Ref. Ares(2025)11474470 - 20/12/2025

NOSTRADAMUS

D2.1 NEEDS-BASED DATA REQUIREMENTS

20/12/2025

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

NOSTRADAMUS

PARTNERS

ERATOSTHENES CENTRE OF EXCELLENCE

DLR Deutsches Zentrum für Luft- und Raumfahrt German Aerospace Center

AgFutura TECHNOLOGIES

CROPT

ITC INOVACIJSKO TEHNOLOŠKI GROZD INNOVATION TECHNOLOGY CLUSTER

AGRIFOOD DIGITAL INNOVATION HUB

DLG

CZU Czech University of Life Sciences Prague

BioSense INSTITUTE

TERRADUE Advancing Earth Science

Kmetijsko gozdarska zbornica Slovenije KMETIJSKO GOZDARSKI ZAVOD MURSKA SOBOTA

FGS

AGRICULTURAL RESEARCH INSTITUTE

ARISTOTLE UNIVERSITY OF THESSALONIKI

UNIVERSITÉ DE GENÈVE INSTITUTE FOR ENVIRONMENTAL SCIENCES

University of St.Gallen

Agroscope

NOSTRADAMUS

Grant Agreement No.: 101134888
Call: HORIZON-CL6-2023-GOVERNANCE-01
Topic: HORIZON-CL6-2023-GOVERNANCE-01-13
Type of action: HORIZON Research and Innovation Actions

D2.1 NEEDS-BASED DATA REQUIREMENTS

Work package	2
Task	Task 2.1: Driving datasets and workflow for assessing EU Food Security and independence Task 2.2: Needs-based data requirements for Data Cubes
Due date	31/12/2025
Submission date	20/12/2025
Deliverable lead	Eratosthenes Centre of Excellence
Version	2.1
Authors	Tomaz Bokan (related to Task 2.1, ITC) Leon Abraham (related to Task 2.1, ITC) Ioannis Varvaris (related to Task 2.2, ECoE)
Reviewers	Konstantinous Panagiotou (AUTH), Georgia Lavranou (HSG), Thomaida Polydorou (ECoE)
Abstract	This deliverable presents the progress of Tasks 2.1 and 2.2 of Nostradamus, and related results up to Month 16 of the project. Task 2.1, completed by Month 8, defines the internal data requirements and workflows needed to develop the four WP4 core modules and specifies the datasets to be provided by the five demonstration sites. Task 2.2 establishes the framework for identifying and continuously updating external data requirements through MAPs, stakeholders, related EU initiatives, and future Third-Party developers. Together, these tasks form the first version of a needs-driven data requirements framework that will guide data acquisition, integration, and Data Cube development

	throughout the project.
Keywords	Data requirements; Data Cubes; Digital agriculture; Core modules; Stakeholder needs; Multi-Actor Platforms; External data framework; Workflows; Earth Observation; IoT data; Ground Measurements; Demonstration Sites; Food security; EU strategic alignment; and Data suitability criteria.

DISCLAIMER

Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

COPYRIGHT NOTICE

© NOSTRADAMUS Consortium, 2025

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorized provided the source is acknowledged.

The Deliverable 2.1 Consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	ERATOSTHENES Centre of Excellence	ECoE	CY
2	Deutsches Zentrum für Luft- und Raumfahrt e.V.	DLR	DE
3	AG Futura Technologii DOOEL Skopje	AGFT	MK
4	Cropt DOO Novi Sad	CROPT	RS
5	ITC – Inovacijsko Tehnološki Grozd Murska Sobota	ITC	SI
6	Česká Zemědělská Univerzita v Praze	CZU	CZ
7	BioSense Institute – Research and Development Institute for Information Technologies in Biosystems	BIOS	RS
8	Terradue Srl	Terradue	IT
9	Université de Genève	UNIGE	CH
10	Universität St. Gallen	HSG	CH
11	Eidgenössisches Departement für Wirtschaft, Bildung und Forschung	WBF	CH

NOSTRADAMUS

EXECUTIVE SUMMARY

Deliverable D2.1 presents the consolidated progress of Tasks 2.1 and 2.2 at Month 16 of the Nostradamus project, providing the foundation for needs-driven data requirements that support the development of the project's digital applications, core analytical modules, and the Data Cube infrastructure. Task 2.1 (completed at Month 8) defines the internal data requirements and associated workflows necessary for the development of the four WP4 core technological modules: (i) drought and vegetation stress monitoring, (ii) crop suitability and yield estimation, (iii) soil fertility and crop nutritional management, and (iv) sustainable pest management. Task 2.1 identifies the driving datasets, variables, spatial and temporal resolutions, and data sources needed for each module. It also establishes the workflow from requirements gathering to data provisioning, ensuring that WP3 can prepare and ingest the necessary open and demonstration-site data for module development across the five case study countries (Cyprus, Germany, Serbia, Slovenia, and Switzerland). These outputs fully support the implementation activities in WP3 and WP4.

Task 2.2 (reported as progress at Month 16, representing the first version of an evolving process that will be updated again at Month 44) extends the requirements process beyond internal datasets by establishing the framework for identifying and continuously updating external data requirements. This includes data needs originating from Multi-Actor Platforms (MAPs), key stakeholder groups, Cluster 6 and EU Missions projects, AKIS platforms and networks, and future Third-Party developers in WP6. Task 2.2 defines the criteria for assessing data suitability (completeness, consistency, accuracy, timeliness, relevance, granularity, normalization, integrity) and introduces the first version of the survey designed to collect external requirements. This framework marks the beginning of a living, iterative process, with systematic external data collection scheduled to intensify after Month 24 as MAPs become fully operational, Third-Party activities progress, and synergy activities with related initiatives continue to evolve.

Together, Tasks 2.1 and 2.2 create a coherent structure for guiding data acquisition, integration, and prioritization across the project. Task 2.1 delivers the internal foundations required for module development, while Task 2.2 prepares the mechanisms and tools for continuous engagement and expansion of the data ecosystem in alignment with stakeholder needs, evolving EU strategies, and upcoming applications.

Deliverable D2.1 therefore provides the first complete version of the Nostradamus data requirements framework, supporting the project's objectives in digital agriculture, food security, and EU independence in agrochemical imports, and ensuring strong interfaces with WP3, WP4, and WP5 for subsequent development stages.

TABLE OF CONTENTS

1 INTRODUCTION 11
1.1 Internal and External Data Requirements 11
1.2 Introduction to Task 2.1 component of Deliverable D2.1 13
1.3 Introduction to Task 2.2 component of Deliverable D2.1 15
1.4 Document structure 16
2 DEMONSTRATION SITE IN GERMANY – DE (DLG) 17
2.1 Location description 18
2.2 Participant 19
2.3 Datasets 19
2.4 Dataset timeline 20
3 DEMONSTRATION SITE IN SERBIA – RS (BIOS, CPT) 20
3.1 Location description 20
3.2 Participant 23
3.3 Datasets 23
3.4 Dataset timeline 24
4 DEMONSTRATION SITE IN SWITZERLAND – CH (WBF) 24
4.1 Location description 24
4.2 Participant 25
4.3 Datasets 25
4.4 Dataset timeline 26
5 DEMONSTRATION SITE IN SLOVENIA – SL (ITC, KGZS) 26
5.1 Location description 26
5.2 Participant 29
5.3 Datasets 29
5.4 Dataset timeline 29
6 DEMONSTRATION SITE IN CYPRUS – CY (ECOE, MARDE) 30
6.1 Location description 30
6.2 Participant 32

6.3 Datasets ... 32
6.4 Dataset timeline ... 33
7 MODULE 1 – EARLY WARNING SYSTEM FOR DROUGHT AND VEGETATION STRESS MONITORING (DLR) ... 33
7.1 Problem / opportunity ... 34
7.2 Benefits and key business objective ... 34
7.3 Technical challenges ... 34
7.4 Module size / metrics ... 35
7.5 Expected impact and KPI's ... 35
7.6 Demonstration plan ... 35
7.7 Data requirements for module ... 36
7.8 Dataset workflow ... 37
8 MODULE 2 – CROP SUITABILITY AND AGRICULTURAL YIELDS ESTIMATION (BIOS) ... 38
8.1 Problem / opportunity ... 38
8.2 Benefits and key business objective ... 39
8.3 Technical challenges ... 39
8.4 Module size / metrics ... 39
8.5 Expected impact and KPI's ... 40
8.6 Demonstration plan ... 40
8.7 Data requirements for module ... 40
8.8 Dataset workflow ... 41
9 MODULE 3 – SOIL FERTILITY AND CROP NUTRITIONAL MANAGEMENT (ITC) ... 42
9.1 Problem / opportunity ... 42
9.2 Benefits and key business objective ... 43
9.3 Technical challenges ... 44
9.4 Module size / metrics ... 44
9.5 Expected impact and KPI's ... 44
9.6 Demonstration plan ... 44
9.7 Data requirements for module ... 44
9.8 Dataset workflow ... 46
10 MODULE 4 – SUSTAINABLE AND EFFICIENT PEST MANAGEMENT (ECOE) ... 46

10.1 PROBLEM / OPPORTUNITY ... 46
10.2 Benefits and key business objective ... 47
10.3 Technical challenges ... 48
10.4 Module size / metrics ... 48
10.5 Expected impact and KPI's ... 48
10.6 Demonstration plan ... 49
10.7 Data requirements for module ... 49
10.8 Dataset workflow ... 50
11 PILOT REQUIREMENTS AND KPI'S ... 51
12 EXTERNAL NEEDS-BASED DATA REQUIREMENTS FRAMEWORK ... 52
12.1 Framework for Prioritising Needs-Based Data Requirements ... 53
12.1.1 Step 1: Identification of Stakeholders ... 54
12.1.2 Step 2: Clarification of Stakeholder Interest and Influence ... 55
12.1.3 Step 3: Agreement on the Engagement Process ... 56
12.1.4 Suitability Criteria for External Data Requirements ... 57
12.1.4.1 Completeness ... 58
12.1.4.2 Consistency ... 59
12.1.4.3 Accuracy ... 59
12.1.4.4 Timeliness ... 60
12.1.4.5 Relevance ... 60
12.1.4.6 Granularity ... 60
12.1.4.7 Normalization ... 61
12.1.4.8 Integrity ... 61
12.1.5 Policy Alignment Benchmarks Informing the Suitability Criteria ... 62
12.1.6 Ranking Matrix for Prioritizing External Data Requirements ... 69
12.2 Preliminary activities ... 71
12.2.1 MAPs – Preliminary Data Requirements Process ... 71
12.2.2 Synergies with EU Projects and Initiatives ... 73
12.2.3 Third-Party Developers under WP6 ... 76
13 CONCLUSIONS ... 77

LIST OF FIGURES

FIGURE 1: NOSTRADAMUS DATA REQUIREMENTS GENERAL CONCEPT 13

FIGURE 1: MAP OF THE IPZ FARM 18

FIGURE 2: LOCATION OF THE IPZ IN GERMANY 18

FIGURE 3: MACRO LOCATION OF VOJVODINA, SERBIA 21

FIGURE 4: MICRO LOCATION OF VOJVODINA 22

FIGURE 5: MACRO LOCATION – SLOVENIA 27

FIGURE 6: MICRO LOCATION - POMURJE REGION 28

FIGURE 7: LOCATION OF THE ACHELEIA EXPERIMENTAL STATION 31

FIGURE 8: MICRO-LEVEL OF ACHELEIA EXPERIMENTAL STATION 32

FIGURE 9: DATASET WORKFLOW FOR THE MODULE FOR DROUGHT MONITORING AND EARLY WARNING 37

FIGURE 10: TIMELINE FOR TASK 4.1 SHOWING WHEN THE REQUIRED DATA WILL BE UTILIZED FOR THE DROUGHT MODULE DEVELOPMENT. 38

FIGURE 11: DATASET WORKFLOW FOR MODULE 2 41

FIGURE 12: TIMELINE DETAILING WHEN THE REQUIRED DATA WILL BE USED 42

FIGURE 13: DATASET WORKFLOW FOR MODULE 3 46

FIGURE 14: MODULE 4 DATASET WORKFLOW 50

FIGURE 15: TIMELINE FOR MODULE 4 MAIN TASKS 51

FIGURE 17: FRAMEWORK FOR PRIORITIING AND ENGAGING STAKEHOLDERS IN THE IDENTIFICATION OF EXTERNAL DATA REQUIREMENTS 53

FIGURE 18: POWER-INTEREST MATRIX TO CLASSIFY NOSTRADAMUS STAKEHOLDERS ACCORDING TO THEIR INFLUENCE ON PROJECT OBJECTIVES AND THEIR LEVEL OF INTEREST IN DATA REQUIREMENTS. 55

FIGURE 19: IDENTIFICATION OF EXTERNAL DATA REQUIREMENTS 71

FIGURE 20: NOSTRADAMUS SURVEY THROUGH A CO-DESIGN PROCESS WITH NOSTRADAMUS CONSORTIUM 72

FIGURE 21: NOSTRADAMUS PARTICIPATION IN RSCY2025 73

FIGURE 22: NOSTRADAMUS PARTICIPATION IN SYNERGY DAYS 2024 74

FIGURE 23: NOSTRADAMUS PARTICIPATION IN LANDSHIFT AND ECO-READY PROJECTS' WORKSHOPS 75

FIGURE 24: ESTABLISHMENT OF AGRICULTURAL SYNERGY CLUSTER 75

LIST OF TABLES

TABLE 1: PILOT KPI'S 51

TABLE 2: KEY EU AND INTERNATIONAL STRATEGIES RELEVANT TO NOSTRADAMUS AND THEIR ASSOCIATED OBJECTIVES AND ACTIONS INFORMING EXTERNAL DATA REQUIREMENTS. 62

TABLE 3: WEIGHTING OF SUITABILITY CRITERIA FOR RANKING EXTERNAL DATA REQUIREMENTS 69

1 INTRODUCTION

The Nostradamus project aims to address critical challenges in European agriculture arising from climate change, resource limitations and the increasing need for sustainable and resilient farming systems. These pressures have intensified the importance of reliable, high-quality and interoperable data for supporting informed decision-making across the agricultural sector. Digital transformation and advanced analytics can only be effective when built upon well-defined and well-structured data requirements, which form the foundation for robust monitoring, modelling and evaluation tools.

Within this context, Nostradamus develops a scalable digital platform that integrates diverse data sources, including satellite observations, IoT measurements, meteorological information, soil data and agricultural management records. The project combines these heterogeneous datasets through data cubes and advanced analytics in order to develop user-oriented applications that enhance climate resilience, improve resource efficiency and strengthen Food Security, Safety, and Sovereignty. Since the accuracy and usability of such applications depend heavily on the quality, suitability and consistency of input data, the identification and organization of data requirements is a central component of the project's technical, scientific and operational work. Nostradamus is validated through demonstration sites located in Cyprus, Germany, Serbia, Slovenia and Switzerland. These sites represent diverse climatic, agronomic and socio-economic conditions and therefore require data strategies that reflect regional variations as well as cross-site harmonization needs. The demonstration sites provide essential in situ measurements and agricultural information, which together with open-access datasets form the basis for the project's analytical modules and digital applications.

1.1 INTERNAL AND EXTERNAL DATA REQUIREMENTS

The transition toward data-driven agriculture requires integrated, interoperable and accessible data ecosystems capable of supporting innovation across diverse farming environments. Breaking traditional data silos is essential for enabling the development of aggregated data infrastructures that can reinforce the mission and vision of projects such as Nostradamus, while also extending value to other initiatives working toward similar goals. When data are fragmented, inaccessible or incompatible, their potential to drive digitalization, improve decision making and strengthen co-governance frameworks remains limited. By contrast, a harmonized and well-structured data environment allows different stakeholders and projects to build upon shared foundations, accelerate technological adoption and maximize the long-term impact of public research investments. This approach is particularly important for supporting competitiveness in agricultural regions with limited connectivity, fragmented information flows or low digital literacy. Establishing clear and robust data requirements is therefore a foundational step for ensuring that the project's outcomes are inclusive, scalable and capable of contributing to broader European ambitions for sustainable and digitally empowered agriculture.

Within this context, a key part of the project is the involvement of Multi-Actor Platforms (MAPs), which bring together farmers, advisors, policymakers, industry partners and researchers. Their input helps define the data needs from an end-user perspective and ensures that the project outputs align with real-world challenges, policy demands and operational requirements. This collaborative approach breaks the isolation of data provision by creating a shared environment where data producers, data users and domain experts can exchange insights and jointly identify missing data, validation needs and opportunities for methodological improvement. Through this continuous interaction, the project strengthens its understanding of data gaps, data quality limitations and emerging data needs.

Beyond stakeholder input, Nostradamus also considers the data requirements identified in other EU-funded projects, European partnerships, thematic clusters and international initiatives that work on related agricultural, environmental and socio-economic topics. The project reviews the external data requirements reported by these initiatives to identify which of them could potentially be supported, complemented or facilitated by the data architecture developed in Nostradamus. This approach allows the project to determine when external needs are aligned with its own objectives and when the Nostradamus data environment can offer added value. It also ensures that the project contributes to a coherent European research landscape by assisting other initiatives when possible. Through this process, Nostradamus supports wider EU strategies and enables the creation of an aggregated and structured data environment that can respond to shared needs across projects and user communities. In addition to these sources, the project will also collect the data requirements submitted by the successful applicants of the two Open Calls (OCs) under WP6, which focus on the development of applications for decision making and policy shaping.

In this broader context, Deliverable D2.1 establishes the internal and external data requirements that will guide the development of Nostradamus modules and applications. The deliverable clarifies what data are needed, where they originate, how they should be formatted and how they can be integrated to ensure interoperability, traceability and usability across the project's technical components. The establishment of clear data requirements is essential for achieving the project's objectives, supporting the development of accurate analytical tools and enabling the delivery of actionable insights to farmers, advisors and public authorities, among others. This introduction therefore situates Deliverable D2.1 within the wider context of the Nostradamus project and highlights the importance of structured, comprehensive and need-driven data architecture for the success of all subsequent activities.

- **Internal data requirements:** refer to all data that originate from within the Nostradamus project environment, which include: a) datasets provided by the five demonstration sites in Cyprus, Germany, Serbia, Slovenia and Switzerland, b) measurements, observations and records generated directly by consortium partners. Internal data support the design, development and validation of the WP4 analytical modules and the associated digital applications.
- **External data requirements:** refer to the types of data needed for project implementation that originate outside the consortium. These include information derived from: a) stakeholder engagement through the MAPs, b) the data requirements identified in other EU-funded projects and initiatives, c) insights from European partnerships and clusters, d) the

data needs submitted by the Third-Party developers funded under the WP6 OCs, e) open-access Earth Observation and, f) meteorological sources and any complementary information that enhances or validates the project results. Integrating these external requirements allows Nostradamus to co-design interoperable solutions, support upscaling activities, and contribute to the broader ecosystem of EU strategies and research efforts.

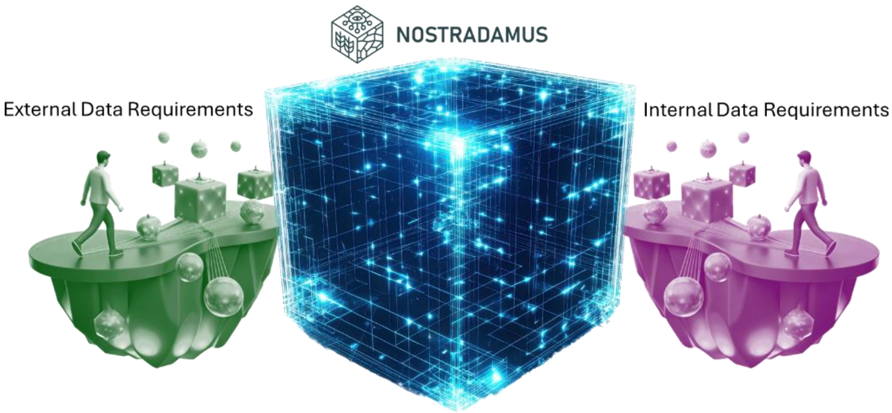


FIGURE 1: NOSTRADAMUS DATA REQUIREMENTS GENERAL CONCEPT

1.2 INTRODUCTION TO TASK 2.1 COMPONENT OF DELIVERABLE D2.1

Task 2.1 of Deliverable D2.1, "Needs-based Data Requirements," plays a foundational role by identifying and structuring the internal data needs essential for the development of the project's four core modules (WP4). These modules are designed to support the agricultural sector with advanced, data-driven tools for drought and vegetation stress monitoring, crop suitability and yield estimation, soil fertility and crop nutrition management, and sustainable pest management. Task 2.1 focuses on defining the internal data inputs required for module development, drawing from diverse sources including remote sensing, IoT sensor networks, meteorological stations, and field observations. These datasets are tailored to five geographically and socio-economically diverse demonstration sites located in Germany, Serbia, Switzerland, Slovenia, and Cyprus. The integration of these site-specific and open datasets will enable the development of innovative digital applications that provide actionable insights to farmers, advisors, policymakers, and other stakeholders. Furthermore, this part of the deliverable ensures that the data requirements are fully aligned with the needs of the modules and the characteristics of the pilot sites, facilitating effective implementation, testing, and demonstration activities throughout the project. By doing so, Task 2.1 sets out the basis for informed decision-making, improved resilience of agricultural systems, and the achievement of the EU's sustainability objectives.

Deliverable D2.1, through the work conducted under Task 2.1, lays out the groundwork for all four technical modules and associated digital applications by systematically identifying, and structuring the internal data they will consume. The specific objectives addressed under Task 2.1 are to:

1. Define and document data sources being relevant for Nostradamus.

   a. Remote-sensing platforms (satellite, UAV).

   b. In-field IoT sensor networks (soil moisture, microclimate, etc.).

   c. Meteorological station records.

   d. Historical and real-time field observations.

   e. Socio-economic and management practice databases.

2. Gather and specify the data requirements for each of the four core modules, leveraging diverse data sources such as remote sensing (satellite/drone imagery), IoT sensor networks, meteorological data, soil analyses, and historical agricultural records (Section 7-10).

   a. **Module 1 - Early Warning System for Drought and Vegetation Stress Monitoring:** Utilizing Earth observation data and advanced anomaly detection, this module aims to provide timely warnings and support adaptive agricultural management strategies.
   b. **Module 2 - Crop Suitability and Agricultural Yields Estimation:** Driven by machine learning algorithms, this module assesses crop suitability and predicts yield potential based on various data inputs, enhancing farmers' decision-making capabilities under changing climate conditions.
   c. **Module 3 - Soil Fertility and Crop Nutritional Management:** Employing satellite data, IoT sensors, and laboratory analyses, this module delivers precise recommendations on fertilizer usage, improving soil health and reducing environmental impacts from nutrient overuse.
   d. **Module 4 - Sustainable and Efficient Pest Management:** This module integrates IoT meteorological data, soil moisture sensors, and satellite observations to create early-warning systems for cereal pest and disease management, significantly reducing agrochemical use.

3. Gather and specify data requirements for the project's digital Applications.

   a. Decision-support dashboards for farmers and advisors.

   b. Reporting tools for policymakers and stakeholder engagement.

   c. Mobile/web interfaces for Multi-Actor Platforms (MAPs).

4. Map and document the relationships between Modules, Applications and demonstration-site data.

   a. Link each Module's input needs to specific data sources available at the five pilot sites (Germany, Serbia, Switzerland, Slovenia, Cyprus).

   b. Define data timelines, formats and harmonization rules for cross-site interoperability.

c. Establish workflows showing how raw data flows through pre-processing into module algorithms and then into end-user applications.

By achieving these core objectives, D2.1 ensures that all subsequent development, testing and demonstration activities are underpinned by a clear, harmonized and needs-driven data architecture—thereby maximizing the impact, reliability and usability of Nostradamus' digital agriculture solutions. Key findings highlight the project's potential to transform agricultural practices through data integration and precision agriculture, resulting in increased crop yields, reduced environmental impacts, improved soil health, and heightened economic competitiveness for farmers. The report emphasizes the necessity of robust technical frameworks to manage and process large volumes of heterogeneous data and recommends continued collaboration between technology experts, agronomic experts, and local stakeholders to refine and validate the proposed solutions. In this deliverable, a detailed workflow is also included for the development of each analytical module, clearly outline their specific data requirements, and determine how data will be provisioned from the five demonstration sites. This approach ensures that module design, data collection, and site-specific validation proceed in a coordinated, transparent manner, laying the groundwork for scalable implementation across diverse agricultural contexts.

1.3 INTRODUCTION TO TASK 2.2 COMPONENT OF DELIVERABLE D2.1

The Task 2.2 component of Deliverable D2.1 focuses on defining, structuring and initiating the process of collection of external data requirements that will complement the internal data requirements collected under Task 2.1 and support the objectives of the Nostradamus modules and applications. Unlike Task 2.1, which concentrates on data originating within the project, Task 2.2 addresses the broader ecosystem of external requirements emerging from stakeholders, MAPs, related EU-funded projects, European partnerships, thematic clusters, Third Parties under WP6 and other relevant initiatives operating across the agricultural, environmental and socio-economic landscapes. Establishing these external data requirements is essential for ensuring that the Nostradamus data hub remains interoperable, forward-looking and aligned with the wider European research and innovation agenda.

A core achievement of Task 2.2 during the first reporting period (up to Month 16) is the development of a comprehensive framework for assessing external data requirements, ensuring that all future needs can be evaluated in a consistent and transparent way. This framework defines the criteria and conditions that must be met for any external data to be considered suitable for use within the Nostradamus ecosystem. These criteria include completeness, consistency, accuracy, timeliness, relevance, granularity, normalization and integrity. By establishing these assessment rules early, the project ensures that future external data requests — whether from MAPs, related projects or Third-Party developers — can be systematically evaluated against the same quality benchmarks.

In addition to establishing the assessment criteria, Task 2.2 developed the overall methodology for determining whether external data requirements fall within the scope of Nostradamus objectives. This includes procedures for evaluating whether specific data types can meaningfully support module development, enhance validation activities, contribute to decision-making workflows, or facilitate upscaling and cross-project interoperability. This methodology ensures that Nostradamus does not accumulate unnecessary datasets but rather focuses on external needs that are aligned with the project's objectives, the anticipated functionalities of its applications and the policy priorities identified under WP1.

During the first iteration of Deliverable D2.1, Task 2.2 also launched the initial phase of external data requirements collection, initiating the process for gathering preliminary inputs from diverse sources. This included the design of a survey to obtain early inputs from the MAPs established under WP1, capturing the needs, expectations and digital capacity challenges identified by farmers, advisors, policymakers and other stakeholders. Furthermore, the project initiated the process for the potential collection of data requirements from relevant EU-funded projects, international initiatives and thematic clusters, focusing specifically on their expressed data needs. This activity aligns with the synergies already established with these projects, and includes communication efforts to inform them about the objective of collecting their potential data requirements. This ensures alignment with broader European efforts and prepares the ground for supporting cross-project synergies where feasible. In addition, Task 2.2 will at a later stage gather the data requirements submitted by the successful Third-Party developers funded under WP6, ensuring that the applications developed through the Open Calls for decision making and policy shaping will be incorporated into the data requirements framework from the outset.

By structuring the assessment criteria, establishing a methodology for alignment and initiating the collection of preliminary external requirements, the Task 2.2 component of Deliverable D2.1 lays the foundation for a dynamic, continuously evolving process of external data integration within Nostradamus. This first iteration provides the basis upon which future updates will expand, with a major revision planned for Month 44.

1.4 DOCUMENT STRUCTURE

The present document is structured as follows:

- **Section 1 – Introduction:** Outlines the purpose and objectives of Deliverable D2.1 and its role within the Nostradamus project.
- **Sections 2 to 6 – Demonstration Sites:** Detailed description of the five pilot sites (Germany, Serbia, Switzerland, Slovenia, and Cyprus), including:
    - Geographical and climatic characteristics.
    - Socio-economic and agricultural context.
    - Existing and required datasets.
    - Data timelines.
- **Sections 7 to 10 – Core Modules:** Description of the four technical modules:

- **Module 1:** Early Warning System for Drought and Vegetation Stress.
- **Module 2:** Crop Suitability and Agricultural Yield Estimation.
- **Module 3:** Soil Fertility and Crop Nutritional Management.
- **Module 4:** Sustainable and Efficient Pest Management.

- **Section 11 – Pilot Requirements and Key Implementation Parameters (KIPs):** Defines the requirements, criteria, and expectations for the demonstration sites.
- **Section 12 – Framework for external data requirements**: Defines the requirements for external data sources and the collection process.
- **Section 13 – Conclusions:** Summarizes the main outcomes of the deliverable and defines the next steps.
- **Annexes:** Include dataset templates, data specifications, and additional technical information to support data harmonization and module development.

This structure ensures a logical flow from general project objectives to specific data needs, facilitating clarity and usability for all project partners and stakeholders. Moreover, the structure and content for the Task 2.1 section in Deliverable D2.1 were defined and formalized in the internal Deliverable D2.1, which was submitted upon the completion of Task 2.1 (Month 8).

2 DEMONSTRATION SITE IN GERMANY – DE (DLG)

The demonstration site is in the Continental biogeographic region covering over a quarter of the European Union (29 % of the EU territory). According to World Development Indicators (WDI) from World Bank, DE has the third highest cereal yield and cereal production in the EU. The country has the fourth-highest covered land area under cereal production. In contrast, DE has one of the lowest fertilizer consumption rates (% of fertilizer production), which is correlated with the implemented national policies and initiatives and financial support to farmers. In addition, most German farm managers have basic training and full agricultural training in AMP, according to Eurostat, while only a smaller percentage have practical experience. An additional criterion for selecting the inclusion of a pilot site in DE is the high percentage of areas covered with low average internet speeds (0-20 Mbps, European Data Journalism Network). Site crops are multifold:

- Cereals: winter wheat, durum, maize.
- Oil crops: rape seed, sunflower.
- Legumes: Pea and Soybean.

2.1 LOCATION DESCRIPTION

The IPZ is in Saxony-Anhalt on the southern edge of the "Magdeburger Börde" region, between Halle and Magdeburg directly on the A14 motorway at the immediate vicinity of the Strenzfeld campus of the Anhalt University of Applied Sciences and the Saxony-Anhalt State Institute for Agriculture and Horticulture (Figure 2). For more information, please visit the IPZ`s webpage (www.dlg-ipz.de/en/ueber-uns/location).

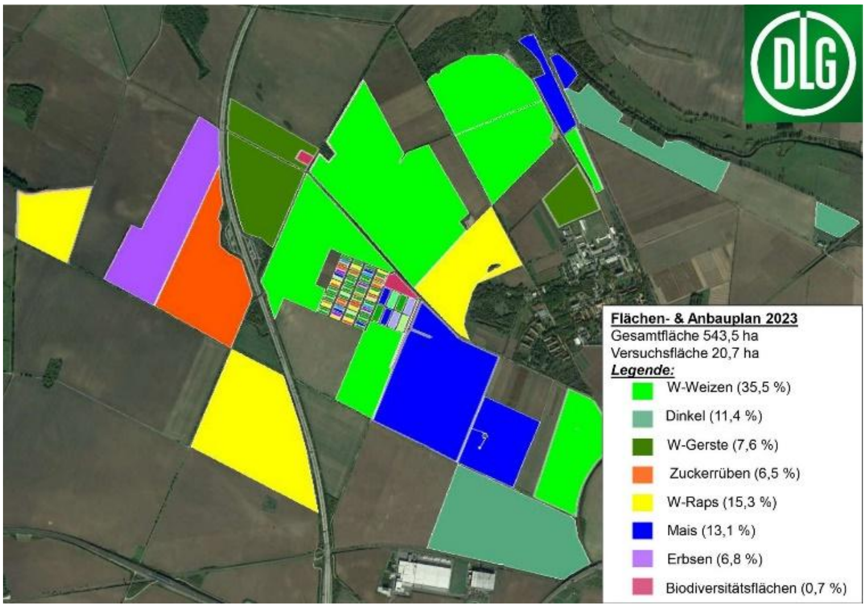


FIGURE 2: MAP OF THE IPZ FARM

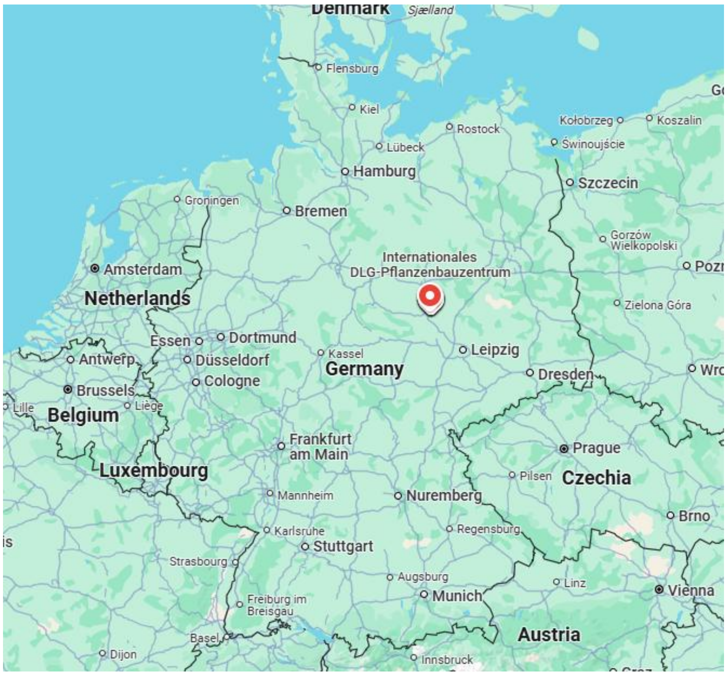


FIGURE 3: LOCATION OF THE IPZ IN GERMANY

The location is characterized by the rain shadow of the Harz Mountains. At an altitude of 80 meters above sea level, the average annual precipitation is 511 mm, and the average annual temperature is 10.1 °C. The predominant soil type is loess-black earth with a usable field capacity of 250 mm and 65-94 soil points. The soil type can be described as silty loam with shares of clay: 22 %, silt: 70 %, sand: 8 %. The inclination is flat and the pH is slightly above neutral. All this information is presented within IPZ`s webpage (www.dlg-ipz.de/en/ueber-uns/location).

Assuring effective agricultural management and services the IPZ is equipped with its own machinery such as tractor(s), ploughs for tillage, cultivation, fertilization, crop protection and harvesting.

It is supplemented by co-operative use of the field trial technology of the site partners Anhalt University of Applied Sciences Campus in Strenzfeld and the State Institute for Agriculture and Horticulture, as well as by external trial service providers.

2.2 PARTICIPANT

In the context of the German demonstration site, Nostradamus stakeholder engagement and management activities build upon the stakeholder groups defined in the Grant Agreement. The Nostradamus Grant Agreement identifies more than 30,000 stakeholders across multiple sectors, including academia, the public sector, the private sector, and practitioner communities, which provide a broad and representative ecosystem for engagement at national and regional levels in Germany. - (i) involvement of scientists (e.g. University of Applied Sciences Anhalt; Federal Research Centre for Cultivated Plants), (ii) public agencies (e.g. State Institute for Agriculture, and Horticulture Saxony-Anhalt, Lower Saxony Chamber of Agriculture), (iii) private sector (e.g. KWS, SKW Piesteritz, Claas), (iv) start-ups (e.g. Biodivers), (v) NPOs (e.g. Bitkom), and (vi) numerous farmers.

2.3 DATASETS

Data collected at the IPZ referring to own projects and regular work are not shared publicly as such but listed in the respective list of the Nostradamus project. However, based on collaborative activities defined in detail as part of WP4. Apart from these projects, the DLG IPZ coordinates and participates in projects funded by public sources. Related data are available and refer to following projects:

ReNuWi - Row-based arable farming with reduced chemical crop protection and promotion of beneficial organisms and arable wild herbs in the field (ReNuWi); Data set: Agronomic data (e.g. inputs, yields, wild herb observations),

IWANA - Integration of valuable arable wild herbs in intensive cash crop cultivation; Data set: Agronomic data (e.g. inputs, yields, wild herb observations),

NaLamKI - Sustainable agriculture with artificial intelligence; Date set: Agronomic data (e.g. inputs, yields, wild herb observations), UAV data of wheat stands to monitor plant performance (e.g. pests), AgriSens DEMMIN 4.0 - Agronomic data (e.g. inputs, yields, wild herb observations), UAV data of wheat stands to monitor plant performance (e.g. pests) and soil moisture,

TRANSFORM - Smart Transformation Labs; DLG as subcontractor, which limits access to related data set referring to cropping data used to develop digital twins of farms. Date sets: Agronomic data (e.g. input factors, yields of wheat and legumes) and irrigation data (subsurface irrigation experiment).

2.4 DATASET TIMELINE

Refers to the data sets and projects mentioned in 2.3. ReNuWi data sets are available. All other projects are in execution requiring an in-depth check to identify and ideally prepare data needed by Nostradamus.

3 DEMONSTRATION SITE IN SERBIA – RS (BIOS, CPT)

Vojvodina is an autonomous province situated in the North of the Republic of Serbia. It is bounded by Croatian, Hungarian, and Romanian borders to the North and the Sava and Danube rivers to the South. It is located within the Pannonian plain and, besides Fruška gora and Vršac hills. Being predominately a flat land with fertile soils, it is where majority of Serbian agriculture takes place, and it was thus chosen as the pilot site for the Nostradamus project. The total area of the province is 21,614 km2, 80% of which is allocated to agriculture. The main crops are maize, wheat, soybean, sunflower and sugar beet, but there are also increasing areas under fruits and vegetables. The climate is predominantly continental, with hot summers and cold winters providing optimal conditions for a wide range of crops. However, due to climate change, droughts and seasonal weather fluctuations have become more frequent. This underlines the importance of efficient water management and application of precision agriculture technologies, which also have the potential to mitigate the negative consequences of intensive agriculture – soil degradation / organic matter decrease, and underground water depletion and pollution.

3.1 LOCATION DESCRIPTION

Macro Location: Vojvodina's Geopolitical and Geographic Context

The autonomous province of Vojvodina (Figure 3) has 1.75M inhabitants and is the most developed area in Serbia. It is a multinational region with 26 nationalities and 6 official languages. Agriculture plays an important role in the province's industry, with a share of over 40% of the total industrial output and 30% of total exports. There are three major rivers crossing its territory, the Sava, Tisa and Danube. The Danube, being the largest river in Central Europe, is the largest waterway in the region and has a direct link to Western Europe on one side, and Black Sea on the other.

FIGURE 4: MACRO LOCATION OF VOJVODINA, SERBIA[1]

Vojvodina is the largest contributor to Serbia's agriculture exports, especially in terms of grains, apples and oils. However, recently, the exports of ICT services and products is slowly surpassing the level of agricultural exports. Novi Sad, the capital of Vojvodina, is the city with the highest number of developers and engineers per capita and the top ICT hub, along with Belgrade. Here, University of Novi Sad with more than 50k students plays an important role in education and innovation.

Vojvodina is located between approximately 44° and 46° north latitude and 18° and 22° east longitude, forming part of the Pannonian Plain in northern Serbia. The region is characterized by a predominantly continental climate, with hot summers, cold winters, and moderate rainfall throughout the year. This climate, combined with the influence of major rivers like the Danube, Tisa, and Sava, creates favorable conditions for agriculture, though occasional droughts and extreme weather events require efficient water management and adaptation strategies. It is located in the typical continental climate zone, and being its typical representative, is the location of NASA'S Land Cover and Land Use study, implemented by three Michigan universities and BioSense Institute.

Micro Location: Vojvodina's Localized Spatial and Functional Characteristics

With its fertile soils, Vojvodina (Figure 4) is considered the breadbasket of this part of Europe. It hosts black soils (chernozem) in the western part, ideal for soybean and maize, sandy soils in the north, suitable for rapeseed and wheat, and a mixture of terrestrial soils elsewhere, suitable for arable crops, fruits and vegetables. Historically, 1 in 5 years are dry, but with the recent effects of climate change, the frequency of droughts has doubled, while the distribution of rainfall has become

[1] https://www.britannica.com/place/Vojvodina

very uneven and unpredictable. Currently, less than 10% of land is irrigated, and due to the change in weather patterns, an increasing number of farmers are opting for center-pivot or linear irrigation in the case of arable crops and drop-by-drop systems in the case of fruits and vegetables.

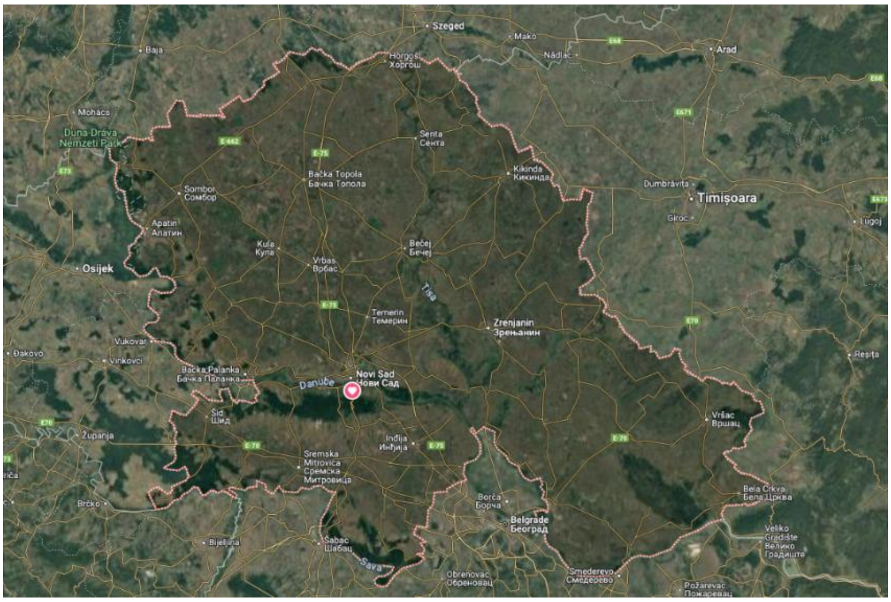


FIGURE 5: MICRO LOCATION OF VOJVODINA

The typical size of the field is around 5 hectares, which is relatively small, with majority of farmers being family farmers. Nevertheless, platforms such as BSI's AgroSense platform and CPT's Space Garden platform are gaining momentum. AgroSense is used at one third of all arable farmland in Vojvodina, while Space Garden is used by the major seed companies, banks, and insurance companies. Digital transformation also has the power to increase the level of uptake of insurance, which currently sits at 5-10% of the total production, with BSI's and CPT's systems being used by major companies, thus slowly increasing this uptake.

Main crops in the region with the corresponding percentage of farmland area are maize (29%), wheat (22%), sunflower (11%), soybean (11%), and others. Due to climate change, crop rotations are changing as well, and farmers are choosing more resilient crops and varieties. In the previous decade, maize and soybean were the most profitable but most susceptible crops to drought, and their areas have shrunk by 15-20%, while wheat, sunflower and controlled, greenhouse production have increased. For this reason, it is very important to reassess crop suitability, produce new maps and optimize the use of inputs.

Geopolitical and Economic Implications

The choice of crops has become influenced by global events a great deal. Commodity prices on Chicago, Paris and Budapest stock exchange markets are strongly influencing local prices, Russia's invasion of Ukraine is disturbing the supply chains in a wider Eastern European region, and the price of inputs is fluctuating due to the high inflation rates and high energy prices. All these are affecting production costs and output prices, and thus the choice of crops.

Similarly, the need for precision agriculture technologies has become pronounced, both due to the instability of the weather and due to the higher prices of inputs. Now, it has become much more reasonable to invest in variable rate application, irrigation systems and infrastructure, as the return on investment has shortened due to the high input costs.

3.2 PARTICIPANT

(i) **Ministry of Agriculture**, including agencies for agricultural payments and agricultural land, and **Provincial Secretariat for Agriculture, Water Management and Forestry**; Achieving better insights in agriculture and more optimal policies.

(ii) **Farmers** – increasing the uptake of precision agriculture technologies and improving resilience to climate change.

(iii) **Agricultural companies** – increasing the uptake of precision agriculture technologies, steering the development of the company and improving resilience to climate change.

(iv) **Banks** – better assessment and insights in agriculture, optimal loan approval policies.

(v) **Insurance companies** – better insights in agricultural production and a better overview of the perils, leading to more optimal operations.

(vi) **AgTech SMEs** – solution providers – better insights in problems, connection with academia, farmers and other stakeholders.

(vii) **Citizens** – better insights and quality information about food production and the impact of agriculture.

3.3 DATASETS

This demonstration site depends on three distinct data sources, which are described below. The datasets are documented in Appendix A.

Dataset 1: IoT data

BIOS has installed dozens of sensors and weather stations around the region in the past 5 years. Data comes at an hourly level and includes parameters such as soil moisture, temperature, precipitation etc.

Dataset 2: Field data

In partnership with agricultural and academic partners, BIOS and CPT collected a database of field operations, with locations, input consumption, yield and other parameters. These datasets are subject to non-disclosure agreements (NDAs).

Dataset 3: Soil data

Soil and terrain data come from various sources, such as SoilGrids, the global soil database, or CGIAR-CSI, the global elevation map. However, there are local databases, some of which were produced by BIOS for the Ministry of Agriculture, which contains Soil Organic Matter (SOM) content at a 10m resolution, or the soil map, which comes in a vector file.

3.4 DATASET TIMELINE

It is available through AgroSense web platform and is updated accordingly.

4 DEMONSTRATION SITE IN SWITZERLAND – CH (WBF)

The Swiss Future Farm (SFF) in Tänikon, Switzerland, is an innovative agricultural hub focused on precision farming, automation, and sustainable practices. It integrates advanced technologies like robotics, digital tools, and data analytics to optimize farming efficiency and reduce environmental impact. The farm serves as a research and collaboration platform, promoting knowledge exchange and innovation in agriculture. It also provides education and training to farmers, showcasing the future of sustainable and tech-driven farming. The SFF encompasses 81 ha of agricultural areas and 65 dairy cows (2/3 Brown Swiss, 1/3 Red Holstein and Holstein Frisian). The SFF 2023 vision is declared as follow: SFF is a Swiss leader in the testing, evaluation and communication of new processes and the use of technologies that support ecological, economical and robust food production.

4.1 LOCATION DESCRIPTION

The Swiss Future Farm is an innovative agricultural project located in Tänikon, Switzerland. It serves as a demonstration and research facility focused on exploring and implementing advanced agricultural technologies and practices. The farm aims to address key challenges in modern agriculture, such as sustainability, efficiency, and the integration of digital technologies.
Key aspects of the Swiss Future Farm include:
Precision Farming: The farm utilizes precision farming techniques, which involve the use of GPS, sensors, and data analytics to optimize field-level management regarding crop farming. This helps in reducing waste, increasing yields, and minimizing environmental impact.

- **Automation and Robotics**: The farm incorporates automated machinery and robotics to perform various tasks such as planting, weeding, and harvesting. This reduces the reliance on manual labor and increases efficiency.
- **Sustainable Practices**: Emphasis is placed on sustainable agricultural practices, including the use of renewable energy sources, efficient water management, and organic farming methods.

- **Digitalization**: The farm leverages digital tools and platforms for farm management, including software for monitoring crop health, soil conditions, and weather patterns. This data-driven approach helps in making informed decisions and improving overall farm productivity.
- **Research and Collaboration**: The Swiss Future Farm serves as a hub for research and collaboration among farmers, researchers, and technology providers. It aims to foster innovation and knowledge exchange in the agricultural sector.
- **Education and Training**: The farm also plays a role in educating farmers and the public about the benefits and implementation of advanced agricultural technologies. It offers training programs and workshops to disseminate knowledge and best practices.

The 81 ha of agricultural areas are split as follows:
- 55 ha arable farming
- 20 ha grassland
- 6 ha biodiversity area

4.2 PARTICIPANT

AGCO International GmbH: Leading manufacturer of high-tech solutions for farmers. Brands: Fendt, Valtra, Massey Ferguson, Precision Planting.
- Arenenberg Agricultural education and extension center of the Canton of Thurgau with three school- and experimental farms.
- GVS Agrar AG Market-leading importer of agricultural machinery in Switzerland. Import, sales and service for all AGCO brands.

4.3 DATASETS

Several datasets are available for accessing national-level data for Switzerland. These datasets are listed below.

- Meteoswiss www.meteoschweiz.ch (weather data, forecasts)
- Agrometeo www.agrometeo.ch (weather data, pest forecast models)
- Swisstopo www.swisstopo.ch (altitude models, geographical infos)
- Soil classification data available on multiple sources
- Stationen | Bewässerungsnetz
- (soil moisture sensor network)
- Field boarders: https://map.geo.tg.ch/

Further, the data related to the SFF can be clustered into 2 parts:
- Field data: Mostly coming from the FIMS.
- Dairy data: all are aggregated in one software called (UNIFORM agri).

Both datasets are documented in Appendix A.

4.4 DATASET TIMELINE

We do not have control over the publicly available datasets (Meteoswiss, Agrometeo, SwissTopo ...). However, for our locally acquired datasets, the timeline is as follows:

- Dairy dataset: this is updated on daily basis, and more details can be found in the annexed table.
- Field dataset: weather station data every 10 minutes, satellite images on weekly basis. On the other hand, the field data do not change often unless there is an operation. Other variables are updated on event basis.

5 DEMONSTRATION SITE IN SLOVENIA – SL (ITC, KGZS)

The demonstration site is located in the Pomurje region in the north-eastern part of Slovenia. The Pomurje region borders Austria, Hungary, and Croatia. The plain along the Mura River is a typical marginal Pannonian plain, with fertile agricultural land stretching along the riverbanks and hilly terrain to the north. Despite covering only 6.6% of Slovenia's territory, this area represents the most significant agricultural landscape in the country, with 22% of fields, 25% of orchards and vineyards, and 8% of meadows and pastures. However, the small farm structure, high land fragmentation, low bargaining power of farmers, and moderate factor productivity (slightly below EU average) imply very low incomes from agriculture for most farmers. Data-based decision support systems would improve the competitiveness of agriculture in Pomurje. Based on the WDI, SL presents the third highest fertilizer consumption across EU countries. Site: crops: wheat, barley, oat, spelt, corn. Instrumentation plan: More than 20 weather stations and IoT devices like soil moisture and temperature sensors are implemented in the Pomurje region but also machinery data will be used.

5.1 LOCATION DESCRIPTION

Macro Location: Slovenia's Geopolitical and Geographic Context

Slovenia is situated in the southeastern part of Central Europe, precisely at the intersection of several major European regions. Its geographic coordinates place it between latitudes 45° and 47° N and longitudes 13° and 17° E. This positioning makes Slovenia a pivotal transit and trade point within the broader European framework. It borders Austria to the north, Italy to the west, Hungary to the northeast, and Croatia along its southeastern and southern borders, with a total land boundary length of 1,370 kilometers.

The country is part of the **Alps-Adriatic region**, a geostrategic area that connects Central Europe with the Mediterranean and the Balkans. Slovenia is geographically diverse, encompassing portions of the Eastern Alps, the Dinaric Alps, the Adriatic Basin, and the Pannonian Plain. Its inclusion in the

Pan-European Transport Network, specifically **Corridor V** (connecting Venice–Trieste–Ljubljana–Budapest–Kiev), enhances its macroeconomic relevance, facilitating freight and passenger movement across Europe.

Slovenia's proximity to major European cities—such as Vienna, Venice, and Zagreb—places it within a core zone of European connectivity. It is also part of the Schengen Area, further solidifying its role as a logistical and infrastructural bridge in Europe. Slovenia's macro location supports a diverse climate regime, where Alpine, Mediterranean, and continental influences converge, impacting its environmental systems and economic activities, including tourism and agriculture.

FIGURE 6: MACRO LOCATION – SLOVENIA

Micro Location: Slovenia's Localized Spatial and Functional Characteristics

At a micro level, Slovenia occupies a relatively compact area of 20,271 square kilometers, characterized by significant geographic and ecological diversity. This localized diversity stems from Slovenia's position at the intersection of four major European geographic units: the Alpine region, the Dinaric Alps, the Mediterranean Basin, and the Pannonian Plain.

The **Julian Alps** dominate the northwest, with Mount Triglav (2,864 meters) serving as the highest peak and a symbol of national identity. The central region is marked by karstic landscapes, featuring extensive subterranean systems such as the **Postojna Cave** and **Škocjan Caves**, which are UNESCO World Heritage Sites. The Mediterranean coastal area, although limited to 46.6 kilometers, includes important ports like **Koper**, a critical node in maritime trade. The eastern regions transition into the fertile plains of the Pannonian Basin, supporting agricultural development and viticulture.

From an administrative perspective, Slovenia is divided into **212 municipalities**, with **Ljubljana** as the political, economic, and cultural center. Located in the Ljubljana Basin, the capital is strategically positioned at the convergence of major national and international road and rail routes. The local urban network is further enhanced by cities such as Maribor, Koper, and Celje, which provide regional hubs for trade, industry, and tourism.

In terms of infrastructure, Slovenia has developed an extensive network of highways (referred to locally as **avtoceste**) and railways, linking internal regions to neighboring countries. Its main international airport, **Ljubljana Jože Pučnik Airport**, enhances air connectivity, while the Port of Koper serves as Slovenia's primary maritime gateway.

Prekmurje is the northeasternmost region of Slovenia, bounded by the Mura River to the west and sharing borders with Austria, Hungary, and Croatia. Traditionally distinct from the rest of the country, it has its own dialect (Prekmurščina), folkloric customs, and cuisine—gibanica (layered cake) being the most famous. The landscape is a gently rolling alluvial plain, intensively farmed for cereals, vineyards, and orchards, punctuated by small woodlands and meandering streams. Key towns include Murska Sobota (the regional center), Lendava with its multicultural heritage, and Gornja Radgona known for wine fairs. Prekmurje's cultural mix—Slovene, Hungarian, Roma—and its borderland history give it a unique identity within Slovenia.

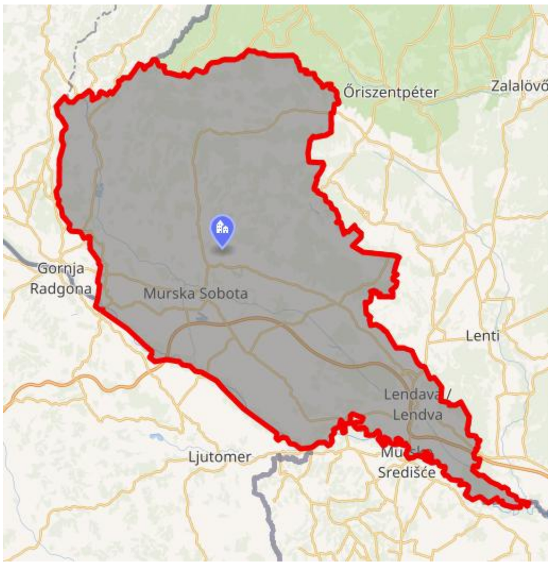


FIGURE 7: MICRO LOCATION - POMURJE REGION

Geopolitical and Economic Implications

Slovenia's micro and macro locations together position it as a critical player in regional and European geopolitics. Its location at the crossroads of Alpine, Adriatic, and Balkan influences facilitates trade, cultural exchange, and international diplomacy. The country benefits from a temperate climate and rich biodiversity, which supports varied industries, including agriculture, forestry, energy, and tourism. Slovenia's well-preserved natural landscapes also provide a competitive edge in sustainable development and eco-tourism.

From a macroeconomic perspective, Slovenia is a vital transit corridor in the European Union, playing a key role in the movement of goods and people across the continent. Its geographic positioning, combined with a developed transport infrastructure, has made it a logistical hub for global supply chains operating in Europe. Furthermore, Slovenia's compact and highly integrated regional geography allows for efficient governance and resource allocation, promoting sustainable development across its diverse terrains.

5.2 PARTICIPANT

(i) Ministry of Agriculture, Forestry and Food.

(ii) Ministry for Digital transformation.

(iii) Institute of Agriculture and Forestry.

(iv) University of Maribor.

(v) Centre for Health and Development Murska Sobota.

(vi) Development Center Murska Sobota.

(vii) Farmers & local food suppliers.

(viii) SMEs – solution providers; (x) Citizens.

5.3 DATASETS

This demonstration site depends on one distinct data source, which is described below. The dataset is documented in Appendix A.

Dataset 1: IoT Dataset

The IoT Sensors dataset contains high-resolution environmental data from a pilot site in Slovenia. It is delivered in JSON format and captures key parameters such as temperature, humidity, air pressure, soil moisture, and leaf wetness, along with detailed timestamps and sensor metadata (including unique IDs, sensor names, and organizational information). Data updates occur every 5–20 minutes, providing near-real-time insights, and the dataset, approximately 1GB in size, is stored in the DIH Agrifood Data Space. Although access is restricted and privacy is maintained through partial anonymization, the dataset is designed for seamless integration with geospatial systems and supports advanced analyses like trend detection and predictive modeling.

5.4 DATASET TIMELINE

It is available through DIH AGRIFOOD Data Space and is updated accordingly.

6 DEMONSTRATION SITE IN CYPRUS – CY (ECOE, MARDE)

The pilot site is about 7 km southeast of Paphos town in the Acheleia Experimental Station (Fig. 1), which is part of the Agricultural Research Institute of Cyprus (ARI), a Center of Excellence in Agricultural Research. ARI is a Department under the Ministry of Agriculture, Rural Development and Environment of the Republic of Cyprus. The vision of the Agricultural Research Institute is to be a model center of knowledge and innovation, and lead Cyprus to a better future by strengthening rural development, improving the quality of life, and ensuring the sustainable use of natural resources. The Agricultural Research Institute conducts research aiming to create and transfer knowledge for the development of the primary sector and to solve problems at the farmer's level. The research results are transferred to stakeholders through modern educational programs and dissemination tools. Its research activity strengthens rural development and contributes to the adoption of a sustainable rural policy and innovation offer.

The Acheleia Experimental Station primarily covers annual grain crops and aromatic plants experiments, and hosts greenhouse facilities. Research at the station also covers viticulture, citrus crops, and tropical-subtropical crops. In addition, there is a screen-house for the maintenance of basic citrus germplasm. In Morokampos area, there is a field of about 7 ha utilized for research activities related to cereals and fresh crops. The Station was / is being used for experiments on:

i. Investigating the impact of the incorporation of plant residues and sludge in cereal crops.
ii. Genetic improvement of barley varieties.
iii. Genetic improvement of Durum wheat varieties.
iv. Evaluation of table grape varieties grafted onto American Rootstocks resistant to *Daktulosphaira vitifoliae* under different environmental conditions.
v. Evaluation of table grape varieties under different environmental conditions, Collection and conservation ex situ of wine and table grape varieties.
vi. Groundnuts varieties evaluation.
vii. Characterization and evaluation of germplasm of durum wheat under Cypriot conditions.
viii. Use of remote sensing for the assessment of the impact of climate changes on crop irrigation needs, especially for crops with high irrigation demands.

6.1 LOCATION DESCRIPTION

Micro Level: The experimental station of Acheleia (Fig. 8) is the largest research station of ARI and is comprised of three main parts, the old station (1), the new station (2) and the Morokampos station (3) with a total area of 51 ha (Fig. 8). The main part of the station is located in Acheleia village in Paphos district, on the southwestern part of Cyprus (Fig. 7). The Station comprises three large plots (Fig. 8). Within the Station there are different types of crops, such as vines, olives, citrus, avocados and other annual and perennial crops. Cereals are grown mainly in the Morokampos area, covering ca. 7 ha (Fig. 8.2), at a seeding rate of 25 kg per hectare. The soil type is clay loam, alkaline with pH ranging from 7 to 8,5 and a deep brown color. The climate of the site is characterized by a mild and wet winter followed by a long dry period. The average annual rainfall in the region is 420 mm, 90% of which falls from October to March. Mean daily minimum/maximum temperatures are 12,4/15,1 °C in January, 16,9/19,3 °C in April and 25,3/34,0 °C in August. Only a small fraction of the land is irrigated.

FIGURE 8: LOCATION OF THE ACHELEIA EXPERIMENTAL STATION

Macro – level: Land used for agriculture in Paphos district covers 26.140 ha, representing 19,5% of the total agricultural land in Cyprus. The Paphos district includes a variety of microclimates due to its varied geography. The coastal areas have warm and relatively dry climate which allows for the growth of avocado, citrus, and grape crops characteristic of the Mediterranean climate. Further inland, the cooler and more humid areas allow the cultivation of potatoes and other vegetables. Regions such as the Troodos Mountains are higher in elevation and experience cooler temperatures, ideal for apples and pears.

Farmers in the Paphos district cultivate annual crops such as potatoes, onions, tomatoes, and other vegetables, while there are areas with extensive banana plantations. In recent years, farmers established several avocado plantations. Wheat and barley production is mostly used for animal feed. Most agricultural activities are small-scale and family-run. The extensive droughts observed in recent years in Cyprus and Paphos have led to a decline in agricultural productivity.

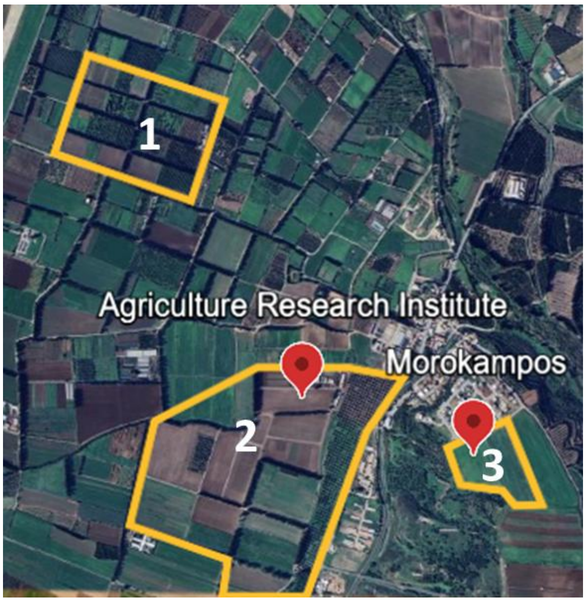


FIGURE 9: MICRO-LEVEL OF ACHELEIA EXPERIMENTAL STATION

6.2 PARTICIPANT

i. Local Farmers: Farmers in proximity to the pilot sites will take part in living labs and learn to use the module.

ii. SME companies are active in agriculture.

iii. Plant Protection Companies: These companies will make use of forecasts based on the phenological phases of the crops to make timely and accurate interventions in problem areas. Their objective is to minimize the application of plant protection products to lessen environmental damage while still maintaining crop health.

iv. Department of Agriculture: The Department will utilize data from the project to create preventive practice during high-risk periods for certain cereal diseases. It will be the contact point for the results of research and the implementation of policies so that farmers can benefit from the research widely.

6.3 DATASETS

Dataset 1: IoT Meteorological data

Meteorological data will be derived from a METER ATMOS 41 meteorological station installed at the demonstration site. ATMOS 41 is a research weather station and measures 12 weather variables: air temperature, relative humidity, vapor pressure, barometric pressure, wind speed and direction, solar radiation, dual precipitation, lightning strike counter and distance. The accuracy of each parameter differs due to the peculiarities of each parameter. The ATMOS 41 comes with an integrated IoT module. Moreover, the data transmission frequency can be adjusted based on the needs (usually 15-minute measurements).

Dataset 2: Soil data

Soil data will be derived from a METER TEROS 12 Advanced Soil Moisture Sensor + Temperature and Electric Conductivity (EC) installed at the demonstration site. Three TEROS 12 will be installed at 10, 30 and 50 centimeters of depth and will transmit data for soil moisture, soil temperature and electric conductivity. The accuracy of each parameter differs due to the peculiarities of each parameter. Moreover, the data transmission frequency can be adjusted based on the needs (usually 15-minute measurements).

Dataset 3: Crop biotic and abiotic stress

Data on abiotic (water and other types of stress) and biotic (pests and diseases) will be collected from the pilot field through visual observation. The presence of symptoms caused by pests, diseases or abiotic stresses (water and nutrient stress) will be recorded with information on the type of symptom, causing organism, symptom severity, area covered in the field and map coordinates.

6.4 DATASET TIMELINE

Data from the meteorological station and soil moisture sensors are recorded on a continuous basis at preset intervals. The data are stored on a datalogger and the IoT meteorological station can directly transfer data to a cloud repository. The frequency of data transmission can be adjusted based on the needs of the module. In addition, there are options for Secure File Transfer Protocol (SFTP) transfer to an SFTP server and then stored to an SQL database. Data can be accessible immediately from the SFTP server or the available cloud repository. Data on biotic and abiotic stress and occurrence will be collected on a weekly basis as needed and will be uploaded on a cloud repository. Data collection will begin with the new growing season and will continue for at least two growing seasons.

7 MODULE 1 – EARLY WARNING SYSTEM FOR DROUGHT AND VEGETATION STRESS MONITORING (DLR)

The early warning system for drought and vegetation stress (Task 4.1) will develop datasets for an early warning system for water stress and drought monitoring, which can be integrated in the Nostradamus cloud-based data hub. The application for drought monitoring and early warning will be based on a combined drought indicator (CDI), comprised of EO information products (surface soil moisture (SSM), vegetation indices (NDVI), land surface temperature (LST)), meteorological data (standardized precipitation index (SPI), forecasts), and IoT data from the agricultural units to track vegetation status, precipitation, soil moisture and evapotranspiration. The application exploits this information to provide spatially explicit early warnings of potential droughts or hydrological stress. The drought alerts, which come with multiple severity levels, will allow farmers and land managers to take appropriate adaptation measures such as adjusting irrigation schedules or planting drought-resistant crops.

7.1 PROBLEM / OPPORTUNITY

Climate change has a severe impact on agricultural production and food security in Europe. Climatic extremes and changing precipitation patterns bear risks for sustainable farmland management. The frequency of droughts has increased in recent years. Detailed spatial information on areas affected by drought and early warnings are required to allow for targeted adaptation and prevention measures.

The drought module from Nostradamus will exploit Earth observation (EO) data and products to provide an application for monitoring and early-warning for vegetation stress and drought conditions. This data can be used to inform decisions on crop selection, irrigation strategies, and other agricultural management practices (AMP). The EO products can also be used to identify areas that are particularly exposed to drought, allowing for targeted interventions and improved preparedness. The drought alerts will come with multiple severity levels, thus allowing farmers and land managers to take appropriate adaptation measures. The application will help to improve crop yields and reduce crop failures, ultimately leading to shielded food security, and environmental sustainability.

7.2 BENEFITS AND KEY BUSINESS OBJECTIVE

The datasets generated by the drought module can provide information about potential droughts or hydrological stress to farmers, insurance companies and other stakeholders. A key business objective of this module focuses on mitigating yield losses for farmers caused by droughts by providing early warning alerts at multiple severity levels, which allow farmers and land managers to take appropriate adaptation measures. Insurance companies and financial institutions can use the data to identify areas that are particularly exposed to drought or to verify which areas are affected by drought events. The data can also be used by government bodies and agricultural associations to support the development of sustainable development plans and recommendations for crop cultivation.

7.3 TECHNICAL CHALLENGES

The combined drought indicator (CDI) approach will be based on anomaly detection for Copernicus Sentinel-2 based NDVI, Sentinel-3 based Land Surface Temperature (LST), Surface Soil Moisture (SSM) of the Copernicus Global Land Service (CGLS) and the Standardized Precipitation Index (SPI) based on European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 data. As for anomaly detection, several years of data must be considered, a large amount of EO data is required.

This implies a technical challenge with respect to storing, handling and processing a large amount of data.

The main challenge for the module development is the proper combination of the information derived from the different input data to allow for differentiation of different degrees of drought severity. Moreover, crop rotation implies a challenge for the identification of vegetation index anomalies. Different crop types also show individual susceptibility to drought in specific seasons or dependent on their crop status. The integration of seasonal crop type maps of the current and previous years and additional agricultural management information (if available from the study sites) would be beneficial to conclude on local drought risk.

7.4 MODULE SIZE / METRICS

The drought module's size and metrics:

- Area covered with mapping: varies dependent on demonstration site
    - E.g. Study area in Germany: ca. 12,100 km^2
- Number of previous years included for analysis: ≥8 years
- Frequency of output dataset generation: bi-monthly
- Spatial resolution of output dataset: 10 m

7.5 EXPECTED IMPACT AND KPI'S

The main impact of the drought module is related to improved and targeted prevention and adaptation measures to droughts. The monitoring and early-warning application for vegetation stress and drought conditions allows farmers to take informed decisions on crop selection, irrigation strategies, and other agricultural management practices. The application will help to improve crop yields and reduce crop failures, ultimately leading to shielded food security, and environmental sustainability.

Key performance indicators (KPIs)

- Automated processing pipeline: 1
- Journal publication: 1
- Conference participation: 1

7.6 DEMONSTRATION PLAN

The plan for module development and demonstration includes the following steps:

- Remote sensing data collection.
- Remote sensing data pre-processing.
- Design of the technological part of the module.
- Development of prototype module.

- Calibration and application of module for one pilot study area.
- Validation of module for pilot study area.
- Further development of module.
- Calibration and application of module for second study area.
- Validation of module for second study area.
- Integration with Data Cube.
- Integration with User Interface.
- Testing and Validation.

7.7 DATA REQUIREMENTS FOR MODULE

The following data are required for the drought monitoring and early warning module.

Remote sensing data (Raster data)

Time-series of remote sensing data are required as core input data for the module. Data from several previous years up to most current data is required for up-to-date anomaly detection. The required input data comprises time-series from Sentinel-2 for calculation of vegetation indices (e.g. NDVI), Sentinel-3 LST (land surface temperature information), and CGLS SSM (surface soil moisture information). The respective data are freely available from ESA/CGLS.

Meteorological data (Raster data)

Meteorological data for several previous years up to most current data are also required as core input data for the module. The data are available from the ECMWF. ERA-5 data are required to calculate the SPI (standardized precipitation index). Additionally, the HRES product is required as it provides meteorological forecast information needed for early warning.

Spatial information input data / maps (raster data, vector data)

Further input data required for the module include a land cover/land use map (seasonally/yearly) to identify agricultural areas. Crop type maps (updated seasonally, also relevant from previous years) are needed for more precise drought monitoring and early warning. A soil type map (constant) can be used to separate different degrees of severity of soil water anomalies for different soil types. This data are requested from the pilot sites.

Vector input data for AOI (polygon vector data)

A polygon vector data file giving the area of interest (AOI) is required to locate the study area. It may include individual agricultural fields. This information is requested from the pilot sites.

Data for calibration / validation (Field data / IoT data / maps / statistics / information)

Additional data for calibration and validation are essential to be able to investigate the quality of remote-sensing based input parameters and to calibrate and validate the module. If possible, the data should cover a larger area or have been collected at several points spread over the pilot area.

To calibrate and validate input and output data time-series, also data measured over longer time-periods and data from previous years are useful. The data are requested from the pilot sites.

Required/useful data for calibration and validation:

- Meteorological station data (Air temperature, precipitation, solar radiation, wind speed, relative humidity).
- Field measurements (soil moisture, soil surface humidity, evapotranspiration, plant chlorophyll content, phenological status, land surface temperature).
- Plant characteristics and agricultural management information (e.g. crop type, sowing data, harvesting date, irrigation management).
- Spatial information/maps or statistical data (e.g. areal damage of crops, yield maps, crop yield statistics, yield failure, information about timing and areas affected by previous droughts).
- Additional information (specific soil characteristics, specific plant characteristics).
- Weather prediction/forecast data (for the meteorological parameters listed above).

7.8 DATASET WORKFLOW

Figure 9.1 gives on overview on the envisaged dataset workflow for the drought module. Figure 9.2 provides a timeline for task 4.1 showing when the required data will be utilized for the drought module development.

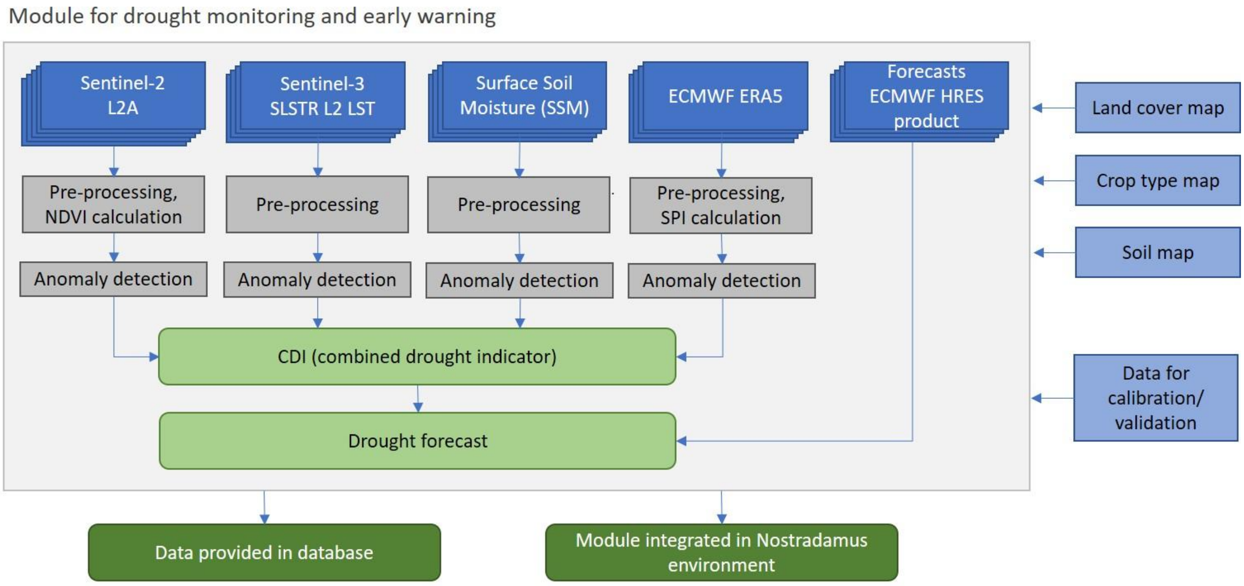


FIGURE 10: DATASET WORKFLOW FOR THE MODULE FOR DROUGHT MONITORING AND EARLY WARNING

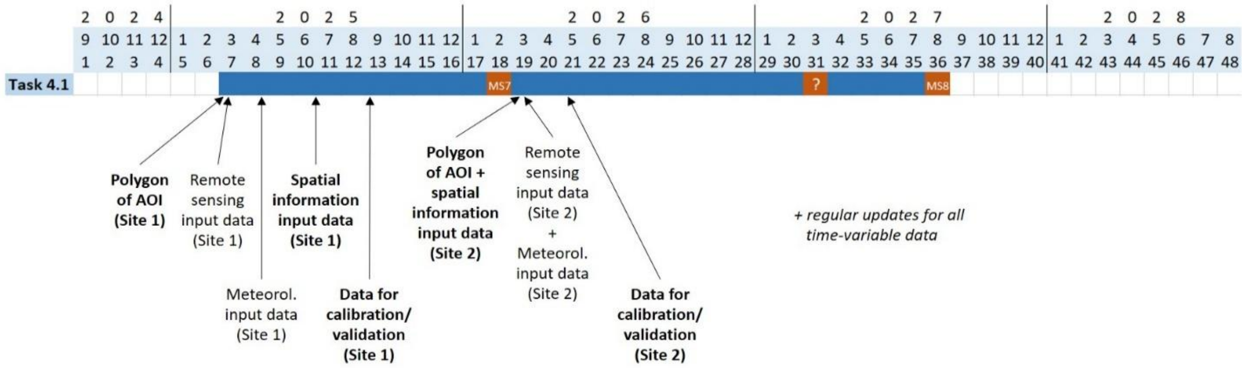


FIGURE 11: TIMELINE FOR TASK 4.1 SHOWING WHEN THE REQUIRED DATA WILL BE UTILIZED FOR THE DROUGHT MODULE DEVELOPMENT.

8 MODULE 2 – CROP SUITABILITY AND AGRICULTURAL YIELDS ESTIMATION (BIOS)

At the heart of this module is a machine-learning driven model that simulates the growth of crops in various environments. It takes soil, weather, elevation, satellite and other data as inputs and tweaks the predicted yield according to the local parameters. Data sources include globally available databases provided by European Space Agency (Sentinel-2 satellite imagery), World Soil Association (SoilGrids), CGIAR-CSI elevation database, Copernicus Climate Change Service and others. The reason behind using globally scalable inputs is because they allow for global scalability of the system. However, as global datasets are often not as accurate as the local ones, we are also relying on local data sources, such as sensor networks, national/regional soil maps and others. In this way, we are facilitating both on-the-cloud scalability and on-the-edge precision, for various stakeholders.

In the development phase of the project, yield prediction models will be trained on the aforementioned open datasets and the proprietary field data including locations, yields and data about field operations. The trained model will be run in multiple weather scenarios, to infer the yield amounts and risks associated with drought, heavy rainfall and other byproducts of climate change.

Testing the system will be done by comparison against the historical datasets, to infer performance metrics, but the resulting maps will be also shown to agronomic experts for comparison with the common knowledge and usual growing practices in the region. A similar methodology for testing will be conducted in other pilot sites where the module will be scaled (outside Serbia). The end results will come in the form of maps of yield potential and crop suitability for the primary region of interest (Vojvodina, Serbia), and for the pilots in which the technology will be replicated.

8.1 PROBLEM / OPPORTUNITY

Climate change is having a severe effect on agriculture in the region. Historically, on average, 1 in 5 seasons had suffered from drought, while others followed a standard pattern with higher rainfall amounts in the spring and autumn and a warm summer with occasional rains that help plants survive through moderately high temperatures. This has, however, changed in recent years. What we observe is that at least 2 out of 5 seasons are hit by droughts, while the rainfall and temperature patterns have become very non-standard. The change in climate and its unpredictability has thus affected agriculture in several ways. The choice of crops and hybrids is leaning to more resilient ones with the ability to survive even in harsh growing conditions and there is a shift happening in growing practices. However, this shift is based on intuition, rather than data. For this reason, the complex interplay between climate, soil and plant growth must be modelled using AI.

8.2 BENEFITS AND KEY BUSINESS OBJECTIVE

Information about crop suitability and yield potential is very interesting, not just for the farmers, but to a variety of stakeholders. Financial institutions such as banks and insurance companies use this kind of data for assessing the risk profile of farmers, and adjusting their offers, premiums and payouts. Agricultural companies, such as traders and distributors, can use this information as well to assess the trends and adjust their strategies, while government bodies can use it to adjust national and regional policies and support farmers in transition to a more resilient type of agriculture. Last but not least, seed companies can use the information provided to steer the development of new varieties and hybrids and for adjusting their sales and distribution

8.3 TECHNICAL CHALLENGES

The main challenge of this module is to connect the soil, weather, elevation and other inputs with the productivity of crops. The machine/deep learning model needs to be trained on field data and predict the yield with sufficient accuracy for implementation in real-world applications. However, there are initial models that show promising performance, and which will be calibrated using new data. The following challenge of scaling up the model to another geography will probably be the most vital one to solve. It will probably require domain knowledge of the season, planting and harvest dates, and agronomic practices, which will need to be incorporated in the model, to align for different yields and different time frames of crop growth. Another technical challenge might be the scale of the system in operation. Producing maps for a large region might require large storage and large processing power for running the algorithms and developing the final maps. To enable this, model pruning, or dimensionality reduction might be applied.

8.4 MODULE SIZE / METRICS

- Number of weather stations: > 10
- Number of parameters measured: 8
- Frequency of data acquisition from weather stations: 1h
- Area covered with mapping: 17,000 km2
- Number of crops analysed: 6
- Number of historical agricultural seasons analysed: > 6

8.5 EXPECTED IMPACT AND KPI'S

Due to climate change, the weather and agricultural seasons have become more unpredictable. That is why expert knowledge has its limits and crop suitability needs to be reassessed, especially in the context of drought and changing rainfall patterns. New maps will help farmers choose more suitable crop rotations, the government to support them in this, agricultural companies to prepare better for the upcoming seasons and financial companies to optimise their strategies.

KPIs:

1. Number of weather stations integrated in the system > 10
2. Yield potential maps produced for >5 crops for the region of interest
3. Crop suitability maps generated for >5 crops for the region of interest

8.6 DEMONSTRATION PLAN

Data collection and IoT device integration: we will collect field data along with data from open data sources and integrate them into a unique database

- **Generation of the machine-learning ready data:** we will perform preprocessing and filtering of the data to overcome the missing values and other problems. Next, we will calculate the input features for the machine learning models, for a clear input-output scheme required by the models.
- **Implementation and testing data cube:** we will run the algorithms and generate the results in form of values for the locations of interest and regional maps for the regions of interest. These will be tested and assessed according to both statistical measurements and by expert analysis, which may lead to model recalibration, hyper-parameter tuning, reassessment of input features or the choice of another model.

8.7 DATA REQUIREMENTS FOR MODULE

Dataset 1: IoT data

The IoT Sensors dataset contains meteorological data from pilot sites in northernmost part of Serbia, Vojvodina captured by weather stations deployed in agricultural environment. It is stored in PostgreSQL database and can be further provided in different data formats such as JSON. It covers key environmental parameters that are influencing plant development including air temperature and humidity, wind direction and speed, UV radiation, illumination and precipitation. Additionally, it contains technical information such as weather station ID and battery status, as well as precise timestamps of weather station measurements. Data samples are generated every hour and transmitted through LORA systems, optimised for sending small batches of data over vast distances with small energy consumption. The entire dataset occupies less than 1GB of memory and has been collected using more than 10 weather stations over 6 years period. It covers an area of 17,000 km^2

with 6 agricultural crops being included. Therefore, it provides a highly valuable resource for advanced data-driven analyses designed to deliver critical solutions for different users.

Dataset 2: Field data

Developing the yield prediction and crop suitability module requires a lot of data for model training. This kind of data has been collected in partnership with agricultural companies, farmers, seed companies, insurance companies and others, and it is subject to NDAs. The dataset consists of a comprehensive list of parameters, which cover all relevant field properties and activities. They can be classified in a few groups, with bold parameters being essential, and others good to have, but not essential.

- Farm info (size of the field, cadastral number, **GPS location**...).
- Sowing (**season/year**, **crop**, variety/hybrid, planting date...).
- Field operations (type, amount and time of application of fertiliser and plant protection products...).
- Harvest (harvest date, **yield**).
- Cost analysis (input costs, operational costs, cost of produce sold...).

Dataset 3: Soil data

Soil and terrain data come from various sources, such as SoilGrids, the global soil database, or CGIAR-CSI, the global elevation map. However, there are local databases, some of which were produced by BIOS for the Ministry of Agriculture, which contains Soil Organic Matter (SOM) content at a 10m resolution, or the soil map, which comes in a vector file.

8.8 DATASET WORKFLOW

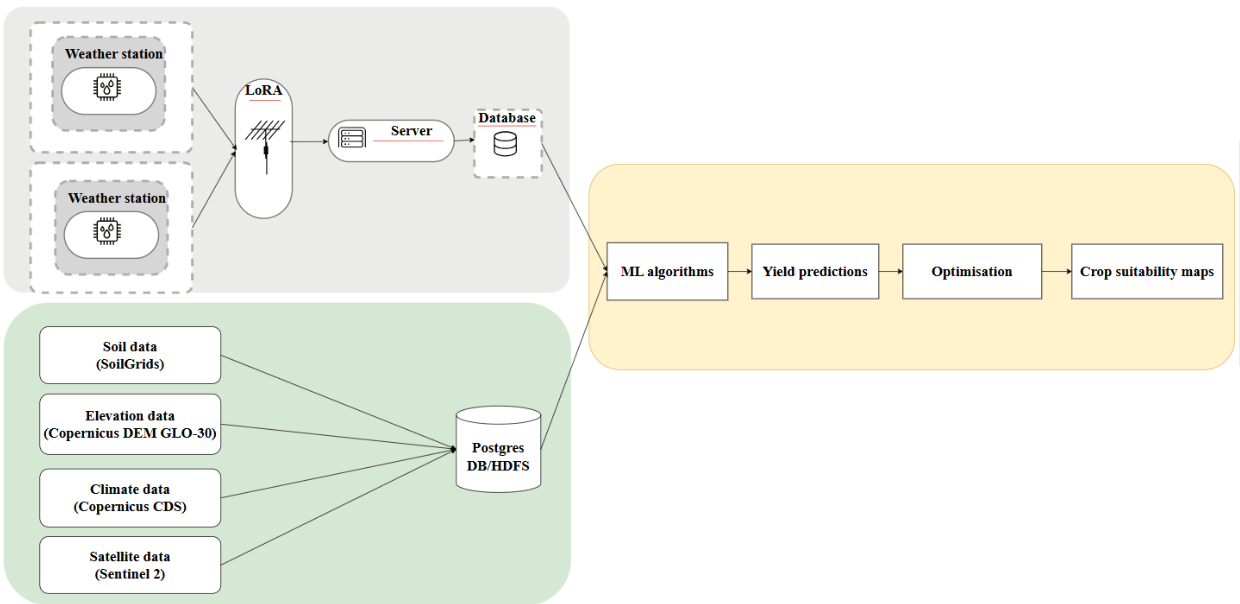


FIGURE 12: DATASET WORKFLOW FOR MODULE 2

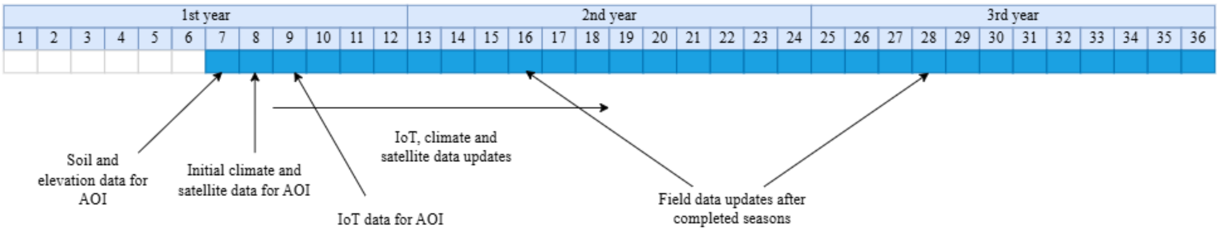


FIGURE 13: TIMELINE DETAILING WHEN THE REQUIRED DATA WILL BE USED

9 MODULE 3 – SOIL FERTILITY AND CROP NUTRITIONAL MANAGEMENT (ITC)

Task 4.3 will develop modules for soil fertility and crop nutritional management utilizing data and datasets to monitor plant growth and make recommendations for crop and site-specific actions like targeted fertilization with right fertilizer type and quantity at optimal time. The task is based on Earth Observation data from satellites and drones, soil and plant related data taken from the field and analyzed in laboratory, sensor data from IoT devices that measure environmental parameters, and other more innovative approaches such as the photons flow in the chlorophyll field. The module will provide farmers and agricultural advisors with recommendations on the appropriate time, amount and type of fertilizer to apply (using advanced precision fertilization equipment), as well as identify hotspots that require special treatment with agrochemical applications. It also will provide information on crop rotation, and other used practices that can help to improve soil fertility and crop nutrition to optimize crop yields and reduce the environmental impact of fertilization and application of Phytopharmaceutical products, aligned also with the CAP objectives for supporting use of precision fertilisation and spraying.

9.1 PROBLEM / OPPORTUNITY

Nutrients such as nitrogen (N) and phosphorus (P) are essential elements for life and an important natural resource. Human activities have significantly altered natural N and P cycles, causing nitrate pollution in drinking water, airborne particulate matter, eutrophication (suffocating life underwater), the loss of biodiversity in lakes, rivers and seas throughout the EU and impacts on soil biodiversity, productivity and higher greenhouse gas emissions. In Europe, surplus N and P in the environment are already exceeding safe planetary boundaries (for N by a factor of 3.3 and for P by a factor of 2). This represents a severe threat to nature and to the Earth's climate. Two thirds of the excessive N and P levels in water originate from fertilisers in agriculture and a third comes from industrial and domestic wastewater.

Development of module for soil fertility and crop nutritional management utilizing data and datasets to monitor plant growth and making recommendations for crop and site-specific actions like targeted fertilization at optimal time and have positive impact to environment. By using AI, the module will be based on EO data from satellites and drones, soil and plant related data taken from the field and analysed in laboratory, sensor data from IoT devices that measure environmental parameters, and other innovative approaches such as the photons flow in the chlorophyll field.

The modules will provide farmers and agricultural advisors recommendations on the appropriate time, amount and type of fertilizer to apply (using advanced precision fertilization equipment), as well as identify hot spots that require special treatment with agrochemical applications. It will also provide information on crop rotation, and other practices to improve soil fertility and crop nutrition to optimize crop yields and reduce the environmental impact of fertilization and application of phytopharmaceutical products, aligned with the CAP objectives.

9.2 BENEFITS AND KEY BUSINESS OBJECTIVE

Benefits and business objectives:

- **Improved user satisfaction by reducing the farmers' costs related to spraying of agricultural areas:** The spraying of large areas, even if it would be enough to spray only hotspots, leads to higher cost related to treatment products itself but also higher cost for fuel and maintenance of machinery due to wear. Some applications are very time sensitive and identified pests or diseases have to be treated very fast to prevent spreading and causing even more damage. If targeted spraying based on satellite or drone images is used, affected areas can be identified earlier and farmers can cover more hot spots in shorter time period and by this prevent the spreading of pest/disease to larger areas. This has direct impact on the produced quantity but also quality of harvested plants.
- **Increased staff satisfaction by reducing their workload:** Due to the high number of fields on single farms and different cultures on parcels, it is not possible for farming advisors or farmers to keep daily track of developments and growth, on each parcel. Algorithms that could identify irregularities and warn the advisors on areas that need special attention are of great value especially for targeted spraying (or similar techniques) as well as for optimizing decision support systems.
- **Improved Soil Health:** Recommendations for targeted fertilization, crop rotation, and other sustainable practices will enhance soil fertility over time, preventing nutrient depletion and maintaining long-term agricultural productivity.
- **Environmental Sustainability:** The project helps mitigate the negative environmental impacts associated with excessive fertilizer use, including reducing nitrogen (N) and phosphorus (P) pollution in water bodies, air pollution from particulate matter, eutrophication, and greenhouse gas emissions.
- **Economic Competitiveness for Farmers:** Improved decision-making support can enhance the competitiveness of farmers, particularly in the Pomurje region, where farm structures are small and fragmented. This can lead to better financial outcomes, improved productivity, and increased incomes.
- **Data-Driven Decision Support:** By leveraging field data, satellite imagery, and IoT sensors, the module can provide real-time, site-specific recommendations that aid farmers in making informed decisions, improving overall farm management efficiency.

9.3 TECHNICAL CHALLENGES

The module seeks to address several technical challenges in developing advanced modules for soil fertility and crop nutritional management. These challenges include the integration of data from diverse sources such as satellites, drones, IoT devices, and soil and plant sensors etc. Achieving real-time data fusion and analytics for precise recommendations on fertilizer application is critical. Additionally, identifying nutrient-deficient hotspots and providing actionable insights on crop rotation patterns require sophisticated machine learning models and algorithms. The system must also address the complexity of optimizing crop yields while minimizing environmental impacts, aligning with the stringent CAP (Common Agricultural Policy) objectives for precision fertilization and spraying. Overcoming these challenges will demand innovations in data processing, IoT networking, and AI-driven decision-making to ensure accuracy and scalability.

9.4 MODULE SIZE / METRICS

- Number of sensors: > 20
- Number of soil tests: > 20
- Number of farm involved in the experiment: > 1
- Number of experiments in the field: > 20

9.5 EXPECTED IMPACT AND KPI'S

- Demonstration events >= 1
- Sources of data to feed Nostradamus Data Cubes >= 1
- Farmers who will participate >= 5

9.6 DEMONSTRATION PLAN

- **Data collection and demonstration environment:** All data sources will be collected and deployed based on dataset description in next section.
- **Generation of the learning data:** ITC will conduct experiments in the fields with domain experts and IoT data from which the generated knowledge will be shared with consortium.
- **Implementation and testing data cube:** ITC will test and (if needed) implement data cubes into DIH AGRIFOOD infrastructure.

9.7 DATA REQUIREMENTS FOR MODULE

Field Data Dataset:

This data set provides comprehensive geospatial and attribute data for agricultural fields. Each field is defined by its polygon geometry, which accurately represents its spatial boundaries. Key technical details include unique identifiers (such as gerk_pid, kmg_mid, and raba_id), and metadata describing field usage (e.g., crop production, greenhouse, or other purposes). Additional spatial attributes encompass the field's area (in ares), perimeter, centroid coordinates, and bounding box parameters. The dataset is available in SQL, JSON, and GeoJSON formats, ensuring compatibility with various geospatial and database systems. It is updated on-demand or upon changes in field metadata, making it a dynamic resource for precision agriculture, spatial analysis, and land management.

Lab Analysis Dataset:

This dataset captures detailed soil analysis data obtained from field samples. It includes multiple layers of technical information: a unique sample identifier (id_vzorca), result identifiers (id_rezultata, id_parametra), and comprehensive measurement details such as the parameter name (parnaziv), measured result (rezultat), measurement unit (enota), and threshold ranges (mejna_vrednost). The data is structured to allow integration with other agronomic systems and supports longitudinal analysis by including timestamps (datum_sprejema). Available in JSON, CSV, and SQL formats, the dataset is updated annually or when new laboratory analyses are performed, enabling researchers and agronomists to monitor soil quality and nutrient dynamics effectively.

IoT Sensor Dataset:

This dataset aggregates high-resolution environmental and meteorological measurements captured by IoT sensors deployed in agricultural settings. Each record includes sensor metadata (e.g., sensor ID, sensor name, and display name), operational status, and organizational details (organization and suborganization identifiers and names). The temporal resolution is high, with data captured every few minutes, and includes precise timestamps (dateCaptured, dateCreated) along with breakdowns by year, month, and day. Measured parameters encompass temperature, air pressure, humidity, soil moisture, leaf wetness, and battery level, among others. The dataset follows a strict JSON schema with defined data types (e.g., UUIDs, integers, doubles) to ensure consistency and reliability. Distributed in JSON format, it supports advanced applications such as the calibration of numerical weather prediction models and the development of soil mapping algorithms, serving as a critical tool for farmers, advisors, and policy makers.

9.8 DATASET WORKFLOW

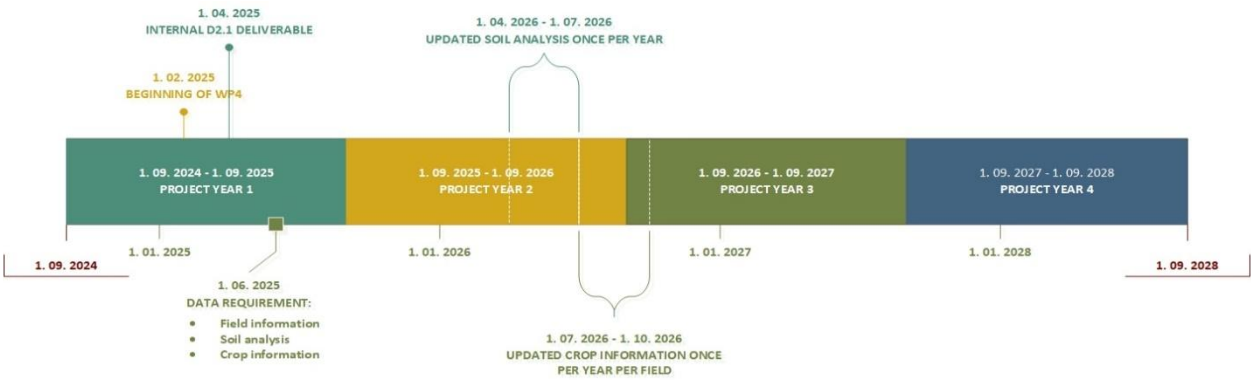


FIGURE 14: DATASET WORKFLOW FOR MODULE 3

10 MODULE 4 – SUSTAINABLE AND EFFICIENT PEST MANAGEMENT (ECOE)

In this module a digital early warning system for cereal pests and diseases will be developed. The digital application will collect near-real-time data from an IoT meteorological station, soil moisture sensors and satellites to provide early warnings to farmers on biotic and abiotic stresses for cereal crops. The use of early warning systems relying on near-real time data and satellite images enables farmers to be proactive and apply control measures at an early stage, leading to more favorable pest management outcomes. Furthermore, early warning systems can lead to a reduction of pesticide use in the fields, while safeguarding yields and leading to an increase in the income of farmers.

The development process of the module incorporates co-design and co-creation with stakeholders. The prototyping, and continuous updating and validation will ensure effectiveness and accuracy in real agricultural conditions. This module is set out as an essential part of the innovation of the pest management practices, and it is also aligned with the European Union's Farm to Fork Strategy that aims at reducing pesticide use by 50% by 2030.

10.1 PROBLEM / OPPORTUNITY

Problem

- Cyprus faces a climate crisis manifested by extended dry spells and extreme weather events.
- The extreme climatic conditions place cereals under substantial biotic and abiotic stresses.
- Farmers increase the use of agrochemicals, such as pesticides and fertilizers, to alleviate the effects of plant stress.
- The EU aims at reducing pesticide use by 50% and fertilizer use by 20% by 2030.
- The EU depends on third countries for pesticide and fertilizer imports

Opportunities.

- Provision of digital solutions (early warning system) to the agricultural industry.
- Early detection and proactive management of pests, diseases and abiotic stresses.

- Avoidance of yield losses due to diseases, pests and abiotic stress.
- Decreasing the use of pesticides and fertilizers.

In recent years Cyprus experiences a climate crisis, exemplified by frequent drought and extreme weather events. The climate of the island, as well as of other parts of the EU is becoming sub-optimal for the growth of certain crops, placing plants under significant abiotic stresses, such as plant water and heat stress. In addition, the changing climate favors pests and diseases that benefit from warmer conditions and experience higher population growth rates. In addition, plants growing under abiotic stress are less defended and more prone to attacks by pests and diseases, which damage crops and reduce yields. Farmers rely on agrochemical inputs, such as pesticides and fertilizers to alleviate or mitigate the effects of stress and protect their yields. The EU is highly dependent on third countries for imports of pesticides and fertilizers, which raises concerns about sustainability and food security in the Union and Cyprus.

Cereal crops cover ca. 50% of the agricultural land of Cyprus. Cereal crops in Cyprus face significant pressures for extreme weather events and pests and diseases. Such biotic and abiotic stresses are effectively contributing to yield losses and lead to dependence on inputs that compromise the environment and the sustainability of agriculture. However, the challenges represent a unique opportunity to manage pests through innovative digital solutions.

The current module will use advanced technologies to provide farmers with near real-time information to predict the impact of biotic and abiotic stresses. In addition, the use of early warning tools will also reduce the use of pesticides and fertilizers, as farmers will be able to predict problems and proceed to solve them. The integration of digital tools (Early Warning System) in agriculture leads to more sustainable pest management with a reduction in pesticide use. This approach helps in the achievement of the EU goals on environmental sustainability. By addressing the above critical issues digital tools will revolutionize pest management strategies ensuring a more sustainable and productive future for agriculture industry in Cyprus and beyond.

10.2 BENEFITS AND KEY BUSINESS OBJECTIVE

Identify biotic and abiotic stress at an early stage

- Provide early warning and management recommendations for cereal abiotic and biotic stresses.
- Reduce reliance on agrochemical inputs.
- Enhance robustness and sustainability of agricultural industry.
- Reduce production costs.

The key business objectives of this module focus on mitigating yield losses for farmers caused by biotic and abiotic stresses with the help of advanced digital solutions. Precision farming practices have transformed the agriculture industry, enhancing its strength and sustainability. Identifying the incidence and spatial distribution of biotic and abiotic stresses enables farmers to apply corrective measures when and where needed instead of blanket applications covering the whole field. Reducing the amounts of applied agrochemicals improves the financial and environmental performance of the agricultural industry and increases its competitiveness. The module will support the decision-making processes of farmers, resulting in benefits to pest management through enabling access to near real-time information, giving farmers the opportunity for data-based decision making and proactive actions that limit damage and improve yields. In addition, it will support the development of pest management strategies that meet EU sustainability goals, allowing farmers to comply with EU regulations, and especially the new Sustainable Use of Pesticides Regulation, currently under discussion.

10.3 TECHNICAL CHALLENGES

Integration of weather, soil and satellite data

- Similar models are not available currently under an open-source license.
- Availability of an adequate volume of data for pest and disease attacks from the pilot site, given the unpredictability of attacks.
- Disentangling biotic from abiotic stress represents a significant scientific challenge because of similar reflectance profiles in multi-band images.
- Upscaling results from one pilot site to several sites in the same and especially other regions.
- Farmers have limited experience using and interacting with the proposed technology.

One of the most crucial challenges is the proper integration of weather, soil and satellite data to identify the upcoming threats on time and therefore inform farmers to take actions. To act on time where the information from real-time sensors is needed, the proper use of IoT technologies is also crucial. Moreover, the connection of climatic and soil data to commonly occurring pests, such as the cereal leaf miner (*Syringopais temperatella* Led.) poses challenges. Due to the climatic conditions of Cyprus, abiotic stresses may also appear and must be distinguished from pest and disease symptoms. Nutritional deficiencies also result in changes to the reflectance signature of crops. To overcome the challenges domain expertise in agriculture, Artificial Intelligence and IoT among others can be applied. Labelled data for AI model training are limited and difficult to collect. Thus, more suitable techniques must be investigated.

To the best of our knowledge, there are no such automated approaches available to support the decision-making activities of farmers. Through this module, we aim to fill the gap. Moreover, the application of the module to a new demonstration site will also be challenging as the conditions differ from area to area. Hurdles will not be only faced during development, but also in the adaptation of such tools for use by farmers. The use of long-standing and empirical approaches with field observations is the main approach by farmers. Moreover, most of the farmers are not trained to use digital solutions. Thus, training sessions and courses may be beneficial.

10.4 MODULE SIZE / METRICS

- Number of sensors: 3
- Number of meteorological stations (integrated sensors for different measurements): 1

10.5 EXPECTED IMPACT AND KPI'S

The expected impact of the module is the provision of data-driven insights to farmers to apply precision farming practices in their fields and act proactively to alleviate abiotic stress and control pests and diseases. Use of the module for pest management is expected to reduce the use of pesticides, while the mapping of the spatial distribution of abiotic stress will lead to a reduction in the need for fertilizer use. In addition, the use of the module is expected to result in higher yields at a lower cost, allowing farmers to reap economic benefits and reduced environmental impacts.

KPIs

- Farmers/farming companies participating in module design: >=5
- Stakeholders participating in model design: >=5
- Automated processing pipeline: 1
- Journal publication: 1
- Conference participation: 1

10.6 DEMONSTRATION PLAN

The demonstration plan consists of the following steps:

- Installation of soil moisture sensors and IoT meteorological station.
- Establishment of pipelines for data acquisition from the demonstration site.
- Satellite data collection.
- In-situ observations and visits.
- Design of the technological part of the module.
- Development of alpha version.
- Development of beta version.
- Integration with Data Cube.
- Integration with User Interface.
- Testing and Validation.

10.7 DATA REQUIREMENTS FOR MODULE

Shapefile/GeoJSON of demonstration parcels/plots.

  - Vector data covering the extent of each demonstration plot for better monitoring through satellite data. The vector data can be used as a standalone file in the format of shapefile, or geojson.

- IoT Meteorological data (as described in 6.3 and appendices)
  - Data from the in-situ meteorological station installed at the demonstration site. The data can be available in different formats, like csv, and SQL. An identifier for date and time (datetime) of acquisition of each record will be available. Moreover, each data record contains the name of sensor and the id of sensor. The station can capture data every 15 minutes (or adjusted according to needs) and in a high resolution.
- Soil sensor data (as described in 6.3 and appendices)
  - Data from in-situ soil sensors installed at the demonstration site. The data can be available in different formats like csv and SQL. Similar to meteorological data, an identifier for date and time (datetime) of acquisition of each record will be available. Soil sensors are capturing high resolution data every 15 minutes.
- Biotic and abiotic stress occurrence data
  - Geolocated observation data on incidence and severity of biotic and abiotic stressors via field observations. Key stressors include cereal leafminer and water stress. Additional stressors include fungal diseases and nutrient stress, depending on occurrence and frequency.

10.8 DATASET WORKFLOW

Satellite data from the Nostradamus Data Cubes along with data from soil sensors and meteorological data from the IoT station will be used daily for the early warning system developed through this module (Figure 15). The data from the demonstration site will be stored in a cloud repository. First, a quality check on the data will take place. Then, the data will be pre-processed if needed and will be prepared for fusion with satellite data in order to feed the processing pipelines. The processing pipeline will include threshold-based approaches using weather and soil parameters that indicate crop stresses. Moreover, AI solutions will be utilized to extract useful information from the different sources that will enhance the early detection of upcoming stresses. In case criteria are met, the information of early warning will be stored. Next, suitable actions will then be taken to inform the farmers. Threshold values will be set for weather and soil parameters that indicate that crops face stress, and if specific criteria are met, the information will be saved. Suitable actions will then be taken to inform the farmers. If needed, data will be saved in databases to create time series of the variables. This could be useful to identify patterns. The implementation timeline is shown in Figure 16.

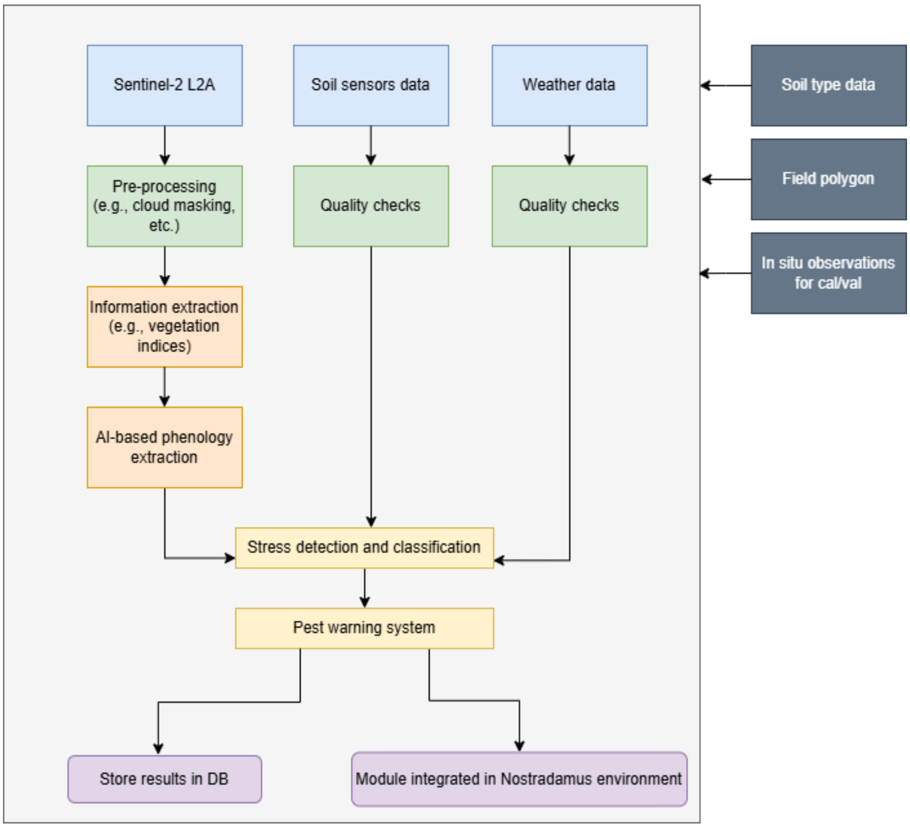


FIGURE 15: MODULE 4 DATASET WORKFLOW

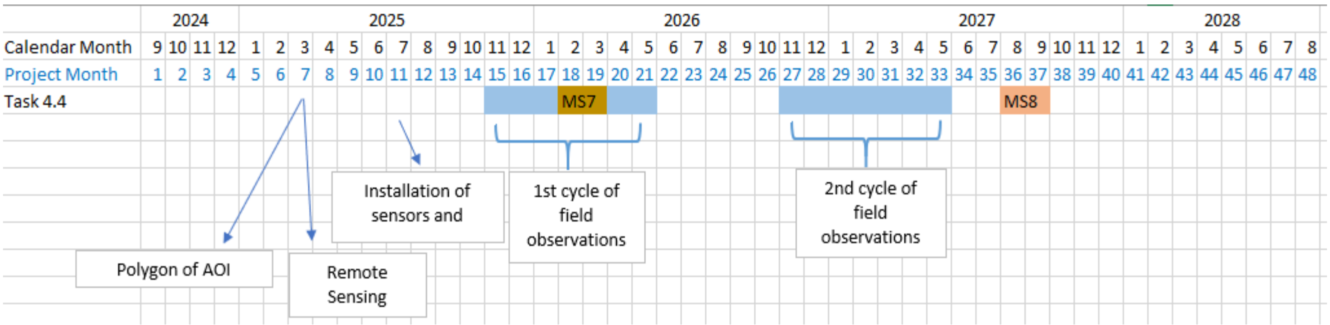


FIGURE 16: TIMELINE FOR MODULE 4 MAIN TASKS

11 PILOT REQUIREMENTS AND KPI'S

The performance of the Nostradamus pilots and demonstration sites will be evaluated through a set of KPIs aligned with project objectives and expected impacts. These KPIs aim to monitor the technical, socio-economic, and environmental performance of the deployed digital applications and modules across diverse biogeographical regions.

TABLE 1: PILOT KPI'S

KPI	Description	Target
KPI-P1	Number of demonstration events in pilot sites	>8
KPI-P2	Participation of stakeholders (policy, farmers, advisors, industry, NGOs, researchers) in demo events	≥10 stakeholders per event
KPI-P3	Number of on-site visits by young scientists (training & knowledge transfer)	≥5 visits with >20 young scientists per visit
KPI-P4	Increase in crop yield due to digital solutions	>15%
KPI-P5	Reduction in pesticide use	>40%
KPI-P6	Reduction in water consumption	>20%
KPI-P7	Reduction in energy consumption in agricultural production	≤15%
KPI-P8	Number of stakeholders engaged in Multi-Actor Platforms (MAPs)	≥5 synergies across the 5 countries
KPI-P9	Technical performance reports per pilot	≥1 per pilot, including assessment of usability, accuracy, and reliability
KPI-P10	Training events delivered (on-site and online) for pilot stakeholders	≥6 events
KPI-P11	Number of surveys collected from pilot participants	>500 surveys total across pilot regions

KPI-P12	Practice abstracts delivered in EIP-AGRI format	≥8
KPI-P13	Number of knowledge exchange workshops and MAP-related activities	≥30 meetings and ≥5 joint activities
KPI-P14	Increase of sustainable and eco-friendly practices adoption among farmers in pilot regions	>5% of farmers adopting sustainable practices
KPI-P15	Demonstration of decision-support impact	≥10% improvement in decision-making speed or accuracy for users compared to baseline

12 EXTERNAL NEEDS-BASED DATA REQUIREMENTS FRAMEWORK

Task 2.2 establishes the foundation for identifying, evaluating and prioritizing the external data requirements that will support Nostradamus in developing interoperable, scalable and policy-aligned digital solutions. As the project operates within a rapidly evolving agricultural, environmental and data governance landscape, external data requirements are not static. They emerge progressively as new actors join the MAPs, as EU-funded projects advance their own work, as policy priorities evolve and as Third-Party developers under WP6 begin designing new applications. For this reason, Task 2.2 places strong emphasis during the first iteration of Deliverable D2.1 on defining a structured, transparent, and repeatable framework for assessing needs-based data requirements. This framework provides a methodological backbone that will guide all external data requirement activities across the project, ensuring continuity, comparability, and alignment from Month 1 to the final update at Month 44.

By establishing this framework early in the project, Task 2.2 ensures that Nostradamus is capable of systematically identifying external data needs as they arise, filtering them according to relevance and feasibility, and integrating them into the broader data architecture developed under WP3 and WP5. The framework supports a proactive rather than reactive approach, allowing the project to anticipate where data needs are likely to emerge, how they may change over time and which actors are best positioned to contribute to or benefit from specific data streams. This is essential given the diversity of agricultural systems, climatic conditions and socio-economic contexts across Europe and the complex interplay between policy, technological innovation and end-user behavior. The Task 2.2 framework also ensures that all external data needs identified through MAPs, Third-Party developers, related EU-funded projects, partnerships, thematic clusters, AKIS platforms and policy actors are analyzed systematically and in a harmonized manner. Each input is validated against the project's objectives, its four analytical modules, its user-oriented applications, and its policy relevance, while also being assessed using clear quality criteria that determine suitability for integration. In addition to methodological rigor, the framework places equal importance on process: it defines how different stakeholder groups will be engaged over time, how inputs will be collected and verified, and how they will be incorporated into subsequent versions of the data requirements catalogue.

Moreover, the framework establishes mechanisms that ensure ongoing dialogue with external actors, enabling Nostradamus to keep pace with emerging priorities, interoperability needs and sectoral transformations. These mechanisms also ensure that external inputs do not remain isolated but are directly channeled into the data workflows and Data Cube specifications developed under WP3 and WP5. As a result, Task 2.2 creates a bridge between stakeholders, external initiatives and internal module development, ensuring that the system remains adaptable, responsive, and strategically aligned with the European agricultural data ecosystem.

12.1 FRAMEWORK FOR PRIORITISING NEEDS-BASED DATA REQUIREMENTS

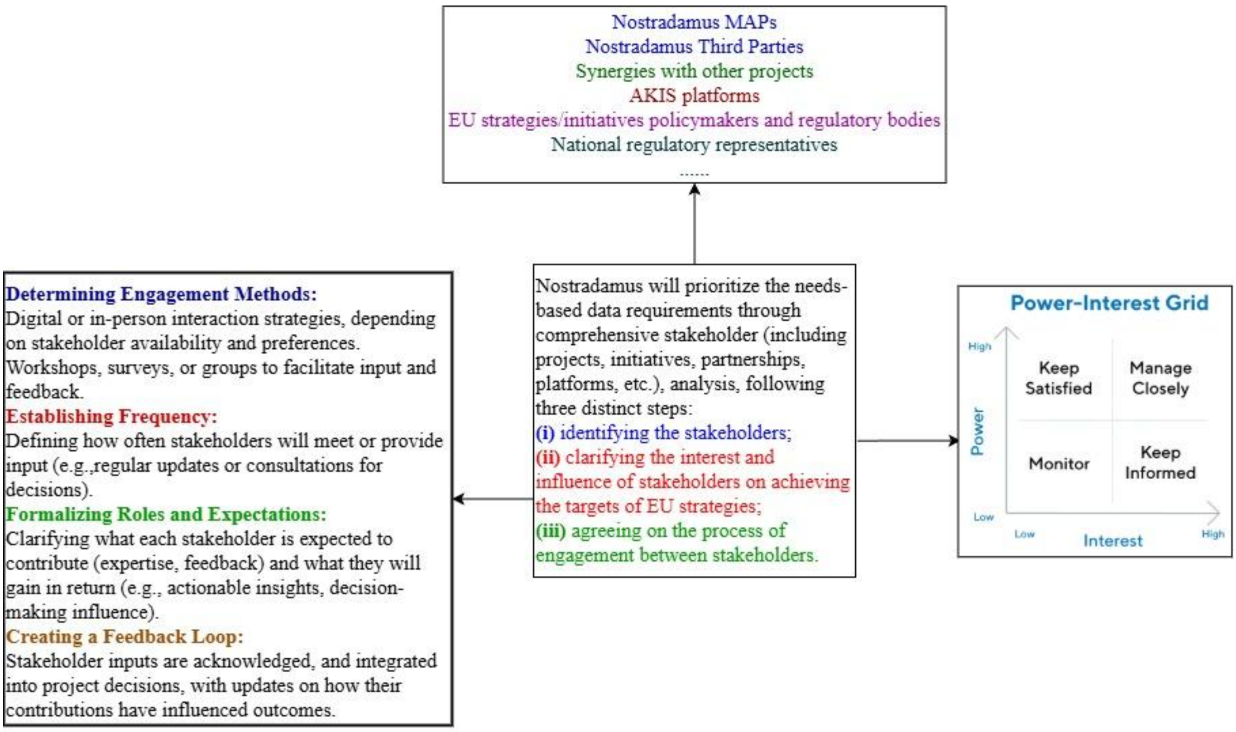


FIGURE 17: FRAMEWORK FOR PRIORITIING AND ENGAGING STAKEHOLDERS IN THE IDENTIFICATION OF EXTERNAL DATA REQUIREMENTS

Nostradamus will prioritise the needs-based data requirements through a comprehensive analysis of all relevant stakeholder categories, including MAPs, Third-Party developers under WP6, EU-funded projects and initiatives, partnerships, platforms such as AKIS, and regulatory or policymaking entities. This prioritisation process ensures that external data needs are addressed in a systematic, transparent and operationally meaningful way. It also guarantees that the data framework evolves in line with technological, regulatory, and user-driven developments across the project's lifetime. The process is therefore structured around three core analytical and procedural steps, each designed to ensure that stakeholder-driven data requirements are captured, validated, ranked and integrated efficiently into the project's data architecture.

- **Step 1:** Identification of Stakeholders

- **Step 2:** Clarification of Stakeholder Interest and Influence
- **Step 3:** Agreement on the Engagement Process

12.1.1 STEP 1: IDENTIFICATION OF STAKEHOLDERS

This step involves mapping all entities whose data needs may influence or be influenced by Nostradamus. The objective is to ensure full visibility of actors who contribute to, benefit from, regulate or depend on the project's data infrastructure. By cataloguing the full landscape of relevant stakeholders, Task 2.2 ensures that the project's data framework reflects the operational, scientific, regulatory and market realities of the agricultural ecosystem. The identification process covers an extensive spectrum of groups, including:

- **MAP participants across the five demonstration sites;** Cyprus, Germany, Serbia, Slovenia and Switzerland, representing the primary end users of Nostradamus tools. This group includes farmers, advisors, agronomists, cooperatives, local authorities and farm-level decision makers whose operational needs and constraints directly shape the requirements for timely, accessible and actionable data.
- **Third-Party successful applicants for the Open Calls under WP6**, who will design and implement digital applications that operate on top of the Nostradamus data environment. Their data requirements represent concrete, real-world innovation needs and ensure that the data framework supports practical deployment, cross-site testing, user-centred design and scalable services for decision making and policy shaping.
- **EU-funded projects including Cluster 6 initiatives** such as ECO-READY, LandShift, OpenAgri, Agrarian and others, which articulate complementary insights, shared objectives and cross-project demands for harmonised, interoperable data infrastructures. Considering their data requirements ensures alignment with broader European research priorities and supports synergies, while enabling Nostradamus to contribute meaningfully to a coordinated EU-wide digital agriculture landscape.
- **AKIS platforms, AgData initiatives and interoperability-focused actors**, which represent structured knowledge systems and data governance efforts across Europe. Their data needs often focus on integration, standardisation, advisory use cases and capacity-building, making them essential for long-term adoption pathways.
- **EU and national policymakers and regulatory bodies** responsible for shaping data governance, CAP monitoring, sustainability reporting and compliance frameworks. Their influence determines the strategic direction of agricultural digitalisation and ensures that the data requirements framework supports regulatory relevance and policy-aligned outcomes.
- **Private-sector actors and service providers** across the agricultural value chain, including input suppliers, remote sensing companies, agro-tech SMEs, data service providers and downstream market actors. Their operational data needs reflect emerging technological trends, innovation potential and commercial scalability.
- **Research organisations, standardisation bodies and data governance experts**, which provide scientific validity, methodological robustness, technical specifications and alignment with open standards, FAIR principles and European data spaces.

By encompassing this broad and diverse range of stakeholder categories, the stakeholder identification phase ensures inclusiveness, prevents the omission of influential actors and establishes an evidence-based foundation for the subsequent prioritisation and engagement steps. This comprehensive mapping guarantees that the data requirements framework reflects both grassroots needs and systemic priorities within the agricultural innovation ecosystem.

12.1.2 STEP 2: CLARIFICATION OF STAKEHOLDER INTEREST AND INFLUENCE

Once stakeholders are identified, each group is assessed according to two strategic dimensions that reflect their role and importance in achieving the objectives of Nostradamus:

- **Interest:** The extent to which stakeholders rely on, benefit from or are directly affected by the data, tools and services developed in Nostradamus. This includes their motivation to engage, their expected use of project outputs, their alignment with the Nostradamus' mission and vision and their dependency on improved data availability, interoperability or decision-support solutions.
- **Power:** The degree to which each stakeholder can support, enable or accelerate the achievement of Nostradamus objectives. This includes their capacity to assist data governance frameworks, promote adoption and scaling of outputs, facilitate regulatory alignment, contribute to technical validation, or provide strategic direction for cross-project collaboration and EU-level priorities.

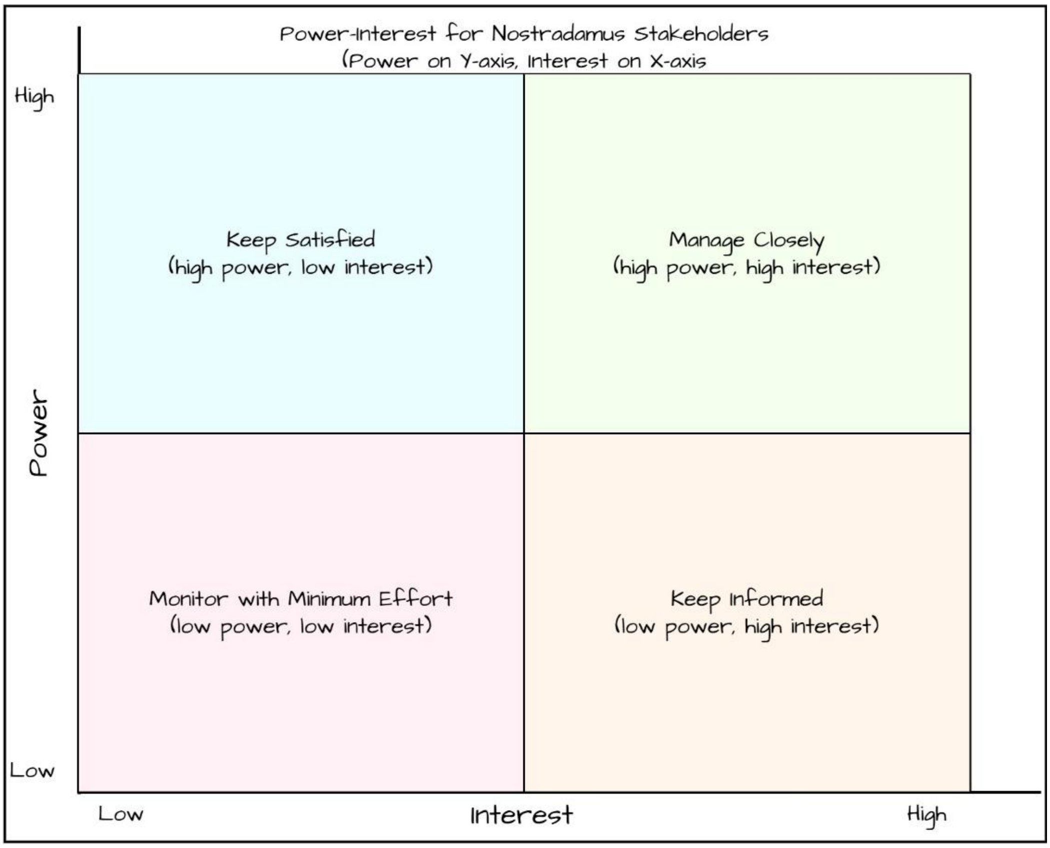


Figure 18: Power-interest matrix to classify Nostradamus stakeholders according to their influence on project objectives and their level of interest in data requirements.

These two dimensions are synthesised using a **power-interest matrix**, which:

- Places stakeholders on an X-axis representing *interest in or benefit from Nostradamus* and a Y-axis representing *power to support or shape project outcomes*
- Identifies four engagement zones:
  - **High power / high interest:** continuous cooperation, co-design and strategic involvement.
  - **High power / low interest:** targeted engagement where their contribution is essential for regulatory, governance or scaling support.
  - **Low power / high interest:** routine consultations, feedback collection and user-driven validation.
  - **Low power / low interest:** monitoring and periodic updates without intensive interaction.

This structured analysis ensures that critical actors—such as policymakers, MAP representatives, Third-Party developers, data governance bodies and EU-level initiatives—remain central to the needs-assessment workflow. It also guarantees that stakeholders with less influence but high interest (e.g., farmers, advisors, SMEs) receive appropriate engagement to ensure relevance, usability and co-creation of the data framework.

12.1.3 STEP 3: AGREEMENT ON THE ENGAGEMENT PROCESS

To operationalise the prioritisation framework, Task 2.2 defines concrete procedures that govern how stakeholders will be engaged, how often, and through which mechanisms. This ensures consistency across WP2 activities and alignment with MAPs (Task 1.1, Deliverables D1.1 - *D1.3 MAA strategic plan*), WP3-6 operations and cross-project synergies (Task 1.3, *Deliverables D1.5 – D1.7 Dissemination, exploitation & communication Plan*).

Determining Engagement Methods

- Selection of digital or in-person interaction formats depending on stakeholder preferences, availability and geographic constraints.
- Use of structured tools such as surveys, interviews, focus groups, and targeted workshops.
- Establishment of dedicated communication channels for verification and feedback, including MAP meetings, technical exchanges and project-project bilateral or clustering sessions.

Establishing Engagement Frequency

- Regular updates aligned with MAP evolution cycles and agricultural activities.
- Periodic consultations in synchronization with Third-Party development phases.
- Scheduled workshops with EU initiatives, partnerships and platforms to align external requirements and ensure consistency with broader European efforts.

- Cross-project review cycles to ensure long-term policy relevance and interoperability with related initiatives.

Formalizing Roles and Expectations

- Clarifying the expected contributions of each stakeholder group (requirements, constraints, validation needs, policy priorities, technical specifications).
- Defining the added value stakeholders receive in return (access to insights, participation in co-design, testing opportunities, alignment with policy processes).
- Ensuring transparency about how external requirements will inform module development, data cube design and analytical integration.

Creating a Continuous Feedback Loop

- Acknowledging and documenting stakeholder inputs in a consistent and traceable manner.
- Validating and integrating relevant requirements into the evolving data framework.
- Communicating how each contribution influenced prioritisation decisions, module design or data collection plans.
- Ensuring iterative refinement of requirements through exchanges with synergetic projects and EU-level initiatives.

This continuous feedback loop ensures that the needs-based prioritisation remains dynamic, adaptive and fully aligned with the evolving realities of the agricultural, environmental and policy landscape throughout the project.

12.1.4 SUITABILITY CRITERIA FOR EXTERNAL DATA REQUIREMENTS

Nostradamus establishes a structured, transparent and repeatable framework for assessing the suitability of external data requirements before they can be integrated into the Nostradamus data environment. Given that the project interacts with diverse sources, it is essential to ensure that all external data needs are evaluated systematically. This ensures that the data incorporated into WP3 processing workflows and WP5 Data Cube specifications is technically sound, relevant, interoperable and aligned with the long-term objectives of the project. To achieve this, Nostradamus applies a set of quality and relevance criteria that every external data requirement must meet. These criteria address both the **substance** of the requirement (what is needed and why) and the **practical feasibility** of supporting it within the scope of Nostradamus.

Each external requirement will be evaluated according to the following eight criteria:

- **Completeness:** Whether the data requirement is clearly defined, unambiguous and accompanied by sufficient contextual information (e.g., purpose, end user, temporal range, spatial scale).
- **Consistency:** Whether the requirement aligns with the overall objectives, technical architecture and workflow logic of Nostradamus.
- **Accuracy:** Whether the expected data type supports reliable module operation, modelling processes or decision-making mechanisms.
- **Timeliness:** Whether the data is available or can be accessed with an appropriate temporal resolution and update frequency for the intended use case.
- **Relevance:** Whether the data contributes meaningfully to module development, digital applications, validation activities or policy evaluation tasks.
- **Granularity:** Whether the spatial, temporal and thematic detail of the requested data matches the needs of the respective use case and is feasible to integrate.
- **Normalization:** Whether the data can be harmonised with WP3 pipelines in terms of formats, metadata, standards and pre-processing requirements.
- **Integrity:** Whether the data is generated, processed or shared in a manner compatible with the project's ethical, legal, governance and quality principles (e.g., provenance, licensing, FAIR compliance).

12.1.4.1 COMPLETENESS

Whether the external data requirement is clearly described and includes all critical contextual information needed by the project. For Nostradamus, a requirement is considered complete when it specifies:

- the intended use case (e.g., drought detection, yield estimation, nutrient assessment, pest risk prediction, agricultural sector competitiveness, reduction of agrochemical inputs, decreased EU dependence on related imports, and enhancement of Food Security, Safety, and Sovereignty).
- the end-user group who would benefit (e.g., farmers via MAPs, advisors, policymakers, Third-Party developers).
- the spatial and temporal coverage.
- the expected format or structure of the data.
- the operational scenario in which the requirement would be applied.

Incomplete or ambiguous requirements cannot be evaluated or operationalised within WP3 or WP5.

12.1.4.2 CONSISTENCY

Whether the requirement aligns with:

- the **core objectives** of Nostradamus, including improving resilience, sustainability and digitalisation of European agriculture.
- the **four WP4 analytical modules** (drought and vegetation stress, crop suitability, soil fertility and nutrition, sustainable pest management).
- the **policy goals** addressed by the project (e.g., Farm to Fork, CAP, European Data Strategy, see Section 12.1.5 Policy Alignment Benchmarks Informing the Suitability Criteria).
- the **technical architecture** of WP3 (data ingestion, harmonisation, processing pipelines).
- the **Data Cube specifications** of WP5 (storage, metadata, interoperability rules).

A requirement that does not reinforce or enable these objectives is considered inconsistent.

12.1.4.3 ACCURACY

Whether the expected data can support reliable decision-making processes. In this project, accuracy refers to:

- the precision of measurements (e.g., soil moisture, weather parameters, pest incidence).
- the fidelity of remote sensing inputs (e.g., atmospheric correction, cloud screening quality, resolution suitability).
- the reliability of field observations or socio-economic inputs.
- the level of uncertainty permissible for module algorithms or policy indicators.

Data with insufficient accuracy could compromise potential exploitation for decision-making and policy-shaping. To ensure operational applicability, Nostradamus will assess accuracy by considering whether the external data requirement specifies:

- A minimum spatial resolution appropriate for the target process (e.g., vegetation stress monitoring; plot-level precision for IoT data).
- A minimum temporal resolution or update frequency, especially for dynamic variables such as drought indicators, pest monitoring, or yield estimation.
- Permissible error ranges or confidence levels (e.g., classification accuracy, sensor calibration tolerances, if available and applicable).

External data requirements lacking sufficient accuracy specifications—or failing to meet thresholds—will be flagged as partially supported, conditionally supported or out of scope, depending on feasibility within the Nostradamus.

12.1.4.4 TIMELINESS

Whether the data can be provided with the temporal resolution and update frequency required by:

- real-time or near real-time processes (e.g., drought early warning, pest monitoring).
- periodic assessments (e.g., seasonal yield estimation, soil nutrient trends).
- annual or multi-year policy reporting tasks tied to EU strategies or CAP monitoring.

Timeliness is essential for operational use cases, particularly for Third-Party applications under WP6, should they choose to exploit them.

12.1.4.5 RELEVANCE

Whether the data requirement offers direct value to at least one of the following:

- WP4 model calibration, validation or operation.
- WP6 application development (farm decision-making or policymaking tools).
- WP3 data ingestion requirements.
- cross-site harmonisation and interoperability.
- MAP-defined user needs.
- EU policy monitoring activities relevant to Nostradamus (e.g., CAP indicators, drought metrics, crop sustainability KPIs).

Data that does not contribute to at least one of these areas is considered low relevance.

12.1.4.6 GRANULARITY

Whether the spatial, temporal and thematic detail requested matches what is necessary and feasible for:

- the agricultural systems represented in the five pilot sites (Germany, Serbia, Slovenia, Switzerland, Cyprus).
- the scale of modelling (plot, farm, regional).
- the resolution of remote sensing and IoT datasets available to WP3.

- the computational capacity of WP5 Data Cubes and available storage within the Nostradamus infrastructure.

Requirements exceeding feasible granularity cannot be supported.

12.1.4.7 NORMALIZATION

Whether the requested data type can be harmonised within the WP3 ingestion and pre-processing pipelines. In Nostradamus, normalization specifically concerns:

- standard formats (GeoTIFF, NetCDF, CSV, JSON, INSPIRE-compliant shapefiles).
- harmonised metadata (e.g., sensor type, acquisition time, coordinate system).
- conversions between coordinate systems or units.
- integration into the Data Cube's multi-dimensional structure.

If normalization is not possible or requires disproportionate effort, the requirement may be deprioritised.

12.1.4.8 INTEGRITY

Integrity assesses whether the data is produced, processed and shared under conditions that are legally compliant, ethically sound and aligned with the governance principles that underpin the Nostradamus data architecture. In this project, integrity refers to whether the data respects:

- **FAIR principles** (Findable, Accessible, Interoperable, Reusable), ensuring long-term usability and alignment with European data-space initiatives.
- **Data provenance and traceability**, including clear documentation of data origin, processing steps, quality controls and versioning.
- **Licensing terms**, identifying whether the data is open, restricted or proprietary, and whether usage rights allow integration into Nostradamus workflows and storage in the project's infrastructure.
- **Ethical and legal compliance**, including the EU General Data Protection Regulation (GDPR) for any personal or sensitive information (e.g., farmer identifiers in MAP surveys, geolocation-linked personal data), as well as relevant national regulations when applicable.
- **Data protection and security requirements**, including secure transfer protocols, controlled-access repositories, authentication mechanisms and need-to-know access controls.

In practical terms, external data fails the integrity criterion when:

- licensing terms prohibit reuse, reprocessing or redistribution.
- provenance is unclear or incomplete, preventing reliability verification.
- data contains personal identifiers without appropriate pseudonymisation, consent or GDPR-aligned safeguards.
- security standards do not comply with WP5 governance or cannot be harmonised with the project's infrastructure.

Any external data requirement that does not satisfy integrity conditions will be categorised as non-supportable until governance, licensing or ethical conditions are clarified and validated.

12.1.5 POLICY ALIGNMENT BENCHMARKS INFORMING THE SUITABILITY CRITERIA

The suitability criteria described above must be interpreted in relation to the strategic, regulatory and policy context in which Nostradamus operates. To ensure consistent evaluation of external data requirements, Nostradamus uses a comprehensive set of EU-level strategies, directives and policy frameworks as reference benchmarks, building directly on the analysis already conducted in Deliverable *D1.4 Obstacles and strategies for EU Food Security and Independence, and Use of Digital Solutions*, and further enriching that baseline with additional strategies. These frameworks define monitoring obligations, reporting requirements, data specifications and thematic priorities relevant to digital agriculture, climate resilience, soil health, biodiversity protection and sustainable food systems. The table below summarises the most relevant strategies and outlines the actions and objectives that influence the Nostradamus data needs and the assessment of external requirements. However, the table may also be updated to reflect new developments in EU strategies or incorporate feedback received from stakeholders.

TABLE 2: KEY EU AND INTERNATIONAL STRATEGIES RELEVANT TO NOSTRADAMUS AND THEIR ASSOCIATED OBJECTIVES AND ACTIONS INFORMING EXTERNAL DATA REQUIREMENTS.

Strategy / Initiative	Objectives	Actions
Farm-to-fork strategy	To facilitate a rapid transition to a sustainable food system that: 1. Has a neutral or positive impact on the environment; 2. Aids in the mitigation of climate change and adapts to its repercussions; 3. Contributes to reversing the trend of biodiversity loss; 4. Ensure food security, guaranteeing that all individuals have access to adequate, safe, nutritious, and sustainable food; 5. Maintains food affordability while generating equitable economic returns, enhancing the	a) Legislative framework for sustainable food systems. b) Contingency plan for ensuring food supply and food security. c) Sustainable food production. d) Promote sustainable food consumption, facilitating the shift towards healthy, sustainable diets.

	competitiveness of the EU supply sector, and promoting fair trade initiatives.	
Zero Pollution Action Plan	The objective is to reduce air, water, and soil pollution to levels that are no longer harmful to public health and natural ecosystems. The targets consist of: 1. Improving water quality and decreasing the release of microplastics into the environment; 2. Improving soil quality through reduction in nutrient losses and the use of chemical pesticides; 3. Reduce the number of EU ecosystems at risk from pollution's impact on biodiversity.	a) Identify and remedy contaminated sites; b) Identify obligations on nutrient management to limit the environmental footprint of farming activities; c) Contribution to the European Green Deal Dataspace to improve data availability; d) Support the development of a high-precision digital model of the Earth, using Copernicus data as a key building block to monitor the state of air, waters, seas, and soil; e) Improved training and educational support on environmental risks; f) Support the implementation of nature-based solutions to mitigate pollution; g) Integrate digital tools (e.g., IoT, AI) for real-time pollution monitoring.
EU chemicals strategy for sustainability	The European Union's chemicals strategy seeks to: 1. Improve the safeguarding of citizens and the environment. 2. Promote innovation in creating safe and sustainable chemical products.	a) Boosting the investment and innovative capacity for the production and use of chemicals that are safe and sustainable by design and throughout their life cycle; b) Promoting the EU's resilience in the supply and sustainability of critical chemicals; c) Establishing a more straightforward "one substance, one assessment" process for chemical risk and hazard assessment.
Circular economy action plan	The objectives in the action plan are designed to: 1. Establish sustainable products as a standard within the EU. 2. Empower consumers and public sector purchasers. 3. Target the most resource-intensive sectors with substantial potential for circularity; 4. Reduce waste; 5. Facilitate circularity.	a) Empowering consumers in the green transition; b) Restriction of intentionally added microplastics and measures on the unintentional release of microplastics; c) Policy framework for bio-based plastics and biodegradable or compostable plastics; d) Harmonized information systems for the presence of substances of concern; e) Making the circular economy work for people, regions and cities; f) Supporting the circular economy transition; g) Improving measurement and policy tools to capture synergies between the circular economy and climate change mitigation and adaptation at EU and national level; h) Reflecting circular economy objectives in revising the guidelines on state aid in the field of environment and energy;

		i) Mainstreaming circular economy objectives in the context of the rules on non-financial reporting and initiatives on sustainable corporate governance and environmental accounting.
EU Biodiversity Strategy for 2030	The biodiversity strategy aims to set Europe's biodiversity on a recovery trajectory and focuses on fortifying societal resilience against potential threats, such as: 1. Climate change repercussions, 2. Forest fires, 3. Food scarcity, 4. Disease outbreaks. 5. Safeguarding wildlife.	a) Establishing a more extensive EU-wide network of protected land areas; b) Set a restoration objective focused on the long-term revitalization of nature within the EU's land, incorporating legally binding restoration targets for habitats and species; c) Introducing measures to tackle the global biodiversity challenge.
Kunming-Montreal Global Biodiversity Framework	The objective is to initiate immediate measures to stop and reverse the decline of biodiversity. This initiative seeks to restore natural ecosystems to benefit humanity and the environment.	a) Plan and Manage all Areas to Reduce Biodiversity Loss; b) Restore degraded ecosystems; c) Conserve land and water. d) Halt species extinction and manage human-wildlife conflicts. e) Ensure sustainable, safe and legal harvesting. f) Reduce the introduction of invasive alien species and minimize their impact. g) Minimize the impacts of Climate Change on Biodiversity. h) Enhance biodiversity and Sustainability in Agriculture and Forestry i) Restore, maintain and enhance Nature's contributions to people j) Increase the sharing of benefits from traditional knowledge. k) Integrate biodiversity in decision-making at every level. l) Businesses assess, disclose and reduce biodiversity-related risks and Negative Impacts. m) Strengthen biosafety and distribute the benefits of Biotechnology. n) Reduce harmful incentives and scale up positive incentives for Biodiversity. o) Strengthening Capacity-Building, Technology Transfer, and Scientific and Technical Cooperation for Biodiversity. p) Ensure that knowledge is available and accessible to guide Biodiversity Action. q) Ensure participation in Decision-Making and access to justice and information related to Biodiversity for all. r) Reduce Pollution to levels that are not harmful to Biodiversity.

EU Forest Strategy for 2030	The strategy articulates a vision to enhance the quantity and quality of forests within the EU while bolstering their protection, restoration, and resilience. Its objective is to adapt European forests to the evolving conditions, including extreme weather events and significant uncertainties associated with climate change.	a) Promoting the sustainable forest bioeconomy for long-lived wood products. b) Developing skills and empowering people for a sustainable forest-based bioeconomy. c) Protecting EU's last remaining primary and old-growth forests. d) Ensuring forest restoration and reinforced sustainable forest management for climate adaptation and forest resilience. e) Re-afforestation of biodiverse forests. f) Provide financial incentives for forest owners and managers to improve the quantity and quality of EU forests. g) Strategic forest monitoring, reporting and data collection. h) Developing a solid research and innovation agenda to improve our knowledge of forests. i) Implementing an inclusive and coherent EU forest governance framework.
European Climate Law	1. Formulate a strategic long-term plan to attain climate neutrality by 2050, integrating all policies in a socially just and economically viable manner. 2. Establish a more rigorous target for the EU to position Europe on a responsible course towards achieving climate neutrality by 2050. 3. Implement a systematic approach for tracking progress and taking additional measures when required. 4. Ensure predictability for investors and economic participants while ensuring the irreversible transition to climate neutrality.	a) A legal objective for the Union to reach climate neutrality by 2050. b) There will be reduction in GHG, with clarity on the contribution of emission reductions and removals. c) Recognition of the need to enhance the EU's carbon sink through a more ambitious LULUCF regulation. d) A process for setting a climate target, considering an indicative GHG budget. e) A commitment to negative emissions after 2050. f) Stronger provisions on adaptation to climate change. g) Strong coherence across Union policies with the climate neutrality objective. h) A commitment to engage with sectors to prepare sector-specific roadmaps charting the path to climate neutrality in different areas of the economy.
EU strategy on adaptation to climate change	The Strategy aims to cultivate a climate-resilient society by deepening the understanding of climate impacts and adaptation solutions. It focuses on enhancing adaptation planning, conducting comprehensive climate risk assessments, accelerating adaptation actions, and bolstering	a) Pushing the boundaries of adaptation knowledge. b) Gathering more accurate climate loss data. c) Enhancing and expanding adaptation knowledge. d) Integrate adaptation into policy, Nature-based solutions and Local adaptation actions.

	global climate resilience. The Strategy outlines three main objectives and suggests a variety of actions to achieve these aims: 1. **Smarter adaptation**: This entails improving knowledge and managing uncertainty. 2. **More systemic adaptation:** Supporting policy development at all levels and all relevant policy fields. 3. **Faster adaptation:** Speed up adaptation implementation across the board.	
Common Agricultural Policy (CAP)	The primary goals of this policy are to: 1. Assist farmers and enhance agricultural productivity, guaranteeing a stable and affordable food supply. 2. Ensure that farmers in the EU can achieve a sustainable livelihood. 3. Address the challenges posed by climate change and promote the sustainable use of natural resources. 4. Maintain the integrity of rural areas and landscapes within the EU.	The CAP uses various actions to bolster the agricultural sector and promote rural development. Among these are: a) Direct payments that provide essential income support, ensuring stability for farmers while compensating them for their environmentally responsible farming practices and delivering public services, such as countryside management, that are generally unremunerated by market systems. b) Implements measures to address adverse situations, including sudden demand. c) Fostering rural development through tailored national and regional programs that respond to rural areas' specific challenges and needs. d) Support for rural development. e) Management and monitoring of the CAP.
Regulation on LULUCF	The Regulation enhances governance, promotes transparency, and strengthens the interconnection between climate mitigation and environmental protection measures. Among the objectives: 1. Establishing ambitious, fair, and binding national targets for net carbon removal in the land use sector. 2. Encouraging synergies between climate mitigation and environmental protection to effectively confront the challenges posed by climate change and biodiversity loss and	a) Reducing emissions resulting from land-use change or cropland. Those include afforestation, reforestation and increasing resilience to risks such as wildfires or pest outbreaks. b) Applying practices such as agroecology, agroforestry, and climate-friendly farming. c) Balanced trade-offs between the objectives of food systems, climate change adaptation and mitigation, and biodiversity conservation.

	achieve climate-neutral land-use sector. 3. Simplifying reporting rules to conform to standards is necessary to enhance transparency. 4. Advanced monitoring technologies, such as satellite data, should be employed to improve monitoring, reporting, and verification of emissions and carbon removals.	
Soil Monitoring Law	The Directive seeks to implement a well-structured and integrated soil monitoring framework across the EU to improve soil health continuously. 1. The overarching objective is to achieve healthy soils by 2050 and to maintain them in a state that allows for the provision of multiple ecosystem services, thereby meeting the environmental, societal, and economic needs of the Union. 2. Mitigate the effects of soil health degradation, climate change and biodiversity loss, enhance resilience, ensure food security, and lower soil contamination to levels no longer considered detrimental to human health and the environment.	a) Soil districts in MS (member states). Reporting by MS. b) Competent authorities by all MS. c) Soil health and land take monitoring framework by all MS. d) Soil descriptors, criteria for healthy soil condition, land take and soil sealing indicators for monitoring and assessment. e) Measurements and methodologies. f) Assessment of the soil health. g) Sustainable soil management. h) Risk-based approach on potentially contaminated sites i) Identification of potentially contaminated sites. j) Investigation of potentially contaminated sites. k) Risk assessment and management of contaminated sites.
Bioeconomy strategy	The strategy for bioeconomy is designed to hasten the transition towards a sustainable bioeconomy in Europe. Its objectives include: 1. Guaranteeing food and nutrition security, fostering the sustainable management of natural resources. 2. Minimizing dependence on non-renewable and unsustainable resources. 3. Mitigating and adapting to climate change, and bolstering European competitiveness.	a) Development and deployment of sustainable bio-based solutions. b) Analyze enablers and bottlenecks for the deployment of bio-based innovations. c) Develop substitutes for fossil-based materials that are bio-based d) Launch pilot actions for developing bio economies in rural areas. e) Support regions and Member States in developing bioeconomy strategies. f) Promote education, training and skills across the bioeconomy. g) Enhance knowledge of biodiversity and ecosystems. h) Monitor progress towards a sustainable bioeconomy. i) Promote good practices to operate the bioeconomy within safe ecological limits.

		j) Enhance the benefits of biodiversity in primary production.
EU Strategy for Data	The European data strategy seeks to establish: 1. A unified data market that enhances Europe's global competitiveness and upholds data sovereignty. 2. To defend and promote European values and rights in the digital world.	The actions are based on four pillars: a) A cross-sectoral governance framework for data access and use. b) Strengthening of Europe's capabilities and infrastructures for hosting, processing, and using data, as well as interoperability. c) Empower individuals by investing in skills and in SMEs. d) Creation of European data spaces in strategic sectors and domains of public interest. e) making available product data and related service data to the connected product or related service user. f) making available data-by-data holders to data recipients. g) facilitating switching between data processing services. h) development of interoperability standards for data to be accessed, transferred and used.
Sustainable Development Goals	1. End poverty in all its forms everywhere 2. End hunger, achieve food security and improved nutrition and promote sustainable agriculture 3. Ensure healthy lives and promote well-being for all 4. Promote sustained, inclusive and sustainable economic growth, full and productive employment 5. Build resilient infrastructure, promote inclusive and sustainable innovation 6. Reduce inequality within and among countries 7. Ensure sustainable production patterns 8. Take urgent action to combat climate change and its impacts 9. Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably combat desertification, and halt and reverse land degradation and halt biodiversity loss 10.	a) Among the targets/ actions to be achieved is included a shortlist of those referring to the SDG6, SDG13, SDG14, SDG15 which have relevance to the Nostradamus project. b) Improve water quality by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials. c) Protect and restore water-related ecosystems d) Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries. e) Integrate climate change measures into national policies, strategies and planning. f) Prevent and significantly reduce marine pollution of all kinds, from land-based activities g) Promote the implementation of sustainable management of all types h) Combat desertification, restore degraded land and soil, including land affected by desertification, drought and floods, and strive to achieve a land degradation-neutral world. i) Ensure the conservation of biodiversity, to enhance their capacity to provide

		essential benefits for sustainable development. j) Take urgent and significant action to reduce the degradation of halt biodiversity loss and protect and prevent the extinction of threatened species. k) Integrate ecosystem and biodiversity values into national and local planning, development processes, poverty reduction strategies and accounts.

12.1.6 RANKING MATRIX FOR PRIORITIZING EXTERNAL DATA REQUIREMENTS

To support transparent prioritization and avoid subjective selection, Task 2.2 applies a quantitative ranking matrix that translates the suitability criteria defined in Section 12.1.4 into a comparable scoring and classification scheme. The purpose of this matrix is to enable consistent decision-making across heterogeneous requirement sources (MAPs, EU initiatives, AKIS platforms, policy actors and WP6 Third Parties), while ensuring that requirements remain aligned with Nostradamus objectives and feasible within the WP3 and WP5 technical and governance boundaries. The matrix does not imply automatic integration of external requirements. Instead, it provides an evidence-based method to determine which requirements are prioritized for further investigation, conditional support, or exclusion.

Each external requirement is scored against the eight criteria using a four-level scale:

- **0** = Not met: criterion is not satisfied or information is missing to assess it.
- **1** = Partially met: criterion is partially satisfied but with limitations or uncertainty.
- **2** = Largely met: criterion is satisfied with minor constraints that are manageable.
- **3** = Fully met: criterion is clearly satisfied and operationally compatible with the project.

Because not all criteria contribute equally to feasibility and impact, Task 2.2 applies weights that reflect the project's technical constraints, policy relevance and governance obligations.

TABLE 3: WEIGHTING OF SUITABILITY CRITERIA FOR RANKING EXTERNAL DATA REQUIREMENTS

Suitability Criterion	Description (reference)	Weight (W)
Completeness	Degree to which the external data requirement is clearly specified and actionable	1
Consistency	Alignment with Nostradamus objectives, WP4 modules, WP3 workflows and WP5 Data Cube architecture	2
Accuracy	Ability of the data to support reliable modelling, decision making and policy-relevant outputs	2
Timeliness	Availability and update frequency of data relative to operational and policy needs	1
Relevance	Contribution to module functionality, application development, validation or policy monitoring	2

Granularity	Suitability of spatial, temporal and thematic resolution for intended use cases	1
Normalization	Feasibility of harmonisation within WP3 pipelines and Data Cube standards	1
Integrity	Compliance with governance, licensing, provenance, FAIR principles and GDPR	2

To ensure minimum viability, the matrix includes gating conditions. If a requirement fails any of the following, it is not eligible for prioritization until addressed:

- **Completeness score < 2**, meaning the requirement is not sufficiently specified to be actionable.
- **Integrity score < 2**, meaning governance, licensing, provenance, or GDPR alignment is insufficient.
- **Consistency score < 2**, meaning the requirement is not aligned with Nostradamus objectives, modules, or technical workflow logic.

Requirements failing gating conditions are classified as "Not supportable until clarified" and are returned for refinement through stakeholder engagement. For requirements passing gating conditions, the total weighted score is mapped to priority classes:

- **High priority (28–36)**: strong alignment and high feasibility for support within project boundaries.
- **Medium priority (19–27)**: relevant but requires additional conditions, harmonisation effort, or partial support.
- **Low priority (10–18)**: limited added value or disproportionate effort relative to project scope.
- **Out of scope (0–9)**: insufficient alignment or feasibility even if partially specified.

In addition to priority level, each requirement receives a feasibility tag:

- **Directly supportable**: can be supported with existing WP3 and WP5 plans.
- **Conditionally supportable**: support depends on explicit constraints being met (format, licensing, resolution, access).
- **Partially supportable**: only certain elements can be supported without compromising scope or feasibility.
- **Not supportable**: cannot be supported under current project constraints or governance conditions.

The ranking matrix is applied iteratively across the project duration and updated as requirements evolve, new actors join MAPs, and synergies with EU initiatives and AKIS platforms mature. Scores and classifications are versioned to preserve traceability and to document how refinements in requirement definitions, governance clarifications, or technical feasibility changes influence prioritization outcomes. For funded Third-Parties (WP6), the assessment and handling of external data requirements follow a differentiated process. As these actors will enter into contractual agreements with the Nostradamus project through Open Calls, the detailed specification of their data needs may evolve as their application concepts mature and technical design choices are refined. Consequently, their data requirements are not treated as open-ended external requests, but as inputs that are examined through a structured consultation process.

12.2 PRELIMINARY ACTIVITIES

Collecting the external data requirements is a task finalized in Month 44, extending across almost the entire duration of the Nostradamus project and implemented in two phases. The first phase (M3–M16) focused on preliminary activities, including the development of methodological aspects for requirement capture. During the first months of the project, Nostradamus initiated the foundational activities necessary to begin collecting, structuring, and validating external data requirements, ensuring that subsequent phases could build on a coherent and well-defined baseline. These activities form the foundation for all subsequent iterations of the needs-assessment framework and ensure that data requirements are aligned with the project objectives, stakeholder expectations, and relevant EU strategies. Although the full collection process will continue throughout the project, the work completed to date establishes the necessary mechanisms, tools, and engagement pathways for systematic data gathering.

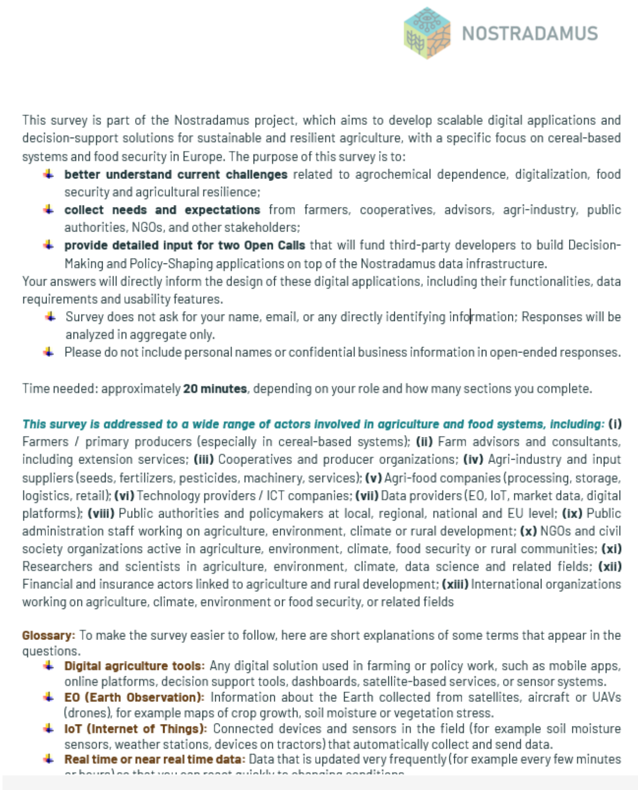
FIGURE 19: IDENTIFICATION OF EXTERNAL DATA REQUIREMENTS

12.2.1 MAPS – PRELIMINARY DATA REQUIREMENTS PROCESS

A first survey is being designed and is currently in the finalization phase (utilizing EU Survey[2]), aiming to collect early-stage data needs from MAP participants during this initial phase of MAP development. The survey targets participants across the demonstration sites described in Deliverable D1.1 (MAA Strategic Plan), as well as potential stakeholders who may have an interest in

[2] https://ec.europa.eu/eusurvey/home/welcome

the project's anticipated outcomes. The survey was also structured to provide an initial roadmap for future Third-Party developers regarding the types of data they may require for the development of the applications foreseen in WP6. In addition, it included a first set of questions aimed at documenting initial external data requirements and identifying potential data gaps that may affect decision-making processes, policy-shaping, or limit the effectiveness of digital tools. The survey captures early indications of stakeholder needs, digital capacity constraints, operational challenges and the types of data required to support decision-making at farm, advisory and policy levels. It further explores expectations regarding temporal resolution, data accessibility and availability, preferred data formats and perceived shortcomings in existing datasets. By initiating this process early, the project begins cultivating a culture of co-governance among MAP members and other stakeholders. Their inputs help define shared priorities, support collective decision-making on data needs, and reinforce the co-creation aspect of Nostradamus by ensuring that data requirements are shaped collaboratively rather than imposed unilaterally. This also strengthens the sense of co-ownership, as stakeholders become active contributors to the design of the future data infrastructure and the digital tools that will rely on it.


NOSTRADAMUS

This survey is part of the Nostradamus project, which aims to develop scalable digital applications and decision-support solutions for sustainable and resilient agriculture, with a specific focus on cereal-based systems and food security in Europe. The purpose of this survey is to:

- **better understand current challenges** related to agrochemical dependence, digitalization, food security and agricultural resilience;
- **collect needs and expectations** from farmers, cooperatives, advisors, agri-industry, public authorities, NGOs, and other stakeholders;
- **provide detailed input for two Open Calls** that will fund third-party developers to build Decision-Making and Policy-Shaping applications on top of the Nostradamus data infrastructure.

Your answers will directly inform the design of these digital applications, including their functionalities, data requirements and usability features.

- Survey does not ask for your name, email, or any directly identifying information; Responses will be analyzed in aggregate only.
- Please do not include personal names or confidential business information in open-ended responses.

Time needed: approximately **20 minutes**, depending on your role and how many sections you complete.

This survey is addressed to a wide range of actors involved in agriculture and food systems, including: **(i)** Farmers / primary producers (especially in cereal-based systems); **(ii)** Farm advisors and consultants, including extension services; **(iii)** Cooperatives and producer organizations; **(iv)** Agri-industry and input suppliers (seeds, fertilizers, pesticides, machinery, services); **(v)** Agri-food companies (processing, storage, logistics, retail); **(vi)** Technology providers / ICT companies; **(vii)** Data providers (EO, IoT, market data, digital platforms); **(viii)** Public authorities and policymakers at local, regional, national and EU level; **(ix)** Public administration staff working on agriculture, environment, climate or rural development; **(x)** NGOs and civil society organizations active in agriculture, environment, climate, food security or rural communities; **(xi)** Researchers and scientists in agriculture, environment, climate, data science and related fields; **(xii)** Financial and insurance actors linked to agriculture and rural development; **(xiii)** International organizations working on agriculture, climate, environment or food security, or related fields

Glossary: To make the survey easier to follow, here are short explanations of some terms that appear in the questions.

- **Digital agriculture tools:** Any digital solution used in farming or policy work, such as mobile apps, online platforms, decision support tools, dashboards, satellite-based services, or sensor systems.
- **EO (Earth Observation):** Information about the Earth collected from satellites, aircraft or UAVs (drones), for example maps of crop growth, soil moisture or vegetation stress.
- **IoT (Internet of Things):** Connected devices and sensors in the field (for example soil moisture sensors, weather stations, devices on tractors) that automatically collect and send data.
- **Real time or near real time data:** Data that is updated very frequently (for example every few minutes

FIGURE 20: NOSTRADAMUS SURVEY THROUGH A CO-DESIGN PROCESS WITH NOSTRADAMUS CONSORTIUM

The survey is being implemented across the countries involved in the five demonstration sites, specifically Cyprus, Germany, Serbia, Slovenia and Switzerland. To broaden the reach of this activity, the survey has also been disseminated through the social media channels of the Nostradamus project. This extended distribution aims to capture additional external data requirements from actors who may not be directly linked to the demonstration sites but whose needs are valuable for establishing a wider baseline of data expectations across Europe. By collecting input from a more diverse set of respondents, the survey contributes to identifying broader patterns, recurring themes and common data needs that extend beyond local contexts. This supports the objective of the Nostradamus project to build a comprehensive, inclusive and representative understanding of external data requirements, while maintaining the principles of co-governance and co-creation that underpin the project's participatory approach. The insights gathered through this process help guide future data integration, prioritization and harmonization activities and reinforce the shared ownership of the evolving data ecosystem that Nostradamus aims to establish.

12.2.2 SYNERGIES WITH EU PROJECTS AND INITIATIVES

Synergies with EU-funded projects, initiatives and thematic networks form a central mechanism through which Nostradamus identifies, exchanges and captures external data requirements. These interactions strengthen the project's capacity to build an inclusive, scalable, and interoperable data ecosystem while ensuring alignment with the wider European research landscape. Throughout the first 16 months, Nostradamus engaged in several strategic events and collaborative activities that enabled the communication of its objectives and the potential collection of external data requirement signals from other projects and initiatives with shared interests.

A key synergy activity was the participation in the RSCy2025 International Conference in Cyprus[3], where Nostradamus co-organised the workshop titled "Advancing Climate Action, Environmental Sustainability, and Food Security through Synergies." This workshop provided an opportunity to present the project's objective of gathering external data requirements and to engage with multiple EU-funded projects and scientific teams working on climate action, food security, environmental sustainability and digital agriculture. Several initiatives expressed interest in learning how Nostradamus collects, prioritises and processes data needs and how their own requirements might benefit from the project's data architecture.

FIGURE 21: NOSTRADAMUS PARTICIPATION IN RSCY2025

[3] https://rscy2025.com/agenda-workshops/

Nostradamus also participated in the Synergy Days 2024[4] event, which brought together more than 30 EU-funded projects, including the sister project OpenAgri. During this event, Nostradamus presented the objective of consolidating external data requirements and explored shared needs with projects operating in areas such as digital agriculture, interoperability frameworks, precision farming services and data-driven policy support. These interactions contributed to building a growing network of initiatives that can benefit from the structured data requirements approach established in Task 2.2.

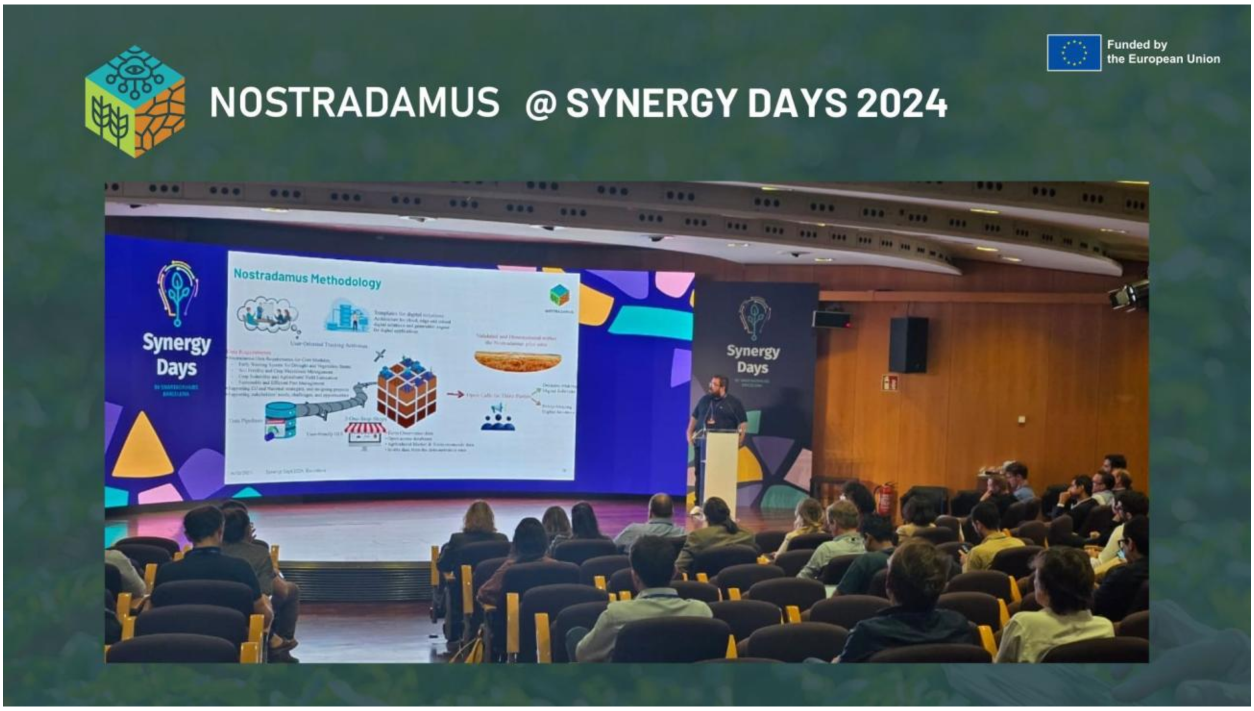


FIGURE 22: NOSTRADAMUS PARTICIPATION IN SYNERGY DAYS 2024

Additional in-person engagement occurred through participation in workshops of projects such as LandShift and ECO-READY, where Nostradamus presented its approach for identifying and assessing external data needs. These exchanges supported the identification of thematic overlaps related to agricultural sustainability, climate resilience, real-time monitoring, and multi-actor data governance. Through these discussions, the project strengthened its understanding of what types of external data requirements may arise from parallel EU initiatives and how these may need to be addressed during future iterations of the deliverable.

[4] https://www.smartagrihubs.eu/synergy-days-2024

FIGURE 23: NOSTRADAMUS PARTICIPATION IN LANDSHIFT AND ECO-READY PROJECTS' WORKSHOPS

A further synergy mechanism was established through the creation of a Hub of Projects with Similar Objectives (AgroTech Hub Cluster), initiated under WP1. The hub currently includes eight projects and will continue to expand, serving as an ongoing collaborative environment where shared challenges, cross-project needs, opportunities for exchanging data requirements, and potential joint activities and synergies can be openly discussed. The hub also provides a communication channel for informing projects about the possibility of submitting their data requirements, thereby supporting their mission and vision while also contributing to wider EU research objectives. As the hub grows, it will play an increasingly important role in enriching the database of external data requirements, supporting replicability, validation and long-term interoperability across initiatives.

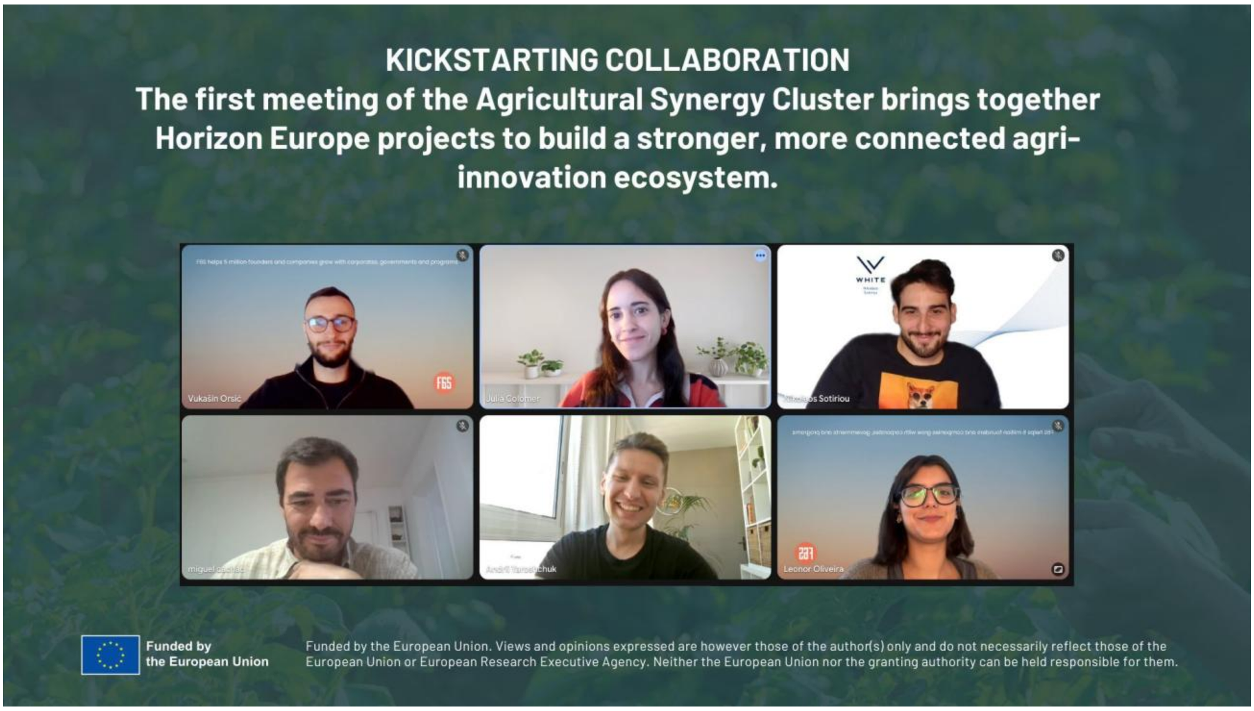


FIGURE 24: ESTABLISHMENT OF AGRICULTURAL SYNERGY CLUSTER

Looking ahead, the next period of Task 2.2 will expand the range of synergies by strengthening engagement with AKIS platforms and interoperability-driven initiatives. These actors have not yet been fully integrated during the first 16 months, but planned activities will ensure their data needs are captured in a structured and systematic manner. This aligns with the broader mandate of Task 2.2, which is to continuously identify and update external data requirements from key actors. All external requirements will be assessed using the suitability criteria defined in this deliverable. This ensures that external requests align with the technical, legal, and ethical parameters of the project and can be feasibly supported by the WP3 workflows and WP5 Data Cubes. During the next phase, Nostradamus will also organize or participate in dedicated online workshops with EU initiatives, partnerships, centres and platforms to reinforce co-creation processes and to ensure that external needs move in the optimal direction for contributing to Green Deal objectives and related policy frameworks. Finally, in collaboration with WP3 and WP5, Task 2.2 will advance efforts to define and validate the formats required for storing and processing external data inputs in the Data Cubes developed under WP5. These planned activities ensure that the synergies initiated in the first 16 months evolve into a coherent, strategic and impactful ecosystem of external data requirements supporting the long-term ambitions of Nostradamus.

12.2.3 THIRD-PARTY DEVELOPERS UNDER WP6

The process for collecting external data requirements from the successful Third-Party applicants funded under WP6 is a central element of the co-creation and co-governance model adopted by Nostradamus. These Third Parties will design and develop digital applications that operate on top of the project's data environment, and their data needs form an important category of external requirements that must be captured, validated and assessed within the evolving framework developed in Task 2.2. At this early stage of the project, it is important to emphasize that the data requirements of the successful applicants will not be automatically integrated into the data workflows of Nostradamus. Instead, they will undergo a structured review to determine their feasibility, their alignment with project objectives and their compatibility with the technical and ethical conditions of the project. Since Third Parties will already commit to specific expected outcomes as part of their Open Call agreements, their data needs are expected to remain broadly consistent with those objectives. Any additional requests beyond what is initially specified will therefore be subject to further discussion with the Nostradamus consortium and evaluated according to their relevance, technical viability and potential added value. Such requests may concern improving the spatial or temporal resolution of the data they require, enhancing the accuracy or robustness of their application outputs, or expanding the applicability of their tools across different agro-climatic contexts. These considerations will be examined collaboratively to ensure that they remain realistic within project boundaries and beneficial for both Nostradamus and the Third Parties involved.

As part of this process, the project will establish dedicated mechanisms to guide Third Parties in expressing their data needs in a structured and standardized manner. These mechanisms will be developed at a later stage and may include templates or guidance documents for specifying spatial and temporal resolution requirements, metadata expectations, preferred data formats, quality constraints, interoperability dependencies and links between datasets and intended functionalities. Their purpose is to support clarity and transparency, not to impose fixed structures prematurely. By organizing the information in a consistent and comparable format, Task 2.2 will be able to assess the requests more efficiently and ensure a fair and systematic evaluation approach. This structured engagement also reinforces the co-ownership dimension of the project, ensuring that Third Parties are not passive data users but active contributors to the evolution of the Nostradamus data ecosystem. Through continuous dialogue, feedback loops and iterative refinement, their inputs will provide valuable insights into emerging innovation needs, market expectations and operational bottlenecks that must be addressed to strengthen the practical relevance and applicability of the project's digital solutions. In this way, the integration of Third-Party perspectives becomes an important mechanism for ensuring that the outcomes of Nostradamus remain future-proof, demand-driven and aligned with both policy priorities and real-world requirements.

13 CONCLUSIONS

The work conducted under Tasks 2.1 and 2.2 forms the foundation of the Nostradamus data ecosystem by establishing the internal and external data requirements necessary for the development, validation, and long-term usability of the project's outcomes. Together, these internal and external components of Deliverable D2.1 ensure that the project adopts a comprehensive, inclusive, and forward-looking approach to data collection, harmonization, and prioritization. This integrated perspective creates the conditions for accurate modelling, effective decision support, scalable deployment, and meaningful policy impact across Europe.

Internal Data Requirements

This deliverable has defined the needs-based data requirements essential for the successful implementation of the Nostradamus project, with a particular focus on aligning technical developments with the real needs of farmers, advisors, and other agri-food stakeholders across the five demonstration sites. The document has systematically identified the specific datasets, data sources, data workflows, and key implementation parameters (KIPs) required for the development of the four core modules: (1) early warning systems for drought and vegetation stress, (2) crop suitability and yield estimation, (3) soil fertility and crop nutrition management, and (4) sustainable pest management. Additionally, the specific characteristics, existing capacities, and challenges of the five pilot sites in Germany, Serbia, Switzerland, Slovenia, and Cyprus were reviewed to ensure that the data requirements are tailored to local agro-climatic and socio-economic conditions.

The deliverable also introduced a harmonized set of Key Performance Indicators (KPIs) for each pilot site, which will serve as the basis for monitoring the effectiveness, adoption, and impact of the developed digital applications during the demonstration and validation phases.

Next Steps

The defined data requirements will be used immediately to guide:

- The development and deployment of data acquisition pipelines and data pre-processing (WP3).
- The development of the Nostradamus platform components and modules (WP3 and WP4).

- The setup and preparation of demonstration activities at pilot sites (WP4).
- The integration of monitoring and impact assessment protocols (WP5).
- The detailed definition of KPIs monitoring methodology and data collection tools.

In parallel, the project partners will continue collaborating with pilot site stakeholders through co-design activities, workshops, and MAPs to validate and refine the requirements defined in this deliverable. The outcomes of this deliverable will also directly inform upcoming deliverables such as the platform architecture (D3.1), module-specific developments (D3.2 to D3.5), and the validation and demonstration protocols (D4.x and D5.x). The work carried out in D2.1 will thus serve as a foundation for the successful implementation, demonstration, and exploitation of the Nostradamus digital solutions.

External Data Requirements

The work carried out under Task 2.2 by M16 has established the conceptual, methodological and procedural backbone for managing external data requirements in Nostradamus. Nostradamus, by M16, has designed a clear framework that defines how external needs will be identified, prioritized, evaluated and, where feasible, supported throughout the project. This includes a structured stakeholder prioritization process using a power–interest perspective, a set of suitability criteria that reflect technical, legal and policy constraints, and an explicit link to relevant EU and international strategies. Taken together, these elements ensure that external data requests are not treated as ad hoc demands but are systematically assessed in relation to the objectives of Nostradamus, the capabilities of WP3 and WP5, and the needs of the four analytical modules and associated applications. By Month 16, the first practical steps to populate this framework have also been initiated. The survey targeting MAP participants and wider stakeholders has begun to generate early signals of external data needs, digital capacity gaps and expectations across the five demonstration countries and beyond. At the same time, active participation in conferences, clustering events and project workshops has opened channels for synergies with EU-funded initiatives, while the establishment of a Hub of Projects with similar objectives provides a long-term mechanism for exchanging requirements and exploring cross-project alignment. In parallel, a process has been outlined for future engagement with Third-Parties under WP6, ensuring that their external data requests will be captured, clarified and assessed in a transparent manner once the OCs are completed. Although most of these activities are still in their initial phase, they already demonstrate that Nostradamus is building the foundations for a coherent, participatory and policy-relevant external data ecosystem.

Next Steps

Building on the foundations established in Phase 1 (M3-M16), the external needs-based data requirements framework will transition into full operational use during Phase 2 (M17-M44). The defined requirements will be applied to:

- Supporting the maturation of MAPs by integrating progressively richer stakeholder inputs and aligning data needs with real-world use cases as MAPs become more populated and structured.
- Guiding Third Parties as they initiate development activities, ensuring that their contributions are interoperable, requirements-driven, and fully embedded within the overarching data framework.
- Strengthening synergies with AKIS platforms and relevant EU initiatives, enabling cross-initiative harmonization, improved data flows, and policy-aligned data sharing practices.
- Documenting the practical application of the framework, including how it is refined in response to stakeholder feedback and operational experience, feeding into the updated version planned for Month 44.
- Contributing to the creation of an interoperable, scalable, and policy-aligned data environment that supports the full spectrum of Nostradamus solutions.

Throughout Phase 2, the framework will continue evolving as a living instrument. Partners will iteratively validate, update, and operationalize the data requirements in collaboration with MAP members, Third Parties, pilot site stakeholders, among others. These interactions will ensure that the framework remains responsive, technically relevant, and fully aligned with the broader European agricultural data landscape, thereby maximizing the long-term impact and usability of the Nostradamus platform and its associated tools.

APPENDIX A

Dataset template for WP4 modules

Data requirements for module X	
Data type: *Iot, Soil data, Remote sensing data, Field data etc.*	**Purpose and relation to the objectives of the project:** *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.* **Data formats:** *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.* **Data Schema:** *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.* **Dataset update frequency:** *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.* **Expected size of the data:** *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.* **Data utility:** *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.* **Data release roadmap:** *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*

	Privacy principles: *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.* **Metadata:** *Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.* **Snippet of the data** *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.* **Existing databases with similar dataset:** *Describe or add hyperlink to existing database with similar or same data as is in this dataset.*

Module 1 data requirements table

Data requirements for module 1	
Data type: **Remote sensing data** (Raster data)	**Purpose and relation to the objectives of the project:** Core input data for the drought monitoring and early warning module. Remote Sensing input data: • Sentinel-2 L2A data • Sentinel-2 cloud probability product • Sentinel-3 SLSTR L2 LST • Copernicus Global Land Service SSM **Data formats:** Georeferenced raster data format (GeoTiff, netCDF, GRID)

	Data Schema: Georeferenced raster data format (GeoTiff, netCDF, GRID) **Dataset update frequency:** • Sentinel-2 L2A data : every 5 days • Sentinel-2 cloud probability product : every 5 days • Sentinel-3 SLSTR L2 LST : daily • Copernicus Global Land Service SSM : daily **Expected size of the data:** The following size of data is expected to be required for covering one pilot study area. • Sentinel-2 L2A data : ca. 800 GB • Sentinel-2 cloud probability product : tbd • Sentinel-3 SLSTR L2 LST : ca. 300 GB • Copernicus Global Land Service SSM : tbd **Data utility:** Core input data for the drought monitoring and early warning module. Data can also be utilized by any stakeholder for performing own analyses. **Data release roadmap:** See dataset update frequency. **Privacy principles:** These area freely available remote sensing data provided by ESA and CGLS. **Metadata:** Information can be derived from data providers • https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2 • https://sentiwiki.copernicus.eu/web/slstr-products#S3-SLSTR-Products-L2-LST-Products • https://land.copernicus.eu/en/products/soil-moisture

	Snippet of the data: N/A **Existing databases with similar dataset:** Available from ESA/CGLS
Data type: **Meteorological data** (Raster data)	**Purpose and relation to the objectives of the project:** Core input data for the drought monitoring and early warning module. Meteorological input data • ECMWF ERA 5 data • ECMWF HRES product **Data formats:** Georeferenced raster data format (GeoTiff, netCDF, GRID) **Data Schema:** Georeferenced raster data format (GeoTiff, netCDF, GRID) **Dataset update frequency:** • ECMWF ERA 5 data : daily • ECMWF HRES product : daily **Expected size of the data:** • ECMWF ERA 5 data : ca. 50 GB • ECMWF HRES product : tbd **Data utility:** Core input data for the drought monitoring and early warning module.

	Data can also be utilized by any stakeholder for performing own analyses. **Data release roadmap:** See dataset update frequency. **Privacy principles:** These area freely available data provided by ECMWF. **Metadata:** Information can be derived from data providers • https://www.ecmwf.int/en/era5-land • https://www.ecmwf.int/en/forecasts/datasets/set-i **Snippet of the data:** N/A **Existing databases with similar dataset:** Available from ECMWF
Data type: **Spatial information input data/ maps (raster data, vector data)** **Vector input data for AOI**	**Purpose and relation to the objectives of the project:** Input data for the drought monitoring and early warning module. **Data formats:** Spatial information input data / Maps • Land cover/land use maps : georeferenced raster file (GeoTIFF) or polygon shapefile • Crop type maps : georeferenced raster file (GeoTIFF) or polygon shapefile • Soil type map (constant): georeferenced raster file (GeoTIFF) or polygon shapefile Vector input data • Polygon of area of interest (AOI) : georeferenced polygon shapefile

	Data Schema: • Georeferenced raster data format (e.g. GeoTiff) • Georeferenced polygon shapefile **Dataset update frequency:** Spatial information input data / maps • Land cover/land use maps : seasonally/yearly • Crop type maps : seasonally • Soil type map : constant/once Vector input data • Polygon of area of interest (AOI) : constant/once **Expected size of the data:** N/A **Data utility:** Input data for the drought monitoring and early warning module. Data can also be utilized by stakeholders. **Data release roadmap:** N/A **Privacy principles:** N/A **Metadata:** N/A **Snippet of the data**

	N/A **Existing databases with similar dataset:** N/A
Data type: **Data for calibration / validation** Field data / IoT data / maps / statistics / information	**Purpose and relation to the objectives of the project:** Additional data required/useful for calibration and validation **(parameters in bold are more relevant than others)** • Meteorological station data (daily) – georeferenced o Air temperature o Precipitation o Solar radiation / Sunshine duration o Wind speed o Relative humidity • Field measurement data (daily/regularly) – georeferenced o **Soil moisture** o **Soil surface humidity** o **Evapotranspiration** o Plant chlorophyll content o Phenological status (e.g. ripening, etc.) o **Land surface temperature** • Plant characteristics / phenological data / agricultural management (seasonally) – georeferenced or for field polygon o **Crop type / Plant type** o **Sowing date** o **Harvesting date / Mowing time** o **Irrigation management: Existence of an irrigation system** o Soil management o Preventive measures that have been undertaken? o Crop types of previous years • Spatial information / Maps o **Areal damage of crops (monthly/seasonally/yearly)** o **Yield maps (seasonally/annually)** o **Information about timing and areas affected/spatial extent of previous droughts (from past years)** • Statistical data (annually) o Crop yield statistics (regional detail, national coverage)

	o Yield failure • Additional information o Soil characteristics for different soil types: light/medium/heavy, AWC (available water capacity) o Plant characteristics: Individual susceptibility to drought in which month/season/at which crop status • Weather prediction / Weather forecast data o For the meteorological parameters listed above **Data formats:** N/A **Data Schema:** N/A **Dataset update frequency:** N/A **Expected size of the data:** N/A **Data utility:** N/A **Data release roadmap:** N/A **Privacy principles:** N/A

	Metadata: N/A **Snippet of the data** N/A **Existing databases with similar dataset:** N/A

Module 2 data requirements table

Data requirements for module 2	
Data type: *Iot, Soil data, Remote sensing data, Field data etc.*	Purpose and relation to the objectives of the project: **It contains in situ meteorological measurements including different environmental indicators essential for a plant development.** Data formats: **JSON, CSV, SQL** Data Schema:

Table: lora_module_type

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_module_type

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants | Row Id | DDL

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME	COLUMN_SIZE
(null)	agrosense	lora_module_type	id	4	serial	10
(null)	agrosense	lora_module_type	name	12	varchar	200
(null)	agrosense	lora_module_type	description	12	text	2147483647
(null)	agrosense	lora_module_type	code	12	text	2147483647
(null)	agrosense	lora_module_type	url	12	varchar	255
(null)	agrosense	lora_module_type	logo	12	varchar	255
(null)	agrosense	lora_module_type	module_order	4	int4	10
(null)	agrosense	lora_module_type	description_eng	12	text	2147483647

Table: lora_device_type

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_device_type

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME
(null)	agrosense	lora_device_type	id	4	serial
(null)	agrosense	lora_device_type	name	12	varchar
(null)	agrosense	lora_device_type	position_type	12	varchar
(null)	agrosense	lora_device_type	code	12	text
(null)	agrosense	lora_device_type	position_type_eng	12	varchar

Table: lora_module

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_module

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME
(null)	agrosense	lora_module	mac_address	12	varchar
(null)	agrosense	lora_module	id_lora_module_type	4	int4
(null)	agrosense	lora_module	name	12	varchar
(null)	agrosense	lora_module	public	-7	bool

Table: lora_measurement

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_measurement

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants | Row Id | DDL

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME	COLUMN_SIZE
(null)	agrosense	lora_measurement	id	4	serial	10
(null)	agrosense	lora_measurement	value	2	numeric	16
(null)	agrosense	lora_measurement	date	93	timestamptz	35
(null)	agrosense	lora_measurement	mac_address_lora_...	12	varchar	100
(null)	agrosense	lora_measurement	id_lora_device_type...	4	int4	10
(null)	agrosense	lora_measurement	id_lora_module_typ...	4	int4	10
(null)	agrosense	lora_measurement	location	1111	geometry	2147483647
(null)	agrosense	lora_measurement	device_on	-7	bool	1
(null)	agrosense	lora_measurement	cumulative_value	2	numeric	16
(null)	agrosense	lora_measurement	valid	-7	bool	1
(null)	agrosense	lora_measurement	url	12	varchar	100

Dataset update frequency:

Hourly.

	Expected size of the data: **Less than 1GB.** Data utility: **Farmers, agricultural advisors, decision and policy makers** Data release roadmap: **Access to IoT data is available.** Privacy principles: **Dataset is anonymized as no data is related to individuals and no exact location of weather stations is shared with the users. It is encrypted for a data transfer and is publicly available to everyone through the AgroSens platform.** Metadata: **Described in Data schema.** Snippet of the data

```
▾ [{sensorTypeName: "Temperatura vazduha", sensorTypeNameEng: "Air temperature", unitMeasure: "°C",…},…]
  ▾ 0: {sensorTypeName: "Temperatura vazduha", sensorTypeNameEng: "Air temperature", unitMeasure: "°C",…}
      date: 1741954322000
      fieldName: "temperature"
      position: 200
      sensorTypeName: "Temperatura vazduha"
      sensorTypeNameEng: "Air temperature"
      unitMeasure: "°C"
      unitMeasureEng: "°C"
      value: 19.8
      windDirection: null
  ▾ 1: {sensorTypeName: "Smer vetra", sensorTypeNameEng: "Wind direction", unitMeasure: "Stepen",…}
      date: 1741954322000
      fieldName: "windDirection"
      position: 200
      sensorTypeName: "Smer vetra"
      sensorTypeNameEng: "Wind direction"
      unitMeasure: "Stepen"
      unitMeasureEng: "Angle"
      value: 200
      windDirection: "S"
  ▾ 2: {sensorTypeName: "Vlažnost vazduha", sensorTypeNameEng: "Air humidity", unitMeasure: "%",…}
      date: 1741954322000
      fieldName: "humidity"
      position: 200
      sensorTypeName: "Vlažnost vazduha"
      sensorTypeNameEng: "Air humidity"
      unitMeasure: "%"
      unitMeasureEng: "%"
      value: 60
      windDirection: null
  ▸ 3: {sensorTypeName: "Brzina vetra", sensorTypeNameEng: "Wind speed", unitMeasure: "m/s",…}
  ▸ 4: {sensorTypeName: "Udar vetra", sensorTypeNameEng: "Wind gust", unitMeasure: "m/s",…}
  ▸ 5: {sensorTypeName: "Količina padavina", sensorTypeNameEng: "Precipitation", unitMeasure: "mm",…}
  ▸ 6: {sensorTypeName: "Osvetljaj", sensorTypeNameEng: "Light", unitMeasure: "lux", unitMeasureEng: "lux",…}
  ▸ 7: {sensorTypeName: "UV zračenje", sensorTypeNameEng: "UV radiation", unitMeasure: "µW/cm2",…}
```

```
▾ {startDate: 1741594809433, mac: "0080000000014A81", name: "Silbaš", moduleAddresName: "Silbaš",…}
    hasMeasurementsRight: false
    lmtDescription: "BioSens meteo stanica"
    lmtDescriptionEng: "BioSens meteo station"
    logo: "https://img1.agrosens.rs/images/sensorLogo/full/bsi.png"
    mac: "0080000000014A81"
    measurements: null
    moduleAddresName: "Silbaš"
    name: "Silbaš"
  ▾ parcels: [{parcelName: "Kudeljara", parcelId: 25680}]
    ▾ 0: {parcelName: "Kudeljara", parcelId: 25680}
        parcelId: 25680
        parcelName: "Kudeljara"
    startDate: 1741594809433
    url: "https://biosens.rs"
```


Existing databases with similar dataset:

This dataset is available through AgroSense web platform: https://agrosense.rs

Data type:

Field data

Purpose and relation to the objectives of the project:

Historical data containing field inputs and yields needed for developing yield prediction and crop suitability maps.

Data formats:

XLSX, CSV, SQL

Data Schema:

Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?	Default
id	integer					
source	text					
subsource	text					
parcel	text					
area	text					
damage_percentag	numeric					
damage_ha	numeric					
latitude	numeric					
longitude	numeric					
cadastre_parcel	text					
cadastre_municipa	text					
municipality	text					
stretch	text					
year	numeric					
crop	text					
seed_production	numeric					
variety	text					
seed_category	text					
ripening_group	text					
seed_manufacturer	text					
previous_crop	text					
seed_units	text					
manure	numeric					
fertiliser_amount	numeric					
n	numeric					
p	numeric					
k	numeric					
pesticide_amount	numeric					
tillage_costs_ha	numeric					
sowing_costs_ha	numeric					
fertiliser_applicatio	numeric					

pesticide_applicati	numeric
irrigation_costs_ha	text
harvest_costs_ha	numeric
seed_costs_ha	numeric
fertiliser_costs_ha	numeric
pesticide_costs_ha	numeric
yield	numeric
price	numeric
revenue	numeric
planting_date	timestamp without t...
harvest_date	timestamp without t...
geom	geometry

Dataset update frequency:

Seasonally / yearly / once

Expected size of the data:

Up to 10 MB

Data utility:

Research and academics, stakeholders and farmers

Data release roadmap:

Updates seasonally, or when new data is otherwise acquired

Privacy principles:

Dataset is available for internal use only, with NDAs signed with the data providers.

Metadata:

Described in Data schema

Snippet of the data

id	source	subsource	parcel	area	damage_percentage	damage_ha	latitude	longitude
0	Source 0	NULL	G-1-1	63.2	NULL	NULL	Lat (NDA)	Lon (NDA)

cadastre_parcel	cadastre_municipality	municipality	stretch	year	crop	seed_production
NULL	NULL	NULL	NULL	2020	Oilseed rape	0

variety	seed_category	ripening_group	seed_manufacturer	previous_crop	seed_units	manure
DK Imperial CL	NULL	NULL	NULL	Soybean	9	0

fertiliser_amount	n	p	k	pesticide_amount	tillage_costs_ha
891	142.12054276	117.15998738	113.48690439	4.794303797	4820.75

sowing_costs_ha	fertiliser_application_costs_ha	pesticide_application_costs_ha	irrigation_costs_ha
3031.25	7666	15084	3000

harvest_costs_ha	seed_costs_ha	fertiliser_costs_ha	pesticide_costs_ha	yield	price	revenue
11232	2809.878323	28433.5443	8236.155063	3.6	39.98	143.928

planting_date	harvest_date	geom
2019-09-17 00:00:00	2020-07-02 00:00:00	Point (NDA)

Existing databases with similar dataset:

N/A

Data type:

Soil data

Purpose and relation to the objectives of the project:

Contains relevant agricultural parameters, such as Soil Organic Matter, elevation etc.

Data formats:

Shapefile, GeoTIFF, SQL

Data Schema:

Georeferenced raster data, georeferenced shapefile data

Id	Name	Alias	Type	Type name	Length	Precision
123 0	klasa4		Integer (32 bit)	Integer	4	0
abc 1	Naziv		Text (string)	String	100	0
123 2	Klasa		Integer (32 bit)	Integer	4	0

Dataset update frequency:

One-time collection

Expected size of the data:

Up to 10GB

NOSTRADAMUS

Data utility:

Farmers, Agriculture advisors, decision and policy makers

Data release roadmap:

One-time update

Privacy principles:

Aggregation

Metadata:

Described in Data schema.

Snippet of the data

klasa4	Naziv	Klasa
221	Humofluvisol	5
226	Humoglej i euglej	6
229	Fluvisol	7
241	Koluvijum	8
245	Arenosol i sirozem na pesku	9
252	Vertisol	10
265	Districni kambisol i mestimic...	11
284	Eutricni kambisol	12
294	Pseudoglej	13
323	Luvisol i zemljista u lesiviranju	14
325	Ranker, sirozem, litosol na sk...	2
326	Ranker, sirozem, litosol na a...	2
328	Kalkomelanosol, sirozem, lit...	15
329	Ranker, sirozem, litosol na se...	2
330	Ranker, sirozem, litosol na p...	2
333	Rendzina, sirozem i litosol na...	16
335	Kalkokambisol i kalkomelan...	17
389	Ranker, sirozem, litosol na gr...	2

Existing databases with similar dataset:

Parts of the dataset are available from SoilGrids and USGS web platforms: https://files.isric.org/soilgrids/latest/data/, https://earthexplorer.usgs.gov/

Module 3 data requirements table

Data requirements for module 3

Data type:

Field data

Purpose and relation to the objectives of the project:

It is used to get metadata from specific field which is important for geospatial operations.

Data formats:

SQL, JSON and GeoJSON

Data Schema:

Column Name	Data Type	Allow Nulls
gerk_pid	int	☐
kmg_mid	int	☐
kmg_mid2nd	int	☑
raba_id	int	☑
raba_opis	varchar(60)	☑
geometry	text	☑
domace_ime	varchar(30)	☑
drzava	varchar(1)	☑
kmg_mid_prevzem	int	☑
area	decimal(20, 4)	☑
perim	decimal(20, 4)	☑
y_min	decimal(15, 3)	☑
x_min	decimal(15, 3)	☑
y_max	decimal(15, 3)	☑
x_max	decimal(15, 3)	☑
y_c	decimal(15, 3)	☑
x_c	decimal(15, 3)	☑
z_avg	decimal(15, 3)	☑
exp_avg	int	☑
fi_avg	decimal(3, 1)	☑
nagib_avg	int	☑
pov_hgo	decimal(20, 4)	☑
pov_do	decimal(20, 4)	☑
pov_po	decimal(20, 4)	☑
d_od	datetime	☑
d_do	datetime	☑
geopard_uuid	uniqueidentifier	☑
polygon	varchar(MAX)	☑
		☐

NOSTRADAMUS

	Dataset update frequency: Is needed if metadata of the field has changed (e.g. polygon of a field has changed). **Expected size of the data:** Up to 1GB **Data utility:** Farmers, agricultural advisords, decision and policy-makers **Data release roadmap:** One time-update or per request **Privacy principles:** Sensitive and personal data will be anonymised or pseudonymised before providing the dataset. No aggregation or minimisation is needed. **Metadata:** • gerk_pid: unique number (id) for field • kmg_mid: unique number (id) for farmed • raba_id: unique number (id) for field usage • raba_opis: unique description for field upsage • geometry: polygon of a field (shape) • domace_ime: name of a field • drzava: country of origin • area: size of a field (ar) **Snippet of the data** [{ "gerk_pid": 561776, "kmg_mid": xxxxxxxxx, "kmg_mid2nd": "NULL", "raba_id": 1100, "raba_opis": "Njiva", "geometry": "POLYGON ((592729.815 165606.51, 592730.78 165626.065, 592728.125 165624.27, 592553.955 165506.645, 592562.53 165493.53, 592728.7 165605.76, 592729.815 165606.51))", "domace_ime": "homename", "drzava": "S", "kmg_mid_prevzem": "NULL", "area": 3252.2346, "perim": "NULL", "y_min":

592553.955, "x_min": "NULL", "y_max": 592730.78, "x_max": "NULL", "y_c": 592642.367, "x_c": 165559.798, "z_avg": 181.719, "exp_avg": 177, "fi_avg": 0.8, "nagib_avg": 1, "pov_hgo": 0, "pov_do": 0, "pov_po": 0, "d_od": "2021-01-22 17:05:32.000", "d_do": "NULL", "geopard_uuid": "B04DE7B2-E565-485C-AE5D-7CCEABE55F34", "polygon": [[[16.20867519, 46.62267067], [16.21086704, 46.6236572], [16.21088175, 46.62366379], [16.21089827, 46.62383954], [16.21086325, 46.62382377], [16.20856587, 46.62278982], [16.20867519, 46.62267067]]] }]

Existing databases with similar dataset:

This data is available in country database: https://rkg.gov.si/vstop/ . Credentials are needed for access, not publicly available. Public access is available: https://rkg.gov.si/GERK/WebViewer/ .

Data type:

Lab analysis

Purpose and relation to the objectives of the project:

It contains basic soil analysis from field.

Data formats:

JSON, CSV, SQL

Data Schema:

ZemljaRezultati (MojeGospodarstvo)
id_vzorca
id_rezultata
id_parametra
parnaziv
rezultat
enota
mejna_vrednost

ZemljaVzorci (MojeGospodarstvo)
id_vzorca
oznaka
id_gerka
gerk_labtesti
datum_sprejema
id_matriksa
matriks
id_stranke
narocnik
nar_sif_rac
kmgmid
unreliable_test

Dataset update frequency:

Once per year or if new analysis is available.

Expected size of the data:

	Less than 1GB **Data utility:** Farmers, agriculture advisory board **Data release roadmap:** Access to soil samples are available. **Privacy principles:** Sensitive and personal data will be anonymised or pseudonymised before providing the dataset. No aggregation or minimisation is needed. **Metadata:** • id_vzorca: unique id for sample • Id_rezultata: unique id for sample results • Id_parametra: unique id for sample measure • parnaziv: description of taken analysis • rezultat: result of the measurement • enota: unit of the measurement • mejna_vrednost: treshold value • oznaka: description of the name of the sample • id_gerka: unique number (id) for field • datum_sprejema: timestamp of measurement • id_stranke: internal unique id for costumer • narocnik: name of the customer • kmgmid: unique id for farmer **Snippet of the data** [{ "id_vzorca": 128246, "oznaka": "name", "id_gerka": xxxxxx, "gerk_labtesti": "4157063,4157029", "datum_sprejema": "2024-04-09 08:40:58.000", "id_matriksa": 560, "matriks": "ZEMLJA-NJIVA", "id_stranke": 16382, "narocnik": "name surname", "nar_sif_rac": "NULL", "kmgmid": xxxxxxxxx, "unreliable_test": 1 }]

<table>
<tr><td></td><td>
[
{
"id_vzorca": 54158,
"id_rezultata": 357577,
"id_parametra": 498,
"parnaziv": "Kalij v zemlji - K",
"rezultat": 58.6,
"enota": "mg/kg",
"mejna_vrednost": "166 - 250"
}
]

Existing databases with similar dataset:

Internal database for lab samples in KGZMS and DIH AGRIFOOD platform.
</td></tr>
<tr><td>Data type:

IoT sensor data</td><td>
Purpose and relation to the objectives of the project:

High-resolution environmental datasets—including in situ meteorological measurements are integral to the calibration and validation of numerical weather prediction models and advanced soil property mapping algorithms.

Data formats:

JSON

Data Schema:

{
"type": "object",
"properties": {
"id": {
"type": "string",
"format": "uuid",
"nullable": true
},
"sensorId": {
"type": "string",
</td></tr>
</table>

	"nullable": true
	},
	"sensorName": {
	"type": "string",
	"nullable": true
	},
	"sensorDisplayname": {
	"type": "string",
	"nullable": true
	},
	"active": {
	"type": "boolean"
	},
	"organizationId": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"organizationName": {
	"type": "string",
	"nullable": true
	},
	"suborganizationId": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"suborganizationName": {
	"type": "string",
	"nullable": true
	},
	"dateCaptured": {
	"type": "string",
	"format": "date-time",
	"nullable": true
	},
	"yearCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"monthCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dayCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dateCreated": {
	"type": "string",
	"format": "date-time",
	"nullable": true
	},
	"yearCreated": {
	"type": "integer",

	"format": "int32",
	"nullable": true
	},
	"monthCreated": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dayCreated": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"temperature": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"airPressure": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"humidity": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"soilMoisture": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"leafWetness": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"batteryLevel": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"status": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	}
	},
	"additionalProperties": false
	}
	
	Dataset update frequency:
	
	Every few minutes, depends on the server and data from IoT sensors.

Expected size of the data:

10GB weather data from ARSO Meteo and 100MB from DIH AGRIFOOD IoT sensors.

Data utility:

Farmers, Agriculture advisords, decision and policy makers

Data release roadmap:

Access to IoT sensor data is available.

Privacy principles:

Sensitive and personal data will be anonymised or pseudonymised before providing the dataset. No aggregation or minimisation is needed.

Metadata:

Described in data schema.

Snippet of the data

{
"id": "922304ea-04a1-4e31-b437-c790b3f3f04a",
"sensorId": "0004A30B012BB395",
"sensorName": "ZT Cold Store 2",
"sensorDisplayname": "ZT Cold Store 2",
"active": true,
"organizationId": 1,
"organizationName": "ITC",
"suborganizationId": 10,
"suborganizationName": "Zelena Tocka",
"dateCaptured": "2024-01-01T00:02:32.867",
"yearCaptured": 2024,
"monthCaptured": 1,
"dayCaptured": 1,
"dateCreated": "2024-01-01T00:02:32.887",
"yearCreated": 2024,
"monthCreated": 1,
"dayCreated": 1,
"temperature": 5.80,
"airPressure": 986.4,
"humidity": 75.63,
"soilMoisture": null,
"leafWetness": null,
"batteryLevel": 2.87,

NOSTRADAMUS

	"status": 0 } **Existing databases with similar dataset:** This dataset is available through DIH AGRIFOOD data marker which is under the umbrella of DIH AGRIFOOD DataSpace: https://dataspace.dih-agrifood.com/

Module 4 data requirements table

Data requirements for module 4	
Data type: *Iot, Soil data, Remote sensing data, Field data etc.*	**Purpose and relation to the objectives of the project:** Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project. **Data formats:** JSON, CSV, SQL **Data Schema:**

Variable name	*Variable type*	*Format*	*nulls*
id	*string*	*uuid*	*true*
sensorId	*string*	-	*true*
sensorName	*string*	-	*true*
organizationId	*integer*	*int32*	*true*
organizationName	*string*	-	*true*
active	*boolean*	-	-
recordedDate	*datetime*	*Date-time*	*true*

	<table><tr><td>*airTemperature*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*humidity*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*vaporPressure*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*barometricPressure*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*windSpeed*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*direction*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*solarRadiation*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*dualPrecipitation*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*soilMoisture*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*soilTemperature*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*electricConductivity*</td><td>*number*</td><td>*float*</td><td>*true*</td></tr><tr><td>*depth*</td><td>*integer*</td><td>*int32*</td><td>*true*</td></tr></table> **Dataset update frequency:** Usually every 15 minutes. **Expected size of the data:** Up to 5GB. **Data utility:** Researchers, academics, stakeholders and farmers. **Data release roadmap:** The data release roadmap will be upon update request.

	Privacy principles: The personal data will be anonymised before something will be shared. **Metadata:** id -> Field id sensorId -> each sensor id sensorName -> each sensor name organizationId -> of which organization get data from the database organizationName -> the name of the organization get data from the database active -> if the field is active **Snippet of the data** N/A **Existing databases with similar dataset:** N/A

Dataset template for demonstration sites and existing databases

Insert name of the dataset	
Name of the dataset	
Pilot site	
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	

NOSTRADAMUS

Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	
Data Schema *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.*	
Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	
Data origin	

Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).	
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	
Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group*	

level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.	
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	
Metadata *Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.*	
Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	
Do you intent to apply the FAIR principles to your dataset ?	

Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available here.	

Demonstration site in Slovenia: Dataset 1

IoT dataset	
Name of the dataset	IoT Sensors
Pilot site	Slovenia
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Restricted access
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	IoT data available for specific location, as is temperature, humidity, air pressure etc.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	IoT data
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	JSON

Data Schema
 	
Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.
	{
	"type": "object",
	"properties": {
	"id": {
	"type": "string",
	"format": "uuid",
	"nullable": true
	},
	"sensorId": {
	"type": "string",
	"nullable": true
	},
	"sensorName": {
	"type": "string",
	"nullable": true
	},
	"sensorDisplayname": {
	"type": "string",
	"nullable": true
	},
	"active": {
	"type": "boolean"
	},
	"organizationId": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"organizationName": {
	"type": "string",
	"nullable": true
	},
	"suborganizationId": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"suborganizationName": {
	"type": "string",
	"nullable": true
	},
	"dateCaptured": {
	"type": "string",
	"format": "date-time",
	"nullable": true
	},

	"yearCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"monthCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dayCaptured": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dateCreated": {
	"type": "string",
	"format": "date-time",
	"nullable": true
	},
	"yearCreated": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"monthCreated": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"dayCreated": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"temperature": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"airPressure": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"humidity": {
	"type": "number",

	"format": "double",
	"nullable": true
	},
	"soilMoisture": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"leafWetness": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	},
	"batteryLevel": {
	"type": "number",
	"format": "double",
	"nullable": true
	},
	"status": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	}
	},
	"additionalProperties": false
	}

Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	It is possible to re-use this data with other geospatial data.
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	Data is collected through various sensors which are under the umbrella of DIH AGRIFOOD and ITC and are publicly available.

Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Invidual, company or other organisation, defined in DIH Agrifood Data Space Dataset Terms of use.
Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	Depends on IoT sensor, between 5-20 minutes.
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Around 1GB
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Farmers, decifion makers, research and academics etc., defined in DIH AGRIFOOD Data Space Dataset Terms of use.
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Partly anonymized.
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per*	Updated depends on IoT Sensors, between 5-20 minutes.

year every January), event-driven updates, one-time update etc.	
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	It is stored in DIH Agrifood Data Space: https://dataspace.dih-agrifood.com/
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	User needs to register on DIH Agrifood and then user can access Datamarket: https://datamarket.dataspace.dih-agrifood.com/apis
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	HTTP Request.
Metadata *Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.*	Described in Data schema.
Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	{ "id": "922304ea-04a1-4e31-b437-c790b3f3f04a", "sensorId": "0004A30B012BB395", "sensorName": "ZT Cold Store 2", "sensorDisplayname": "ZT Cold Store 2", "active": true, "organizationId": 1, "organizationName": "ITC", "suborganizationId": 10, "suborganizationName": "Zelena Tocka", "dateCaptured": "2024-01-01T00:02:32.867", "yearCaptured": 2024, "monthCaptured": 1, "dayCaptured": 1, "dateCreated": "2024-01-01T00:02:32.887", "yearCreated": 2024, "monthCreated": 1, "dayCreated": 1,

	"temperature": 5.80, "airPressure": 986.4, "humidity": 75.63, "soilMoisture": null, "leafWetness": null, "batteryLevel": 2.87, "status": 0 }
Do you intent to apply the FAIR principles to your dataset ? *Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available here.*	**Partly.**

Demonstration site in Serbia: Dataset 1

IoT dataset	
Name of the dataset	IoT Sensors
Pilot site	Serbia
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Restricted access
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	IoT data available for specific location: air temperature and humidity, wind direction and speed, UV radiation, illumination and precipitation
Data types	IoT data

<table>
<tr><td>Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.</td><td></td></tr>
<tr><td>Data formats

Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.</td><td>JSON, CSV, SQL</td></tr>
<tr><td>Data Schema

Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.</td><td>
Table: lora_measurement
AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_measurement
Info | Columns | Data | Row Count | Primary Key | Indexes | Grants | Row Id | DDL
<table>
<tr><th>TABLE_CAT</th><th>TABLE_SCHEM</th><th>TABLE_NAME</th><th>COLUMN_NAME</th><th>DATA_TYPE</th><th>TYPE_NAME</th><th>COLUMN_SIZE</th></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>id</td><td>4</td><td>serial</td><td>10</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>value</td><td>2</td><td>numeric</td><td>16</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>date</td><td>93</td><td>timestamptz</td><td>35</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>mac_address_lora_...</td><td>12</td><td>varchar</td><td>100</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>id_lora_device_type...</td><td>4</td><td>int4</td><td>10</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>id_lora_module_typ...</td><td>4</td><td>int4</td><td>10</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>location</td><td>1111</td><td>geometry</td><td>2147483647</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>device_on</td><td>-7</td><td>bool</td><td>1</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>cumulative_value</td><td>2</td><td>numeric</td><td>16</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>valid</td><td>-7</td><td>bool</td><td>1</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_measurement</td><td>url</td><td>12</td><td>varchar</td><td>100</td></tr>
</table>
Table: lora_module_type
AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_module_type
Info | Columns | Data | Row Count | Primary Key | Indexes | Grants | Row Id | DDL
<table>
<tr><th>TABLE_CAT</th><th>TABLE_SCHEM</th><th>TABLE_NAME</th><th>COLUMN_NAME</th><th>DATA_TYPE</th><th>TYPE_NAME</th><th>COLUMN_SIZE</th></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>id</td><td>4</td><td>serial</td><td>10</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>name</td><td>12</td><td>varchar</td><td>200</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>description</td><td>12</td><td>text</td><td>2147483647</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>code</td><td>12</td><td>text</td><td>2147483647</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>url</td><td>12</td><td>varchar</td><td>255</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>logo</td><td>12</td><td>varchar</td><td>255</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>module_order</td><td>4</td><td>int4</td><td>10</td></tr>
<tr><td>(null)</td><td>agrosense</td><td>lora_module_type</td><td>description_eng</td><td>12</td><td>text</td><td>2147483647</td></tr>
</table>
</td></tr>
</table>

Table: lora_device_type

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_device_type

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME
(null)	agrosense	lora_device_type	id	4	serial
(null)	agrosense	lora_device_type	name	12	varchar
(null)	agrosense	lora_device_type	position_type	12	varchar
(null)	agrosense	lora_device_type	code	12	text
(null)	agrosense	lora_device_type	position_type_eng	12	varchar

Table: lora_module

AS PROD/Databases/agri/Schemas/agrosense/Tables/lora_module

Info | Columns | Data | Row Count | Primary Key | Indexes | Grants

TABLE_CAT	TABLE_SCHEM	TABLE_NAME	COLUMN_NAME	DATA_TYPE	TYPE_NAME
(null)	agrosense	lora_module	mac_address	12	varchar
(null)	agrosense	lora_module	id_lora_module_type	4	int4
(null)	agrosense	lora_module	name	12	varchar
(null)	agrosense	lora_module	public	-7	bool

Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	It is possible to re-use this data with other geospatial data.
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	Historical and real time data are being collected as a part of the internal BioSense Institute project from the institute's weather stations. Dataset is publicly available through AgroSens web platform, yet not available for direct download.
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public*	BioSense, the public Research and Development Institute for IT in biosystems, is the owner of the weather stations and corresponding data.

institution, collaborative ownership, open data ownership etc.	
Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	Hourly.
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Less than 1GB.
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Farmers, Agriculture advisors, decision and policy makers.
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Dataset is anonymized as no data is related to individuals and no exact location of weather stations is shared with the users. It is encrypted for a data transfer and is publicly available to everyone through the AgroSens platform.
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	Hourly.
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud*	This dataset is available through AgroSense web platform: https://agrosens.rs/

storage services, Github or Gitlab, Local server od database, public data archive etc.	
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	N/A
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	Dataset is available through web access and REST API.
Metadata *Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.*	Described in Data schema.

Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	[{sensorTypeName: "Temperatura vazduha", sensorTypeNameEng: "Air temperature", unitMeasure: "°C",…},…] 0: {sensorTypeName: "Temperatura vazduha", sensorTypeNameEng: "Air temperature", unitMeasure: "°C",…} date: 1741954322000 fieldName: "temperature" position: 200 sensorTypeName: "Temperatura vazduha" sensorTypeNameEng: "Air temperature" unitMeasure: "°C" unitMeasureEng: "°C" value: 19.8 windDirection: null 1: {sensorTypeName: "Smer vetra", sensorTypeNameEng: "Wind direction", unitMeasure: "Stepen",…} date: 1741954322000 fieldName: "windDirection" position: 200 sensorTypeName: "Smer vetra" sensorTypeNameEng: "Wind direction" unitMeasure: "Stepen" unitMeasureEng: "Angle" value: 200 windDirection: "S" 2: {sensorTypeName: "Vlažnost vazduha", sensorTypeNameEng: "Air humidity", unitMeasure: "%",…} date: 1741954322000 fieldName: "humidity" position: 200 sensorTypeName: "Vlažnost vazduha" sensorTypeNameEng: "Air humidity" unitMeasure: "%" unitMeasureEng: "%" value: 60 windDirection: null 3: {sensorTypeName: "Brzina vetra", sensorTypeNameEng: "Wind speed", unitMeasure: "m/s",…} 4: {sensorTypeName: "Udar vetra", sensorTypeNameEng: "Wind gust", unitMeasure: "m/s",…} 5: {sensorTypeName: "Količina padavina", sensorTypeNameEng: "Precipitation", unitMeasure: "mm",…} 6: {sensorTypeName: "Osvetljaj", sensorTypeNameEng: "Light", unitMeasure: "lux", unitMeasureEng: "lux",…} 7: {sensorTypeName: "UV zračenje", sensorTypeNameEng: "UV radiation", unitMeasure: "µW/cm2",…} {startDate: 1741594809433, mac: "0080000000014A81", name: "Silbaš", moduleAddresName: "Silbaš",…} hasMeasurementsRight: false lmtDescription: "BioSens meteo stanica" lmtDescriptionEng: "BioSens meteo station" logo: "https://img1.agrosens.rs/images/sensorLogo/full/bsi.png" mac: "0080000000014A81" measurements: null moduleAddresName: "Silbaš" name: "Silbaš" parcels: [{parcelName: "Kudeljara", parcelId: 25680}] 0: {parcelName: "Kudeljara", parcelId: 25680} parcelId: 25680 parcelName: "Kudeljara" startDate: 1741594809433 url: "https://biosens.rs"
Do you intent to apply the FAIR principles to your dataset ? *Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available <u>here</u>.*	Partly.

Demonstration site in Serbia: Dataset 2

Field data dataset	
Name of the dataset	Field data
Pilot site	Serbia

Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Internal use only
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	The dataset includes historical field data, with farm information (location, area, etc.), field operations, sowing and harvest dates, yields, and other relevant agricultural parameters.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	Field data, Geospatial data
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	XLSX, CSV, SQL

Data Schema *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.*	see schema below

Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?	Default
id	integer					
source	text					
subsource	text					
parcel	text					
area	text					
damage_percentag	numeric					
damage_ha	numeric					
latitude	numeric					
longitude	numeric					
cadastre_parcel	text					
cadastre_municipa	text					
municipality	text					
stretch	text					
year	numeric					
crop	text					
seed_production	numeric					
variety	text					
seed_category	text					
ripening_group	text					
seed_manufacturei	text					
previous_crop	text					
seed_units	text					
manure	numeric					
fertiliser_amount	numeric					
n	numeric					
p	numeric					
k	numeric					
pesticide_amount	numeric					
tillage_costs_ha	numeric					
sowing_costs_ha	numeric					
fertiliser_applicatic	numeric					

pesticide_applicati	numeric
irrigation_costs_ha	text
harvest_costs_ha	numeric
seed_costs_ha	numeric
fertiliser_costs_ha	numeric
pesticide_costs_ha	numeric
yield	numeric
price	numeric
revenue	numeric
planting_date	timestamp without t...
harvest_date	timestamp without t...
geom	geometry

Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	It is possible to re-use this data with other geospatial data.
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	This data was acquired through collaboration with various stakeholders (farmers, agricultural companies, seed companies, insurance companies).
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Partner companies (NDA signed)

Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	Annually, after current season's harvest, or when new data is otherwise acquired.
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Up to 10MB
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Research and academics, data scientists and analysts, business analysts and strategists
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Privacy of individuals included in the dataset is protected by NDAs signed with the stakeholders who provide the data.
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	The dataset is updated after new harvest data becomes available, or if a new data source is added.
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	The dataset is stored on local servers and in local databases.

<table>
<tr><td>Link to dataset

If the dataset is available online, provide a direct link to its repository or download page.</td><td>N/A</td></tr>
<tr><td>Data access

Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.</td><td>N/A</td></tr>
<tr><td>Metadata

Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.</td><td>Described in Data schema.</td></tr>
<tr><td>Snippet of the data

Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.</td><td>
<table>
<tr><td>id</td><td>source</td><td>subsource</td><td>parcel</td><td>area</td><td>damage_percentage</td><td>damage_ha</td><td>latitude</td><td>longitude</td></tr>
<tr><td>0</td><td>Source 0</td><td>NULL</td><td>G-1-1</td><td>63.2</td><td>NULL</td><td>NULL</td><td>Lat (NDA)</td><td>Lon (NDA)</td></tr>
</table>
<table>
<tr><td>cadastre_parcel</td><td>cadastre_municipality</td><td>municipality</td><td>stretch</td><td>year</td><td>crop</td><td>seed_production</td></tr>
<tr><td>NULL</td><td>NULL</td><td>NULL</td><td>NULL</td><td>2020</td><td>Oilseed rape</td><td>0</td></tr>
</table>
<table>
<tr><td>variety</td><td>seed_category</td><td>ripening_group</td><td>seed_manufacturer</td><td>previous_crop</td><td>seed_units</td><td>manure</td></tr>
<tr><td>DK Imperial CL</td><td>NULL</td><td>NULL</td><td>NULL</td><td>Soybean</td><td>9</td><td>0</td></tr>
</table>
<table>
<tr><td>fertiliser_amount</td><td>n</td><td>p</td><td>k</td><td>pesticide_amount</td><td>tillage_costs_ha</td></tr>
<tr><td>891</td><td>142.12054276</td><td>117.15998738</td><td>113.48690439</td><td>4.794303797</td><td>4820.75</td></tr>
</table>
<table>
<tr><td>sowing_costs_ha</td><td>fertiliser_application_costs_ha</td><td>pesticide_application_costs_ha</td><td>irrigation_costs_ha</td></tr>
<tr><td>3031.25</td><td>7666</td><td>15084</td><td>3000</td></tr>
</table>
<table>
<tr><td>harvest_costs_ha</td><td>seed_costs_ha</td><td>fertiliser_costs_ha</td><td>pesticide_costs_ha</td><td>yield</td><td>price</td><td>revenue</td></tr>
<tr><td>11232</td><td>2809.878323</td><td>28433.5443</td><td>8236.155063</td><td>3.6</td><td>39.98</td><td>143.928</td></tr>
</table>
<table>
<tr><td>planting_date</td><td>harvest_date</td><td>geom</td></tr>
<tr><td>2019-09-17 00:00:00</td><td>2020-07-02 00:00:00</td><td>Point (NDA)</td></tr>
</table>
</td></tr>
<tr><td>Do you intent to apply the FAIR principles to your dataset ?

Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available <u>here</u>.</td><td>Partly.</td></tr>
</table>

Demonstration site in Serbia: Dataset 3

Soil dataset	
Name of the dataset	Soil dataset
Pilot site	Serbia
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Restricted access
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	This dataset includes various soil characteristics that are relevant to agricultural production, ranging from Soil Organic Matter maps and mechanical properties of soil to soil classes.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	Soil and environment data, Geospatial data
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	Shapefile, GeoTIFF, Postgres

<table>
<tr>
<td>Data Schema

Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.</td>
<td>
<table>
<tr><th>Id</th><th>Name</th><th>Alias</th><th>Type</th><th>Type name</th><th>Length</th><th>Precision</th></tr>
<tr><td>123 0</td><td>klasa4</td><td></td><td>Integer (32 bit)</td><td>Integer</td><td>4</td><td>0</td></tr>
<tr><td>abc 1</td><td>Naziv</td><td></td><td>Text (string)</td><td>String</td><td>100</td><td>0</td></tr>
<tr><td>123 2</td><td>Klasa</td><td></td><td>Integer (32 bit)</td><td>Integer</td><td>4</td><td>0</td></tr>
</table>
+ georeferenced raster data
</td>
</tr>
<tr>
<td>Re-use of existing data

Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.</td>
<td>It is possible to re-use this data with other geospatial data.</td>
</tr>
<tr>
<td>Data origin

Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).</td>
<td>SOM maps and vector maps of soil classes were produced by BioSense Institute. Other data is publicly available, coming from global data sources such as SoilGrids and CGIAR-CSI.</td>
</tr>
<tr>
<td>Data owner

Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.</td>
<td>BioSense, the public Research and Development Institute for IT in biosystems, is the owner of the SOM maps and soil classification vector maps. SoilGrids maps are a global soil data product generated at ISRIC — World Soil Information (https://www.isric.org/explore/soilgrids). CGIAR-CSI SRTM maps are originally produced by NASA (https://srtm.csi.cgiar.org).</td>
</tr>
<tr>
<td>Dataset update frequency

Specify how oftenthe dataset is refreshed or updated, eg. Real-</td>
<td>One-time collection</td>
</tr>
</table>

time, daily, weekly, monthly, one-time collection, as needed etc.	
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Up to 10GB.
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Farmers, Agriculture advisors, decision and policy makers
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Aggregation
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	One-time update
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	SoilGrids and CGIAR-CSI data are publicly available on their respective online portals. Soil classification vector map is available on Cropt's Space Garden platform which requires user credentials.
Link to dataset	SoilGrids: https://files.isric.org/soilgrids/latest/data/ CGIAR-CSI: https://earthexplorer.usgs.gov/

<table>
<tr><td>If the dataset is available online, provide a direct link to its repository or download page.</td><td></td></tr>
<tr><td>Data access

Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.</td><td>Dataset is available through web access and REST API.</td></tr>
<tr><td>Metadata

Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.</td><td>Described in Data schema.</td></tr>
<tr><td>Snippet of the data

Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.</td><td>
<table>
<tr><th>klasa4</th><th>Naziv</th><th>Klasa</th></tr>
<tr><td>221</td><td>Humofluvisol</td><td>5</td></tr>
<tr><td>226</td><td>Humoglej i euglej</td><td>6</td></tr>
<tr><td>229</td><td>Fluvisol</td><td>7</td></tr>
<tr><td>241</td><td>Koluvijum</td><td>8</td></tr>
<tr><td>245</td><td>Arenosol i sirozem na pesku</td><td>9</td></tr>
<tr><td>252</td><td>Vertisol</td><td>10</td></tr>
<tr><td>265</td><td>Districni kambisol i mestimic...</td><td>11</td></tr>
<tr><td>284</td><td>Eutricni kambisol</td><td>12</td></tr>
<tr><td>294</td><td>Pseudoglej</td><td>13</td></tr>
<tr><td>323</td><td>Luvisol i zemljista u lesiviranju</td><td>14</td></tr>
<tr><td>325</td><td>Ranker, sirozem, litosol na sk...</td><td>2</td></tr>
<tr><td>326</td><td>Ranker, sirozem, litosol na a...</td><td>2</td></tr>
<tr><td>328</td><td>Kalkomelanosol, sirozem, lit...</td><td>15</td></tr>
<tr><td>329</td><td>Ranker, sirozem, litosol na se...</td><td>2</td></tr>
<tr><td>330</td><td>Ranker, sirozem, litosol na p...</td><td>2</td></tr>
<tr><td>333</td><td>Rendzina, sirozem i litosol na...</td><td>16</td></tr>
<tr><td>335</td><td>Kalkokambisol i kalkomelan...</td><td>17</td></tr>
<tr><td>389</td><td>Ranker, sirozem, litosol na gr...</td><td>2</td></tr>
</table>
</td></tr>
<tr><td>Do you intent to apply the FAIR principles to your dataset ?

Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility,</td><td>Partly.</td></tr>
</table>

Interoperability and Reusability. More explanation about FAIR principles is available here.	

Demonstration site in Cyprus: Dataset 1 (Meteorological data)

IoT Meteorological Dataset	
Name of the dataset	Meteorological data
Pilot site	Cyprus
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Restricted Access
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	IoT module for measures 8 weather variables (e.g Air Temperature, relative humidity, vapor pressure etc.)
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	IoT Data
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX,*	JSON, CSV, SQL

JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.	
Data Schema *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.*	{ "type": "object", "properties": { "id": { "type": "string", "format": "uuid", "nullable": true }, "sensorId": { "type": "string", "nullable": true }, "sensorName": { "type": "string", "nullable": true }, *"organizationId": { "type": "integer", "format": "int32", "nullable": true }, "organizationName": { "type": "string", "nullable": true },* "active": { "type": "boolean" }, "recordedDate": { "type": "datetime", "format": "date-time", "nullable": true }, "airTemperature": { "type": "number", "format": "float", "nullable": true }, "humidity": { "type": "number",

	"format": "float",
	"nullable": true
	},
	"vaporPressure": {
	"type": "number",
	"format": "float",
	"nullable": true
	},
	"barometricPressure": {
	"type": "number",
	"format": "float",
	"nullable": true
	},
	"windSpeed": {
	"type": "number",
	"format": "float",
	"nullable": true
	},
	"direction": {
	"type": "number",
	"format": "float",
	"nullable": true
	}
	"solarRadiation": {
	"type": "number",
	"format": "float",
	"nullable": true
	}
	"dualPrecipitation ": {
	"type": "number",
	"format": "float",
	"nullable": true
	}
	}

Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis,*	They can be re-used in statistical and analytical depends on data structure and compatibility.

predictive modeling, machine learning etc.	
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	The data are collected through sensors in real time. The data are stored in ZENTRA Cloud and they can be used by integrated with datasets for GIS, predictive modelling and environmental analysis.
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Individual, Organization, Project depend on who is deploying and operating the weather forecasting based on ZENTRA Cloud terms of use because the data is stored on them.
Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	The data transmission frequency can be adjusted based on the needs (usually 15 minutes measurements).
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Around 5GB.
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Researchers, academics, stakeholders and farmers.
Privacy principles	Partly anonymized.

Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.	
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	Based on the needs, usually update every 15 minutes.
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	It is stored on ZENTRA Cloud.
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	N/A
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	REST API, FTP access
Metadata *Provide detailed information about the dataset's structure, conten and context helping other stakeholders to understand its purpose, component and characteristics.*	sensorID, sensorName, recordedDate
Snippet of the data *Provides a small sample or example of the dataset's content, giving the*	N/A

stakeholders a preview of its structure, fields and data types.	
Do you intent to apply the FAIR principles to your dataset ? *Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available here.*	**Partly**

Demonstration site in Cyprus: Dataset 2 (Soil data)

Soil data Dataset	
Name of the dataset	Soil data
Pilot site	Cyprus
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Restricted Access
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	IoT module for measures advanced Soil Moisture and Temperature.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	IoT Data
Data formats	JSON, CSV, SQL

Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.	

Data Schema
 	
Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.
	{
	"type": "object",
	"properties": {
	"id": {
	"type": "string",
	"format": "uuid",
	"nullable": true
	},
	"sensorId": {
	"type": "string",
	"nullable": true
	},
	"sensorName": {
	"type": "string",
	"nullable": true
	},
	"active": {
	"type": "boolean"
	},
	"recordedDate": {
	"type": "datetime",
	"format": "date-time",
	"nullable": true
	},
	"soilMoisture": {
	"type": "number",
	"format": "float",
	"nullable": true
	},
	"soilTemperature": {
	"type": "number",
	"format": "float",
	"nullable": true
	},
	"electricConductivity": {
	"type": "number",
	"format": "float",
	"nullable": true
	},

	"depth": {
	"type": "integer",
	"format": "int32",
	"nullable": true
	}
	},
	}

Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	They can be re-used in statistical and analytical depends on data structure and compatibility.
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	The data are collected through soil moisture sensor in real time. The data are stored in ZENTRA Cloud and they can be used for specific analysis or used cases.
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Individual or Organization depend on who owns and operates the soil moisture sensor.
Dataset update frequency *Specify how oftenthe dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	The data transmission frequency can be adjusted based on the needs (usually 15 minutes measurements).

Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Around 5GB.
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Researchers, academics, stakeholders and farmers.
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Partly anonymized.
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	Based on the needs, usually update every 15 minutes.
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	It is stored on ZENTRA Cloud.
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	N/A
Data access *Explain the methods available for accessing the dataset, eg. Web*	REST API, FTP access

access, REST API, FTP access, downloadable apckage, database query etc.	
Metadata *Provide detailed information about the dataset's structure, content and context helping other stakeholders to understand its purpose, component and characteristics.*	sensorID, sensorName, recordedDate
Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	N/A
Do you intent to apply the FAIR principles to your dataset ? *Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available here.*	**Partly**

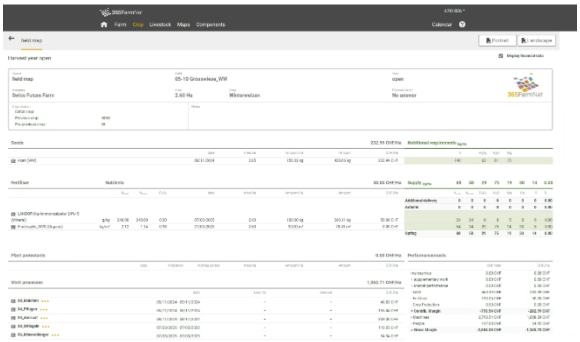
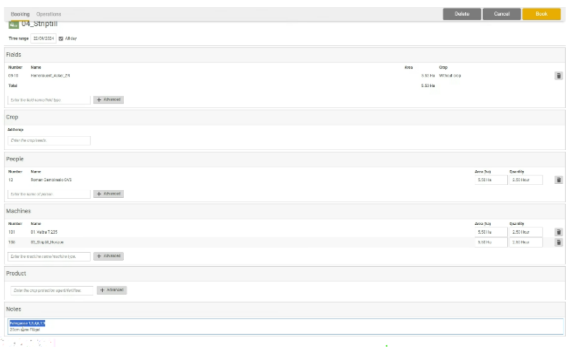
Demonstration site in Switzerland: Dataset 1 – Field Data

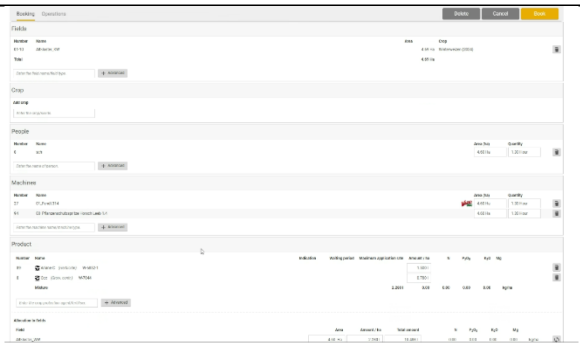
SFF Field Data	
Name of the dataset	SFF Field data
Pilot site	Swiss Future Farm (WBF – Agroscope)
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Access etc.*	Through Barto software/API (FMIS, which was the former 365farmnet) Through Exatrek software/API Other dataset sources standardized across Switzerland but very needed to the core modules: - Meteoswiss www.meteoschweiz.ch (wheather data, forecasts) - Agrometeo www.agrometeo.ch (wheater data, pest forecast models)

	- Swisstopo www.swisstopo.ch (altitude models, geografical infos) - Soil classification data available on multiple sources - Stationen | Bewässerungsnetz - (soil moisture sensor network) - Field boarders: https://map.geo.tg.ch/
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	Nationwide data portal and network can provide a quick and standardized way to access the data needed for the 4 modules. Local datasets acquired at farm level can give a more granular control. However, the challenge remains in the continuous availability of the data and whether all of them are useful to the current modules, given that these modules are mainly developed based on nation-wide available data (such as Remote sensing). On the other hand, we would like to use these data as requirements for the pilot-site-specific modules we envision in the context of the open call. We will be happy to provide more information, context and details about these applications that are tailored to the Swiss context in a later stage.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	Data contains variables related to field and soil data. Namely, Mineral fertilizer (kg/ha), Slurry (cubic meters per ha), Field tillage, Operations, Plant protection products, Crops and yields, Nmin, multispectral images, Various soil scans, Grazing
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	CSV, Barto-API
Data Schema *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If the schema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.*	We could not have access to the schema from Barto yet, but this is possible at a later stage

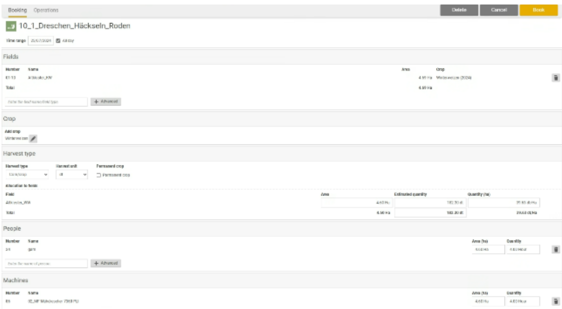
Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	Most relevant data are available. A clear concept is needed, which indicators need to be calculated, to be able to determine how data can be accessed.
Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	For Barto: data are entered manually on smart phones throughout the year. For exatrek: data are logged directly from the ISO-bus of the tractors
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Indvidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Swiss Future Farm, beside personal data, no restrictions are in place regarding production data.
Dataset update frequency *Specify how often the dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	As needed (for Barto), on-navigation (for exatrek) Very few field data need to be updated frequently as they do not change (Executed work, applied products etc.) Weather data are available in 10 min Intervals and satellite images weekly
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Less than 1 GB
Data utility *Identify the primary users and beneficiaries of the dataset, eg.*	Farmers, agronomists, research and academics, data scientists, companies developing tools for farmers.

Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.	If data exists in aggregated form at a country level, then the data can be useful for: Decision and policy makers, analysts and strategists
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal informations), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	NA
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	Barto: Event-driven updates (Depending on the operation) Exatrek: Automatic logging once the ignition is turned on in the tractor
Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	Stored in the cloud of the suppliers
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	All links are provided above. For Batro specific data sets, we would need to check again with Barto
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	Same as Link to Dataset
Metadata *Provide detailed information about the dataset's structure, content and context helping other stakeholders to*	The below dataset is provided through the Barto software. This software contains a lot of modules related to farm management. An extensive list can be found here: Tous les modules numériques | barto et 365FarmNet | Barto

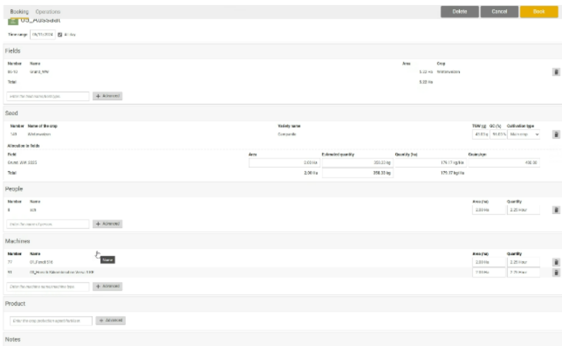
understand its purpose, component and characteristics.	The snippets presented below are screenshots from the interface itself. Additionally, we also have the exatrek software that stores data related to agricultural machines (tractors) such as: diesel usage, speed, location and navigation, RPMs, activity, … We also show a snippet of these data below.
Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	Snippet of Mineral and Slurry Fertilizers: - These data exist by field (~40-50 depending on the season)  - Field Tillage and further work processes:  - An example of the field tillage:  - An example for the Plant protection:



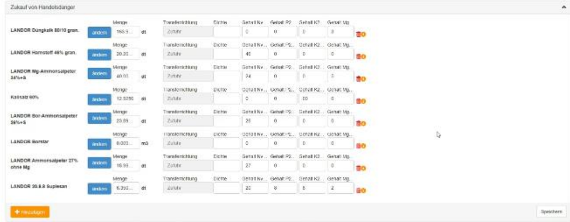
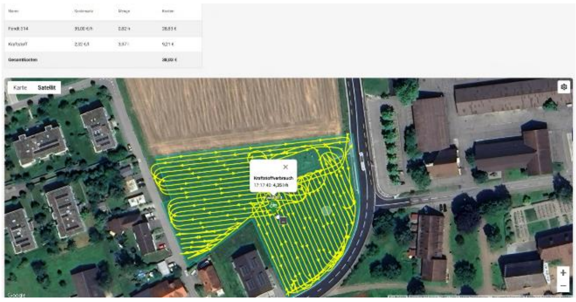
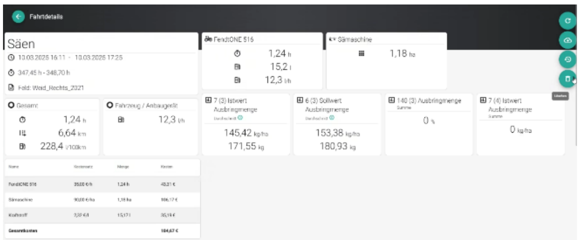
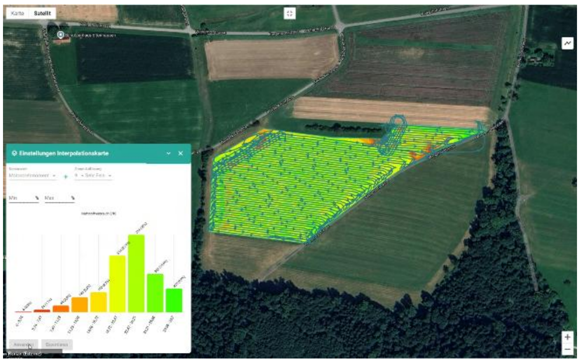
- Crops and yield example:



- Mulch Sowing:



- The following measures are sporadically available. In fact, depending on the experiment, we can have them: Nmin, Multispectral images, various soil scans.
- Mineral Fertilizers Used on the Farm level:

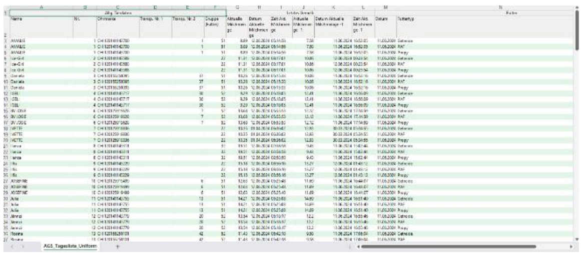
	 (unfortunately, it's in German). - Snippets from exatrek   
Do you intent to apply the FAIR principles to your dataset ? *Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available [here](#).*	We do not know how to answer this as it's not clear whether these data are proprietary for SFF of both the software provider and the farm only

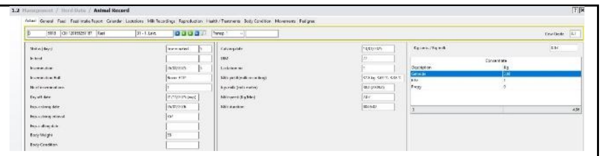
Demonstration site in Switzerland: Dataset 2 (Dairy data)

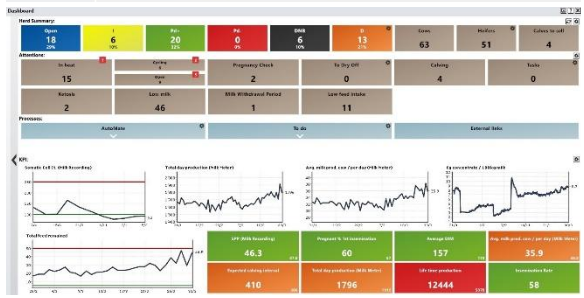
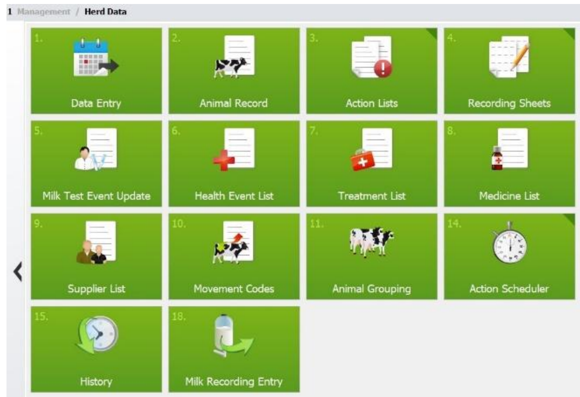
SFF Diary Farm Data	
Name of the dataset	SFF Dairy Farm Data
Pilot site	Swiss Future Farm (WBF – Agroscope)
Accessibility *Define the level of access permitted for this dataset. Possible categories include Open Access, Restricted Access, Confidential, Internal Use Only, Licensed Acess etc.*	Through Uniform Software, Dynamica generale (mixed feed amounts), Smaxtex (data are acquired and stored but not validated)
Purpose and relation to the objectives of the project *Explain the purpose and relation of this dataset for this project so that all stakeholders understand the contribution of this dataset to project.*	These datasets extend beyond our immediate 4 core modules to establish foundations for future Dairy and Animal Farm monitoring capabilities. A part can be used to the 4 cores modules (such as local weather data), but they reflect our vision of project's evolution, particularly as we prepare compelling animal behavior monitoring proposals for the upcoming cascade call that will leverage and expand upon our established technological framework.
Data types *Describes the primary nature and category of data contained within this dataset, eg. Geospatial data, IoT data, Demographic data, Soil and environment data, Field data etc.*	**Uniform:** The dataset contains information about the dairy cattle. Namely, the following variables are present: Concentrated feed (in Kg), Concentrate feed intake of the cows, Milk production (kg) + Protein and Fat percentage, Animal weight (kg), Age of the cattle, reproduction data (calving, insemination, nb of calves, gestation, pregnancy and conception rate, reproduction status), lactation (SPP value, days in milk), animal count (by type), heath data. **Dynamica Generale:** Per group of cows: Roughage and Composition of roughage (from the mixer) **Calculated:** Farmyard manure quantity (using swiss standards) **Smaxtex:**

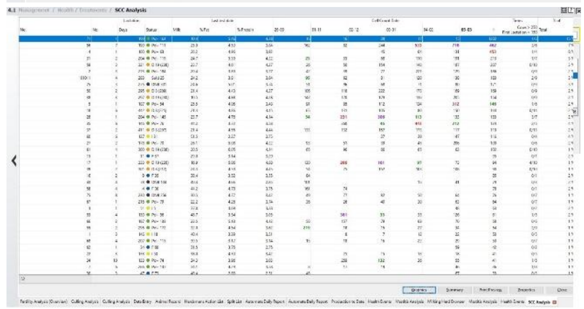
	Water consumption dairy cattle barn (MVS) (Calculated) Activity from an accelerometer inside a pill inside a cow Temperature Localization (feeding area, lying area, ...) **Measured:** Electricity consumption dairy barn, Water consumption EVS m3, Electricity consumption Emission test barn (EVS) **Weather Station:** Temperature, humidity, wind speed **Barto** Identification of the animals in the barn according to the Swiss animal register **Swiss Herdbook** Data of the monthly milk analyses
Data formats *Specifies the file format(s) used to store the dataset, eg. CSV, XLSX, JSON, Shapefile, GeoJSON, XML, Plain text, SQL etc.*	CSV or ADS
Data Schema *Provide the data schema or visual representation of the dataset's organization, including tables, fields, relationship and data types (if available). If shema is large or complex attach to the end of document or provide a link to external file. Diagrams can be created using tools like draw.io, Microsoft Visio etc.*	Basically, the data schema reflects timeline activities related to cows on the farm (E.g., what happens with cow X at a time T). We give more details in the snipet section below.
Re-use of existing data *Explains the potential for re-using this dataset in various analytical, statistical, or operational contexts, eg. Data is compatible with other GIS systems or data is structured and robust to support trend analysis, predictive modeling, machine learning etc.*	Most relevant data are available from the milking robot (De Laval) being extracted by Uniform. On the other side, data can be extracted from the Swiss Herdbook and Swiss Animal Register

Data origin *Provide details or origins of how the data is obtained/created, eg. Data is created as part of current project, data was sourced from existing dataset, data is publicly available, collaborative data (collected with other stakeholders), derived data (transformed of derived from other sources, like aggregations, calculations etc).*	Mostly logged from the milking system (DeLaval and Fullwood milking systems) For other variable, refer to the data types
Data owner *Explain who is the owner of the data, which is critical for data governance, eg. Invidual, organization or company, government or public institution, collaborative ownership, open data ownership etc.*	Swiss Future Farm
Dataset update frequency *Specify how often the dataset is refreshed or updated, eg. Real-time, daily, weekly, monthly, one-time collection, as needed etc.*	Usually, daily basis. Maximum frequency of extraction is 8 times a day. Weight: on weighing
Expected size of the data *Estimate of the dataset size, eg. Up to 10 MB, around 100 GB etc.*	Up to 100 MB per daily export
Data utility *Identify the primary users and beneficiaries of the dataset, eg. Farmers and agronomists, decision and policy makers, research and academics, business analysts and strategists, data scientists and analysts etc.*	Farmers, agronomists, research and academics, data scientists, companies developing tools for farmers. If data exists in aggregated form at a country level, then the data can be useful for: Decision and policy makers, analysts and strategists
Privacy principles *Explain the measures taken to protect the privacy of individuals within the datasets, eg. Anonymization (removed all personal information), aggregation (summaries or averages at a group level rather than individual records), data minimization (only essential data is collected and stored), encryption, access control etc.*	Not applied because there is no information about people involved.
Data release roadmap *Overview of the dataset's planned updates, eg. Regular additions, annual expansion (updated once per year every January), event-driven updates, one-time update etc.*	Uniform: daily or 8 times a day Dynamica: syncs twice a day with cloud only if there is an event (data to be transferred – for example, when they weigh something)

Repository *Location or platform where the dataset is stored, eg. Online Data portal, institutial repository, Cloud storage services, Github or Gitlab, Local server od database, public data archive etc.*	Stored locally on the SFF storage because they have the desktop version of the software
Link to dataset *If the dataset is available online, provide a direct link to its repository or download page.*	No link
Data access *Explain the methods available for accessing the dataset, eg. Web access, REST API, FTP access, downloadable apckage, database query etc.*	Currently, the only way to access the data is to look at the automatically exported CSV files. However, the account manager told us that they will provide an API access soon (date undetermined)
Metadata *Provide detailed information about the dataset's structure, content and context helping other stakeholders to understand its purpose, component and characteristics.*	
Snippet of the data *Provides a small sample or example of the dataset's content, giving the stakeholders a preview of its structure, fields and data types.*	All the data can be exported as CSV file with the below shown format (unique format for any extraction):  This picture shows the cows by name, number, ID and group, and some other information such as milk yield. Although we show above what are the variables that can be viewed, we show a few screenshots of the UNIFORM software used at the farm to give a more granular idea on what is possible and how our staff access them:

<table>
<tr>
<td></td>
<td>



</td>
</tr>
<tr>
<td>Do you intent to apply the FAIR principles to your dataset ?

Explain whether the dataset adheres to the FAIR principles – Findability, Accessibility, Interoperability and Reusability. More explanation about FAIR principles is available <u>here</u>.</td>
<td>We do not know how to answer this as it's not clear whether these data are proprietary of both the software provider and the farm only</td>
</tr>
</table>