs and standard download services, provide efficient ways to obtain datasets. The RIE also offers a Python library that simplifies access to datasets and metadata, making it easier for users to work with the data. The overall goal is to provide transparent and efficient access, fostering better data sharing and collaboration.

For interoperability, the data follows common standards such as RDF and ISO/OGC standards, ensuring compatibility across different platforms. Standardized metadata is used to describe datasets consistently, and APIs are provided for accessing and processing the data. These APIs adhere to well-established standards, facilitating integration with various tools and platforms. The RI environment includes a set of tools and services that help users process and analyze data efficiently, ensuring that data can be seamlessly integrated and used in diverse applications.

To promote data reusability, clear licensing terms and rich metadata are essential. The use of open standards and formats ensures that datasets can be reused across different research domains. Additionally, datasets are accompanied by detailed documentation, including information about their provenance and usage recommendations.

4.2. Making data openly accessible

ScaleAgData embraces both Responsible Research and Responsible Innovation to promote systemic change in the practices of the RI systems towards inclusiveness, “opening up” and sharing. The initiative aims to promote the co-development of research-based responsible innovations by including a wider spectrum of responsible innovation processes beyond just those that are research-driven. Responsible Innovation aligns with economic, social, ethical, and environmental goals, encompassing material aspects such as technologies and resources, organizational aspects including management, networking, and policy, as well as discursive dimensions like ideas and narratives. The European Commission (EC) has set the ambition of the European Open Science Cloud (EOSC) as a trusted environment for hosting and processing research data to support the EU's global leadership in science. In alignment with the principle of making research 'as open as possible,' ScaleAgData wholeheartedly supports the objectives of the Open Science Policy under Horizon Europe. Accordingly, appropriate open science actions will be implemented as an integral part of the project's methodology, as outlined below.

ScaleAgData aims to use the Open Research Europe scholarly fully open access publishing service for Horizon Europe to facilitate rapid publication times and ensure outputs that uphold research integrity, reproducibility, and transparency while enabling open science actions. Open access to all ScaleAgData publications will be ensured for all interested individuals, supporting either self-archiving ('Green' open access) or open access publishing ('Gold' open access). These publications will be made available primarily through the project's website, as well as through EU open access publishing services such as EOSC and OpenAIRE. Additionally, dissemination will extend to platforms frequently used by research communities, including ResearchGate and Academia. To enhance the accessibility and uptake of resources generated within the project, ScaleAgData will employ reference data models and provide interface-based data access and exchange according to well-established standards like those set by the Open Geospatial Consortium (OGC).

To formalize its commitment to open access, ScaleAgData has included open access clauses in both the Consortium Agreement (CA) and the Grant Agreement (GA). For every publication released under the project, a machine-readable copy of the manuscript will be made available on the project website within six months of publication. Furthermore, the project's Dissemination Leader will leverage EC open access portals and tools to ensure that project knowledge is widely accessible to the public.

Open access clauses are added to both the Consortium Agreement (CA) and the Grant Agreement (GA). For every publication, no later than 6 months after the publication is released, a machine-readable copy of the manuscript will be available on the website. Finally, and to make project knowledge publicly available, the Dissemination leader will exploit EC open access portals and tools. Open access to research data will be provided whenever it does not conflict with the protection of results. Shared data will be deposited in well-recognized open data repositories, including Open Research Europe, Zenodo, GEOSS, AgroSTAC, and the GEOGLAM in situ data repository. Additionally, whenever applicable and unprotected, open access will also be granted to the software required to validate the scientific conclusions presented in the project's publications.

The RIE is equipped with capabilities to manage access permissions for each dataset used or produced within the project. These access permissions are defined at both the project-wide level and the user-specific level, accommodating different stages of the research process. Deliverable D4.1 provides an overview of access modalities to the data as well as the data catalogue. In a nutshell, everyone with a valid user in the RIE will have access to its data catalog. Via a dedicated SDK, those users will be able to search and browse all ScaleAgData data collections. By default, those data collections will be freely accessible to all RIE users but there might be specific datasets that, due to their nature, are only available to authorized users defined by the data owner. If that is the case then the user can open a ticket in the RIE helpdesk and request access to this dataset, providing information on the reason why he/she is requesting this access. The request will be sent to the data owner, and after their evaluation, a reply will be given via the open ticket, to the user.

✓ **Workspace Organisation and User Access**

The RIE operates as a shared but flexible workspace for consortium partners:

- A common catalogue and metadata repository supports data discovery
- A shared archive stores selected datasets and products
- Users may also maintain personal working spaces for data processing

At the same time, the use of the RIE is not mandatory for all partners or workflows. Partners may:

- Work within the RIE when collaboration or shared access is beneficial
- Use their own institutional infrastructures where more appropriate

This hybrid approach ensures adaptability to diverse technical environments across the consortium. Access to the RIE is restricted to project partners, in line with its role as a controlled research and development environment.

✓ **User Workflows and Data Access Mechanisms**

Users interact with EO data in the RIE through two main entry points:

RIE Catalogue

The catalogue provides a web-based interface for dataset discovery, allowing users to:

- Search datasets using metadata filters
- Browse datasets per Research and Innovation Lab (RIL)
- Access dataset metadata, spatial footprints, and download links

Each dataset entry includes links to:

- The RIE archive (for internally hosted datasets)
- External platforms (for externally hosted datasets)
- GitHub scripts (for externally processed datasets)

Jupyter Notebook Environment

The notebook environment enables programmatic access via the SDK. Users can:

- Retrieve available data collections
- Filter datasets by parameters such as time, location, and product type
- Convert metadata into structured formats (e.g. pandas DataFrames)
- Download and process datasets locally

Typical workflow steps include:

1. Searching for datasets using SDK methods
2. Selecting datasets based on metadata
3. Accessing download links
4. Downloading and unpacking data
5. Processing data within custom scripts

In terms of Earth Observation (EO) data, ScaleAgData Task 3.3 provides analysis-ready EO data (from different sources) via the RIE (Task 4.3) to the RILabs and supporting partners (Task 5.2). This EO data is and will continue to be used for the development of new products & services in the RILabs. For an overview of all raw EO data as well as derived EO products, the EO data provider, their use, and how they are made available, please see deliverable D3.3.

4.3. Making data interoperable

The initial goal of ScaleAgData is for all partners to align their efforts to ensure that the data generated and exchanged within the project is interoperable. This alignment will facilitate the seamless exchange, sharing, and reuse of data across various research institutions, organizations, and stakeholders involved in agricultural research and innovation. Ensuring data interoperability is essential for breaking down data silos and promoting the efficient and effective use of data within the project's data ecosystem, as well as between different data spaces and domains.

To achieve interoperability, several foundational steps and strategies will be followed. Where possible, the consortium will utilize open data standards, community-agreed schemas, and well-defined data models. By leveraging established and widely adopted standards such as those developed by OGC (Open Geospatial Consortium) and other relevant bodies, the project ensures that its data will be aligned with best practices in the industry. These standards are crucial for supporting data exchange and integration across various platforms and services.

Data will be modeled using controlled vocabularies, standardized encodings, and ontologies. By aligning with existing data models like the Demeter Agriculture Information Model (AIM) and other relevant ontologies such as AGROVOC, ENVO, and SAREF4AGRI, the data becomes interoperable and easily integrated into broader ecosystems. The use of standard models ensures that the data adheres to common structures and terminologies, making it understandable and reusable across different domains. Furthermore, these data models are not static but will be extendable to accommodate additional needs that may arise as the project evolves. This flexibility allows for the addition of new data types, concepts, or fields that may emerge from the specific requirements of ScaleAgData's research, ensuring ongoing adaptability.

In cases where project-specific technical datasets are difficult to align directly with existing standards, the consortium will explore the possibility of mapping these datasets to widely recognized vocabularies and standards. Data harmonization, through mapping techniques, allows data that might not directly align with a common model to be translated into a more standardized format, thereby

ensuring interoperability. This approach is particularly useful in cases where legacy systems or specialized datasets do not natively fit into the desired standardized models.

To ensure that data can be exchanged smoothly across systems, ScaleAgData will prepare the necessary API interfaces to handle the designed data schemas. These APIs will follow standardized protocols such as REST and leverage machine-readable formats like JSON or XML for data exchange. Additionally, schema validation functionalities will be incorporated to ensure that data exchanged via APIs adheres to the required formats, making it easier to detect errors early and maintain consistency throughout the system.

In line with the principles of data discoverability, all datasets in the ScaleAgData RIE will be cataloged using standard metadata models such as Dublin Core and ISO19115. These metadata models provide a standardized way to describe datasets, making it easier to understand their content, context, and intended use. By providing detailed metadata, the data will become more findable and interpretable, thus enabling users and systems to better discover, access, and use the data. This metadata will also allow datasets to be discoverable through standard catalog interfaces, including Catalogue Services for the Web (CSW) and openSearch, which will make it easy to query the data catalog for relevant datasets. These standards ensure full interoperability with external systems that also adhere to similar metadata interfaces, broadening the scope of data exchange and enhancing the project's ability to interconnect with other initiatives and ecosystems.

Beyond the internal alignment of ScaleAgData partners, ensuring that the data generated and shared within the project can be integrated with other external systems is a priority. The adherence to common open standards, the use of metadata models, and the creation of interoperable API interfaces all contribute to this goal. By aligning with widely used data formats, standards, and metadata models, ScaleAgData ensures that its data will be usable not just within the project but also in the wider agriculture and IoT ecosystems.

In conclusion, making data interoperable in ScaleAgData is about much more than just technical compatibility. It involves a thoughtful approach to data modeling, the use of open standards, the mapping of proprietary formats to common vocabularies, and the provision of standardized metadata. By taking these steps, ScaleAgData will ensure that the data produced, shared, and reused within the project is easily discoverable, understandable, and usable across various systems and platforms, promoting long-term value and fostering collaboration across the agricultural data ecosystem.

As mentioned in previous sections of this document, all datasets, used as input or produced as output, catalogued in the RIE will follow a standard metadata model based on GML schema for EO products and will be discoverable via CSW and openSearch, assuring full interoperability with systems that use similar standard interfaces. These discovery functionalities are encapsulated in the RIE within the provided SDK, making the work of the user easier in order to access them via their Python scripts.

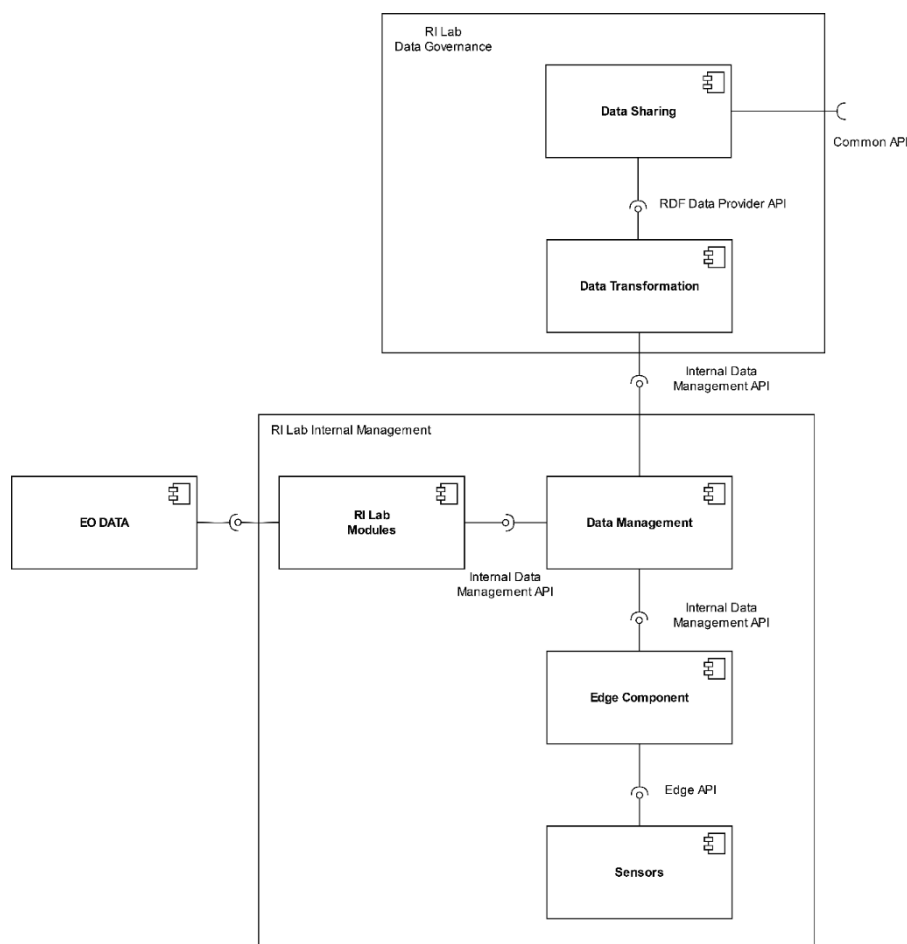


Figure 13: Functional view of the architecture of a node of ScaleAgData system

4.4. Increase data re-use

The generated data will be well-documented, and they will have clear license and provenance information. By default, open access data will be made available to the public for reuse following appropriate licensing schemes (i.e., Open Data Commons licenses). The CA will also optimize the creation of networks, interaction, exchange, and access to internal information/technologies and maximize the technology transfer.

The scripts generated by the different WP3 and WP4 partners in the RIE will be available in dedicated ScaleAgdata GIT repository available in <https://github.com/ScaleAGData>. At this stage the repository will be private but during the last phase of the project it will be open for external stakeholders to have access to those scripts and test them.

✓ Data Sharing and Integration

EO datasets used within the project may originate from:

- EO and technology providers (WP3 partners)
- Research and Innovation Labs (RILs)

To support reuse and interoperability:

- Partners are encouraged to share datasets via the RIE
- Data ingestion is managed centrally, with support from Deimos
- Metadata entries ensure consistent documentation and discoverability

The ingestion of project datasets is ongoing and is expected to be largely completed by June 2026, with continued updates focusing on final products generated by the RILs.

✓ **Public Outputs and Project Legacy**

At project completion, the publicly accessible outputs are more limited and focus on dissemination and reuse, including:

- A catalogue of input and output datasets
- A public GitHub repository containing software components, scripts, and applications

The RIE itself remains an internal environment, while the project legacy is reflected in its documented datasets, tools, and results.

4.5. Allocation of resources

The cost of making data FAIR is covered by multiple work packages.

- WP1 takes the lead in drafting the open science and data management plan and received resources to this end to have a first version at month 6 (D1.2) and an updated version later in the project at month 26 (D1.3).
- WP2, Task 2.4 contributes to fair data access and use by developing governance frameworks that support equitable data sharing within and across RIL verticals. By leveraging ScaleAgData results, other project outcomes, and EU legislation, it ensures that data transactions are structured, transparent, and aligned with regulatory and operational needs. Through the analysis and mapping of existing governance structures, WP2 identifies gaps and refines frameworks to facilitate responsible data management. In its second iteration, Task 2.4 D2.2 tailors these frameworks to contribute enhancing fairness, accessibility, and value creation within data ecosystems.
- WP3 focuses more on data governance, architecture and integration as well as the edge processing enabling technologies, real time data processing and privacy. Resources have been allocated to this work package for the development of edge processing enablers, NRT processing, data rate transmission automatic adaptation, data fusion from Copernicus and other missions, preprocessing and the creation of datasets to be used by the RILabs, as well as the development of components needed for the data technologies, and the creation of a blueprint architecture to be used in the RILabs.
- WP4 received resources to focus on the Research and Innovation environment. To this end, a centralized data store is implemented, where input datasets are made available to the techniques/product development teams, and accessible via dedicated Jupyter notebooks. This virtual lab is based in the Deimos EO thematic exploitation solution, services4EO, a collaborative EO Exploitation Platform relying on a set of core IT and geospatial-based application development modules with standardized interfaces, which can be chained together to produce a number of different services easily tailored by the platform user.
- WP 5 has resources to focus on integration and knowledge sharing. It focuses on integrating ScaleAgData services to existing applications as well as designing an approach for sharing knowledge with a larger external target audience. This will be applicable for sharing soft facts concerning experience gained and lessons learnt as well as for sharing tangible knowledge/project results, considering the FAIR principles.

- WP6 will focus both on the exploitation and capacity building of products and services, as well as on the IPR management and business models.

The consortium will handle all project's collected and generated data and required resources to ensure FAIR management and effective and secure storage. The consortium has the knowledge and tools to make data FAIR, and appropriate resources have been foreseen to cover costs for data harmonization, integration, sharing and publication. During the implementation of the project, the consortium members will collect data in various forms, i.e., photos, videos, electronic documents. For the purpose of documentation, data management and governance rules and regulations will be applied with respect to data storage and security.

5. Data security

Data security is a fundamental priority within the ScaleAgData project, as it ensures that the collected and generated data is handled with the highest level of protection and compliance with relevant legal and ethical standards. All data, both input and output, will be curated, stored, and preserved in secure data repositories that adhere to the most up-to-date quality, security, privacy standards, and applicable legislation within the European Union, particularly focusing on the General Data Protection Regulation (GDPR). Compliance with GDPR is essential for safeguarding the privacy rights of individuals and ensuring that data handling practices are transparent, accountable, and respectful of privacy concerns.

To ensure the long-term preservation and integrity of the data, the project will implement a comprehensive data security system based on best practices and industry standards. This system will serve as a reference model for data curation and preservation, establishing a secure environment for storing and managing project data throughout its lifecycle. The reference architecture will also incorporate robust access control mechanisms to ensure that only authorized individuals can access sensitive or confidential information, while data sharing will be governed by clear and transparent policies.

Access control policies within the ScaleAgData ecosystem will be designed to be versatile and fine-grained, allowing for the detailed assignment of permissions based on the specific needs and requirements of the project. These policies will enable the secure sharing of data between partners while safeguarding private or sensitive information. For instance, certain datasets may have specific restrictions or licensing conditions, which will be enforced through the use of access controls. These mechanisms will regulate access to resources by defining who can view, modify, or share data, ensuring that data sharing aligns with the partners' agreements and the project's ethical guidelines.

In addition to protecting data through access control, the project will ensure that any data containing personal data or related to privacy, as well as any intellectual property rights (IPR), will be handled with the utmost care. When such data is involved, it will only be stored in the project's secure repository if stakeholders have provided explicit consent for its inclusion. In compliance with privacy laws and ethical standards, this sensitive data will not be publicly disclosed or shared without the stakeholders' permission unless an alternative agreement has been reached. By placing strict guidelines around the management of personal data and IPR, ScaleAgData aims to create an environment of trust, where stakeholders feel confident in the handling of their data.

Furthermore, while the project's primary goal is to enable data interoperability and open sharing, the system will balance openness with necessary restrictions to ensure that privacy and data protection rights are respected. This means that the data will be made available in accordance with the stakeholders' preferences, ensuring that the privacy of individuals is preserved, and intellectual property is protected. For datasets that do not contain sensitive personal information or violate IPR, the project will explore ways to share data in open and transparent ways, adhering to the principles of open science while maintaining strong privacy protection.

As a whole, ScaleAgData's approach to data security is designed to foster a secure, compliant, and ethical data environment. The use of secure repositories, clear access control policies, and rigorous privacy protections ensures that the project will meet both the technical and regulatory requirements for data management. By integrating these security measures with the project's data sharing and collaboration goals, ScaleAgData will support the responsible and secure use of data across all partners while promoting trust and collaboration. This framework will allow for the safe, compliant exchange and reuse of data across research institutions and organizations, all while safeguarding the privacy, rights, and interests of stakeholders involved.

The RIE includes an identity management solution that provides Federated Identity Management functions that include data access authorization. Identity Management is accessible online via state-

of-the-art standards promoted via OGC innovation programmes (i.e., OAuth, Open ID Connect and User Managed Access) and a browser that provides a user-friendly interface for managing the access to user resources. As mentioned above, this component will allow certain datasets only to be accessible to a set of authorized users, to be defined by the data owner.

The secure management of project data throughout their lifecycle will be achieved by using a strong data protection strategy. In this context, the consortium partners will determine specific security controls to apply in each phase, evaluating the level of compliance during project evolution. Apart from privacy and sensitive data management, another important aspect for data security is data recovery. Thus, additional data recovery strategies will be developed for protecting data from loss that can take place due to several events (natural disasters, software bugs and infrastructure failures).

6. Ethical aspects

Within the ScaleAgData project, Artificial Intelligence (AI) will be applied for integrating sensor data in the data products on agro-environmental conditions, as well as for innovative solutions to enhance the characteristics of the input data (EO and others). The algorithms will involve AI techniques, such as machine learning (ML), for fusion of raw data, fusion between EO and in situ data as well as the development of services for the RILabs. In the past few years, ethical questions associated with the usage of ML have been the subject of academic and public scrutiny. The EC has recently published guidelines on “Requirements of Trustworthy AI”. Based on these guidelines and the active experience of the consortium researchers and experts in data technologies, remote sensing and decision-making R&D, that the ScaleAgData consortium confirms that there are no ethics concerns in the project developments.

The project will not collect human data such as self-identified ethnicity, or age from the participants. The fundamental principles outlined in the Charter of European Fundamental Rights complemented by the GDPR, such as human dignity, the integrity of the person, the protection of personal data to ensure privacy, will be fully respected and promoted in the project.

Management of personal data by the consortium members and participating stakeholders will follow GDPR guidelines. Personal data will be collected and processed only if, and to the extent, necessary. Prior to any interaction, respondents and participants of interviews will receive information on what will happen with their personal data, and what their rights are in this respect. If participants think that certain information should not be used, or if they consider some information incorrect, they can contact the designated data protection officers with their requests at the visibly marked contact e-mail address. If the participants consider that their personal information has not been handled in a correct way, they will also have the right to make a complaint.

The project will not include any research on vulnerable populations. The machine learning and artificial intelligence parts of the project will apply to remote sensing prototypes and will not apply to population monitoring. Our AI systems will not subordinate, coerce, deceive or manipulate people, and will not create attachment or stimulate addiction. All datasets and processes associated with AI & machine learning decisions will be well communicated and appropriately documented. Best possible efforts will be made to avoid unfair bias. No possible risk or harm is anticipated.

Annexes

Annex I: KPIs and their status

OBJ 1 - Developing innovative approaches for collecting in-situ data and applying data technologies by actively involving stakeholders in the agricultural production process.

KPI	WPs involved	Target	Status at M36
KPI.O1: Implementation of sensor and sampling innovations	2,3,5	All 6 RILabs	Achieved, <ul style="list-style-type: none"> – Sensor and sampling innovations deployed across all 8 sub-RILabs: IoT meteorological + soil-sensor networks and YaraSense/Phytech stations (RIL1, LV+IL) – IoT weather stations + farm logs + nanoparticle gas-sensor (RIL2a, GR) – 10 weather stations N. Italy + soil/farm-operations data (RIL2b) – eDWIN platform with weekly agrophage observations + meteorological data (RIL2c, PL) – RGB camera + gravimetric soil-moisture/EC scanning on potato harvester (RIL3, BE/NL) – ICCS-planned soil sampling (~19 ha, ~90–100 SOC samples) + UAV-VTT hyperspectral (RIL4, BE+GR) – biomass + LAI + fAPAR sampling + 4 eddy-covariance flux towers (RIL5, ES+IT) – harvester-based + OEM telemetry on dairy grasslands (RIL6, DE).
KPI.O2: Applicability of sensors: Investigation of field-based sensors in different agri-ecological zones or agri-environmental systems and investigation of hyperspectral sensor payload able for drone observation.	2,3,5	≥ 3 field-based sensors (electrochemical, optical and piezoelectric biosensors) ≥ 1 hyperspectral sensor	Achieved, <ul style="list-style-type: none"> – Field-based sensors investigated across contrasting agri-ecological zones: IoT + in-soil sensors with VNIR + thermal airborne (RIL1, LV+IL) – nanoparticle electrochemical pesticide sensor on tomato pilots Farsala (RIL2a, 2024+2025) – gravimetric soil-moisture + NDWI + EC + RGB harvester camera (RIL3) – harvester-based sensors on grassland dairy farms (RIL6). Hyperspectral payload: VTT camera deployed on UAV at two locations (Flanders + C. Macedonia, 2024+2025) and tested for soil/potato discrimination on harvester (RIL3, Aug 2025)./
KPI.O3: Data processing architecture: Testing and application of edge analytics and federated learning technology in processing platforms	3,5	1	Achieved, <ul style="list-style-type: none"> – Edge analytics: YaraSense station with on-device processing (data compression, outlier detection, OTA firmware, μPython/WASM scripting) deployed in RIL1 (LV+IL) and reused in RIL2a (REST integration pattern) – pesticide sensor treated as edge-deployable component classifying spray events (RIL2a).

			<ul style="list-style-type: none"> – Federated learning: Flower.ai framework implemented within the RIE across EV ILVO + AUTH instances for distributed SOC modelling on Flanders + C. Macedonia data without centralisation (RIL4).
--	--	--	---

OBJ 2 - Enabling and promoting data sharing along the entire data value chain

KPI	WPs involved	Target	Status at M36
<p>KPI.04: RILabs adopt and demonstrate data sharing best practices. The sensor data is readily available according to the metadata and governance models</p>	2,3,5,6	6 RILabs	<p>In progress,</p> <ul style="list-style-type: none"> – Data-sharing best practices implemented across all RILabs: shared internal repository LV↔IL (RIL1) – AIM operational backbone with anonymisation, access control and provenance, plus bilateral ICCS + LUKE exchanges (RIL2a/2b) – AIM standardisation for meteorological data (RIL2c) – CNH harvester API publicly launched Q4 2024 + AVR potato data shared with UGent/VITO/LUKE (RIL3) – Demeter AIM ontology + standardised metadata via SOC API + federated learning (RIL4) – EURAC ML model repository (MLR/RF/GPR/CatBoost) shared with IFAPA + carbon-flux datasets to DME (RIL5) – automated milk-quality data provision correlated with EO/meteorological data (RIL6).
<p>KPI.05: Implementation of a best practice document on the different governance models, with drawbacks and benefits for the different stakeholders.</p>	2,3,6	Not defined in GA; set to 1	<p>Achieved and will continue,</p> <p>Data governance is covered by D2.2, D3.1 and D3.4 (with a planned revision scheduled for 13 May 2026).</p>
<p>KPI.06: Workshop(s) to highlight the importance of data sharing can be organized.</p>	2,3,6	Not defined in GA; set to 1	<p>Achieved,</p> <ul style="list-style-type: none"> – Project-wide workshops: (a) AIM Capacity Building Webinar for RILabs (Nov 2024, NP-led) – (b) ScaleAgData Workshop on Governance Models for Data Sharing & Agricultural Data Space (19 Nov 2025) – (c) "Data Spaces in Practice" at European Synergy Days Rotterdam (Oct 2025, with CEADS / 4Growth / Data4Food) – (d) "Scaling Innovation in Agriculture through Data, AI and Federated Learning" at Synergy Days Rotterdam (Oct 2025, with GEORGIA + AGRARIAN). All RILabs participated

			– RIL5 + RIL6 also held cooperative-stakeholder data-sharing discussions during regional demonstrators
--	--	--	--

OBJ 3 - Demonstrating how the readily-accessible, locally-collected sensor data can be scaled to agri-environmental data products at the national, regional or European level.

KPI	WPs involved	Target	Status at M36
KPI.07: Development of improved or new data products that provide information on the agri-environmental conditions, are available to the different stakeholders, both within the RILabs as to other stakeholders and cover the different environmental axes (water, air, soil and living organisms)	3,4,5	≥ 10 data products	In progress, <ul style="list-style-type: none"> – 4 water-status/yield products at field + satellite level (RIL1) – agri-data L1→L2→L3 aggregation tool + pesticide-detection product (RIL2a) – Horta grano.net® DSS with new VRF functionality (RIL2b) – monitoring dashboard combining disease models + EO NDVI + precipitation indicators + field observations (RIL2c) – 4 RIL3 products (potato yield via PRESTO – winter wheat yield via LUKE/UGent/CNH digital twin – winter wheat N-uptake – potato tare via RGB harvester) – regional federated-AI SOC product + field-scale UAV SOC product (RIL4) – gap-filled LAI Bolzano + parcel-level grassland yield + flux-tower yield maps + NPP Pedroches (RIL5) – regional dairy productivity + milk Q&Q deviation + new milk-quality forecast (RIL6). Various stages of validation finalization expected Q2–Q4 2026.

OBJ 4 – Demonstrating the benefit of the improved monitoring capacities of the agri-environmental conditions in a precision farming context

KPI	WPs involved	Target	Status at M36
KPI.08: The improved data products of RILabs are used to optimize the services to farmers, providing information or a service that is not	2,5,6	≥ 5 RILabs	In progress, <ul style="list-style-type: none"> – Improved data products optimise farmer services in all RILabs: irrigation-decision support in peppermint/quinoa via threshold-based + ET-driven water-deficit modelling using sensor + EO fusion (RIL1) – evidence-based farm and regional indicators consumed by cooperatives + advisors + policy

feasible with the in-situ sensors alone.			<p>makers, plus automated spray-event detection (RIL2a)</p> <ul style="list-style-type: none"> – variable-rate N application via EO-NDVI integration in Horta DSS (RIL2b) – advisor dashboard integrating disease models + EO + field observations (RIL2c) – harvester-yield maps + sub-parcel yield/N-uptake forecasts + tare-correction in wet harvest (RIL3) – federated SOC prediction across distributed datasets without centralisation (RIL4) – grassland productivity + kg/ha DM → Animal Units conversion for sustainable grazing (RIL5) – milk-quality forecasting demonstrator UI for cooperative planning across 4,500+ farms (RIL6). Various stages of progress in ≥ 5 RILabs
KPI.O9: The improved data products can be used in a precision farming context outside of the RILabs, where the farmer(s) don't have access to sensor technology.	2,5	≥ 1 demonstration	<p>Achieved,</p> <ul style="list-style-type: none"> – RIL3: PRESTO few-shot learning + EO-based ML framework generates wheat and potato yield maps for any field — including farmers without harvester yield sensors – tare-weight correction extends use to wet-harvest conditions – LUKE digital twin demonstrated for winter wheat N-uptake and yield – supports VRT diagnosis of recurrent low/high-yielding zones. RIL4: federated AI for regional and national SOC mapping enables monitoring outside pilots without centralised data sharing or local sensor infrastructure. – Further demonstrations planned in the final year

OBJ 5 – Demonstrating the benefit of upscaled regional datasets for the agricultural sector in general

KPI	WPs involved	Target	Status at M36
KPI.O10: The improved data products are evaluated by different users (both from within as outside the consortium) on their relevance and useability	5	min. 4	<p>In progress,</p> <ul style="list-style-type: none"> – External evaluation across RILabs: Crop Lab Q1 2025 online showcase 40+ participants combining RIL2a/2b/2c (RIL2) – preliminary potato + wheat + tare products presented at Yield Demo Feb 2025 with positive farmer/advisor/researcher feedback, ECPA Barcelona June 2025 (RIL3) – Q1 2025 federated SOC demo + interactive stakeholder session covering research, farming, agribusiness, policy, technology (RIL4) – Feb 2025 Grasslands Productivity demos in Italy and Spain with regional stakeholder feedback (RIL5) – cooperative-department evaluation of milk-quality demonstrator UI (RIL6). Feedback collected via Mentimeter + surveys + structured discussions. Final iteration products

			will be presented in 2026 with named external evaluators per RILab
KPI.O11: Number and feedback of the project's co-engagement activities such as workshops, webinars with national or EU policy makers, or with other liaised projects or initiatives such as the "Partnership of Agriculture of Data"	2,5,6	Not defined in GA; set at 15	<p>Achieved</p> <ul style="list-style-type: none"> – Project-wide: HE-funded MoUs with AgriDataValue + CrackSense (2023–2024) – meetings with OpenAgri + Foodscaleshub (Dec 2024), Farmtopia (Mar 2025) – SmartAgriHubs Synergy Days (Oct 2024 Barcelona + Oct 2025 Rotterdam) – AgData Partnership launch + kick-off (Sept–Oct 2025) – GEOGLAM/EuroGEOSS contributions – EGU 2025 Vienna – EUROGEO workshops 2023 + Krakow 2024 + The Hague 2025 – Data Spaces Symposium Warsaw (Mar 2025) – ADRF25 Stavanger (Sept 2025). Per RILab: Agro-Israel 2025 lecture + on-farm demos (RIL1) – GeoAI Camp Prague + QuantiFarm Piacenza Oct 2025 (RIL2) – Yield Demo Feb 2025 + ECPA Barcelona + GEO ODOK China Sept 2024 (RIL3) – Mission Soil week Brussels Nov 2025 + IGARSS 2024 + EBDVF Budapest Oct 2024 (RIL4). – LPS June 2025 + Grasslands demos Italy/Spain Feb 2025 + RAQRS VII Valencia Sept 2024 (RIL5). – AgGateway Berlin Mar 2025 + Agritechnica Hannover Nov 2025 (RIL6).

OUTCOME 1 – Strengthening capacities for smart farming, and thus to enhance the environmental and economic performance of the agricultural sector

KPI	WP involved	Target	Status at M36
KPI.I1: number of newly developed Smart Farming solutions.	3, 5	8	Achieved Eight SF solutions delivered: <ul style="list-style-type: none"> – (1) IoT drought/irrigation monitoring + (2) EO-based upscaling of soil moisture and ET (DHI) for peppermint LV + quinoa IL (RIL1). – (3) agri-data L1→L2→L3 aggregation tool + (4) nanoparticle pesticide-detection tool with spray-event validation (RIL2a) – (5) harvester yield estimates incl. CMAAS digital twin + (6) tare estimates (RIL3) – (7) EO-based regional SOC mapping service via federated learning at 10 m for Flanders + C. Macedonia (RIL4) – (8) grassland biomass product fusing weather + biophysical sensors + Sentinel-1/2 + flux-tower data, supporting drought-adaptation recommendations (RIL5).
KPI.I2: Number of regions where those solutions were tested.	5	10	Achieved <ul style="list-style-type: none"> – 11 regions across 5 countries: Latvia + Israel (RIL1, 2 distant agri-ecological zones) – 3 NUTS-1 regions across Crete, Thessaly and Central Macedonia covering 5 NUTS-2, 7 NUTS-3 and 8 commune-level areas (RIL2a) – Belgium + Netherlands (RIL3) – Flanders + Central Macedonia (RIL4) – Pedroches Andalucía + Bolzano/Trento provinces (RIL5).
KPI.I3: Number of engaged possible users for all SF related data products.	2, 5, 6	16	In progress <ul style="list-style-type: none"> – 9 engaged users across SF-relevant RILabs: SIA Field and Forest (LV, organic peppermint farming) + Israeli quinoa enterprise (RIL1). – 3 Crete/Thessaly farmer cooperatives + advisors (RIL2a). – 2 interested users from Feb 2025 Yield demonstration event (RIL3) – Flemish Department of Agriculture + MARVIC and CAFAMORE carbon-credit research projects (RIL4). Engagement continues – final-year demonstration round in 2026 will close the gap to target.

OUTCOME 2 – Strengthening capacities for agri-environmental (climate) monitoring, particularly of soil and crop conditions

KPI	WPs involved	Target	Status at M36
-----	--------------	--------	---------------

<p>KPI.I4: Number of agri-environmental soil and crop parameters covered.</p>	<p>3, 5</p>	<p>11</p>	<p>Achieved</p> <ul style="list-style-type: none"> – 11 parameters: soil water status + yield (RIL1) – Sentinel-2 vegetation indices NDVI/NDRE/LAI/NDWI/PSRI + S2-derived crop classification (RIL2a), – wheat fertilisation indicator from weather + soil data (RIL2b), – pest-risk from field observations + meteorological + S2 indices (RIL2c) – yield (RIL3) – Soil Organic Carbon at regional + field scale (RIL4) – LAI + fAPAR with derived NPP (RIL5) – temperature (RIL6)
<p>KPI.I5: Number of regions where these solutions will be tested.</p>	<p>5</p>	<p>11</p>	<p>Achieved</p> <ul style="list-style-type: none"> – 16 regions tested: Latvia + Israel (RIL1) – Crete + Thessaly + Central Macedonia across NUTS-1/2/3 (RIL2a) – part of Emilia-Romagna N. Italy (RIL2b) – Poland (RIL2c) – Belgium + Netherlands + France out-of-calibration validation (RIL3) – Flanders + Central Macedonia (RIL4) – Pedroches Andalucía + Bolzano/Trento provinces (RIL5) – Cuxhaven/Stade N. Germany + North-Rhine-Westphalia W. Germany (RIL6).
<p>KPI.I6: Number of engaged possible users for each soil and crop related data product.</p>	<p>2, 5, 6</p>	<p>22</p>	<p>Achieved</p> <ul style="list-style-type: none"> – Engaged users across soil/crop products include: SIA Field and Forest LV + Israeli quinoa partner (RIL1) – Crete/Thessaly cooperatives + advisors (RIL2a) – advisor users on WODR Wielkopolska reference fields via eDWIN platform (RIL2c) – 2 interested users from Feb 2025 Yield demo (RIL3) – Flemish Dept of Agriculture + ESDAC + MARVIC + CAFAMORE researchers (RIL4) – 4 Pedroches pilot farms + COVAP cooperative + CONDIFESA Bolzano consortium + 4 NE-Italy test farms + 147 insured forage-production farms via index-based insurance in 2025 (RIL5) – dairy-cooperative department users (RIL6)

OUTCOME 3 – Provision of inputs to the work of the Horizon Europe candidate partnership “Agriculture of Data” and the potential Research & Innovation mission on soil health

KPI	WPs involved	Target	Status at M36
KPI.I7: Number of project agriculture data products that are of interest for the “Agriculture of Data” partnership.	3, 4, 5	14	In progress <ul style="list-style-type: none"> – 4 products mapped to AgData partnership at M36: potato yield + winter wheat yield estimates (RIL3) – gap-filled grasslands LAI + NPP products providing actionable productivity proxies for farm management, resource efficiency and evidence-based policy (RIL5). Mapping of remaining products in progress – AgData Partnership formally kicked off in Oct 2025, completion of mapping planned for the final period.
KPI.I8: Number of developed services and data that are of interest to the Research and Innovation mission on soil health.	3, 4, 5	3	Achieved <ul style="list-style-type: none"> – 3 services aligned with Mission "A Soil Deal for Europe": yield variability maps relevant for soil-status diagnosis (RIL3) – federated-learning SOC modelling + privacy-preserving governance + standardisation, aligned with the Soil Health Data Space – equally relevant for Carbon Credit market (RIL4) – grassland productivity maps providing spatial/temporal soil-functioning indicators for degradation detection, recovery monitoring and sustainable land management (RIL5).

IMPACT 1 – Innovative governance models enabling sustainability and resilience notably to achieve better informed decision-making processes, societal engagement and innovation

KPI	WPs involved	Target	Status at M36
KPI.I9: Number of defined governance and business models.	2, 6	3	In progress, Initial 6 KERs (one per RILab) narrowed down to 2 high-potential Key Exploitable Results through the Horizon Results Booster service. Comprehensive business models under development with Exploitation Intentions, Market Definition, Value Proposition and Risk Assessment defined. Business models will be presented in D6.5 (M44).
KPI.I10: Number of stakeholders involved in the definition of the governance and business models.	2, 6	40	More than 40 stakeholders engaged: 2 farming companies (RIL1) <ul style="list-style-type: none"> – ~16 stakeholders via the eDWIN advisor platform (RIL2c) – harvester partners AVR + CNH (RIL3) – cooperative departments + carbon-credit research projects (RIL4) – dairy cooperative departments (RIL6) – plus stakeholders from the 25 co-design workshops (KPI.I17), the Nov 2024 AIM Webinar and the Nov 2025 Governance Models workshop.

IMPACT 2 – Green Deal related domains benefit from further deployment and exploitation of Environmental Observation data and products

KPI	WPs involved	Target	Status at M36
KPI.I11: Number of Green Deal objectives and actions contributed to.	5	3	7, achieved 7 Green Deal objectives/actions contributed to: (i) Farm to Fork sustainability monitoring via EO regional indicators (pesticide/water/yield intensity) <ul style="list-style-type: none"> – (ii) CAP eco-schemes via LAU/NUTS-tagged L3 aggregates with confidence flags – (iii) Green Deal food security/resilience via scalable agri-env monitoring from farm logs + EO (RIL2a) – fertilisation optimisation (RIL3) – SOC monitoring supporting climate-neutral EU targets via Mission Soil + Carbon Credit market – federated interoperable soil-health products aligned with reduced-input VRF (RIL4) – food security via regional dairy productivity + milk-quality forecasting (RIL6).

IMPACT 3 – A strengthened Global Earth Observation System of Systems (GEOSS)

KPI	WPs involved	Target	Status at M36
KPI.I12: Number of project-specific data collections available in NextGEOSS	4	14	4 data collections (RIL1) registered in the ScaleAgData Catalogue (built as a NextGEOSS extension at scaleagdata.nextgeoss.eu): <ul style="list-style-type: none"> – Sentinel 2/3 Evapotranspiration – Sentinel 2/3 Soil Moisture – Sentinel 2/3 Soil Moisture – Sentinel 2/3 Derived Crop CP (RIL1) More data will be added in the NextGEOSS in the future, aiming to have a subset of data from all products of all RILabs.
KPI.I13: Number of engaged GEO activities that are testing or providing feedback on project outputs.	6	n.d. (1)	GEO Symposium and Open Data & Open Knowledge (ODOK) workshop in China, Sept 2024 — sensor-data-enhanced yield-estimation methodologies presented (RIL3). ScaleAgData also contributes to GEOGLAM Essential Agricultural Variables and Data Coordination working groups, and participated in EUROGEO 2025.

IMPACT 4 – Sustainability performance and competitiveness in the domains covered by Cluster 6 are enhanced through further deployment of digital and data technologies as key enablers

KPI	WPs involved	Target	Status at M36
KPI.I14: Number of Cluster 6 areas of intervention contributed to.	3, 4, 5	5	5, achieved Contributes to 5 Cluster 6 areas of intervention (Food Systems, Soil Health & Food, Environmental Observation, Agriculture, Forestry & Rural Areas, and European Partnerships).
KPI.I15: Number of Cluster 6 related policies and strategies contributed to.	3, 4, 5	3	5, Achieved Contributes to 5 key EU policies and strategies (Common Agricultural Policy, European Green Deal, Farm to Fork Strategy, Mission “A Soil Deal for Europe”, and the European Climate Pact).

IMPACT 5 – More informed and engaged stakeholders and end users including primary producers and consumers thanks to effective platforms such as Agriculture Knowledge and Innovation Systems (AKIS)

KPI	WPs involved	Target	Status at M36
KPI.I16: Number of engaged users.	6, 2	40	<ul style="list-style-type: none"> – More than 40 users engaged across project demonstrations and workshops: 2 farming companies LV+IL (RIL1) – 2 farmers from Feb 2025 Yield demo (RIL3) – COVAP cooperative + CONDIFESA Bolzano consortium (RIL5) – dairy-cooperative department users (RIL6) – plus Crop Lab Q1 2025 showcase (40+ participants) and participants from the 25 co-design workshops, AIM webinar, Governance workshops and Synergy Days events.
KPI.I17: Number of co-design workshops.	2	12	<p>25, achieved</p> <p>25 co-design workshops conducted M1–M36, each (sub-)RILab participating in at least 3:</p> <ul style="list-style-type: none"> – (a) three co-design series in Iteration 1 with internal RILab focus – (b) Iteration-2 internal review workshops Apr–May 2025 with all RILabs and technology provider – (c) targeted stakeholder demonstration workshops Q1–Q3 2025 (Yield demo Feb 2025, Grasslands Productivity demos Feb 2025, Crop Lab RIL2 showcase Q1 2025 plus other lab-specific demos).
KPI.I18: Number of capacity building webinars and workshops.	6	6	<p>6, achieved</p> <ul style="list-style-type: none"> – 6 capacity-building events: 2 finance & insurance webinars (Jan + Feb 2025) – 2 internal WP3/WP4 results webinars (Sept 2024) – 2 workshops at European Synergy Days Rotterdam (Nov 2025) – "Data Spaces in Practice" + "Scaling Innovation in Agriculture through Data, AI and Federated Learning".

IMPACT 6 – Strengthened EU and international science-policy interfaces to achieve the Sustainable Development Goals (SDGs)

KPI	WPs involved	Target	Status at M36
KPI.I19: Number of SDG indicators the project is directly contributing to.	3, 4, 5	5	5, achieved Directly contributes to 5 SDG indicators spanning 5 SDGs (SDG 2: Zero Hunger, SDG 9: Industry & Innovation, SDG 12: Responsible Consumption [Indicators 12.2.1 & 12.3.1], SDG 13: Climate Action, and SDG 15: Life on Land).
KPI.I20: Number of SDG targets the project is indirectly contributing to.	3, 4, 5	2	7, achieved Indirectly contributes to 8 SDG targets spanning 7 SDGs (SDG 1: No Poverty, SDG 2.4: Sustainable Food Production, SDG 3: Good Health, SDG 6.4: Water-Use Efficiency, SDG 12.4: Chemical Management, SDG 13.2: Climate Action, and SDG 15.3: Land Degradation Neutrality).

Annex II: Risks, mitigations and state of play

Project Execution Risks – Part A					
RISK	Low involvement of RILabs members and inadequate progress of specific RILabs	Identification of unrealistic schedules	Losing critical partners at crucial point of the project	Budget inappropriately assigned to support tasks, especially related to RILabs	Travel restrictions due to Covid-19.
WP	WP 1,2,3,4,5,6	WP 1	WP 1,2,3,4,5,6	WP 1	WP 1,2,3,4,5,6
MITIGATION MEASURES	Enrolment of RILabs members has been performed considering, among others, their high interest in the project work and their ability to cover all project aspects. Continuous monitoring of the activity and engagement of RILabs members is foreseen in tasks T2.1-T2.2-T4.4-T5.1 and T5.3. Activation of the networks of the partners in case a RILabs member does not progress adequately.	Continuous monitoring ensures that possible delays are recognized in early stages, ensuring timely and effective implementation of necessary corrections in the work plan.	Given the good reputation of all consortium partners, this possibility is considered to be very unlikely. The General Assembly would decide whether the uncovered project activities can be carried-out by other partners. If not, then another partner would be recruited.	Re-distribution of budget according to importance for the project: re-negotiating the value-for-money with WP leaders/partners in order to achieve a better-balanced output against payment.	All project activities have been designed to be able to take place with minimal traveling to other countries, building on lessons learnt since the pandemic started.

STATE OF PLAY	<p>Frequent RILab meetings are taking place as well as progress reporting within the RILabs coordinated by the WP5 lead.</p> <p>Through the rolling plan, matching/use of technologies by the RILabs is being monitored.</p> <p>Milestone Meetings also allow monitoring the state of play of the RILabs and their collaborations with technology providers.</p>	<p>Organization of ExBo meetings (first monthly and now bi-monthly) to monitor progress and ensure deliverable are prepared and timely submitted</p>	<p>To avoid losing Deimos UK's expertise in federated learning, firstly a consortium agreement amendment took place to facilitate inclusion of DMU in the project. When the door ended up being completely closed due to lack of perspective for UKRI funding for DMU, the other Deimos affiliates took over the role of DMU and a Grant Agreement amendment further facilitated this. This risk could also materialize for UGent in the next reporting period (in relation to the Israel – Gaza conflict). Discussions have already taken place with the PO on various options available to minimize the impact should this materialize.</p>	<p>Regular meetings are being held with the RILabs, such as the bi-monthly ExBo meeting, the WP5 meeting as well as RILab specific meetings. In addition, every 6 months internal financial monitoring is taking place for entire consortium. No request for re-distribution of budget has been made by any partner yet but should this materialize, it would be consolidated in an amendment in the 3rd year of the project.</p>	<p>The amount of in-person scheduled meetings in the project lifecycle has been minimized (3 in total). In addition, the milestone meeting and review meeting are being clustered together when possible.</p>
Project Execution Risks – Part B					
RISK	<p>Not a steady adaptation of the work schedule of the project</p>	<p>Deployment and demonstration activities don't lead to expected results to performance</p>	<p>Insufficient budgets available due to high inflation rates</p>	<p>Possible delivery delays - due to the lack of raw materials - for sensors, edge, and telecommunication devices</p>	<p>Dissemination & communication activities do not reach the targeted groups. Project results lack the intended visibility</p>
WP	<p>WP 1,2</p>	<p>WP 5</p>	<p>WP 1,2,3,4,5,6</p>	<p>WP 3,5</p>	<p>WP 4,5</p>
MITIGATION MEASURES	<p>The project uses a rolling plan and perform rolling plan activities within WP2 as a way to link those with the co-design Phase. WP2 collaborates with WP1 on decisions related to adaptation.</p>	<p>Before deployment and demonstration, a careful assessment of the environment conditions is performed, so that the causes of poor performance can be identified early in the project and used for the optimization of the system, and the prevention of future system failures.</p>	<p>The unexpectedly rising prices force the consortium to adapt to the current uncertain societal environment. Personnel costs, for example in Belgium are by law linked to the evolution of inflation. The consortium will try to find more efficiencies in every aspect of its daily way of working. These risks will be mitigated to a minimum, in consultation with the consortium, by measures that increase efficiency and reduce costs as much as possible. These may include: assigning employees with a lower wage cost (but within the required qualifications for the project), and especially by focusing on maximum effectiveness: more efficient/less meetings with a limited number of participants, shorter meetings, bundling meetings, on-line meetings to avoid travelling time, meetings at the EC to reduce costs; agile reporting that is good enough to streamline the progress of the project. However, depending on the inflation rate, additional measures may be needed to ensure a proper execution of the tasks. These may include reducing the amount of experimenting in the RILabs as well as the number of data products, or increasing the available budget to ensure all foreseen work can be done.</p>	<p>The consortium members closely monitor the market and any change that may impact the needed consumables/equipment. The consortium will be working and experimenting with some alternative electronic boards or even some custom designs as a contingency plan.</p>	<p>The comprehensive Dissemination, Exploitation & Communication Plan and Package (D6.1) identifies key stakeholders, ambassadors, multipliers etc., and defines clear strategies to reach them. The Plan is updated regularly, reflecting potential required changes in communication channels or messages.</p>
STATE OF PLAY	<p>The Rolling plan is being used to manage the necessary and requested 'adaptations' to the project activities. The plan is being managed together with the ScaleAgData Coordinator and WP leads and is presented quarterly at Executive Board meetings attended by all WP leads as well as RILab leads. This enables the interdependencies between the project's subsystems/WPs to be coordinated, critical issues, technical risks and actions to</p>	<p>While the demonstration phase started in M25, preliminary mitigation measures were implemented proactively to reduce potential risks. These measures involve detailing the demonstrators of all RILabs and</p>	<p>A milestone and review meeting have been bundled to reduce travel costs. Free Flemish government meeting room facility was used for the last Milestone meeting to reduce costs. Only three (instead of 4) in-person meetings are scheduled within the project lifecycle to reduced costs. As 3 out of the four advisors identified at the start of the project left their respective organization or retired, the</p>	<p>WP6 has not yet started developing business models. Future actions include defining possible business models with clear roles and responsibilities, establishing a baseline for future commercial agreements, and communicating project results and recommendations to</p>	<p>Website and Social Media channels actively provide updates on project results. Two external stakeholder events have taken place to make initial connection with intended stakeholders. The demonstrator events that took place in February 2025</p>

mitigate them to be identified and monitored.	their respective environments in the context of the preparation of the RILab action plans. The action plans amongst other include evaluation criteria, assessment results, and potential updates, all of which contribute to better preparation of the RILabs for their demonstrations.	Advisory Board will not be re-established to save costs.	policy makers through a policy brief. In the Water Productivity RILab, IES faced the challenge that initially planned soil monitoring sensors (OneSoil weather and soil stations) were not produced anymore and analogous low-cost (<500 euro per station) solution was not available in the market. Therefore, other more classical options for soil and weather parameter monitoring in test fields were evaluated. This resulted in the set-up of one meteorological sensor and four soil sensor nodes which proved still sufficient for lab demonstrator purposes as well as fits within planned budget (IoT soil sensors).	included RILab specifically targeted groups and stakeholders. Website and social media visits as well as newsletter subscriptions are being monitored. D6.1 (drafted in June 2023) has already been updated through D6.2 which was drafted in June 2024 and a new update is foreseen towards the end of 2025. During the first reporting period, project partners participated in over 40 events during the first 18 months to create visibility among key stakeholders and create synergies.
---	---	--	---	---

Technological Risks					
RISK	Specifications too high level or innovations don't satisfy the requirements	Sentinel-1A failure	Intended new technologies for agri-environmental monitoring do not deliver	Development of solutions faces unexpected technical issues that are hard to resolve with the given resources	Design of results that do not foreseen the need for collaboration and interoperability between the RILabs but also between the different vertical domains (soil, crop, livestock etc).
WP	2,3,4,5	3,4,5	3,4	3,4	2,3
MITIGATION MEASURES	The co-design approach (WP2), with the iterative design phases and foreseen rolling plan, are in place to ensure a constant evaluation and refinement of the proposed innovation needed by the RILabs and other users, and the capacities of the technology providers.	As Sentinel-1B failed in 12/2021, only the Sentinel-1A remains to provide SAR data at 10m resolution. Some of the foreseen data augmentation methods (fusion S-1 and 2 for cloud filling), as well as data products (e.g. harvest detection) rely on this satellite, although neither are critical to the project. In addition, no issues are foreseen for S-1A, and an early launch for Sentinel-1C (early 2023) has been announced, meaning S1 data will be available for the critical phases of the project	WP3 will monitor closely the related progress and detect potential problems in time, suggesting other technical solutions. Often, multiple technical solutions are envisaged to be provided to the RILabs, allowing them to evaluate the most adequate methodology and technology for their needs.	An iterative, agile development approach will be used to detect technical challenges immediately & solve them quickly. In addition, multiple technical service providers are in the consortium that work together to provide the requested solutions. This also works as a buffer for a certain number of technical issues.	In the co-design phase the needs were identified and the technical requirements (functional and non-functional) that support the interoperability were defined, after which WP3 identified necessary components and interactions. Additionally the governance models were defined that follow the openDEI design principles for data spaces to support the interoperability between the vertical domains of the RILabs but also the Horizon Europe candidate partnership "Agriculture of Data".
STATE OF PLAY	A rolling plan was developed. This is a "living" document visualising and reporting the research activities (cfr. deployment scenarios and innovation areas) for the different research partners and technical providers to enable easy adaptation to the latest developments and innovations in the field of sensors and sensor data	The launch of Sentinel 1C has taken place and will provide an effective back-up for the current Sentinel 1A.	Due to DMU issues with UKRI funding, some delay occurred with regards to the set up of the methodological framework for federated learning. ILVO and Auth have mitigated this by setting up their own federated approach. Currently the team from Deimos Space already recovered this delay and is already working with ILVO and AUTH to produce a first version of the federated learning component based	The sensor development for the VTT camera is progressing as planned and was ready by May 2024. However, there have been issues with manufacturing the Fabry-Perot Interferometer filter, meaning that initially, the instrument will only be able to operate	A co-design approach has been employed to gather information for each RILab and the vertical domain they represent. Technical requirements have been identified, including both functional and non-functional aspects, as well as specific data-sharing needs within and between RILabs, to support interoperability. Additionally, innovative governance frameworks have been

	in the public and private sectors.		on the requirements of those two partners DHI also experienced delay in the development of data integration techniques. This is now postponed to the second iteration. In the first iteration, they focused on providing basic soil moisture and ET products	within a reduced wavelength range of approximately 650-950 nanometers. The project proposal stipulates customization of the hyperspectral camera to align with project requirements. Although no specific wavelength range is predetermined, efforts will be directed towards maximizing the range while maintaining other instrument specifications. A key aspect in achieving this goal is the fabrication of the FPI filter with sufficient small gap between mirror surfaces. R&D for that is an actively ongoing activity at VTT. A potential mitigation strategy involves adjusting the operational range of the camera according to the observations made during the initial year of measurements. While this would not increase the total wavelength range, for the RILab it would potentially increase the usefulness of the data recorded by the instrument.	aligned with recent developments in data spaces and governance, and necessary tools have been identified.
--	------------------------------------	--	--	--	---

Other Risks				
RISK	The potential users of the RILabs vertical domains, due to business restrictions, facing difficulties to access and use ScaleAgData methodologies and data products.	Extreme weather during the 2022/23 growing season hindered the first field campaign.	Notable difference in data and technology maturity between research partners and RILabs. Some partners have extensive experience in developing data products and automating data processes, while others are less experienced and focus mainly on exploring innovative technologies for improved monitoring in competitive and sustainable agriculture. This difference in maturity presents challenges during co-design workshops and data collection	Diverging expertise: The diverse expertise among technical partners and RILabs involved in WP3—spanning sensor development, edge processing, and data fusion—creates a risk of misalignment in technical approaches and priorities. This misalignment can lead to integration challenges.
WP	WP 2,6	WP 5	WP2	WP3

MITIGATION MEASURES	<p>ScaleAgData aims to support further development of interoperable data spaces for each vertical domain. WP2 governance models will support the establishment of trust between the actors. WP6 will deliver business models to support the establishment of interconnecting value networks for the RILabs vertical domains, foreseen in parallel for a suitable definition of the IPR rights.</p>	<p>Extend the field collection for an additional year (2024/25) in order to complete the necessary dataset.</p>	<p>All relevant project activities are closely monitored with a specific focus on tracking progress within the RILabs. Information is continuously collected as the RILabs progress, and ongoing efforts ensure flexibility in adapting tools to elicit lab-specific requirements that meet evolving needs. Additionally, fine-tuning and updating the collected requirements are undertaken to ensure they remain relevant and align with the evolving objectives of both the RILabs and the project.</p>	<p>To address these risks, it is crucial to establish early and ongoing collaboration among all partners. Regular integration meetings should focus on aligning data flow, transmission rates, and real-time processing needs. Incremental prototyping and cross-partner validation should be used to test and refine data delivery capabilities, helping to identify and resolve issues before full-scale integration. High-level architecture alignment meetings will ensure that all system components are consistent with the overall design. Additionally, building in buffer time for unexpected challenges and appointing a dedicated team for architecture coordination will facilitate early conflict resolution, ensuring that both data delivery and architectural milestones are achieved as planned.</p>
STATE OF PLAY	<p>This has not materialized yet as too early in the project cycle</p>	<p>Field collection campaign was extended for an additional year</p>	<p>Project activities are being closely monitored, with a focus on tracking progress in the RILabs. Information is collected continuously, and tools to elicit lab-specific requirements are being adapted to meet the labs' evolving needs. Requirements are regularly updated to stay relevant and align with the project's objectives.</p>	<p>Collaboration among many partners through meetings focusing on integration was set up and will continue to take place. In addition, the maturity of the project and the partners has already improved the misaligned. WP3 lead has also organized bilateral meetings with RILabs to identify and support their difficulties and challenges in development and/or implementation of specific building blocks. Accordingly, for specific labs databases for storing data were designed supporting integration and ensuring data integrity.</p>