D1.6 FINAL DATA MANAGEMENT PLAN

Project: Monitoring of Environmental Practices for Sustainable Agriculture Supported by Earth Observation

Acronym: ENVISION

Document Information

Grant Agreement Number	869366	**Acronym**	ENVISION
Full Title	Monitoring of Environmental Practices for Sustainable Agriculture Supported by Earth Observation		
Start Date	1st September 2020	**Duration**	36 months
Project URL	https://envision-h2020.eu/		
Deliverable	D1.6 Final Data Management Plan		
Work Package	WP1: Management		
Date of Delivery	**Contractual**	M32	**Actual**	M32
Nature	ORDP	**Dissemination Level**	Public
Lead Beneficiary	DRAXIS		
Responsible Author	Stavros Tekes (DRAXIS), Anastasios Karakostas (DRAXIS)		
Contributions from	AGROAPPS, NOA, EV ILVO, ITC, ETAM		

Document History

Version	Issue Date	Stage	Description	Contributor
D0.1	12/4/2023	Draft	Draft for review	DRAXIS
D0.2	26/4/2023	Draft	Review feedback	AgroApps
F1.0	02/5/2023	Final	Final version	DRAXIS

Disclaimer

This document and its content reflect only the author's view, therefore the EASME is not responsible for any use that may be made of the information it contains!

CONTENT

Executive Summary

The purpose of the current is to present the third, and final, version of the project Data Management Plan (DMP).

This version lists the various datasets that have been collected, processed and/ or produced by the ENVISION project and presents the main data sharing and the major management principles that have been followed. Furthermore, it is incorporated all the critical changes, such as any external factors that might had impact on the data management within the project and might affect it, even after the project lifetime.

The deliverable is structured in the following chapters:

Chapter 1: Introduction – Includes an introduction to the deliverable

Chapter 2: DMP Components in ENVISION – Includes a description of the datasets along with the documented changes and additional information.

1   Introduction

The Deliverable D1.6 Final Data Management Plan represents the third version of the DMP of the ENVISION project. ENVISION is an Innovation Action project funded under the H2020 programme of the EC that will last 36 months.

The current deliverable follows the FAIR template, and the various identified datasets are analysed, while answers are provided about how the data have been managed within the project and is described how it are provided as much open and re-usable data as possible from the execution with the purpose of facilitating to others the re-use of such data.

Each dataset is defined, modified, and described, and information is provided on to which extent it is standard compliant, and how the data are available, accessible, interoperable and re-usable, and corresponding procedures for the preservation and the data management.

The Work Packages that have not occurred any changes are also included in this deliverable.

2 DMP Components in ENVISION

2.1 DMP Components in WP1 – Management (DRXS)

DMP Component	Deliverable Title
Data Summary	Contact details of the project partners. Databases containing all the necessary information regarding the project partners. The project partners data are stored on a simple table (excel file) and it is stored on the ENVISION dropbox folder, with the following fields: • Organisation • Name • Email Furthermore, consortium meetings have been conducted remotely every month (first Tuesday of the month) to discuss the project progress and address any important issue. Most of the meetings have been conducted using Google Meet. Meetings have been prepared after each meeting and are stored on the ENVISION dropbox folder (docx. format). Furthermore, an excel file has been created (ENVISION_Actionlist), including the following fields: • Event/ Source • Relation to WP • Description • Owner • Contributor • Deadline • Status (Done, On Progress, Delayed) • Comments • History The expected size of the docx. is not applicable. Moreover, WP leaders have sent input on how they handle and process the data produced/generated and/ or collected during the project. Presentations, agenda, and the participants list of each plenary meeting or review meeting have been collected and kept. Interviews with new Advisory Board (AB) members have been contacted using Google Meet and the recordings have been kept. Lastly, two project events have been held; the first one was the 1st review meeting, and the second one was the project meeting. The former was held online using the WebEx application, and the latter physically in Athens. The material of these meetings (agenda, presentations, recordings) is stored on the project's dropbox folder and on coordinator's internal server.
Making data findable, including	The data with regards to the remote meetings as well as the

provisions for metadata	plenary and review meetings are stored on DRXS server and in the ENVISION dropbox folder. The data are not directly accessible from outside. Moreover, these data will not be made available to third parties.
	However, input provided with regards to the data management, as well as LHCs and AB members are available through the respective deliverables (D1.3 Initial Data Management Plan, D1.4 Intermediate Data Management Plan, D1.5 Intermediate report on Lighthouse Customers and Advisory Board feedback and actions taken, D1.6 Final Data Management Plan, as well as D1.7 Final report on Lighthouse Customers and Advisory Board feedback and actions taken). The dissemination level of these deliverables is public, and they are available in the project's website (https://envision-h2020.eu/), dropbox folder and in Zenodo[1] (https://zenodo.org/communities/envision/) through the Digital Object Identifier (DOI):
	• D1.3 Initial Data Management Plan: DOI: https://doi.org/10.5281/zenodo.6121858
	• D1.5 Intermediate report on Lighthouse Customers and Advisory Board feedback and actions taken: DOI: https://doi.org/10.5281/zenodo.6303595
	• D1.7 Final report on Lighthouse Customers and Advisory Board feedback and actions taken: DOI: https://doi.org/10.5281/zenodo.7885912
	The naming conventions used for these data are:
	• Data_WP1_1_Data_Management_Plan
	• Data_WP1_2_Lighthouse_Customers
	• Data_WP1_3_Advisory_Board_Members
	As part of any stored data, metadata have been generated which include sufficient information with appropriate keywords to help external and internal users to locate data and related information.
Making data openly accessible	The datasets are not publicly available.
	All the data are publicly available as part of the aforementioned deliverables and through the ENVISION website, dropbox folder and Zenodo.
Making data interoperable	N/A
Increase data re-use	Data are publicly available as part of the aforementioned deliverables and are accessed and re-used by third parties indefinitely without a license.
Allocation of resources	Resources have been allocated according to the project plan and WP1 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The data are collected for internal use in the project, and not intended for long-term preservation. No personal information will be kept after the end of the project. Furthermore, DRXS

---

	pays special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European, and international framework, and the European Union's General Data Protection Regulation (GDPR) 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.2 DMP Components in WP2 – Commercial Service Requirements (URDG)

DMP Component	Deliverable Title
Data Summary	WP2 gathers the user needs from Paying Agencies (PAs), Certification Bodies (CBs) and other end users of the ENVISION platform and services. This provides the developers with a set of prioritised needs which feed in the tasks described in WP3 and WP4. In addition, WP2 identifies potential problems that may hinder the adoption of ENVISION by PAs and CBs (internal to the organisation weaknesses, and external threats). This is achieved through a series of workshops and surveys with ENVISION stakeholders. To investigate current service provision WP2 has distributed an online questionnaire-based survey through the Qualtrics platform, which has been completed by 36 PAs and 18 CBs. Furthermore, it performed 8 online workshops, 2 per business case (CAPO with 12 participants, LV with 5 participants, OCS with 3 participants, and NPA with 8 participants), together with ENVISION development teams (total of 10 participants); the qualitative information obtained through these workshops was sourced through the online platform Miro and is stored in Pdf and Excel formats. WP2 also engaged with the ENVISION business case and development stakeholders through 2 online workshops (1 for each group) for the weighting and prioritisation of the user needs (User Stories as described also further below).
	WP2 also collects data regarding gender considerations integrating the gender dimension within the project by clearly articulating the gender relevance to all WPs, estimating how research findings apply to the specific needs of all sexes, ensuring data disaggregation on sex and gender related issues and considering all intersecting factors. It also tackles dissemination issues of the project and its results.
	To generate and record data within WP2 regarding the user needs, the online platform Miro has been used and information has been extracted in a pdf form whenever necessary. No sensitive information has ever been requested for any participant through this platform, particularly for attendees outside of the consortium (i.e., farm managers, advisory board members). The data have the form of "user stories" which provide information in respect to the characteristics of specific roles about of the end users (IT experts, administrator, coordinator, inspector, etc.), their needs associated to the services provided by the ENVISION platform and a description of why they need it. In addition, online workshops and co-production meetings taking place during tasks 2.2 and 2.3 of WP2 have been recorded. WP2 is monitoring the coproduction process and collecting feedback throughout. Specifically, 2 hybrid workshops (remote and in-person participants) have been performed in the first in-

person consortium meeting (Thessaloniki, May 2022) to reflect and evaluate progress in relation to the established User Stories, with 24 participants in total recording qualitative and quantitative responses via the Mentimeter platform and through open-ended discussions. In addition, WP2 runs a monthly questionnaire-based survey regarding feedback on coproduction activities, which all ENVISION stakeholders are expected to complete each month (either individually or through representatives of stakeholder groups). The survey collects quantitative information (i.e., Likert scale scores) and has collected 30 responses thus far. In the second in-person consortium meeting (Athens, November 2023), WP2 performed a Delphi consensus survey that consisted of 2 online questionnaires and a hybrid workshop. A total of 27 participants recorded their qualitative and quantitative responses via Qualtrics and open-ended discussions (i.e., information stored in Word and Excel documents). The aim of this Delphi approach was to identify challenges and disruptive factors in ENVISION coproduction and ways to improve future coproduction frameworks.

WP2 has also a series of planned activities for the collection of additional feedback on coproduction and to supplement the ENVISION development process with further insights of stakeholders external to the consortium. Specifically, the Qualtrics survey that explored current service provision was circulated to all Panta Rhei 2023 PA participants, and 2 coproduction workshops will be performed to explore ways to improve the ENVISION mobile application features and functionalities from the perspective of external to ENVISION software developers (5 developers in May 2023) and farmers (10 CAPO affiliated farmers in June 2023). To further reflect on coproduction and conclude on progress in relation to User Stories, WP2 will hold a series of workshops in association with WP5 (business case evaluation) with each business case including LEAF, in September 2023.

Any recordings of WP2 coproduction related activities, are stored initially on the Microsoft Streams service that is provided by the University of Reading and is accessible only by ENVISION – WP2 team members adhering to all data protection and ethical guidelines that the University of Reading's policies propose. Then they are uploaded to the relevant ENVISION Dropbox folder (and removed from Microsoft Streams), where they will remain for the project lifetime before they will be removed entirely. The videos serve the purpose of the detailed capture of information and in addition these offer the chance to revisit the workshops and summarise the outcomes. The data are extracted from end-users participating in an online survey and to a series of online workshops and consultations and are summarised into several excel spreadsheets that contain information relating to the

four business cases. The size of the video recording will be relevant to the duration of the meeting. These data are useful to software developers and researchers who wish to understand the "needs of key stakeholders" that form the core of the development process of platforms and tools for remote monitoring of environmental activities.

The data regarding the task 2.5 will be collected by all partners using excel files. As these personal data will include sensitive information, attention will be given so that no persona might be identified once processed. Additionally, to safeguard the privacy of each partner within their organisations and consortium, this information will be sent directly to the responsible processor, who is the Project's Ethic Manager and the Information Security Officer of her organisation. The data will be used to produce statistics and no other information will be included and the excel files will not be preserved.

Finally, WP2 has conducted a survey to investigate farmer awareness and engagement to coproduction activities. The survey was constructed on Qualtrics XM, which will also be the main portal for distribution, through a secure link that was provided to Paying Agencies (PAs) and Certification Bodies (CBs) across Europe. The PAs and CBs distributed the survey link to contacts (farmers) through their applicants/clients lists, following the same randomised approach. No sensitive and identifiable information was requested through the survey and WP2 did not have access to the applicant/client list of contacts for any of the organisations involved. If necessary, PAs/CBs will be allowed to translate, and transpose and share the survey through other means (i.e. Google forms) but for any such step they will be required to provide a detailed description to WP2 of the methodological steps followed, people involved, and actions taken to ensure that the identity of participants is not compromised. The outcomes of this survey will be stored as Excel files and will be processed in SPSS and R. We have the ability to encrypt and password-protect any relevant file, although this will not be necessary since they will not contain any sensitive information for any of the parties involved. To date, 118 responses have been collected from farmers of 7 different countries (unbalanced sample from UK, Greece, Cyprus, Belgium, Lithuania, Serbia, Bosnia and Herzegovina).

Making data findable, including provisions for metadata	The data produced either via the online survey or the workshops with the end-users (video recordings included) will not be identifiable to the individual respondent. However, there are metadata that allow the institution to be identified. These data are stored on URDG's servers. Regarding the data for tasks 2.5, no information will be available to third parties directly and the provided excel files by the partners will be deleted after the process of the

	information that will be included in the respective deliverable. The outcome of the aforementioned information is available in D2.1 Review of current services provision, D2.2 Report of customer requirements from ENVISION services, D2.3 Gender Situation Analysis and Needs Assessment, D2.6 Draft_Report of co-production of ENVISION services and D2.7 Report of co-production of ENVISION services. The dissemination level of these deliverable is public and they are available in the project's website, dropbox folder and in Zenodo through the DOI: • D2.1 Review of current service provision: DOI: https://doi.org/10.5281/zenodo.4564201 • D2.2 Report of customer requirements from ENVISION services: DOI: https://doi.org/10.5281/zenodo.4564653 • D2.3 Gender Situation Analysis and Needs Assessment: DOI: https://doi.org/10.5281/zenodo.4564344 • D2.4 Guidelines on legal and ethical issues: DOI: https://doi.org/10.5281/zenodo.6121934 • D2.5 Privacy Risk Assessment for ENVISION: DOI: https://doi.org/10.5281/zenodo.6122040 • D2.6 Draft Report of co-production of ENVISION services: DOI: https://doi.org/10.5281/zenodo.7885893 The naming conventions used for these data are: • Data_WP2_1_PA_survey • Data_WP2_2_User_needs • Data_WP2_3_Co-production • Data_WP2_4_Gender_Considerations • Data_WP2_5_Farmers_survey A thematic analysis will be conducted to identify suitable keywords that could allow future scenarios and metadata will include the date of data collection, the source (interview/workshop/video recording), the organisation and role of individual in the organisation, and contact information (whenever appropriate/ available i.e. in the case of ENVISION partners).
Making data openly accessible	The datasets are not publicly available. All the data are made publicly available as part of the aforementioned deliverables and through ENVISION website, dropbox folder and Zenodo. Furthermore, the data regarding user needs are available through scientific publications in a summarised form.
Making data interoperable	The data are a series of user needs and statistics that have been and will be analysed and reported in a format that will be easy to share between stakeholders and be interpreted by any party.
Increase data re-use	The data with regards to gender considerations will not be

	available for re-use. The data regarding the user needs will be available once the papers are published. Some of the initial data have been reported in an early publication relating to user needs (2021/11) and the co-production methodology will be made available in a publication after the end of the project (2022/23). Most of the data in WP2 is project-specific but the co-production methodology will be useable for other third parties after the end of the project and the data will remain re-usable as long as it is useful. The questionnaires and workshops as well as the excel files regarding gender considerations task, received ethical clearance and the data were double checked for prevention of errors.
Allocation of resources	Resources have been allocated according to the project plan and WP2 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The data have been collected for internal use in the project, and not intended for long-term preservation. No personal information will be kept after the end of the project. Furthermore, UREAD pays special attention to security and respect the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679. More specifically data regarding gender considerations were collected in an anonymized manner. No crude data were presented in the deliverable, only the results of statistical analysis and as such cannot lead to the identification of a person. Lastly, ETAM has appointed a specific Information Security Officer and they implement an accredited Information Security Management System according to EN ISO 27001:2013.
Ethical aspects	The user need methodology has received ethical clearance from the School of Agriculture, Policy and Development (University of Reading) to ensure that the participants are protected. In addition, the collection of the gender and sex related issues has also received ethical clearance from ETAM and an Information Security Officer has been appointed since the data collected are considered sensitive data. Therefore, the data raise no legal issues.
Other issues	N/A

2.3 Earth Observation data products (NOA)

2.3.1 Task 3.2 Auxiliary data collection (Lead Partner: DRXS, Contributors: NOA, NPA, LV, CAPO, OCS, EV ILVO, INOS, AgroApps)

DMP Component	Deliverable Title
Data Summary	The initial data that have been collected in WP3, cover the needs to implement/ develop and calibrate the initial desired products in all business cases. In order to collect the raw data from the business users (raster and vector format), an SFTP service was proposed ensuring end-to-end file transfer encryption, since it follows the SSH protocol. Moreover, data have been collected through the ENVISION platform, during the business cases' implementation phase. The origin of the data for WP3, are from: • Farmers' declarations, along with access to the Land Parcel Identification System (LPIS), and VHR imagery has been provided by the PAs, • Farmers' declarations, along with access to the Geoserbjia has been provided by the CB, • Auxiliary data concern farmers' personal and farm information and shapefiles containing farm location, • Laboratory results of soil. Data products assist to calibrate, validate and feed ENVISION's processing algorithms to attain the highest information quality possible for improved remote monitoring and decision-making services to PAs, CBs, and other users. Auxiliary data include: • LPIS • Farmers' past declarations • Remote sensing results • Layers • Laboratory results Data gathered are approximately 20 GB.
Making data findable, including provisions for metadata	The auxiliary data are available to all technical partners. Data are stored in a file under the responsibility of the data controllers/ handlers/ processors (NOA, EV ILVO, AgroApps) and labelled with the work package, country of origin and the type of data. The naming convention used will be: Data_WP3_1_Country_of_Origin_Type_of_data
Making data openly accessible	The data will be kept closed until the end of the project due to handling of personal data and will not be allowed and disclosed to be used by any third party. Anonymized and summarized data will be available in any public deliverable (D3.2 Catalogue on auxiliary data and available repositories to be incorporated, D3.3 Data Products initial report, D3.4 Data products validation report) or through any other relevant publications relating to dissemination and exploitation

	purposes. The dissemination level of these deliverable is public, and they are available in the project's website, dropbox folder and in Zenodo through the DOI:
	- D3.2 Catalogue on auxiliary data and available repositories to be incorporated: DOI: 10.5281/zenodo.6121987 - D3.3 Data products initial report: DOI: https://doi.org/ 10.5281/zenodo.6303627 - D3.4 Data products validation report: DOI: https://doi.org/10.5281/zenodo.6320349 The raw data that have been provided, to the technical team will not be publicly available to third parties and will be provided under a confidentiality agreement (CA) between the involved partners.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	Resources have been allocated according to the project plan and WP3 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The data have been collected for internal use in the project, and not intended for long-term preservation. No personal information will be kept after the end of the project. Furthermore, NOA, EV ILVO, AgroApps pay special attention to security and respect the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.3.2 Task 3.3 Analytics on Vegetation and Soil Index Time-Series, Task 3.4 Cultivated crop type maps, Task 3.5 Grassland mowing events detection (Lead Partner: NOA)

DMP Component	Deliverable Title
Data Summary	The tasks undertaken in this project have generated various products that cater to the implementation and validation of all the desired products across different business cases. All the collected data are available in both raster and vector formats, and they are used as inputs for WP4. **EO data** We have gathered multi-year time-series of Sentinel-1 and Sentinel-2 images for nationwide coverage of Lithuania and Cyprus. The raw optical spectral bands (B02-B12) have been used for the reconstruction of other remote sensing indices such as NDVI, NDWI, NDMI, SAVI, PSRI, and NBR. SAR backscatter estimation has also been carried out. **Vector data** We have used multi-year LPIS data for Cyprus and Lithuania, which were provided by the respective PAs and project partners. They have been used to train and validate the

machine learning (ML) algorithms. The vector data are available in a shapefile (.shp) format, representing a collection of polygons depicting parcel boundaries.

Raster data

Apart from EO data, we have transformed the aforementioned vector data into a raster format and imported them into the datacube to enhance the implemented algorithms. An ids product has been processed and generated for eight different buffer parameters (-5m, -3m, -1m, 0m, +3m, +5m, +10m). Assuming two business cases, the following estimation is provided:

- For Lithuania: 14 tiles * ~20 MB per tile * 8 different buffers * 2 years ≈ 4.5 GB
- For Cyprus: 2 tiles * ~20 MB per tile * 8 different buffers * 2 years ≈ 640 MB

Size of input EO data

The pre-processing of Sentinel-2 images generated images from 10 MB to 250 MB for each band, while the generation of Sentinel-1 backscatter produces products with sizes ranging from 1.1 GB to 3 GB. Assuming two business cases, the following estimation is provided:

For Lithuania:

- Sentinel-1 (Backscatter): 4 orbits * ~230 images per year * 2 years * ~1.1 GB per image ≈ 2 TB
- Sentinel-2: 14 tiles * ~180 GB per tile (depending on the orbit) per year * 2 years ≈ 5 TB

For Cyprus:

- Sentinel-1 (Backscatter): 2 orbit * 5 images per month * 12 months * 2 years * 1.1 GB ≈ 260 GB
- Sentinel-2: 2 tiles * ~175 GB per tile (depending on the orbit) per year * 2 years ≈ 700 GB

This estimation represents the worst-case scenario for the two years processed. In many cases, the optical images are covered by clouds or sea, so we crop out clouds and other unnecessary parts of the images to reduce the real image size.

Initial development

The data/products generated in these tasks provide Earth Observation (EO) products with crucial information on:

- Analytics on Vegetation and Soil Index Time-series: Producing indicators for vegetation status and health, stubbles burning identification, harvest detection, natura2000 illegal activity detection, phenological metrics, soil exposure and run-off risk assessment.
- Cultivated crop type maps.
- Grassland mowing detection product.
- Datacube API

These products have been used in each pilot implementation. The data and products assist in calibrating, validating, and feeding ENVISION's processing algorithms to achieve the

	highest information quality possible for improved remote monitoring and decision-making
Making data findable, including provisions for metadata	To improve the findability of data within the ENVISION platform, the following actions have been taken: **Training data:** The development team has imported the training data to the ENVISION platform's server, along with related metadata that describes the data structure and methodology used to collect them. To ensure that the data are easily discoverable, a comprehensive search feature has been implemented on the platform, enabling users to search for specific data sets based on keywords, categories, and other metadata. **User Input Data:** To ensure that users' input data are compatible with the platform, it needs to comply with the field requirements of the platform. This includes vector multi-polygon files in .shp format with valid geometry and a compatible projection system. To facilitate the uploading and querying of user data, clear instructions on the accepted formats and requirements are provided on the platform. **Raw Satellite Data:** Raw satellite data used for feature extraction are stored on the platform's operational database, accompanied by the relevant metadata following the original name conventions. However, it is not available and accessible to partners and hence are not open for reuse. To enable the discovery of these data, the relevant metadata are included in the search function. **Output Data:** The output data are accessible only to the registered partners who made the request and are available as layers via a Geoserver's web mapping service (WMS). To ensure that the output data are discoverable, clear instructions on how to access the data are provided on the platform. **INSPIRE Metadata:** To improve the discoverability of EO-based geospatial products, INSPIRE metadata are created for all the data that have been generated in the lifetime of the project. These metadata include information on the data's origin, structure, and methodology, making it easier for users to search for and discover relevant data sets. **Data Storage:** All data are stored on the CreoDIAS servers and labelled with the work package, country of origin, and the type of data. To facilitate the discovery of specific data sets, a unique identifier is assigned to each EO data. For the added value products, a unique identification separated by an underscore (_) is appended to the file name. Versioning identifiers are also appended at the end of the product's name, starting with 0 (_v00, v01, .., vnn).
Making data openly accessible	Access to data is restricted to registered users who have made a query (PAs), meaning only the partners who have requested specific data have granted access to the resulting products. Collected imagery and extracted features, as well as

	generated indices, are not available to the rest of the partners or users, and only the development and technical teams have access to them. Regarding user input data, apart from the registered user, only the development and technical teams have access to these data. To access the output data, registered users only need a web browser and internet access. The data and products are made accessible through an API on top of a Postgres database for parcel-based results, or on top of the Data Cube for pixel-based results, via a web framework such as Django. Spectral Indices and EO-based classification objects are made available to users. The data can be accessed using standard tools and does not require any special software. A user can create scripts to access and query the database and retrieve relevant datasets. All data and associated metadata are deposited in CreoDIAS Virtual Machines, which are highly secure and reliable. The data are available through standard protocols, such as RESTful APIs and web mapping services, ensuring easy and flexible access to the data. The metadata for all geospatial products that are generated during the project's lifetime are created using INSPIRE standards. This ensures that the data are properly documented, and can be easily discovered and used by interested parties.
Making data interoperable	Interoperability is a crucial aspect of managing EO information, and it can be achieved by leveraging widely adopted geospatial standards. In this project, PostGIS, Geoserver, and Open Data Cube open-source tools will be used for the management of EO information, ensuring accessibility to a wider audience. To enable discovery/view services, OGC WMS and WFS services have been implemented. This allows users to access and visualize the output data, which are available in either GeoTIFF or JSON format, along with associated metadata. Users can access these data through either the GeoServer or a RESTful API or both. For process services, the tasks rely on the ODC processing API, which are blended with various other Python scripts to provide Analysis Ready Data. To ensure interoperability, the INSPIRE protocol is used for metadata descriptors, providing a standard for geospatial data. This ensures that the data and metadata are discoverable, accessible, and usable by a wider community.
Increase data re-use	Appropriate licensing agreement will be required for data access after the project's conclusion, which will be defined through the business model during the course of the project. The EO-based products will be usable by third parties through RESTful API, but only for those parties who are part of the project and during the lifespan of the project.

	No particular data quality assurance process is followed, and no relevant warranties will be provided.
Allocation of resources	Resources have been allocated according to the project plan and WP3 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	CreoDIAS servers are managed by the IT department. They are regularly backed up and secured. All servers are hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction enforces the secure storage of data. Furthermore, CreoDIAS pays special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679. The CREODIAS Platform cloud security relies on OpenStack's centralized authentication and authorization model managed by the OpenStack Identity Service (Keystone). Keystone manages Tenants (Environments), Projects, Users, user Roles, service Catalogues and service access Policies. Every cloud management operation (such as mounting a volume or accessing object storage) performed by a User or an application through the Dashboard or through the API must first be checked for validity with Keystone. The Keystone security model is further described in Keystone Architecture.
Ethical aspects	N/A
Other issues	N/A

2.3.3   Task 3.6 Soil condition monitoring (Lead Partner: EV ILVO)

DMP Component	Deliverable Title
Data Summary	EV ILVO used free available on the cloud Sentinel 2 image collections governing the region of Flanders. Afterwards, we assessed indices like vegetation, bare soil and soil moisture and applied them with a cloud masking to create a cloudless bare soil collection for agricultural areas. EV ILVO used the cloudless bare soil collection to select soil sampling locations for the soil campaign. EV ILVO collected 171 soil samples within the 1 Quarter of 2021 and performed SOC lab measurements to assess the top-soil organic carbon values and other soil parameters. After that, EV ILVO built a SOC prediction model using reflections from all bands of Sentinel 2 coming from the Flemish cloudless bare soil collection and the results of the SOC measurements as input and output data. Additionally, we included an extra input parameter, which represents the soil association using a map of the soil associations in Flanders (available open data set). EV ILVO applied this model to the cloudless bare soil collection

+ soil association map and created a SOC layer (% of SOC) of the Flanders Region. EV ILVO delivers this layer in GeoTIFF format, with a spatial resolution of 10m and the average top Soil Organic Carbon (% of SOC) for each agricultural parcel in Flanders for the declaration period of 2021. The agricultural parcels come from the Flemish LPIS provided by the Agricultural Department of Flanders (LV) and exist as open-source data.

For the whole process, we make use of the following data sets:

- Soil associations in Flanders (Bodemassociaties dataset) https://www.geopunt.be/catalogus/datasetfolder/c4f51b28-51bf-4189-8e98-72b94ae8da13
- Sentinel 2 L2[2] with cloud masks[3] from 25 May 2018 until 25 May 2022.
- ESA world land cover 10 m.[4]
- Lab measurements result from 171 collected soil samples.
- LPIS data 2021.

Furthermore, the outputs (SOC products) are:

- Data Product A: A raster file (GeoTIFF) with a pixel spatial resolution of 10 m by 10 m contains top-soil Soil Organic Carbon estimations (% of SOC) divided into organic carbon classes. The SOC GeoTIFF raster is about 1.6 GB. This file is provided and presented in the Envision Platform as a background layer map. The data product is expected to be updated every five years to follow the frequency of the MRV systems. However, we will update the product to cover the submission period of 2022.
- Modelling Metadata that present and evaluate the accuracy of the SOC modelling by using the RMSE (Root Mean Square Errors)[5] and graphs that illustrate the accuracy per input parameter.
- A detailed description of the protocol we followed to collect the sample data and perform the lab measurements.
- Data Product B: A vector file (shapefile format) with the Flemish LPIS 2021 agricultural parcels and a field containing the average top-soil Soil Organic Carbon value. The parcels aggregate the Top Soil Organic Carbon Information using the average value, including pixels coming from the raster file. Parcels without sufficient coverage are excluded. The vector file will

---

[2] https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR#description
[3] https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks
[4] https://esa-worldcover.org/en
[5] Expected for the calibration RMSEC, cross-validation RMSECV and prediction set RMSEP. RPD and R2 are also used to evaluate the accurancy of the model.

	be published in the Envision platform as a vector layer. The estimated total size of the SOC vector product is 100 MB.
Making data findable, including provisions for metadata	EV ILVO delivers the SOC products to the ENVISION repository. - The raster file has the following naming description: TopSoilSOCmap_Region_StartPeriod_EndPeriod_ModelDataSet.geotiff - The metadata description exists in a pdf file and contains the range of the S2 satellite data set, the model data set, the model accuracy, the methodology, the data owner, and the usage rights. - A shapefile with agricultural parcels. The SOC products are published through the ENVISION platform using web services. EV ILVO will deliver both Data products A and B at ZENODO, together with the modelling metadata using the OGC GeoPackage 1.3.1, to support interoperability.
Making data openly accessible	The 171 soil lab measurements and the SOC models belong to EV ILVO (intermediary products). After the end of the project, we will expose the Lab measurements through the DjustConnect platform, giving them the ability to farmers to control their data. The data products produced in the context of the SOC service will be published on the ENVISION platform for use by LV (final product). In parallel, we will post the data products at ZENODO using OGC Geopackage format.
Making data interoperable	As explained, the SOC product will be exposed through the ENVISION platform. The ENVISION platform uses Geoserver, which means we can publicise the data using WMS and WFS services. In parallel, we will post the data at ZENODO using OGC Geopackage format, supporting FAIRness.
Increase data reuse	EV ILVO will give full permission to use and re-use the SOC products to LV. After arrangements, lab measurements of the SOC will be available to third parties. In terms of data quality assurance, there are processes in place that include the modelling process and lab measurements.
Allocation of resources	Resources have been allocated according to the project plan, and WP3 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The ENVISION platform stores the versions of the SOC GeoTIFF, which means EV ILVO inherits the security level of the platform. EV ILVO backs up the data products (A and B) to ensure fast and safe recovery of the EO final products. Internal data management procedures have been applied to the intermediary products. Furthermore, EV ILVO pays special attention to security and respects the privacy and

	confidentiality of the 'user's personal data by fully complying with the applicable national, European and international framework and the European Union's GDPR 2016/679. For the code we are using ILVO Githubs and after the end EV ILVO will provide access to specific users.
Ethical aspects	N/A
Other issues	N/A

2.3.4   Task 3.7 Crop growth Monitoring and identification of organic farming practices (Lead Partner: AgroApps, Contributors: DRXS)

DMP Component	Issues to be addressed

Data Summary	**Task 3.7 Crop growth Monitoring and identification of organic farming practices**
	 Lead Partner: AgroApps Contributors: DRXS
	 Task 3.7 delivers a fully automated **Organic crop identification service**, which aims at identifying whether a particular crop type declared as organic is classified as such, **based on a traffic light system**. The service exploits a number of EO derived indicators, emerging from Sentinel-2 multispectral data, to ensure effective monitoring of the crop condition variability and phenological stages, in both space and time.
	 For the initial development phase of the service (algorithm training and validation) parcels geospatial data and metadata concerning agricultural practices and cropping information have been provided by the Organic Certification body of Serbia. Afterwards, EO data have be extracted for the agricultural parcels and processed to produce raster layers that have been used as predictors in Machine Learning (ML) Classifier algorithms necessary for farming practice identification.
	 These raster layers include valuable crop related information such as Vegetation Indices (VIs), texture analysis features and attributes, resulting after crop phenological analysis. Specifically:
	 • **Normalized Difference Vegetation Index (NDVI):** Mosaicked – Gap Filled with temporal interpolation on 5 day intervals (72 layers per year)
	 • **NDVI 1st Derivative & 2nd Derivative:** Application of Savitzky Golay derivative filter to the NDVI timeseries (72 layers per year)
	 • Phenology Metrics (6 Layers per year – Main Growth Cycle)
	     o Max Positive Gradient - Growth Slope
	     o Starting Date (Sowing)
	     o NDVI Plateu Initialization Date
	     o NDVI Plateu Termination Date
	     o Ending Date (Harvesting)
	     o Max Negative Gradient - Senecense Slope
	 • Crop Growth Stages Duration (5 Layers per year – Main Growth Cycle)
	     o Growth – Concave

- o Growth – Convex
- o Plateu
- o Senecense – Convex
- o Senecense – Concave
- GLCM Texture metrics on NDVI (72 layers per year)
  - o Entropy
  - o Variance
  - o Homogenity

The algorithm development output is stored in the core component of the service. In operational mode, the user defined input data (parcels of interest) are stored in the platform's operational database and utilized for the EO data query from CreoDIAS. After appropriate processing, the output maps (organic or possibly non-organic flag-map) are provided to the Certification Bodies on a WMS.

This task also developed a universal methodology for organic crop yield estimation with the use of EO data, for various types of crops. The methodology that has been followed is based on the assimilation of EO derived VIs into crop growth models that calculates total crop biomass production and crop yield.

EO data

The EO data have been collected by satellite constellations carrying optical and radar instruments; ESA Sentinel-2A/B mission provides multispectral images with an average 5-day revisit. Both optical and radar data have been acquired from the CreoDIAS platform. Atmospherically corrected Sentinel-2 Level-2A images have been retrieved and in cases where Level-2A images were not available, Level-1C have been retrieved and processed with sen2cor algorithm in order to become Level-2A products.
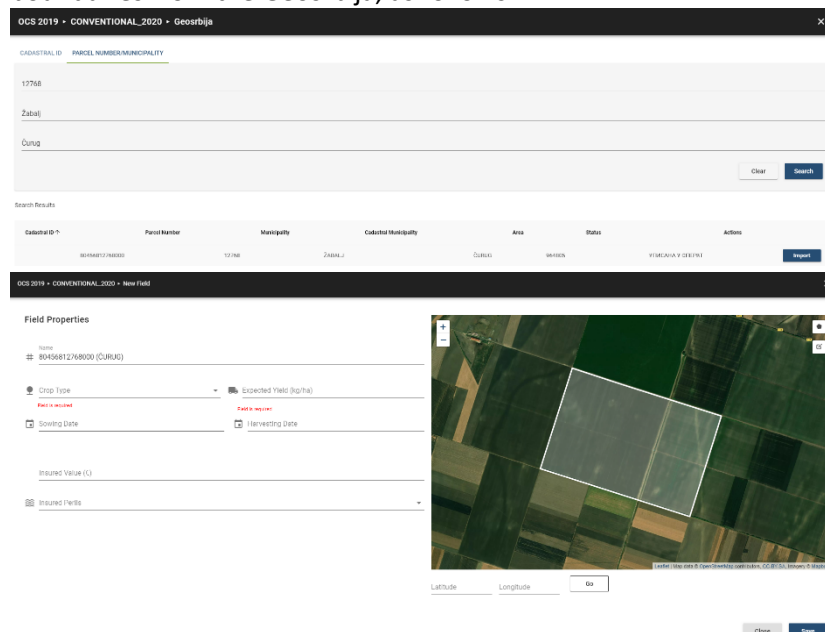
Vector data

The historical data defining the area of interest have been provided by OCS and have been used for training and validating the ML algorithms. For the initial assessment of the service, a number of pre-pilot data have been provided. The provided parcels were archived from 2016 until 2020. The parcel data included the following information; Municipality, Cadastral municipality, Cadastral Number, Area (ha), Crop category, Crop Variety, sowing date, harvesting date, Average yield (t/ha), Status of parcel, Status of crops, Cultivation Year.

Batches of excel sheets contained the required information originating from the Serbian CB, as presented in the following figure:

Municipality	Cadastral muncipality	CADASTRAL No	Area (ha)	Crop category	Crop	Variety	Sowing date	Harvesting date	Average y	Status of parcel	Status of crops	Year
Žabalj	Čurug	12768	62,3105	CEREALS	barley	NS Nonius	10-17.09.2015	June	4.0	organic	organic	2016
Žabalj	Čurug	12768	34,17	CEREALS	barley	NS Nonius	15-19.10.2015	June	4.0	organic	organic	2016
Žabalj	Čurug	9840	33,867	CEREALS	barley	Amorosa	September	June	4,5	organic	organic	2017
Žabalj	Čurug	9840	33,867	CEREALS	barley	Amorosa	September	June	4,5	organic	organic	2017
BELA CRKVA	Bela Crkva	725	5,5426	CEREALS	barley	Sandra	17.10.2016.	June	4,0	organic	organic	2017
BELA CRKVA	Bela Crkva	725	10	CEREALS	barley	Sandra	17.10.2016.	June	1,8	organic	organic	2017
Žabalj	Gospođinci	5658/1	60,2269	CEREALS	barley	NS Nonius	September	June	5,0	organic	organic	2018
Žabalj	Gospođinci	5658/1	60,2269	CEREALS	barley	NS Nonius	September	June	5,0	organic	organic	2018
BELA CRKVA	Dobričevo	795/1	10,806	CEREALS	maize		3381	17.-19.04.2016.	September	3.5	organic	organic	2016

By using the Municipality, Cadastral municipality, and Cadastral Number the polygons in a vector format (shapefile, .shp) was retrieved from Geosrbija, the Serbian LPIS system. Access to Geosrbija was granted by the relative Serbian Ministry for the purpose of the ENVISION project. AgroApps developed the infrastructure to collect the polygons data depicting parcel

boundaries from the Geosrbija, as follows.

A further pre-process of the initial data was applied due to the significant small area and elongated shape of the cultivated parcels. A unification process was performed to neighbouring parcels that have common borders – having always in mind to unify same crops, transition phase, status. The result was to form parcels of larger area that would provide a sufficient number of pixels and assist in achieving a fairly successful discrimination between Organic and Conventional crops. After this unification the total number of parcels was eventually reduced but the average parcel size increased;

A total so far of available parcels is;

- 1429 wheat (776 organic, 653 conventional)
- 1225 maize (172 organic, 1053 conventional)
- 429 Soybean (213 organic, 216 conventional)
- 1055 Sunflower (643 organic, 412 conventional)

At a pre-analysis stage input parcel data were checked for false declarations. There were many cases of wrong crop type declaration, as well as wrong crop practice declaration. Two independent outlier detection methodologies were applied,

- Visual inspection of NDVI timeseries
- Anomaly Detection with SVM classifier using the NDVI predictors.

The conclusions of both methods were combined, and as a result the parcel samples were further reduced to the most trustworthy subset:

- 533 wheat (350 organic, 183 conventional)
- 30 maize (26 organic, 4 conventional)
- 335 Soybean (146 organic, 189 conventional)
- 335 Sunflower (146 organic, 189 conventional)

Size of input data:

One Sentinel-2 MSI L1C/L2 raw image including all bands in .zip format is 600MB. Considering that the project's pilot phase lasted two years and that both summer (maize, sunflower, soybean) and

	winter (wheat) crops have been monitored, satellite imagery were available for 24 months. For Sentinel-2 L2A, BOA: 600MB x 52 days x 8 tiles, equals 248 GB. This is a gross estimation of the requirements during the pilots, presenting the worst-case scenario. The amount of input raw data required for the initial assessment of the service, considering pre-pilot data from 2016, is 496 GB for Sentinel-2.
Making data findable, including provisions for metadata	Training and validation pilot vector data have been imported by the development team and have been hosted at the platform's server. Related metadata describe the data structure and methodology used to collect them. Once uploaded to the platform, only the development and technical teams have access to these data. Regarding the users' input data, those need to comply with the field requirements of the platform for a successful database query; vector multi-polygon files in .shp format with valid geometry and compatible projection system. Raw satellite data that have been used for feature extraction are stored on the platform's operational database accompanied by the relevant metadata following the original name conventions. They are not available and accessible to partners and hence they are not open for reuse. The output will be accessible only to the registered partners who made the request, and it is available as two layers in a WMS; one layer presenting the flag-map and one layer presenting the yield estimations. INSPIRE metadata have been created for all the EO-based geospatial products that will be generated in the lifetime of the project. All data, associated metadata and documentation have been deposited into the web server and are available through RESTful API and Geoserver's web mapping service (WMS).
Making data openly accessible	Only Certification Bodies (CBs) that made the query (registered users) has access to the produced organic/non-organic flag-maps and parcel yield estimations. Collected imagery and the extracted features are not available to rest of the partners or users and only the development and technical teams have access. Regarding the user's input data, apart from the registered user, only the development and technical teams have access to these data. Only web browser and Internet access are needed for the registered users to access the output data.
Making data interoperable	The classification output raster data, on pixel scale, have been saved in GeoTiff format with associated metadata and are not immediately accessible. The classification result, on parcel level scale, has been managed by a PostGIS database and is accessible through GeoServer GIS server application, as a WMS layer. INSPIRE protocol provides the typical standard for geospatial data and it has been used for metadata descriptors.
Increase data re-use	Appropriate licensing agreement is required for data access after the project's conclusion, which has been defined through the business

	model during the course of the project. The EO-based products are usable by third parties through RESTful API, but only for those parties who are part of the project and during the lifespan of the project.
Allocation of resources	Resources have been allocated according to the project plan and how the WP3 resources are allocated. No additional costs are foreseen for making this dataset FAIR.
Data security	All data have been stored on the platform's server and on a separate storage server, both with backup procedures. These servers are managed by the AgroApps IT department. AgroApps fully complies with the applicable national, European, and international framework, and the European Union's General Data Protection Regulation 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.4 DMP Component in WP4 – ENVISION service (DRXS)

2.4.1 System Architecture

DMP Component	Deliverable Title
Data Summary	Functional and non-functional aspects, technical capabilities, components descriptions and dependencies, Application Programming Interface (API) descriptions, information flow diagrams, internal and external interfaces, software and hardware requirements and testing procedures related data specified and validated among the ENVISION technical and business cases partners. Technical requirement reports have been created in order to describe the aforementioned procedures and requirements for all the business cases. These reports have been the basis upon which the system has been developed and further modified.
Making data findable, including provisions for metadata	The reports have been stored in DRXS server and are not directly accessible from outside. However, these data are both discoverable and accessible to the public through the D4.1 Architecture and Services Specifications report, since its level of dissemination is public. The deliverable is accessible via the project's website, dropbox folder and Zenodo: • D4.1 Architecture and Services Specifications report: DOI: https://doi.org/10.5281/zenodo.6121914 The naming convention used is: • Data_WP4_1_System_architecture. As part of any stored data, metadata have been generated, which include sufficient information with appropriate keywords to help external and internal users to locate data.
Making data openly accessible	All data are made publicly available as part of the D4.1 Architecture and Services Specifications report.
Making data interoperable	N/A
Increase data re-use	Data are made publicly available as part of the D4.1 Architecture and Services Specifications report and be re-used by third parties indefinitely without a license.
Allocation of resources	Resources have been allocated according to the project plan and WP4 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The data collected and stored for internal use in the project and not intended for long-term preservation. Furthermore, DRXS pays special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.4.2 ENVISION platform

DMP Component	Deliverable Title
Data Summary	Various data, like farm information, shapefiles containing farm location, layers will be generated via the platform. All of these data will be useful in order the ENVISION services and products to function properly and provide accurate information. These data have been saved in the ENVISION central database. All user actions (login, logout, visits on specific parts of the platform, visualization of maps, etc.) are logged and kept in the form of text file. This log is useful for debugging purposes. Reports containing information on user devices (which browsers and mobile phones) as well as number of mobile downloads (taken from play store for android downloads and app store for mac downloads) are useful for marketing and exploitation purposes, as well as decisions regarding the supported browsers and operating systems. Furthermore, files are exported and only the registered users will have access to them.
Making data findable, including provisions for metadata	The data are not directly accessible from outside. An overview of the ENVISION platform's functionalities and preliminary results will be available, discoverable and accessible to third parties, since the dissemination level of the respective deliverables D4.2 Initial version of ENVISION platform, D4.3 Integrated and validated version of the ENVISION platform and D4.4 Final version of ENVISION platform is public. Furthermore, through these deliverables, architecture updates will be available to third parties. • D4.2 Initial version of ENVISION platform: DOI: https://doi.org/10.5281/zenodo.6122302 • D4.3 Integrated and validated version of the ENVISION platform: DOI: https://doi.org/10.5281/zenodo.6303613 The naming convention used is: • Data_WP4_2_ENVISION_platform Every action on the platform produces meaningful metadata such as time and date of data creation or data amendments and are saved along the services results to enhance the discoverability of the results. The database is not discoverable to other network machines operating on the same LAN, VLAN with the database server or other networks. Therefore, only users with access to the server (ENVISION technical team members) are able to discover the database.
Making data openly accessible	Only registered users and administrators have access to the data. The data produced by the platform are personal data and will not be shared with others without user's permission. No open data will be created as part of ENVISION. The database is only accessible by the authorised technical

DMP Component	Deliverable Title
	team.
Making data interoperable	N/A
Increase data re-use	ENVISION has been integrated with third parties' applications, currently being used by PAs and/ or CBs, in order to re-use information already inserted in those systems and to import the results of the ENVISION services into their own systems. The raw data are not publicly available. Finally, the farmer RESTful API has been the backend system that drives the Mobile client application that is used by the farmers. The ENVISION platform is an open-source solution.
Allocation of resources	Resources have been allocated according to the project plan and WP4 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	All platform generated data are saved on the ENVISION database server. Encryption has been used to protect personal user data like emails and passwords. The ENVISION platform offers a secured way to collect users' data, providing a Graphical User Interface for uploading files via HTTPs calls that are additionally authorized by the OpenID Connect layer which is on top of the OAuth 2.0 protocol. If there is need for updates, the old data have been overwritten and all actions have been audited in detail and a log is kept, containing the changed text for security reasons. The system is weekly backed up and the backups are kept for 3 days. All backups are hosted on a remote server to avoid disaster scenarios. All servers are hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure of data. The ENVISION platform will not keep personal data and other information after the end of the project. Furthermore, DRXS pays special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.4.3   Maps produced by the EO data

DMP Component	Deliverable Title
Data Summary	One of the main offerings of the ENVISION platform is the generation of maps, based on the produced EO services, that can assist the PAs and CBs to increase their efficiency. Specifically, layers have been presented on the top of the maps depicting the outcomes of the remote sensing as well as layers from other resources (Natura sites, etc.).

	The types of the maps might differ but some indicative types for vectors are ESRI, shapefiles, GeoJSON, GML, etc. and for raster is GeoTiff. Similarly, the size might also vary a lot, from 1KB to 10GB.
Making data findable, including provisions for metadata	All the registered users have access to the maps. The users are able to identify the maps by their distinctive name. Meaningful metadata have been produced as a result of every action (time and date of data creation or data amendments, actions that took place, service that produced map, crop type of depicted farm). The naming convention used is: • Data_WP4_3_Maps
Making data openly accessible	Maps that have been produced are not openly accessible. Users should sign in to access the produced maps. The maps and the metadata have been made available for use by the ENVISION system through the secure API that has been created. The raw data, used for the generation of the maps' layers, that are stored in the ENVISION database are only accessible by the authorised technical team.
Making data interoperable	Maps have been saved in standard formats that are commonly used through OGC services.
Increase data re-use	Maps that have been produced during the project are offered to anyone who requests them. After the completion of the project, these data will only be available to users who will buy the respective services.
Allocation of resources	Resources have been allocated according to the project plan and WP4 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	All data generated by the platform are saved on the ENVISION server. DRXS pays special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.4.4 OCTOPUSH

DMP Component	Deliverable Title
Data Summary	OCTOPUSH is an integrated satellite and weather derived agricultural software service, which collects earth observation, geospatial, weather, in-situ, and other-referenced data, and applies appropriate processing algorithms and responds with ready-to-use results. It accepts Web requests form gridded datasets that will be ingested into the storage layer of GeoServer, which feeds the maps presented to the end user. In general, OCTOPUSH acts as a gateway to a plethora of agricultural services that enriches the geospatial visualization

	of the Service Providers results.
	The dataset provided by OCTOPUSH to ENVISION are:
	• Raster files of NDVI and Yield Estimation
	• Mean, stdev and median timeseries of NDVI and Yield Estimation
	• Classified raster with values 1=organic, 0=non-organic
Making data findable, including provisions for metadata	The collected/ processed and generated data are not publicly available and only the registered users have access to the results served by the ENVISION platform.
Making data openly accessible	The datasets are not publicly available.
Making data interoperable	N/A
Increase data re-use	Appropriate licensing agreement will be required for data access after the project's conclusion, which will be defined through the business model during the course of the project. The EO-based products will be usable by third parties through RESTful API, but only for those parties who are part of the project and during the lifespan of the project.
Allocation of resources	Resources have been allocated according to the project plan and WP4 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	OCTOPUSH is service gateway that provides a Web API via HTTPS, which supports Transport Layer Security (TLS) encryption. This ensures that the data-transfer is end-to-end encrypted. On top of that it authorizes the HTTPS requests by validating the API token, a mandatory parameter on all the HTTPS calls, against the Authorization Server registry. This process verifies that the identity making the request, is authorized to receive the particular set of data.
Ethical aspects	N/A
Other issues	N/A

2.4.5 DataCube

DMP Component	Deliverable Title
Data Summary	ENVISION DataCube leverages the power of the Open Data Cube (ODC) software to manage massive amounts of geospatial data, providing control of Analysis Ready Data. Its Python Application Programming Interface (API) enables easy querying and extraction of indexed data. The ENVISION DataCube also utilizes the Python XArray library for efficient calculations and analysis. The back-end services of ENVISION are built on top of the DataCube, allowing for cloud-optimized GeoTIFFs with any time, space, and band dimensions, and for results to be stored directly in the ENVISION database. ENVISION DataCube exploits the Open Data Cube (ODC) software aiming at managing massive amount of geospatial data by allowing the full control of Analysis Ready Data. It comes with a Python Application Programming Interface (API) enabling querying and extraction of indexed data. The

	ENVISION DataCube also utilizes the Python XArray library for efficient calculations and analysis. The back-end services of ENVISION are built on top of the DataCube, allowing for: • Cloud optimized GeoTIFFs in any time, space and band dimension • Results, which are directly stored in ENVISION database On top of the ENVISION DataCube, web services have been created. Specifically, the use of an open data cube and a Django rest API, combined with a user-friendly web page allow users to ask and get the indexed data in several formats. The provision of time series plots and graphs makes the data more accessible to users who may not have the technical expertise to work with raw data. These visualizations help users to understand trends and patterns in the data, making it easier for them to draw insights and make informed decisions. The web page provides an intuitive and user-friendly interface, making it easier for users to interact with the data and gain insights into their research questions.
Making data findable, including provisions for metadata	The collected/ processed and generated data will not be publicly available and only the registered users will have access to the results served by the ENVISION platform.
Making data openly accessible	The datasets are not publicly available. Regarding the processed Sentinel data, the use of a Django REST API and a web page greatly enhances the accessibility of sentinel data. This means that users can easily and quickly access data based on their needs and interests. With the use of a bounding box and time range, users can refine their search, making it more efficient and accurate. This not only saves time but also ensures that the user gets the data they need, thereby improving their experience.
Making data interoperable	Furthermore, the use of an open data cube and a Django rest API along with the web form, where users are able to use for requesting directly data, promote interoperability, making it possible for different software applications to exchange and use the data. This means that users can access the data using different applications, regardless of the software or operating system used. This eliminates the need to use proprietary software or formats, which can limit the usability of the data. Interoperability is crucial in data sharing and collaboration, as it allows researchers, scientists, and other users to work together effectively.
Increase data re-use	Appropriate licensing agreement will be required for data access after the project's conclusion, which will be defined through the business model during the course of the project. The EO-based products will be usable by third parties through RESTful API, but only for those parties who are part of the project and during the lifespan of the project.
Allocation of resources	Resources have been allocated according to the project plan and WP4 allocated resources. No additional costs are foreseen

	for making this dataset FAIR.
Data security	DataCube exploits a PostgreSQL/PostGIS database. The latter comes with a built-in user permissions system designed around the concept of roles. Thus, any user is assigned to a role with a login attribute. The external users will be assigned to specific roles that do not update the database except from the possible posting of requests to it for retrieving data from the DataCube
Ethical aspects	N/A
Other issues	N/A

2.5 DMP Components in WP5 – Business cases implementation and evaluation (EV ILVO)

DMP Component	Deliverable Title

Data Summary	The aim of this WP is to deploy and evaluate the ENVSION data products and services by testing them under different conditions and according to each PA's and CB's specifications and requirements.
	• The business value and acceptance,
	• The performance, usability, and effectiveness of these products and services,
	• Their impact at an economic, environmental and societal level
	have been evaluated during the project period.
	Detailed implementation guidelines have been produced to assure the smooth and uniform implementation across the business cases and outline common features and practices along with tailored information for each of the business cases. Furthermore, a comprehensive business cases action plan has been implemented providing in detail all the operations that should be performed such as data resources and data providers, business stakeholders, suggested potential actors and their role, potential ENVISION products and services users, dates of the beginning and end of the demonstration, and calendar of performance evaluation and feedback reporting, responsibilities and tasks of the platform and the providers.
	In addition, the evaluation criteria have been defined to measure the success and the impact of the delivered solutions.
	Lastly,
	• D5.4 Intermediate BC Implementation Report was developed based on collected information/data from;
	   o 10 (2 per BC) submitted internal progress reports reflecting the current status and progress of implemented activities,
	   o notes of regular progress meetings,
	   o and emails from WP5.
	• A concrete and detailed report "D5.5 Intermediate Report on The Evaluation of Services" has been produced evaluating Envision products and services in terms of business value, acceptance and performance, usability, effectiveness and impact at economic and social levels. The relevant data were generated and captured through;
	   o Workshops (physical or via the online platform Mural) to review and discuss the content of the revised indicators and criteria.
	   o Surveys to evaluate business value, acceptance, and performance, usability and effectiveness of Envision products and

services. Survey data was collected using an online survey tool (google form).

Survey answers:

- ✓ Likert scales, asking respondents whether they agree or disagree with the statement. Each option is given a score, which can be used to quantitatively analyse the results.
- ✓ Open-ended questions, asking respondents to provide their own answers.
- ✓ Closed questions asking respondents to provide data that are later analysed, normalised and processed.
  - o The interview data were collected and stored using digital recording (MS Teams meetings) only if the interviewees allowed it. In case of refusal, notes of the interviews were taken according to agreed formats and standards.
  - o Template to collect estimated KPI values for economic and social impact assessment.

Both reports aim to reveal key experiences and feedback for Envision solutions midway through BC implementation. Thus, the report's output will be able to support the Envision data product and services to reach the required maturity and meet specific customer needs.

During the process

- All data has been cleaned and validated to guarantee its accuracy and completeness. This entails finding and fixing any mistakes or discrepancies in the data, including fixing typos, finding missing or duplicate data, and checking data entry.
- The data collected were analysed and reviewed. Data analysis procedures are documented in detail within the respective deliverables to ensure that the analysis is transparent.
- Workshops, surveys, and interviews were conducted in accordance with ethical guidelines, including obtaining informed consent from participants.
- Mainly the main documents and formats that have been used to collect and generate the necessary data were the templates agreed in the D5.1 Implementation Guidelines and the D5.2 Action Plan.
- During the project, all the reports produced are stored in the project's collaborative file storage folder (Dropbox).
- All data is in doc./ docx./ xlsx. and pdf format.

Making data findable, including provisions for metadata	The raw data that have been collected in WP5 are not publicly available as they include confidential and personal data. The results are available through the respective deliverables,

since their dissemination level is public.

- D5.1 Implementation Guidelines: DOI: https://doi.org/10.5281/zenodo.6122511
- D5.2 Business cases action plan: DOI: https://doi.org/10.5281/zenodo.6310554
- D5.3 Evaluation criteria: DOI: https://doi.org/10.5281/zenodo.7874593
- D5.4 Intermediate Business Case Implementation Report: DOI: https://doi.org/10.5281/zenodo.7874601
- D5.5 Intermediate Report on the Evaluation of Services: DOI: https://doi.org/10.5281/zenodo.7874616

The naming conventions used are:

- Data_WP5_1_Guideliness
- Data_WP5_2_Action_Plan
- Data_WP5_3_Evaluation_Criteria
- Data_WP5_4_Reports

Making data openly accessible	The raw data that have been collected in WP5 are not publicly available as they include confidential and personal data. However, the results are available to third parties through the public deliverables.
Making data interoperable	N/A
Increase data re-use	The data that have been collected and processed during this WP are exclusively for analytical and statistical purposes and will not be re-used.
Allocation of resources	Resources have been allocated according to the project plan and WP5 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	The data that have been collected for internal use in the project and not intended for long-term preservation. The data are stored on EV ILVO servers. EV ILVO fully complies with the applicable national, European and international framework, and the European Union's General Data Protection Regulation 2016/679.
Ethical aspects	Confidentiality Agreement document has been prepared specifying the main purpose of the data collected and/ or generated within WP5, i.e., these data will be neither available to third parties nor discoverable and accessible to the public, since the parties disclosed to each other information and documentation, which is proprietary and confidential or otherwise generally not available to the public.
Other issues	N/A

2.6 DMP Components in WP6 – Commercialisation and exploitation (ETAM)

DMP Component	Deliverable Title
Data Summary	The purpose of the data collection in WP6 is to support commercialization and exploitation of the ENVISION products and platform, to define the business models for sustainable

	growth and to satisfy the needs for the collaboration with other EU projects.
	Several meetings with relevant EU projects have been performed in order to establish communication and explore possible ways of fruitful collaboration. The meetings were held with the Microsoft Teams Suite. The relevant material (emails, agenda, presentations, recordings) has been generated and collected.
	The data that have been collected and/ or generated within WP6 are represented the foreground knowledge, derived from the experienced based on the project implementation and the intangible data and results of the project, such as: business modeling information, outcomes, know-how, etc.
	The expected size of the data is not applicable, as the size is not a meaningful measure.
Making data findable, including provisions for metadata	The data with regards to the Business models are stored on ETAM servers and are not directly accessible from outside. Moreover, these data will be neither available to third parties nor discoverable and accessible to the public, since the dissemination level of the respective deliverables is confidential. Similarly, the data with regards to business plan and exploitation strategy.
	The dataset derived from the meetings with the relevant EU projects is not directly accessible by third parties. However, the main information and the outcomes of these meetings are presented in the respective deliverables and are accessible through the project's website and Zenodo:
	• D6.1 Collaboration with EU projects and initiatives (1): DOI: https://doi.org/10.5281/zenodo.6122094
	• D6.5 Collaboration with EU projects and initiatives (2): DOI: https://doi.org/10.5281/zenodo.7874631
	Moreover, the data collected for the Market analysis will be available through the public deliverables:
	• D6.2 Market Outlook Analysis: DOI: https://doi.org/10.5281/zenodo.6122356
	Regarding the data generated/ collected for the roadmap for the incorporation of EO-based monitoring in environmental assurance standards as well as for the incorporation of ENVISION in LEAF Marque will be publicly available since the dissemination level of the respective deliverables is public and they will be accessible either through the project's website or Zenodo.
	The naming conventions used are:
	• Data_WP6_1_Business_Plan
	• Data_WP6_2_Business_Models
	• Data_WP6_3_Exploitation
	• Data_WP6_4_ Collaboration_with_EU_projects
	• Data_WP6_5_Market_Analysis
	• Data_WP6_6_LEAF_Marque

Making data openly accessible	Data are publicly available as part of public deliverables and through the ENVISION website, dropbox folder and Zenodo. The other datasets are not publicly available.
Making data interoperable	N/A
Increase data re-use	Data that are publicly available through public deliverables are accessed and re-used by third parties indefinitely without a license.
Allocation of resources	Resources have been allocated according to the project plan and WP6 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	ETAM has established and is successfully implementing an information security management system (ISMS) in accordance with the requirements of the international standard ISO/IEC 27001:2013. Information security policies (including access control, secure storage and recovery) and an information security risk assessment process are in place. Furthermore, ETAM respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

2.7 DMP Components in WP7 – Dissemination and Communication (ITC)

DMP Component	Deliverable Title
Data Summary	The aim of the data collected and/ or generated within WP7 is to develop and implement an effective dissemination and communication strategy. The data that have been collected are statistics related to the project website, social media like LinkedIn, twitter, etc. for tracking the progress and improve the communication and dissemination activities. Reports have been collected from the partners regarding their performed dissemination activities through Jotform. Furthermore, personal data of newsletter subscribers have been collected (i.e. emails) and contact data of relevant project stakeholders. The expected size of the data is not applicable, as the size is not a meaningful measure. The data will be only available to the project partners.
Making data findable, including provisions for metadata	The data with regards to the dissemination and communication strategy and activities are publicly available and accessible by third parties since the dissemination level of the respective deliverables is public. These deliverables are accessible through the project's website, and Zenodo: • D7.1 Dissemination and Communication Plan: DOI: https://doi.org/10.5281/zenodo.4564222 • D7.2 Intermediate report on dissemination activities: DOI: https://doi.org/10.5281/zenodo.6303599 The naming convention used is: • Data_WP7_1_Activities Regarding the personal data derived from newsletters or other sources, such as webinars, clustering events, they are/will not be publicly available and only project partners will have access to them after request to the responsible controller. The naming convention used are: • Data_WP7_2_Personal_data No metadata will be generated.
Making data openly accessible	The dissemination and communication activity data are publicly available through the public deliverables and can be accessed and re-used by third parties indefinitely without any restrictions.
Making data interoperable	N/A
Increase data re-use	The dissemination and communication activity data are publicly available through the public deliverables and can be accessed and re-used by third parties indefinitely without any restrictions.
Allocation of resources	Resources have been allocated according to the project plan and WP7 allocated resources. No additional costs are foreseen for making this dataset FAIR.
Data security	All data will be stored on ITC servers. Furthermore, ITC pays

	special attention to security and respects the privacy and confidentiality of the users' personal data by fully complying with the applicable national, European and international framework, and the European Union's GDPR 2016/679.
Ethical aspects	N/A
Other issues	N/A

End of Document