Sintetic
SINGLE ITEM IDENTIFICATION FOR FOREST PRODUCTION, PROTECTION AND MANAGEMENT

D1.3
Data Management Plan
2nd version

Project Acronym: SINTETIC

Project name: Single item identification for forest production, protection, and management Europe

Call ID: HORIZON-CL6-2022-CIRCBIO-02-06-two-stage
(Harnessing the digital revolution in the forest-based sector)

Duration: 48 months

Starting date: 01/06/2023

Work Package: WP1

Task Number: T1.2

Lead beneficiary: CNR - Consiglio Nazionale delle Ricerche, alias National Research Council

Contributing beneficiary(ies): All Partners

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101082051

Dissemination level

[X] PU- Public: must be available in the website

[ ] CO- Confidential: Only for members of the Consortium and the Commission Services

Document history

Edition	Date	Status	Author
Version 1	28/08/2025	DMP update 2nd version Uses cases data flow	Tiziana De Filippis (CNR) Manuela Corongiu (LAMMA)
Version 2	18/09/2025	Review	Cristiano Guadagnino (BLUEB)
Version 3	27/09/2025	Second Revision	Tiziana De Filippis (CNR) Manuela Corongiu (LAMMA)
Version 4	30/09/2025	Final Revision	Tiziana De Filippis (CNR) Diana Bolopo (CTFC)

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101082051

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Acronym / Abbreviation	Meaning / Full text
AB	Advisory Board
API	Application Programming Interface
ARBOREAL	Arboreal AB
ASEMFO	Asociación Nacional de Empresas Forestales
BLUEB	Bluebiloba Startup Innovativa SRL
CC	Creative Commons
CNR	Consiglio Nazionale delle Ricerche or National Research Council
CSV	Comma-separated values
CTFC	Forest Science and Technology Centre of Catalonia
DMP	Data Management Plan
Dn.m	Deliverable number
DS	Demonstration Site
DOI	Digital Object Identifier
EC	European Commission
EOS	European Organisation of the Sawmill Industry
FAIR	Findable, Accessible, Interoperable, Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
HE	Horizon Europe
INNO	Innorenew Coe Center Odlicnosti Za Raziskave In Inovacije Na Podrocju Obnovljivih Materialov In Zdravega Bivanjskega Okolja
IPR	Intellectual Property Rights
LAMMA	Environmental Monitoring and Modeling Laboratory for the Sustainable Development Consortium
OGC	Open Geospatial Consortium

OTME	Otmetka Holding Ab
PMT	Project Management Team
RDA	Research Data Alliance
re3data	Registry of Research Data Repositories
SEP	Standard Ethics Protocol
SGDR	Sui Generis Data Right
SILVA	SILVADOR Company S.R.L
SIMTRO	Simtrona, Razvojna Dejavnost, D.O.O.
TL	Task Leader
UNITBV	Universitatea Transilvania Din Brasov
VM	Virtual Machine
WP	Work Package
WPn	Work Package number
WPL	Work Package Leader

INTRODUCTION

Deliverable D1.3 represents the second version of the DMP, following the initial version D1.2. As the project has progressed over 28 months, it is important to review and update the DMP to ensure it aligns with the project's ongoing development. Thus, the D1.3 DMP has been created to address the evolving needs of the project.

The ambition of SINTETIC is to define, prototype and demonstrate a complete solution for a digital platform dedicated to comprehensive forest value chain data management and protection of individual trees. Each tree will be linked to the data produced along the supply chain, with the unprecedented capacity to span from the forest inventory to the final sawnwood products. It will allow to relate yield and quality output of any process in the value chain with all the previous steps, including historical climatic data, silvicultural treatments, and forest stand descriptors.

The DMP describes what data will be collected, processed, or generated and what methodologies and standards will be applied. It also defines if and how this data will be shared and made open and how it will be curated and preserved.

The plan describes all relevant aspects, such as the definition of the data sources, data flows, security and privacy assurance, data features, data availability, data certification, data openness and/or availability, data storage, file formats, and standards to be adopted for the project needs. The DMP will outline how the data will be managed, shared, and preserved using FAIR (Findable, Accessible, Interoperable and Re-usable) data principles.

The DMP identifies the main data to be generated within SINTETIC, outlining the handling of research data during the project and how and what parts of the datasets will be openly shared.

The present deliverable also reflects the provisions established by the project contracts and complements the project exploitation, dissemination and IPR procedures and decisions defined in different deliverables (D7.1, D 6.1).

The DMP will evolve along the project implementation. It will start as a general overview of the data management strategy at the early stage. At the beginning of the project, the DMP outlines how data will be collected or generated, and how they will be organized, stored, and shared according to the FAIR principles. A detailed and completed life-cycle organization and description of data and related flows will be developed at the final stage. Recognizing that not all data types may be accessible at the start of the project, if any changes occur to the SINTETIC project due to the inclusion of new data sets, changes in consortium policies or external factors, the DMP will be updated. These revisions will ensure alignment with the data generated and user requirements identified by SINTETIC consortium participants.

This document is intended for internal and external use, describing the mechanisms that SINTETIC will implement to ensure all public data follow the FAIR data management principles. The current version presents the status and planning at month twenty-eight of the four-year project.

DMP purpose and principles

The purpose of this data management plan for the SINTETIC project is to ensure that all data produced during the project is properly traceable, available, and handled with the utmost quality assurance. Additionally, the plan aims to make the research data FAIR as it is formulated in the H2020 guidelines manual[1].

Since the SINTETIC project involves enterprises, an Intellectual Property Protection plan (D6.8) has been developed alongside the project management plan (PMP) to find a balance between data accessibility and IPR rights.

A Data Management Plan is a document that describes the data management life cycle. The plan will cover all the important aspects related to the research data, such as defining the data sources, data flows, ensuring security and privacy, describing data features, availability, openness, and storage, and specifying file formats and standards. The Data Management Plan will also include information on how to handle research data during and after the project ends. Additionally, the DMP will have information on all the research products and results.

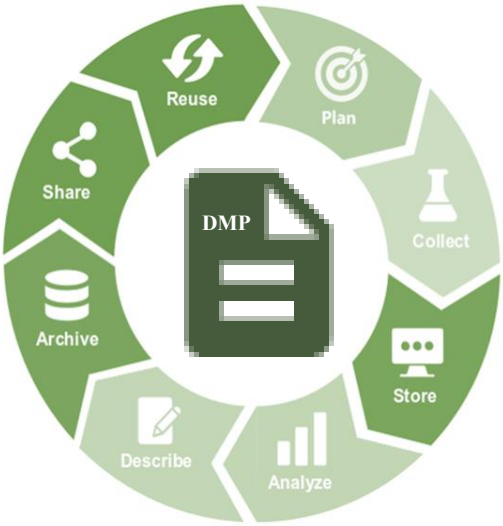


Figure 1: Data management life cycle

The plan is tailored specifically for SINTETIC forestry data to enable traceability throughout the entire wood value chain. DMP is based on an iterative approach, where the next steps are defined by and based on the feedback received from the end-users and stakeholders involved, first of all the project's partners.

The document will be updated three times during the project (Figure 2), in M6, M28, and M42 (D1.2, D1.3, D1.4). This first version was created during the initial months of the project (D1.2, M6). Regular updates will be required, especially after significant changes, such as new datasets or policy changes, including information on licensing of data, their availability, re-use and the reproducibility of research outputs.

[1] European Commission, Directorate-General for Research and Innovation. " Guidelines on FAIR Data Management in Horizon 2020." Version 3.0, 26 July 2016.

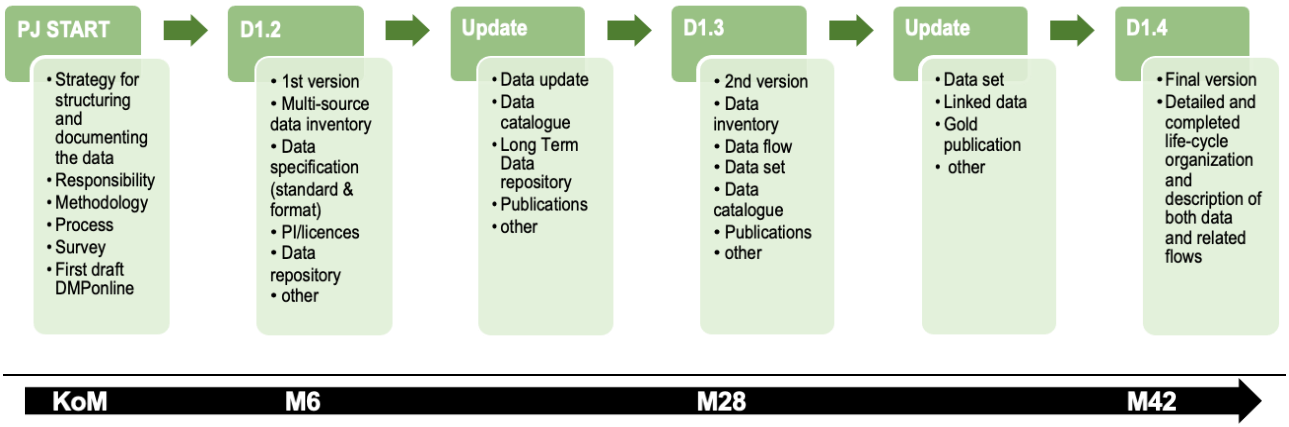


Figure 2: DMP overview versions and deliverables

The updates will focus on the SINTETIC DMP strategies, with a particular emphasis on the aspects of data findability, accessibility, interoperability, and reusability.

The DMP is a living document that will evolve along the project implementation. It will serve as a guide for other project's work packages and the implementation of the Geodatabase. It is important to keep it updated as needed. For more background and additional help, we refer to the OpenAIRE project[2] and its software tools, such as DMPTool. These tools provide guidance through a set of relevant questions.

DMP official versions will be stored on the project online repository as PDF files. An editable word copy of the latest version will also be stored to facilitate revision and update of the already identified datasets and policies. CNR will be in charge of updating the document and its annexes, uploading them on the repository and notify the consortium through the project mailing list system.

Relation with other project activities

The SINTETIC project is divided into seven Work Packages, as shown in Figure 3. Creating a Data Management Plan (DMP) is a crucial part of Work Package 1 to define system requirements, data model and technical aspects (formats, standards, relationship etc.). Task 1.2 is responsible for managing and coordinating the project data effectively. Although the Data Management Plan is a part of WP1, it is connected to all research WPs that collect (WP3) and generate data (WP2, WP4, WP5). Closed relations are also fundamental to the tasks of WP7 and WP6, which design the guidelines and rules for IP, copyright, and level of dissemination and exploitation of SINTETIC data and products. The information gathered from all WPs will be helpful for updating and versioning documents during and after the project's completion. All Work Packages will follow the procedures outlined in this document for data management and quality assurance.

[2] https://www.openaire.eu/how-to-create-a-data-management-plan

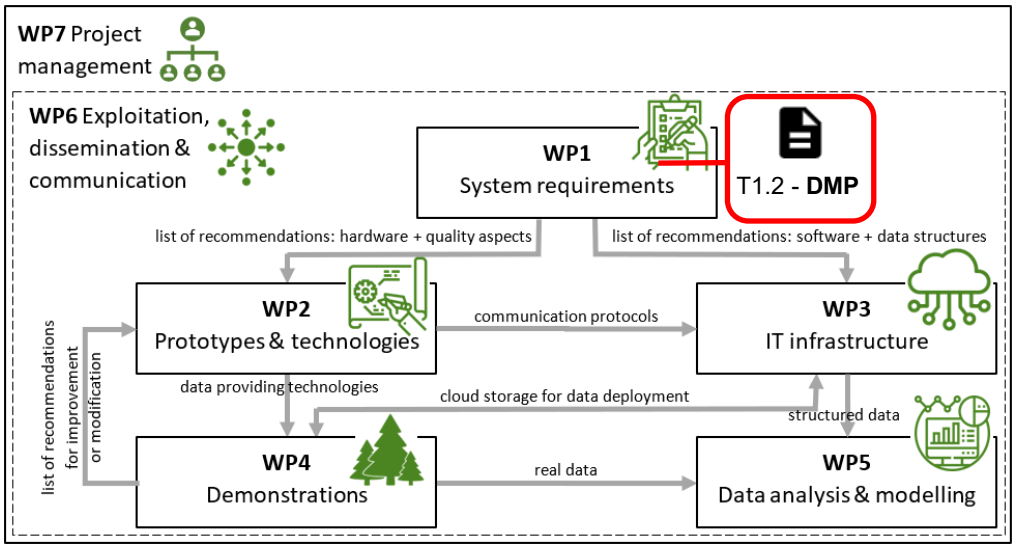


Figure 3: SINTETIC WPs structure

Document structure and updates

The document structure we have outlined follows the Horizon 2020 Data Management Plan template[3], a structured guideline consisting of a set of questions we need to answer with appropriate detail for the project's scope. The template is divided into six core sections, each addressing essential aspects of data management thoroughly:

1. Data Summary
2. FAIR Data Principles
3. Allocation of Resources
4. Data Security
5. Ethical Aspects
6. Other Research Outputs

This structured approach ensures a comprehensive exploration of data management aspects, ethical considerations, resource allocation, and the handling of diverse research outputs, all in alignment with the guidelines set forth by Horizon 2020.

This document presents the first update to the Data Management Plan (DMP) for the SINTETIC Project. It serves as an important review to ensure that the DMP aligns with the project's ongoing development on timber traceability system.

The GeoDB (Geodatabase) is central to the SINTETIC project, serving as the primary data repository for structured, semi-structured, and unstructured data. It provides the foundational infrastructure for data management, enabling seamless integration, traceability, and analysis of forestry data across various activities and stakeholders.

The second version of the DMP will therefore be closely tied to the progress of implementing the Project Geodatabase, which will be populated with data collected during the project's Demonstrations (WP4) and Analysis Services (WP5).

[3] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

DATA SUMMARY

Data Summary provides a brief description of the information to be gathered and the nature, scope, and scale of the data that will be generated or collected. For the first version of this plan, we have collected information from work packages tasks (GA_DoA) and the project data survey (T1.1) to provide the detailed responses required by the HE DMP template to describe the life cycle of SINTETIC data. The second version of this document report the advances in SINTETIC data collection, storage and management.

At this stage of the project only the example of data structure and formats have been tested to guarantee the data flow of GeoDB implemented for the timber traceability system, due the early stage of planning and data collect from DEMO site. On these premises the data are categorized on the basis of SINTETIC DEMO (WP4) and Geodatabase structure and entities relationships.

The purpose of data collection and generate by the project is to ensure that all data produced during the project is properly traceable, available, and handled with the utmost quality assurance.

Types of data managed in SINTETIC

There is a huge variety of data types used and collected in the timber value chain for the traceability system within SINTETIC project. Data can be data described and classified in many different ways[4], for example, based on their content, form, mode of data collection, digital or non-digital nature, primary or secondary nature, raw or processed, how they are created. Furthermore, in the contemporary data landscape, there is a new category of data known as Big data, which can come from various sources such as images (aerial, satellite, UAV, laser scanning) or crowdsourcing data from mobile devices.

Using the whole ecosystem of data types, this project will generate structured, unstructured and documented data to set up and demonstrate a traceability system for trees, logs and boards, it is based on ICT through a central database (T3.1) covering all the forest-based value chain during the project's life. The project aims to create a strong value chain traceability system for forest products by integrating various technologies and data. To achieve this goal, SINTETIC work packages will utilize different and heterogeneous data sources identified during the project life time (D 1.2 D 1.5) and Geodatabase implementation (D3.1).

[4]https://www.ugent.be/en/research/datamanagement/why/rdm-explained.htm

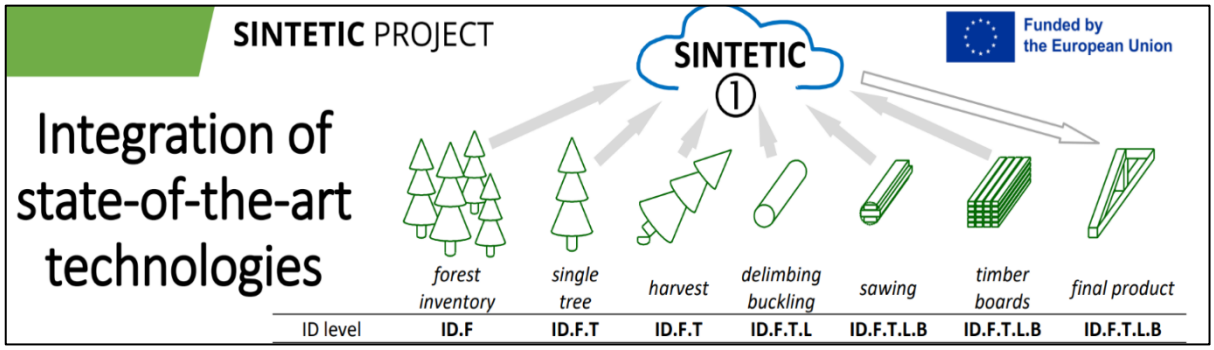


Figure 4: Data flow from different phases of forest traceability system (source SINTETIC DoA)

The conceptual data flow (Figure 4) has been redesigned within WP3 (T3.1), to better support the physical implementation of the SINTETIC Geodatabase and the development of the web application.
Data are collected from various sources, such as forest inventory tools, harvesting equipment, and sawmill sensors, as noted in the SINTETIC data workflow (see Annex 1). This data is then processed through APIs (Application Programming Interfaces), ensuring a seamless flow of information as well as data interoperability and traceability.
The data workflow is structured into three primary categories: 1) operational data collection and integration (WP2 & WP4), 2) database management and visualization (WP3), and 3) analytical services and data processing (WP5). The GeoDB serves as the backbone, supporting both operational and analytical data, while the frontend application provides users with a unified interface to visually interact with geospatial datasets.

1. Operational Data Collection and Integration

The data from various sources produced during the WP2 and WP4 activities are pre-processed and stored in the GeoDB. Data are so categorized on the base of process phase and digital devices in forest value chain as follows:

- **Forest Property Management:** Data from the ForestSharing platform (BLUEB) provides ownership records, land boundaries, and classifications. This data is integrated into the GeoDB via an API.
- **Tree and Log Measurement:** The Forest Scanner app (ARB) collects detailed forestry data (e.g. tree diameters, heights, log dimensions). This data is processed by ForestHQ (TREE) and shared with the GeoDB.
- **Harvesting Operations:** Harvester machines capture production data (e.g. log size, volume, location). This data is analysed by ForestHQ and integrated into the GeoDB.
- **SINTETIC harvester operations:** SINTETIC Harvester captures harvester production data, but also, RFID tag for each log, and wood quality indicators from the sensor installed on the machine. This data is analysed by ForestHQ and integrated into the GeoDB.
- **Sawmill Process:** Microtec sensors (MICRO) monitor logs and boards size and quality. The data is exported as JSON files, parsed by the SINTETIC Data Parsing Service (TREE), and integrated into the GeoDB.

2. Analytical Services and Data Processing

Remote Sensing Data are considered as a valuable data source in forest inventory, illegal logging analysis (WP5), land cover changes, climate characterizations, and wood quality modelling. SINTETIC services (T3.2) use ERA5 Land reanalysis dataset and NDVI time series from Sentinel 2 for the two main tasks:

- **Illegal Logging Detection:** Satellite data is analyzed by the Illegal Logging Analysis service (CNR). Results are imported into the GeoDB as non-structured files (e.g., PNG, JSON, CSV) for further interpretation and visualization.
- **Climatic Data Analysis:** Meteorological data is processed by the Climatic Data Analysis module (CNR). Summaries of results are stored in the GeoDB as non-structured files (e.g., PNG, CSV) for use alongside other forestry data.

Data derived from climate and illegal logging services are imported and managed by APIs in the GeoDB as attached files to forest operation. For more details see API documentation at
https://geodb-staging.sintetic.iit.cnr.it/api#tag/Climate-Data-Attachments
https://geodb-staging.sintetic.iit.cnr.it/api#tag/Vegetation-Data-Attachments

At this stage of the project, the SINTETIC GeoDB is structured to manage the following resources (entities), as they appear on the main staging documentation page: geodb-staging.sintetic.iit.cnr.it.

• Organizations • Users • Auth (authentication and token, login, URL signature, ecc.) • Tree Species • Forest Property Owners • Forest Properties • Forest Operations • Forest Operation Attachments • Climate Data Attachments • Certificates • Forest Operation Certificates • Subcompartments • Vegetation Data Attachments • Sample Plots • Sample Plot Attachments • Tree Processors • StanForD Attachments • Tree Log Product Definitions • Cutting Instruction Sets • Cutting Instruction Items • Trees	• Tree Attachments • Tree Measurements • Tree Logs • Tree Log Attachments • Tree Log Measurements • Tree Log Locations • Sawmill Companies • Sawmill Sites • Sawlines • Sawline Attachments • Sawmill Batches • Tree Log Sawmill Measurements • Board Classes • Boards • Quality Indicators • Quality Indicator Items • Quality Indicator Aggregations • Quality Indicator Aggregates • Harvesting Reports • Inventory Reports • Sawmill Tree Log Reports • Tree Log Reports

Figure 5: List of resources (entities) structured in the SINTETIC GeoDB

Documentation on the relational database structure and data management APIs is available at geodb-staging.sintetic.iit.cnr.it.

The SINTETIC GeoDB currently occupies 128 GB of storage. This capacity will be expanded to 5 TB to accommodate the storage and backup of LIDAR data as well. Additionally, 200 GB has been allocated for the Virtual Machine (VM) responsible for storing and managing the SINTETIC data catalogue.

3. Data not included in the SINTETIC GeoDB

Heterogeneous data from external sources are identified for their re-use in SINTETIC activities (WP5; WP4) regarding mainly: Earth observation data and global observation systems; Climate datasets (e.g. ECMWF and NOOA); Digital Elevation Model; Infrastructure's maps from national agencies providing data; Environmental maps; Forestry and LandCover-LandUse maps; Forest inventories (including NFI – National Forest Inventories - one's), ecosystem data, cadastral information, sectorial studies. These data are mainly derived from external services and will not be archived in the SINTETIC Traceability System. The GeoDB will only manage the output data from the Climate and Spectral Analysis and Illegal Logging Detection services (T3.2 and T5.3). This output data will be uploaded to the GeoDB as attached files using the specific APIs (T3.1).
In addition, publications, reports, and datasets related to project activities, such as extra demos or data for training (e.g., illegal logging, wood quality) will be archived in the CKAN[5] catalogue (Figure 6) with agreed-upon access permissions. This type of data is not strictly part of the prototype of GeoDB traceability system and are managed separately.

[5] https://ckan.org

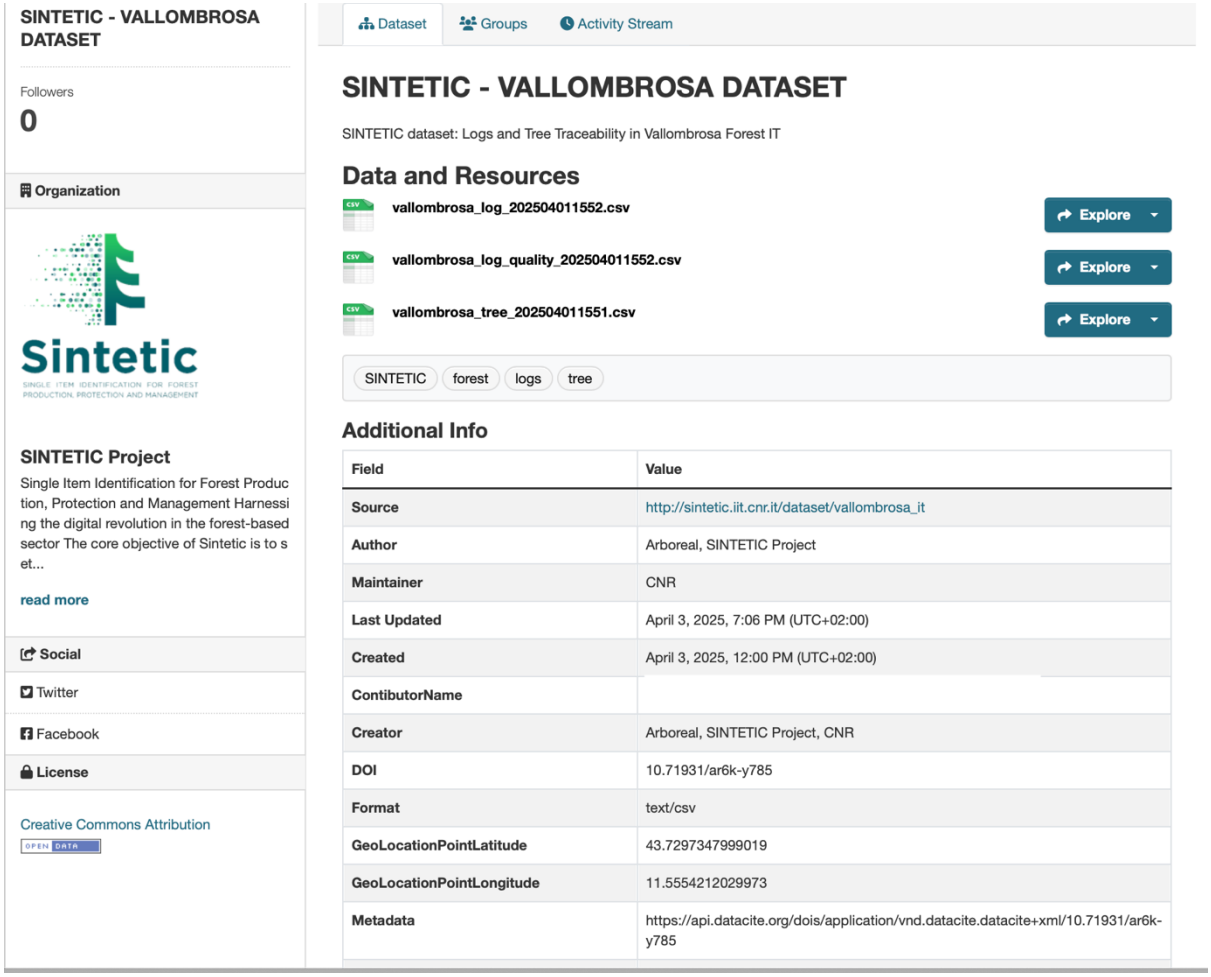


Figure 6: A Pre-demo dataset in the SINTETIC catalogue

Data format

Providing data and documentation in preferred archival and re-use formats facilitates more timely and efficient processing and release. Preservation formats should be platform-independent and non-proprietary to ensure they will be usable in the future.

SINTETIC will generate higher-level datasets based on the sources and WPs' activities. The formats used in internal data storage and processing are distinct from those data sets recommended for facilitating reuse through deposit in external, trusted repositories.

Data to the SINTETIC GeoDB is uploaded only via its API, and it has to follow specific formats and authentication methods tailored to each resource.

The SINTETIC GeoDB APIs primarily accept JSON payloads for data creation and update. Each resource (e.g., Tree Species, Forest Property, Forest Operation) has its own required JSON structure.

The APIs rely on standard JSON data types—string, integer, number, boolean, object, and array. They also support domain-specific formats, including GeoJSON for spatial data and enumerations for constrained value sets. Temporal data is represented as dates or datetimes in string form, and spatial filtering uses arrays for bounding boxes.

For a comprehensive list of all resources and their specific data formats, please refer to the SINTETIC GeoDB API documentation on https://geodb-staging.sintetic.iit.cnr.it/.
For resources that support attachments (e.g., Forest Operations, Sample Plots), it's possible to upload files such as CSVs, PNG images, or spreadsheets. The APIs allow to associate these files with specific records (https://geodb-staging.sintetic.iit.cnr.it/docs/attachment-files).
Data downloaded from the SINTETIC GeoDB are usually returned in JSON format (or sometimes CSV or other formats depending on the endpoint).

Data origin and purpose

The SINTETIC traceability system will use ICT and geomatics tools to capture data and track its flow from sensors deployed throughout the timber value chain.
Data are provided from different sensors by machines/apps/operators on forest stands, trees, logs, and sawn wood, in the whole value chain. WP2, WP4 and WP5 are the packages involved in the collection, processing, and analysis of data for the deployment and test of SINTETIC traceability system.

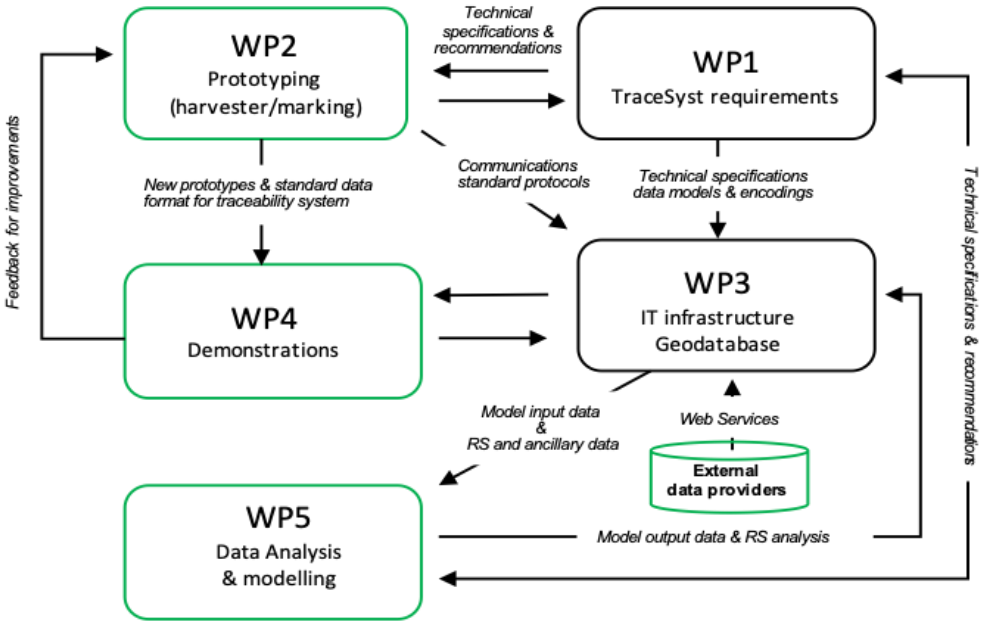


Figure 7: WPs involved in data provisions (green) and data flow toward the SINTETIC Geodabase

The prototypes developed in WP2 are devoted to data collection along the forest traceability chain, to delivering a Marking and tracking system for trees and timber products (D2.3), as well as the use of the smartphone app is addressed to timber measuring, grading and tracking (D2.4); a prototype of a forest harvester will implement tracking functions and quality sensors (D2.5) as well as a Quality assessment and traceability system will be prototyped in sawmills (D2.6).
The SINTETIC prototypes after deployment will be extensively tested during a series of demonstrations (WP4). All

data collected from prototypes' tests (WP2) and Demos (WP4) will gradually populate the data storage (unstructured data in their native formats) and the SINTETIC Geodatabase (WP3) (structured data[6]), according to guidelines and specifications defined in WP1 and T3.1. Data from ground truthing measurements (e.g. with traditional tools on logs), collected along the WP4's tasks in the demo sites, will also secure the data needed for modelling in Task 5.1 and Task 5.2 (e.g. selection of harvest sites to compare different silviculture treatments) and wood quality models and applied service (WP5).

In the first release of the Data Management Plan (DMP), we provided a descriptive overview of the data sources for each task in WP2, WP4, and WP5. In this second version, the datasets generated by each SINTETIC demonstration (DEMO) are identified and described within their respective workflows, along with the services required to populate the GeoDB and support the development of SINTETIC applications and Graphical User Interfaces (GUIs).
Each demonstration involves collecting specific datasets that pertain to the development, integration, testing, and feedback of prototypes and systems aligned with the goals of the SINTETIC project.
Although the demo tasks are currently underway in this stage of the DMP, the related data flows have been defined, and the APIs have been implemented in T3.1. Use Case UML diagrams (Annex 2) have been created for each demo, from 1 to 10. Partners are actively engaged in their development to enhance monitoring and testing related data flows and services throughout various forest operations of each WP4 demo. This work facilitates the transfer of information toward the SINTETIC GeoDB and the related GUI.

Data from external data providers

Cartographic data, national forest inventory data, satellite data, climatological data, and other complementary data required for analysis will not be stored in the SINTETIC GeoDB. Users will need to download this type of open data directly from external services, such as Copernicus Climate Services or national cartographic portals, or access it through OGC Web services.
The GeoDB will only manage the output data from the Climate and Spectral Analysis and Illegal Logging Detection services (T3.2 and T5.3). This output data will be uploaded to the GeoDB as attached files using the specific API (T3.1).

Data utility

After conducting an external data survey (T1.1), we found that stakeholders outside our project have a growing demand for comprehensive and accessible forestry-related data. The survey revealed that industries dependent on timber sourcing, forest organizations, and research institutions are interested in accessing detailed traceability system data for their respective analyses and decision-making processes.

Specifically, these stakeholders expressed a need for:

[6] *As structured data is intended the geodatabase physically implemented, including the related services*

- **Timber Sourcing Industries**: Companies involved in timber sourcing need to access detailed traceability system data to enhance their supply chain management. This includes information on the origin, quality, and sustainability of timber, which is crucial for meeting regulatory requirements and ensuring responsible sourcing practices.
- **Forest Organizations:** Organizations responsible for forest management need to leverage the traceability system data for monitoring and improving forest health. The detailed inventory and timber data would enable them to make informed decisions about silviculture treatments, assess the impact on timber quality, and enhance overall forest management strategies.
- **Research Institutions:** Academic and research institutions have shown interest in the traceability system data for conducting in-depth analyses. The data provides valuable insights into the relationships between forestry practices, climate conditions, and the quality of timber. This is vital for advancing research in the field of forestry, ecosystem services, and climate impact on forests.

The possibility to relate highly detailed inventory and timber data to climate and forest management records would also allow to better understand the results of silviculture treatments on timber quality, forest health, and ecosystem services. This comprehensive dataset serves as a valuable resource for various stakeholders, aligning with the principles of openness, accessibility, and transparency advocated by the SINTETIC project.

Making data Findable, Accessible, Interoperable and Re-usable (FAIR)

SINTETIC embraces the FAIR (Findable, Accessible, Interoperable and Reusable) principles, adopting Open Data and Open Access practices for the data collected or produced within the project. The FAIR data principles are a set of guidelines that help make data more accessible and usable. They ensure that data can be discovered through catalogues or search engines, is accessible through open interfaces, is compliant with standards for interoperable data processing, and, therefore, can be easily reused. By making data FAIR, we can ensure that it is a valuable resource that can be easily accessed and used by anyone who needs it.

As Open Science guiding principles, research data (e.g., publications and datasets) will be openly available through public repositories, such as the Open Access Infrastructure for Research in Europe (OpenAIRE), to facilitate dissemination, validation, and increase the potential use of research results. The data will be licensed (e.g. CC BY, ODC BY) to permit the most comprehensive re-use possible.

Data Findability

To make data findable, it is important to give it a clear and unique name or identifier. Additionally, it should be accompanied by descriptive information about what the data contains and where they should be stored. This will help both humans and computers easily locate the data they need.

This section provides guidelines on how to ensure data findability within the SINTETIC project. It outlines the importance of assigning clear identifiers (preferably DOI[7]), depositing datasets in trusted repositories (Zenodo[8], OpenAIRE[9]-compliant, institutional), and describing them with standardized metadata (DataCite[10], ISO19115[11], Dublin Core[12]).

During the project, the SINTETIC Geodatabase will be used to archive and manage data. When the project is completed, each research team must deposit and describe their datasets in SINTETIC data repository or in an identified public data repositories (Trusted Repository) as long as they can assign persistent unique identifiers to the deposited items. Valid and machine-readable DOIs (Digital Object Identifiers) are recommended to be used to make datasets produced by SINTETIC citable for publication.

Through the OpenAIRE EU portal, it is possible to find and choose a repository to deposit or publish the research data in Open Access. It's possible to choose the appropriate trusted repository where to deposit the project research datasets between two typologies: 1) OpenAIRE compliant repository registered in OpenDOAR, re3data and FAIRsharing, and 2) Zenodo.

[7] https://www.doi.org
[8] https://zenodo.org
[9] https://www.openaire.eu
[10] https://datacite.org
[11] https://www.iso.org/standard/53798.html
[12] https://www.dublincore.org

OpenAIRE allows choosing among Institutional Repository, Data Repository, Thematic Repository, Publication Repository, Software Repository, while Zenodo is indicated as a multidisciplinary repository (catch-all). The chosen data repositories should support standard descriptive metadata to ensure datasets indexing and discoverability by machines. Zenodo is a recommended repository that satisfies these essential requirements. Metadata of Zenodoare compliant with DataCite Metadata Schema[13] minimum and recommended terms. Metadata of each record is indexed and searchable directly in Zenodo search engine immediately after publishing and sent to DataCite servers during DOI registration and indexed there. However, other repositories, including institutional partners' repositories, must support Dublin Corea and DataCite Metadata Schema.
SINTETIC will also provide all relevant documentation explaining the data collection and analysis procedures, including codebooks, methodologies, and other necessary information. This will allow a clear understanding of the project findings and enable reproducibility and validation through specific information or code. All documents such as reports, guidelines, papers, diagrams, etc., will be archived in the trusted repository and assigned a DOI.
The DMP, in the last planned releases, will identify standard rules for naming and versioning each datasets category to improve data visibility, discoverability, citation, and permanent online tracking.
Actually, datasets are not yet available because they are still linked to the ongoing data collection operations of the DEMOs. Specific datasets with DOI are considered as final products.

Data findability is provided also by identification of associated metadata. Metadata provides information about the characteristics of a dataset and can include a variety of information types. Descriptive metadata, for instance, contains details such as the title, abstract, author, and keywords, and is mainly used to identify and locate a dataset. Administrative metadata is another type of metadata that comprises details such as the license, intellectual property rights, creation date, access control, and so on.
In SINTETIC, the definition of a set of metadata elements is necessary to allow identification of the vast amount of information managed for which metadata is created, its classification and identification of its geographic location and temporal reference, quality and validity, conformity with implementing rules on the interoperability of spatial data sets and services, constraints related to access and use, and the organization responsible for the resource.
Descriptive information about what the data contains related to forest operations, forest properties, trees, and logs includes details such as geospatial location, operation type, status, and other relevant characteristics that are included in the GeoDB resources tables. Some tables include specific fields for metadata, such as `additional_info` (in JSON format) for additional information. Semi-structured data uses formats like JSON and XML, which include self-descriptive metadata.
Standard metadata schema (e.g. DataCite metadata schema, ISO19115, etc.) will be adopted for different data set generated from the Geodatabase or other SINTETIC services to be explored, like CKAN[14], the open-source data catalogue that have been customized (Figure 7) for the project (https://catalog.sintetic.iit.cnr.it).

13 https://schema.datacite.org/
14 https://ckan.org/

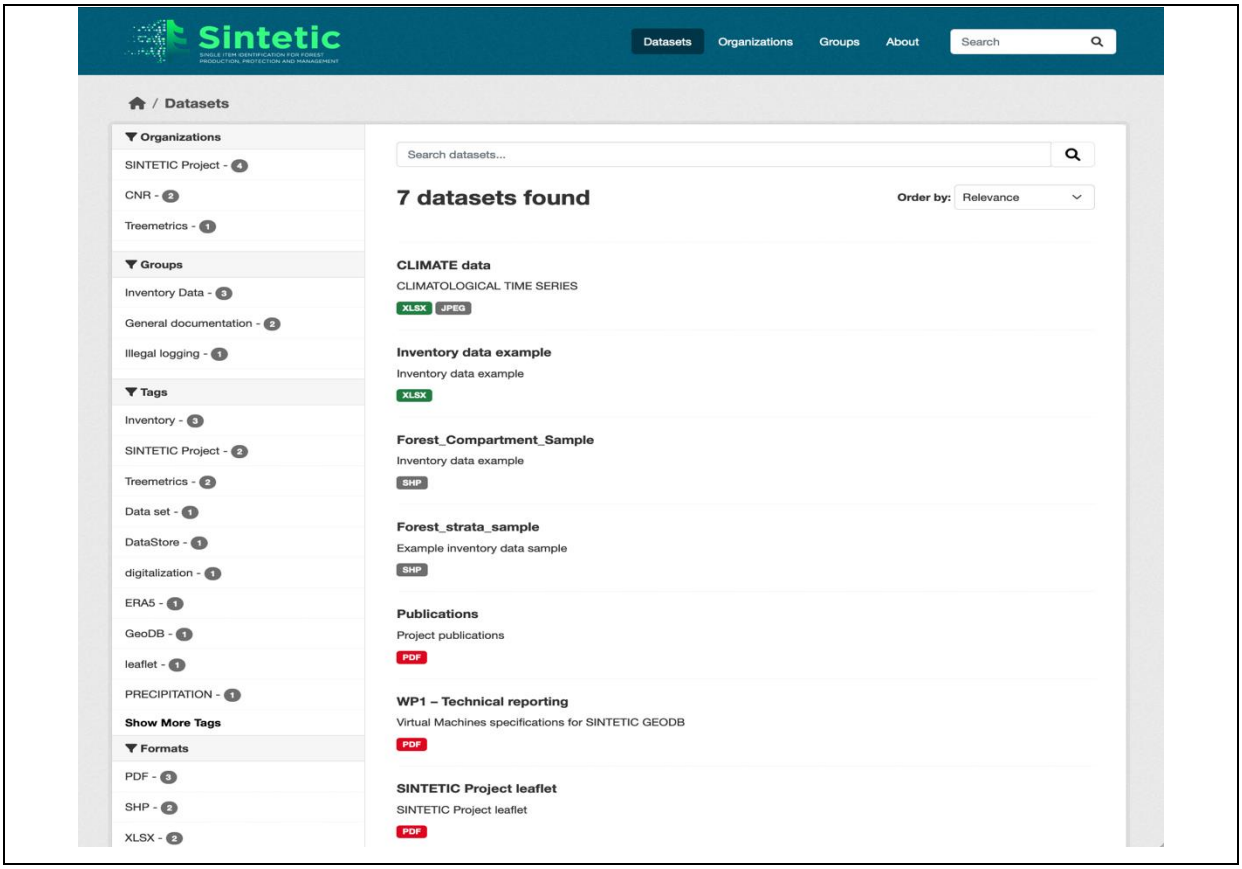


Figure 8: SINTETIC catalogue

The RDA Metadata Standard Catalogue[15] will be the reference in the choice of the appropriate metadata data standard that will be adopted for each typology of research data produced by SINTETIC. Annex III reports the main proprieties required by DataCite metadata proposed for all generic datasets to describe the data that will be produced along the project life. During the process of creating dataset, curators are asked to provide metadata following the schema and guidance[16].

Data Accessibility

SINTETIC adheres to the guiding principle of making research data openly accessible as much as possible to allow for dissemination and validation. This approach maximizes the potential for research results to be reused. All files will be converted to standard and well-documented open formats, and datasets deposited will include all relevant documentation and explanations. Exceptions to this principle will only be made in cases where restrictions on data access or sharing is necessary. The principle of "as open as possible, as closed as necessary" will be considered in such cases.

15 https://rdamsc.bath.ac.uk/

16 Reference: DataCite Metadata Working Group. (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. DataCite e.V. https://doi.org/10.14454/3w3z-sa82

Data should be easily accessible to anyone who needs it, without unnecessary restrictions such as passwords or complicated access procedures. However, restrictions on data access or sharing may be necessary in certain circumstances:

- when collected data belongs to third party which have denied permission for sharing them on account of confidentiality and proprietary issues;
- protection of personal data of key informants involved in surveys, focus groups, interviews;
- if making the data available could potentially jeopardize the project's main aim.

Any other legitimate reasons for restricting data access will be explained in the accessibility details relating to each specific dataset.

SINTETIC will adopt all possible and legitimate actions and strategies to allow data sharing including:

- seeking explicit copyright permissions from third-party data owners for reusing, reproducing, and distributing the collected data when required. Specific agreements with data owners will be pursued in such cases.
- prioritizing the utilization of standard open formats or self-descriptive formats for both internal and external and purposes.
- providing all relevant documentation and explanation for the data and the datasets, including the procedures adopted to obtain them, versioning, and software for reading data in case of non-standardize formats.
- obtaining the consent of citizens and stakeholders involved in focus groups and anonymizing and aggregating the data of interviews or brainstorming or in evaluation activities, typically carried out within WP4 and WP6 tasks.

in case of copyright on raw data derived, collected, or elaborated from pre-existing databases or from other original sources (i.e., papers, journal articles, book chapters, reports, video and audio sources), collected data will be made available if the reproduction and sharing are allowed by expressed permission of the right holders or by applicable copyright exceptions and exemptions.
If there is any data that fall under some of the restrictions described above and it cannot be made shareable, the EU allows complete closure or restricted access to them. The SINTETIC DMP in the final version will report the versions or part of datasets that cannot be freely shared providing the specific motivations as per GA.

The research data from the SINTETIC project will be deposited in the CNR data server (WP3, T3.1), that should be registered as an institutional data repository in re3data.org.
Accessibility level to the data server has been defined in D3.1 which specifies how to manage the access to project partners and their members. At the time of presentation of results in scientific peer-reviewed publications, researchers will deposit the project dataset that can be shared in a trusted data repository (where a DOI is assigned to the dataset uploaded) to guarantee their discoverability, access, and preservation beyond the project end. Such repositories support open licenses and different access levels. Cross-linking between publications and the relevant datasets freely accessible will guarantee Gold Open Access[17] Publications deliverables. Finally, they adopt descriptive metadata standards as required by the OpenAIRE Guidelines and allow cross-linking between

[17] https://open-access.network/en/information/open-access-primers/green-and-gold

publications and the relevant datasets. The specific teams responsible for a specific dataset are responsible for the management in the repository of their choice.

Before the end of the project, the consolidated datasets and related publications will be deposited in the SINTETIC Data Server, which is registered in re3data.org and in DataCite as an institutional data repository and accessible through the SINTETIC catalogue. The catalogue will assign a DOI and standard metadata to each resource. Datasets and other research outputs may also be deposited in trusted repositories, such as Zenodo or other thematic repositories indexed on the OpenAIRE platforms, in order to ensure their usability and long-term preservation. In such cases, the SINTETIC catalogue will provide a link to the external resource together with its corresponding DOI. In general, Zenodo is recommended for the open dissemination and preservation of research data by research teams that do not have access to suitable institutional, national, or disciplinary repositories, or are not required to use their institutional repositories.

Data Interoperability

Interoperability is another key aspect of making data FAIR. Data should be structured in a way that allows it to be combined and used with other data. This can be achieved by using standard formats and definitions that different systems can understand. To enable seamless data integration across domains, communities, platforms and tools, a combination of a standardized data model, a standardized transport protocol and a standardized data access interface is required.

As the amount of data increases during the digital traceability system within the timber supply chain, finding ways to ensure interoperability and manage large volumes of data becomes crucial. Open Geospatial Consortium (OGC) models and data exchange services are sufficient to handle the available data. However, this element will be reviewed during the project. For now, data interoperability will be ensured by complying with internationally adopted standards. A specific task is to explore the present data formats and define the best options to harmonize the data interchange (WP1, Task 1.1). Final data formats will comply with the most common standards used in the sector, such as the StanForD for timber procurement data, to facilitate its use by all the actors involved.

Data interoperability also requires specific ontologies and vocabularies for data and metadata to make project data understandable and usable. So, each SINTETIC dataset will be associated with specific keywords derived from Thesauri and controlled vocabularies[18] to enhance semantic discoverability. At this stage, the Consortium will refer to DCAT[19] vocabulary, designed to facilitate interoperability between data catalogues published on the Web. By using DCAT to describe datasets in data catalogues, publishers can increase discoverability and enable applications to use metadata from multiple catalogues easily.

Data Re-use

Finally, data should be well-documented and clear so that others can understand and use it for different purposes. When data is reusable, it saves time and effort for researchers and helps avoid duplication of work.

---

[18] *Controlled vocabularies are standardized and organized arrangements of words and phrases and provide a consistent way to describe data. Metadata creators assign terms from vocabularies to improve information retrieval.*

[19] https://www.w3.org/TR/vocab-dcat-1/

The DMP is a living document, periodically updated, including also information on licensing of data, their availability, re-use and the reproducibility of research outputs. The data will be licensed to permit the widest re-use possible (e.g. CC BY, ODC BY). As the project involves enterprises, an Intellectual Property Protection plan will be developed in parallel to the PMP, to agree on a proper balance between data accessibility and IP rights -both being among the first outputs of the project.

ETHICAL ASPECTS

The SINTETIC project adopts ethical guidelines that regulate the overall execution of the project activities. Despite the different legislation regarding ethical aspects in European countries, SINTETIC guarantees the ethical principles reported in the Project Management Plan (D7.2) during all project activities.

All research activities carried out within the project will comply with the Article 19 of the Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021, establishing Horizon Europe - the Framework Programme for Research and Innovation about ethics, and with the European Code of Conduct for Research Integrity (ALLEA 2023; The European Code of Conduct for Research Integrity – Revised Edition 2023. Berlin. DOI 10.26356/ECOC).

SINTETIC operates in line with the principle of serving the public interest, demonstrating a commitment to democratic collaboration, and conducts its functions with integrity, fairness, impartiality, political neutrality, and ideological neutrality. The project actively advocates for ecological preservation, energy conservation, and overall environmental protection, prioritizing the safeguarding of the environment for future generations.

All personal data collected within the SINTETIC project from questionnaires, interviews, surveys and focus groups are carefully protected in compliance with relevant national data protection legislation of the EU member states implementing the European directive 95/46/EC and with the procedures defined by the European Code of Conduct for Research Integrity.

It is important to evaluate ethical issues on a case-by-case basis. If researchers want to share sensitive and confidential data in an ethical manner, they can adopt certain strategies such as obtaining informed consent that includes consent for data sharing, protecting privacy by anonymizing the data, and considering controlling access to the data through measures like embargoes or access/licensing terms and conditions.

DATA SECURITY

Ensuring data security involves taking measures to protect data during all stages of a project and after its completion. Raw data containing participants' personally identifiable information will be rigorously protected during collection and processing. Effective security practices include access controls such as passwords, encryption, backup power systems, and defences against viruses and unauthorised access.

Data access

Accessibility of research data and the terms of use are two crucial issues that need to be addressed in a research data repository. To define the general accessibility of a research data repository, we need to consider three categories representing different access levels. Firstly, there is access to the research data repository, which defines whether users can access the database in general. Secondly, there is access to the research data sets in the repository. Thirdly, there is access to upload research data sets to the repository. The access to each level can be open, restricted, or closed. Open means that there are no access barriers. Restricted means that external users can overcome access barriers. Closed means that external users cannot overcome access barriers. Embargoed access applies only to the level of access to the research data sets. It means that external users cannot overcome access barriers until the data is released for open or restricted access. Table 1 illustrates the main access levels (General Accessibility, Data Set Accessibility, Upload Access), each with options for Open, Restricted, or Closed access. Additionally, it includes a specific category for Embargoed Access, which applies only to the level of access to the research data sets.

Table 1 Accessibility levels for research data repository

Access Level	Access to Repository	Access to Data Sets	Upload Access to Repository
General Accessibility	Open / Restricted / Closed	Open / Restricted / Closed	Open / Restricted / Closed
Data Set Accessibility	-	Open / Restricted / Closed	-
Upload Access	-	-	Open / Restricted / Closed
Embargoed Access (Data Sets)	-	Embargoed	-

Data archiving and preservation

Digital data must be actively managed over time to ensure they will always be available and usable. This is important in order to preserve and protect shared scientific heritage as technologies change. Depositing data

resources with a trusted digital archive can ensure that they are curated and handled according to good practices in digital preservation. By depositing data with a trusted repository (e.g. OpenAire, Zenodo, Institutional repository, DataCite), the SINTETIC project will ensure that the research data are migrated to new formats, platforms, and storage media as good practice requires. CNR will host and maintain the Geodatabase in its server farm, which is registered as data repository in re3data.org, through the whole duration of SINTETIC and will also guarantee its maintenance, activity and public accessibility for at least 5 years after the end of the project (WP3).

ALLOCATION OF RESOURCES

Assessment of project cost for data management

As part of the SINTETIC project, the Geodatabase will be used as the primary data storage system to share data among project partners. Some data sets will also be shared with external users. The expenses associated with activating and maintaining this system throughout the project will be managed within the WP1 and WP3 project budget, including the FAIR expenses.

Achieving FAIR data standards requires a dedicated allocation of researchers' time and investments in infrastructure. CNR staff time has been assigned to cover the expenses for organizing data and documentation for archiving purposes and maintaining the data repository and Geodatabase for 5 years after the end of the project.

Furthermore, there are no costs for long-term storage and preservation of publicly shareable data, even if a trusted repository such as Zenodo is chosen. Data storage and curation costs are not imposed, and they have an expected lifespan of at least 20 years.

Project tasks and responsibilities related to data management

The T1.2 leader will be responsible for data management and updates to DMP and later releases. The WP3 Task 3.1 leader maintains the SINTETIC Geodatabase, data server, updates the configuration and APIs.

However, the creators of individual datasets are responsible for managing them. These creators are usually team leaders directly involved in organizing and collecting research data. To identify themselves, researchers and personnel involved in dataset creation will use the unique persistent identifier ORCID[20]. This identifier links a researcher's identity with their research activities and products, and it is free of charge. In the future release of DMP, a table summarizing the contacts of the research team leaders responsible for each dataset will be added.

To appropriately credit all personnel involved in data creation and management activities, a list of roles will be adopted in the project. The provisional list will include Data Collector (such as survey conductors, interviewers, or a person who runs and manages a sensor or model), Producer (person responsible for preparing data to be shared in a specific format), Project Member (a researcher indicated in the GA), Researcher (person assisting co-authors with research, data collection, processing, and analysis but is not part of the team indicated in the GA), and Research Group (the name of a research institution or group that contributed to the dataset)

All WPs use and produce data in different formats generated from devices, instruments and users. The number of data set produced/used by SINTETIC is huge and with no standard metadata.

[20] https://orcid.org

WPs (and leader of tasks) will be requested to provide an initial effort to identify data (used and produced) and their description according to FAIR principles and T1.2 recommendations. If necessary, when required, specific focus groups will be organized (e.g. licence, IP, restrictions …) whit a full involvement of partners.

CONCLUSIONS

This second version of DMP outlines the essential principles and guidelines for managing data in the SINTETIC project. As a living document, it will be reviewed and modified as needed throughout the project's duration to reflect changes in its progress. During the life cycle of the SINTETIC project, a wide range of data types will be collected from various sources, including forestry machinery, remote and proximal sensors, smartphone apps, and other technologies used in timber traceability systems. The datasets will be as open as possible and as closed as necessary, with a focus on sound forestry data management for best research practices, creating value, and promoting knowledge and technology beyond the project.

To ensure that the data is well-managed, archived, and preserved, SINTETIC's partners will be encouraged to adhere to sound data management practices. Updates to the Data Management Plan will involve revising the research data repository where data will be collected and shared. Gradual generation and collection of descriptions of the dataset and research data will also be done. This deliverable will be updated (D1.4) in M42.

References

Jäckel, D and Lehmann, A. 2023. Benefits and Challenges: Data Management Plans in Two Collaborative Projects. *Data Science Journal*, 22: 25, pp. 1–7. DOI: https://doi. org/10.5334/dsj-2023-025

ARGOS

https://argos.openaire.eu/home

Data Repository reg.

https://www.re3data.org/repository/r3d100010108

https://fairsharing.org/10.25504/FAIRsharing.NzKTvN

https://fairsharing.org/

OpenAIRE EU

https://catalogue.openaire.eu/home

https://explore.openaire.eu/search/dataprovider?pid=r3d100010108

note: It's possible to register in OpenAire data repositories that are already in re3data.org

National Forestry Database

http://service.re3data.org/repository/r3d100010108

https://phaidra.univie.ac.at/search?sortdef=created%20desc&q=DATA%20MANAGEMENT%20PLAN&page=1&pagesize=10

https://phaidra.univie.ac.at/o:1603940

Center for International Forestry Research (CIFOR)
Dataverse Catalogue
https://data.cifor.org/dataverse/s?q=FOREST

CIFOR Harvested Dataverse - Center for International Forestry Research (CIFOR)
https://dataverse.harvard.edu/dataverse/cifor_harvested?q=&types=datasets&sort=dateSort&order=desc&page=2

The National Forestry Database (NFD) is Canada's Compendium of Forestry Statistics,
http://nfdp.ccfm.org/en/data/harvest.php

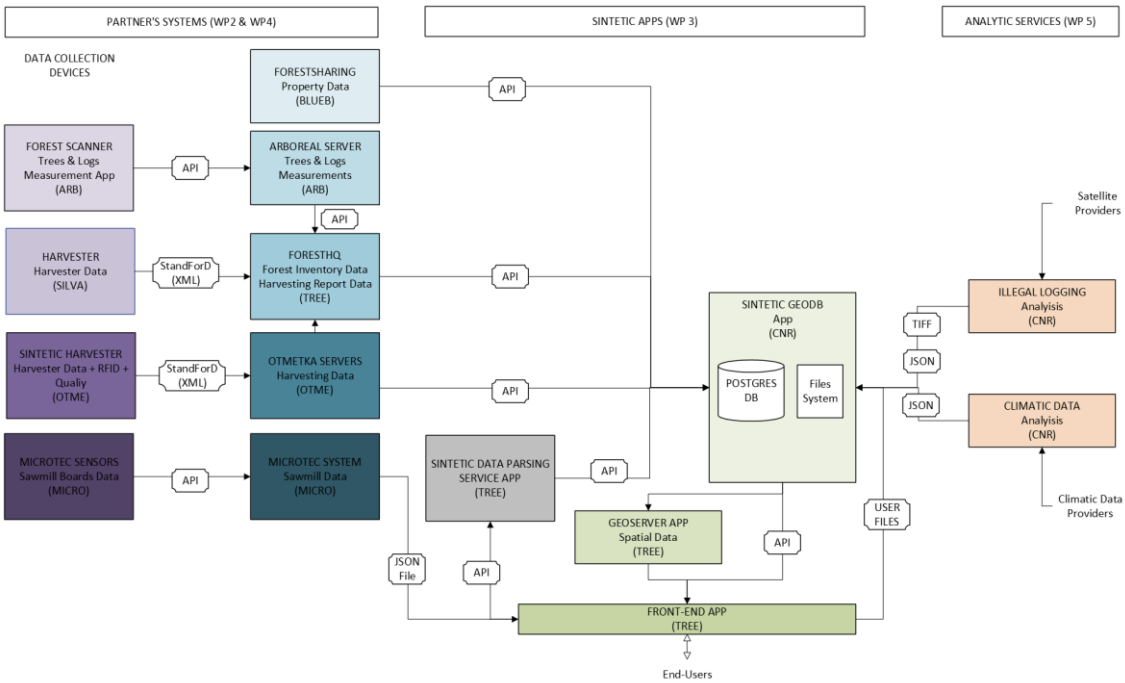
DMP EXAMPLES

• CORDIS / EuropeanCommission DMP Examples,2021, collection of over 800 completed DMPs from approved EU projects
• ZenodoDMP collection•Digital Curation Centre (DCC) DMP collection
• Data Management Plan Catalogue LIBER(AssociationofEuropean Research Libraries

https://dmptool.org/
DMPTool is a service of the California Digital LibraryOpens in a new window, a division of the University of California Office of the PresidentOpens in a new window. Version: v4.1.8

D1.3 Annexes

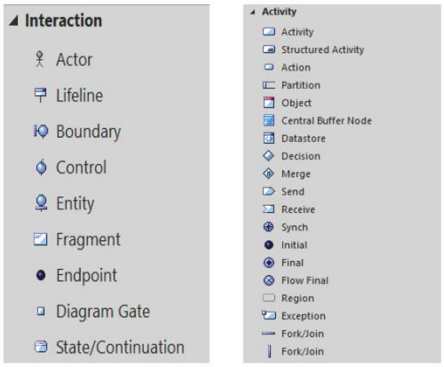
Annex 1 – SINTETIC data workflow defined in D 3.1

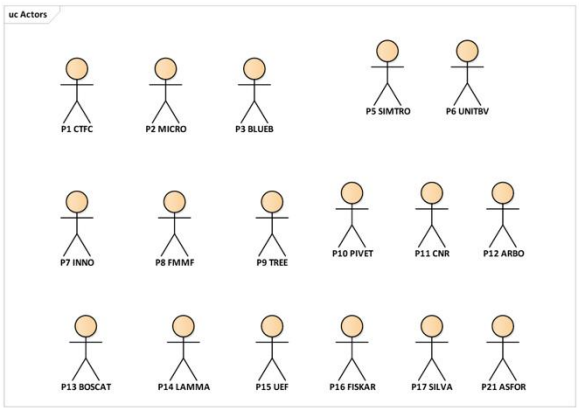


Overall data workflow around the SINTETIC GeoDB. (Extract from Deliverable 3.1 - Geodatabase of all data and metadata provided by sensors and forest machinery)

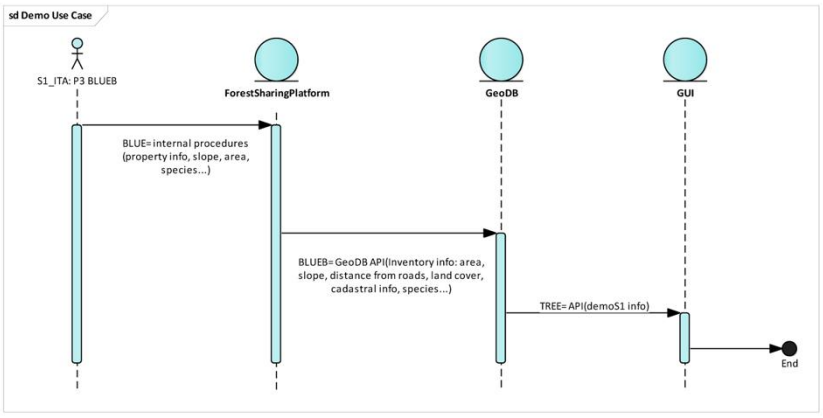
Annex 2 – WP4 DEMO data flow

WP4 – Demo Use Cases vs Services

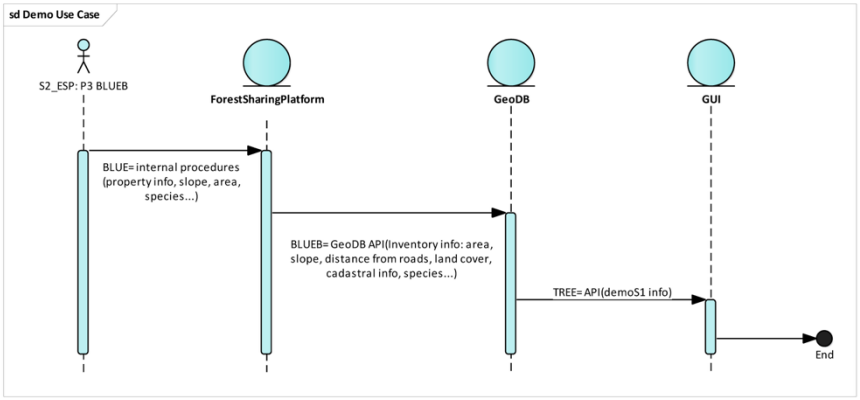




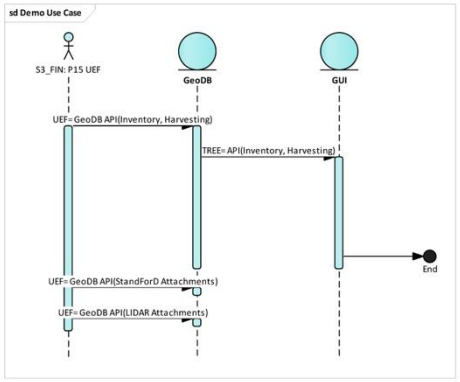
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
1	BLUB	ITA	Shape of interest area (Lat Long)	T3.1	ForestSharing	Blueb loba implemenation	Blueb loba implemenation - GEODB API T3.1 PUT in GeoDB property and inventory information	Property Details Invnetory data (BLUB)	PUT: property and addtional information : property boundaries, distance from the road, soil health, slope of the land, plant species, etc. Additional materials on the forest sharing platform (videos, photos, etc.) will not be transferred to the GeoDB.	BLUB	no intermediate data storage required	data tranmission contact



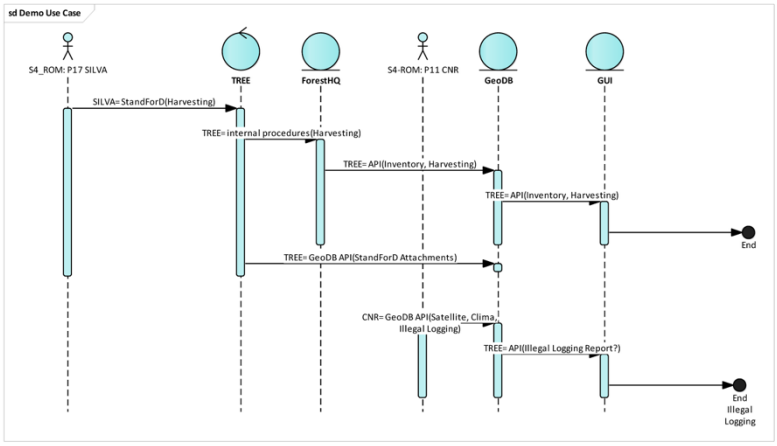
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
2	BLUB	ESP	Shape of interest area (Lat Long)	T3.1	ForestSharing	Blueb loba implemenation	Blueb loba implemenation - GEODB API T3.1 PUT in GeoDB property and inventory information	Property Details Invnetory data (BLUB)	PUT: property and additional information : property boundaries, distance from the road, soil health, slope of the land, plant species, etc.. Additional materials on the forest sharing platform (videos, photos, etc.) will not be transferred to the GeoDB.	BLUB	no intermediate data storage required	data tranmission contact



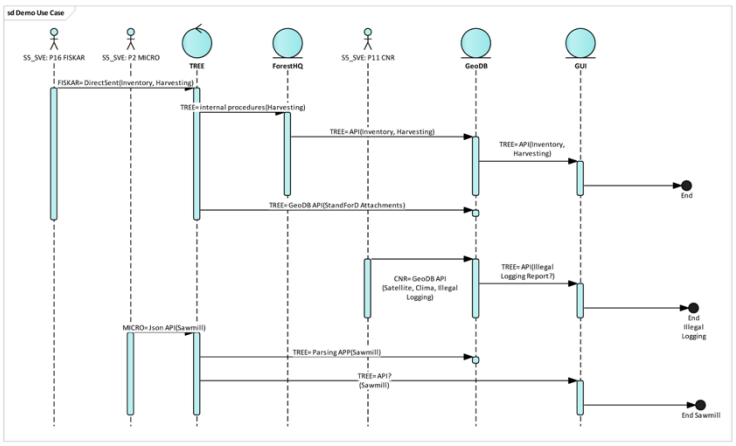
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
3	UEF_KONE	FIN	Shape of interest area? (Lat Long)?	T3.1 T2.5	Inventory- LIDAR Harvesting	UEF implemenation	UEF implemenation - GEODB API T3.1 PUT into GeoDB structured (inventory and harvesting?) PUT LIDAR Files as unstructured PUT StandForD as unstructured data	1) Terrestrial Lidar Point clouds 2) Inventory Data (from analysis and reports) 3) HARVESTER Mechanised StandForD files 4) Havesting Data (from analysis and reports)	UEF is in very early stage of development	Carlos provides StandForD/LIDAR services?	Lidar ? STAC catalogue ? Size 20 GB?	



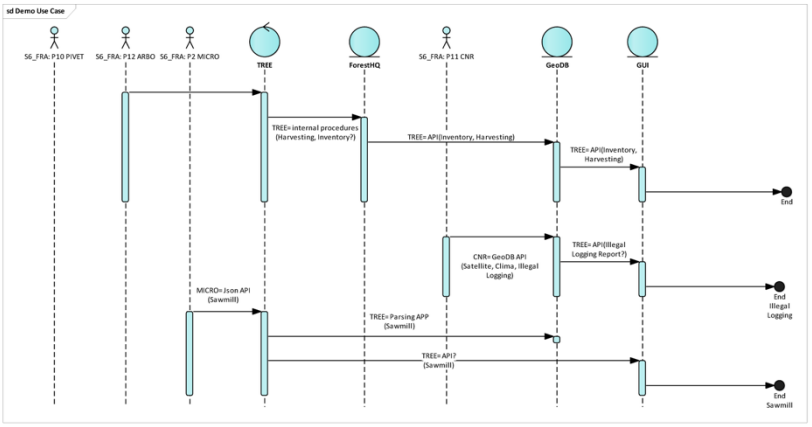
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	4 ASFOR	ROM	Shape of interest area? (Lat Long)?, UniBrasov ?	T3.1 T3.2 T3.3	Inventory (historical data) ? Harvesting CNR Illegal Loging Service	ForestHQ harvester ForestHQ inventory?? CNR Illegal Loging Service	Illegal Logging Service - GeoDB API 3.1 GETinto GeoDB Unstructured (Images and illegal logging report) ForestHQ - GeoDB API 3.1 PUT intoGeoDB structured Inventory info ? PUT into GeoDB structured harvesting data PUT StandForD files as unstructured data ? PUT other operational files as unstructured data	1) Harvester Mechanised StandForD files 2) Illegal logging outputs (raster and json)		Leandro Provides NDVI/StandForD Services Historical data by model and demo data	ASFOR Historical data? STAC catalogue, NDVI from CNR	



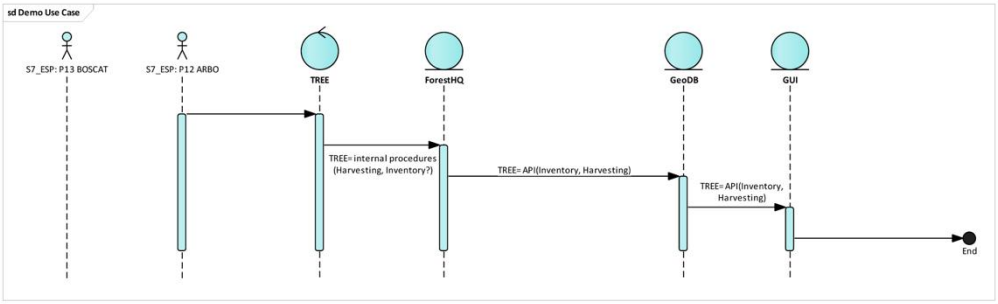
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	5 FISKAR	SVE	Existing Inventory already defined? shared into the GeoDB?	T5.1 T5.2 T3.2 T3.3	Inventory Harvesting Sawmill	TREE Sawmill service (Microtec) ForestHQ (harvesting) ? Arboreal? CNR Illegal Loging Service	TREE Sawmill service - GeoAPI T3.1 PUT into GeoDB structured PUT into GeoDB JSON files unstructured ForestHQ - GeoAPI T3.1 PUT intoGeoDB structured Inventory info PUT into GeoDB structured harvesting data PUT StandForD files as unstructured data? PUT other operational files as unstructured data	1) Microtec Sawmill JSON 2) Inventory data? 3) HarvesterMechanised StandForD?	Inventory measured within demo Arboreal, LiDAR or already done Type of comparison needed? Quality? Sawmill JSON S provided by Microtec and parsed by Treemetrics	?		



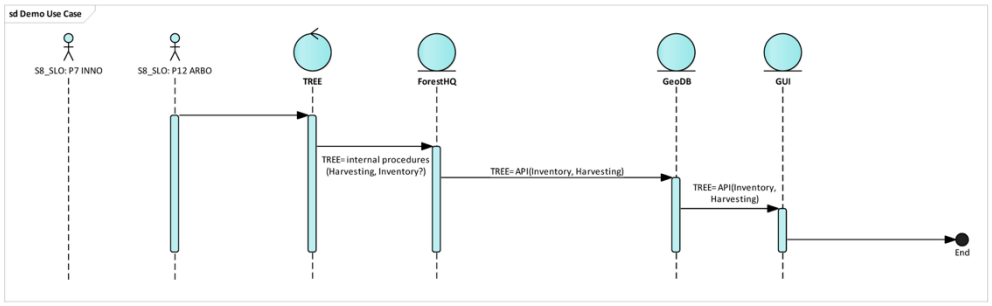
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	6 PIVET	FR	Demo Shape area? Inventory data?	T5.1 T5.2 T5.3 T3.2	Inventory? Harvesting Sawmill	TREE Sawmill service ForestHQ ? Arboreal?	TREE Sawmill service - GeoAPI T3.1 PUT into GeoDB structured PUT into GeoDB JSON files unstructured ForestHQ - GeoAPI T3.1 PUT intoGeoDB structured Inventory info PUT into GeoDB structured harvesting data PUT StandForD files as unstructured data? PUT other operational files as unstructured data Illegal Logging Service - GeoDB API 3.1 PUT into GeoDB Unstructured (Images and illegal logging report)	1) Sawmill json ? 2) Inventory data? 3) HarvesterMechanised StandForD?		Microtec+ CNR? Tree		



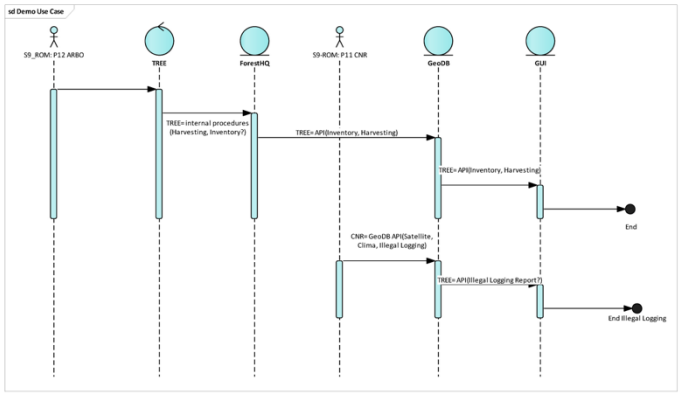
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	7 BOSCAT	ESP	Demo Shape area? Inventory data?	T2.4 T3.3	Inventory Harvesting	Arboreal ForestHQ Harvesting Import Data Service	Arboreal manual Harvesting - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured manual harvesting info Arboreal Inventory - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured Inventory info Manual harvesting data file (form) - GeoDB API T3,1 PUT into GeoDB manual ADDITIONAL harvesting info	1) Manual Harvesting structured 2) Manual harvesting data file? 3) Inventory data	PUT: manual Harvesting as structured data PUT: additional Manual Harvesting as structured/unstructured/intermediate data? PUT: StandForD as unstructured Data ? PUT: inventory as structured data?	ARBOREAL/ TREE	NO additional manual as structured data into GeoDB ?put them as intermediate out the GeoDB?	



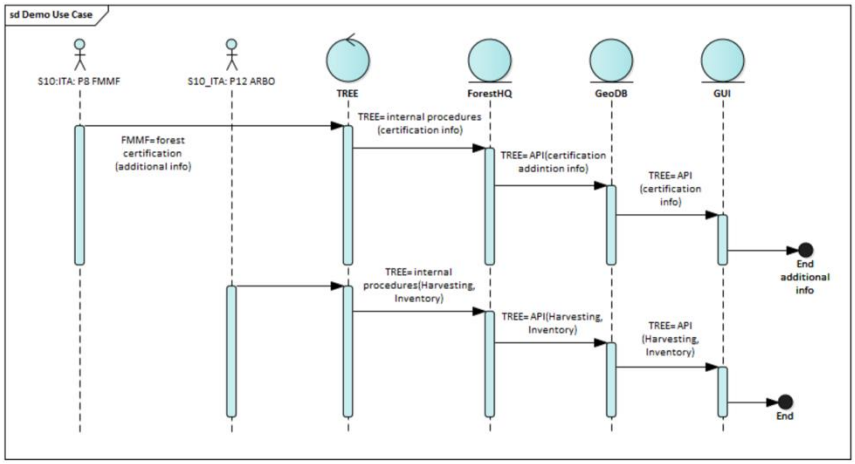
WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	8 INNO	SLOV	Demo Shape area? Inventory data?	T2.1 T2.2 T3.2 T3.3	Inventory ? Harvesting ?	Arboreal ? ForestHQ ? Harvesting Import Data Service ?	Arboreal manual Harvesting - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured manual harvesting info Arboreal Inventory - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured Inventory info Manual havesting data file (form) - GeoDB API T3,1 PUT into GeoDB manual ADDITIONAL harvesting info	1) Manual Harvesting structured 2) Manual harvesting data file? 3) Inventory data	PUT: NDVI into STRUCTURED or UNSTRUCTURED GeoDB?	ARBOREAL/ TREE		



WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	9 ASFOR	ROM	Demo Shape area? Inventory data?	T5.3 T2.1 T2.2 T3.2 T3.3	Inventory Harvesting (manual) IllegalLogging	Arboreal ForestHQ Harvesting Import Data Service	Illegal Logging Service - GeoDB API 3.1 GETinto GeoDB Unstructured (images and illegal logging report) Arboreal APP manual Harvesting API PUT into GeoDB structured manual harvesting info Arboreal manual Harvesting - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured manual harvesting info Arboreal Inventory - ForestHQ - GeoDB API T3,1 PUT into GeoDB structured Inventory info Manual havesting data file (form) - GeoDB API T3,1 PUT into GeoDB manual ADDITIONAL harvesting info	1) Manual Harvesting structured 2) Manual harvesting data file? 3) Inventory data Satellite time series data (raster or json?)	PUT: into GeoDB manual ADDITIONAL harvesting info where? Illegal Logging, requires structured standforD, who do this? ADDITIONAL manual harvesting requires a StandforD-like structure? who do this?	ARBOREAL/ TREE /CNR?		



WP4 Demo	Partner	Country	DEMO AREA	TASK involved	Sevices Required	Partner services	Service	dataset	Comments	Who	STORAGE	note
	10 FMMF	ITA	Demo Shape area? wchich data into GeoDB?	T3.1 T2.4?	Arboreal Inventory (historical)? Manual Harvesting Illegal Logging	Bluebiloba/ForestSharing ForestHQ? (import inventory data)? Harvesting Import Data Service? Quality indicator for logs? CNR Illegal Loging Service	Illegal Logging Service - GeoDB API 3.1 GETinto GeoDB Unstructured (Images and illegal logging report) Quality DATA PUT into GeoDB structured data about "FMMF certified" quality indicator (YES/NO) Historical Inventrory - ForestHQ? - GeoDB API T3,1 PUT into GeoDB structured Inventrory info Manual havesting data file (form) - GeoDB API T3,1 PUT into GeoDB manual ADDITIONAL harvesting info	Historical inventory data	How PUT historical inventory data in DB ?? How PUT into GeoDB structured data about "FMMF certified" quality indicator (YES/NO)	BLUB ? ARBOREAL/TREE?		



Annex 3 – Data Cite metadata schema

ID	*Property*	*Obligation*
1	Identifier (with mandatory type sub-property)	M
2	Creator (with optional given name, family name, name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	Publication Year	M
6	Subject (with scheme sub-property)	R
7	Contributor (with optional given name, family name, name identifier, and affiliation sub-properties)	R
8	Date (with type sub-property)	R
9	Language	O
10	Resource Type (with mandatory general type description sub-property)	M
11	Alternate Identifier (with type sub-property)	O
12	Related Identifier (with type and relation type sub-properties)	R
13	Size	O
14	Format	O
15	Version	O
16	Rights	O
17	Description (with type sub-property)	R
18	Geolocation (with point, box, place, and polygon sub-properties)	R
19	Funding Reference (with name, identifier, and award related sub-properties)	O
20	Related Item (with identifier, creator, title, publication year, volume, issue, number, page, publisher, edition, and contributor sub-properties)	O

DataCite Mandatory (M) Recommended (R) and Optional Properties (O)

Annex 4 - Elements of DPM monitoring

Element	*Description*	*SINTETIC Mapping*
Data description	A description of the information to be gathered; the nature and scale of the data that will be generated or collected.	Expected Data
Existing data	A survey of existing data relevant to the project and a discussion of whether and how these data will be integrated.	Data Survey
Format	Formats in which the data will be generated, maintained, and made available, including a justification for the procedural and archival appropriateness of those formats.	Data Format specifications and Dissemination
Metadata	A description of the metadata to be provided along with the generated data, and a discussion of the metadata standards used.	Data Format, metadata schema and Dissemination
Storage and backup	Storage methods and backup procedures for the data, including the physical and cyber resources and facilities that will be used for the effective preservation and storage of the research data.	Data Storage and Preservation of Access
Security	A description of technical and procedural protections for information, including confidential information, and how permissions, restrictions, and embargoes will be enforced.	Data Format and Dissemination
Responsibility	Names of the individuals responsible for data management in the research project.	Roles and Responsibilities
Intellectual property rights	Entities or persons who will hold the intellectual property rights to the data, and how IP will be protected if necessary. Any copyright constraints (e.g. copyrighted data collection instruments) should be noted.	Data Format and Dissemination
Access and sharing	A description of how data will be shared, including access procedures, embargo periods, technical mechanisms for dissemination and whether access will be open or granted only to specific user groups. A timeframe for data sharing and publishing should also be provided.	Data Storage and Preservation of Access
Audience	The potential secondary users of the data.	Data Format and Dissemination
Selection and retention periods	A description of how data will be selected for archiving, how long the data will be held, and plans for eventual transition or termination of the data collection in the future.	Data Format and Dissemination
Archiving and preservation	The procedures in place or envisioned for long-term archiving and preservation of the data, including succession plans for the data should the expected archiving entity go out of existence.	Data Storage and Preservation of Access
Ethics and privacy	A discussion of how informed consent will be handled and how privacy will be protected, including any exceptional arrangements that might be needed to protect participant confidentiality, and other ethical issues that may arise.	Data Format and Dissemination
Budget	The costs of preparing data and documentation for archiving and how these costs will be paid. Requests for funding may be included.	SINTETIC GA
Data organization	How the data will be managed during the project, with information about version control, naming conventions, etc.	Geodatabase and data management plan
Quality Assurance	Procedures for ensuring data quality during the project.	Data flow and quality control