Hi-Drive

Designing Automation

Deliverable D4.2 /

Data for Evaluation

Version: 1.0

Dissemination level: PU

Lead contractor: BMW

Due date: 31.07.2022

Version date: 11.10.2022

This project has received funding from the
European Union's Horizon 2020 research and innovation
programme under grant agreement No 101006664.

Hi-Drive

Document information

Authors

Felix Fahrenkrog – BMW Group

Satu Innamaa – VTT

Hendrik Weber – ika, RWTH Aachen

Esko Lehtonen – VTT

Teemu Itkonen – VTT

Anastasia Bolovinou -ICCS

Susanne Reithinger – BMW Group

Erik Svanberg - Chalmers

Lennart Vater – ika, RWTH Aachen

Marcel Sonntag – ika, RWTH Aachen

Johannes Hiller – ika, RWTH Aachen

Charlotte Fléchon – PTV Group

Jeroen Hogema – TNO

Sytze Kalisvaart – TNO

Saeed Rahmani – TU Delft

Yee Mun Lee – University of Leeds

Natasha Merat – University of Leeds

Henri Sintonen – VTT

Barbara Metz – WIVW

Andreas Landau – WIVW

Coordinator

Aria Etemad

Volkswagen AG

Berliner Ring 2

38440 Wolfsburg

Germany

Phone: +49-5361-9-13654

Email: aria.etemad@volkswagen.de

Project funding

Horizon 2020

DT-ART-06-2020 – Large-scale, cross-border demonstration of connected and highly automated driving functions for passenger cars

Contract number 101006664

www.Hi-Drive.eu

Legal Disclaimer

Table of contents

List of figures

List of tables

Executive summary

The Hi-Drive project is developing several automated driving functions (ADFs), technology enablers (technological tools—software, hardware and methodology—that have the potential to enable new ADFs and/or upgrade existing ones), and approaches in the user domain to overcome today's challenges in deploying higher levels of vehicle automation across European countries. These advancements should enable longer and more consistent automated drives in the future. Looking beyond the technical, the Hi-Drive project will also test and evaluate the developed technologies and the approaches related to them. Thus, data needs to be detailed, specified, recorded, and exchanged between the Hi-Drive partners. The work described in this deliverable forms the basis for this process.

The document gives a brief introduction to the Hi-Drive evaluation concept (section 2), which will be discussed in greater depth in future deliverables. This deliverable focuses on the data requirements covering data collection and storage concept in Hi-Drive (section 3). The introduction depicts:

- the overall data structure and content, which is closely linked to the evaluation process and builds upon the iterative process of the FESTA methodology from research question to data needs,
- the relevant stakeholders involved, and
- the core principles behind the storage and exchange of data in Hi-Drive.

To store the recorded data in an efficient way in the databases, five different data categories have been defined in Hi-Drive. These are:

- **Experiment metadata**: data intended to provide an overview of the Hi-Drive experiments (e.g., overall driven distance with AD in one experiment).
- **Questionnaire data**: data intended to collect information from subjective responses to questionnaires from human participants experiencing AD either as users in Hi-Drive experiments or as other road users.
- **Performance indicator data**: data for use in the technical assessment to answer the Hi-Drive research questions. This data will be derived from the information collected, processed, and stored during the Hi-Drive tests (e.g., average speed in car-following driving scenarios).
- **Time series data** and **aggregated time series data**: data intended to be used for the impact assessment and the Hi-Drive scenario and edge case database. The aggregated time series data is a modified form of the time series data, which will allow the Hi-Drive

partners to provide and access information that is less sensitive from an IP standpoint (e.g., velocity over time during a critical driving scenario).

These data categories will be uploaded in the different Hi-Drive databases:

- Consolidated database (CDB): The CDB will include three separate instances for *Users* evaluation (questionnaire data), *Effects* evaluation (performance indicator data) and "Experiment Metadata" (Experiment metadata).
- Driving scenario database (DSDB): The DSDB contains data characterising driving scenarios independently from research questions (time series data and aggregated time series data).
- Edge case database (ECDB): The ECDB is intended to contain driving scenarios that may have extreme parameters (time series data and aggregated time series data).

Next to the data categories, a list of signals to be logged by Hi-Drive partners in the experiments has been defined. For this purpose, first the required information in each data category was defined. Next, the required signals per data category were derived, and in the third step they were combined into a single signals list. The signals, which typically consist of information logged over time, are used to derive the required information for the five data categories. The signals list has been discussed with the relevant data providers.

In this context, it is important to underline that the list provides the complete set of signals required for the evaluations in *Users* and *Effects* evaluations. However, the partners are not obliged to store all the signals in their experiments, but rather to focus on the signals that are relevant to their technical development and experiment. Which signals need to be stored in an experiment depends on the scope of the experiment, the tested Hi-Drive enabler(s), the ADF, and the test environment. The work that is reported in this deliverable marks only the starting point. Other work packages will use the result of this deliverable and continue the work by preparing the required tools and databases, logging the data, and using the data in the evaluation.

1 Introduction

1.1 The Hi-Drive project

Connected and Automated Driving (CAD) has become a megatrend in the digitalisation of society and the economy. CAD has the potential to drastically change transportation and to create far reaching impacts. SAE level 3 (L3) automated functions were piloted in Europe by the L3Pilot project in 2017–2021 (L3Pilot consortium 2021). Hi-Drive builds on the L3Pilot results and advances the European state-of-the-art from SAE L3 'Conditional Automation' further up towards 'High Automation'. This is done by demonstrating in large-scale trials the robustness and reliability of CAD functions under demanding and error-prone conditions with special focus on:

- Connected and Automated Vehicles (CAV) travelling in challenging conditions covering variable weather and traffic scenarios and complex infrastructure,
- connected and secure automation providing vehicles/their operators with information beyond the line of sight and on-board sensor capabilities,
- complex interaction with other road users in normal traffic,
- factors influencing user preferences and reactions including comfort and trust—and eventually through a wide consumer acceptance of AD resulting in purchase and use, enabling viable business models for AD.

The project's ambition is to extend the CAD's operational design domain (ODD) from the present situation, which frequently demands taking over control of the vehicle by a human driver. As experienced in the EU flagship pilot project L3Pilot, on the way from A to B, a prototype AV will encounter a number of ODD factors, leading to fragmented availability of the AD function. Hi-Drive addresses these key challenges which are currently hindering the progress of vehicle automation. The concept builds on reaching a widespread and continuous ODD, where automation can operate for longer periods, and the interoperability is assured across borders and brands. Hi-Drive strives to extend the ODD and reduce the frequency of take-over requests (TOR) by selecting and implementing technology enablers leading to highly capable CAD functions, operating in diverse driving scenarios including, but not limited to, urban traffic and motorways. The removal of fragmentation in the ODD is expected to give rise to a gradual transition from conditional operation towards higher levels of automated driving (AD).

The work started in July 2021 with the collection and description of the different enabling technologies, ODDs, and automated driving functions. When testable functions and use cases were defined, research questions and hypotheses were formulated, leading to specification of data needed for recording and evaluation and recording of vehicle and driver behaviour.

Evaluation will focus on four areas: 1) users; 2) technical impact; 3) impact assessment (on safety, efficiency, environment, mobility, transport system, and society); and 4) societal impact. Furthermore, these assessments serve as input to determine whether the socio-economic benefits outweigh the costs. The project also engages in a broad dialogue with the stakeholders and general public to promote the Hi-Drive results. Dissemination and communication are boosted by a demonstration campaign to show project achievements.

Overall, Hi-Drive strives to create a deployment ecosystem by providing a platform for strategic collaboration. Accordingly, the work includes an EU-wide user education and driver training campaign and series of Codes of Practice (CoP) for the development of automated driving functions and Road-Testing Procedures, while also leading the outreach activities on standardisation, business innovation, extended networking with interested stakeholders, and coordinating parallel activities in Europe and overseas.

1.2 Overall implementation plan for Hi-Drive

The FESTA methodology was designed to be applied to field-operational tests (FOTs) with market-ready products (see Version 8 of the FESTA Handbook by FOT-Net, CARTRE & ARCADE (2021)). Therefore, it does not fully apply to studies with prototypical automated driving functions (ADFs). Thus, some adjustment of the FESTA implementation plan, described as the "FESTA-V" structure, was needed to accommodate testing of AD.

Figure 1.1 illustrates the FESTA implementation plan adapted for Hi-Drive. The plan is divided into three phases: (I) prepare, (II) operate, and (III) evaluate. In the beginning of the preparation phase (I), ADFs, the technology enablers, and their use cases and associated test scenarios across multiple test environments (test track, open road, simulation) are described in detail. Then, an initial list of research questions is set up and organised as high-, medium- and low-level questions. The state of the art is summarised for topics covered by these research questions. The feasibility of each research question is checked next in terms of data availability, suitability of the experimental design and procedures, availability of research tools, methods and external data sources, and availability of resources (e.g., project duration and human) required.

Next, the performance indicators and other data with which the research questions are answered, and the evaluation tools, are defined and calibrated. Based on these requirements for evaluation, five lists—one for different data categories[1] —with the required information are defined. In the following step, the five lists with the required information are merged into one signals list which specifies all the signals needed. Next to the signals list, a common data format (CDF) applicable to the project evaluation is specified for them. The data to be shared for evaluation is agreed with the data providers. Various databases and data tools are set for data processing and storage.

The experimental design and procedures are set to test highly automated driving and its technology enablers, and to provide data on them for evaluation. The plans for all operation sites are approved between the site owners and those setting the methodology for evaluation.

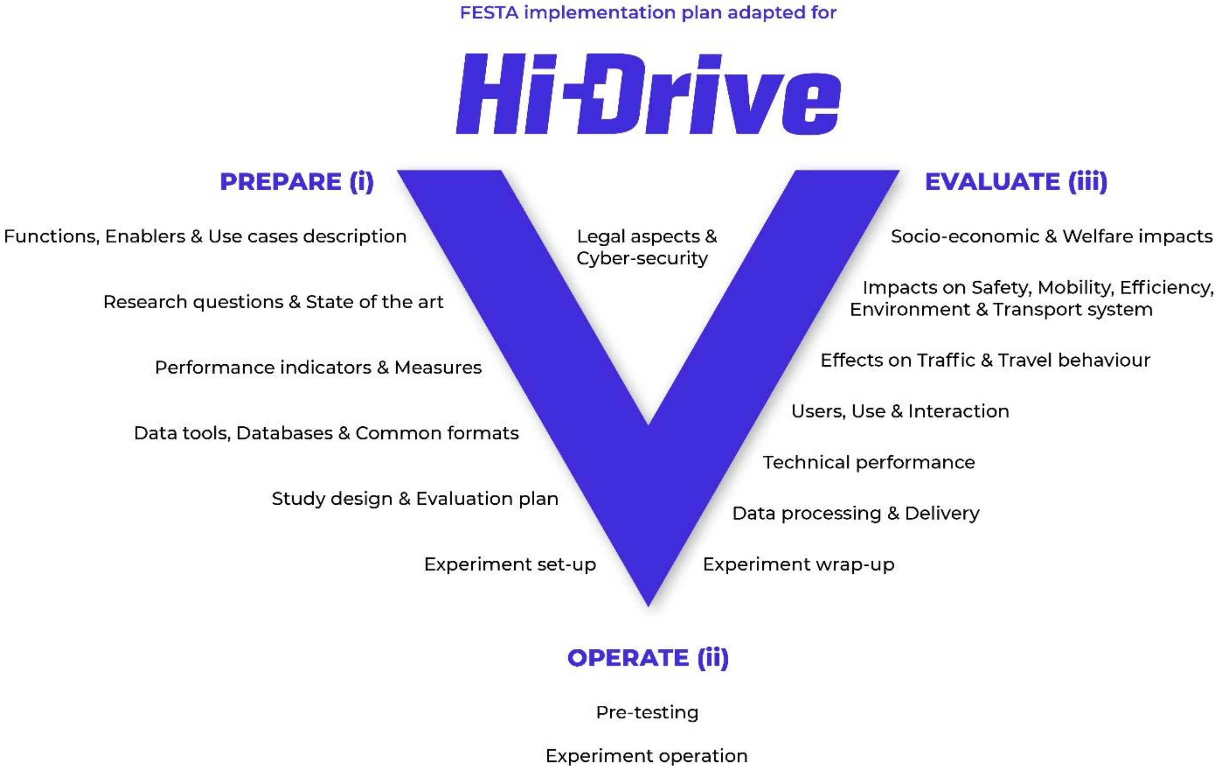

Figure 1.1: FESTA implementation plan adapted for Hi-Drive.

An evaluation plan is set for each research question to agree on who is responsible for what, to specify the methods, tools, and data to be used, scenarios to be addressed, and to plan the dependencies, i.e., linking the inputs and outputs as well as their timeline.

The operation phase (II) starts with the pre-testing step. It involves running all the phases of the project on a small scale to ensure that all the processes and tool chains function as

[1] The data categories are closely linked to the different databases which will become the tool for making the data available for evaluation.

intended. Once everything is confirmed to work as intended, the experiment operation begins. This phase involves the actual data collection.

The evaluation phase (III) starts with the data delivery as part of the experiment wrap-up. In this phase, it is also important to report all the deviations from the plan and any system updates made during the data collection phase. The data are converted to CDF, processed, and delivered to the evaluation team.

In the effect evaluation, technical performance of the tested technology is assessed. User evaluation focuses on the users, usage, and interaction. Effects on traffic and travel behaviour are assessed together with their societal impacts on safety, mobility, efficiency, and environment and later scaled up to European level. The final step is to assess the socio-economic and welfare impacts.

1.3 Activity objective, scope, and structure of the deliverable

The work within Hi-Drive is structured into subprojects. The objectives of the *Methodology* subproject (SP) 4 are to:

- Specify the Hi-Drive research questions for both Users and Effects evaluation, how they will be addressed, and the related data needs.
- Agree on CDF for provision of different datasets.
- Agree on experimental design and procedures for testing and evaluation of automated driving functions and related enablers in challenging environments.
- Reconsider the theoretical background and impact mechanisms to build a multidisciplinary evaluation methodology, covering not only the expected positive impacts on safety, comfort, and the environment but also the unintended and possibly negative impacts on users and the transport system.
- Refine the state-of-the-art methods to address user and human-factor aspects of high-level driving automation, and facilitate understanding of possible effects on the transport system level, addressing travel behaviour, safety, efficiency and emissions.
- Provide Lessons Learned from the methodology point of view.

This deliverable reports on the activities of work package (WP) 4.4 *Data requirements*. The purpose of the WP is to ensure that the required data for the evaluation and the project's databases are logged and provided by the Hi-Drive experiments. The data will be used for multiple purposes in the project, for instance to support the development of technology enablers and to collect scenarios for Hi-Drive's scenario database. But the most important purpose is to ensure that the data are suitable for the evaluation. Therefore, this WP belongs to SP4 *Methodology*, which will pave the way for the evaluation by preparing the definition of the Hi-Drive research questions, experimental design, and evaluation methodology.

This report is structured as follows: First is an introduction to the evaluation approach (chapter 2), followed by an overview of the data collection and database concept of Hi-Drive (chapter 3). Chapter 4 describes in greater detail the different data categories that will later be stored in the Hi-Drive databases. A further important feature due to the variety of experiments in Hi-Drive is the storage of context data. Here, two kinds of information—experiment metadata & contextual data—are of importance. The relevant tables are provided in Annex 1 and Annex 2. Chapter 5 reports on the overall Hi-Drive signals list. It must be noted that not all signals need to be stored by all partners. The final chapter 6 draws the conclusions of the report and provides an outlook on the next steps.

1.4 Objectives of the work

This report provides information for the Hi-Drive project’s internal stakeholders (i.e., partners providing data to the different databases) and external stakeholders regarding which data are stored within Hi-Drive and in what way. This deliverable therefore reports on the data, and their structure, that will be used for analysis in the Hi-Drive project.

The project’s analyses are conducted as part of the Hi-Drive subprojects SP6 *Users*, SP7 *Effects*, and SP8 *Outreach*. Subproject SP2 *Enabler* will further report on the assessment of the developed enablers. As subprojects SP2 *Enabler*, SP5 *Operation*, and SP6 *User* will mainly provide the data needed for the analysis, WP4.4 has discussed the needs and required signals primarily with those subprojects. This report describes the outcome of those discussions.

It should be noted that the report reflects the project’s status as of August 2022. In a project that runs for over 4 years, changes might occur that have implications on the data logging, data concepts, and individual signals.

2 Hi-Drive Evaluation and data concept

2.1 General information

The Hi-Drive project is investigating ADFs in multiple directions and aims to extend the ADF's ODD to allow for longer and more frequent usage of the functions. Within the project, the aim of SP4 *Methodology* is to develop and define the overall methodology for the evaluation at the end of project. This methodology will serve as a guideline that helps to focus the research done within the various subprojects on the overall project goals. Furthermore, applying a harmonised and scientifically sound methodology will guarantee that the work conducted within the project leads to reliable and valid knowledge. The methodology developed by SP4 will be applied by the evaluation subprojects (*Effects* and *User*) to answer the project's research questions.

Within the project, a variety of different experiments will be conducted. These experiments differ in both their type and scope. For a comprehensive evaluation, it is necessary to combine the findings of the different experiments. The basis for the evaluation is the data collected in the experiments. Here, a distinction must be made between the evaluation scope (technical, user, and impact assessment) and the test aggregation level (see Figure 2.1)

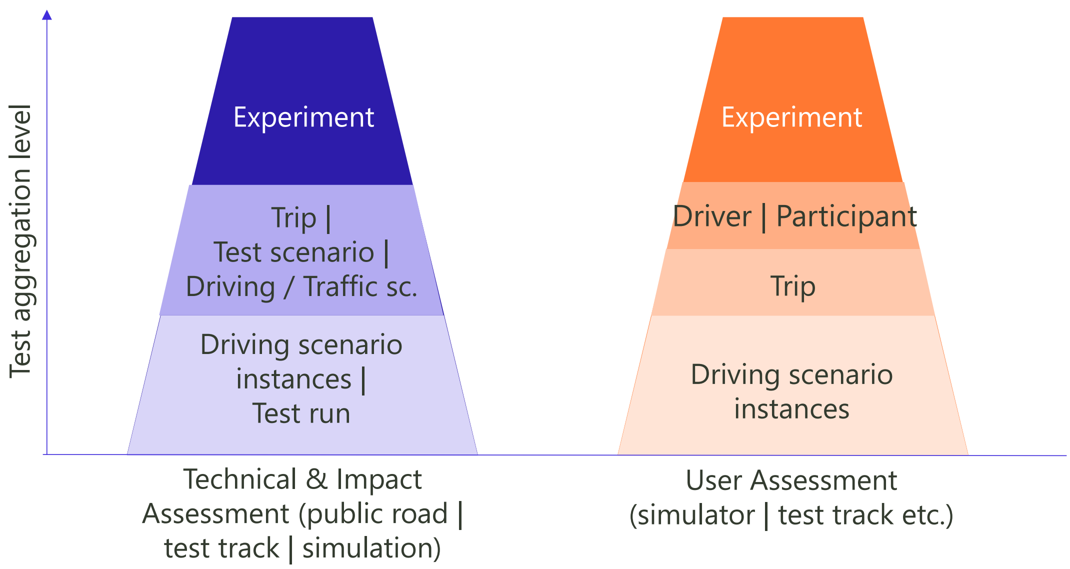

Figure 2.1: Evaluation scope per test environment and the reported test level.

The definitions used are given below:

- **Experiment:** An experiment consists of a series of test runs/trips to investigate a common aspect (ADF, enabler, user). It is made up of several test runs/trips that are conducted under comparable circumstances (e.g., same test route). Experiment types include open road, test track, driving simulator, simulation model, etc.

- **Trip:** Includes the sequence from the vehicle ignition key being turned on until it is turned off (even if the vehicle is not moving during this time frame). Particularly relevant in case of test on public roads.
- **Test scenario:** Description of sequence of triggers, events, and actions among use case entities (ego vehicle, other traffic participants, etc.) in order to reach a use case goal. Test scenarios are often related to the test track test.
- **Driving scenario:** Driving scenarios describe the development of a situation within a traffic context in which at least one actor performs a (pre-)defined action and/or the driving scenario is triggered by a (predefined) event. The action or event is specified without the definition of concrete parameters. The influenced actor may be either the ego vehicle (e.g., performing a lane change or a minimum risk manoeuvre) or another traffic participant (e.g., a lane change in front of the ego vehicle). The event triggering the driving scenario can be a change in road infrastructure (e.g., an end of lane or a change in speed limit) or an external obstruction (e.g., an obstacle on the road). An example is a lane change.
- **Driving scenario instances:** A driving situation represents a single segment in time that is assigned to a certain kind of driving.
- **Traffic scenario:** Traffic scenarios describe a larger traffic context by covering a longer period of time and longer road sections with certain traffic characteristics. One traffic scenario may include different (not predefined) driving scenarios. An example: a 3-lane motorway section of length 10 km with 2 motorway entrances and exits, a speed limit of 130 km/h, traffic volume of 4 000 vehicles/h/direction, 10% of heavy vehicles and a time period of 1 hour.

2.2 Evaluation in Hi-Drive

There are two main pillars of evaluation activity within Hi-Drive: 1) *Users* and 2) *Effects*. The *Users* evaluation focuses on acceptance, usage, user experience, and interaction with externals. *Effects* evaluation addresses how technology enablers enhance the AD performance and contribute to defragmentation of the ODD, how traffic and travel behaviour are affected by high-level automation, and what the societal impacts are of these changes.

Both evaluation pillars are linked, since user-related data also provide input for socio-economic and impact assessment, especially mobility impacts. Effect evaluation aims to consider the effects identified on user behaviour and acceptance when determining the impact of automated driving.

Evaluation activities within Hi-Drive are based on five experiment types (Figure 2.2). These include on-road studies in which the ADF and/or technical enabler is tested on public roads. More safety-critical tests and development related tests will be performed in closed/controlled environments, typically on test tracks. These tests play an important role for technical enablers. User focused studies will involve different test environments, and alongside real test environments such as public roads and test tracks, driving simulator studies and surveys will have an important part. Finally, simulations will be used for impact assessments and enabler tests.

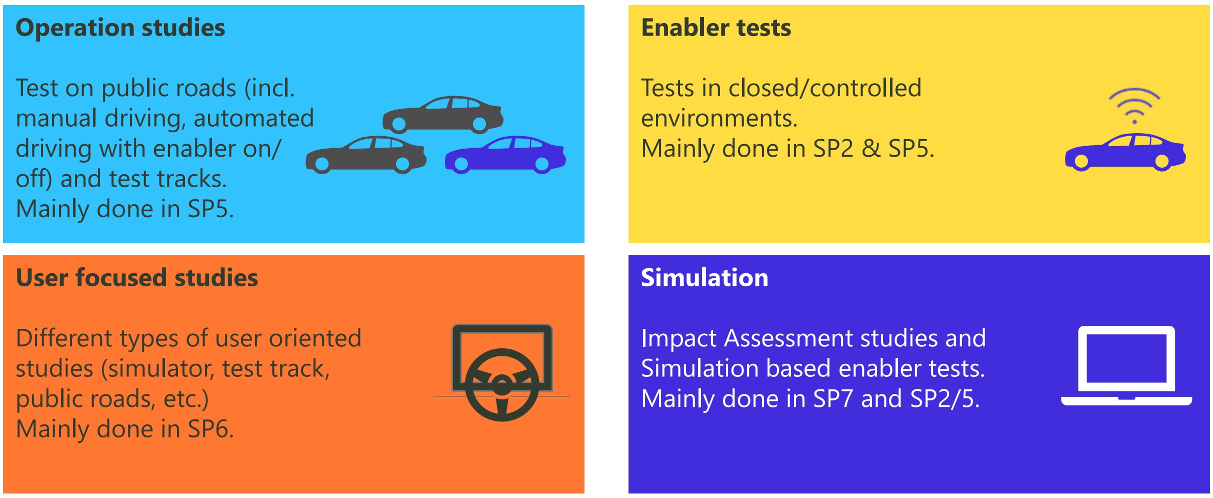

Figure 2.2: Hi-Drive Experiment types.

More details on the different evaluation areas are provided in Section 2.2.2.

2.2.1 Research questions

The work of SP4 *Methodology* started, according to FESTA-V (see Figure 1.1), with the definition of the research questions. A first version of research questions became available in December 2021. The list of questions is based on the description of the expected work and an extensive state-of-the-art review. Based on these initial activities, altogether six areas of research questions were identified that need to be addressed within Hi-Drive to achieve the overall goals of the project. Two areas deal with the technical evaluation of the new technological approaches (enablers) and their impact on AD performance and on driving. Three areas focus on user related topics. The last area deals with estimating the future impacts of AD on our lives and on society.

Within each area, high, medium, and low-level research questions are defined that are again linked to specific performance indicators. The first draft version of research questions served as a direct input to the work described within this deliverable. Based on the research questions, performance indicators had been defined from which the data needs were derived. The process has an iterative character due to the complexity of the addressed areas. Thus, the performance indicator list is preliminary. To ensure that the final list of research questions does not describe a wish list but rather something feasible within the scope and reach of Hi-Drive, the final research questions and their performance indicators will be checked again against the data requirements defined in this deliverable. This last feasibility check is a necessary step to ensure that the evaluation work can be executed as planned.

2.2.2 Hi-Drive Evaluation

Within SP7 *Effects*, data collected within the various operations in Hi-Drive are analysed to answer research questions pertaining to the technical effects of enablers and their wider impacts.

The **technical assessment** focuses on the performance evaluation of Hi-Drive ADFs in combination with enablers. To assess research questions, related hypotheses are defined. To answer them, performance indicators (PIs) are defined, which fall under two main research areas: namely AD performance (including questions on safe, efficient, and comfortable driving behaviour) and AD availability (including questions on ODD extension and defragmentation). The PIs will be derived based on logged data and statistical analyses. A preliminary list of PIs is shown in Table 4.1. What makes the Hi-Drive dataset unique and interesting is that it will contain a big pool of diverse driving scenarios, PIs, and experiment-related metadata based on the input from pilot drives on both public roads and controlled test tracks as presented in Figure 3.2.

The **user assessment** analyses users' behaviour and experience of CAVs. The assessment covers both drivers inside the vehicle and external road users, including vulnerable road users, drivers of other vehicles, and teleoperators. For this purpose, research questions will be investigated using methods such as driving and pedestrian simulators, head-mounted displays, test tracks tests, real-world observations, questionnaires, and interviews. User assessment will be performed within SP6 *Users*, and includes understanding user acceptance and awareness, human-like driving and user comfort, user monitoring and related HMI, and interactions with other road users.

The **impact assessment** evaluates the effect of a technology on traffic. In Hi-Drive, the scope will be on assessing ADFs in combination with the developed technical enablers. The impact assessment investigates five different areas, namely safety, efficiency, environment, mobility, and transport system. The impact assessment relies mainly on simulation conducted for the different areas. Later, the simulation results are scaled up to determine the impact on European level. Both steps require several input data. These input data are provided by external sources (e.g., accident databases) as well as by the technical assessment of Hi-Drive experiments. The Hi-Drive experiments are used to parametrise the simulation of the impact assessment and to validate the simulation outcome.

2.2.3 Databases

Various databases are used in the evaluation process as an access point to the researchers working on answering the different research questions set by the project. Data may either be used to answer certain research questions directly by means of statistical analysis or serve as input to other activities within the evaluations. In principle, three different common databases were defined in the project proposal phase (cf. Table 2.1). However, for the three databases, multiple instances may exist. Furthermore, each database may have different access rights, such that certain users can access different data elements.

Table 2.1: Hi-Drive databases.

Database	Purpose
Consolidated Database (CDB) for • *User* Evaluation • *Effects* Evaluation • Experiment Metadata	The main purpose of the consolidated database is the answering of research questions. Based on statistical methods, data contained in the database can be used to answer research questions or to confirm certain assumptions used in other assessments (e.g., supporting assumptions about the automated vehicles' desired headway distances for the impact assessment). Data in the CDB is typically represented in tabular format. Three different instances of CDBs are foreseen. The CDB for the User evaluation concerns entries per user. The CDB for the Effects evaluation concerns data per trip, test scenario, or driving scenario instance. The experiment metadata database has a special role since this database concerns the management of the testing process, not the actual vehicular/user data ("project dashboard"). The purpose of this database is to provide an overview of the status of the project and the progress of individual partners. Therefore, in contrast to the other databases, the contributing partners will be explicitly named next to the provided information. For the other databases, the information on who contributed which data will not be stored.
Driving Scenario Database (DSDB)	The DSDB contains data characterising driving scenarios. The DSDB enables evaluation partners to investigate what certain driving scenarios recorded during drives and then extracted from the data look like. Use cases of the DSDB could be selecting driving scenarios for simulation and selecting data for parameterising of simulation models. DSDB data are independent of the research questions.

Database	Purpose
Edge Case Database (ECDB)	The ECDB is intended to contain driving scenarios that may have extreme operating parameters or unusual combinations of operating parameters as well as scenarios that may not match the catalogue of predefined scenarios. The data contained in this database can be viewed as a dataset containing only unusual instances of driving scenarios.

While the CDB will store single values, the DSDB and ECDB will store multiple data points over time in different driving scenarios; both databases can be used for the SP7 *Effects* analysis. Their structure might be similar, but the scope will be different. The DSDB will focus on driving scenarios in a broader sense, while the ECDB will only focus on edge cases that contain extreme occurrences of the defined driving scenarios. Furthermore, the edge case scenarios may contain extreme events not considered by the defined driving scenarios.

2.3 Relation to previous projects

The Hi-Drive methodology is based on the L3Pilot methodology, adapting it to the Hi-Drive scope and extending it to evaluation areas not covered by L3Pilot. Enhancements are also made to overcome certain limitations in the L3Pilot evaluation.

Among the main differences between L3Pilot and Hi-Drive is the complexity of the experiments. L3Pilot involved large-scale but relatively homogeneous experiments (Pilots), which ensured sufficient comparable data for evaluation and made harmonisation relatively easy. Hi-Drive, on the other hand, will have heterogeneous experiments (particularly in terms of number and type of ADFs, enablers, users), which is a challenge with respect to data content, quantity, and harmonisation.

In L3Pilot, data was provided for specified site-specific analysis partners in a common data format (CDF), which was specified based on the data needs of the project’s research questions. L3Pilot extracted driving scenarios from the data and derived PIs for them. For some purposes like for instance the safety impact assessment, these PIs were found not to be sufficient. Therefore, Hi-Drive has specified more data requirements to serve different data use purposes, such as the scenario and edge case database.

An additional important input is the data sharing framework v1.1 from the FOT-Net Data and CARTRE projects (2019).

3 Data collection and database concept in Hi-Drive

3.1 Hi-Drive data collection process

The comprehensive approach of Hi-Drive in terms of assessing ADFs and enablers poses a challenge for the evaluation, since a huge quantity and variety of experiments and data need to be processed and combined. Therefore, a process for data collection and handling as well as common principles are required. The process of data collection in Hi-Drive[2] is illustrated in Figure 3.1. Typically, the top process steps are done by the *Operations* partners while the bottom process steps involve the evaluation partners of *Users* and *Effects*.

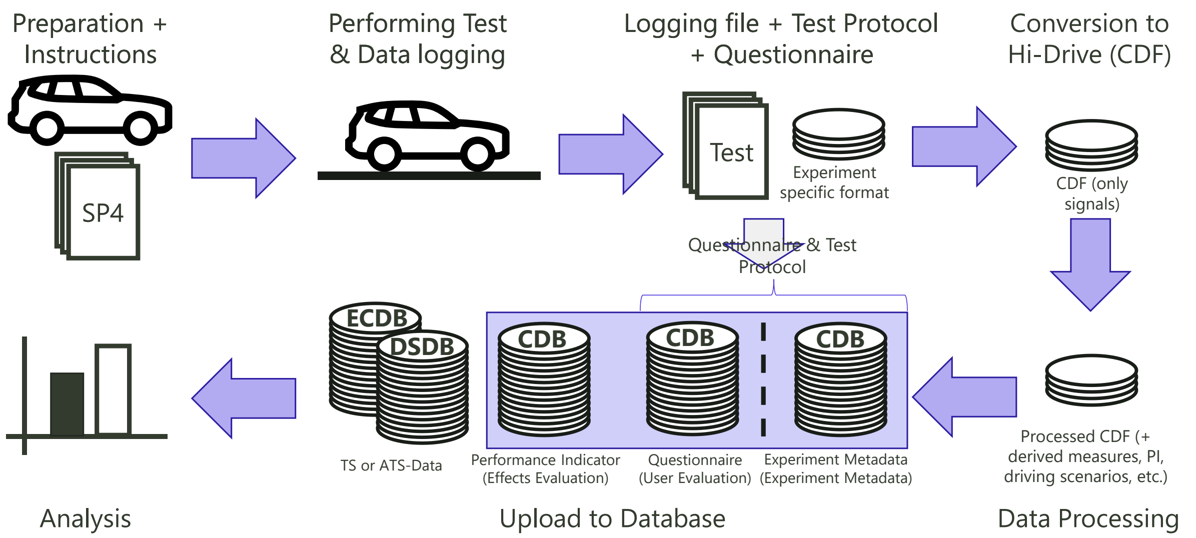

Figure 3.1: Process from data requirements to the evaluation results.

The process starts with the preparation of demonstrator vehicles by the vehicle owners and of instructions on the experiments. Both aspects also concern the issue of data. While the experiment instructions are being developed by SP4—including this deliverable on data requirements—the vehicle preparation is being done within SP3 and must ensure that the information requested in this deliverable can be provided using the data loggers in the vehicles. The next phase concerns the execution of the experiments combined with the logging data (vehicle data, test protocol, questionnaire). This data is typically stored in the formats that are defined by the experiments' operators.

[2] In the following, the focus is on data collection for vehicle tests. Similar data will be collected in *Users* experiments (e.g., driving simulators studies). The process is similar overall, but in most cases the entire process from (simulator) experiment preparation to analysis will be conducted by one partner.

To enable joint evaluation, the data need to be converted into a single data format, for processing feasibility. This is achieved by the next step, which is done either by the experiment operator or by the analysis partner. The format used here builds upon the CDF developed as part of L3Pilot (Koskinen et al. 2021 and Hiller 2021). Once the data is in the CDF, the data can be processed with the scripts and tools developed within Hi-Drive. Part of this step concerns the calculation of derived measures (signals that could not be logged directly but need to be calculated), PIs, and driving scenarios. For the questionnaire data and test protocol, these steps may be minor. Therefore, for these types of data, the figure displays a direct link to the related database instances (experiment metadata and questionnaire), skipping the vehicular data processing steps.

Once the data is processed, it is uploaded to the Hi-Drive databases, which are defined by the different data categories collected during Hi-Drive (see 3.4.1). The evaluation subprojects SP6 *Users* and SP7 *Effects* will use the databases to conduct their work and assess the ADFs in combination with the developed enablers over the different experiments. This whole chain for data processes—logging, conversion to the single data format, calculation of derived measures, PIs and driving scenarios, upload to the databases, and use for evaluation—should be verified as part of the pre-testing to ensure that everything works as intended.

3.2 Principles for data storage and exchange

The experience of previous projects—starting with euroFOT (Kessler et al. 2012) up to L3Pilot (Bellotti et al. 2020, Bellotti et al. 2020b)—has shown that data sharing is always a sensitive topic and relies on trust between the involved partners. It is also clear that legal requirements, such as the EC General Data Protection Regulation (GDPR, European Parliament 2016), must be followed by the partners and the project. For this reason, the legal requirements and principles of data storage and exchange within Hi-Drive have been defined as follows:

- All partners who are conducting experiments contribute with their data to the evaluation within Hi-Drive.
- Maximise the use of collected data / utilisation of an experiment for data collection. This means always check what else the data could be used for. For example, an experiment within SP6 *Users* could also contribute to the edge case database with relevant driving scenarios in SP7 *Effects*. This requirement derives directly from the aim to optimise to use of the different Hi-Drive experiments.
- In contrast to previous projects that had a rather homogeneous type of experiment (e.g., L3Pilot experiments reported by Andreone et al., 2021), here it is necessary to share more

information about the experiments to achieve good understanding among the evaluation partners of the tests conducted within the project. This understanding is important to ascertain which data sources can be combined for meaningful evaluation. Therefore, information on the experiment, such as metadata and, if possible, pictures (respecting GDPR), shall be provided.

- The logged and derived data shall be provided in CDF to enable the usage of common evaluation tools.
- The partner should deliver data as early as possible. In the case of long-term experiments, data should be delivered continuously (e.g., every 2 months). This requirement derives from the experience in previous projects that the data tool chain needs to be checked and validated with real data. The earlier the data is available, the earlier these checks can be done.
- The defined signals and data requirements focus primarily on the needs of this project. The data may be possibly used for later purposes as well, provided that the consortium agreement is respected. However, this is not within the scope of WP 4.4 *Data requirements*. Since including external requirements would make the complexity of requirements overwhelming and the task over-constrained, we do not attempt to satisfy future interest.
- The data defined is used to assess ADFs in combination with enablers in general. The goal is to provide the public with a general picture of automated driving and to improve the continuity of ODDs. Hi-Drive does not intend to benchmark between ADF implementations, and the CDF is set up to prevent such benchmarking.
- The collected data should not enable reengineering of the developed ADF and enablers. Therefore, the way the data has been defined is to prevent reengineering. In addition, it is also common sense among partners that Hi-Drive data should not be misused for such a purpose.

3.3 Data stakeholders

The logged data will be later used in several tasks of the project. Therefore, an overview of the relevant stakeholders in the work of WP4.4 *Data requirements* and their scope of interest is given below.

- SP4 *Methodology*: The data requirements are part of the overall evaluation methodology. Since the evaluation to be carried out in SP6 *Users* and SP7 *Effects* relies on the data collected during the project, this WP is an essential part of the evaluation methodology. On the other hand, WP4.4 needs input from the other work packages of SP4 (namely

WP4.6 *Method for user evaluation* and WP4.7 *Method for effects evaluation*) to define the data needs.

- SP6 *Users*: User assessment is one of the subprojects that carries out evaluation activities on a larger scale. Next to the studies carried out within SP6, the subproject wishes to use the data collected in the Operations subproject (SP5). One important data category for the evaluation in SP6 will be questionnaire data.
- SP7 *Effects* – Technical assessment: The technical assessment is part of the *Effects* subproject that focuses on its evaluation of ADFs in combination with technical enablers. To assess the ADF research question, related hypotheses are defined. To answer them, PIs are derived and compared. Thus, this subproject also requires for its work the data collected within the Operations subproject.
- SP7 *Effects* – Impact assessment (Safety, Efficiency, Environmental, Mobility, Social Economic) (SP7): Another assessment within the *Effects* subproject is the impact assessment that analyses the effect of automated driving with enablers on European road traffic. For the impact assessment, signals over time that are collected in different driving scenarios are important. Those signals are used as input for the conducted simulations. These driving scenarios can be collected in different experiment environments (driving simulator, simulation, test track or operation on public roads).
- SP7 *Effects* – Scenario/Edge Case Analysis: The scenario/edge-case analysis is also part of the *Effects* subproject. For this analysis, similar data are needed as for the impact assessment, but the evaluation scope is different.
- Progress of the project (Project Management, SP2, SP5, SP6): There is interest among project management, as well as within the subprojects conducting experiments, to gain an overview of the partners' progress. Thus, information should be provided on which partner has conducted which type of experiment and what the status is.
- Enabler development & assessment (SP2): The Enabler subproject focusing on the development of technical enablers within Hi-Drive has an interest in using data for the development and a more detailed assessment of the enablers. The interest in data is as for the impact assessment signals over time. However, for the analysis, also PIs will be calculated. Input for the work in SP2 will come both from experiments conducted within SP2 and from the Operations subproject.
- General Public, Researchers & the European Commission (EC): Besides internal interest, there is also external interest in the Hi-Drive data. While the primary interest of the EC can be expected to be monitoring the progress of the project, the public is also interested in

the overall results. For public research, the edge case / scenario data might be of relevance.

3.4 From data collection to data storage

The general concept on data collection and storage for Hi-Drive is presented here. Since many partners of Hi-Drive have been involved in the L3Pilot project before, in which huge amounts of data were shared among the partners, it was decided to start from the L3Pilot concept. The concept has been developed further to fit the needs of Hi-Drive while considering the lessons learned from L3Pilot. In the following sections, the different data categories, the aspect of meta-information, and the data flow are described.

3.4.1 Introduction to the Hi-Drive data categories

For the storage of Hi-Drive information collected from the Hi-Drive experiments, five data categories have been defined. The categories reflect the needs of the stakeholders as described in Chapter 3.3 and are outlined in Figure 3.2.

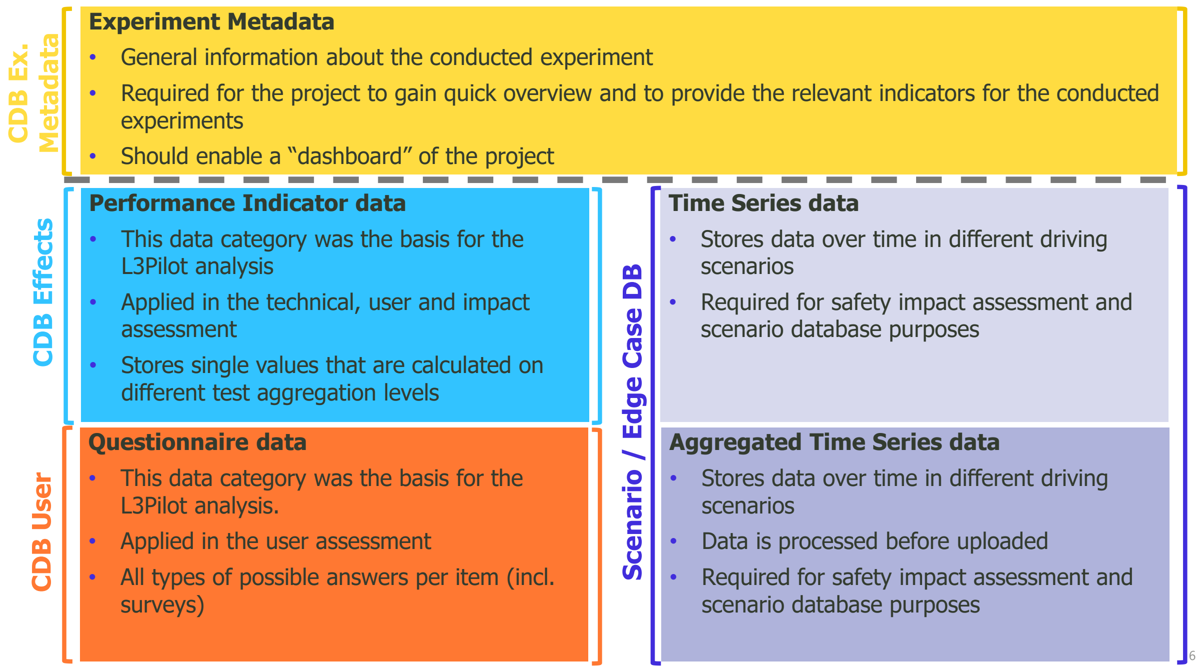

Figure 3.2: Hi-Drive data categories and their purpose.

The figure also shows their application within the Hi-Drive project. Back in L3Pilot, questionnaire data and PIs were already stored in common databases. This approach is continued in Hi-Drive, even if data structures, the research questions and types of experiments differ. Moreover, Hi-Drive adds three additional data categories: time series, aggregated time series, and experiment metadata. Time series data were logged during

L3Pilot but not stored in a common project database. The aggregated time series create an additional option to provide time series in a more anonymised way. The experiment metadata addresses the need for an overview of the conducted experiments. All Hi-Drive data categories are presented in detail in Chapter 4.

3.4.2 Meta-information for the evaluation

The evaluation within Hi-Drive deals with the challenge that several different ADFs and enablers need to be assessed. In addition, the tests for ADFs and enablers are conducted in different test environments. Therefore, it is important to have meta-information that shows what and under which conditions certain experiments have been conducted. This information was missing in the L3Pilot project and posed a challenge already back then, even though the experiment setting was quite similar between different test sites.

Therefore, it has been decided to include meta-information for the Hi-Drive project to enable the partners involved in the evaluation to develop a better understanding of the data. In this context it needs to be considered that the data in the database normally contains no information on which demonstrator vehicle has provided the data. Considering further the additional requirement to gain a project overview, it has been decided to include two types of meta-information. These are the "experiment metadata" and the "contextual data".

- **Experiment Metadata:** This data category aims to provide an overview of the status of the experiments conducted within Hi-Drive. The data will be provided at experiment level and once an experiment is finished. For longer experiments the data should be provided in regular time frames. The data should allow the project to run a dashboard of conducted experiments. In contrast to all the other databases in the project, the partners will be explicitly named in this database.
- **Contextual Data:** The second meta-information type is contextual data. The purpose of contextual data is to provide information about the conducted experiment together with the different data categories. In contrast to the experiment metadata, this information is not an own data category. It is additional information included in the other data categories (questionnaire, PI, time-series and aggregated time-series data). The contextual data should inform the evaluation under which conditions the data has been gathered for the analysis.

3.4.3 Data flow from signals to databases

The data is logged during the Hi-Drive test drives, independent of whether the test drive is a trip on public roads, a test on a test track, or a simulation run. The process from individual signals to the information stored in the Hi-Drive databases is visualised in Figure 3.3.

As stated earlier, alongside the project-relevant data, also dedicated information specific to one experiment will be stored, if needed. The processing and storage of the data for this case is rather up to the individual partner, and therefore not further discussed here.

The project-relevant signals defined in the Hi-Drive signals list (see Chapter 5) are first processed then transferred to the databases. The processing could differ per data category and reported test level. For instance, the process A will calculate a performance indicator trip level, while processing B will generate the time series data in one driving scenario. The processing for the PIs will produce a series of single values, while the processing of the time-series data (combining aggregated and original time-series data) will output values over time. Contextual data will be transferred to the database together with this information. The contextual data will be derived from the vehicle signal and the test protocol. These data are reported depending on the analysis scope on trip level (public road test), test run level (test track test), driving scenario level (public road, driving simulator or simulation test), or driver level (driving simulator test).

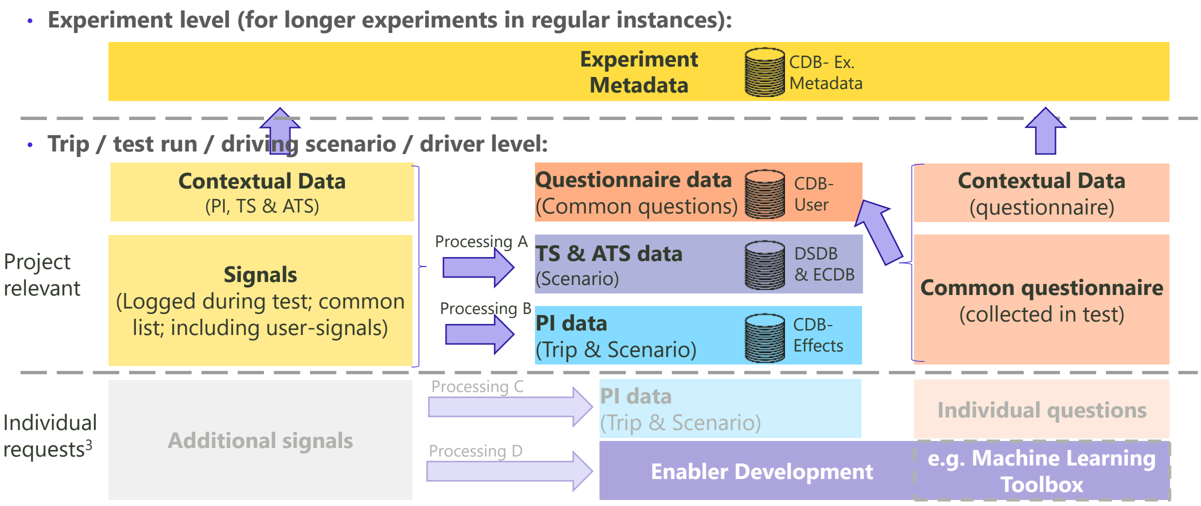

Figure 3.3: Information to be logged and provided per Hi-Drive data category to the related databases.

Questionnaire data follows a similar processing, ending up in the questionnaire database. The report is expected to happen only on trip and driver level.

In the end, the contextual data of both input types—signals and questionnaire—are used to provide the information for the experiment metadata database. Here, the data is only reported at experiment level.

3.5 Statement regarding bilateral exchange of data in the project

This report focuses on the data required to analyse the ADFs developed within Hi-Drive combined with their developed enablers. Thus, it describes the common grounds for evaluation.

However, it should be noted that certain Hi-Drive partners might need additional data if they wish to dive deeper into specific aspects or require it for development of the ADF. Logging this data is obviously possible, but these signals/data are not covered by this report. It is also possible that the partners share the data with other partners on a bilateral or multilateral basis. However, this data will not be transferred to the project-level Hi-Drive databases.

4 Hi-Drive data categories

4.1 Contextual data

The five Hi-Drive data categories were introduced in Section 3.4.1. This chapter describes the data categories in greater detail. A general description is also given for each category, including its scope and the information provided.

The purpose of contextual data is to provide information about the conducted experiment. Contextual data is not a data category *per se* but is provided for all data categories (except for experimental metadata) as additional input for the analysis. Particularly, it is relevant for the following data categories: questionnaire, PI, time series, and aggregated time series data.

One example of contextual data would be the experiment environment (test track, public road, or simulator study). Without it, an evaluation analyst would not know whether the average speed (one example of a PI) of two experiments can be compared or merged. Contextual data should thus enable the evaluation partners of SP6 & SP7 to decide which data can be combined and which needs to be excluded from the analysis if it does not match the other data. Contextual data should also enable further analysis, like investigating the effect of an ADF or enabler for different driver types or in different regions.

Depending on the data category, contextual data is provided on either (see also Figure 2.1):

- driving scenario level (technical PI, aggregated time series or time series data),
- test run level (technical PI, aggregated time series or time series data),
- driver level (questionnaire or user PI data), or
- trip level (technical & user PI data)

Information provided in the contextual data concerns, for instance:

- the country or region in which the experiment took place—to check for regional effects,
- the purpose of the experiment—to understand the primary focus of the experiments,
- the road type on which the test took place—to identify the ADF type,
- the test environment—to group similar experiment types,
- the participant type—to distinguish between different driver types,
- information about the status of systems (ADF, enabler, ADAS)—to identify baseline and treatment tests,

- weather, lighting & road conditions—to investigate the performance of the ADF and enablers in different environmental conditions,
- age & gender of the test participant (only for naïve, non-professional drivers)—to investigate the effect on different user groups, and
- different information related to execution of the experiment—to group similar experiments and check for the context of the results.

The list of all stored information of the contextual data is provided in Annex 2.

4.2 Experiment metadata

4.2.1 General information

The main purpose of this data category is to enable a quick overview of all conducted Hi-Drive experiments and tests. The data category should allow to analyse quickly and easily what the project partners have tested (e.g., how many kilometres have been travelled, how many persons took part in the tests). Thus, this category should provide all the information required for a "dashboard of the project". The "experiment metadata" category should be reported for all the experiments (test run, survey, laboratory test, simulation etc.) conducted within Hi-Drive (namely SP2, SP5, SP6, and SP7). The data should be provided independently of whether the ADF and/or enabler is on or off.

The relevant stakeholder for this data category is the project management, including the leaders of the subprojects *Enabler*, *Operations*, and *Users*. This data category has high relevance for WP5.4 *Operations*, which coordinates the experiments in the subproject Operations. These are also the partners to whom the data should be available, at least. Who else should have access to it is under discussion.

Experiment metadata should be uploaded together with the other data categories to the database. For the user studies this means that this data category should be uploaded once an experiment is finished. For experiments that run over longer time periods (e.g., on public roads), data should be uploaded in regular time frames (e.g., every 2 months) and after finishing an experiment. The experiment metadata should also contain information about conducted pre-tests, if feasible.

This data category will be reported on experiment level. It should combine the individual tests (test runs on a test track or trips in pilot studies) of one experiment. Most information in this data category requires only the storage of a single value. Only a few items require the storage of a histogram (e.g., of driven velocities). The biggest challenge is to get the data quickly and easily into the database. Here it should be noted that the experiments covered are quite different from each other and that automated calculation might not be feasible in all cases.

Contextual data and experimental metadata contain similar information (see sections 4.1 and 4.2.2, or Annex 1 and Annex 2). These overlaps are required because contextual data are not provided for experimental metadata, and the data are stored differently in different databases. For experimental metadata, the partner who provided the information will be specified differently from the other data categories. To ensure that the relevant information is available for each analysis, information like e.g., the country of the experiment must be provided for both the contextual data and experiment metadata.

4.2.2 Data

Information stored for the data category "experiment metadata" includes e.g.:

- Start & End date of the experiment
- Status Experiment finished (yes/no) to cover multiple updates
- Partner that conducted the experiment and the location
- Primary focus of the experiment, such as user assessment, technical evaluation, impact assessment
- Experiment type, such as test on public roads, test on test tracks, driving simulator study
- ADF and technical enablers tested in the experiment
- Time of day when the tests were conducted
- Number of test runs under specific circumstances (weather, temperature, lighting conditions, etc.)
- Travelled distance and driving time under different circumstances, in different operation modes and on different road types and in total
- Addressed use cases
- Number of encountered driving scenarios
- Number of activations and deactivations of ADF/technical enabler
- Histogram of vehicle's velocity in different operation modes

The list of all stored experiment metadata is given in Annex 1.

4.3 Questionnaire data

4.3.1 General information

The purpose of the questionnaire data is to collect information on how the test participants experienced the ADFs or how the general public reacted to the presented ADF concepts. The collected subjective evaluations will be used to address the research questions in the subprojects *Users* (SP6) and *Effects* (SP7) (e.g., mobility impacts or willingness to pay for ADFs). The questionnaire data can also be used to validate the conclusions drawn based on the vehicle data. For example, if an enabler reduces jerkiness of driving, from the questionnaire data it is possible to see whether that is reflected in experienced comfort, trust, or perceived safety.

Questionnaire data will be collected in the subproject *Operations* (SP5). A common questionnaire for that purpose will be designed in WP4.6. The user studies in SP6 will use the same questionnaires when applicable. There can also be specific questionnaires that are used only by one or a few of the studies.

4.3.2 Data

Below is a preliminary list of topics in the common questionnaire designed for the SP5 Operations studies. The questionnaire has two parts: the pre-drive questionnaire, to be completed before experiencing the ADF, and the post-drive, to be completed after the experiment. The post-drive questionnaire can be filled in multiple times by the same driver if the experiment makes repeated measurements. In this case, the user ID and the trial number should be stored in the contextual data.

The draft pre-drive questionnaire contains, for instance, questions or scales composed of multiple questions on:

- Sociodemographic background information
- Current mobility behaviour
- Driving experience by passenger car and professional driving experience
- Experience with ADAS
- Willingness to use automated driving systems before experiencing them
- Perceived safety of driving in general
- General personality, sensation seeking and technological readiness questionnaires
- Tendency to experience motion sickness when being a passenger

The post-drive questionnaire contains, for instance, questions or scales composed of multiple questions on:

- Perceived usefulness and satisfaction
- Willingness to use
- Expectation to travel more with ADF
- Experience with the system (comfort, demand, fun, tiredness, motion sickness) including take-over situations
- Trust in ADF
- Perceived safety of ADF
- Non-driving related activities during automated driving
- Willingness to pay
- Value of travel time

4.4 Performance indicator data

4.4.1 General information

The purpose of the "performance indicator" data category is to collect the relevant (scalar) values for answering the Hi-Drive research questions. Thus, this data category is mainly of relevance for the subprojects *Effects* (SP7) and *Users* (SP6). It is also partly for the assessment of technical enablers within subproject *Enabler* (SP2), which will dive deeper into the technical aspects of the enablers.

To calculate the PIs, the input from pilot drives on public roads and test tracks will be used, as well as data from user tests and simulations. Once uploaded into the database, the data should be accessible to the data provider (at least their own data) and the evaluation partners. The PIs should be uploaded on a regular basis in the case of long-lasting (e.g., over several months) experiments. For shorter experiments, the PIs will be uploaded once the experiment is finished. Since some PIs are calculated for specific driving scenarios (see above), they depend on both the detected driving scenarios and their definition. Thus, it is possible to calculate the PIs only once these definitions are set. Furthermore, it would be beneficial to be able to recalculate the PIs if the implementation of the scenario detection changes during the project. The uploaded data will contain the contextual information as defined in Chapter 4.1 and Annex 2. The data provider will not be named explicitly.

The PI data will be reported for drives with ADF active and not active and the ADF with enabler passive/active/off. Reporting of pre-test data is not needed. PI data should be reported on driving scenario, test run, driver, and trip level, depending on the PI and the relevant research question.

4.4.2 Data

The final list of the Hi-Drive PIs will be available once the final list of research questions is available. This deliverable is scheduled for project month 18 (end of 2022) and therefore uses the preliminary list of PIs.

The following table presents a selection of important PIs for the technical assessment in the *Effects* subproject along with the required signals. An overview of the PIs and the level at which they will be reported will be provided in the upcoming Hi-Drive deliverables on the evaluation methodology.

Table 4.1: List of selected performance indicators and required signals (a list of abbreviations and acronyms is provided on page 50)

Performance Indicator	Required Signals
N(THW<Threshold)/scenario	Distance to another vehicle, velocity ego vehicle, signals for driving scenario detection3
std(THW)	Distance to another vehicle, velocity ego vehicle
min(THW)	Distance to another vehicle, velocity ego vehicle
mean(THW)	Distance to another vehicle, velocity ego vehicle
N(TTC<Threshold)/scenario	Distance to another vehicle, velocity ego vehicle, velocity of another vehicle, signals for driving scenario detection
Min(TTC)	Distance to another vehicle, velocity ego vehicle, velocity other vehicle
N(ax<Threshold)/scenario	Longitudinal acceleration ego vehicle, signals for driving scenario detection
std(ay)	Lateral acceleration ego vehicle
max(abs(ay))	Lateral acceleration ego vehicle
sum(ax^2) / km,	Longitudinal acceleration ego vehicle, travelled distance ego vehicle
mean(ax)	Longitudinal acceleration ego vehicle
std(ax)	Longitudinal acceleration ego vehicle
min(ax)	Longitudinal acceleration ego vehicle

[3] The list of driving scenarios is still under discussion and will be provided later. Once the list is available, the relevant signals can be identified. However, it is expected that the current list of signals in section 5.2 will cover the driving scenarios.

Performance Indicator	Required Signals
max(ax)	Longitudinal acceleration ego vehicle
sum(jerk^2)	Longitudinal acceleration ego vehicle
std(v)	Velocity ego vehicle
mean(v)	Velocity ego vehicle
mean(dist)	Distance to another vehicle (longitudinal, lateral, absolute)
std(dist)	Distance to another vehicle (longitudinal, lateral, absolute)
min(dist)	Distance to another vehicle (longitudinal, lateral, absolute)
sum of positive accelerations/km	Longitudinal acceleration ego vehicle, travelled distance ego vehicle
%distance(v>speed limit)	Travelled distance ego vehicle, velocity ego vehicle
std(rel speed) to relevant vehicles	Velocity ego vehicle, velocity other vehicle
min(rel speed) to relevant vehicles	Velocity ego vehicle, velocity other vehicle
max(rel speed) to relevant vehicles	Velocity ego vehicle, velocity other vehicle
mean(rel speed) to relevant vehicles	Velocity ego vehicle, velocity other vehicle
N(abs(lateral position)>Th)/ N driving scenario	Lateral position ego vehicle, signals for driving scenario detection
Percentage of time(abs(lateral position)>Th)	Lateral position ego vehicle, driving time
std(lateral position)	Lateral position ego vehicle
mean(lateral position)	Lateral position ego vehicle
min(gap behind)	Distance to another vehicle (rear)
%situations managed per category (environmental condition)	System status, scenario managed status, driving time, travelled distance ego vehicle, signals for environment condition detection (weather, lighting, road)
%situations managed per category (road infrastructure)	System status, scenario managed status, driving time, travelled distance ego vehicle, signals for detection of infrastructure (e.g. tunnel, pedestrian crossing, merging area)
%situations managed per category (driving scenario)	System status, scenario managed status, driving time, travelled distance ego vehicle, signals for driving scenario detection
Frequency of TOR per trip	TOR status, driving time, travelled distance ego vehicle,
Frequency of TOR per category (environmental condition)	TOR status, driving time, travelled distance ego vehicle, signals for environment condition detection (weather, lighting, road)

Performance Indicator	Required Signals
Frequency of TOR per category (road infrastructure)	TOR status, driving time, travelled distance ego vehicle, signals for detection of infrastructure (e.g. tunnel, pedestrian crossing, merging area)
Frequency of TOR per category (driving scenario)	TOR status, driving time, travelled distance ego vehicle, signals for driving scenario detection
Percentage of time (AD available in ODD/overall) / per environmental condition	ADF status, driving time, travelled distance ego vehicle, signals for environment condition detection (weather, lighting, road)
Percentage of time (AD available in ODD/overall) / per road infrastructure	ADF status, driving time, travelled distance ego vehicle, signals for detection of infrastructure (e.g. tunnel, pedestrian crossing, merging area)
Percentage of time (AD available in ODD/overall) / per driving scenario	ADF status, driving time, travelled distance ego vehicle, signals for driving scenario detection

4.5 Time series data

4.5.1 General information

The "time series" data category is intended to provide the project partners with more detailed data at driving scenario level without enabling reengineering or benchmarking of the ADF functions. The time series data category is closely linked with the "aggregated time series" category, which offers the Hi-Drive partner the possibility to handle more sensitive data. It is expected that the time series data will be limited in most cases to baseline data collected from manual driving, unless partners explicitly wish otherwise. The main relevance of the data category is for stakeholders in subproject *Effects* (SP7) who are involved the detection of scenarios or edge cases of the scenario databases and who engage in simulation of specific scenarios (driving and traffic scenario). The usefulness of these two data categories depends on each other since both will be used in the *Effects* evaluation.

The list of signals reflects several categories of information: driver behaviour, system status, movement of the ego vehicle itself, and the relation of the ego vehicle to its surrounding objects. The signals will be in a common format, with a unified sample rate and agreed-upon units, obscuring any specific information on the vehicle's sensors. Contextual metadata for each driving scenario, such as weather or lighting conditions, is also needed.

The time series data will be provided at the level of driving scenarios, and the length of the series depends on them. The driving scenarios will be defined later in the project and will influence the exact usage of the time series data. The data will be used mostly during the latter parts of the project, but if they are available shortly after recording, the work could advance more quickly.

4.5.2 Data

In the table below are the most relevant signals for the time series data category.

Table 4.2: Signals that should be logged for the time-series data category.

Ego vehicle kinematics include:	Driver and system behaviour includes:	Driver and system behaviour includes:	For objects surrounding the ego vehicle:
• Velocity(m/s) • Longitudinal acceleration (m/s^2) • Lateral acceleration (m/s^2) • Yaw rate (rad/s) • Position in lane (m)	• Steering wheel angle (rad) • Brake pedal position (%) • Throttle pedal position (%) • Active gear (gear number) • Safety driver intervention (seconds from scenario start) • Indicator status (category) • Brake light status (category) • ESP status (on/off)	• ABS status (on/off) • ACC status (category) • Activation of AEB (on/off) • ADF status (category) • Lane keeping status (category)	• Object id (integer) • Object class (category) • Relative position to ego x (m) • Relative position to ego y (m) • Relative longitudinal speed to ego (m/s) • Relative lateral speed to ego (m/s) • Lane assignment (left/right/ahead/behind) • For the "leading" object ahead (in addition to signals for object surrounding the ego vehicle): • Time gap to lead vehicle (s) • Longitudinal acceleration (m/s^2) • Lateral acceleration (m/s^2) • Position in lane (m) • Brake light status (yes/no)

Additional information that is considered for the time series data is:

- Photo/video (low resolutions) of scene (-)
- Intervention by safety driver (category)
- Id of communicating vehicle in case of V2X communication (-)

4.6 Aggregated time series data

4.6.1 General information

Aggregated time series data serves a similar purpose to that of time series data (section 4.5). However, the main difference is that the data stored is represented in a more aggregated format, which omits some details of the collected input data for confidentiality reasons. Like the time series category, also the aggregated time series will be used for the Scenario and Edge Case databases. One example of aggregated time series data is given in Figure 4.1. In this example, the resolution of the original velocity signal is reduced.

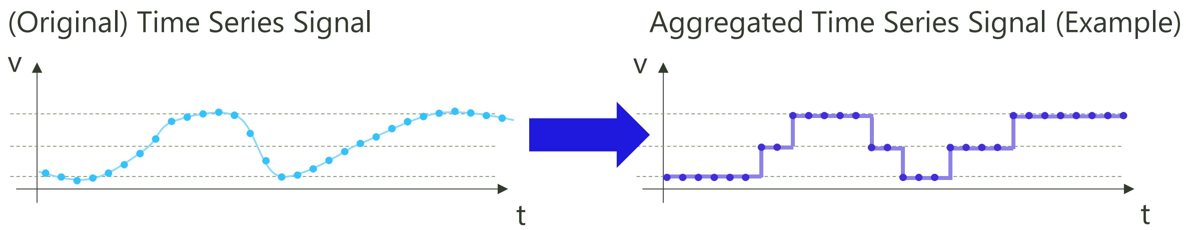

Figure 4.1: Example of an original time series signal converted into an aggregated time series signal by reducing the signal's resolution.

By leaving out these details, the data would not enable any reengineering of the behaviour of the ADF with which the data were recorded. Still, the data should contain enough information to enable the planned evaluations by providing information on the temporal development of a driving scenario. Aggregated time series data is thus a data category considered mainly for drives with an active ADF. Moreover, aggregated time series data bears a greater potential of combining data from different data sources, like on-road and simulator studies.

For aggregating of the collected data, different options will be considered:

- Reducing the resolution of signals
- Sampling down the frequency at which signals are stored
- Applying a low-pass filter
- Giving signal values at characteristic points of other signals (e.g., distance to lead vehicle at the maximum decelerations, velocity at the lowest positive time-to-collision value
- For complex driving scenarios, giving signal values at transitions between sub-scenarios of which the complex scenario can be constructed: for example, for an overtaking scenario giving signal values at the start of the following phase, the start and end of the lane change to the left, and when passing the vehicle being overtaken.

- Curve fitting by applying a parameterised description of a curve which can represent certain signals in a defined scenario, e.g., a sinusoidal shape for the steering angle over a lane change
- Adding noise to the data
- Aggregating over multiple instances of a driving scenario, e.g., giving the mean steering angle signal over multiple lane changes

For aggregating the behaviour of objects surrounding the ego vehicle, dedicated methods will be considered such as:

- Occupancy grids around the ego vehicle, which does not give detailed information on the individual objects but makes it possible to judge how much room for action the ego vehicle has in a given scenario
- Applying the method presented in (Weber et al. 2019), which defines certain "roles" for surrounding vehicles in a scenario which can be described with a smaller parameter set than providing all the objects in a scenario
- Fitting parameterised trajectories to the behaviour of objects within a scenario, e.g., parameterising the lateral position of another vehicle performing a cut-in by means of a 5th order polynomial

To decide which aggregation methods are appropriate, the aspects to be considered are:

- Can the aggregation method be applied to all types of signals (discrete, continuous)?
- Does the aggregation method induce any unwanted side effects (e.g., time shift when applying a low-pass filter)?
- What parameters need to be set to apply the aggregation method, and do these need to be decided per individual system?
- Does the aggregation method sufficiently abstract the data to avoid identification of the data supplier or reengineering of the ADF or ADF performance?

The detailed considerations of which aggregation methods will be applied will be discussed by the work packages on evaluation methods and the driving scenario database.

4.6.2 Data

The basic signals list for aggregated time series data is similar to the information contained in time series data as listed in section 4.5.2. The most relevant signals to be stored in the aggregated time series data are given in Table 4.3.

Table 4.3: Signals that should be logged for the aggregated time-series data category.

Ego vehicle state:	**Signals given for all objects:**	**Additional signals for leading object:**	**Additional information:**
• Aggregated velocity (m/s) • Aggregated longitudinal acceleration (m/s²) • Aggregated lateral acceleration (m/s²) • Aggregated position in lane (deviation from lane centre in m) • Aggregated yaw rate (rad/s) • Aggregated indicator status (category) • Aggregated brake light status (category) • Aggregated AD system status (category) • Aggregated technical enabler status (category	• Aggregated relative x- and y-position (m) • Aggregated relative longitudinal and lateral velocity (m) • Aggregated lane assignment for objects within a range (category) • Aggregated object class (-)	• Aggregated relative lateral and longitudinal acceleration (m/s²) • Aggregated position in lane (m) • Aggregated brake light status (category)	• Aggregated photo/video (low resolutions) of scene (-) • Aggregated intervention by safety driver (category) • Aggregated id of communicating vehicle in case of V2X communication (-)

5 Hi-Drive Signals

5.1 Process from data categories to signals list

This chapter provides the Hi-Drive signals list. Before presenting the actual list, the process that WP4.4 has taken to get from the different data categories to the signals list is described.

First, for each data category (see Chapter 4) the required information (i.e., what is needed to do the evaluation) was defined by a group of experts. In some cases, the work was done completely within WP4.4 (time series, aggregated time series and experiment metadata), while in other cases (performance indicator and questionnaire) the work was supported by other work packages (WP4.3 *Research Question*, WP4.6 *Method for the user evaluation*, SP6 *Users*, WP4.7 *Method effects evaluation*). Once the lists of information for each data category were available, the signals needed to derive the required information were defined. In the next steps, the signals were combined and reviewed by the WP4.4 partners. Following this review, meetings were held with the different data providers (partners running experiments in the subprojects Enabler, Operation or User). At these meetings it was checked for each data provider which signals are available. Based on their feedback, WP4.4 prepared the final signals list. Signals from the original list that were only available for a very limited number of partners have been removed from the list. Nevertheless, the partners can still log these signals if they are of importance for their work in Hi-Drive (e.g., development of a technical enabler).

5.2 Hi-Drive signals list

The Hi-Drive signals list is given in Table 5.1. All signals are required for the analysis in Hi-Drive. However, it must be noted that not all ADFs and enablers are evaluated in all categories. Therefore, it is not mandatory for each demonstrator vehicle to provide all the signals. The question of which signals are required from which ADF and enabler is addressed individually for each demonstrator vehicle/experiment. The table does not imply that every signal needs to be directly measured during the tests. In addition, experiments might log additional dedicated signals which are not submitted as time series data in the CDF. Signals that cannot be directly measured may also be provided after processing of the logged data. Logging of the signals should be done at least at a frequency of 10 Hz. Format and conversion of the signals are derived by the CDF working group.

Furthermore, some of the signals mentioned in Table 5.1 are not actually signals over time (e.g., age of test participant or country), as they remain constant for most of the test. This is especially the case for experiment information, which is required mainly for the experiment metadata and contextual data. It could even be that the information is derived simply from a test protocol. Nevertheless, the information is needed for the evaluation in Hi-Drive and these signals are therefore included in the list to ensure that the information is recorded during the test if relevant for the combination of the ADF and enabler under test.

Table 5.1: Hi-Drive signals list.

No	Signal	Signal Group	Description	Unit
1	Id of transmitting object	Connectivity	Id of communicating object	Number
2	Latency	Connectivity	Time difference between sent and received information	ms
3	Package loss	Connectivity	Number of actual received packages divided by number of sent packages	%
4	Active gear	Driver Input	Current driven gear	Number
5	Brake pedal position	Driver Input	Position of brake pedal, % of range	%
6	Brake pedal pressure	Driver Input	Brake system pressure, % of min/max range to normalise the signal	%
7	Safety driver intervention	Driver Input	Indicates any intervention by safety feature (takeover, steering, braking, accelerating)	[1/0] or Categories
8	Steering wheel angle	Driver Input	Steering wheel angle	Rad
9	Throttle pedal position	Driver Input	Position of throttle pedal, % of range	%
10	GNSS position latitude	Ego vehicle kinematics	GNSS longitudinal position	WGS 84
11	GNSS position longitude	Ego vehicle kinematics	GNSS lateral position	WGS 84
12	GNSS Quality	Ego vehicle kinematics	GNSS Quality	TBD
13	Heading angle (vehicle)	Ego vehicle kinematics	Heading of ego vehicle (axis)	Rad
14	Lateral Acceleration	Ego vehicle kinematics	Lateral acceleration (Y-axis, measured in CoG)	m/s^2
15	Lateral speed in lane	Ego vehicle kinematics	Lateral speed relative to the lane marking	m/s

No	Signal	Signal Group	Description	Unit
16	Longitudinal Acceleration	Ego vehicle kinematics	Longitudinal acceleration (X-axis, measured in CoG)	m/s^2
17	Odometer	Ego vehicle kinematics	Vehicle odometer reading	km
18	Position in Lane	Ego vehicle kinematics	Lateral position in lane (measured to centre of vehicle)	m
19	Velocity	Ego vehicle kinematics	Velocity of ego vehicle as reported by the ABS/wheel sensing module	m/s
20	X-position	Ego vehicle kinematics	Relative x-position (longitudinal position) of ego vehicle towards the start of the scenario/trip	m
21	Yaw Rate	Ego vehicle kinematics	Yaw rate of vehicle	rad/s
22	Y-position	Ego vehicle kinematics	Relative y-position (lateral position) of ego vehicle towards the start of the scenario/trip	%
23	Lighting Condition	Environment	Ambient light level at location of ego vehicle. Requires signal from light sensor. Measured as the natural log of the raw value (because the range is huge)	ln(lux) / Categories
24	Rain sensor	Environment	Rain sensor signal	Categories
25	Road condition	Environment	Current road condition at location of ego vehicle	Categories
26	Temperature	Environment	External ambient temperature at location of ego vehicle	°C
27	Weather information	Environment	Current weather at location of ego vehicle (requires information from rain sensor, wipers)	Categories
28	Age (Subject)	Experiment information	Only for non-professional drivers	Categories or Number
29	Country	Experiment information	Country of test site	Categories
30	Participant ID	Experiment information	Unique ID for each driver; important to map answers in the questionnaire	ID

No	Signal	Signal Group	Description	Unit
31	Driving Scenario	Experiment information	Current detected driving scenario	Categories
32	Gender	Experiment information	Only for non-professional drivers	Categories
33	Number of trials/runs per driver	Experiment information	Number of trials/runs a test person has already conducted	Number
34	Role of driver	Experiment information	Role of the test person during test	Categories
35	Purpose of experiment	Experiment information	Purpose of the experiment	Categories
36	Questionnaire status	Experiment information	Describing if and when questionnaires are conducted	Categories
37	Seat position	Experiment information	Seat position of test person during the test	Categories
38	Status "unusual things"	Experiment information	Flag to mark test in which something went wrong and to indicate tests that should not be considered for the assessment	[0/1]
39	Status practice drive	Experiment information	Status of current test drive – Practice drives should not be considered for the analysis	[0/1]
40	Test scenario	Experiment information	Currently tested scenario	Categories
41	Trial or test run number	Experiment information	Number of test runs/trials in the experiment	Number
42	Type test person	Experiment information	Type of test person	Categories
43	Use Case	Experiment information	Currently tested Hi-Drive use case	Categories
44	Traffic Light Status	Infrastructure detection	Status of an external traffic light	Categories
45	Infrastructure detection – in XY	Infrastructure detection	Determine that the vehicle is currently in a certain infrastructure situation – which situation is relevant depends on the ADF and enabler (XY: intersection, construction site, tunnel, motorway entry, passing a motorway entry)	[1/0]

No	Signal	Signal Group	Description	Unit
46	Infrastructure detection – in ODD	Infrastructure detection	Detection that the infrastructure related ODD condition is fulfilled	[1/0]
47	Distance to XY	Infrastructure detection	Distance from the current location of the ego vehicle to the start of an infrastructure condition (XY: intersection, construction site, tunnel, end of lane, motorway entry, passing a motorway entry)	m
48	Video (driver)	Other	Video feed of driver	-
49	Video (Front)	Other	Video feed from front looking camera	-
50	Lane marking type	Road	Type of lane markings	Categories
51	Lane width	Road	Width of lane in which the ego vehicle is currently driving	m
52	Number of lanes	Road	Number of lanes at current location of ego vehicle in driving directions	[-]
53	Road Type	Road	Road type according to Hi-Drive classification, to be discussed	Categories
54	Speed Limit	Road	Speed limit at current location of ego vehicle	m/s
55	Driver angle head n/ Head tracking	Status Driver	Heading position (yaw, roll & pitch angle)	Categories | Orientation on 3 axes
56	Driver attention	Status Driver	Status driver attention	Categories
57	Status hands-on-wheel detection	Status Driver	Status hand of detection	Categories
58	Crash detection	System status	Status crash	[1/0]
59	Enabler interacting with ADF	System status	Describes whether the technical enabler is interacting with the ADF while the ADF is operational	[0/1]
60	Enabler Status	System status	Status of enabler	Categories
61	Status ABS Active	System status	Whether the ABS system is intervening	[1/0]
62	Status ACC	System status	Status ACC	Categories

No	Signal	Signal Group	Description	Unit
63	Status ADF	System status	Status ADF	Categories
64	Status AEB	System status	Status AEB	Categories
65	Status Brake Light	System status	Status whether brake lights of ego vehicle are on or off	[1/0]
66	Status ESC Active	System status	Whether electronic stability control / traction control system is intervening	[1/0]
67	Status Horn	System status	Status horn	[1/0]
68	Status within ODD	System status	Whether ego vehicle is/isn't in the ODD of the ADF	[1/0]
69	Status lane keeping	System status	Status lane keeping	Categories
70	Status ADF Level ≤2	System status	Whether any ADF level ≤2 is active	Categories
71	Status MRM	System status	Status minimal risk manoeuvre	[1/0]
72	Status Take Over Request active	System status	Whether a take-over request to return control to the driver is active	[1/0]
73	Status Turn Indicator	System status	Status turn indicator	[1/0]
74	Status Windscreen wipers	System status	Status windscreen wiper	[1/0]
75	Time since start	Time	Time since start of trip/run/experiment	s
76	UTC Time	Time	Universal Time Coordinated	Date/Time
77	Object lane assignment	Traffic object (all)	Lane assignment of object x	Categories
78	Longitudinal Acceleration Object	Traffic object (all)	Longitudinal acceleration of another vehicle	m/s^2
79	Object class	Traffic object (all)	Classification of object type	Categories
80	Object id	Traffic object (all)	ID of object x (should be constant during driving scenario)	[-]
81	Object source tag	Traffic object (all)	Tag showing if the detection of object x is based on one or	Categories

No	Signal	Signal Group	Description	Unit
			multiple sensor sources (incl. V2X)	
82	Relative lateral velocity to ego	Traffic object (all)	Relative y-velocity to object x (lateral direction) measured from ego vehicle	m/s
83	Relative velocity speed to ego	Traffic object (all)	Relative x-velocity to object x (longitudinal direction) measured from ego vehicle	m/s
84	Relative position to ego x	Traffic object (all)	Relative x-distance to object x (longitudinal direction) measured from ego vehicle	m
85	Relative position to ego y	Traffic object (all)	Relative y-distance to object x (lateral direction) measured from ego vehicle	m
86	Distance to lead vehicle	Traffic object (Lead vehicle)	Longitudinal distance to lead vehicle	m
87	Lateral acceleration lead vehicle	Traffic object (Lead vehicle)	Longitudinal acceleration of lateral object	m/s^2
88	Longitudinal acceleration lead vehicle	Traffic object (Lead vehicle)	Longitudinal acceleration of lead object	m/s^2
89	Velocity of lead vehicle	Traffic object (Lead vehicle)	Absolute velocity of lead vehicle	m/s
90	Time gap to lead vehicle	Traffic object (Lead vehicle)	Time gap from ego vehicle to lead vehicle	s

6 Conclusions and outlook

6.1 Conclusion

Starting with a brief introduction to the Hi-Drive evaluation concept, this deliverable described the data requirements of the project. This includes the data collection process, relevant stakeholders, and a description of the core principles for data storage and exchange in Hi-Drive. Next was a detailed description of the Hi-Drive data categories; overall, five categories plus the contextual data have been defined as follows:

- Experiment metadata: data intended to provide an overview of the Hi-Drive experiments and their progress
- Questionnaire data: data intended to collect information obtained from the Hi-Drive questionnaires
- Performance indicator data: data to be used to assess the Hi-Drive research questions. This data will be derived from the information stored during the Hi-Drive tests and will be uploaded to the CDB
- Time series and aggregated time series data: data intended to be used for impact assessment and the Hi-Drive scenario and edge case database. The aggregated time series data is a modification of the time series data, which should allow the Hi-Drive partners to provide information in a more anonymised way.

In addition, contextual data will be provided alongside the data in the four categories. Its purpose is to provide the evaluation partners with any necessary background information about the evaluation.

A list of signals to be logged by the Hi-Drive partners in the experiments was defined. The signals are used as the source to derive the required information in the five data categories. The signals list was discussed with the various data providers. In this context, is important to underline that the list provides the complete set of signals that will be used for the evaluation in Hi-Drive. The partners are not obliged to store all the signals in their experiments but will focus rather on the signals relevant to them. Which signals need to be stored in an experiment depends on the scope of the experiment, the technical enabler, the ADF, and the test environment.

The general concept presented in this deliverable regarding the data requirements (i.e., for defining data content and common format) of Hi-Drive will be reviewed from time to time and further elaborated as needed. This iterative process follows along with developments in the project and is necessary to react to them. Some flexibility of the concept is therefore required to ensure an efficient evaluation.

One challenge encountered during preparation of general concept has been the diversity of Hi-Drive technologies (enablers, ADFs and approaches in the user domain) being developed and investigated during the project. It was nearly impossible to consider every signal that a certain technology might require. The compromise was to consider in the signals list those signals that can be provided by multiple partners, and to focus primarily on those that will be needed later for the evaluation in the Hi-Drive subprojects *User* and *Effects*. The partners do not have to log every signal on the list but only those that are relevant for their ADF and enabler. For the technical evaluation of certain enablers, additional signals might be required. Here the partners are asked to decide in cooperation with the evaluation partner which additional signals need to be logged, as experts on the enablers will know best which signals best describe a certain measure.

6.2 Outlook on evaluation

This deliverable provides specification of data usage in Hi-Drive and different data categories needed for evaluation activities in the project. Data agreed upon with the data owners to be shared for evaluation provides the basis for feasibility checking of the initial list of research questions in terms of related data needs.

The data specifications of this deliverable serve also as input to the overall evaluation plan that is prepared for the evaluation in *Users* and *Effects*. In this evaluation plan, the methods are set for different research questions. Furthermore, it is described which scenarios, functions, and enablers will be addressed by whom and based on data from where.

An important consideration to be taken when specifying the driving scenario and edge case databases is the selection of appropriate aggregation methods for aggregated time series data. In this context, preselected approaches for aggregating time series data will be discussed with the different stakeholders to judge the suitability and parameterisation of the approach for different signals. Later, the selected approaches need to be applied.

6.3 Outlook on data management

As discussed in this deliverable, multiple data categories are available for the different assessments and evaluations within Hi-Drive. As none of these data categories is analysed by one partner alone, solutions for storing this data need to be found and defined.

As mentioned earlier, the work begun within L3Pilot towards data harmonisation and sharing, especially as regards the CDF, will continue with Hi-Drive. To enable the further use of data already recorded within L3Pilot (where appropriate), the data format will be kept compatible with the L3Pilot data where possible.

Building further upon the successful work of L3Pilot, it makes sense also to consider the data storing and sharing options developed within L3Pilot for further use in Hi-Drive. This will include the database technology employed within L3Pilot for sharing of PIs, which will be extended to cover additional PIs necessary for the evaluations within Hi-Drive. The L3Pilot database technology already included a preliminary option of storing questionnaire data. This will be extended in this project to handle the Hi-Drive questionnaires.

As for contextual data, this was already present in earlier versions of the database used, albeit not to the extent to which it is now defined and used within Hi-Drive. Therefore, methods and tools will need to be adapted and developed to integrate this data into the evaluation workflow.

Experiment metadata, and its usage for experiment overview and control to this extent, is new to the Hi-Drive consortium. However, it is good practice, and established methods of data visualisation can be employed. The focus here will need to be on the option to continuously update the database with incoming changes from the experiments, to ensure the provision of an adequate and up-to-date overview of the data.

For the two other data categories, time series and aggregated time series, the data storage options will need to be closely evaluated and assessed regarding the appropriate database architecture, while also considering ease of data access and usage by the partners later on. The amount of data stored here is different compared to the PIs. Since this use case was not really in the focus of L3Pilot (the database could handle the use case, but it was not applied), other projects such as PEGASUS, VV Methods, StreetWise, etc. can provide valuable input.

References

Andreone, L., Borodani, P., Pallaro, N., Tango, F., Bellotti, F., Weber, H, Altpeter, B., Reimer, F., Griffon, T., Sauvaget, J. L., Geronimi, S., Page, Y., Guerineau, T., Willemotte, G., Mäkinen, T. (2021). Pilot Reporting Outcomes. L3Pilot deliverable D6.5, Version 1.0.

Bellotti, F., Berta, R., De Gloria, A., Osman, N. (2020). Data delivery to evaluation & common data set for future research. L3Pilot deliverable D6.4. Version 1.0.

Bellotti, F.; Osman, N.; Arnold, E.H.; Mozaffari, S.; Innamaa, S.; Louw, T.; Torrao, G.; Weber, H.; Hiller, J.; De Gloria, A.; Dianati, M.; Berta, R. (2020b). Managing Big Data for Addressing Research Questions in a Collaborative Project on Automated Driving Impact Assessment. *Sensors* 2020, *20*, 6773. https://doi.org/10.3390/s20236773.

European Parliament (2016). Regulation (EU) 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). REGULATION (EU) 2016/679, https://eur-lex.europa.eu/eli/reg/2016/679/oj.

FOT-Net Data and CARTRE (2019). Data sharing framework v1.1, January 2019. https://www.connectedautomateddriving.eu/wp-content/uploads/2021/09/Data-Sharing-Framework-v1.1-final.pdf

FOT-Net, CARTRE & ARCADE (2021). FESTA Handbook. version 8, September 2021. 227 p. https://www.connectedautomateddriving.eu/wp-content/uploads/2021/09/FESTA-Handbook-Version-8.pdf.

L3Pilot consortium (2021). Final Project Results. L3Pilot deliverable D1.7, Version 1.0.

Weber, H., Bock, J., Klimke, J., Roesener, C., Hiller, J., Krajewski, R., Zlocki, A., Eckstein, L. (2019). A framework for definition of logical scenarios for safety assurance of automated driving, Traffic Injury Prevention, 20:sup1, S65-S70, DOI: 10.1080/15389588.2019.1630827.

Hiller, J., (2021). Data Handling and Sharing. Presentation at L3Pilot Final Event during the ITS World Congress 2021, https://l3pilot.eu/fileadmin/user_upload/Downloads/Final_Event/13102021/L3Pilot_Final_Event_presentation_08_Johannes_Hiller_20211013.pdf.

Kessler, C., Etemad, A. (2012). FOT Data. euroFOT Deliverable D6.8, Version 1.4, https://www.eurofot-ip.eu/download/library/deliverables/eurofotsp620121212v14dld68_fot_data.pdf.

Koskinen, S., Christen, F., Hiller, J., Neila, N., Svanberg, E., Kremer, M., (2021). Guidelines and lessons learned on Pilot Tools and Data. L3Pilot deliverable D5.2, Version 1.0.

List of abbreviations and acronyms

Abbreviation	Meaning
AD	Automated Driving
ADAS	Advanced Driver Assistance Systems
ADF	Automated Driving Function
AV	Automated Vehicles
Ax	Longitudinal Acceleration
Ay	Lateral Acceleration
CAD	Connected and Automated Driving
CAV	Connected and Automated Vehicle
CDB	Common Database
CDF	Common Data Format
CoG	Centre of Gravity
CoP	Code of Practice
dist	Distance between two vehicles (measured from bumper to bumper)
DSDB	Driving Scenario Database
EC	European Commission
ECDB	Edge Case Database
FOT	Field Operational Test
GDPR	General Data Protection Regulation
GNSS	Global Navigation Satellite System
IP	Intellectual Property
L3	SAE Level 3 "Conditional Automation"
Max	Maximum
Min	Minimum
MRM	Minimal Risk Manoeuvre
N	Number (count) of events
ODD	Operational Design Domain
PI	Performance Indicator
rel speed	Relative speed between two vehicles
SP	Subproject
std	Standard derivation
THW	Time Headway
TOR	Take-Over Request
TTC	Time To Collision
WP	Work Package

Annex 1 Hi-Drive experiment metadata

Table 6.1: Hi-Drive experiment metadata information.

Information	Category	Abbreviation	Unit
Start Date		Date_start	yyyy/mm/dd
End Date		Date_end	yyyy/mm/dd
Name of Experiment		Experiment_Name	Name
Is the experiment finished?		Experiment_status	Binary
Number of uploads for this experiment		Experiment_upload	Integer
Company		Experiment_Company	Acronym
Primary Focus of experiment	ADF (SP5)	Experiment_Focus_SP5	Binary
	Enabler (SP2)	Experiment_Focus_SP2	Binary
	User (SP6)	Experiment_Focus_SP6	Binary
	Effects / Technical Evaluation (SP7)	Experiment_Focus_SP7_Effects	Binary
	Impact Assessment (SP7)	Experiment_Focus_SP7_Impact	Binary
	Outreach (SP8)	Experiment_Focus_SP8	Binary
Overall number of test runs in experiment		Experiment_tests	Integer
Time of tests in experiment	Number of tests 0 – 6	Time_day_0_6	Integer
	Number of tests 6 – 12	Time_day_6_12	Integer

Information	Category	Abbreviation	Unit
	Number of tests 12 – 18	Time_day_12_18	Integer
	Number of tests 18 – 24	Time_day_18_24	Integer
Duration of tests in experiment	Number of tests under 15 min	Duration_15	Integer
	Number of tests less than 30 min	Duration_30	Integer
	Number of tests less than 1 h	Duration_60	Integer
	Number of tests less than 1 h 30 min	Duration_90	Integer
	Number of tests less than 2 h 0 min	Duration_120	Integer
	Number of tests more than 2 h	Duration_120plus	Integer
Country in which experiment took place		Country	[2-letter country code]
Location		Location	Text
Experiment scope	Number of tests for manual driving (baseline) data collection	Tests_manual_baseline	Integer
	Number of tests for enabler testing – Connectivity and digital infrastructure	Tests_enabler_connectivity	Integer
	Number of tests for enabler testing – high precision positioning	Tests_enabler_positioning	Integer
	Number of tests for enabler testing – cybersecurity	Tests_enabler_cybersecurity	Integer

Information	Category	Abbreviation	Unit
	Number of tests for enabler testing – ML Techniques	Tests_enabler_ML	Integer
	Number of tests for ADF Baseline	Tests_ADF_Basline	Integer
	Number of tests for testing ADF without enabler	Tests_ADF_wo_enabler	Integer
	Number of tests for testing ADF with enabler	Tests_ADF_w_enabler	Integer
	Number of tests for training of test participant	Tests_training_participant	Integer
	Number of tests for developing enabler	Tests_developing_enabler	Integer
	Number of tests for developing ADF	Tests_developing_ADF	Integer
	Number of tests for user – User Acceptance & Awareness	Tests_User_Acceptance	Integer
	Number of tests for user – Human-like driving and user comfort	Tests_User_Comfort	Integer
	Number of tests for user – User monitoring and related HMI	Tests_User_Monitoring	Integer
	Number of tests for user – Interaction with other road users	Tests_User_Interaction	Integer
	Number of (simulation) tests – Safety impact assessment	Tests_Sim_Safety	Integer
	Number of (simulation) tests – Environmental impact assessment	Tests_Sim_Environmental	Integer
	Number of (simulation) tests – Traffic efficiency impact assessment	Tests_Sim_ Traffic_efficiency	Integer

Information	Category	Abbreviation	Unit
	Number of (simulation) tests – Enabler evaluation	Tests_Sim_enabler	Integer
	Number of responding persons to experiment questionnaire	Tests_Questionnaire	Integer
	Number of responding persons survey	Tests_Survey	Integer
Experiment type	Tests on public road	Experiment_type_Road_public	Binary
	Wizard of Oz (public road)	Experiment_type_WoOz_public	Binary
	Test track test / Closed environment	Experiment_type_Track	Binary
	Wizard of Oz (test track)	Experiment_type_WoOz_track	Binary
	(Virtual) Simulation	Experiment_type_virt_Sim	Binary
	Simulator (Human)	Experiment_type_Simulator	Binary
	Simulation (Hardware)	Experiment_type_hard_Sim	Binary
	Survey	Experiment_type_Survey	Binary
	Focus Group	Experiment_type_Focus_Group	Binary
Involved enablers	Vehicle-to-Vehicle Communications (direct)	Enabler_V2V_direct	Binary
	Vehicle-to-Vehicle Communications (cellular network)	Enabler_V2V_cellular	Binary
	Vehicle-to-Infrastructure and Infrastructure-to-Vehicle Communications	Enabler_V2I	Binary
	Vehicle to Cloud (Edge and Core)	Enabler_V2Cloud	Binary

Information	Category	Abbreviation	Unit
	Threat analysis and risk assessment	Enabler_TARA	Binary
	Recommendations and V2X cyber-risk mitigation techniques	Enabler_Cybersecurity	Binary
	Geo-referenced cloud services	Enabler_Cloud	Binary
	Sensor fusion for localisation	Enabler_Localization	Binary
	Positioning relying on ranging signals	Enabler_Positioning	Binary
	CADF ML Toolkit	Enabler_ML_Toolkit	Binary
	CADF ML perception, object detection and classification	Enabler_ML_Perception	Binary
	CADF ML decision-making	Enabler_ML_decision	Binary
	CADF ML Driver Monitoring	Enabler_ML_monitoring	Binary
	None	Enabler_none	Binary
Number of tested persons – gender	Overall	Participant_all	Integer
	Male	Participant_male	Integer
	Female	Participant_female	Integer
	Unknown	Participant_unknown	Integer
Number of tested persons – age	16 - 25	Participant_16_25	Integer
	25 - 50	Participant_25_50	Integer

Information	Category	Abbreviation	Unit
	50 - 75	Participant_50_75	Integer
	>75	Participant_75plus	Integer
Safety Driver present (Driver Seat)	Driver Seat	Safety_driver_driver	Binary
	Other seat in car	Safety_driver_other	Binary
	Outside vehicle (remote control)	Safety_driver_outside	Binary
	No safety driver	Safety_driver_none	Binary
Number of subjects / tested persons on/at	Driver seat	Participant_driver	Integer
	Any other seat in car	Participant_other	Integer
	Outside vehicle (other/external road users / remote control)	Participant_outside	Integer
	No subjects / tested persons	Participant_none	Integer
Temperature	Number of tests below -10°C	Temp_below_-10	Integer
	Number of tests between -10°C and 0°C	Temp_-10_0	Integer
	Number of tests between 0°C and 10°C	Temp_0_10	Integer
	Number of tests between 10°C and 20°C	Temp_10_20	Integer
	Number of tests between 20°C and 30°C	Temp_20_30	Integer
	Number of tests above 30°C	Temp_30plus	Integer
Road Condition	Number of tests on dry roads	Road_dry	Integer

Information	Category	Abbreviation	Unit
	Number of tests on wet roads	Road_wet	Integer
	Number of tests on snowy roads	Road_snowy	Integer
	Number of tests on icy roads	Road_icy	Integer
	Number of tests on dry roads & ADF on	Road_dry_ADF	Integer
	Number of tests on wet roads & ADF on	Road_wet_ADF	Integer
	Number of tests on snowy roads & ADF on	Road_snowy_ADF	Integer
	Number of tests on icy roads & ADF on	Road_icy_ADF	Integer
	Number of tests on dry roads & ADF on & enabler on	Road_dry_ADF_Enabler	Integer
	Number of tests on wet roads & ADF on & enabler on	Road_wet_ADF_Enabler	Integer
	Number of tests on snowy roads & ADF on & enabler on	Road_snowy_ADF_Enabler	Integer
	Number of tests on icy roads & ADF on & enabler on	Road_icy_ADF_Enabler	Integer
Weather / Precipitation Conditions	Number of tests without rain	Weather_no_rain	Integer
	Number of tests with light rain	Weather_light_rain	Integer
	Number of tests with rain	Weather_rain	Integer
	Number of tests with heavy rain	Weather_heavy_rain	Integer
	Number of tests with light snowfall	Weather_light_snow	Integer

Information	Category	Abbreviation	Unit
	Number of tests with snowfall	Weather_snow	Integer
	Number of tests with heavy snowfall	Weather_heavy_snow	Integer
	Number of tests with fog	Weather_fog	Integer
	Number of tests in cloudy conditions	Weather_cloudy	Integer
	Number of tests in sunny conditions	Weather_sunny	Integer
	Number of tests without rain & ADF on	Weather_no_rain_ADF	Integer
	Number of tests with light rain & ADF on	Weather_light_rain_ADF	Integer
	Number of tests with rain & ADF on	Weather_rain_ADF	Integer
	Number of tests with heavy rain & ADF on	Weather_heavy_rain_ADF	Integer
	Number of tests with light snowfall & ADF on	Weather_light_snow_ADF	Integer
	Number of tests with snowfall & ADF on	Weather_snow_ADF	Integer
	Number of tests with heavy snowfall & ADF on	Weather_heavy_snow_ADF	Integer
	Number of tests with fog & ADF on	Weather_fog_ADF	Integer
	Number of tests in cloudy conditions & ADF on	Weather_cloudy_ADF	Integer
	Number of tests in sunny conditions & ADF on	Weather_sunny_ADF	Integer
	Number of tests without rain & ADF on & enabler on	Weather_no_rain_ADF_Enabler	Integer
	Number of tests with light rain & ADF on & enabler on	Weather_light_rain_ADF_Enabler	Integer

Information	Category	Abbreviation	Unit
	Number of tests with rain & ADF on & enabler on	Weather_rain_ADF_Enabler	Integer
	Number of tests with heavy rain & ADF on & enabler on	Weather_heavy_rain_ADF_Enabler	Integer
	Number of tests with light snowfall & ADF on & enabler on	Weather_light_snow_ADF_Enabler	Integer
	Number of tests with snowfall & ADF on & enabler on	Weather_snow_ADF_Enabler	Integer
	Number of tests with heavy snowfall & ADF on & enabler on	Weather_heavy_snow_ADF_Enabler	Integer
	Number of tests with fog & ADF on & enabler on	Weather_fog_ADF_Enabler	Integer
	Number of tests in cloudy conditions & ADF on & enabler on	Weather_cloudy_ADF_Enabler	Integer
	Number of tests in sunny conditions & ADF on & enabler on	Weather_sunny_ADF_Enabler	Integer
Lighting conditions	Number of tests in daylight	Lighting_daylight	Integer
	Number of tests at dawn	Lighting_dawn	Integer
	Number of tests at dusk	Lighting_dusk	Integer
	Number of tests at night	Lighting_night	Integer
	Number of tests in daylight & ADF on	Lighting_daylight_ADF	Integer
	Number of tests at dawn & ADF on	Lighting_dawn_ADF	Integer

Information	Category	Abbreviation	Unit
	Number of tests at dusk & ADF on	Lighting_dusk_ADF	Integer
	Number of tests at night & ADF on	Lighting_night_ADF	Integer
	Number of tests in daylight & ADF on & enabler on	Lighting_daylight_ADF_Enabler	Integer
	Number of tests at dawn & ADF on & enabler on	Lighting_dawn_ADF_Enabler	Integer
	Number of tests at dusk & ADF on & enabler on	Lighting_dusk_ADF_Enabler	Integer
	Number of tests at night & ADF on & enabler on	Lighting_night_ADF_Enabler	Integer
Addressed Use Cased	UC1	UC_name1	Binary
	UC2	UC_name1	Binary
	…	UC_name…	Binary
	UC n	UC_namen	Binary
Number of encountered driving scenarios (Overall)	Driving Scenario 1	DS_name_1	Integer
	Driving Scenario 2	DS_name_2	Integer
	…	DS_name_…	Integer
	Driving Scenario n	DS_name_n	Integer
Number of encountered driving scenarios (Manual Driving)	Driving Scenario 1	DS_name_1_manual	Integer

Information	Category	Abbreviation	Unit
	Driving Scenario 2	DS_name_2_manual	Integer
	…	DS_name_…_manual	Integer
	Driving Scenario n	DS_name_n_manual	Integer
Number of encountered driving scenarios (ADF Driving)	Driving Scenario 1	DS_name_1_ADF	Integer
	Driving Scenario 2	DS_name_2_ADF	Integer
	…	DS_name_…_ADF	Integer
	Driving Scenario n	DS_name_n_ADF	Integer
Number of encountered driving scenarios (ADF Driving with enabler)	Driving Scenario 1	DS_name_1_ADF_Enabler	Integer
	Driving Scenario 2	DS_name_2_ADF_Enabler	Integer
	…	DS_name_…_ADF_Enabler	Integer
	Driving Scenario n	DS_name_n_ADF_Enabler	Integer
Travelled distance	Overall	Distance_overall	km
	Motorway (Overall)	Distance_motorway	km
	Rural road (Overall)	Distance_rural	km
	Urban road (Overall)	Distance_urban	km
	Urban motorway (Overall)	Distance_urban_motorway	km

Information	Category	Abbreviation	Unit
	Test track (Overall)	Distance_test_track	km
	Overall & ADF on	Distance_overall_ADF	km
	Motorway & ADF on	Distance_motorway_ADF	km
	Rural road & ADF on	Distance_rural_ADF	km
	Urban road & ADF on	Distance_urban_ADF	km
	Urban motorway & ADF on	Distance_urban_motorway_ADF	km
	Test track & ADF on	Distance_test_track_ADF	km
	Overall & ADF on & enabler on	Distance_overall_ADF_Enabler	km
	Motorway & ADF on & enabler on	Distance_motorway_ADF_Enabler	km
	Rural road & ADF on & enabler on	Distance_rural_ADF_Enabler	km
	Urban road & ADF on & enabler on	Distance_urban_ADF_Enabler	km
	Urban Motorway & ADF on & enabler on	Distance_urban_motorway_ADF_Enabler	km
	Test track & ADF on & enabler on	Distance_test_track_ADF_Enabler	km
Driving time	Overall	Distance_overall	s
	Motorway (Overall)	Distance_motorway	s
	Rural Road (Overall)	Time_rural	s
	Urban Road (Overall)	Time_urban	s
	Urban motorway (Overall)	Time_urban_motorway	s
	Test track (Overall)	Time_overall_ADF	s

Information	Category	Abbreviation	Unit
	Overall & ADF on	Time_overall_ADF	s
	Motorway & ADF on	Time_motorway_ADF	s
	Rural road & ADF on	Time_rural_ADF	s
	Urban road & ADF on	Time_urban_ADF	s
	Urban motorway & ADF on	Time_urban_motorway_ADF	s
	Test track & ADF on	Time_test_track_ADF	s
	Overall & ADF on & enabler on	Time_overall_ADF_Enabler	s
	Motorway & ADF on & enabler on	Time_motorway_ADF_Enabler	s
	Rural road & ADF on & enabler on	Time_rural_ADF_Enabler	s
	Urban road & ADF on & enabler on	Time_urban_ADF_Enabler	s
	Urban motorway & ADF on & enabler on	Time_urban_motorway_ADF_Enabler	s
	Test track & ADF on & enabler on	Time_test_track_ADF_Enabler	s
Number of activations in Experiments		Activations	Integer
Number of Deactivation in Experiments	Overall	Deactivation_overall	Integer
	After TOR	Deactivation_TOR	Integer
	By human driver	Deactivation_human	Integer
	By safety driver	Deactivation_safety	Integer

Information	Category	Abbreviation	Unit
Number of TOR in Experiments		TOR	Integer
Number of MRM in Experiments		MRM	Integer
Number of interventions by safety driver		Safety_Driver_intervention	Integer
Operation time ADS	Number of ADF activations with operation time under 1 min	ADF_Time_1	Integer
	Number of ADF activations with operation time under 5 min	ADF_Time_5	Integer
	Number of ADF activations with operation time under 10 min	ADF_Time_10	Integer
	Number of ADF activations with operation time under 15 min	ADF_Time_15	Integer
	Number of ADF activations with operation time under 20 min	ADF_Time_20	Integer
	Number of ADF activations with operation time under 30 min	ADF_Time_30	Integer
	Number of ADF activations with operation time under 40 min	ADF_Time_40	Integer
	Number of ADF activations with operation time under 50 min	ADF_Time_50	Integer

Information	Category	Abbreviation	Unit
	Number of ADF activations with operation time under 60 min	ADF_Time_60	Integer
	Number of ADF activations with operation time under 90 min	ADF_Time_90	Integer
	Number of ADF activations with operation time under 120 min	ADF_Time_120	Integer
	Number of ADF activations with operation time over 120 min	ADF_Time_120plus	Integer
Operation time ADS	Number of ADF activations with operation time under 1 min & enabler on	ADF_Time_1_Enabler	Integer
	Number of ADF activations with operation time under 5 min & enabler on	ADF_Time_5_Enabler	Integer
	Number of ADF activations with operation time under 10 min & enabler on	ADF_Time_10_Enabler	Integer
	Number of ADF activations with operation time under 15 min & enabler on	ADF_Time_15_Enabler	Integer
	Number of ADF activations with operation time under 20 min & enabler on	ADF_Time_20_Enabler	Integer
	Number of ADF activations with operation time under 30 min & enabler on	ADF_Time_30_Enabler	Integer
	Number of ADF activations with operation time under 40 min & enabler on	ADF_Time_40_Enabler	Integer

Information	Category	Abbreviation	Unit
	Number of ADF activations with operation time under 50 min & enabler on	ADF_Time_50_Enabler	Integer
	Number of ADF activations with operation time under 60 min & enabler on	ADF_Time_60_Enabler	Integer
	Number of ADF activations with operation time under 90 min & enabler on	ADF_Time_90_Enabler	Integer
	Number of ADF activations with operation time under 120 min & enabler on	ADF_Time_120_Enabler	Integer
	Number of ADF activations with operation time over 120 min & enabler on	ADF_Time_120plus_Enabler	Integer
Speed Histogram	Overall	v_hist_X	bins 0:5:250
	ADF on	v_hist_X_ADF	bins 0:5:250
	ADF & enabler on	v_hist_X_ADF_Enabler	bins 0:5:250
Travelled distance on roads with X number of lanes	1 lane	Lane_1	km
	2 lanes	Lane_2	km
	3 lanes	Lane_3	km
	4 lanes	Lane_4	km
	>4 lanes	Lane_4_plus	km
	1 lane & ADF on	Lane_1_ADF	km
	2 lanes & ADF on	Lane_2_ADF	km

Information	Category	Abbreviation	Unit
	3 lanes & ADF on	Lane_3_ADF	km
	4 lanes & ADF on	Lane_4_ADF	km
	>4 lanes & ADF on	Lane_4_plus_ADF	km
	1 lane & ADF on & enabler on	Lane_1_ADF_Enabler	km
	2 lanes & ADF on & enabler on	Lane_2_ADF_Enabler	km
	3 lanes & ADF on & enabler on	Lane_3_ADF_Enabler	km
	4 lanes & ADF on & enabler on	Lane_4_ADF_Enabler	km
	>4 lanes & ADF on & enabler on	Lane_4_plus_ADF_Enabler	km
	Adjacent sidewalk and/or cyclist lane	Sidewalk	km
	Adjacent sidewalk and/or cyclist lane & ADF on	Sidewalk_ADF	km
	Adjacent sidewalk and/or cyclist lane & ADF on & enabler on	Sidewalk_ADF_Enabler	km
Number of encountered crossing traffic sections	Intersections without signs (Overall)	Intersection_no_signs	Integer
	Intersections with static signs (Overall)	Intersection_signs	Integer
	Traffic light-controlled intersections (Overall)	Intersection_traffic_light	Integer
	On-ramps (Overall)	On_ramp	Integer
	Off-ramps (Overall)	Off_ramp	Integer
	Passing on ramps (Overall)	On_ramp_passing	Integer
	Construction zones (Overall)	Construction	Integer

Information	Category	Abbreviation	Unit
	Roundabouts (Overall)	Roundabout	Integer
	Pedestrian crossings (Overall)	Pedestrian_crossing	Integer
	Traffic light	Traffic_Light	Integer
	Tunnel	Tunnel	Integer
	Intersection without signs & ADF on	Intersection_no_signs_ADF	Integer
	Intersection with static signs & ADF on	Intersection_signs_ADF	Integer
	Traffic light-controlled intersection & ADF on	Intersection_traffic_light_ADF	Integer
	On-Ramp & ADF on	On_ramp_ADF	Integer
	Off-ramp & ADF on	Off_ramp_ADF	Integer
	Passing on ramp & ADF on	On_ramp_passing_ADF	Integer
	Construction zone & ADF on	Construction_ADF	Integer
	Roundabout & ADF on	Roundabout_ADF	Integer
	Pedestrian crossing & ADF on	Pedestrian_crossing_ADF	Integer
	Traffic light & ADF on	Traffic_Light _ADF	Integer
	Tunnel & ADF on	Tunnel_ADF	Integer
	Intersection without signs & ADF on & enabler on	Intersection_no_signs_ADF_Enabler	Integer
	Intersection with static signs & ADF on & enabler on	Intersection_signs_ADF_Enabler	Integer

Information	Category	Abbreviation	Unit
	Traffic light-controlled intersection & ADF on & enabler on	Intersection_traffic_light_ADF_Enabler	Integer
	On-Ramp & ADF on & enabler on	On_ramp_ADF_Enabler	Integer
	Off-ramp & ADF on & enabler on	Off_ramp_ADF_Enabler	Integer
	Passing on ramp & ADF on & enabler on	On_ramp_passing_ADF_Enabler	Integer
	Construction zone & ADF on & enabler on	Construction_ADF_Enabler	Integer
	Roundabout & ADF on & enabler on	Roundabout_ADF_Enabler	Integer
	Pedestrian crossing & ADF on & enabler on	Pedestrian_crossing_ADF_Enabler	Integer
	Traffic light & ADF on & enabler on	Traffic_Light _ADF_Enabler	Integer
	Tunnel & ADF on & enabler on	Tunnel_ADF_Enabler	Integer
Travelled distance on roads with speed limit	<25 km/h	Distance_25_less	km
	30 km/h	Distance_25_35	km
	40 km/h	Distance_35_45	km
	50 km/h	Distance_45_55	km
	60 km/h	Distance_55_65	km
	70 km/h	Distance_65_75	km
	80 km/h	Distance_75_85	km
	90 km/h	Distance_85_95	km
	100 km/h	Distance_95_105	km

Information	Category	Abbreviation	Unit
	110 km/h	Distance_105_115	km
	120 km/h	Distance_115_125	km
	130 km/h	Distance_125_135	km
	>135 km/h	Distance_135plus	km
	<25 km/h & ADF on	Distance_25_less_ADF	km
	30 km/h & ADF on	Distance_25_35_ADF	km
	40 km/h & ADF on	Distance_35_45_ADF	km
	50 km/h & ADF on	Distance_45_55_ADF	km
	60 km/h & ADF on	Distance_55_65_ADF	km
	70 km/h & ADF on	Distance_65_75_ADF	km
	80 km/h & ADF on	Distance_75_85_ADF	km
	90 km/h & ADF on	Distance_85_95_ADF	km
	100 km/h & ADF on	Distance_95_105_ADF	km
	110 km/h & ADF on	Distance_105_115_ADF	km
	120 km/h & ADF on	Distance_115_125_ADF	km
	130 km/h & ADF on	Distance_125_135_ADF	km
	>135 km/h & ADF on	Distance_135plus_ADF	km
	<30 km/h & ADF on & enabler on	Distance_25_less_ADF_Enabler	km
	30 km/h & ADF on & enabler on	Distance_25_35_ADF_Enabler	km

Information	Category	Abbreviation	Unit
	40 km/h & ADF on & enabler on	Distance_35_45_ADF_Enabler	km
	50 km/h & ADF on & enabler on	Distance_45_55_ADF_Enabler	km
	60 km/h & ADF on & enabler on	Distance_55_65_ADF_Enabler	km
	70 km/h & ADF on & enabler on	Distance_65_75_ADF_Enabler	km
	80 km/h & ADF on & enabler on	Distance_75_85_ADF_Enabler	km
	90 km/h & ADF on & enabler on	Distance_85_95_ADF_Enabler	km
	100 km/h & ADF on & enabler on	Distance_95_105_ADF_Enabler	km
	110 km/h & ADF on & enabler on	Distance_105_115_ADF_Enabler	km
	120 km/h & ADF on & enabler on	Distance_115_125_ADF_Enabler	km
	130 km/h & ADF on & enabler on	Distance_125_135_ADF_Enabler	km
	>135 km/h & ADF on & enabler on	Distance_135plus_ADF_Enabler	km
Number of accidents during experiment		Number_accidents	Integer
Number of safety critical situations/incidents during experiment	(Subcategories might be added depending on the definition of safety critical situations)	Number_safety_critical	Integer

Annex 2 Hi-Drive contextual data

Table 6.2: Overview on Hi-Drive contextual data per data category and the level at which they should be reported.

			Level to be reported		Relevant for …			
Contextual Information	Unit	Category	Test run / Trip / Driver Level	Driving Scenario Level	Question-naire	Performance Indicator	Time Series	Aggregated Time Series
Country of the test	Categories	Name of country or multi-country (in case of e.g. online survey)	x	x	x	x	x	x
Test environment	Categories	Pilot on public road Test track Simulator Computer Simulation Online Survey Wizard of Oz Simulation (Hardware test) Other?	x	x	x	x	x	x
Reporting level	Categories	Trip Test scenario Driver / Participant Driving Scenario instance / Case Test (run)	x	x	x	x		

			Level to be reported		Relevant for …			
At what time of day did test /driving scenario start?	hh:mm		x	x	x	x	x	x
At what time of day did test /driving scenario end?	hh:mm		x	x	x	x	x	x
On which date did test /driving scenario start?	yyyy/mm/dd		x	x	x	x	x	x
On which date did test /driving scenario end?	yyyy/mm/dd		x	x	x	x	x	x
Purpose of experiment is	Yes/No (per item)	Effects (ADF) Assessment (SP5) User Assessment (SP6) Enabler Assessment (SP2) Safety Impact assessment (SP7) Database (SP7) Outreach (SP7)	x	x	x	x		

			Level to be reported		Relevant for …			
Road type (if applicable)	Categories	Motorway, Urban Motorway, Urban, Rural Roads, Transition from motorway to urban, Parking	x	x	x	x	x	x
Driving scenario	Yes/No (per item)	Entire trip or Scenario from list [Car Following, Approaching, Cut-in, Cut-out, Activation ADF, Deactivation ADF, Incident, Take-over scenario, MRM, Roundabout, Junction, Tunnel, Other, …]	x	x		x	x	x
Tested enabler(s)	Yes/No (per item)	[TBD according to SP2]	x	x	x	x	x	x
Weather conditions	Categories	1: sunny, 2: cloudy, 3: light rain, 4: rain, 5: heavy rain, 6: light snow, 7: snow, 8: heavy snow, 9: fog	x	x	x	x	x	x
Lighting condition	Categories	1: daylight, 2: dawn, 3: dusk, 4: night	x	x	x	x	x	x
ADF active	[Yes/No]	Yes/No	x	x	x	x	x	x
ADF type	[Yes/No]	Checklist [motorway, urban, …]	x	x	x	x		

			Level to be reported		Relevant for …			
Assisted Driving function (<= L2) active	[Yes/No]	Yes/No	x	x	x	x	x	x
Enablers active	Yes/No per item	Categories according to SP2	x	x	x	x	x	x
Enablers interacting with ADF	[Yes/No]	Yes/No	x	x	x	x	x	x
Video logging (front) present?	[Yes/No]	Yes/No	x	x		x	x	x
Video logging (driver) present?	[Yes/No]	Yes/No			x	x		
Gaze data logged?	[Yes/No]	Yes/No	x	x	x	x		
Total driving distance ego vehicle (exposure)	[m]		x	x	x	x	x	x
Total driving duration ego vehicle (exposure)	[s]		x	x	x	x	x	x
Duration of experiment	Categories	One-day session; multiple days	x		x	x		

			Level to be reported		Relevant for …			
Was there a practice drive? (yes/no)	[Yes/No]		x		x	x		
Setting of the practice drive (How was the practice?)	Categories	None Only manual AD included Enablers included	x		x	x		
Where was the practice drive conducted?	Categories	None Same route as the test Different route from the test	x		x	x		
Participant type	Categories	Professional driver Non-professional driver – company employee non-professional driver – external participant Online participant	x	x	x	x	x	x
What was the role of the participant?	Categories	Driver Operator (outside vehicle) Observer (in vehicle) Observer (outside vehicle)	x	x	x	x		
Where was the participant seated?	Categories	Driver seat Co-driver Rear seat Outside vehicle	x	x	x	x		

			Level to be reported		Relevant for ...			
Age (only naïve drivers)	Categories	Categories	x (only normal driver)	x	(x)	x		
Gender (only naïve drivers)	Categories	Male | Female | Diverse	x (only normal driver)	x	(x)	x		
How long did the participants experience AD?	[s]		x		x	x		
When were questionnaires done?	Yes/No per item	Categories (pre, in, post)	x		x	x		
Were participants allowed to engage in non-driving tasks?	Categories	Categories needed: 1) No non-driving activities, 2) Free choice of activities, 3) Specific activity	x		x	x		
Was there a Safety Driver? (yes/no)	[Yes/No]		x		x	x		
Where in the vehicle did the safety driver sit?	Categories	Driver's seat, passenger's seat, left rear, middle rear, right rear	x		x	x		
What was the role of the Safety Driver?	Categories	Categories: Handled exceptional situations only, Driver handling all ToR, WoOz driver	x		x	x		

			Level to be reported		Relevant for ...			
Was an experimenter present?	[yes /no]		x		x	x		
Where did the experimenter sit?	Categories	Driver seat, passenger seat, left rear, middle rear, right rear, simulator control room	x		x	x		
Did any unusual things (e.g., technical issues) occur during the trip/scenario?	[Yes/No]		x	x	x	x	x	x
Traffic Condition	Categories	Number of vehicles per lane and hour, Categories: jammed, high, medium, low	x	x		x	x	x
Number of lanes of road	Numeric			x		x	x	x
Driving duration in ODD (ego vehicle)	[s]		x	x	x	x	x	x
Driven distance in ODD (ego vehicle)	[m]		x	x	x	x	x	x
AEB (warning or intervention) active?	[Yes/No]		x	x	x	x	x	x

			Level to be reported		Relevant for ...			
Use Case	Categories	Categories according to SP3 definition	x	x	x	x	x	x
Test scenario	Categories		x	x	x	x	x	x
ID Test person	[-]	Unique and GDPR compliant	x		x	(x)		
Number of trials/runs per driver	[-]		x	x	x	x		