PRE-
LARGE-SCALE
DEMONSTRATION
PROJECT

Interim Report

Draft Data Concept

Version: 1.0 Interim Report

Dissemination level: PU

Lead contractor: ika & BMW

Due date: 31.07.2026

Version date: 31.07.2026

This project has received funding from the
European Union's Horizon Europe research and innovation
funding programme under grant agreement No 101269485.

Document information

Authors	
Marcel Sonntag – ika, RWTH Aachen University Felix Fahrenkrog – BMW Group Pre-LSDemo Partners of Task 4.2 and Task 5.3	
Coordinator	
Aria Etemad Volkswagen AG Berliner Ring 2 38440 Wolfsburg Germany Phone: +49-5361-9-13654 Email: aria.etemad@volkswagen.de	
Project funding	
Horizon Europe CL5-2025-04-D6-02 — Preparing for large-scale CCAM demonstrations (CCAM Partnership) – Societal Readiness Pilot Contract number 101269485 www.Pre-LSDemo.eu	

Legal Disclaimer

The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The 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. Although efforts have been coordinated, results do not necessarily reflect the opinion of all members of the Pre-LSDemo consortium. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains.

© 2026 by Pre-LSDemo Consortium

Note of usage of AI

Generative AI (e.g. Microsoft CoPilot) was used solely to improve the linguistic quality and readability of the document. No AI-generated content was used to develop scientific findings, analyses, or conclusions. All content was reviewed and approved by the authors.

Revision and history chart

Version	Date	Comment
0.1	18.05.	First draft by ika and BMW
0.5	17.07.	Version for internal review
1.0	31.07.	Final Version of the interim report

Table of contents

List of figures

List of tables

Executive summary

This document provides the draft data concept for Pre-LSDemo based on the initial discussions within the project. The draft data concept serves two purposes. First, it establishes the foundation for future work within the two data-related tasks of Pre-LSDemo. Second, it supports proposals for the HORIZON-CL5-2026-10-D6-01 “Flagship-pilot: large-scale demonstrations of CCAM” call by addressing the topic of data collection and sharing appropriately. Thus, the data concept focused on input to a research / demonstration project with vehicles, which in most cases will be equipped with dedicated data loggers. The data concept is limited to this topic.

It is evident that data collection and sharing will play a central role in any large-scale demonstration project to, among others, enable demonstration tracking and evaluation. Therefore, these aspects need to be considered from the very beginning of the project.

The document is structured into three chapters. The first chapter discusses general considerations regarding stakeholder roles, data formats, data access, data storage, and links to related projects. The second chapter introduces the proposed data groups and discusses the challenges associated with sharing them. Data groups represent logical clusters of signals or information elements that belong together from an information perspective. The final chapter describes 27 different data purposes and analyses the relevance of the various data groups for each purpose.

The data purposes represent different options from which a large-scale project can choose and determine which are relevant to its objectives. These data purposes were identified through a survey conducted within the Pre-LSDemo project. The input for the different sections was collected through several consultations with Pre-LSDemo partners, including a large survey conducted during the first workshop and more detailed inquiries within the data-related work packages.

While the first section is relevant for all projects, Sections 2 and 3 are not intended to be read sequentially from beginning to end. Instead, they are intended to be consulted according to the data purposes that are relevant to a specific project. For example, if a project intends to investigate an in-vehicle technology, the suggested data groups are as follows:

- Relevant for data purpose: Experiment Status, Ego Vehicle Movement, Ego GNSS Information, Surrounding Objects Information, Abstract Map, Digital Map, Driver Input, Driver Status, AD & Technology Enabler Status, AD & Technology Information, AD & Technology Communication, AD & Technology Sensor, Environmental Status, Driving Scenario Status, Energy Consumption
- Unclear but are potentially relevant: Freight information, Vehicle & System Status, Operation, Video
- Not needed: External Statistics, Driver / Passenger Information, Passenger, Subjective Data

In a second step, it should be assessed how likely it is that a particular data group can be shared by the data provider with a data user. The answer depends on the category of stakeholders with whom the data shall be shared. In the example above, the data would be shared with consortium partners responsible for evaluation, i.e. an exclusive group of partners within the consortium. The tables below present the expected level of complexity to share this data group in a first example.

Sharing with exclusive group of partners	Data groups that are relevant	Data groups that are potentially relevant
Possible	Experiment Status, Ego Vehicle Movement, Ego GNSS Information, Surrounding Objects Information, Abstract Map, Digital Map, Driver Input, Driver Status, AD & Technology Enabler Status, AD & Technology Information, AD & Technology Communication, AD & Technology Sensor, Environmental Status, Driving Scenario Status, Energy Consumption	Freight information, Operation
Some issues expected		Vehicle & System Status, Video
Difficult		

For the technical evaluation of in-vehicle systems, sharing the required data with an exclusive group of consortium partners should, according to the opinion of the Pre-LSDemo partners, be feasible. Only the sharing of the data groups Vehicle & System Status and Video may present certain challenges, should these data groups be required.

In a second example, data sharing with external stakeholders is considered. The opinion of the Pre-LSDemo partners regarding the ease of sharing the different data groups is presented in the table below.

Sharing with external partners	Data groups that are relevant	Data groups that are potentially relevant
Possible	Energy Consumption	Operation
Some issues expected	Experiment Status, Ego Vehicle Movement, Ego GNSS Information, Surrounding Objects Information, Abstract Map, Digital Map, Driver	Freight information, Vehicle & System Status

	Input, Driver Status, AD & Technology Enabler Status, AD & Technology Communication, Environmental Status, Driving Scenario Status	
Difficult	AD & Technology Information, AD & Technology Sensor	Video

In this case, more challenges are expected in general. For some data groups, particularly AD & Technology Information, AD & Technology Sensor, and Video, data sharing is expected to be considerably more difficult. However, if this data needs to be shared, measures must be identified to enable sharing under conditions that are acceptable to the data providers. This process requires discussions among the parties involved. Possible solutions may include specific agreements between partners, anonymization, and/or the aggregation of information.

1 Introduction

The two main objectives of this report are to be the basis for future work within the two data-related tasks of Pre-LSDemo and to support proposals of the HORIZON-CL5-2026-10-D6-01 call on large scale demonstration of CCAM in defining the data topic. Against this background, it has been decided in the data responsible tasks of the CSA Pre-LSDemo to release an intermediate report. The intermediate report has become necessary in the opinion of the authors, since the due date of the originally planned milestone report is after the deadline of the call.

At the same time, it is clear that the topic data will and must play a central role for any large-scale demonstration project. Data will link the demonstration activities across the EU with each other as well as with supportive and evaluation activities. It could be even said that data and the data pipeline that covers next to the data collection the data processing, the quality check and storage will be the nerve system of the project.

The term "data" is used in a general sense and covers different types of (basic) information (signals, aggregated information, derived measures, questionnaire answers, etc.) as well as different contexts for the information, such as vehicle (e.g., velocity), infrastructure (e.g., network load), user (e.g., acceptance), operation (e.g., failure rate) and service (e.g., number of passengers) information.

The objective of the final data concept of Pre-LSDemo will be to provide answers to the following questions:

1. What data should be collected?
2. How should data be logged, processed (e.g. quality check) and stored?
3. Who needs access to what data, and when?
4. What are the current gaps and future needs?

The draft data concept in this document focuses on questions 1 and 3. It is important to note that this report focuses on input to research / demonstration project with vehicles, which in most case will be equipped with dedicated data loggers – and does not discuss the data sharing of customer vehicles.

Figure 1-1 gives an overview of the approach to define the data concept for a CCAM large-scale demonstration project. The process starts with defining the data purposes that might be relevant for a large-scale demonstration project. The data purposes are different applications for which data from the demonstrations are required. In a wider sense they represent different possible applications, from which a large-scale project can pick suitable ones and will receive its requirements in terms of data. Examples for data purposes are evaluation of user comfort, develop End-2-End AI systems, and derive Project KPIs. In the second step, the relevant data for each data purpose is identified. Instead of discussing individual data, the approach focuses on data groups that combine different data of the same kind. For each data group, it is then

determined whether it is required, optional or not needed for the defined data purposes. This leads to a draft data group list that is used for further discussion with data user stakeholders (T4.2) and data provider stakeholders (T5.3) regarding the technical requirements as well as the feasibility to provide the data. This discussion will be done after the draft data concept to be considered in the final data concept. Once a compromise between requirements and feasibility is found, the final data concept will be published in deliverable D5.1, which will contain the data list including minimum set of data to be shared in the project, additional data that is required per data purpose as well as proposals for the technical logging concept, databases and tools for the data handling.

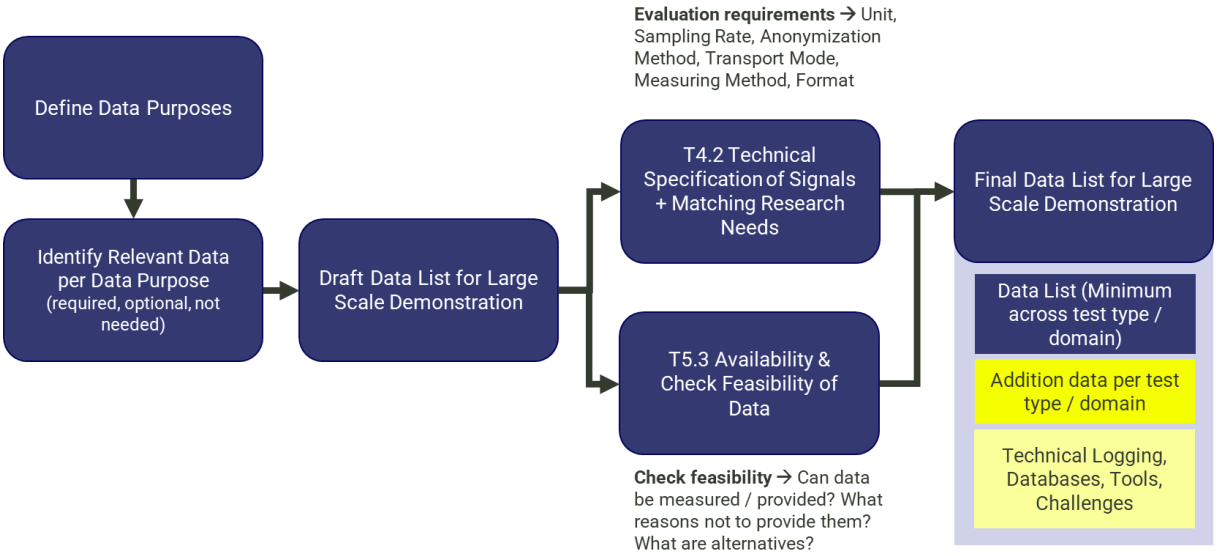

Figure 1-1: Approach to define the data concept for a large-scale demonstration project.

This document is structured in three main parts. It starts with general aspects of the data concept, including initial thoughts on roles, data formats, storage, access to data, and links to related research activities. The second part presents the data groups, describing the main categories of data that may be relevant for a large-scale demonstration project and outlining their characteristics, examples, sharing criticality, and logging aspects. The third part focuses on data purposes and explains for which objectives the data may be used, such as evaluation, monitoring, validation, database creation, or technology development, together with the related data requirements, access needs, and mapping to the relevant data groups.

1.1 Roles

It is important to be clear on the roles and responsibilities with respect to data, since a functioning data concept depends not only on the data itself, but also on the clear definition of responsibilities and interactions between the different actors involved. The focus of the roles defined in this data concept is on practical roles for data handling within a large-scale demonstration project. Other roles, e.g., related to GDPR (EU 2016) obligations are out of scope for this data concept as they are to be defined in individual data management plans. The basic

roles with respect to data processing in a large-scale demonstration project for CCAM is given in Figure 1-2.

Figure 1-2: Different roles in the data process of a large-scale demonstration project for CCAM.

The first main role are the data providers, who are generating or logging the data. Typically, the data providers are within the project and are running the demonstrations. OEMs, service providers, logistics companies, and public transport companies are examples for potential data providers. However, data providers can also be outside of the project and provide additional data that are needed to complete the picture (e.g. in safety impact assessment crash data are required, which are only available in crash databases).

The second main role are data users, who utilize the data in terms of analysing and evaluating it. The collected data can also be utilized for the future development of a system. Examples for data users are researchers and developers. Although the primarily focused data users are typically within the project, it needs to be considered that the data of the project are also relevant for persons outside of the project. The interaction with data users outside of the project need to be defined in the project in close collaboration with the individual data owners.

In addition, there is a third role with respect to data, which covers the data processing and governance. The tasks of this role are to technically maintain the data, perform quality checks, process for data aggregation, setup the data storage, organise access rights, ensure alignment between data providers and data users, and support a consistent operation of the data process. Here, partners with the technical capabilities and resources are required, these can either be industrial or research organisations. However, this role can only be fulfilled by a partner within the project – only certain sub-tasks might be sub-contracted. As some raw data provided by the data providers might contain GDPR-related data or be data critical for other reasons, such as company IP, bilateral contracts between data processors and data providers might need to be drawn up.

1.2 Data Structure

In a large-scale demonstration project, the definition of a suitable harmonized data structure is a critical enabler for the interoperability, exchange, and long-term reuse of the recorded data. A key design principle is the use of neutral and widely accepted standard formats, such as CSV (Comma-Separated Values) or JSON (JavaScript Object Notation), for structured data. These formats are text-based, human-readable, and supported by nearly all programming environments. This makes them suitable for heterogeneous research consortia. On the other hand, they reach limitations in efficient structuring and memory usage when being used for extensive data with complex structures. For these cases, other formats such as HDF5 (Hierarchical Data Format, Version 5) or MCAP might be used.

However, a single unified data format will not be sufficient to capture the full spectrum of data generated in the large-scale demonstration efficiently. For instance, CSV provides a simple and universal structure for tabular data, allowing easy exchange and inspection, while JSON offers greater flexibility to represent hierarchical or nested information, such as complex vehicle states or environmental descriptions. Importantly, both formats can be easily converted into other representations (e.g., databases, binary formats, or domain-specific data models), which ensures future compatibility and avoids vendor lock-in. Multimedia data such as video streams from onboard cameras impose fundamentally different requirements compared to structured sensors or metadata. Video data is usually stored in binary formats such as MP4 or AVI, which are optimized for efficient storage and playback of large volumes of image data but are not human-readable and cannot be integrated directly into text-based formats like CSV or JSON. In conclusion, a multi-format data strategy is required.

A pragmatic approach is to separate data types according to their characteristics while maintaining logical consistency through synchronization mechanisms. One synchronization mechanism is the consistent time stamp in all data formats. Structured data (e.g., vehicle signals, scenario descriptors, timestamps, or evaluation metrics) of short sequences can be stored in simpler data formats, whereas high-bandwidth sensor data such as video, LiDAR, or radar streams are stored in dedicated binary formats or container files (e.g., ROS bag or MCAP). Container formats are commonly used in robotics and automated driving to record time-synchronized multi-sensor data, including images and vehicle states, within a unified logging framework. At the same time, these formats often provide export capabilities into neutral formats, ensuring interoperability across different partners.

In addition, clear documentation of data structures, naming conventions, and synchronization (e.g., timestamps or unique identifiers linking video and metadata) is essential to ensure consistent interpretation across all stakeholders and to allow consistent usage from internal and external partners. This highlights to prioritise using already standardized data formats as well as data models, which have been developed and adopted before in other research projects.

In this regard, the CCAM SYNERGIES project provides a directly relevant reference, having operationalised these requirements within its data quality assessment methodology by defining

mandatory metadata fields covering dataset summary, administrative information, structural metadata including data field descriptions with units and sampling rates, and data quality policy, see for instance section 2.3.5 in Alakkad et al. 2025 .

In general, the alignment of data structures among (CCAM) projects can support the exchange of data and generation of knowledge to the benefit of the involved projects.

1.3 Data Access

The collected and structured data are expected be of interest to many stakeholders. Therefore, it is wise to define access roles from the early stage of the project. It might be even a topic to be discussed in the proposal phase to ensure a clear situation for all stakeholders from the beginning and to allow to address the topic properly also in the consortium agreement of the large-scale demonstration project.

Table 1-1: Examples roles and required access based on Bellotti et al. 2022

Name	Description	Required Access
System Admin	The system admin has the possibility of creating different databases in a multi-tenancy installation, and (only in case of need) has full access to all data entries. If pseudonymization is employed, the admin cannot resolve IDs (i.e., relating a data entry with its vehicle owner or driver)	Full access. Management of database and all users.
DB Admin	The DB admin can configure the DB (e.g., setting up the supported data types, users and rights) and (only in case of need) has full access to all their data entries. If pseudonymization is employed, the admin cannot resolve IDs (i.e., relating a data entry with its vehicle owner or driver)	Full access to a single DB tenant. Management of database and all users for single DB tenant.
Providers	Providers are typically the data owners. They can upload data and retrieve only their own data.	Full access (upload & download) to the own data.
Analysts	Analyst users cannot upload data but can see all the data of their typology. Different types of typologies can be defined. For instance, in L3Pilot three typologies were defined, matching the type of relevant data: Technical and Traffic analysts, Impact analysts, User analysts.	Full access (download only). Download anonymised data according to the needs.
Consortium Users	Member of the projects who want to analyse or use data for development and / or own non-project related analysis.	Access to data commonly shared within the consortium.

External User	External user that needs /requests data for a purpose outside of the large-scale project. There might be quite different interests. Therefore, it is hardly possible to specify the access right. Nevertheless, the process how to decide on the access request shall be clarified.	To be defined by the project depending on interaction with other projects.

The design of data access mechanisms in CCAM projects may take into consideration the strategic guidelines promoted by the EC, such as the CCAM Data Sharing Framework (DSF)[1] as a valuable reference for implementation phases. The DSF was recently published under the European coordination action FAME (continuing the work from previous projects like FOT-Net, CARTRE, and ARCADE), offering a structured methodology to address the legal, ethical and organizational challenges of sharing data from driving trials and operational environments.

From a technical perspective, the DSF highlights the emerging concept of Data Spaces as the technology framework to implement its paradigm of data sovereignty, mutual trust, and secure exchange. This federated approach and its practical implementations (like FAME's Test Data Space, or the upcoming FUSE4CCAM platform[2]) provide a basis benchmark for secure data sharing, showing potential benefits for CCAM large scale demonstrators.

1.4 Processing and Storage

The success of large-scale demonstration projects for CCAM depends on the efficient handling of vast and heterogeneous data generated across demonstration sites. Test vehicles equipped with multimodal sensor suites continuously produce massive data volumes which requires robust and scalable data storage and processing infrastructures. Efficient data pipelines are essential to ingest, preprocess, and harmonize data from multiple sources. Furthermore, compliant data management strategies—aligned with established principles such as FAIR (Findable, Accessible, Interoperable, Reusable)—ensure that the collected data can be effectively shared, reused, and preserved over the project lifecycle and beyond. Within the context of this project, the data storage and processing framework therefore provides the backbone for transforming collected data into project results, supporting both the technical development and the large-scale evaluation of CCAM.

[1] CCAM Data Sharing Framework: available at the Connected Automated Driving Portal - https://www.connectedautomateddriving.eu/data-sharing/ccam-data-sharing-framework/
[2] FUSE4CCAM is a project proposal elected for funding under the EU call HORIZON-CL5-2026-01-D6-06 - https://cordis.europa.eu/programme/id/HORIZON_HORIZON-CL5-2026-01-D6-06

A generic concept for data processing and storage in a large-scale demonstration project for automated driving is given in Figure 1-3.

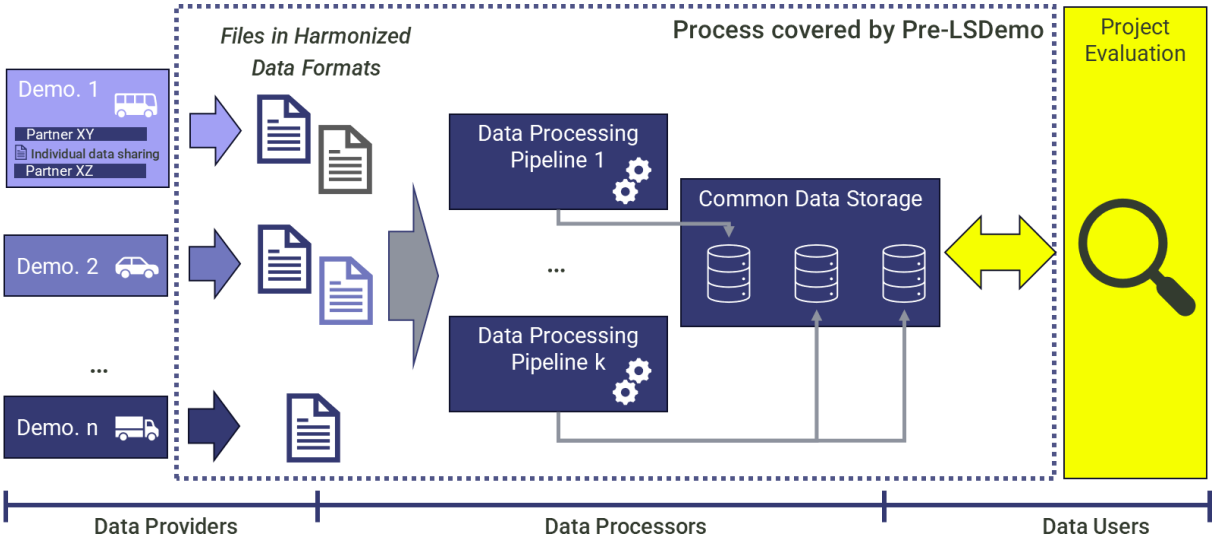

Figure 1-3: Generic concept for data processing and storage in a large-scale demonstration project for CCAM.

It is showing the different demonstrations that contribute with their data. The data are stored using harmonized data format. Different demonstrations might produce different data, which require own data formats (e.g. vehicle data vs. questionnaire data). For the different data, different data pipelines might be applied leading to a common data storage. Due to the different nature of data, it is rather unlikely that all data will be stored in one big database. It is rather expected that the database will consist of multiple instances which are dealing with different datatypes. At the same time, it needs to be recognized that there might be also bilateral or multilateral data sharing for individual demonstrations. These collaborations are not covered by the general data processing and storage pipeline. These collaborations need to be discussed individually. However, they can only be an addon and will not replace the central data approach.

Recommendations for the data processing and storage pipeline in a large-scale demonstration project for automated driving are:

- Data storage needs to be accessible for multiple partners considering the access right concept of the project. This includes also the potential access from users outside the project.
- The data storage must be able to handle the defined data formats.
- To avoid benchmarking of system, data / data providers shall be anonymised as far as possible. At the same time, there might be needs to identify after upload who provided the data (e.g. deletion requests by individuals outside of the project).

- The data logging in the demonstration is likely to use vehicle or service specific data logging solutions, since it is unlikely that all demonstrations will use the same data logging equipment. Thus, it is likely that the data in first instance is provided in different format. However, it is strongly recommended to switch as soon as possible to harmonized data formats, since it enables the usage of comment processing and storage scripts.
- Quality checks need to be implemented in the data pipeline to ensure a minimum level of quality and to avoid too heterogeneous data.
- Since the project will also collect personal data, GDPR requirements in terms of security and as well as in data handling (e.g., deletion of data in case of request) must be implemented.
- Since the project is funded by the European Commission, the data handling and storage shall be located in the EU. In any case European laws and regulation (e.g. GDPR) need to be adhered.
- The data processing itself might be conducted locally at the data providers or processors, as for some data it is not possible to share it with the whole consortium before anonymization or data aggregation.
- It is strongly recommended to set up the data processing and storage in the earliest phases of the project. Ideally, data processing and storage pipeline is set up in the first 6 month of the project. The reason for this is that the pipeline needs to be tested and finalized prior to the arrival of the first data from the demonstrations. Otherwise, there is the risk that data are stored at multiple locations and data handling for the involved partners gets more difficult.
- Using open-source tools and programming languages for the processing chain is preferred as it is likely that many different partners will implement and use it. To avoid issues with licenses for commercial tools, all components of the processing chain should be open.

1.5 Links to Other Research Projects

Due to its designated central role within the CCAM partnership and the fact that it will be a publicly funded project by the EU, the large-scale demonstration project is called upon to establish links to research activities. On the one hand, it is wise to continue the work of or to build upon previous research. On the other hand, the large-scale project and its outcomes can foster the activities of the other research projects.

Therefore, the large-scale demonstration project shall investigate in terms of data possible links to the following projects:

- IPCEI-AI project: projects to develop an AI-based foundation model and needs data for this purpose.
- EU-funded CCAM project SYNERGIES develops a federated European Scenario Dataspace for the virtual testing and validation of CCAM systems. Alignment with SYNERGIES would enable data reuse and access to established quality frameworks and tools. By mid-2027, SYNERGIES will deliver a tool marketplace with 50+ tools for data processing, scenario generation, and virtual validation
- EU projects founded in the call HORIZON-CL5-2026-01-D6-06 Federated CCAM data exchange platform.
- Previous research projects, such as ULTIMO (2026), MODI (2026), Hi-Drive (2026) and L3Pilot (2026).
- Previous data concepts / frameworks, such as Hidalgo et al. 2025, the CCAM Data Sharing FAME (CCAM 2024), the Connected Automated Driving Data Sharing website (2026) and the FESTA Handbook (2021).

2 Data Groups

2.1 Introduction

To facilitate efficient collaboration in the Pre-LSDemo project and to ensure a common understanding across its partners, Pre-LSDemo introduces the concept of data groups as an intermediate abstraction layer for discussion and alignment of data content. Data groups represent a logical cluster of signals or information elements that belong together from perspective of provided information. This should enable a focus on the higher-level concepts instead of individual signal definitions. By this, it is intended to streamline communication, reduce complexity, and avoid premature discussions at raw signal level, which can be too detailed and context specific. Detailed specifications and signal-level harmonization are intended to for the later stage of the project, in which they can be addressed in a more mature context. In total 23, data groups have been identified, and a short characteristic is given of each group in the results of the data sharing inquiry.

In general, there are some aspects to be considered when sharing data related to research activities. These aspects are:

- Using data for the re-engineering of systems / components need to be prevented as far as possible.
- Data is not intended to perform benchmarking on individual implementations of technologies by different partners. Technology should be assessed on general level regarding the performance and capabilities.
- For personal data, GDPR needs to be considered. This can limit the data sharing. It is important to discuss the challenges and measures for this as early as possible in the project.
- Video data is quite useful for many data purposes. Besides GDPR concerns the handling and recording comes with certain challenges. Typically, the data amount is quite high. It is not always allowed to record video (e.g. in factories or areas with critical infrastructure).
- Raw data from sensors, from HD maps, etc. may be subject to license restrictions.

To discuss which data groups can be shared between different stakeholders an expert inquiry among the Pre-LSDemo partners who are in the corresponding work package has been done. Partners were asked to rate each data type whether it can be shared with one of the four stakeholder groups. The stakeholder groups are:

- External partners: Any partner or group of partners that is not within the consortium.
- Consortium partners: Any partner or group of partners that is in the consortium.

- Exclusive group of partners (within the consortium): Partners that belong to defined sub-group of the consortium. A typical example are partners who need to do the evaluation in the project and are in same work package.
- Bilateral partner: A partner with whom the data provider has a special agreement to share data. These can be for instance a partner supporting the development, execution of demonstrations or the data processing of a demonstration.

The partners were asked to answer just for those data groups for which they feel confident to answer. Therefore, the number of received feedbacks by the Pre-LSDemo partners per data group (status June 2026) vary between 6 and 9. In a last step the feedback was reviewed by dedicated meeting with the Pre-LSDemo partners. The results per data group is presented in the following.

The results will serve as the basis for further discussions in Pre-LSDemo. The focus of theses discussion will be to break down the data groups further to identify signals and information that are most critical and to discuss potential measure for those to forester data sharing among partners. Based on this, the consortium of a large-scale demonstration project can determine which data groups are applicable for the objectives of their project. Then, for these groups, they can define which sharing is required and look up where potential issues might arise to solve them as early as possible. Thus, the following sections are not primarily made for reading them sequentially, but rather to use them as a lookup for data groups of interest for the individual project.

Important Remarks:

1. It is marked per data group for the four mentioned stakeholders' groups whether the data sharing is expected to be possible (no issues expected), some issues are expected to occur (issue needs to be solved), or data sharing is difficult to implement (several issues need to be solved).
2. It needs to be noted that although data sharing is expected to be difficult or issues are expected, it does not mean that it is impossible. It rather indicates that a large-scale demonstration project should pay high attention to this matter in case the data should be shared with the corresponding stakeholder group. There are multiple potential solutions (e.g., limit access right, anonymisation of data, payment, special agreements). The most appropriate / the consensus need to be decided by the project itself.
3. It is recommended to address issues related to data sharing prior to the start of the project.
4. Statements with respect to the expected feasibility of sharing signal are only related to a large-scale demonstration project. In other context, the feasibility to share data might be rated differently.

2.2 External Statistics

This data groups deals with external statistics that are used to complete missing data in the project. External statistics might be needed for assessment (e.g. impact assessment) or the definition of system or component requirements. This data can be helpful for:

- providing a broader context for interpreting experimental results,
- supporting safety and impact assessment using real-world reference data,
- complementing in-vehicle and operational data with external references.

Example data of this data groups are crash databases such as GIDAS (German In-Depth Accident Study) (GIDAS 2026), traffic flow data or prices of components.

Table 2-1: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	N/A	1, 2
	Consortium	Some issues expected	1
	Exclusive group	Possible	1
	Bilateral	Possible	1
Logging Frequency		N/A	Typically, no time series data

Comments:

1. Since this data group deals with external data and can cover quite different data, it needs always to be checked per data source whether the information can be shared with other partners and under which conditions. There are certainly data source, which cannot be shared directly with any partner, as they might be bound to specific licenses (e.g., access to GIDAS database is limited to certain partners). For the given answer, it is presumed that data sharing is possible.
2. Some partner stated that data sharing with externals will be an issue. It must be also considered that data sharing of external input under the project is not expected to be real use cases, since such data would typically not be collected by the project.

2.3 Experiment Status

The Experiment Status data group provides high-level metadata that describes the configuration, identification, and progression of an experiment within a large-scale demonstration project. It enables consistent tracking and interpretation of data across multiple test sites, use cases, and partners. The data do not change continuously during a trip but rather define the context in which the data were collected. The data can be useful for

- structuring and organizing experiments across large and heterogeneous set of demonstrations,
- check and review the status of certain experiments,
- supporting evaluation and comparison between different configurations, sites, or technologies.

Typical elements in this data group include identifiers such as Experiment ID, Run ID, and test condition labels (e.g., baseline vs. treatment), which allow each dataset to be clearly assigned to a specific experimental setup.

Table 2-2: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1, 2
	Consortium	Some issues expected	2
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Once per trip	Typically, no time series data

Comments:

1. One or more partners stated that this sharing is difficult.
2. Concerns exist for certain data (e.g. location, time) that they enable the possibility of identifying certain partners.

2.4 Ego Vehicle Movement

The Ego Vehicle Movement data group captures the dynamic motion state of the ego vehicle during an experiment. It includes time-dependent, quantitative signals that describe how the vehicle moves and evolves in space over time. The data can be useful for:

- reconstructing vehicle trajectories and motion profiles,
- evaluating automated driving performance and control behaviour,
- detecting and analysing driving scenarios and events,
- linking vehicle dynamics to other data groups such as environment, driver input, or system status.

Typical elements comprise variables such as speed, acceleration, heading, and yaw rate, which together provide a detailed representation of the vehicle's longitudinal and lateral motion behaviour.

Table 2-3: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Medium (e.g. 10 Hz)	

2.5 Ego GNSS Information

The Ego GNSS Information data group captures the absolute geographic position of the ego vehicle based on satellite navigation systems. It provides the spatial reference needed to locate the vehicle within the real-world environment during an experiment. The data can be useful for:

- georeferencing vehicle trajectories and experimental data,
- validating localization performance of automated driving systems,
- linking ego vehicle data with map, environment, and infrastructure information,
- assessing positioning accuracy and reliability in different operational conditions.

This data group generally consists of time-dependent positioning data, which can be used to reconstruct the driven route and align vehicle data with maps, infrastructure, and external context information.

Typical elements include latitude, longitude, and GNSS quality indicators, which describe both the position of the vehicle and the reliability of the positioning signal. Additional related signals (e.g., heading or positioning quality metrics) may further support the interpretation of the vehicle's location.

Table 2-4: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	1

	Exclusive group	Possible	1
	Bilateral	Possible	1
Logging Frequency		Low (e.g. 10 Hz)	

Comments:

1. It needs to be noted that ego GNSS Information data are very likely to be GDPR relevant since they provide position and time of the vehicle. This aspect needs to be considered, and measures need to be taken for this data group to adhere to the GDPR.

2.6 Surrounding Objects Information

The Surrounding Objects Information data group captures the state and behaviour of objects in the environment around the ego vehicle, including vehicles, vulnerable road users (VRUs), and other relevant traffic participants. Data can be provided e.g. by vehicle sensors, V2X information sent from other vehicles and road beacons. The data can be useful for:

- understanding interactions between the ego vehicle and other traffic participants,
- reconstructing traffic scenarios and situational context,
- evaluating perception and tracking performance of automated driving systems,
- supporting safety and behaviour analysis, particularly in critical or complex situations.

This data group typically consists of time-dependent, object-based information, where multiple objects are tracked simultaneously over time. It forms a key representation of the dynamic traffic situation encountered during an experiment.

Typical elements include relative position (e.g., distance in longitudinal and lateral direction), relative velocity, and object classification, which together describe how surrounding objects are located and moving with respect to the ego vehicle. Data from sensors can be subject to license restrictions.

Table 2-5: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Medium (e.g. 10 Hz)	

2.7 Abstract Map

The Abstract Map data group provides a structured representation of the static road and infrastructure environment in which the ego vehicle operates. It contains map-based information that describes the road layout and relevant traffic rules independently of the vehicle's sensor perception. This data group can be helpful for

- providing contextual information about the road network and infrastructure,
- enabling interpretation of driving behaviour with respect to traffic rules and road geometry,
- facilitating scenario identification and classification.

Typical elements include attributes such as road type, speed limits, number of lanes, and other abstracted infrastructure characteristics, which define the driving context at a given location.

Table 2-6: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Possible	1
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Low (e.g. 1 Hz)	

Comments:

1. Concerns exist for certain data (e.g. location, time) that they enable the possibility of identifying certain partners.

2.8 Digital Map

The Digital Map data group provides a detailed and structured representation of the road environment based on digital map sources, typically including high-definition (HD) maps used by automated driving systems. This data group can be useful for:

- providing detailed road geometry and infrastructure information,
- supporting planning and decision-making of automated driving systems,
- validating map-based functions and sensor fusion approaches.

Typical elements include precise lane geometries, road boundaries, traffic signs, and other map features, which go beyond abstract attributes by offering a geometrically accurate and enriched description of the driving environment. This information is often derived from a combination of GNSS, sensors, and mapping processes to support high-precision localisation. HD map data can be subject to license restrictions.

Table 2-7: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Low (e.g. 1 Hz)	

2.9 Driver Input

The Driver Input data group captures the control actions applied by the human driver to the vehicle during an experiment. It represents how the driver actively influences vehicle behaviour through direct interaction with the driving controls. This data group can be useful for:

- providing baseline for designing driving scenarios and model driver behaviour,
- identifying take-over actions and driver interventions in automated driving,
- comparing human driving performance with automated system behaviour,
- supporting safety and human factors evaluation.

Typical elements include signals such as steering wheel angle, brake pedal position, throttle (accelerator) input, and other driver interventions, which describe the driver's control commands at a given point in time.

Table 2-8: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	

Logging Frequency	Medium (e.g. 10 Hz)

Comments:

1. One or more partners stated that this sharing is difficult.

2.10 Driver Status

The Driver Status data group captures the state and condition of the human driver during an experiment, providing information about the driver's readiness, attention, and overall ability to supervise or control the vehicle. This data group can be helpful for:

- assessing driver readiness and take-over capability in automated driving scenarios,
- analysing human factors such as attention, workload, and fatigue,
- understanding driver–system interaction, especially in transitions between automated and manual control,
- supporting safety evaluation, particularly for Level 2–3 systems where driver supervision is required.

Typical elements may include indicators related to driver attention, gaze behaviour, engagement, fatigue, or availability for take-over, which describe whether and how the driver can interact with the vehicle at a given moment.

Table 2-9: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	1
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Medium (e.g. 10 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.

2.11 Driver / Passenger Information

The Driver / Passenger Information data group captures general characteristics and contextual information about the occupants of the vehicle, including both the driver and any passengers involved in an experiment. This data group can be useful for:

- characterising the participant population in large-scale demonstrations,
- supporting analysis of user-related effects, such as comfort, acceptance, and behaviour,
- enabling segmentation and comparison across different user groups,
- providing context for interpreting driver behaviour and system interaction data.

Typical elements may include attributes such as demographic information (age, gender etc.), role (e.g., driver, passenger, safety driver), experience level, or study-related classifications, which help to describe who is participating in the experiment.

Table 2-10: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Difficult	1
	Consortium	Some issues expected	1, 2
	Exclusive group	Possible	1, 2, 3
	Bilateral	Possible	1
Logging Frequency		Once per Trip	Typically, no time series data

Comments:

1. One or more partners stated that this sharing is difficult.
2. Some data (e.g. age) are GDPR relevant. This could lead – if not treated appropriately – to issue. It needs to be carefully considered which data is shared with whom under which agreements to ensure the compliance with GDPR.
3. It is presumed that for a smaller group of partners different measure with respect to GDPR can be taken. However, in principle the same concerns exist as for the stakeholder group “Consortium”.

2.12 Freight information

he Freight Information data group captures contextual and operational information related to transported goods within automated driving experiments, particularly for logistics and commercial use cases. This data can be useful for:

- characterising logistics-specific use cases and operational conditions,
- analysing the impact of load on AD performance and behaviour,
- supporting evaluation of automated driving in freight and logistics scenarios,
- enabling comparison across different transport missions and load configurations.

Typical elements may include attributes such as type of goods, load characteristics (e.g., weight, volume), loading conditions, and mission-related information (e.g., delivery route or logistics operation type), which describe the cargo carried by the vehicle and its relevance for the experiment.

Table 2-11: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Possible	1
	Exclusive group	Possible	1
	Bilateral	Possible	1
Logging Frequency		Once per Trip	Typically, no time series data

Comments:

1. Data need to be treated in compliance with GDPR.

2.13 Passenger

The Passenger data group captures information related specifically to passengers within the vehicle, focusing on their presence, role, and interaction with the automated driving system during an experiment. This data can be helpful for:

- analysing passenger-related aspects of automated driving, such as comfort and acceptance,

- understanding in-vehicle interaction in passenger use cases (e.g., robotaxi, shuttle services),
- supporting user-focused evaluation across different occupant roles.

Typical elements may include attributes such as number of passengers, seating position, role (e.g., test participant), and interaction or usage behaviour, which help to describe how passengers are involved in the experimental context.

Table 2-12: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Once per Trip	Typically, no time series data

Comments:

1. One or more partners stated that this sharing is difficult.

2.14 AD & Technology Enabler Status

The AD & Technology Enabler Status data group captures the current operational state of the automated driving system (AD) and its associated technology enablers during an experiment. It provides information on whether specific functions or supporting technologies are active, inactive, or operating under certain conditions. These data are useful for:

- understanding when and how automated driving functions are active,
- analysing transitions between manual and automated control,
- evaluating the performance and reliability of technology enablers,
- supporting interpretation of vehicle behaviour in relation to system state.

Typical elements include indicators such as the ODD status, activation status of the automated driving function, availability of enablers, system modes (e.g. MRM or enabler status), or transitions between different operational states, which describe how the system is functioning at a given time. Data from this group can be subject to license restrictions.

Table 2-13: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Some issues expected	
	Bilateral	Possible	
Logging Frequency		Medium (e.g. 10 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.

2.15 AD & Technology Information

The AD & Technology Information data group captures descriptive and technical information about the automated driving system (AD) and its associated technology enablers used within an experiment. It provides a structured overview of the system components, configurations, and capabilities beyond their real-time operational status. The data group can be useful for:

- documenting the technical characteristics and configuration of the system under test,
- supporting interpretation of performance and behaviour in relation to system design,
- providing traceability and reproducibility of experimental results.

This data group typically consists of static metadata, representing the technical setup of the system in a demonstration rather than its real-time behaviour.

Typical elements may include ODD description, system type, function descriptions, configuration parameters, used software or hardware versions, and characteristics of the deployed technology enablers. Data from this group can be subject to license restrictions.

Table 2-14: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Difficult	1
	Consortium	Possible	1

	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Once per Trip	Typically, no time series data

Comments:

1. One or more partners stated that this sharing is difficult.

2.16 AD & Technology Communication

The AD & Technology Communication data group captures information on the properties of data exchanged between the automated driving system, external systems, and other entities during an experiment. It represents how the vehicle communicates with its environment and other systems to support automated driving functions. The data group can be helpful for:

- analysing the role of communication in automated driving behaviour and performance,
- evaluating connectivity-based functions such as V2X and cooperative driving.
- supporting validation of communication reliability and its impact on system performance.

Typical elements may include bandwidth information, latency, packet loss and transmitted messages (e.g. hazard information). Data from this group can be subject to license restrictions.

Table 2-15: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		High (e.g. 100 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.

2.17 AD & Technology Sensor

The AD & Technology Sensor data group captures information provided by the sensors used by the automated driving system and its technology enablers during an experiment. It represents the raw or pre-processed perception data that forms the basis for environment understanding. The data group can be useful for:

- analysing perception capabilities and sensor performance,
- provide input for future developments of AD and sensor technologies,
- supporting object detection, tracking, and environment modelling,
- evaluating sensor fusion approaches and robustness under different conditions.

Typical elements may include data from cameras, radar, LiDAR, ultrasonic sensors, and other perception systems, as well as derived sensor outputs such as detected objects (e.g. raw information or object list), signal quality indicators, or sensor-specific measurements.

This data group typically consists of high-frequency, time-dependent data, reflecting the continuous sensing of the vehicle's surroundings. Data from this group can be subject to license restrictions.

Table 2-16: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Difficult	1,2
	Consortium	Some issues expected	1,2
	Exclusive group	Possible	1,2
	Bilateral	Possible	2
Logging Frequency		High (e.g. 100 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.
2. Requires more detail discussions regarding the development requirements, since there are different levels of sensor / sensor raw data. Information from some sensor (e.g. radar) or at a certain level of detail might be considered as IP relevant and cannot be shared for that reason.

2.18 Vehicle & System Status

The Vehicle & System Status data group captures the operational condition and internal state of the vehicle and its onboard systems during an experiment. It provides information on how the vehicle and its technical subsystems are functioning at a given point in time. This data group can be useful for:

- monitoring the operational condition,
- identifying system faults, warnings, and abnormal behaviour,
- supporting interpretation of automated driving performance in relation to vehicle's status,
- ensuring traceability and validation of experimental results.

Typical elements may include indicators such as vehicle signals (e.g., indicator status, horn), system states (e.g. ACC, ESC), warnings, faults, and availability of vehicle functions.

This data group typically consists of time-dependent status signals, reflecting changes in vehicle and system conditions throughout a trip or test run.

Table 2-17: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Medium (e.g. 10 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.

2.19 Operation

The Operation data group captures contextual and operational information about how the vehicle is deployed and used within the experiment. It describes the organisational and functional setup of the mobility service or test activity in which the automated driving system operates. The data group can be helpful for:

- providing context on how and where automated driving is applied,

- characterising use cases and deployment scenarios across different domains,
- supporting evaluation of operational performance and scalability.

Typical elements may include attributes such as type of operation (e.g., public transport, logistics service), operational location, service characteristics, and timing information (e.g., operation duration, delay).

Table 2-18: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Possible	1
	Consortium	Possible	
	Exclusive group	Possible	
	Bilateral	Possible	
Logging Frequency		Low (e.g. 1 Hz)	

Comment:

1. In general, the data group is considered as not critical. However, there are data (e.g., location related information) that could be used to identify certain partner / demonstrations site. This could make the sharing to these data difficult.

2.20 Environmental Status

The Environmental Status data group captures information about the external environmental conditions in which the ego vehicle operates during an experiment. It describes the surrounding physical context that can influence both vehicle behaviour and system performance. This data group can be useful for:

- understanding the influence of environmental factors on automated driving performance,
- evaluating system robustness under varying weather, lighting, and road conditions,
- supporting scenario interpretation and classification,
- providing context for data used for the development of AD and its components.

Typical elements include parameters such as light conditions, weather conditions, temperature, and road surface conditions, which characterise the environment at a given moment.

Table 2-19: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Possible	1, 2
	Consortium	Possible	2
	Exclusive group	Possible	2
	Bilateral	Possible	2
Logging Frequency		Low (e.g. 1 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.
2. Environmental condition might allow track where the vehicle is operating. Therefore, the information could be GDPR relevant.

2.21 Driving Scenario Status

The Driving Scenario Status data group captures the current classification and occurrence of driving scenarios during an experiment. It indicates which type of scenario the ego vehicle is experiencing at a given point in time. This data group can be helpful for:

- identifying and segmenting relevant driving scenarios within recorded data,
- supporting scenario-based analysis and evaluation of automated driving systems,
- detecting critical or edge-case situations for safety assessment,
- linking vehicle, environment, and system data to specific traffic contexts.

Typical elements include indicators for scenarios such as lane changes, car-following, critical situations, or other predefined or detected events (e.g. edge case, critical situation, crash), which describe the dynamic traffic context in a structured way.

Table 2-20: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1
	Consortium	Some issues expected	
	Exclusive group	Possible	

	Bilateral	Possible
Logging Frequency		Medium (e.g. 10 Hz)

Comments:

1. One or more partners stated that this sharing is difficult.

2.22 Video

The Video data group captures visual recordings from cameras deployed on the vehicle or in its surrounding environment during an experiment. It provides representation of the driving situation as perceived visually by the camera. Video data can be helpful for:

- providing ground truth and contextual reference for interpreting other data sources,
- supporting validation of perception, detection, and scenario classification,
- the interpretation of driving scenarios,
- enabling qualitative analysis of driving behaviour and interactions.

Typical elements include footage from front, rear, or interior cameras, as well as potentially infrastructure-based cameras, which document the traffic environment, road conditions, and interactions with other road users.

This data group consists of time-dependent, high-volume data, often recorded continuously or event-based, depending on the experimental setup. Depending on video quality, license plates and faces can be identifiable, which is GDPR relevant.

Table 2-21: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Difficult	1
	Consortium	Some issues expected	1
	Exclusive group	Some issues expected	
	Bilateral	Some issues expected	
Logging Frequency		Medium / High (e.g. 25 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.

2.23 Subjective Data

The Subjective Data data group captures self-reported perceptions, opinions, and experiences of participants involved in the experiment. It reflects how users subjectively evaluate automated driving functions, rather than objective measurements from sensors or vehicle systems. This data group can be useful for:

- assessing user acceptance and trust in automated driving systems,
- understanding perceived comfort, safety, and usability,
- complementing objective performance data with human experience,
- supporting human factors and societal readiness evaluation.

Typical elements include data collected through questionnaires, interviews, ratings, or feedback, covering aspects such as comfort, perceived safety, trust, usability, and acceptance of the automated driving system.

Table 2-22: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1,2
	Consortium	Some issues expected	2
	Exclusive group	Possible	2
	Bilateral	Possible	2
Logging Frequency		Once per Trip	Typically, no time series data

Comments:

1. One or more partners stated that this sharing is difficult.
2. Data needs to be handled in compliance with GDPR.

2.24 Energy Consumption

The Energy Consumption data group captures information about the energy usage of the vehicle and its automated driving system during an experiment. It provides insight into how much energy is required to operate the vehicle under different driving and operational conditions. This data group can be helpful for:

- evaluating the efficiency of automated driving functions and vehicle operation,
- assessing the impact of automation on energy usage and sustainability,
- Supporting development of optimized driving strategies and system design.

Typical elements may include metrics such as fuel or electrical energy consumption, power usage of vehicle subsystems, battery state (e.g., state of charge), and energy efficiency indicators, which describe the energy demand of the vehicle and its components.

Table 2-23: Information about data group

Category		Answer	Comment
Feasibility for sharing			
	External	Some issues expected	1, 2
	Consortium	Some issues expected	1, 2
	Exclusive group	Some issues expected	2
	Bilateral	Some issues expected	2
Logging Frequency		Medium (e.g. 10 Hz)	

Comments:

1. One or more partners stated that this sharing is difficult.
2. For the interpretation of the data the context under which the data is recorded is essential to avoid misinterpretation. Therefore, the data group requires further data and should not be analysed alone.

2.25 Summary on Expected Feasibility of Sharing Data

The following table summarises the expected feasibility of data sharing per data group. Please look also for the individual concerns with respect to the feasibility of data sharing the sections above.

Table 2-24: Overview data sharing feasibility per data group.

Data group	External	Consortium	Exclusive group	Bilateral
External Statistics	N/A	Some issues expected	Possible	Possible
Experiment Status	Some issues expected	Some issues expected	Possible	Possible
Ego Vehicle Movement	Some issues expected	Some issues expected	Possible	Possible
Ego GNSS Information	Some issues expected	Some issues expected	Possible	Possible
Surrounding Objects Information	Some issues expected	Some issues expected	Possible	Possible
Abstracted Map Data	Some issues expected	Possible	Possible	Possible
Digital Map	Some issues expected	Some issues expected	Possible	Possible
Driver Input	Some issues expected	Some issues expected	Possible	Possible
Driver / Passenger Information	Difficult	Some issues expected	Possible	Possible
Freight information	Some issues expected	Possible	Possible	Possible
Passenger	Some issues expected	Some issues expected	Possible	Possible
AD & Technology Enabler Status	Some issues expected	Some issues expected	Possible	Possible
AD & Technology Information	Difficult	Possible	Possible	Possible
AD & Technology Communication	Some issues expected	Some issues expected	Possible	Possible
AD & Technology Sensor	Difficult	Some issues expected	Some issues expected	Possible
Vehicle & System Status	Some issues expected	Some issues expected	Possible	Possible
Operation	Possible	Possible	Possible	Possible
Environmental Status	Possible	Possible	Possible	Possible

Data group	External	Consortium	Exclusive group	Bilateral
Driving Scenario Status	Some issues expected	Some issues expected	Possible	Possible
Video	Difficult	Some issues expected	Some issues expected	Possible
Subjective Data	Some issues expected	Some issues expected	Possible	Possible
Energy Consumption	Some issues expected	Some issues expected	Some issues expected	Possible

3 Data Purposes

3.1 Introduction

The data purpose describes the intended use of the data recorded within a large-scale demonstration project. For example, the data purpose "Project KPIs" aims to use collected data to derive an overview of project achievements, such as the total mileage driven with and without automated driving. In the following, multiple data purposes are presented. These data purposes were collected through an inquiry conducted within the Pre-LSDemo project. In a broader sense, the data purposes represent different options from which a large-scale demonstration project may choose, depending on its objectives and requirements. Once the relevant data purposes have been selected, their mapping to the corresponding data groups indicates which data need to be collected within the project.

The approach for the defining the data-purposes followed a three-stage approach. In the first stage, an inquiry was conducted among the Pre-LSDemo partners involved in the data-related work packages to identify relevant data purposes. The presented list may not be exhaustive, but it covers the purposes considered most relevant for large-scale demonstration projects in automated driving.

The second stage consisted of a survey carried out during the first Pre-LSDemo stakeholder workshop. Participants were asked to assess the relevance of the data groups (see Chapter ●) for specific data purposes. Each participant was assigned three randomly selected data purposes. In total, 77 participants completed the questionnaire. The results were calculated as weighted average scores based on the responses received. However, these results represent a general opinion, and not all participants could be expected to have expert knowledge of the data purposes assigned to them.

Therefore, in the third stage, the same questions were answered by Pre-LSDemo experts from the data-related work packages. Experts were asked to respond only to those data purposes for which they felt sufficiently confident to provide an informed assessment. As a result, the number of responses varied across data purposes, ranging from five responses for "System-of-Systems & Logistics Integration" to fourteen responses for "Driving Scenarios". For most data purposes, responses from nine or more experts were obtained.

The final determination is primarily based on the expert assessments. Where a high level of agreement among experts was observed, the corresponding results are presented directly in this document. In cases where expert opinions did not converge sufficiently, the respective topics were discussed in greater detail within the work package, considering the broader stakeholder opinions obtained during the first Pre-LSDemo workshop.

Each data purpose is discussed individually, following a common structure. The first section provides a short introduction to the respective data purpose and assigns it to one or more of the three overarching categories: "Evaluation", "Monitoring / Validation / Database", and "Design /

Development". The second section examines whether the data purpose requires baseline-driving data, automated-driving data, or both. This distinction has direct implications for study design and data-logging requirements. Finally, the data groups that, in the opinion of the Pre-LSDemo experts, should be available for the respective data purpose are presented. The question of which stakeholder groups should have access to the data associated with a data purpose was discussed within Pre-LSDemo. This discussion was conducted independently of the question of how likely it is that the data can be shared with these groups (see Chapter ●). However, access requirements are often highly dependent on the specific questions in the data purpose. Therefore, only cases for which a clear consensus was identified are reported in this document.

For easier navigation the data purposes are linked below:

- Societal Readiness
- User Comfort
- User Acceptance
- User Interaction (in vehicle)
- User Interaction (outside vehicle)
- Technology (in vehicle)
- ODD Assessment
- Onboard Sensing
- Infrastructure Sensing
- Behaviour Planning
- V2X Function
- Develop End-to-End AI
- Public Dataset AD Development
- Operation Monitoring
- Driving Scenarios
- Safety – Crashes
- Define Reference Driver
- System-of-Systems & Logistics Integration

- Technology (outside vehicle)
  - Public Dataset AD Development
  - Reliability, Availability & Maintainability
- Technology Positioning
  - Reference Data for Ego Vehicle
  - Regulatory & Compliance Readiness
- Table 3-24: Mapping data groups and data purpose.
  - Public Dataset AD Development
  - Development of digital services

Data Group	Relevant for data purpose?
External Statistics	No
Experiment Status	Yes
Ego Vehicle Movement	Yes
Ego GNSS Information	Yes
Surrounding Objects Information	Yes
Abstract Map	Yes
Digital Map	Yes
Driver Input	Maybe
Driver Status	No
Driver / Passenger Information	No
Freight information	No
Passenger	No
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	Maybe
Vehicle & System Status	No
Operation	No
Environmental Status	Yes
Driving Scenario Status	Maybe
Video	No
Subjective Data	No
Energy Consumption	No

- Impact Assessment

- Reference Data for Ego Vehicle	- Project KPIs	- Control centre / remote operator	

3.2 Societal Readiness

Data used to determine the changes in societal readiness based on the project developments. The data helps the project to understand how society (people, communities, stakeholders etc.) feel about and perceive the new technology — e.g., whether they know about it, trust it, accept it, and see it as useful or problematic.

All stakeholder (public, consortium etc.) should have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-1: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-2: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-3: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	No	
Ego GNSS Information	No	
Surrounding Objects Information	No	
Map	No	
Driver Input	No	
Driver Status	No	

Driver / Passenger Information	No
Freight information	Yes
Passenger	No
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	No
Vehicle & System Status	No
Operation	No
Environmental Status	Yes
Driving Scenario Status	No
Video	No
Subjective Data	Maybe
Energy Consumption	Yes

3.3 User Comfort

Data used to evaluate user's comfort while using the technology. Thus, data is required to understand how comfortable people feel while using the technology, for example whether the experience feels smooth, safe, and pleasant, or whether it causes discomfort or unease. It can be assessed for instance by means of vehicle data (e.g., accelerations) or questionnaire.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the data type, required data types, who needs to have access to data and the mapping to data groups is given in the tables below.

Table 3-4: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-5: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-6: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Maybe	
Surrounding Objects Information	Maybe	
Abstract Map	Yes	
Digital Map	No	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	Yes	
Freight information	No	
Passenger	Yes	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	No detailed (e.g., sensor) information is required, but general information about the AD (e.g. ODD)
AD & Technology Communication	No	
AD & Technology Sensor	No	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Maybe	
Subjective Data	Yes	
Energy Consumption	No	

3.4 User Acceptance

Data used to evaluate user's acceptance of technology (e.g., are users willing to use the technology, trust it, and would choose to use it in everyday life?), e.g. by means of questionnaire after use.

All stakeholder (public, consortium etc.) should have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-7: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-8: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-9: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Maybe	
Surrounding Objects Information	Maybe	
Abstract Map	Maybe	
Digital Map	No	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	Yes	
Freight information	No	
Passenger	Yes	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	No detailed (e.g., sensor) information is required, but general information about the AD (e.g. ODD)
AD & Technology Communication	No	
AD & Technology Sensor	No	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	No	
Subjective Data	Yes	
Energy Consumption	No	

3.5 User Interaction (in vehicle)

Data used to evaluate the user interaction inside the vehicle with a technology, e.g. by means of questionnaires and observations. Basically, data is required that allow to understand how people use and interact with the features inside the vehicle, such as buttons, screens, or voice controls, and whether these are easy or difficult to use.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-10: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-11: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-12: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	No	
Surrounding Objects Information	No	
Abstract Map	Maybe	
Digital Map	No	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	Yes	
Freight information	No	
Passenger	Yes	

AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	No detailed (e.g., sensor) information is required, but general information about the AD (e.g. ODD)
AD & Technology Communication	No	
AD & Technology Sensor	No	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Yes	
Subjective Data	Yes	
Energy Consumption	No	

3.6 User Interaction (outside vehicle)

Data used to evaluate the user interaction outside the vehicle with a technology, e.g. by means of questionnaires and observations. Basically, data are required that helps us understand how people outside the vehicle (including VRUs other traffic participants) interact with it, such as whether they notice it, understand its signals, and feel confident about how it behaves.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-13: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-14: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-15: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	No	
Surrounding Objects Information	Yes	
Abstract Map	Maybe	
Digital Map	No	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	No detailed (e.g., sensor) information is required, but general information about the AD (e.g. ODD)
AD & Technology Communication	No	
AD & Technology Sensor	Maybe	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Yes	
Subjective Data	Yes	
Energy Consumption	No	

3.7 Technology (in vehicle)

Data used to determine the technical effects of an in-vehicle technology, e.g., faster reaction, earlier detection of objects, changes in gap keeping, less critical driving. The data should enable us to check whether the vehicle's technology works properly, behaves as expected while driving or improves the operation with AD.

Technology data are not supposed to be shared with the public to avoid benchmarking.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-16: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-17: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-18: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	Maybe	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Yes	
AD & Technology Sensor	Yes	
Vehicle & System Status	Yes	
Operation	Maybe	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Maybe	Is helpful, but GDPR and other restriction need to be considered
Subjective Data	No	
Energy Consumption	Yes	

3.8 Technology (outside vehicle)

Data used to determine the technical effects of a technology, that is placed outside of the vehicle and/or relies on communication, e.g., earlier detection of objects, less critical driving. The data should enable us to check whether the vehicle's technology works properly, behaves as expected while driving or improves the operation with AD.

Technology data are not supposed to be shared with the public to avoid benchmarking.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-19: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-20: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-21: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	Maybe	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	

AD & Technology Communication	Yes
AD & Technology Sensor	Yes
Vehicle & System Status	Yes
Operation	No
Environmental Status	Yes
Driving Scenario Status	Yes
Video	Yes
Subjective Data	No
Energy Consumption	Maybe

3.9 Technology Positioning

Data used to assess the quality, reliability, and degradation patterns of GNSS signals in operational environments. Includes position accuracy, signal losses, multipath effects, and correction data performance.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-22: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-23: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-24: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	

Digital Map	Yes
Driver Input	Maybe
Driver Status	No
Driver / Passenger Information	No
Freight information	No
Passenger	No
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	Maybe
Vehicle & System Status	No
Operation	No
Environmental Status	Yes
Driving Scenario Status	Maybe
Video	No
Subjective Data	No
Energy Consumption	No

3.10 Impact Assessment

Data used to determine the technology's impact on the traffic safety, efficiency, mobility, etc. Data is required to enable simulations (parameterization & validation) and to scale up results on European level.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-25: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-26: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-27: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	No	
Driver Input	Yes	
Driver Status	No	
Driver / Passenger Information	Yes	
Freight information	Yes	
Passenger	Yes	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Maybe	
AD & Technology Sensor	Maybe	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	No	
Subjective Data	Yes	
Energy Consumption	Yes	

3.11 Reference Data for Ego Vehicle

Data gathered from external spectators like surrounding vehicles or infrastructure as a reference for, e.g., localisation and sensor validation.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-28: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-29: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-30: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	No	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Maybe	
AD & Technology Sensor	Yes	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Yes	
Driving Scenario Status	Maybe	
Video	Maybe	
Subjective Data	No	
Energy Consumption	No	

3.12 ODD Assessment

Data should allow to identify where and when a future AD technology can be used. For this purpose, the information about ODD elements (traffic, environmental, digital, infrastructure data) occurrence and the potential usage is required.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-31: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-32: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-33: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Yes	
AD & Technology Sensor	No	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Yes	
Subjective Data	No	
Energy Consumption	No	

3.13 Onboard Sensing

Data is used to develop new / improve existing onboard sensing technology. The data need to help developers to understand what the vehicle can detect in real world conditions and how this detection can be made more accurate and reliable - in particular in situations in which the technology is limited today such as adverse weather conditions (e.g. by proving images with ground-truth information).

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-34: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-35: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes2

Table 3-36: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Maybe	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Maybe	

AD & Technology Sensor	Yes
Vehicle & System Status	Maybe
Operation	No
Environmental Status	Yes
Driving Scenario Status	Maybe
Video	Yes
Subjective Data	No
Energy Consumption	No

3.14 Infrastructure Sensing

Data is used to improve sensing technology (e.g., roadside detection modules, collective perception nodes, and ground truth references) that is installed in traffic infrastructure, such as roads or intersections, by helping developers understand how well this infrastructure can detect vehicles, pedestrians, and traffic situations in real world conditions.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-37: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-38: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-39: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Maybe	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	

Digital Map	Maybe
Driver Input	No
Driver Status	No
Driver / Passenger Information	No
Freight information	No
Passenger	No
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	Yes
Vehicle & System Status	No
Operation	No
Environmental Status	Yes
Driving Scenario Status	Maybe
Video	Yes
Subjective Data	No
Energy Consumption	No

3.15 Behaviour Planning

Data is used to help develop / extend the decision-making logic / trajectory planning of the vehicle, so it can choose appropriate driving actions—such as slowing down, stopping, turning, or changing lanes—based on what is happening in traffic.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-40: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-41: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-42: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	Yes	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Yes	
AD & Technology Sensor	Yes	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Yes	
Subjective Data	No	
Energy Consumption	No	

3.16 V2X Function

Data is used to develop technologies that allows vehicles to exchange information with other vehicles, traffic infrastructure, and road users, so they can better understand what is happening beyond their immediate surroundings. Example V2X-enabled functionalities are cooperative perception, traffic information, safety messages, and manoeuvre coordination.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-43: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development

X

Table 3-44: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-45: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Maybe	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Yes	
AD & Technology Sensor	Yes	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Maybe	
Driving Scenario Status	Yes	
Video	No	
Subjective Data	No	
Energy Consumption	No	

3.17 Develop End-to-End AI

Data is used to develop AI systems that can process information from sensors and directly turn it into driving actions, helping the vehicle learn how to respond appropriately in different traffic situations.

At least, the consortium partners as well as an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-46: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-47: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-48: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	No	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Yes	
AD & Technology Sensor	Yes	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Yes	
Driving Scenario Status	Yes	

Video	Yes
Subjective Data	No
Energy Consumption	No

3.18 Public Dataset AD Development

Data is used to create and improve (openly) available datasets that help researchers and developers build, test, and compare AD technologies under realistic traffic conditions.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-49: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-50: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-51: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Maybe	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	Maybe	
Passenger	No	
AD & Technology Enabler Status	Yes	

AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	Yes
Vehicle & System Status	Yes
Operation	No
Environmental Status	Yes
Driving Scenario Status	Yes
Video	Yes
Subjective Data	No
Energy Consumption	No

3.19 Project KPIs

Data is used to measure and track the progress of the project, helping to understand whether activities are running as planned and whether project goals are being achieved. Example key performance indicators for a project are driven distance, driven distance in AD, number of transported passengers.

At least, the consortium partners as well as an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-52: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-53: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-54: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	

Ego Vehicle Movement	Yes
Ego GNSS Information	Yes
Surrounding Objects Information	No
Abstract Map	Yes
Digital Map	No
Driver Input	No
Driver Status	No
Driver / Passenger Information	Yes
Freight information	Yes
Passenger	Yes
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	No
AD & Technology Sensor	No
Vehicle & System Status	Maybe
Operation	Yes
Environmental Status	Yes
Driving Scenario Status	Yes
Video	No
Subjective Data	No
Energy Consumption	Maybe

3.20 Operation Monitoring

Data is used to monitor the day-to-day operation of vehicles and transport services, helping to understand where vehicles are operating, how they are being used, and whether everything is functioning as expected.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-55: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-56: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-57: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Maybe	
Driver Input	No	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	Yes	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Yes	
AD & Technology Sensor	Maybe	
Vehicle & System Status	Yes	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	Maybe	
Video	No	
Subjective Data	No	
Energy Consumption	Maybe	

3.21 Driving Scenarios

Data for detection of driving scenarios (e.g. lane change, car following) including edge cases and other challenging scenarios. Data is mainly used to store them in a database to facilitate development and evaluation of AD.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-58: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-59: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-60: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Yes	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	No	
Freight information	Maybe	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Maybe	
AD & Technology Sensor	Maybe	
Vehicle & System Status	Yes	
Operation	No	
Environmental Status	Yes	
Driving Scenario Status	Yes	
Video	Yes	
Subjective Data	No	
Energy Consumption	No	

3.22 Safety – Crashes

Data is used to detect and record traffic crashes during the operation, helping to understand when crashes occur and to analyse them to improve safety. Data is mainly used to store them in a database to facilitate development and validation of AD.

At least, partners with a bilateral agreement shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-61: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-62: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-63: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Maybe	
Driver Input	Yes	
Driver Status	Yes	
Driver / Passenger Information	Yes	
Freight information	Yes	
Passenger	Maybe	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Maybe	

AD & Technology Sensor	Yes
Vehicle & System Status	Yes
Operation	No
Environmental Status	Yes
Driving Scenario Status	Yes
Video	Yes
Subjective Data	Maybe
Energy Consumption	No

3.23 Define Reference Driver

Data describe how human drivers drive on public roads. Data can be used to define reference driver for comparison in safety validation or homologation process (i.e. competent & careful driving) across Europe.

At least, an exclusive group of consortium partners (e.g. evaluation partners) shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-64: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-65: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-66: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Maybe	
Ego Vehicle Movement	Yes	
Ego GNSS Information	No	
Surrounding Objects Information	Yes	

Abstract Map	Yes
Digital Map	No
Driver Input	Yes
Driver Status	Yes
Driver / Passenger Information	Yes
Freight information	Yes
Passenger	Maybe
AD & Technology Enabler Status	No
AD & Technology Information	No
AD & Technology Communication	No
AD & Technology Sensor	No
Vehicle & System Status	Maybe
Operation	No
Environmental Status	Yes
Driving Scenario Status	Yes
Video	No
Subjective Data	No
Energy Consumption	No

3.24 System-of-Systems & Logistics Integration

Data used to determine the performance of the AD system within the wider transport and logistics ecosystem (e.g. mission success, fleet orchestration interactions, infrastructure dependencies), which is collected during operational testing to assess end-to-end system integration.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-67: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-68: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Yes	Yes

Table 3-69: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Maybe	
Ego Vehicle Movement	Maybe	
Ego GNSS Information	Maybe	
Surrounding Objects Information	No	
Abstract Map	No	
Digital Map	No	
Driver Input	No	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	Yes	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Yes	
AD & Technology Sensor	No	
Vehicle & System Status	No	
Operation	Yes	
Environmental Status	Maybe	
Driving Scenario Status	Yes	
Video	No	
Subjective Data	No	
Energy Consumption	Yes	

3.25 Reliability, Availability & Maintainability

Data used to determine the robustness and operational stability of the AD system over time (e.g. system uptime, degradation events, recovery times), which is collected during extended real-world operation to assess fleet readiness, fallback strategies, and feasibility of long-term deployment.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-70: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development

X

Table 3-71: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-72: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Maybe	
Abstract Map	Maybe	
Digital Map	No	
Driver Input	No	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	Maybe	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Maybe	
AD & Technology Communication	Yes	
AD & Technology Sensor	Maybe	
Vehicle & System Status	No	
Operation	Yes	
Environmental Status	Yes	
Driving Scenario Status	No	
Video	No	
Subjective Data	No	
Energy Consumption	Maybe	

3.26 Regulatory & Compliance Readiness

Data used to determine the degree to which the AD system complies with applicable regulations, permit conditions, and approval requirements (e.g. ODD compliance, incident

reporting, minimal risk manoeuvres), which is collected during real-world operational testing to support auditability and regulatory decision-making.

At least, partners with a bilateral agreement shall have access to data of this purpose.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-73: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
X		

Table 3-74: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-75: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Yes	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Yes	
Digital Map	Maybe	
Driver Input	No	
Driver Status	No	
Driver / Passenger Information	No	
Freight information	Yes	
Passenger	No	
AD & Technology Enabler Status	Yes	
AD & Technology Information	Yes	
AD & Technology Communication	Yes	
AD & Technology Sensor	No	

Vehicle & System Status	Yes
Operation	No
Environmental Status	Yes
Driving Scenario Status	Yes
Video	No
Subjective Data	No
Energy Consumption	No

3.27 Development of digital services

Data used to develop digital services such as predictive maintenance-as-a-service, smart insurance or congestion prediction.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-76: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
		x

Table 3-77: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
Maybe	Yes

Table 3-78: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment
External Statistics	Maybe	
Experiment Status	Yes	
Ego Vehicle Movement	Yes	
Ego GNSS Information	Yes	
Surrounding Objects Information	Yes	
Abstract Map	Maybe	
Digital Map	Maybe	
Driver Input	Yes	
Driver Status	No	

Driver / Passenger Information	Yes
Freight information	Maybe
Passenger	Yes
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	No
Vehicle & System Status	Yes
Operation	Yes
Environmental Status	Yes
Driving Scenario Status	Yes
Video	Maybe
Subjective Data	No
Energy Consumption	Yes

3.28 Control centre / remote operator

Data used to develop or apply remote operation support for automated vehicles. It includes potential triggers for the need of different kinds of remote operation (from monitoring and assistance to taking over control) as well as data required to design the interaction like communication with passengers, latencies, etc.

Details about the relevant topics, required data types and the mapping to data groups is given in the tables below.

Table 3-79: Mapping data purpose to topics.

Evaluation	Monitoring / Validation / Database	Design / Development
	X	X

Table 3-80: Required data types.

Requires Baseline Data (manual driving)	Requires AD Data
No	Yes

Table 3-81: Mapping data groups and data purpose.

Data Group	Relevant for data purpose?	Comment

External Statistics	Maybe
Experiment Status	No
Ego Vehicle Movement	Yes
Ego GNSS Information	Yes
Surrounding Objects Information	Yes
Abstract Map	Yes
Digital Map	Yes
Driver Input	No
Driver Status	No
Driver / Passenger Information	No
Freight information	Maybe
Passenger	Maybe
AD & Technology Enabler Status	Yes
AD & Technology Information	Yes
AD & Technology Communication	Yes
AD & Technology Sensor	No
Vehicle & System Status	Yes
Operation	Yes
Environmental Status	Yes
Driving Scenario Status	No
Video	Yes
Subjective Data	No
Energy Consumption	No

References

Alakkad, F., Khastgir, S., Kolbay, B., Garcia, S., et al. (2025). Methodology for Data Quality Assessment. Synergies Deliverable D3.1. Available at: SYNERGIES D3.1 (accessed 27.7.2026).

Bellotti, F., Berta, R., Capello, A., De Gloria, A., et al. (2022). Databases Structure. Hi-Drive Deliverable D5.2.

Connected Automated Driving Data Sharing (2026). Website: Data Sharing - Connected Automated Driving (accessed 27.7.2026).

EU (2016). Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). EU Regulation 2016/679. Available at: http://data.europa.eu/eli/reg/2016/679/oj (accessed 27.7.2026).

FAME (2024). CCAM Data Sharing Framework. Version 2. Available at: EUROPEAN COMMISSION (accessed 27.7.2026).

FESTA (2021). FESTA Handbook. Version 8 September 2021. Available at: FESTA-Handbook-Version-8.pdf (accessed 27.7.2026).

GIDAS (2026). German In-Depth Accident Study. Website: Home | GIDAS (accessed 11.07.2026).

Hi-Drive (2026). Project Website: Hi-Drive Deployment of Higher Automation (accessed 27.7.2026).

Hidalgo, J., Luján, C., Rodriguez, P., Karahasanović, J. (2025). Recommendations for European framework for testing on public roads. FAME Deliverable D5.3. Available at: FAME-DEL-D5.3-RecommendationsEUFrameworkTestingPublicRoads_v1.3.pdf (accessed 27.7.2026).

L3Pilot (2026). Project Website: L3Pilot: L3Pilot (accessed 27.7.2026).

MODI (2026). Project Website: MODI Project (accessed 27.7.2026).

ULTIMO (2026). Project Website: Home | ULTIMO Project (accessed 27.7.2026).

List of Abbreviations and Acronyms

Abbreviation / Acronym	Meaning
ACC	Adaptive Cruise Control
AD	Automated Driving
AVI	Audio Video Interleave
CCAM	Cooperative, Connected and Automated Mobility
CSA	Cooperate Support Action
CSV	Comma-Separated Values
DB	Database
DSF	Data Sharing Framework
EC	European Commission
ESC	Electronic Stability Control
EU	European Union
GDPR	General Data Protection Regulation
GIDAS	German In-Depth Accident Study
HDF	Hierarchical Data Format
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LiDAR	Light Detection and Ranging
MRM	Minimal Risk Manoevure
ODD	Operational Design Domain
ROS	Robot Operating System