Sintetic
SINGLE ITEM IDENTIFICATION FOR FOREST PRODUCTION, PROTECTION AND MANAGEMENT

D1.2
Data Management Plan

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 or National Council for Research

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	30/11/2023	Working draft for internal review	Tiziana De Filippis (CNR)
Version 2	11/12/2023	Updated draft internally reviewed	Alex Poveda (TREE)
Version 3	18/12/2023	Updated draft after contribution from partners	Gemma Arnau (CTFC)
Version 4	28/12/2023	Final Revision	Antonio Ruano (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
ARBOREAL	Arboreal AB
ASEMFO	Asociación Nacional de Empresas Forestales
BLUEB	Bluebiloba Startup Innovativa SRL
CC	Creative Commons
CNR	Consiglio Nazionale delle Ricerche or National Council for Research
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
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
WP	Work Package
WPn	Work Package number
WPL	Work Package Leader
UNITBV	Universitatea Transilvania Din Brasov

INTRODUCTION

This document is the first version of the Data Management Plan (DMP) for the SINTETIC project. Its purpose is to outline how research data will be handled during and after 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 will outline 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 six 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 online manual *(ref:).*

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 IP rights.

A Data Management Plan is a document 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 DMP have been create using the official generic template[1] from the European Commission (version of 2021-05-05) that describes in 10 (sub) sections the topics that need to be covered. The template is a set of questions that we should answer with a level of detail appropriate to the project. Relevant questions useful for the first version of the DMP has been included in the data survey forms created within Task 1.1.

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. It is also recommended to provide an update along with periodic reporting. 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 and its software tools, such as

[1] https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf

DMPTool. These tools provide guidance through a set of relevant questions. We can also refer to the guidelines on FAIR Data Management in Horizon 2020 for additional help.

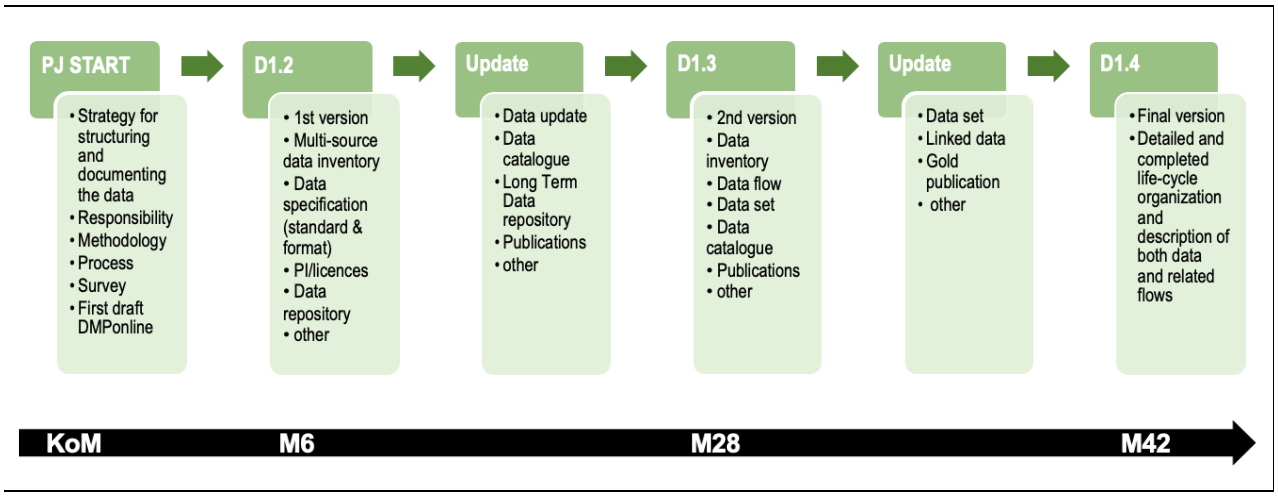

Figure 2: DMP versions and deliverables

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. If during the project life cycle, a new dataset is identified, partners can submit a new form through SINTETIC Teams platforms, notifying the PMT. CNR will then 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.

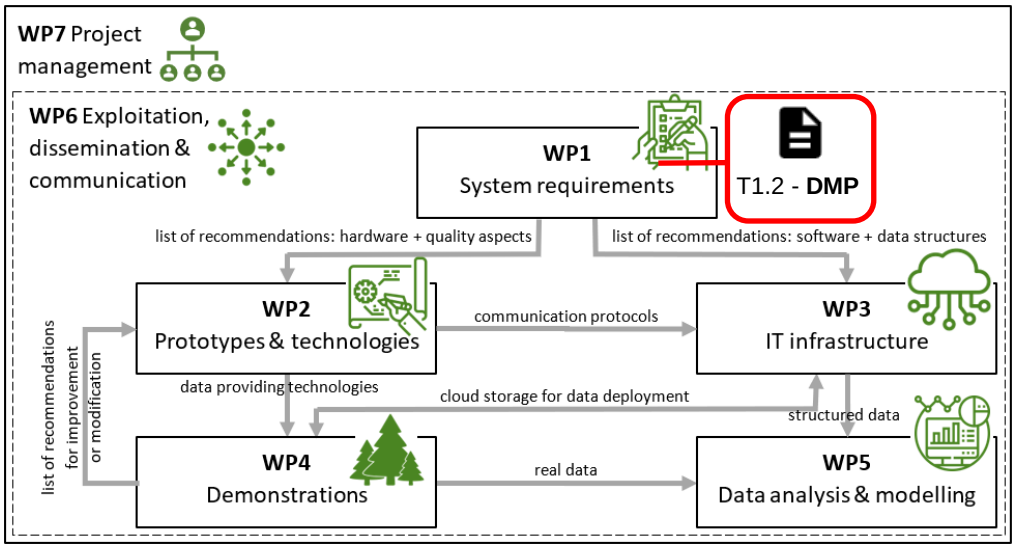

Figure 3: SINTETIC WPs structure

Document structure

The document structure we have outlined follows the Horizon 2020 Data Management Plan template[2], 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:** This section outlines the intended purpose of re-using existing data and covers the types, formats, and anticipated sizes of generated or re-used data. It also includes information about data origin, provenance, and potential utility beyond the project's life cycle.
2. **FAIR Data Principles:** Focusing on the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR), this part underscores the importance of making data easily discoverable by using identifiers, metadata, and search keywords. It highlights the need for accessible trusted repositories for both data and metadata, and the use of vocabularies, standards, and formats to enhance interoperability and promote data reusability. It also emphasizes the significance of comprehensive documentation for facilitating data re-use.
3. **Allocation of Resources:** This section assesses the costs of making data FAIR, including considerations like storage, archiving, etc. It identifies the responsible entities for managing data and ensures strategies for long-term preservation.
4. **Data Security:** This segment addresses data security provisions, recovery mechanisms, secure storage, and transfer protocols. It emphasizes the use of trusted repositories to ensure long-term data preservation.
5. **Ethical Aspects:** This part identifies ethical or legal implications related to data sharing. It references specific ethics deliverables and the ethics chapter within the Description of Action (DoA).
6. **Other Research Outputs:** This section covers digital outputs such as software, workflows, models, etc., as well as physical outputs, including new materials and samples.

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

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.

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 this initial 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 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 and in many different ways[3], 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 and documented data to set up and demonstrate a traceability system for trees, logs and boards based on ICT through a central Geodatabase 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, which can be categorized (Table 1) based on the data flows (Figure 4) from different phases of forest traceability system (GA_DoA).

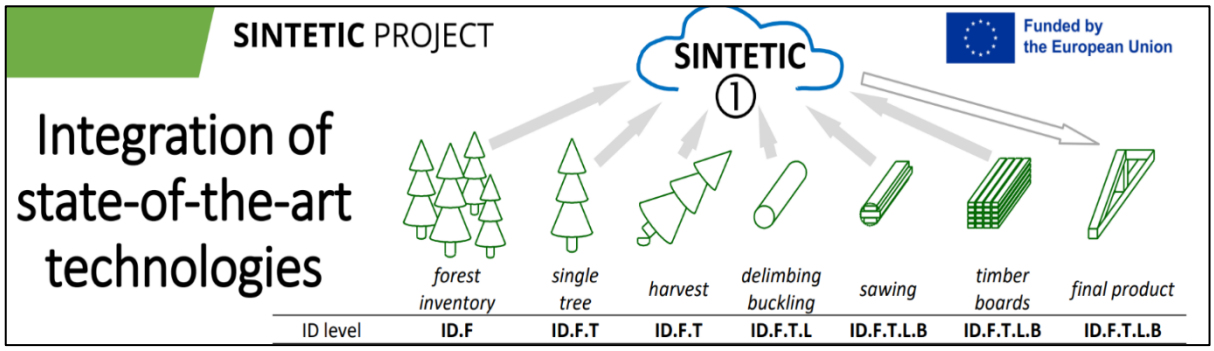

Figure 4: Data flow from different phases of forest traceability system

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

Each category aims to provide specific details about the datasets, their characteristics, and their context for better understanding and utilization within SINTETIC Project.

Table 1 Overview of data categories for the forest product traceability system

DATA CATEGORY	DESCRIPTION
Standing Tree	- Field Measurements: Data of physically measured trees - Data from Modelling: Taper/volume models and inventory inference data - Remote sensing and remote measurements
Bucking Simulation	- Data generated from breaking down trees into logs during bucking simulation
Harvest & Bucking	- Processor Head Info - Manual Felling and Bucking Info
Harvest (Roadside)	- Forwarder Info - Roadside Measurements
Sawmill (Storage)	- Entry Dockets - Yard Inventory
Sawmill (Saw)	- Saw Machines (+sensors) - Data
Timber Boards	- Data related to storage of timber boards
Final Product	- Uncertain if information will be available for this category

A further data aggregation has been done by macro-phases of traceability system as reported in Table 2 and used to collect information from internal partners involved in one or more of these stages.

The project partners filled out questionnaires online to furnish information on the current data and technologies being used, any obstacles they were facing, and their future expectations. These questionnaires address both general information and technical data descriptions, ensuring that data within the SINTETIC platform will be well-documented, reliable, and secure. The questionnaire results (Annex I) provided valuable insights for DMP and other tasks, including system specifications (WP1 T1.3;) and geodatabase implementation (WP3 T3.1 T3.2).

Table 2 Main sector selected for the first internal SINTETIC data survey

Traceability phases (questionnaire)	Significant answers	Organization Types	Partners	Tasks involved
Inventory Data	*8*	□ Research Institution □ Forest Management □ Harvesting / Haulage Contractor □ Private Forest Owner □ Public company/app creator □ National Ass.private forest company &contractors	□ UEF □ ARBOREAL □ BLUEB □ ASEMFO □ SILVA □ TREE □ CNR	T1.1 T1.2, T1.2, T2.1, T2.2, T2.3, T2.4, T3.2, T3.3, T3.4, T4.1, T4.2, T5.2, T5.3, T5.4 T6.2, T6.3, T6.4 T7.1, T73

Traceability phases (questionnaire)	Significant answers	Organization Types	Partners	Tasks involved
			□ UNITBV	
Harvesting Data	*5*	□ Research Institution □ Harvesting / Haulage Contractor □ Forest Management □ Machinery Manufacturers □ Manufacturers	□ UEF □ INNO □ OTME □ SIMTRO □ TREE	T2.1, T2.2, T2.3 T3.3 T4.1 T.5.1
Sawmill Data	*3*	□ Harvesting / Haulage Contractor □ Machinery Manufacturers □ Manufacturers	□ SIMTRO □ OTME □ UEF	T2.1, T2.3, T2.4 T3.3 T4.3
Remote sensing	7	□ Research Institution □ Manufacturers □ Harvesting / Haulage Contractor	□ CNR (2 Dep.) □ LaMMA □ UEF □ SIMTRO □ TREE □ UNITBV	T1.2, T1.3, T2.1, T2.2, T2.3, T2.4, T3.2, T3.3, T4.1, T4.2, T4.3, T5.2, T5.3, T5.4 T6.1, T6.2, T6.3 T7.1, T73

In this Deliverable first version (D1.2) partners' answers are included synthetically in the different section of this report referring to the data that will be collected through the timber traceability phases. Remote Sensing Data are considered as a valuable data source in forest inventory, illegal logging analysis, land cover changes, climate characterizations so a specific questionnaire have been submitted to internal partners to evaluate data volume and the consistency of Big Data that will be used.

In the majority of SINTETIC research activities (WP4), the partners will tend not to re-use existing data, due to the need to address specific project questions (new prototypes and sensors), but rather to carry out ad hoc measurements (WP2, WP4) for generating the needed information on digital and integrated forest traceability system. Instead, 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. ECWF and NOOA); Digital Elevation Model; Infrastructure's maps from national agencies providing data; Environmental maps; Forest cover land use maps; Forest inventory, National Forest inventories (IFN), ecosystem data, cadastral information, sectorial studies.

One of the primary objectives of DMP is to establish a project data inventory, continuously updated during the project cycle. During the process of creating dataset (WP2, WP4, WP5) curators are asked to provide or update the Data Inventory Table (Annex II), providing a list of the datasets and expected sizes produced or being produced within WPs' activities. The updated Inventory Tables will be annexed to the new versions of DMP (D1.3, D1.4) and used as reference for implementing the general structure and data organization of the SINTETIC data catalogue (CKAN[4] catalogue) of geodatabase (T1.2, T3.1).

The next version of DMP will include information about the expected size of datasets and details on geodatabase modelling (T1.3). During this initial phase, approximately 2 Terabytes of space have been allocated for virtual machines that will be used for data storage and the SINTETIC geodatabase.

[4] https://ckan.org

Data format

SINTETIC will generate higher-level datasets based on the sources and WPs' activities. These datasets will be considered as SINTETIC products. The partners who have contributed to each product will be shared owners of the products. Table 3 summarizes the different formats actually used by SINTETIC partners. The formats used in internal data storage and processing are separated from those suggested for facilitating re-use through deposit in trusted repositories.

Table 3 Summary of data formats used in SINTETIC

Type of data	Current data formats	Formats suggested for sharing reuse and preservation
Numerical or \textual tabular data (SPREEDSHEET FILE)	Microsoft Excel (.xls/.xlsx/ODS/)	Comma-separated values (.csv)
Textual data	Microsoft Word (.doc/.docx)	Plain text data, ASCII, Rich Text Format (.rtf) or text (.txt) Documentation will be provided as PDF
Database	DB, sqlite, Pikle, PostGreSQL	
Raster data	Hdf5, NetCDF4, GeoTiff, GRIB, ECW, ERDAS Imagine, JPEG2000, ESRI Grid,WMS, GeoRef. raster images (jpeg, BMP)	Hdf5, NetCDF 4, GeoTiff,
Vector data	ESRI shapefile, DBF, GeoJSON, KML, KMZ, WFS, Spatialite,GeoPackage, GML,	ESRI shapefile, DBF, GeoJSON, KML, GML
Images	PNG, JPEG, BMP	PNG, JPEG,
Lidar	Point cloud (LAS, ASCII, XYZ, PLY)	Point cloud (LAS, ASCII, XYZ, PLY)
Others formats (Data harvesters /sensors)	StanForD (e.g. pri, etc.)	StanForD 2010 (new release)

Depositing data and documentation in formats preferred for archiving and re-use can make the processing and release of data faster and more efficient. Preservation formats should be platform-independent and non-proprietary to ensure they will be usable in the future.

Data origin and purpose

SINTETIC traceability system will be bases on ICT technologies and geomatics tools to generate the data identification and their flow from sensors used in whole timber value chain.

Data are provided from the different sensors by machines/app/operators on forest stands, trees, logs and sawnwood in the value chain. The WP2, WP4 and WP5 are the packages involved in collects, processing, analyse data for the deployment and test of SINTETIC traceability system.

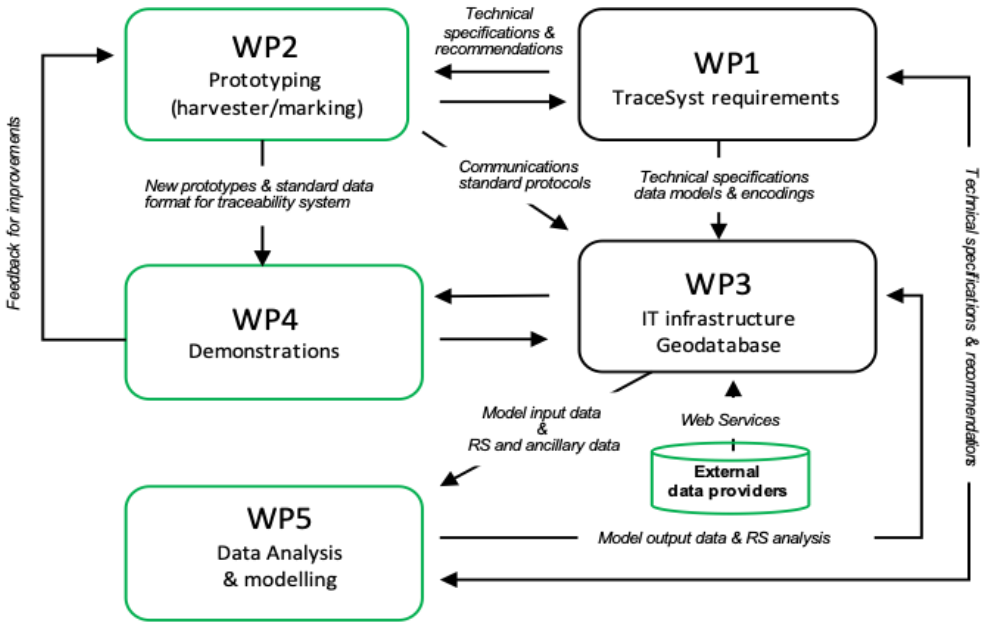

Figure 5 WPs involved in data provisions (green) and data flow in SINTETIC Geodatbase following the WP1 specifications.

The prototypes developed in WP2 are devoted to data collect along the forest traceability chain devoted to delivery a Marking and tracking system for trees and timber products (D2.3); Smartphone-APP for timber measuring, grading and tracking (D2.4); The prototype of forest harvester with tracking functions and quality sensors (D2.5) and a Quality assessment and traceability systems in sawmill (D2.6).

The SINTETIC prototypes after deployment will be extensively tested during series of demonstrations (WP4). All data collected from prototypes' tests (WP2) and Demos (WP4) gradually will populate the data storage (unstructured data in their native format) and SINTETIC geodatabase (WP3) (structureted data[5]) following the guidelines, specifications defined in WP1. Data from ground truthing measurements (e.g. with traditional tools on logs) collected in the 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).

For this first realise of DMP a descriptive origin of data in each WP2, WP4 and WP5 tasks is reported:

Task 2.1 - ID Detection Systems – *(the optical/electronic system marking and tracing system)*

- Collection of data related to marking trees and processed roundwood with unique IDs readable via digital sensors.
- Data on the development of manual and machine-based marking systems, including portable and fixed systems for automatic identification throughout the value chain.
- Data on ID detection in sawmills to identify roundwood using punched codes or RFID tags.
- Pre-test results, characteristics, and operative feedback from demos and upgrades applied in M24 (D2.1) and M36 (D2.2).

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

Task 2.2 - Smartphone APPs for Manual Operations

- Data collected through a smartphone app integrated with log ID tracking service.
- Usage data of built-in smartphone sensors like color cameras, LiDAR, GPS for forest inventory, stem geometry, and roundwood volume measurements.
- Selection and testing data of portable sensors for timber grading and early wood quality assessment.
- Validation data comparing in-field derived data with laboratory measures.
- Flexible database data customization, including environmental services and forest certification data.

Task 2.3 - Forest Harvester Prototype - *the harvester equipped with LiDAR and sensorized processor head (with marking system and wood quality sensors)*

- Data on the development of a functional prototype of an innovative forest harvester with timber quality sensors and LiDAR.
- Data on the integration of sensors with customized software for real-time data processing, providing value recovery suggestions and data reduction.
- Solutions for data transfer from harvester to the SINTETIC Geodatabase.

Task 2.4 - Sawmill Sensors

- Collection of quality-related data at various stages of log transformation in sawmills.
- Data on the integration of roundwood identification technologies with pilot installations in partner sawmills.
- Development and testing of a prototype system for all-boards traceability.
- Integration of image-based tracking system with single-item ID marking.
- Cross-linking diverse databases associated with the Geodatabase.

Each task involves the collection of specific dataset related to the development, integration, testing, and feedback of prototypes and systems within the SINTETIC project's objectives.

Table 4 will be continuously updated during the project life with data specification that are registered during test and demo.

Table 4 Data set expected from new prototypes of SINTETIC traceability system.

Task	Dataset expected	Description (included attributes)	Metadata schema (filled standard adopted)
T2.1. ID detection systems			
T2.2. Smartphone APPs for manual operations			
T2.3. Forest harvester prototype			
T2.4. Sawmill sensors			

Task 4.2 - Demo Execution

- **Forest and Industry Demonstrations**:
  - Forest characterization pre-harvest using traditional methods and SINTETIC Smartphone APP.
  - Marking trees/logs, interrelating with the database for quality, process, and descriptive data.
  - ID detection of roundwood entering industrial transformation and tracking boards throughout the process.

- Deliverables: Documenting demo evaluations and feedback (D4.3, M36) and Final Report (D4.4, M45).

Task 4.3 - Enhancement of Forest Value and Management

- **Integration of Platforms**:
  - Deployment of digital tools in Italy and Spain for forest inventory, harvest operations, product tracking, and revenue distribution among landowners.
  - Demonstration of a citizen-science mobile APP for wildfire risk assessment and digital timber product certification.
  - Development of a user-friendly tool for forest owners and video tutorials.
  - Development of a dedicated section in the Forest Sharing Platform addressing forest management attitudes and FAQ section based on consortium know-how and stakeholder interactions.

Each task in Work Package 4 involves planning, execution, and enhancement of forest management through demonstrations, digital tools deployment, and integration between platforms, aiming to optimize forest inventory, management, and stakeholder engagement.

Table 5 Data set expected from WP4

Task	Dataset expected	Description (included attributes)	Metadata schema (filled standard adopted)
T4.2 Demo Execution			
T4.3 Enhancement of Forest Value and Management			

Task 5.1. Wood quality modelling: boards-roundwood-tree interrelations

- Individual-tree wood quality (WQ) models output data (timber quality, tree structure, wood anatomy, yield and characteristics of final boards).

Task 5.2. Mobile LiDAR data analysis for remaining trees

- LiDAR data for inventory of remaining trees
- LiDAR data to identify the occurrence of deadwood
- LiDAR data for CO2 inventory
- Terrestrial Laser scanning inventory (D5.2)

Task 5.3. Illegal logging alert service (demo site RO)

- Satellite remote sensing datasets (Copernicus Sentinel 1&2)
- Synthetic Aperture Radar (SAR)
- Forest cover changes

Table 6 Data set expected from WP5

Task	Dataset expected	Description (included attributes)	Metadata schema (filled standard adopted)
T5.1 Wood quality modelling			

Task	Dataset expected	Description (included attributes)	Metadata schema (filled standard adopted)
T5.2 Mobile LiDAR data analysis for remaining trees			
T5.3 - Illegal logging alert service			

Sensors Data

The sensors used in the forestry sector, particularly in line with Communication Technology (ICT) digital technologies, encompass a range of specialized tools. As investigated in D1.1 the sensor types used in forest value chain are described in Table 7.

Table 7 Sensor type used in forest sector

Sensor type	Description
Satellite-Borne Sensors	Devices utilizing data captured by Earth-orbiting satellites. They gather extensive forest information on a large scale, monitoring territories and conducting extensive forest inventories. The sensors capture various data, such as vegetation health, land cover changes, and forest density, aiding in comprehensive analysis and management.
Light Detection and Ranging (LiDAR) Sensors	Versatile tools deployed on aerial carriers (like drones or airplanes) or as Terrestrial Laser Scanners (TLS). They emit laser pulses and measure their return time after hitting objects, providing highly detailed 3D maps of forested areas. This aids in precise measurements of tree heights, canopy structures, and terrain modeling.
Harvesting Machinery Sensors	Integrated into harvesting machinery, these sensors play a pivotal role in positioning and measuring logs. They generate digital data in the StanForD format, a standardized format used in forestry. This data is crucial for various purposes including invoicing, operational management, and inventory tracking.
Optical and X-ray Sensors	Predominantly utilized in sawmill industries, these sensors are used for assessing timber characteristics and optimizing the sawing process. Optical sensors analyze wood properties like density, moisture content, and quality, while X-ray sensors provide detailed internal structural information, enabling efficient wood processing and reducing waste.

These sensors collectively form an integral part of modern forestry practices, aiding in sustainable management, efficient operations, and resource optimisation. One of the SINTETIC objective is to transform/modify the StanForD format for all sensors used in timber traceability systems to enhance the interoperability of data derived from each stage of value chain.

Sensor data for the traceability system will be modelled in the Geodatabase (T1.3, T3.2, T3.3) to enhance integration of different data sources at different temporal and spatial scale and support the development of processing services (T3.2).

Data from external sources

Table 8 External data

External data	Source	URL
Cartographic data	Data provided by the "Catastro" (National Land Register/Cadastre) and the "SIGPAC" (Geographic Information System of Agricultural Plots) produced by the "MAPA – Ministerio de Agricultura, Pesca y Alimentación" (Ministry of Agriculture), and by the "PNOA - Plan Nacional de Ortofotografía Aérea" (National Aereal Orthophotography Plan) from the "IGN-Instituto Geográfico Nacional" (National Geographic Institute) of Spain.	https://www.sedecatastro.gob.es/ https://www.mapa.gob.es/es/agricultura/temas/sistema-de-informacion-geografica-de-parcelas-agricolas-sigpac-/visor-sigpac.aspx https://pnoa.ign.es/
Other data	Inventario Nacional Forestal - IFN" (Forest National Inventory) carried out by "MITECO - Ministerio para la Transición Ecología y el Reto demográfico" (Ministry of Environment), every 10 years, the last one been the IFN4, which started in 2008. El "Mapa Forestal de España - MFE" (Spanish Forest Map) at 1: 25.000 scale is the cartographic basis of the IFN4 inventory	https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales https://www.miteco.gob.es/es/cartografia-y-sig/ide/
Traceability tools	FORTRA (Forest Traceability) x Galician timber sector blockchain technology	https://fortra.xunta.gal/fortra/paxina/fortra?_referer=%252Ffortra&locale=en
Documents	Estrategia Forestal Española Horizonte 2050 - EFE" (Spanish Forestry Strategy Horizon 2050), published by MITECO in 2022.	https://www.miteco.gob.es/en/biodiversidad/temas/politica-forestal/planificacion-forestal/politica-forestal-en-espana/pfe_estrategia_forestal.html
Satellite data	Sentinel 1 C-SAR (C-band Synthetic Aperture Radar)	https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/data-products https://www.sentinel-hub.com/

Data utility

After conducting an external data survey (T.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:

- **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 like the Open Access Infrastructure for Research in Europe (OpenAire) to allow for dissemination and validation and to increase the potential of research results' use. 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. This will help both humans and computers easily locate the data they need.

During the project, the SINTETIC geodatabase will be used to archive and manage data. However, when the project is completed, each research team must deposit and describe their datasets in identified public data repositories (Trusted Repository) and their Institutional repositories, if available, 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 find the appropriate trusted repository where to deposit the project research datasets between two typologies: 1) OpenAire compliant repository registered in OpenDOAR, re3data and FAIRsharing 2) Zenodo.

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 Zenodo[6] are compliant with DataCite[7] Metadata Schema 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,

6 https://zenodo.org/

7 https://schema.datacite.org/

must support Dublin Core[8] and DataCite Metadata Schema. At the end of the project, datasets will be deposited in the trusted repository Zenodo and made visible through the OpenAIRE[9] portal.
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 repots, guidelines, papers, diagrams, etc., will be archived in the trusted repository and assigned a DOI.
The DMP, in upcoming planned releases, will identify standard rules for naming and versioning each datasets category to improve data visibility, discoverability, citation, and permanent online tracking. In this first version, only a general recommendation for WPs datasets file naming is reported to facilitate the internal dataflow (input/output of each task), system specifications and final products identification (specific datasets are considered as final products).
• For dataset file(s)
DATASET_SINTETIC_WPn_Tn.m_ content specifications_date (YYYYMMDD)_vn.file extension.

• Associated to datasets, relevant documentation explaining data collection procedures and analysis (such as codebooks, users' manuals, methodologies, etc.) are provided in the form of a human readable README file.
README_SINTETIC_WPn_Tn-m_ content specifications_date (YYYYMMDD)_vn.file extension.

Where WPn means "work package number" Tn.m is the "task number", and ver identifies the "version number" (in case of data revisions or updates)

Table 9 Discoverability level

Findable	How	ID	Where
Data repository	Re3data.org	Data repository DOI	Trusted repository
Data sets		Data set DOI	(Zenodo, DataCite, Institutional data rep., thematic repositories)
Metadata	Standard Metadata schema (e.g. DataCite, ISO19115)	ID	DATA Catalogue
Other data (publication, technical documents, user guide, manuals)	Standard Metadata schema	DOI	(Zenodo, DataCite, Institutional data rep., thematic repositories)

8 https://dublincore.org/
9 https://www.openaire.eu/

Metadata are not usually edited by data producers in forestry traceability chain (ref: questionnaire answers) and no standard schema are applied during the inventory phase to describe in detail the data. 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. Standard metadata schema (e.g. DataCite metadata schema, ISO19115, etc.) will be adopted for different data generated and managed by geodatabase and will be explored using an open source a data catalogue (e.g. Ckan[10] data catalogue) that will be implemented (T.1.2 T3.1) within the project.

The RDA Metadata Standard Catalogue[11] will be the reference in the choose 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[12].

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 data access restrictions or sharing is necessary. The principle of "as open as possible, as closed as necessary" will be considered in such cases.

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.

10 https://ckan.org/

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

12 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

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 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 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), that can be registered as an institutional data repository in re3data.org. if compliant with the minimum requirements.
Accessibility level to data server will be defined in T3.1 where will be specify 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 be guarantee Gold OA Publications deliverables. Finally, they adopt descriptive metadata standards as required by the OpenAIRE Guidelines and allow cross-linking between 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 dataset and related publications can also deposit in the trusted repository as Zenodo or other thematic repositories findable on OpenAIRE platforms so that any issues surrounding the usability of the data can be resolved. Generally, Zenodo can be recommended for open dissemination and preservation of research data by all research teams that do not have suitable institutional, national, or disciplinary data repositories or are not bound 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[13] to enhance semantic discoverability. At this stage, the Consortium will refer to DCAT[14] 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.

The DMP will be 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.

[13] *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.*

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

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 8 illustrates the three 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 10 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 registered as data repository in re3dataorg, through the whole duration of SINTETIC and will guarantee could be also 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 maintain the data repository and geodatabase 5 years after the end of the project.
However, there will be no charges for depositing and preserving publicly shareable data in the long term, as the selected repositories, such as Zenodo, do not impose fees for archiving and data curation and 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 data server and updates the configuration and web services.
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 ORCID22. 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 includes 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.
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.
Detailed involvements of partners and definition specific focus groups (e.g. licence, IP, restrictions …)

CONCLUSIONS

This first 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 (D1.3) in M28.

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

D1.2 Annexes

Annex I – Report on SINTETIC internal questionnaires

INVENTORY

Partners' response	**8**	UEF	Harvesting / Haulage Contractor
		Arboreal	Public company - app creator
		Bluebiloba Startup Innovativa SRL	Forest Management
		ASEMFO National Association of private forest companies and contractors	National Association of private forest companies and contractors
		Silvador Company	Private Forest Owner
		Treemetrics	Forest Management
		CNR - National Research Council	Research Institution
		Transilvania University of Brasov	Research Institution

	Nr. answers [15]	**Answers**
Overall Inventory Reports	7	- Field measurements (e.g., DBH, Heights) - More than once per month - More than 5 years - NA - Yearly - Yearly - More than once per year - More than once per month
Tree Measurements	6	- More than once per month - More than 5 years - NA - Yearly - Yearly - More than once per year - More than once per month
Single Tree Location	5	- More than once per month - NA - NA - Yearly

[15] significant responses

	Nr. answers 15	Answers
		- Yearly - More than once per year - More than once per month
Single Tree ID	5	- More than once per month - NA - NA - Yearly - Yearly - Yearly - More than once per month
Timber Quality Information	3	- More than once per month - NA - NA - Yearly - NA - NA - More than once per month
Point Clouds from Terrestrial Sensors	6	- More than once per month - More than 5 years - NA - Yearly - Yearly - More than once per year - More than once per month
Aerial or Satellite Images to Determine the Forest Canopy Area	6	- More than once per year - Every 2-5 years - NA - Yearly - Yearly - More than once per year - More than once per month
Tree-Level Remote Inventory: Aerial or Satellite Data to Estimate the Tree Size	5	- More than once per year - NA - NA - Yearly - Yearly - More than once per year - More than once per month
Stand-Level Remote Inventory: Aerial or Satellite Data to Estimate the Stand Volume	4	- More than once per year - NA - NA - Yearly - NA - More than once per year - More than once per month
Other Inventory Data	3	- ecosystem data (biodiversity, water regulation, etc.), user data, cadastral information data - Please see the information included in Comments - data from the drone flight
Dataset Size	7	- 10 gb, mainly point clouds - About 1Gb excluding imaging data from sensors (Lidar, aerial, satellite, etc.) - Please see the information included in Comments - xlsx, less than 10 MB

	Nr. answers 15	Answers
		- Around 10 Mb per inventory - 10Mb - Difficult to say at this time.
Data Format.	7	- Static documents (e.g., PDF) - Database (DB, sqlite), Images (e.g., PNG, JEPG, etc) - Static documents (e.g., PDF), Database (DB, sqlite), TIF, ECV, point clouds - NA - Text files (e.g., txt, csv, tsv, etc), Spreadsheets files (e.g., XLS, XLSX, ODS, etc.), Documents (e.g., DOC, DOCX, etc), Database (DB, sqlite), shp files - Text files (e.g., txt, csv, tsv, etc), Spreadsheets files (e.g., XLS, XLSX, ODS, etc.), Database (DB, sqlite) - Text files (e.g., txt, csv, tsv, etc), Spreadsheets files (e.g., XLS, XLSX, ODS, etc.), Documents (e.g., DOC, DOCX, etc), Database (DB, sqlite), Pickle, Spreadsheets files (e.g., XLS, XLSX, ODS, etc.), Database (DB, sqlite)
Origin of the Data	8	- Manual measurements or analysis - The data are provided by the client - Manual measurements or analysis, Automated collection by remote sensors or devices, The data are provided by the client, From public data repository - From public data repository, Please see the information included in Comments - Manual measurements or analysis, Automated collection by remote sensors or devices - Manual measurements or analysis, Automated collection by remote sensors or devices, The data is collected by an external company on demand, The data are provided by the client - Manual measurements or analysis, Automated collection by remote sensors or devices - Manual measurements or analysis, Automated collection by remote sensors or devices, The data are provided by the client
Data Structure	8	- K - Sample plots with individual tree data. Everything is connected to a sample plot. - https://drive.google.com/file/d/1-uOa06qMfOLT8rtYKjRD2P9pdoMkSlaT/view?usp=sharing - Please see the information included in Comments - http://www.mmediu.ro/app/webroot/uploads/files/Manual%20Sumal%20Avize%20-%20Agenti.pdf - https://docs.foresthq-en.treemetrics.com/2-measurement-and-valuation/2-3-measurement-valuation-results - Plot_id ID Stump ID Tree Species Tree type Social Pos. Deadwood type Decay class DBH Height (m) Tree microhabitats (TreMs) - Difficult to say at this time.
Metadata	7	- No - No - Please see the information included in Comments - Yes - We do not follow a metadata standard, however it would be to store information about, date of measurement, person who measured, date of analysis and date of update - No

	Nr. answers 15	Answers
		- Yes but not standardized.
Data Level of Processing	8	- Raw data from sensor or measurements - Raw data from sensor or measurements, Data aggregation (report), Data generated from models or algorithms - Raw data from sensor or measurements, Data aggregation (report), Data generated from models or algorithms - Please see the information included in Comments - Data generated from models or algorithms - Raw data from sensor or measurements, Data aggregation (report), Data generated from models or algorithms - Raw data from sensor or measurements, Data aggregation (report), Data generated from models or algorithms, Dataviz and InfoGraphics - Raw data from sensor or measurements, Data aggregation (report), Data generated from models or algorithms
Sintetic Tasks	6	- Task 5.2 - Task 2.2 - Task 1.1, Task 1.2, Task 1.3, Task 2.1, Task 3.1, Task 3.2, Task 3.3, Task 4.1, Task 4.3, Task 5.4, Task 6.2, Task 6.3, Task 6.4, Task 7.1, Task 7.3 - NA - NA - Task 2.2, Task 4.3 - Task 3.2, Task 5.3 - Task 1.2, Task 1.3, Task 2.1, Task 2.2, Task 2.3, Task 2.4, Task 3.2, Task 4.1, Task 4.2, Task 5.3, Task 5.4
New Datasets	7	- U - Pointcloud data - XYZ dataset - Integration of foreign country data (outside of Italy) - Please see the information included in Comments - S - NA - A updated dataset of inventory data - Not at this time.
Visualization	8	- Table - Map - Map - Please see the information included in Comments - I do not want to visualise this data in SINTETIC website - Table, Map - Text, Table, Map, File - Text, Table, Map, File
Required Level of Processing	8	- Only storage and/or visualisation - Data to be used as input to models or algorithms in more complex analysis - Only storage and/or visualisation, Data aggregation, reporting and/or filtering - Please see the information included in Comments - Data to be used as input to models or algorithms in more complex analysis - Only storage and/or visualisation, Data aggregation, reporting and/or filtering - Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis, No further storage or processing is required

	Nr. answers 15	Answers
		- Only storage and/or visualisation, Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis
Download	7	- CSV - TIFF, WMS/WFS, SQL - Please see the information included in Comments - Yes - CSV - Csv - XLS, CSV, LiDAR-based files
Data Access	8	- Access restricted to my organization only - Open to public - Some data restricted to my organization only, some data restricted to the consortium members - Data access restricted to the consortium members - Open to public - Data shared with approved external stakeholders - Open to public - Open to public
Data License	7	- Open data usage license such as CC-By attribution, ODBL, etc - Some data can have an open data license, other data will be private - I do not have a preference - Private Data License Agreements - I do not have a preference - Open data usage license such as CC-By attribution, ODBL, etc - Open data usage license such as CC-By attribution, ODBL, etc
Inventory Devices	8	- I - Mobile devices with application for virtual measurement (E.g. Arboreal, Trestima, etc), Mobile devices for data log - Traditional measurement tools (callipers, hypsometer, tapes), Measurment sensors (e.g. terrestrial laser scanner), Mobile devices for data log - Please see the information included in Comments - Traditional measurement tools (callipers, hypsometer, tapes), Internet of Things (IoT) devices, Mobile devices with application for virtual measurement (E.g. Arboreal, Trestima, etc), Mobile devices for data log - Traditional measurement tools (callipers, hypsometer, tapes), Measurment sensors (e.g. terrestrial laser scanner), Internet of Things (IoT) devices, Mobile devices with application for virtual measurement (E.g. Arboreal, Trestima, etc) - Traditional measurement tools (callipers, hypsometer, tapes), Measurment sensors (e.g. terrestrial laser scanner), Dedicated GPS devices - Traditional measurement tools (callipers, hypsometer, tapes), Measurment sensors (e.g. terrestrial laser scanner), Mobile devices with application for virtual measurement (E.g. Arboreal, Trestima, etc), Mobile devices for data log, Dedicated GPS devices
Inventory Devices Description	7	- iPhone 7,8,X, 11, 12, 13, 14, 15 Pro and not pro. Ipad of different models - All devices supporting QField - Please see the information included in Comments - IPhone 14 PRO - https://docs.foresthq-en.treemetrics.com/2-measurement-and-valuation/2-3-measurement-valuation-results

	Nr. answers 15	Answers
		- Vertex III fo Height and calliper for DBH. - Haglof, Arboreal, ZebRevo, iPhone Pro Max
Inventory Traceability	8	- U - Device providing single tree identification and location (GNSS) - Device providing single tree identification and location (GNSS) - Please see the information included in Comments - Paint marking, Device providing single tree identification and location (GNSS), Wood marking hammers - Paint marking, Device providing single tree identification and location (GNSS), Device with RFID or other electronic tags - Paint marking, Device providing single tree identification and location (GNSS), Wood marking hammers - Paint marking, Wood marking hammers
Inventory Cost	8	- 50-100€/ha - 0-50€/ha - 0-50€/ha - NOT APPLY - 0-50€/ha - 50-100€/ha - 100-150€/ha - >150€/ha
Devices Cost	8	- 100-500€/year - >1000€/yea - 0-100€/year - NOT APPLY - 100-500€/year - 0-100€/year - 500-1000€/year - >1000€/yea
Comments	1	- Our response to the above requested information is as follows: 1. Our associated forest companies and contractors use the following as a source of information, to carry out forestry management works and projects: • Cartographic data Data provided by the “Catastro” (National Land Register/Cadastre) and the “SIGPAC” (Geographic Information System of Agricultural Plots) produced by the “MAPA – Ministerio de Agricultura, Pesca y Alimentación” (Ministry of Agriculture), and by the “PNOA - Plan Nacional de Ortofotografía Aérea” (National Aereal Orthophotography Plan) from the “IGN-Instituto Geográfico Nacional” (National Geographic Institute). The accesses are shown here: https://www.sedecatastro.gob.es/ https://www.mapa.gob.es/es/agricultura/temas/sistema-de-informacion-geografica-de-parcelas-agricolas-sigpac-/visor-sigpac.aspx https://pnoa.ign.es/ All this information is transferred and implemented in tools through the “GIS – Sistema de Información Geográfica” (Geographic Information System)

	Nr. answers 15	Answers
		• Rest of the data/information Provided by the “Inventario Nacional Forestal - IFN” (Forest National Inventory) carried out by “MITECO - Ministerio para la Transición Ecología y el Reto demográfico” (Ministry of Environment), every 10 years, the last one been the IFN4, which started in 2008. El “Mapa Forestal de España - MFE” (Spanish Forest Map) at 1: 25.000 scale is the cartographic basis of the IFN4 inventory. The accesses are shown here: https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales https://www.miteco.gob.es/es/cartografia-y-sig/ide/ 2. Some of our associated companies and forest administrations of some Autonomous Communities have developed or are developing digital tools and Apps, for forestry management works and projects, based on the information sources just mentioned, together with other data from field work, specific to the work being carried out. 3. A project to be highlighted is FORTRA (Forest Traceability) a digital tool promoted recently by the Xunta de Galicia that is made available to forest companies and consumers with the aim of providing traceability to wood products, recording the operations carried out by all the companies that form part of the transformation process, from the forest to the market product. This Forest Traceability tool, with blockchain technology, offers a network to digitally connect the entire value chain of the Galician timber sector. The access is shown here. As well in English https://fortra.xunta.gal/fortra/paxina/fortra?_referer=%252Ffortra&locale=en 4. ASEMFO, to carry out reports and projects related to the forestry sector, uses the same database sources (Public Data Repository), mainly the National Forest Inventory (IFN); and sectorial studies, such as the “Estrategia Forestal Española Horizonte 2050 - EFE” (Spanish Forestry Strategy Horizon 2050), published by MITECO in 2022. The access is shown here: https://www.miteco.gob.es/en/biodiversidad/temas/politica-forestal/planificacion-forestal/politica-forestal-en-espana/pfe_estrategia_forestal.html

HARVESTING

Partners' response	**5**	UEF	Harvesting / Haulage Contractor
		Innorenew CoE	Research Institution
		Simtrona d.o.o.	Manufacturer
		OTMETKA	Machinery Manufacturers
		Treemetrics	Forest Management

	Nr. answers [16]	**Answers**
Harvester Machine Production Data	5	- Daily - Weekly - Real-time - Real-time - Daily
Harvester Machine Products Specifications	5	- Daily - Weekly - Real-time - Real-time - Occasionally
Harvester Machine Calibration Data	5	- Weekly - Once per operation - Daily - Real-time - Occasionally
Other Harvester Machine Standford Data	5	- Daily - Weekly - Real-time - Real-time - Never
Fowarder Files	5	- Weekly - Weekly - Real-time - Real-time - Never
Roadside Measurements	5	- Weekly - Weekly - Real-time - Real-time - Never
Other Harvesting Data	1	- GPS

[16] significant responses

	Nr. answers 16	Answers
Dataset Size	4	- Don't know. - 100Mb - Unknown - 1mb/ha
Data Format	5	- StandForD2010 - Not using SandForD - Not using SandForD - StandForD2010 - StandForD2010
Data Structure	0	
Data Ownership	5	- Data owned by my organisation - Data owned by 3rd party - Data owned by my organisation - Depends on the contract. - Data owned together with 3rd party
Harvester Machines	5	- Scorpion - Currently, we have no harvester. Planning to use SP Maskiner Harvester head SP 661 LF - ? - Eco Log 590 and Rottne H21 - We support any brand using StandForD2010 or prior StandForD standard
Tree Marking	5	- No system in place - No system in place - Tree/Log identification - No system in place - Tree coordinates
Traceability System In Harvesting Machine	5	- No system in place - No system in place - Tree/Log identification - No system in place - Other
Manual Logging	5	- No system in place - No system in place - Tree/Log identification - No system in place - Tree coordinates
Traceability System In Forwarder	5	- No system in place - No system in place - Tree/Log identification - No system in place - Other
Roadside	5	- No system in place - No system in place - Tree/Log identification - Other - Other
Transport	5	- No system in place - No system in place - Tree/Log identification - Other - Other

	Nr. answers 16	Answers
Other Harvesting Traceability	2	- Planning to implement solutions within the project. manual ID readings
Traceability Cost	1	- 0.6€/log; app 2€/m3
Sintetic Tasks	5	- Task 4.1 - Task 2.1, Task 2.2 - Task 2.1, Task 4.1 - Task 2.1, Task 2.3 - Task 2.3, Task 5.2
New Datasets	5	- J - The datasets should contain parameters to immediately help to increase yield while bucking on the side, and provide good value while tracking the logs at the log yard and sawmill. - Log ID, Log properties - OTMETKA creates real time database from HPR-files. - Data coming for LiDAR installed in the machine. The data will analysed in situ and transferred using StandForD standard
Visualization	5	- Table - Text, Table, Map, Picture - Text, Table, Map, File - Map, Through our service FindMyTimber - Table, Map
Required Level of Processing	5	- No further storage or processing is required - Data to be used as input to models or algorithms in more complex analysis - Data to be used as input to models or algorithms in more complex analysis - No further storage or processing is required - Data aggregation, reporting and/or filtering
Download	4	- Txt - Csv - xls, csv - XLS, CSV
Data Access	5	- Data shared with approved external stakeholders - Data shared with approved external stakeholders - Data shared with approved external stakeholders - Data access restricted to the consortium members - Data shared with approved external stakeholders
Data License	4	- Open data usage license such as CC-By attribution, ODBL, etc - I do not have a preference - Open data usage license such as CC-By attribution, ODBL, etc - I do not have a preference
Comments	0	

SAWMILL

Partners' response	**5**	UEF	Harvesting / Haulage Contractor
		Simtrona d.o.o.	Manufacturer
		OTMETKA	Machinery Manufacturers

	Nr. answers [17]	**Answers**
Sawmill Log Reception Data	3	- Board measurement, quality and grading - Real-time - Real-time
Sawmill Yard Inventory	2	- Real-time - Real-time
Log Measurement, Quality, Optimisation And Cut Patterns	2	- Real-time - Real-time
Row Board Measurement, Quality And Grading4	2	- Weekly - Real-time
Other Sawmill Data	0	
Dataset Size	2	- 100M - Unknown
Data Format	3	- 2222-11-11 - .csv, xml - Unknown
Data Structure	3	- G - http://ask4tree.simtrona.si/; type in 12345677890123456 - Unknown
Data Level Of Processing	3	- Raw data from sensor or measurements - Raw data from sensor or measurements - Data generated from models or algorithms (e.g., to train models predictions)
Traceability System At Reception	3	- Forest property or location traceability per log pile - No system in place - No system in place

[17] significant responses

	Nr. answers 17	Answers
Traceability System In The Yard	3	- Forest property or location traceability per log pile - No system in place - No system in place
Traceability System In The Production Facilities	3	- Forest property or location traceability per log pile - No system in place - No system in place
Sawmill Traceability	3	- 5 - No - Unknown
Sawmill Traceability Cost	3	- 5 - ? - Unknown
Sintetic Tasks	3	- Task 3.3 - Task 2.4, Task 4.3 - Task 2.1, Task 2.3
New Datasets	3	- G - ID back traceability - Unknown
Visualization.	3	- Table - Charts - Charts
Required Level Of Processing	3	- Only storage and/or visualisation - Data to be used as input to models or algorithms in more complex analysis - Data to be used as input to models or algorithms in more complex analysis
Download	3	- F - Csv - xls, csv
Data Access	3	- Access restricted to my organization only - Data shared with approved external stakeholders - Access restricted to my organization only
Data License	2	- I do not have a preference - Open data usage license such as CC-By attribution, ODBL, etc
Comments	0	

REMOTE SENSING

Partners' response	**7**	UEF	Harvesting / Haulage Contractor
		Simtrona d.o.o.	Manufacturer
		Treemetrics	Forest Management
		CNR - National Research Council	Research Institution
		Consorzio LaMMA	Research Institution
		Transilvania University of Brasov	Research Institution
		CNR	Research Institution

	Nr. answers [18]	**Answers**
Satellite. Low Resolution Images	4	- More than once per year - Never/NA - More than once per month - Never/NA - More than once per month - More than once per month
Satellite. Medium Resolution Images	3	- More than once per year - Never/NA - More than once per year - Never/NA - More than once per month - Never/NA
Satellite. High Resolution Images	4	- Yearly - Every 2-5 years - More than once per year - Never/NA - More than once per month - Never/NA
Drone/Uav Images Taken By External Surveyors	2	- More than once per year - Never/NA - More than once per month - Never/NA - Never/NA - Never/NA
Drone/Uav Images Taken By My Organisation	4	- More than once per year - Never/NA - More than once per month - Never/NA - More than once per month - Yearly
Aerial Images Taken By On Request For My Organisation	3	- More than once per year - Never/NA - Yearly - Never/NA - More than once per month - Never/NA
Aerial Images Available In Public Repositories	4	- More than once per month - Never/NA - Yearly - Never/NA

[18] significant responses

	Nr. answers 18	Answers
		- More than once per month - Yearly
Aerial Images Available In Private Repositories	1	- More than once per month - Never/NA - Never/NA - Never/NA - Never/NA - Never/NA
Aerial Or Uav Lidar Data	3	- Every 2-5 years - Never/NA - More than once per month - Never/NA - More than once per month - Never/NA
Satellite Sar Data	4	- Every 2-5 years - Never/NA - More than once per year - Never/NA - More than once per month - More than once per month
Other Remote Sensing Data	2	- Terrestrial Lidar - MODIS Landsat 8
Digital Terrain Models	2	- NA - NA - NA - Created by my Organisation - NA - Created by my Organisation - NA
Digital Elevation Modes	2	- NA - NA - NA - Created by my Organisation - NA - Created by my Organisation - NA
Distribution Maps	5	- NA - Acquired from an External Provider - Created by my Organisation - Acquired from an External Provider - NA - Created by my Organisation - Created by my Organisation
Infrastructure Maps	4	- NA - Acquired from an External Provider - NA - Acquired from an External Provider - NA - Created by my Organisation - Acquired from an External Provider
Environmental Maps	4	- NA - Acquired from an External Provider - NA

	Nr. answers 18	Answers
		- Acquired from an External Provider - NA - Created by my Organisation - Acquired from an External Provider
Forest Cover Map	5	- Acquired from an External Provider - Created by my Organisation - Created by my Organisation - NA - Created by my Organisation - Created by my Organisation
Vegetation Indexes Map	3	- NA - NA - Created by my Organisation - NA - Created by my Organisation - Created by my Organisation
Land Use Maps	4	- Acquired from an External Provider - NA - Created by my Organisation - NA - Created by my Organisation - Acquired from an External Provider
Climatic Data	4	- NA - NA - Acquired from an External Provider - Created by my Organisation - Created by my Organisation - Created by my Organisation
Other Spatial Data	1	- Climate dataset produced by ECWF and NOOA
Dataset Size	6	- 100Mb - RGB 0.5m image (0.3 Mb/ha), Canopy cover (0.1MB/ha), Other raster info (0.1MB/ha), Forest Boundaries (0.1MB/ha), Strata (0.1MB/ha) - 3 Tb - 7 GB, per year, per 18 variables - Difficult to estimate at this time. - PostGIS Raster Data (actually 1.2 TB ca)
Dataset Structure	2	- GeoTIFF format, usually 0.5m, 16 or 8 bits - https://erddap.lamma.toscana.it/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=bolam+t2m+1994+
Metadata	6	- no - We do not use a metadata standard, but it would be good to store data regarding acquisition data, processing date and sensor - No - https://erddap.lamma.toscana.it/erddap/info/BOLAM_1994_30fe_4ebe_68f5/index.html - Yes, but not standardized. - ISO 19115, DataCite Metadata schema
Level Of Processing	7	- Manual digitalisation - Raw data from sensors (e.g., images) - Raw data from sensors (e.g., images), Manual digitalisation, Basic data processing and feature extraction

	Nr. answers 18	Answers
		- Raw data from sensors (e.g., images), Manual digitalisation, Basic data processing and feature extraction, Complex models or algorithms such as Machine Learning and AI - Complex models or algorithms such as Machine Learning and AI - Raw data from sensors (e.g., images), Manual digitalisation, Basic data processing and feature extraction, Complex models or algorithms such as Machine Learning and AI - Raw data from sensors (e.g., images), Basic data processing and feature extraction, Complex models or algorithms such as Machine Learning and AI
Thematic Data Acquisition	6	- National Forestry Service (WMS/WFS) - Other national agencies providing data - NA - Other national agencies providing data, NASA, Copernicus, Google Earth Engine - National Environmental Protection Agency (WMS/WFS), Other national agencies providing data, Copernicus, Tuscany Region and local Authorities, EUMETSAT - Other national agencies providing data, NASA, Copernicus - NASA, Copernicus
Dataset Urls	2	- https://apps.sentinel-hub.com/eo-browser/; https://google.earthengine.app/; https://www.usgs.gov/media/files/landsat-8-9-olitirs-collection-2-level-1-data-format-control-book; - https://geoportale.lamma.rete.toscana.it; https://erddap.lamma.toscana.it
Cost	6	- 0 - No - NA - No - 1000/year - Not sure. - We haven't any licenses.
Data Size	7	- 0 - ? - Up to 1Mb/ha. RGB 0.5m image (0.3 Mb/ha), Canopy cover (0.1MB/ha), Other raster info (0.1MB/ha), Forest Boundaries (0.1MB/ha), Strata (0.1MB/ha) - I don't know. It depends on the applications - 7 GB, per year, per 18 variables - Not sure at this time. - We are not able to calculate the data size in units. Our remote sensing datasets cover all Mediterranean Basin and European countries.
Raster Format	7	- GeoTIFF – TIFF - GeoTIFF – TIFF - GeoTIFF – TIFF - GeoTIFF - TIFF, ECW, Esri grid, IMG – ERDAS IMAGINE, JPEG2000, WMS, NetCdf, GRIB - GeoTIFF - TIFF, JPEG2000, WMS - GeoTIFF - TIFF, ECW, Esri grid, IMG – ERDAS IMAGINE, JPEG2000, World file Georeferencing a raster image file (e.g. JPEG, BMP), WMS - GeoTIFF - TIFF, Esri grid, World file – Georeferencing a raster image file (e.g. JPEG, BMP)
Vector Format	7	- Shapefile (SHP) - Shapefile (SHP)

	Nr. answers 18	Answers
		- Shapefile (SHP), GeoJSON - Shapefile (SHP), KML or KMZ, WFS, GeoJSON, Spatialite, GeoPACKAGE - Shapefile (SHP), KML or KMZ, GML, WFS, GeoJSON, Spatialite, GeoPACKAGE - Shapefile (SHP), KML or KMZ, GML, WFS, GeoJSON, Spatialite, GeoPACKAGE - Shapefile (SHP), GeoJSON
Lidar Format	7	- LAS - Point cloud (LAS, ASCII, XYZ) - Not needed - Point cloud (LAS, ASCII, XYZ) - Point cloud (LAS, ASCII, XYZ) - LAS, ASCII, XYZ, PLY and others compatible with CloudCompare and eventually with Open 3D - Point cloud (LAS, ASCII, XYZ)
Other Formats	1	- NetCDF, Grib, Specific DB formats, for example postgresql
Sintetic Tasks	7	- Task 2.2 - Task 2.1, Task 2.2, Task 2.3, Task 2.4, Task 4.1, Task 4.2, Task 4.3 - Task 1.3, Task 2.2 - Task 3.2, Task 5.3 - Task 5.3 - Task 1.2, Task 1.3, Task 3.2, Task 4.1, Task 4.2, Task 5.3, Task 5.4, Task 6.1, Task 6.2, Task 6.3 - Task 1.3, Task 3.1, Task 3.2, Task 3.3, Task 5.2, Task 5.3
New Datasets	6	- K - Raw data of logs - NA - Vegetation index Time series - BOLAM: 18 meteorologic/climate variables, hourly per year - Identification of forest cut-off areas. - not defined yet
Visualization	7	- Text - Text, Table, Map, File - Map - Text, Table, Map, File - Text, Table, Map, File - Text, Table, Map, File - Text, Table, Map, File
Required Level Of Processing	7	- Only storage and/or visualisation - Only storage and/or visualisation, Data aggregation, reporting and/or filtering - Only storage and/or visualisation - Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis, No further storage or processing is required - Only storage and/or visualisation, Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis - Only storage and/or visualisation, Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis

	Nr. answers 18	Answers
		- Data aggregation, reporting and/or filtering, Data to be used as input to models or algorithms in more complex analysis
Download	7	- K - TIF, Image (PNG, JPEG) - TIF, SHP - TIF, SHP, WMS/WFS - TIF, SHP, WMS/WFS - TIF, SHP, WMS/WFS, Image (PNG, JPEG) - TIF, SHP, Image (PNG, JPEG)
Data Access	7	- Data shared with approved external stakeholders - Open data access - Data shared only with approved external stakeholders - Open data access - Open data access - Open data access - Open data access
Data License	6	- I do not have a preference - Depending on the EO data provider restrictions - Open data usage license such as CC-By attribution, ODBL, etc - Open data usage license such as CC-By attribution, ODBL, etc - Open data usage license such as CC-By attribution, ODBL, etc - Open data usage license such as CC-By attribution, ODBL, etc
Comments	0	

Annex II – Data inventory table

PARTNER short name	Sintetic								
PID [Persistent Identifier]	**NAME*** [name of the dataset]	**DESCRIPTION** [brief explanation of the data]	**TYPE** [Type of the dataset]	**FORMAT*** [physical data format]	**SIZE** [MByte]	**ORIGIN/PROVENANCE** [APP/DATA/SENSOR/SERVICES]	**url** [web address]	**NEW/EXISTING** [new/existing]	**DATA UTILITY** [WP/Deliverable]

AVAILABILITY [open/restricted]	**METADATA** [yes/no]	**DATA CATALOGUE** [url]	**STRENGTH/OPPORTUNITY NOTES**	**WEAKNESS/THREAT NOTES**	**ETHICAL ISSUES ?** [yes/not]

Annex III – 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 Properties (O)

Annex IV- 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