Data Management Plan

Project deliverable D1.3

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101076963. UK participant IRAP is supported by UKRI grant number 10139277. UK participant Agilysis is supported by UKRI grant number 10157029.

DELIVERABLE ADMINISTRATIVE INFORMATION

PROJECT ACRONYM	CAMBER
PROJECT START DATE	01/01/2025
PROJECT DURATION (MONTHS)	36
GRANT	101076963
CONSORTIUM COORDINATOR	EIRA
TITLE OF THE DELIVERABLE	D1.3 – Data Management Plan (DMP)
WORK PACKAGE	WP1 – Project Management
TYPE OF DELIVERABLE	DMP — Data Management Plan
DISSEMINATION LEVEL	PU - Public
STATUS – VERSION, DATE	Draft – V1, 26/06/2025
SUBMISSION DATE	30/06/2025

AUTHORS/CONTRIBUTORS

Name	Organisation	Contribution
Andreas Hula	AIT	Lead Contributor

Roland Spielhofer	AIT	Dataset Description Contributor
Lisa De Winter	BMOB	Dataset Description Contributor
Katia Pagle	ICCS	Dataset Description Contributor
Ali Pirdavani	UH	Dataset Description Contributor
Dario Babić	FPZ	Dataset Description Contributor
Richard Owen	AGIL	Dataset Description Contributor
Jan Hendrik van Petegem	SWOV	Dataset Description Contributor
Monica Olyslagers	IRAP	Dataset Description Contributor
Sandra Viera	LNEC	Dataset Description Contributor
Itziar Urbieta Izquierdo	VICOM	Dataset Description Contributor

QUALITY CONTROL

	Name	Organisation	Date
Peer review	Ralf Willenbrock, Euripides Sakellariou	ERTICO	23/06/2025
Peer review	Nathan Harpham, Suzanne Coles	AGILYSIS	23/06/2025

VERSION HISTORY

Version	Date	Author
0.1	04/04/2025	Internal Draft
0.2	11/04/2025	Internal Draft after first Meeting
0.3	06/05/2025	Internal Draft after receiving Dataset Descriptions
0.4	28/05/2025	Internal Draft after receiving Final Dataset Contributions
0.5	05/06/2025	Document ready for internal review
1.0	27/06/2025	Final Version after Peer review

APPROVED FOR SUBMISSION BY

Name	Organisation	Approval date
Olivera Rozi	EIRA	30/06/2025

LEGAL DISCLAIMER

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101076963. UK participant IRAP is supported by UKRI grant number 10139277. UK participant Agilysis is supported by UKRI grant number 10157029. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. The information in this document is provided “as is”, and no guarantee or warranty is given that it is fit for any specific purpose. The Camber project Consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

TABLE OF CONTENTS

LIST OF TABLES

PROJECT EXECUTIVE SUMMARY

The Connected and Adaptive Maintenance for Safer Urban and Secondary Roads project ('CAMBER') aims to develop and demonstrate improved safety monitoring across urban and secondary rural road networks through real-time data feedback into road maintenance systems and proven low-cost interventions.

Performance metrics based on new-generation data sources will provide road managers up-to-date information on safety issues, damage, and routine maintenance and upgrade needs. Data collated from a range of sources, such as telematics, vehicle and smartphone sensors, and road user feedback, will feed into safety assessment models to flag what measures are required to ensure a safe road environment for all road users, including road-user minority groups with varying design needs, such as powered two-wheelers (PTW).

CAMBER will support this through much-need research and testing of low-cost road safety interventions and low-impact maintenance techniques, including those for vehicles with advanced driver-assistance systems (ADAS). The approaches will be demonstrated on urban and road networks in five European countries. CAMBER's economically-sound solutions and new knowledge will be communicated through established networks to European road managers, policymakers and industry to support the decision-making and investment needed for more efficient maintenance for safer urban and secondary roads.

Social Media:

@camber-project

@CamberProject

For further information please visit camber-project.eu

DELIVERABLE EXECUTIVE SUMMARY

The CAMBER Data Management Plan (DMP) satisfies Deliverable 1.3 (D1.3) for the CAMBER project. This document contains the first version of the DMP which will be reviewed and updated as needed in project delivery, with an introduction to the purpose of the document to be found in section 1. The first iteration of the DMP lists data sets foreseen to be used in the project up to the beginning of month 6 (M6) in section 2, which is the main content of D1.3. Section 3 outlines the FAIR (Findable, Accessible, Interoperable and Reusable) principles and what next steps (T3.1, T3.2) in the project will deal with their implementation. The document also points to the implementation of data protection guidelines in CAMBER in section 4 and the protection of personally identifiable data in section 5 and the acknowledgment and mitigation of risks with regards to data risks in CAMBER in section 6, within two other deliverables (the Risk Management Plan D1.2 and Ethics Management Plan D1.6). The next DMP deliverables will deal with the implementation of data sharing within the project (D1.4) and making data reusable and publicly accessible beyond the project (D1.5).

The DMP from D1.3 to D1.5 describes how the CAMBER project will manage the datasets that will be gathered or created during the course of the project. In addition, the DMP details the terms of best practice for handling metadata and storage in order to ensure that data is findable, accessible, interoperable, and reusable (FAIR). The scope of the CAMBER Data Management Plan (DMP) is to detail how research data will be handled during and after the end of the project.

LIST OF PARTNERS

Acronym	Organisation	Country
EIRA	EVROPSKI INSTITUT ZA OCENJEVANJE CEST – EURORAP	SI
SWOV	STICHTING WETENSCHAPPELIJK ONDERZOEK VERKEERSVEILIGHEID SWOV	NL
LNEC	LABORATORIO NACIONAL DE ENGENHARIA CIVIL	PT
AIT	AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH	AT
FPZ	SVEUCILISTE U ZAGREBU FAKULTET PROMETNIH ZNANOSTI	HR
UH	UNIVERSITEIT HASSELT	BE
ICCS	EREVNITIKO PANEPISTIMIAKO INSTITOUTO SYSTIMATON EPIKOINONION KAI YPOLOGISTON	EL
BMOB	BE-MOBILE	BE
VICOM	FUNDACION CENTRO DE TECNOLOGIAS DE INTERACCION VISUAL Y COMUNICACIONES VICOMTECH	ES
ERTICO	EUROPEAN ROAD TRANSPORT TELEMATICS IMPLEMENTATION COORDINATION ORGANISATION - INTELLIGENT TRANSPORT SYSTEMS & SERVICES EUROPE	BE
ETRIK	ANAPTYXIAKI ETAIREIA DIMOU TRIKKAION ANAPTYXIAKI ANONYMI ETAIREIA OTA	EL
MITMA	MINISTERIO DE TRANSPORTES Y MOVILIDAD SOSTENIBLE	ES
iRAP	INTERNATIONAL ROAD ASSESSMENT PROGRAMME	UK
AGIL	AGILYSIS LIMITED	UK

LIST OF ABBREVIATIONS AND ACRONYMS

Acronym	Meaning
AI	Artificial Intelligence
CAMBER	Connected and Adaptive Maintenance for Safer Urban and Secondary Roads project
CAN	Control Area Network
DMP	Data Management Plan
DPIA	Data Privacy Impact Assessment
EC	European Commission
EN	European Standard
FAIR	Findable, Accessible, Interoperable and Reusable
FCD	Floating Car Data
FVD	Floating Vehicle Data
GA	Grant Agreement
GDPR	Generalised Data Protection Regulation
GIS	Geoinformation System
GPS	Global Positioning System
Hz	Hertz
IPR	Intellectual Property Rights
ISO	International Organization for Standardization
KoM	Kick-off Meeting
KPI	Key Performance Indicator
M	(Project) Month
OBD	On Board Diagnose
PII	Personally Identifiable Information

RFID	Radio Frequency Identification
T	Task
TE	Traffic Events
WP	Work Package
YOLO	You Only Look Once

1 INTRODUCTION

1.1 DATA MANAGEMENT PLAN PURPOSE

The data management plan (DMP) outlines the procedures and records for handling data within the CAMBER project to promote robust data management practices. These methods comply with the Horizon Europe Guidelines on FAIR data management (European Commission 2016) and uphold data management standards that align with legislation such as the GDPR, while also incorporating best practice principles.

The primary objectives of the DMP include outlining the methods for managing both operational and research data throughout the project and beyond its duration. It also specifies how project datasets will adhere to the FAIR principles—making data Findable, Accessible, Interoperable, and Reusable—while acknowledging any necessary restrictions under the principle of being “as open as possible, as closed as necessary.” Additionally, the plan establishes how resources will be allocated for data management in alignment with the overall workplan and deliverables to facilitate ongoing review and enhancement of the plan. Lastly, it sets forth procedures to ensure data security during the project and to guarantee proper data preservation.

1.2 DATA MANAGEMENT PROCEDURE

The primary audience for the DMP is the project partners, to promote best practices in data handling. However, as a public document, it may also be accessed by broader stakeholders or other relevant parties when necessary.

Responsibility for adhering to the data management plan, along with any associated guidelines and deliverables, lies with each partner, who must also contribute to its ongoing review and improvement. Partners are required to inform the coordinator and work package leaders of any changes in the data they collect or process throughout the project, ensuring the DMP remains current. This communication takes place within WP1, specifically under task T1.4. As the project advances, any modifications to data, processing methods, or practices may lead to updates of this document and the overall data management plan, enabling continuous refinement of data management procedures.

2 RESEARCH DATA SUMMARY

This section lists the datasets currently (M1-M5) foreseen to be collected throughout the project. Changes are expected to arise from the outputs of WP2 (data sets to be identified to be important for impact assessment or state of the art KPIs) and will inform following versions of the DMP (D1.4 and D1.5).

The inventory of data will be extended in particular in tasks 3.1, 3.2, 3.4 and throughout the project, through WPs 2,3,4 and 5 (data needs to be identified in T2.3, T2.4, T2.5 as well as T4.2, T4.4 and WP5 test cases at the pilots). This overview provides the data collectors/contributors and the type of data collected in the project. It is augmented by an Excel-List of the datasets on the project sharepoint in the task 1.4 folder. The fields of that Excel-List are

- Project Partner - Who will provide the data to the project
- Dataset Label - A label to be used in the project. The naming convention will follow the activities in 3.1 and 3.2
- Data Title – A short descriptive title
- Type of Data - A short descriptive comment on the data structure
- Source/Sensor - What source/sensor the data is collected from
- Data Description – a longer description of the data
- Data Format – a formal description (like ".csv") of the data type
- Access or Processing Restrictions – Note on whether access restrictions are foreseen
- For Testsites – for which testsites the data set has relevance
- Observations, Comments - Short notes of the providers to the CAMBER consortium
- Descriptions – Longer explanations of the intended use

Below we show the Dataset Label, Data Title, Type of data and Data description to give a brief overview. As more data sets become relevant to the project, they will be included in the next version of the DMP (D1.4 at M18 and D1.5 at M36) and the Excel-List may be made publicly available. For this first version of the DMP the data set descriptions provided by the partners are included below.

All data sets are stored with the partners providing them currently. Tasks 3.1 and 3.2 will develop a concept together with the DMP to review the storage and manage access for the project partners (D1.4) and beyond the project (D1.5). Naming conventions for the data sets will follow in Tasks 3.1 and 3.2 to be included in D1.4 (currently "tbd" still serves as a placeholder).

2.1 MEASUREMENT VEHICLE DATA (AIT)

PURPOSE IN CAMBER

These Roadlab data sets provide ground truth data for several pilots (North Holland, Cascais, Trikala and Croatia), as well as material for building digital twins (point clouds) in WP4.

The MoProVe data provides localised driving dynamics data from a motorcycle probe vehicle to be used in risk assessment and impact estimation (for instance in WP5).

Data will initially be generated and stored at AIT. Access will be discussed in T3.1.

ROADLAB DATA (AIT)

The high-performance survey vehicle Roadlab is equipped with state-of-the-art sensors, satellite navigation and laser scanner/camera technology. It records all relevant properties of the road surface and road environment with the highest quality and accuracy and can even be used on cycle paths and side roads thanks to its compact design. Thanks to its variable design concept, it is easy to integrate a wide range of sensor technologies, which subsequently enable AI-supported data evaluation and assessment.

Modular equipment:

1. Highest precision in the detection of road geometry and localization of road objects thanks to the accurate positioning and attitude system
2. Dead reckoning for uninterrupted positioning even in the event of satellite shadowing
3. 4K video systems (anonymized)
4. 360° panoramic camera (anonymized)
5. Laser scanner for capturing the road corridor
6. Laser scanner for capturing the road surface
7. Flexible platform for sensor integration

The RoadLAB has been used for:

1. Recording and evaluation of cycle paths
2. Condition assessment of municipal road networks (surface distress, rutting, longitudinal evenness)
3. AI-supported evaluation of surface distress
4. Video documentation with 360° camera and 4K camera systems
5. Determination of alignment parameters (gradient and crossfall, curve radius)
6. Inventory of traffic signs, road markings, etc.
7. Checking the clearance profile
8. Determination of lane widths
9. Detailed 3D roadway models in OpenCRG (a fileformat) and FBX (a fileformat)

The high-precision recording of the road condition and, if necessary, the road environment provides infrastructure operators with an important decision-making basis for sustainable and cost-efficient

maintenance planning and an increase in road safety. The measurements will be carried out without interfering with moving traffic.

In CAMBER the Roadlab is foreseen to collect ground truth data on selected corridors of the pilot sites and provide data for digital twin creation.

TABLE 1 — ROADLAB DATA

Dataset Label	Data Title	Type of Data	Data Description
CAMBER_AIT_Laz	3 D Pointcould Data	Geolocated 3D Points	Data acquired to accurately represent the surrounding road infrastructure
CAMBER_AIT_Lon	Longitudinal Evenness	Time Series of Values	Per Meter value of longitudinal evenness
CAMBER_AIT_Surface_Distress	Surface Distress	Time Series of Values	Surface Distress and Rutting
CAMBER_AIT_Road_Geometry	Road_Signage	Time Series of Values	Curvature and other indicators on road geometry

MOTORCYCLE PROBE VEHICLE DATA (AIT)

The Motorcycle Probe Vehicle (MoProVe) is a KTM 1290 Super Adventure motorcycle equipped with a RACELOGIC data recorder that allows to access the in-vehicle CAN-Bus, providing (wheel)speed, (x-,y-,z-) acceleration data as well as rotational angle speeds (yaw-,pitch- and roll-rate) and obtains GPS localization at 100 Hz. The data collection and test rides are handled by experienced AIT employees. Several rides on each test track are necessary to obtain the amount of data for an assessment of the track.

TABLE 2 — MOTORCYCLE PROBE VEHICLE DATA

Dataset Label	Data Title	Type of Data	Data Description
MoProVe_tbd	Date/Time	Driving Behaviour	Date and Time information of testrides, synchronized with the parameter data at 100 Hz
MoProVe_tbd	GPS Data	Driving Behaviour	GPS-derived latitude and longitude output at 100 Hz frequency

MoProVe_tbd	(x,y,z-accelerations)	Driving Behaviour	accelerations at 100 Hz collected from an in-vehicle accelerometer
MoProVe_tbd	(yaw-,roll-,pitch-rates)	Driving Behaviour	rotational angle speeds at 100 Hz collected from an in-vehicle gyroscope

2.2 DRIVING SIMULATION AND EYE TRACKING DATA (UH)

PURPOSE IN CAMBER

The driving simulator data is to be used in impact estimation of interventions.

Data will initially be generated and stored at UH. Access will be discussed with the DMP for D1.4.

DRIVING SIMULATOR DATA

The Driving Simulation and Eye-Tracking Dataset is collected using a high-fidelity experimental setup composed of the STISIM Drive® 3 simulator in combination with a Tobii Pro eye-tracking system. The simulator records detailed driving behaviour data including vehicle speed, position (X, Y), acceleration, steering input, braking, throttle use, and event triggers at a configurable sampling frequency (commonly set to 30–100 Hz). The raw driving data is stored in .DAT format and later exported as .CSV or .TXT for analysis.

In parallel, visual attention data is gathered through the Tobii Pro eye-tracker, which records gaze position, fixation duration, pupil size, and blink events with a sampling rate of up to 120 Hz. Synchronization between driving and gaze data enables precise analysis of driver behaviour in response to various roadway scenarios.

All simulations are conducted in a controlled laboratory environment by trained staff, with participants following a structured driving scenario. Data collection is carried out following informed consent procedures and is approved by the ethical committee at Uhasselt. Each participant completes both pre- and post-experiment questionnaires to provide contextual information and subjective feedback. The questionnaire data is also used to collect relevant demographic, perceptual, and attitudinal information, enhancing the understanding of driver behaviour before and after the simulation. Multiple test runs per scenario are performed to ensure data reliability and robustness for behavioural analysis.

TABLE 3 — HASSELT DATA

Dataset Label	Data Title	Type of Data	Data Description
DS_Log	STISIM Drive® 3 Raw Driving Data	Quantitative (log)	Raw log of driving behaviour: speed, lane position, brake/throttle usage, etc.

DS_Metrics	Calculated Driving Performance Metrics	Processed quantitative	Aggregated indicators like average speed, reaction time, lane deviation, etc.
PreQ	Pre-Experiment Questionnaire Responses	Qualitative/Quantitative	Participant background: age, gender, driving experience, prior simulator exposure, physical well-being, self-reported past driving behaviour, etc.
PostQ	Post-Experiment Questionnaire Responses	Qualitative/Quantitative	Subjective feedback: perceived difficulty, realism, comfort, qualitative intervention evaluation, physical well-being, safety attitudes, etc.
ET_RawData	Eye-Tracking Data from Tobii Pro	Quantitative (time series)	Gaze coordinates (X, Y), timestamps, fixation duration, pupil size, blink rate, and AOI (Areas of Interest) hits. Synchronization with STISIM Drive® 3 timestamps needed. Sampling rate: typically 60–120 Hz depending on hardware.

2.3 CROATIAN PILOT DATA (FPZ)

PURPOSE IN CAMBER

At the Croatian pilot several data collections methods for additional data on road infrastructure and road condition. The exact extent of data usage is to be revisited as interventions testing at the pilots is specified in more detail.

Data will initially be generated and stored at FPZ. Access will be discussed in T3.1.

DYNAMIC RETROREFLECTOMETER

Dynamic retroreflectometer (ZDR 6020) is a device for measuring night-time visibility (retroreflection) of road markings. Device is mounted on a measuring vehicle and allows continuous measurement of the nighttime visibility of road markings while the vehicle is in motion. Device is measuring in accordance to geometry defined in EN 1436 which implies that the retroreflection of the luminous beam from the tested surface is measured at an angle of 2.29 °, with an entry angle of 1.24 ° and at a distance of 30 m for low beams. Measuring area of the device is 1000 x 880 mm. The device is operated by two certified operators and can measure retroreflection at the normal driving speeds.

RETROREFLECTOMETER FOR ROAD SIGNS

Zehntner ZRS 6060 retroreflectometer is a hand-held device that contains an internal light source and a photoreceptor. During the testing, the devices are placed on the road sign surface in order to exclude the influence of daylight and must give retroreflection results for all geometries (input and viewing angles) prescribed by the standard HRN EN 12899-1. Coefficient of retroreflection of road sings will be measured using geometry that implies an observation angle of 20‘ and entrance angle of 5°. On every sings at least three measurements of each colour on a sign (except black) will be taken. As relevant value of the coefficient of retroreflection average value of this three measurements will be taken.

MOBILEYE 630

The data relating to lane marking and road sign detection will be recorded using a Mobileye 630 system implemented in the testing vehicle. Using image processing chips, a Mobileye camera enables high-performance real-time image processing of different objects on roads. For the purpose of the project, data related to the view range and the quality of markings and detection accuracy of road signs will be used.

RFID TRANSPONDER, READER AND ANTENNAS

Passive RFID tags that contain encoded information will be used. The technical specifications of the RFID transponder are shown in table below:

RFID readers are devices that power and wirelessly communicate with RFID tags, delivering data about the read RFID tags to a central management system. To facilitate the automatic reading of road signs, readers are installed in vehicles, enabling continuous RFID tag reading. Given that road signs, due to the structural characteristics of the road and surrounding area, are often located several meters from the edge of the roadway, it is essential to use a long-range RFID reader (with a reading range between 12 m and 15 m) to ensure reading of signs placed at varying distances and locations, which is crucial for maintaining the accuracy of the database updates. Within the project, the Speedway Revolution R420 UHF reader will be used. The technical specifications of the RFID reader are shown in table below:

Antennas are connected to the RFID reader to read the RFID tags. Generally, antenna types are categorized based on how they emit signals: linear and circular. Linearly polarized antennas emit a signal linearly from the antenna and have only one power field. In comparison, circularly polarized antennas emit RF waves in a circular shape and have two fundamental power fields with identical amplitudes that differ in phase by 90°. This specifically means that the wave of one power field is at its maximum value at the same time the other is at its minimum value.

In general, linear antennas, due to their narrower radiation beam, allow for the reading of RFID tags at greater distances, provided that the RFID tags are precisely positioned, making them an optimal solution for systems where the orientation of RFID tags is fixed and predictable. Conversely, circular antennas, due to the nature of their polarization, are less sensitive to the orientation of RFID tags, but their reading range is shorter compared to linear antennas.

Due to the various positions of road signs (either to the right or above the roadway) and the different distances at which the signs are placed, four antennas—two circular and two linear will be used. The technical specifications of the antennas are presented in the table below:

EYE TRACKING DATA

Tobii Pro Glasses 2 (potentially 3) will be used for evaluation of new and innovative interventions aimed at eliciting driving behaviour. This is noninvasive eye tracking method with which the subjects can move freely in a real environment. The glasses contain two eye-tracking cameras for each eye, and four sensors (gyroscopes and accelerometers). The front-facing camera records the scene in front of the subject with a resolution of 1920 × 1080 pixels, with a viewing range of greater than 160◦ horizontally and 70◦ vertically. The glasses contain a microphone for recording sounds and comments. The device allows tracking specific eye behaviour, i.e. number and duration of fixations and saccades, number and duration of gazes to specific areas of interest, pupil diameter etc.

TABLE 4 — CROATIA PILOT DATA

Dataset Label	Data Title	Type of Data	Data Description
tbd	Retroreflectivity_markings	Quality of road markings	Night-time visibility (retroreflection) measurements on the basis of 50 m interval using dynamic measuring device which is based on the measuring geometry - retroreflection angle of 2.29° and entrance angle of 1.24° at a distance of 30 m
tbd	Retroreflectivity_signs	Quality of road signs	Night-time visibility (retroreflection) measurements of each colour (except black) on road signs using standard geometry - observation angle of 20‘ and entrance angle of 5°
tbd	Mobileye_markings	Mobileye’s quality readings and view range	Mobileye’s quality readings and view range calculated on 50 m interval
tbd	Mobileye_signs	Mobileye reading of road signs	Mobileye detection and accuracy of road signs readings

tbd	RFID readings	Detection of road signs	"Transponder: Passive UHF RFID tag Class 1, second generation - Protocol: EPC Global Class 1 Gen2 ISO 18000-6C; Operational frequency: EU 865 - 868 MHz; Active transpoders will also be explored

2.4 FLITZMEISTER APP-DATA (BMOB)

PURPOSE IN CAMBER

Flitsmeister app data is to provide dynamic data to CAMBER including events data and traffic data to be used in impact assessment (WP5 and T3.5).

Data will initially be generated and stored at BMOB. Access will be discussed in T3.2.

FLITSMEISTER DATA

The floating car data (FCD) set, provided by Be-Mobile, contains historical aggregated travel times per minute for a to-be-defined time period. Each record includes a unique SegmentId, a UTC TimeStamp marking the observation time, average Speed in km/h, TravelTimeMs indicating travel time in milliseconds and a LocalTimeStamp providing the corresponding local time. The SegmentId corresponds with a 50 meter road segment on the Be-Mobile basemap based on OpenStreetMap. The dataset enables analysis of traffic flow and congestion patterns across road segments. Be-Mobile is an expert in collecting, processing and mapmatching floating car data. To collect smartphone-based telematics, we will use Be-Mobile's navigation companion app Flitsmeister which has a user base of more than three million monthly users in the Benelux region.

The traffic event dataset, provided by Be-Mobile, contains aggregated traffic events related to poor road surface conditions, based on user feedback submitted via the Flitsmeister mobile app over a to-be-defined time period. Each entry includes a type_id identifying the event type, a unique event id, timestamps for creation and last update (created_at, updated_at), and the event's geographic coordinates (latitude, longitude). The bearings attribute indicates the direction of traffic event. This data supports analysis of road surface issues and user-reported hazards across locations.

TABLE 5 — B-MOBILE DATASETS

Dataset Label	Data Title	Type of Data	Data Description
BMOB_FCD	FCD	Traffic information	Aggregated traveltimes per minute with attributes: SegmentId;TimeStamp;Speed;TravelTimeMs;LocalTimeStamp

BMOB_TE	Road user feedback	Traffic events (TE)	Aggregated traffic events concerning bad road surface with attributes: type_id;id;created_at;updated_at;latitude;longitude;bearings

2.5 CASCAIS PILOT DATA (LNEC)

PURPOSE IN CAMBER

Cascais Pilot data to be provided by LNEC covers data on the pilot road network, relevant events and accident data. The data is essential to the Cascais pilot and impact assessment thereon.

Data will initially be generated and stored at LNEC. Access will be discussed in T3.1 and T3.2.

ACCIDENT DATA

- Type of data - Accident characteristics
- Source/Sensor - ANSR - Portuguese Road Safety Agency
- Data Description - Information on recorded accident characteristics
- Data Format - .csv
- Access or Processing Restrictions - No Restrictions as Anonymized
- Observations, Comments - Processed Data
- Descriptions - This data will be used to map accident risk in Cascais Municipality.

The accident data is collected with Boletim Estatístico de Acidente de Viação (BEAV), the official statistical report used in Portugal to register and characterize road traffic accidents. The BEAV is completed by law enforcement authorities and provides standardized, detailed information on the circumstances in which accidents occur, as well as on the individuals and vehicles involved. In order to ensure consistency and focus the analysis on relevant events, the BEAV applies specific inclusion criteria. It includes all accidents that occur on public roads or areas accessible to public traffic, involving at least one vehicle and resulting in personal injury and/or material damage. These include collisions between vehicles, between vehicles and pedestrians, animals or obstacles, as well as single-vehicle accidents and accidents resulting from attempted suicides. Conversely, the BEAV excludes accidents occurring on private roads or in areas without public access, as well as those resulting from confirmed suicides, unless they cause harm to third parties. Accidents caused by natural events, such as sudden illness or death of the driver, and intentional acts such as homicides, are also excluded, except when they lead to injury or damage to uninvolved persons. The information recorded in the BEAV is structured in six sections that cover the identification of the accident, the external circumstances, the nature of the collision, the vehicles and drivers involved, and the resulting consequences. This data forms the basis for national road safety statistics and supports the development of evidence-based policies and preventive measures.

GPS DATA AND ROUTE EVENTS

- Type of data - Driving Behaviour
- Source/Sensor - Mobileye/CAN-Bus
- Data Description - GPS-derived latitude and longitude
- Data Format - .csv
- Access or Processing Restrictions - No Restrictions as Anonymized
- Observations, Comments - raw Data
- Descriptions - Four Mobileye devices will be installed on four buses servicing two bus lines in Cascais Municipality. These devices will collect data on location and special events (accelerations, obstacles, interactions with other road users) during each journey.

Four Mobileye devices will be installed on public transport buses operating within the Cascais Municipality, with each device assigned to a different vehicle. The selected buses service two specific bus lines, ensuring repeated coverage of defined urban and suburban routes. Throughout each journey, the Mobileye systems will continuously collect georeferenced data, capturing both the precise location of the vehicle and a series of "special events" related to driving dynamics and interactions in traffic. These events include instances of sudden acceleration or deceleration, detection of fixed or moving obstacles, and interactions with vulnerable road users such as pedestrians and cyclists. The data gathered will enable the identification of recurring patterns, critical points in the network, and potential risk factors associated with driver behaviour, infrastructure conditions, and the overall traffic environment.

REGISTERED CASCAIS MUNICIPALITY ROAD NETWORK ASSETS

- Type of data - GIS
- Source/Sensor – Cascais Municipality
- Data Description - GIS layers with information on road assets from Cascais Municipality
- Data Format - .shp
- Access or Processing Restrictions - No Restrictions
- Observations, Comments - Processed Data

This data will be used to correlate accident risk with data from Mobileye and other collected data (for assessing the effects of pilot interventions). This GIS is the framework for mapping any events during the project.

The GIS (Geographic Information System) data available for Cascais Municipality includes a wide range of information relevant to traffic analysis and road infrastructure. This data comprises both traffic volume records and detailed inventory data categorized into several thematic areas.

In terms of traffic-related infrastructure, the GIS includes the location and classification of horizontal road markings, traffic signs, directional signs, traffic lights, and parabolic mirrors. This information supports the understanding of regulatory and warning elements present along the road network.

Waste collection infrastructure is also mapped, including the distribution of regular waste containers, recycling stations, and oil disposal units, which can influence curbside activity and roadside obstructions.

Additional public space equipment is registered, such as road delineator posts, benches, concrete barriers, pedestrian protection barriers, safety guardrails, continuous guardrails, and speed bumps. These elements are important for assessing both road user safety and public space design.

Moreover, detailed typological information about road segments is available, including the classification and naming of roads, minimum, average, and maximum roadway widths, platform widths, road gradients, and average surface areas. The dataset also contains contextual urban features such as the average building height, the existence of green areas and trees, as well as the presence of parking and public transport services.

TABLE 6 — LNEC DATASETS

Dataset Label	Data Title	Type of Data	Data Description
tbd	Accident data	Accident characteristics	Information on recorded accident characteristics
tbd	GPS Data	Driving Behaviour	GPS-derived latitude and longitude
tbd	Route events	Driving Behaviour	Information on types of events during bus journeys
tbd	Road network	GIS	GIS layers with information on road assets from Cascais Municipality

2.6 TRIKALA PILOT DATA (ICCS/ETRIK)

PURPOSE IN CAMBER

Trikala pilot data provides the essential data sets for the use cases of the Trikala pilot site (WP5).

Data will initially be generated and stored at ETRIK and with ICCS. Access will be discussed in T3.1 and T3.2.

TRIKALA DATASETS

The application will make combined use of smartphone's sensors such as accelerometers and GPS sensor to detect and classify irregularities of the road surface while bicycling (mainly used in the bike lanes network of the city).

TABLE 7 — TRIKALA DATASETS

Dataset Label	Data Title	Type of Data	Data Description
tbd	GPS Data	Bike Path Classification	Identify path to classify / Identify paths to examine
tbd	Video recording	Bike Path Classification	Path classification from video analysis with object detection (YOLO)
tbd	Gyroscope Data	Bike Path Classification	Location identification from analysing raw data from sensor
tbd	Accelerometer Data	Bike Path Classification	Location identification from analysing raw data from sensor
tbd	GPS Data	Bicycle Location and Movements	Identify paths to examine
tbd	Accelerometer Data	Bicycle Location and Movements	Location identification from analysing raw data from sensor

2.7 PILOT DATA SPAIN (VICOM/MITMA)

PURPOSE IN CAMBER

This pilot data provides the essential data sets for the use cases of the Spanish pilot site (WP5). Here this also includes the essential data to form digital twins for the Spanish pilot, as at this site RoadLAB data is not foreseen to be collected.

Data will initially be generated and stored at VICOM. Access will be discussed in T3.1 and T3.2.

DATASET AT SPANISH PILOT

CARLOTA (CAR Learn On Road Think Autonomous) is a Toyota Prius 1.8 HSD Executive vehicle, instrumented and operated by Vicomtech. It serves as a prototype platform for deploying and testing CCAM (Cooperative, Connected and Automated Mobility) solutions, including advanced data recording capabilities.

The vehicle is equipped with 8 fisheye cameras providing a full 360º field of view, and 3 driver monitoring cameras to track driver behaviour. Additionally, it features a 360º horizontal field-of-view LiDAR sensor for 3D scene capture, a Nav550 DGPS, a NeoGLS V2X OBU, and a Quectel 5G modem.

All components are orchestrated by an onboard PC (Intel i7-9700E, 64 GB RAM, 2× Nvidia 2080Ti GPUs), running RTMaps middleware, which enables synchronized data acquisition and recording. For remote or autonomous driving tasks, the vehicle's CAN-Bus is accessible via an actuator interface known as the CAN-BUS Move Box.

2.8 NORTH-HOLLAND PILOT DATA (SWOV)

PURPOSE IN CAMBER

This pilot data provides the essential data sets for the use cases of the North Holland pilot site (WP5).

Data will initially be generated and stored at SWOV. Access will be discussed in T3.1 and T3.2.

NORTH-HOLLAND SITE DATA

For the North-Holland pilot site, several current datasets on road characteristics and road assets will be collected, merged and extended. Additional data will be collected for ground truth measurements on selected rural and urban roads using for example AIT's Roadlab and MoProVe.

Road characteristics and assets data will be utilized for a safety review of the road network. The safety review will include the usage of the iRAP and CycleRAP methodologies. The network will encompass both secondary roads and physically seperated cycle tracks outside build-up areas and roads and cycle tracks inside build-up areas. The datasets will include spatial datasets (gpkg ea.) and attribute datasets by text files (.csv ea) and spreadsheets (.access, xlsx, ea) linking to the spatial datasets. Besides road characteristics and assets data, traffic flow data (AADT and/or aaht) for these roads and cycle tracks will be collected. These form the static datasets of the North Holland pilot site.

For the North Holland pilot site, we forsee several dynamic datastreams. North Holland is involved in several innovative initiatives to improve the management of the road network, from asset management, traffic management and road safety to maintenance and smart mobility. The Province is involved in the national-led Road Monitor (ROMO) project which will deliver vehicle sensor and events data from several vehicle manufacturers and service providers for a series of specific use cases, relevant to the Camber project. Amongst which are use cases for road surface quality monitoring and the identification of road safety hazard hot spots. At this stage the use cases are in tender and the risks to data collection have been added to the risk management log of CAMBER. Early summer information will come forth containing which use cases will come available for research. Furthermore, B-Mobile is active in the

North-Holland region and will be collecting vehicle probe data and link travel time data. These dynamic datasets will be utilised to identify hazard locations and possible mitigating safety measures and improvements to assets and safety management.

Aside from collecting dynamic data, B-mobile may provide dynamic data to traffic control centres or road users containing warnings and or advices aiming to mitigate or prevent hazardous circumstances. These warnings and advices will be logged as well as vehicle probe data to monitor possible effects on behavioural changes by road users.

Finally, the aim is to use the datacollection by AIT's Roadlab and MoProVe to calibrate the dynamic datastreams for relevant triggers to road safety and asset management processes. For example, to calibrate what (amount of) vehicle events data on insufficient lane marking can provide the road administrator with a trigger for bad lane marking quality impeding the effectiveness of systems as LKA and LDW.

TABLE 8 —NOTH-HOLLAND DATASETS

Dataset Label	Data Title	Type of Data	Data Description
tbd	Road characteristics	GIS data	Spatial data on roads and surroundings
tbd	Road Assets	GIS data, Attributes data, life cycle data, ea	Data on type, quality, state, maintenance ea on roads assets
tbd	Deceleration events	Vehicle probe data and Vehicle events data	Data on decelleration events
tbd	Vehicle probes	Vehicle probe data	Vehicle probe data
tbd	Vehicle sensor data	OBD and Telematics data	Processed vehicle sensor data
tbd	Vehicle events data	Vehicle sensor events data	Events logging of vehicle sensors, such as road surface condition, harsh braking events and opponent type, ea
tbd	Vehicle flows	Live and or AADT traffic flow data	Traffic flows per link. Live and or ex post estimates. Different possible sources available.
tbd	Travel times	Live road link travel times	Travel time conditions per link

tbd	Simulator study	Driving behaviour	Data generated by a simulator study, including metrics, eyetracking data and questionnaire responses

2.9 AGILYSIS DATASETS (AGIL)

PURPOSE IN CAMBER

AGIL foresees the use of supplier data in impact assessment at multiple pilot sites (WP5 and T3.5).

Data will initially be managed by AGIL. Access will be discussed in T3.1 and T3.2.

GPS-BASED VEHICLE DATA

We will seek to obtain or procure Floating Vehicle Data (FVD) from a supplier or suppliers that will enable the pilot projects to assess vehicle traffic and speeds on secondary roads. This data is currently supplied by companies such as Here, Inrix, TomTom, Google and others, including data aggregators who obtain information directly from vehicle OEMs. The primary use cases for access to this information are to identify average (mean) and 85th percentile speeds and potentially other information relating to the speed profile of vehicles. Data may also indicate relative traffic flows / movements and potentially turning frequency or origin / destination. This data is likely to be either aggregated to specific road sections using an underlying road network, or produced as 'raw data' which would include individual GPS tracks for vehicles. If this is provided then it will be anonymised prior to receipt and will not contain any personally identifiable information including make, model and colour of vehicle, or the close vicinity of start and end points of journeys.

VEHICLE SENSOR DATA

The Vehicle Sensor Data (VSD) may be obtained for a number of purposes and the specific requirements have not been confirmed at this time. There are a number of sensor outputs that would be of use to the project, and which we understand there is some availability within the market. The first category is related to vehicle movements but goes beyond simple GPS data about average speeds and movement patterns. This would include much more granular data in terms of spatial and temporal references, plus the inclusion of "XYZ" data. XYZ data typically refers to three-dimensional coordinates, representing the location of a point in space. When used in the context of orientation, the Z-axis is often associated with the vertical or up direction, and the other two axes (X and Y) are used to define the horizontal plane. This information cannot accurately be obtained from GPS systems unless they are operating with a high frequency and accuracy. This data would enable assessments of harsh braking and other driver behaviours. Other VSD comes under the classification od "ADAS" Advanced Driver Assistance Systems. This information would include alerts triggered due to driver behaviour, the presence of other vehicles, or external factors such as weather. There is already some data being supplied through the Data For Road Safety project which demonstrates its potential, but this will need to be explored to see how it can be used operationally in this project. Some vehicle OEMs are already demonstrating the supply and use of this data for harsh braking events and there remains the possibility that other warning could be geocoded and analysed for secondary road networks.

TABLE 9 — AGILYSIS DATASETS

Dataset Label	Data Title	Type of Data	Data Description
tbd	Traffic Data	Floating Vehicle Data	Includes information on vehicle speeds, and probe counts. May be 'raw data' or pre-processed and matched to existing road sections.
tbd	ADAS data	Connected Vehicle Data	Includes information on vehicle systems, specifically anonymized data related to activations.

2.10 PROJECT PARTICIPANT DATA (EIRA/IRAP)

In addition to the above, we mention that personally identifiable data on participation in project meetings is collected and managed, in close coordination with the Ethics Management of CAMBER, at EIRA. Consent has been obtained for storing this data under the Ethics guidelines.

3 IMPLEMENTATION OF FAIR GUIDELINES

This section is based on the standard H2020 template for FAIR data management as outlined in the "Guidelines on FAIR Data Management in Horizon 2020" (European Commission, 2016). All collected and generated data will be prepared for external use and made findable, accessible, interoperable, and reusable, except where specific restrictions apply due to an opt-out.

The implementational details and concrete storage means will be outlined in D1.4 as the range of data sets to be collected and the extend of the data sets becomes clear in more detail through the outcomes of T2.3, T2.4, T2.5 (forming an impact assessment concept) and T4.1, T4.2, T4.4 (forming the needs of simulation and digital twin building), as well as the specific pilots to be performed in WP5. These needs will also be identified and documents through the activities in T3.1 and T3.2 which will also work closely with the DMP.

3.1 LIMITATIONS TO APPLYING FAIR

Project partners may choose to opt out of managing data according to FAIR principles. Throughout the project, all instances of FAIR data opt-out will be monitored along with the legitimate reasons for each decision. These opt-out choices will be documented in the D1.4 and D1.5 reporting deliverables to ensure continuous tracking. These reports will specify how FAIR principles are managed for each dataset and note any applicable opt-out exemptions.

- Data may be excluded from FAIR principles for various reasons, including but not limited to protecting data privacy, such as limiting data sharing to safeguard sensitive information related to the project, stakeholders, or participants, particularly concerning contact details and restricted management information.
- Another reason could be to protect intellectual property rights and support project exploitation, where existing data may have protections or limitations that prevent full disclosure, or where project findings might generate opportunities for exploitation that require limiting full FAIR disclosure to avoid harm.
- Contractual protections may also necessitate opt-outs, for example, when use case or scenario-related data collected in pilots is only accessible to consortium members under specific contractual terms that restrict further sharing, especially when data is owned or controlled by third parties in certain regions.
- Lastly, ethical considerations may lead to opt-outs if releasing data could potentially cause societal harm, as identified through stakeholder interactions, such as in cases where specific road data might negatively impact regional activities like road enforcement, public safety, security, or resilience.

3.2 ACHIEVING FINDABILITY, INCLUDING PROVISIONS FOR METADATA

Activities within WP3, specifically Tasks 3.1 and 3.2, will involve gathering metadata to facilitate making the data findable. These activities, starting from M6 will be synchronized with the M6 to M18 activities of the DMP and their outcome will be part of D1.4, the second version of the data management plan. This metadata will also be incorporated into any public data releases, such as those made available on the project website, to assist users in locating relevant data. These tasks will develop a standardized naming convention for files to ensure consistency and easy identification of folders, files, and deliverables. The naming system will include project identifiers and version information. Additionally, all deliverables and documents will feature document control details in accordance with the guidelines set out in this document.

CAMBER plans to use Digital Object Identifiers (DOIs) for datasets in a public repository, specify naming conventions for datasets, and use rich metadata with open data formats (csv, xml).

3.3 ACHIEVING ACCESSIBILITY

Throughout WP3 and the entire project, all project data will be reviewed to determine if it can be made openly accessible, taking into account any necessary opt-outs due to privacy, intellectual property rights, contractual obligations, or ethical considerations. Data containing Personally Identifiable Information (PII) may initially be held by individual partners before being aggregated or anonymized for project use. Shared data will be accessible to all project partners through the project repositories or a shared drive folder authorized for use by all participants. Initially, data sets collected at the pilot sites will be stored at pilot site operated data infrastructure and will be made accessible through specific links, if partners have a processing need. This sharing structure and the processing needs will be documented in D1.4 and the data infrastructure is foreseen to be outlined in tasks 3.1 and 3.2 (both starting in M6). Later on, data to be made accessible publicly and beyond the end of the project can be deposited in public repositories like Zenodo. The exact modalities of this will be worked out for D1.5 the final version of the DMP, due in M36 of the project.

Whenever possible, data will be stored in common, shareable formats—such as CSV, DOCX, XLS, and PDF—to ensure accessibility without requiring specialized software. In cases where specialized tools are necessary, such as for risk framework analysis and modelling, these tools will be made available within the project as needed, this includes in particular necessary scripts to preprocess the data or make more readily accessible the outcomes in standard programming languages like R or Python.

3.4 ACHIEVING INTEROPERABILITY

The project aims to adopt standard file formats and naming conventions to facilitate interoperability. Metadata collection will follow a unified approach, which will be defined in Tasks 3.1 and 3.2 using standard file formats (like .csv, .docx, .xls and .pdf). This process will be supported by ongoing activities within WP3 throughout the duration of the project.

3.5 ACHIEVING REUSABILITY

Several tasks within the project focus specifically on enhancing data formats (T3.1, T3.2, T3.3, T3.4) as well as data storage, handling, and accessibility (T3.1, T3.2, T3.4). As part of WP3, data reuse is tracked along with other data attributes to maximize the potential for data to be reused effectively. CAMBER aims that all data producers will license their data to allow the widest reuse possible (e.g., using licenses from opendefinition.org).

3.6 ENSURING DATA MANAGEMENT RESOURCES

This section outlines the resources needed to support FAIR data management. The majority of activities related to data collection, preservation, curation, and management are already budgeted within the consortium through various project tasks and activities. Each partner is responsible for the data they generate. Data obtained from field trials, simulator experiments, or other sources is stored and archived by the respective partners at their own facilities, where they are also responsible for its management and security. Additionally, data cleaning and preparation for processing or further analysis are considered integral parts of data collection. Tasks 3.1 and 3.2 will develop procedures and logistics for handling large datasets.

As the task leader for the DMP, AIT maintains an overview of all datasets collected throughout the project and will coordinate with partners to ensure timely adherence to FAIR commitments. Tasks 3.1 and 3.2 will discuss the potential for project-wide specific roles for data management beyond individual partner responsibilities besides the DMP management and include the outcome in D1.4.

4 DATA SECURITY HANDLING

To ensure compliance with legislation, a dedicated GDPR and Data Protection Impact Assessment (DPIA) process has been established under Task 1.5 and documented in D1.6 (Ethics Management), following GDPR best practices and record-keeping requirements. The process can be found in section 2 of the Ethics Management Plan (especially 2.3). This process aims to enhance monitoring and ensure adherence to regulations during project-related data processing. It includes maintaining a list of data controllers to facilitate the issuance of data protection notices, reviews, or inquiries through designated legal data controller contacts (see 2.4).

With regards to Authentication, Authorization, Accounting, Confidentiality, Communication Security (e.g., HTTPS, IPsec), Data Integrity (e.g., hash functions, digital signatures), and Availability i.e. the modalities of access, more details is to follow from tasks 3.1 and 3.2, taking into consideration data management standards like ISO 27001 and the outcome will be listed in D1.4. The specific requirements for PII are however to be found in 3.3.5 in D1.6 (Ethics Management Plan).

When special considerations are necessary to protect data during the creation or collection of new data, these will be addressed in Task 1.5 and D1.6 i.e. the DMP management will report to the Ethics Committee about new data sets and evaluate potential PII status.

This approach guarantees that any Personally Identifiable Information (PII) is either consented for release and reuse or properly redacted before sharing or further processing.

5 ENSURING ETHICAL DATA USE

Ethical management, although related to data management, is detailed separately in the Ethics Management Plan (D1.6). The ethical assessment will happen alongside the listing of datasets in the DMP.

6 RISK MANAGEMENT

CAMBER has defined that risk management related to data is handled throughout the project and will also be captured in the project risk register. Deliverable 1.2 Risk Management Plan and Risk Management Register will enable the consortium to keep track of the risks and their associated management.