AUTOPILOT

Grant Agreement Number: 731993

Project acronym: AUTOPILOT

Project full title: AUTOmated driving Progressed by Internet Of Things

D.4.1

Methodology for Evaluation

Due delivery date: 31 January 2018

Actual delivery date: 30 January 2018

Organization name of lead participant for this deliverable: TNO

Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the GSA)	
RE	Restricted to a group specified by the consortium (including the GSA)	
CO	Confidential, only for members of the consortium (including the GSA)	

Project funded by the European Union's Horizon 2020 Research and Innovation Programme (2014 – 2020).

Document Control Sheet

Deliverable number:	D4.1
Deliverable responsible:	TNO
Work package:	4 - Evaluation
Editor:	Bart Netten

Author(s) – in alphabetical order		
Name	**Organisation**	**E-mail**
Aittoniemi, Elina	VTT	elina.aittoniemi@vtt.fi
Barnard, Yvonne	UNL	y.barnard@leeds.ac.uk
Federley, Maija	VTT	maija.federley@vtt.fi
Gaitanidou, Lila	CERTH	lgait@certh.gr
Karagiannis, Georgios	HUA	georgios.karagiannis@huawei.com
Kolarova, Viktoriya	DLR	viktoriya.kolarova@dlr.de
Lenz, Olivier	FIA	olenz@fia.com
Netten, Bart	TNO	bart.netten@tno.nl
Pont Rañé, Jordi	IDIADA	jordi.pont@idiada.com
Van den Boom, Benedikt	FIA	euprojects@fia.com
Willenbrock, Ralf	TSY	ralf.willenbrock@t-systems.com

Document Revision History			
Version	**Date**	**Modifications Introduced**	
		Modification Reason	**Modified by**
0.1	19-04-2017	First Draft D4.1 for TMT	All
0.2	18/10/2017	First Draft IR4.1	All
0.3	20/11/2017	IR4.1 version for GA	All
0.4	14/12/2017	update	All
0.5	12/01/2018	Version for peer review	All
0.6	22/01/2018	Update	All
1.0	31/01/2018	Version submitted to the EC	ERTICO

Abstract

This deliverable presents the methodologies for the evaluation of the piloted use cases; for technical evaluation and the assessments of user acceptance, quality of life and the business impact.

The FESTA methodology is applied and enhanced for evaluating the added value of the Internet-of-Things (IoT) to improve Cooperative and Automated Driving (AD). The main research question to be evaluated is defined as "What is the added value of IoT for AD?" This central question is refined for all four evaluation perspectives in more detailed research questions, hypotheses and key performance indicators, measurements and log data from the pilots, and in evaluation methods. The methodologies provide the starting point for implementation and execution in the evaluation tasks in the next preparation and piloting phases.

The evaluation methodologies are tailored for the scale and scope of the pilot sites and implementations of the use cases. The common research focus in the evaluation methodologies on the concepts and criteria that are most common among pilot sites and use cases maximises the synergy and coherence between the evaluation tasks. Potential improvements of IoT to accelerate,

enhance or enable automated driving functions and services will be evaluated and assessed collaboratively from all four perspectives. The methodologies will be extended for additional use case or pilot site specific evaluation criteria during the coming phases.

This deliverable provides guidelines, requests and requirements for pilot test scenarios and data provisioning that will be needed as input for evaluation. This is input for the specification and data management of the pilots.

Legal Disclaimer

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The above referenced consortium members shall have no liability to third parties for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law. © 2017 by AUTOPILOT Consortium.

Abbreviations and Acronyms

Acronym	Definition
AD	Automated Driving or Cooperative and Automated Driving
AVP	Automated Valet Parking
BI	Business Impact
EC	European Commission
FESTA	Field opErational teSt supporT Action
GA	Grant Agreement
IoT	Internet of Things
PO	Project officer
QoL	Quality of Life
TE	Technical Evaluation
UA	User Acceptance
VRU	Vulnerable Road User
WP	Work Package

1 Table of Contents

List of Figures

List of Tables

Executive Summary

This deliverable presents the methodologies for the evaluation of the piloted use cases for technical evaluation, user acceptance assessment, and assessments of the impacts on quality of life and business.

This deliverable concludes the work of Task 4.1 “Evaluation requirement and methodology” and provides the starting point for implementation and execution of the specific evaluation tasks; Task 4.2 to Task 4.5.

The FESTA methodology is applied and extended for using the Internet of Things (IoT) for Cooperative and Automated Driving (AD). For each of the evaluation methodologies, research questions, hypotheses, performance indicators and measures are defined. The methodologies are detailed per evaluation task in sections 0- �.

Input for the methodologies is provided by the project objectives and specifications of the use cases, pilot sites and technologies for AD functions and services, IoT, communication, security and privacy from work packages 1, 2 and 3.

Outputs of this deliverable are two living documents on Project Place (document management system of AUTOPILOT) of which the current versions are attached in Annex 14:

- Research questions, hypotheses and indicators for the four evaluation tasks.
- Requirements for log data and data quality to be provided from the pilots via the central data management server.

For each of the evaluation methodologies, the required inputs are defined:

- Section 0 summarizes the guidelines and requirements for pilot test scenarios and activities for users and stakeholders, as input to Task 3.1.
- Section 0 provides requirements for data provisioning via the central data management server, as input to Task 3.4.

The pilot sites, the automated driving functions and services, IoT platforms, devices and cloud services are still being developed and adapted. At times, the input is still high-level for the development of evaluation methodologies. In additional workshops, storyboard sessions and discussions with pilot sites and use case developers, the scope and focus of the evaluation methodologies have been refined. The approaches chosen at the end of this Task 4.1 are stated at the beginning of each evaluation task in sections 0 – �.

The common focus for evaluation is defined in a central research question:

“What is the added value of IoT for AD?”

that will be answered from the central hypotheses:

- IoT is *accelerating* the development and deployment of automated driving functions,
- IoT is *enhancing* the functionality or performance of automated driving functions,
- IoT is *enabling* new automated driving functions.

The evaluation methodologies are refining the research question and hypotheses for objectives and concepts that are most relevant and common to pilot sites and use cases, as defined in Table 1.

Table 1: Evaluation focus and objectives

Evaluation	Objectives and focus
Technical	Technical improvements of IoT in functionality and performance of automated driving modes, functions and services. Improvements are evaluated on: • positioning, localisation, manoeuvring and navigation • data communication and data management • environment detections • impact of IoT on safety • security and privacy requirements
User Acceptance	• Formulate IoT-related improvements for automated driving functions based on user feedback. • Determine, whether there are improvements or added value in automated driving functionalities with and without the assistance of the IoT regarding user acceptance.
Quality of Life	• Explore how IoT in automated driving meets personal mobility needs • Explore the improvements in transport system efficiency with various penetration rates of IoT devices and automated driving vehicles. • Explore the contribution of IoT to traffic safety improvements • Explore the contribution of AD and IoT to citizens' well-being
Business Impact	• Evaluate the cost benefit and cost-effectiveness of the AUTOPILOT exploitable results, i.e. the IoT accelerated, enhanced or enabled automated driving systems. • Evaluate the impact of exploitable results to the market in terms of creating new products and customers, and establishing a new stakeholder ecosystem.

The evaluation methodologies will be extended with pilot site or use case specific criteria later. In the next phase of pilot preparation, the pilot test scenarios will be refined in collaboration with Task 3.1, following the implementations and adaptations of automated vehicles, IoT devices, IoT platforms and cloud services. The data provisioning process and data requirements will be refined in collaboration with Task 3.4. The methodologies for evaluation, and the research questions, hypotheses, indicators and data requirements of Annex 14 will be refined and extended accordingly.

2 Introduction

2.1 Purpose of the document

The purpose of this deliverable is to define the evaluation methodologies that have been developed in Task 4.1 “Evaluation requirements and methodology”.

In the remainder of the AUTOPILOT project, the evaluation tasks will start, and the evaluation methodologies in this deliverable will be implemented, refined and executed. The next evaluation tasks to start are:

- Task 4.2 – Technical Evaluation
- Task 4.3 – Business Impact Assessment
- Task 4.4 – Quality of Life Impact Assessment
- Task 4.5 – User Acceptance Assessment

Task 4.6 – “Legal issues” will also start at the same time. Task 4.6 will not implement an evaluation methodology, but instead investigate any legal issues that arise from piloting and the other evaluation tasks.

Task 4.1 has implemented the FESTA methodology to develop the evaluation methodologies based on the input from work packages 1, 2 and 3; use cases, functional and technical specifications for data communication, IoT platforms and architectures, pilot descriptions and storyboards for piloting the use cases. An initial extension to FESTA is described for evaluating the added values of the Internet-of-Things (IoT) for Automated Driving (AD). The FESTA extensions will be developed further during AUTOPILOT evaluations and included in the final evaluation report.

The main objective is to develop the common focus for evaluation to ensure coherence between the evaluation methodologies. The evaluation artefacts are defined in spreadsheets on Project Place and are living documents that will be updated and extended throughout the project. The spreadsheets are attached as “14.1 Annex 1 – Research Questions, Hypothesis, and Indicators” and “14.2 Annex 2 – Data Requirements” to this report. The final versions will be the basis for the final evaluation report.

The main outputs to other tasks in AUTOPILOT are summarized in two sections:

- Section 0 with guidelines and requirements for pilot test scenarios and activities for users and stakeholders, as input to Task 3.1.
- Section 0 provides requirements for data provisioning via the central data management server, as input to Task 3.4.

2.2 Intended audience

This deliverable is mainly intended as a working document for internal use in AUTOPILOT, i.e. for partners:

- To implement and execute the evaluation tasks T4.2 – T4.5.
- To develop pilot test scenarios in Task 3.1.
- To develop the central data management server to provide input for evaluation in Task 3.4.
- To provide input to business exploitation in Task 5.3.

2.3 Terminology

User	Users are understood here in a broader definition as "*anyone who uses the AUTOPILOT functions and services*". This definition is congruent with the approach taken in the unpublished position paper by the CARTRE thematic interest group [7].
Other road users	Road users that are indirectly affected by the use of the AUTOPILOT technology (i.e. in the single use cases), e.g. cyclist, pedestrian, drivers of conventional vehicles; this group can be also interpreted as a part of the stakeholder groups
Measurand	Parameter or property intended to be measured in a unit.
Measurement	Operation to determine the value or quantity of a measurand at a given time
Position	Absolute position of an object in WGS'84 or GPS coordinates in latitude, longitude, and optionally with an altitude.
Location	Relative position of an object on the road defined by lane number, lateral road or lane offset, and optionally with a map matched position with a longitudinal offset to a road reference point, or road identifier.

2.4 Structure of the report

Section 0 gives an overview of the FESTA methodology and how this is implemented in AUTOPILOT to develop the methodology for evaluation described in the following sections. The FESTA methodology will be extended for piloting IoT and Automated Driving functions, and section 0 gives an outlook for the next two project years.

Section 0 describes the common approach to evaluation, including the common and shared research questions and hypotheses.

Sections 0 to ⍰ present how the common approach is developed into the methodologies for each of the evaluation tasks. The comprehensive list of research questions, hypotheses, indicators, data measures and quality criteria are collected in two spreadsheets in Annex 14.

Section 0 summarizes guidelines and requirements from sections 0 to ⍰ for defining pilot test scenarios and activities for users and stakeholders, as input to Task 3.1.

Section 0 provides requirements for data provisioning via the central data management server, as input to Task 3.4.

3 FESTA Methodology

The original FESTA handbook was produced in 2008 by the FESTA consortium (Field opErational teSt supporT Action, 2007–2008). The FOT-Net and FOT-Net 2 consortia updated this handbook several times in order to take into account the lessons-learned from the many FOTs that have been conducted since, and the insights and ideas shared between experts in workshops, international workshops, seminars and stakeholder meetings. The latest version, version 6 [1], was produced end 2016 by the FOT-Net Data consortium. The handbook, and other information ([2][3][4][5]), are available at www.fot-net.eu.

The FESTA methodology is summarised below and in Figure 1. There are several steps, which although described in a linear way, are performed in iteration. The V-shape shows the dependencies between the different steps in the left- and right-hand side of the V. The steps can be summarised as:

- Defining the study : Defining functions, use cases, research questions and hypotheses
- Preparing the study: Determining performance indicators, study design, measures and sensors, and recruiting participants
- Conducting the study: Collecting data
- Analysing the data: Storing and processing the data, analysing the data, testing hypotheses, answering research questions
- Determining the impact: Impact assessment and deployment scenarios, socio-economic cost benefits analysis

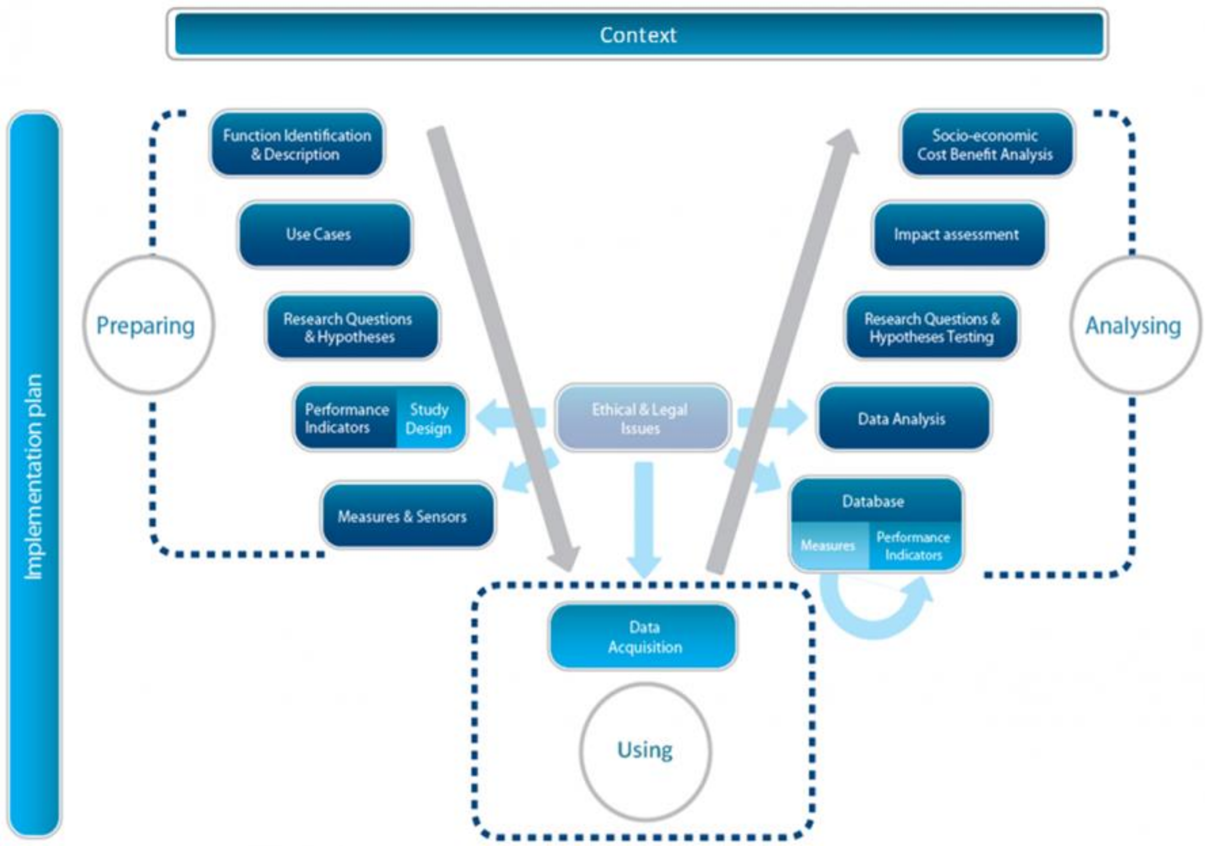

Figure 1: FESTA-V

There are several steps that are of importance for setting up the evaluation framework for AUTOPILOT. Only the left-hand site of the FESTA V is described here ("preparing"), the next step will be the set-up of the data analysis and impact assessment.

3.1 Project Objectives and Context

The horizontal bar on top of the diagram summarises the context in which the pilot site tests are supposed to take place. This defines the objectives at a high level which the different pilot sites want to reach. In AUTOPILOT we will be looking at both the specific contexts, like for example urban driving, but also the overall goal of AUTOPILOT to prove the Internet-of-Things approach to automated driving. A good description of the context may clarify the way in which we will keep the objectives into focus, both at project and pilot site levels.

3.2 Function Descriptions and Use Cases

This concerns the specific forms of automation the pilot sites are going to test, and the situations in which the automated vehicles will operate.

The process to go from functions to hypotheses is as follows:

1. Selecting the functions to be tested
2. Defining the connected use cases to test these functions
3. Identifying the research questions related to these use cases
4. Formulating the hypotheses associated with these research questions, and
5. Linking these hypotheses to the corresponding performance indicators.

The use cases normally have the form of a textual description, explaining how the automation will work and in which circumstances. The use cases are further specified with situations and scenarios. For example, it may be a description of how Automated Valet Parking will work (driver leaves car at parking sport, car gets information about available parking space, drives toward the nearest empty space etc. etc.)

The use case description will be generated by the pilot sites, in interaction with WP4.1 members. In FESTA the focus is on specific functions, for example lane departure warning, but in AUTOPILOT the automated vehicles are a system with multiple functions, which are supposed to work in a wide range of situations, so a complete use case description may not be feasible. Probably some prototypical uses cases have to be selected and described.

3.3 Research Questions and Hypotheses

This is one of the most important steps, as the research questions will drive the testing. In the FESTA approach, the research questions are grouped under five impact areas: Efficiency, Environment, Mobility, Safety and User Uptake. For automation, impact areas may be even wider such as security, health, land use etc.

In the experience of FOT-Net the formulation of research questions is an elaborate and iterative process, taking both a top-down approach (start with impact areas) and bottom-up (start with use-cases).

From research questions hypotheses can be formulated. The definition of a hypothesis [1] is:

“A specific statement linking a cause to an effect and based on a mechanism linking the two. It is applied to one or more functions and can be tested with statistical means by analysing specific performance indicators in specific scenarios. A hypothesis is expected to predict the direction of the expected change.”

Usually a large number of research questions and hypotheses is generated during a workshop. The most difficult part is selecting a limited set of research questions and hypotheses. As automated vehicles have many functions it remains to be seen how easy it is to formulate and select hypotheses at a detailed level.

3.4 Performance Indicators

Performance indicators are quantitative or qualitative indicators, derived from one or several measures, agreed on beforehand, expressed as a percentage, index, rate or other value, which is monitored at regular or irregular intervals and can be compared to one or more criteria. During the process of developing hypotheses, it is important to choose appropriate performance indicators that will allow answering the hypotheses, but that will also be obtainable within the budget and other limitations of the project. Performance indicators are based on measures. FESTA distinguishes four types: Direct Measures, Indirect Measures, Self-Reported Measures, and Situational Variables.

FESTA provides a Performance Indicators-Measures-Sensors matrix, see [6].

3.5 Study Design

The study design describes the experimental design, the participants, the environment, and piloting procedures.

3.6 Measures and Sensors

On basis of the previous steps, it can be determined what needs to be measured and how, e.g. collect background data, logging data from sensors and application software, and questionnaires. In addition, data in the form of manually or automatically transcribed data and reductions of collected data is also considered sensor acquired data (but with a manual sensor—the analyst). In FESTA, all the data sources mentioned above are considered sensors. Subsequently all data can be acquired, stored, and processed in a generalised way.

3.7 Ethical and Legal Issues

These aspects have to be worked out for each pilot site, as regulation and approval procedures may vary amongst countries. However, FESTA and related documents from FOT-Net provide support.

3.8 The FOT Implementation Plan (FOTIP)

The FOTIP serves as a checklist for planning and running FOTs.

3.9 Data Sharing

If we want to be able to share data between organisations in the consortium, but also to be able to continue analyses after the project, also by other parties, data sharing has to be taken into consideration from the very start. The ‘FOT-Net Data’ Data Sharing Framework provides elaborated guidelines (fot-net.eu/Documents/data-sharing-framework).

4 Evaluation Approach in AUTOPILOT

This section presents the overall approach to evaluation in AUTOPILOT. It is based on the project objectives and is the common starting point for defining the methodologies for all evaluation sub tasks in the next sections.

4.1 What is the added value of IoT for Automated Driving?

The objectives of the AUTOPILOT project are to define and implement an IoT architecture for Automated Driving (AD), and to realize IoT-based AD use cases. The main research question to answer in the evaluations of the pilots is "What is the added value of IoT for Automated Driving in the piloted Use Cases?" The main hypotheses to test, qualify and quantify the added value are:

- IoT is *accelerating* the development and deployment of automated driving functions.
- IoT is *enhancing* the functionality or performance of automated driving functions.
- IoT is *enabling* new automated driving functions.

Potentially IoT devices can provide information on other vehicles, emergency and heavy good vehicles, stationary and illegally parked vehicles, etc.. IoT devices may also provide information on vulnerable road users such as pedestrians, bicyclists and motorbikes, or wheel chairs. A vehicle's host sensors and ITS-G5 communication can also provide similar information within the range of the sensors or communication. 'Similar' is interpreted as information of similar type, contents and quality. IoT can accelerate for example with a cheaper solution, by increasing the penetration rate of probed devices, or extending the 'range of view' for *similar* information.

If the quality or contents of IoT data is better than existing data, then the AD functionality can be enhanced, and performance can be improved. IoT data may provide more information directly from other road users or obstacles for example, or may provide more accurate positioning, localisation or navigation information.

Whether IoT or IoT data is accelerating or enhancing AD may not always be clear to distinguish a priori. It depends on the existing equipment and infrastructure of use case implementations, which may differ between pilot sites for example. The evaluations should test and classify this later. Fortunately, similar test scenarios can be defined for both types of hypotheses; with a baseline scenario for the existing situation without IoT data, and comparative evaluations of test scenarios with IoT data.

The third type of hypotheses requires different test scenarios as the pilot system can only be tested with IoT data source. Hence the added value of IoT can be assessed on feasibility for example. A baseline scenario without IoT would not be meaningful or executable, and a comparative evaluation against a 'without IoT' baseline is not possible.

The hypotheses on the added value of IoT and AD will be tested from the following evaluation tasks:

- Technical evaluation of the improvements in functionality and performance of automated driving due to IoT. Also, the effects of IoT are evaluated that may impact the safety, security and privacy of AD.
- User acceptance evaluation to assess IoT based improvements for automated driving modes and services, and recommendations to accelerate or enhance their acceptance.
- Impact of IoT on automated driving and its impact on the quality of life, including traffic safety and efficiency, personal mobility, environmental effects and well-being.

- Impact of IoT on the business for automated driving for example by decreasing the time-to-market, making it more affordable, and identifying new markets and stakeholder ecosystems.

Table 2 gives some examples of IoT data sources for use cases and their a priori hypothesised added value for automated driving. Note that the hypotheses from D1.1 are filtered:

- An IoT data source is considered accelerating if similar data is already available to cooperative or automated vehicles.
- An IoT data source is considered enhancing if the functionality or performance of automated driving modes or services can be enhanced.
- An IoT data source is considered enabling if the use case cannot be demonstrated at all without the IoT data, and hence a baseline test without IoT data sources is not possible.

Table 2: Examples of the added value of IoT data per use case

<table>
<tr><th>Use Case</th><th>Pilot site</th><th>Accelerating</th><th>Enhancing</th><th>Enabling</th></tr>
<tr><td rowspan="3">Automated Valet Parking</td><td>Vigo</td><td rowspan="3"></td><td rowspan="3">obstacle detection, reduced parking time and (re)routing, available parking space assignment</td><td>positioning and (re)routing for indoor navigation</td></tr>
<tr><td>Finland</td><td rowspan="2"></td></tr>
<tr><td>Brainport</td></tr>
<tr><td rowspan="4">Urban Driving</td><td>Finland</td><td rowspan="4">smoother speed adaptations from traffic and traffic light status info</td><td rowspan="3">Vulnerable Road User (VRU) detection</td><td rowspan="4"></td></tr>
<tr><td>Versailles</td></tr>
<tr><td>Livorno Port</td></tr>
<tr><td>Brainport TU/e campus</td><td>+ lecture schedules, weather</td></tr>
<tr><td rowspan="2">Highway Pilot</td><td>Livorno highway</td><td colspan="2">more (accurate) detection and warning of pot holes, puddles, road works</td><td></td></tr>
<tr><td>Brainport highway</td><td colspan="2">+ rocks, bumps, broken down vehicles, foreign objects</td><td></td></tr>
<tr><td rowspan="2">Platooning</td><td>Versailles</td><td rowspan="2">smoother speed adaptations and lane selection from traffic info</td><td>+ reduce distance between vehicles</td><td>+ organisation</td></tr>
<tr><td>Brainport highway</td><td>improved positioning and localisation</td><td>platoon planning</td></tr>
<tr><td rowspan="2">Car Sharing</td><td>Versailles</td><td rowspan="2">Better pick-up/drop-off (including vehicle availability)</td><td rowspan="2">free parking space monitoring + better fleet management</td><td rowspan="2"></td></tr>
<tr><td>Brainport TU/e campus</td></tr>
</table>

The IoT devices are enhancing the Automated Valet Parking in the Finland and Brainport sites, because an optimal route to an available parking place is provided a priori to the automated vehicle and re-routing is provided to avoid any obstacles elsewhere on the parking area thereby potentially reducing parking times. In a baseline test scenario without IoT data, the automated vehicles would still be able to navigate and avoid obstructions on the parking area. In the indoor parking site of Vigo however vehicle positioning and navigation would not be feasible without IoT information.

IoT data is accelerating or enhancing urban driving use cases. Traffic light status information can already be provided to connected vehicles, which could be accelerated through IoT. IoT also enhances the automated driving functionality by providing additional information on detected vehicles, weather or schedules of lectures on the TU/e for example.

In the highway pilot use cases, IoT data is used to accelerate or enhance the detection of, and warning for, road hazards. Baseline scenarios can be defined for manual driving, in which drivers avoid road hazards, and road side detectors or cooperative or connected vehicles to detect evasive manoeuvres, adapt on-board maps and exchange hazard warnings.

In the platooning use case, vehicles are automatically following a leader that may be driven automatically or manually. IoT data enables the optimisation of platoon planning, including discovery of platoon members, platoon formation and management. Additional IoT data is used to improve vehicle positioning (RTK-GPS) and localisation accuracy (HD-maps). IoT data is also sourced to smooth automated driving behaviour for traffic conditions, merging traffic or traffic control information on the use of extra lanes or priority and green light at intersections.

In the Car Sharing use cases, car sharing users can use IoT cloud services to better match car sharing services and optimise route planning based on traffic information. Additionally, car sensor data may be used to monitor free parking spaces during piloting.

4.2 Common research questions and hypotheses

The main hypotheses on the added value of IoT to accelerate, enhance or enable AD are largely defined from a technical perspective. From the other perspectives, the distinction between the IoT and AD component and their contribution to the added value may not be obvious. Users, for example, may be subject of evaluation to compare systems without IoT to systems with IoT. However, users may not be able to distinguish any difference when IoT is accelerating AD, or able to compare newly enabled functionalities. For quality of life and business impact assessments, the integrated piloted ‘product’ or ‘service’ is subject of evaluation rather than the underlying technology per se.

The added value of IoT also depends on the existing reference implementations of use cases at pilot sites, and integration of use cases and existing infrastructure at pilot sites. Hence, variations need to be made in the refinements of the main research question and hypotheses per evaluation task, as well as use cases and pilot sites.

The objective for evaluation is to answer the research questions that are common to most use cases. Hence, evaluation seeks commonality in research objectives, hypotheses and evaluation methodologies where possible across pilots, use cases, or the underlying automated driving modes, functions and services. The approach for evaluation is to refine the main research question and hypotheses in three levels of priority:

1. Common to all or most use cases, pilot sites and evaluation tasks.
2. Specific to a use case or pilot site, and most relevant to an evaluation task.

Following subsections identify the most common research questions. These, and the more specific research questions, will be described for the respective evaluation tasks in the following sections.

4.2.1 Environment detections

The important added value of IoT to accelerate or enhance AD in all use cases, and piloted on all sites, is the use of IoT cloud services and IoT device data to acquire or enhance information from the environment, such as obstacles and road hazards, other road users, traffic information and environmental conditions.

From a technical perspective the environmental data may enhance or enable environmental detections for example for VRU or pothole detection, traffic control and status. Many hypotheses can be defined on the improved functionality and performance of the functions and services in the vehicles that are expected to yield improvements in automated driving such as smoother driving, earlier adaptation to traffic situations and more efficient routing. Technical evaluation also considers research questions and hypotheses on technical criteria that determine the feasibility or identify issues in the achieved improvements, such as data management, data communication, safety, security and privacy.

When the AD functionality and performance is enhanced, the user may be able to observe the improvements, and user acceptance can be evaluated for example on factors like the perceived usefulness, ease of use, comfort, trust, safety and security. Can the perceptions be matched with the technical improvements, and how can user perceptions be fed back to improve the technology, implementations and pilot test scenarios?

Improved environmental information should improve the quality of life in terms of traffic safety, transport system efficiency, and mobility, on both a personal and societal level. The effects on a personal level, i.e. the piloted scenarios for the users and test drivers, are also evaluated from a technical and user acceptance perspective. Quantitative results from these evaluations can also be used for the assessment of the societal benefits.

Technical improvements not only affect user acceptance and impact the quality of life, it should also enable new products, markets, stakeholder ecosystems, all of which are subject of business impact assessment.

4.2.2 Positioning, localisation, manoeuvring and navigation

An important added value of IoT to accelerate, enhance or enable AD is the use of IoT cloud and data services to improve the accuracy of positioning, localisation, navigation or routing.

From a technical perspective, the performance using existing vehicle sensors and maps can be compared with the performance while using for example for RTK-GPS, HD maps, parking spot information or routes to available parking spots received from IoT cloud services and data sources. The general hypotheses are that IoT enabled position and localisation should improve the smoothness of driving, manoeuvring and lateral behaviour, while navigation and routing should be more efficient and avoid more obstacles and delays. The performance of in-door positioning and navigation enabled by IoT for Automated Valet Parking in Vigo will also be evaluated. Underlying technical functionalities and services are also evaluated, for example the data management and communication to find and retrieve IoT data.

These technical improvements should also result in improvements for evaluation of user acceptance, quality of life and business impacts in a similar manner as described for the environmental detections above.

4.2.3 Communication

Communication functionality is a general necessity for all use cases, both the ad-hoc or V2X communication and the centralised network communication to for example IoT platforms and devices [D1.7]. The added value for AD may come primarily from optimising communication performance and business impact such as costs and markets, which may also affect user acceptance and adoption.

4.3 Iterative Approach to Evaluation

4.3.1 Commonality in use cases and pilots

Five use cases are implemented, tested and demonstrated on six AUTOPILOT sites. Not all use cases are implemented on all pilot sites. More importantly, the implementations of the use cases at various pilot sites will be different, for example by using different types of IoT devices and platforms, different communication technologies, different automated vehicles, different public road and traffic environments, and different combinations of services. Consequently, use cases will be piloted in different test scenarios and with different impacts.

The challenge is to identify the commonalties in the implementation and usage of IoT as the basis for the methodologies for evaluation. The commonalties have been identified and discussed in several workshops with evaluators and use case developers from all pilot sites, amongst others based on the use case descriptions and storyboards (see [Storyboards] on Project Place). A set of common research questions and hypothesis, as described above and in following sections, has been selected by the evaluation team as the basis for evaluation. It should be emphasised that this will be refined and adapted as needed in close cooperation with other work packages in AUTOPILOT during developments and piloting. Updates will be documented in Annex 14.1 and the final deliverable on Evaluations at the end of the project.

4.3.2 Scale of pilots

Most likely, the pilot sites have only capacity for a limited number of automated vehicles. For legal and sometimes practical constraints, most of the AUTOPILOT applications and services will be tested not by "regular" (naïve) users, but by trained drivers and company employees, thus preventing both a large sample size for each use case per pilot site and a representative sample. Where possible, naïve users and other stakeholders will also be invited to participate in the tests, either as drivers accompanied by a test driver, or as passengers. The possible scale of the pilot tests, number of test runs, and participation of users will be detailed in the next phase, as part of the pilot test scenarios. The following sections will present requests and requirements from various evaluation perspectives, and a short list for pilot test specifications is provided in section 0.

4.3.3 Iterative approach to piloting and evaluation

Evaluation will follow the iterative approach to pilot testing and demonstrations. Figure 2 schematically shows the iterative approach in which every pilot test activity is supported by technical data analyses and evaluations.

The first phase is the preparation phase for pilots before actual piloting starts. This includes the pilot readiness verifications and validations of pilot site adaptations. These activities will also be used by the evaluation team to test and verify the tooling for technical evaluations, and also to give first feedback to the pilot sites, for example on data provisioning and pilot test scenarios. This phase is concluded with a report on first findings from technical evaluations (D4.2).

A few demo events are foreseen in conjunction with pilot test sessions during which workshops can be organised for discussions, (de)briefings and interviews with test drivers, regular end-users and other stakeholders. This is important to collect subjective feedback for user acceptance, the impact on the quality of life and business, and to collect potential legal issues.

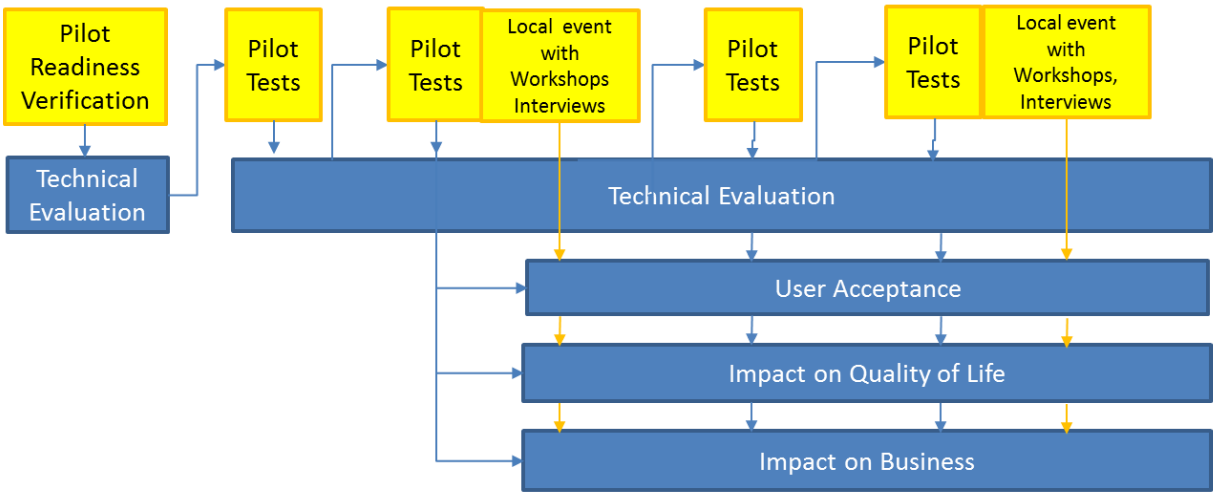

Figure 2: Iterative approach to piloting and evaluation

4.4 Managing and sharing evaluation data and results

The objective in Task 3.4 is that all pilot site data for evaluation is centrally managed, stored and accessible to all partners in AUTOPILOT. The objective for evaluation is to access all pilot site data that is needed for evaluations in a common format from this central system. This includes the logged data from test vehicles, IoT platforms, cloud services and additional IoT devices, situational data and the collected surveys and questionnaires (Figure 3). This includes the input from all pilot sites and use cases and for all test scenarios and test runs. Consequently, the input for evaluation should not be accessed directly from pilot site data management systems or IoT platforms or cloud services, and that no conversions are needed from their proprietary log formats into common data formats.

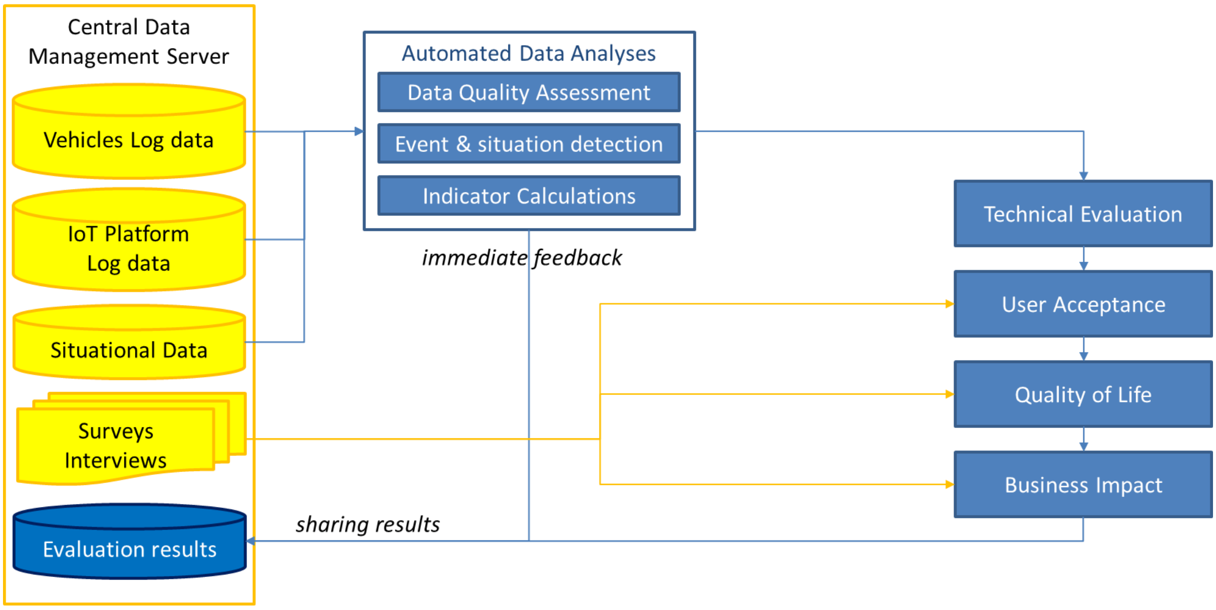

Figure 3: Managing and sharing evaluation data and results

The initial set of requirements for input data from the central data management server, including log data parameters, data quality criteria and data provisioning, is defined in section 0 and Annex 14.2. All data analysis tools for evaluation will be developed or adapted to use the common data formats.

These data requirements define for example the input for analysis of the data for a completed pilot test runs or session. The analysis includes the assessment and validation of the quality of the collected data, detection of relevant events and situations during the pilot test, and the calculation of predefined indicators. These analysis steps can be executed automatically after provisioning of all data to provide immediate feedback on the successfulness of the test run or session to both the pilot sites and evaluation tasks.

The central data management system can also be used to share intermediate and final evaluation results with other data analysts, evaluators, pilot sites and other partners within AUTOPILOT. This assumes that evaluators can also write data and evaluation results back to the central data management system. Formats and specifications for writing and accessing evaluation results are not included in this deliverable and will be defined later in the project.

5 Technical Evaluation Methodology

This section presents the methodology for technical evaluation that will be conducted in Task 4.2. The first sections define the objectives and approach to develop the technical evaluation methodologies. The second part presents the implementation of the methodology by means of examples of the research questions, hypotheses, performance indicators and metrics by means of examples. The full specification is provided in Annex 14. The third part describes a first approach to use case specific evaluations of events and situations expected in the pilot test scenarios. The fourth section defines the data that needs to be collected from the pilot sites. Finally, the principal differences between the pilot readiness verification (Task 2.5) and this Technical Evaluation is clarified.

5.1 Objectives and methodological approach

The main objective is to evaluate how IoT could offer potential improvements to automated driving. The potential improvement is measured by the improvement in the technical functionality and performance of automated driving functions, driving modes and the connected and automated driving services. The following subsection defines the concepts for measuring the improvements.

The potential improvements may be affected or restricted by safety, security and privacy requirements, hence their impact must be evaluated:

- IoT may impact automated driving safety, which is a hard pre-condition for automated driving. The impact of IoT on automated driving modes, functions and services should also be measured and evaluated.
- Security and privacy requirements may impact the usage of IoT. The objective is to measure the security and privacy level achieved by the implementation and usage of IoT.

The outcome for the main objective triggers secondary objectives to provide the technical evaluation results to other tasks and work packages (cf. Figure 3):

- The test runs, events and situations in which the technical systems functioned and performed successfully are the relevant input to the user acceptance, quality of life, and business impact assessments.
- Failures and anomalies where the IoT or AD are not successful provide events and situations as input for improvements for the systems, infrastructure or test scenarios in following iterations of piloting.

The technical evaluation methodology for the main objective follows the methodology described in section 0. Research questions, hypotheses and indicators are defined in the next subsections to measure and evaluate the technical improvements from the data logged on vehicles, IoT platforms and cloud services.

The secondary objectives will be implemented in the evaluation tools that will be developed later in Task 4.2, through the central data management tools (Task 3.4), and in workshops and meeting with partners from other work packages. To support the iterative approach to piloting, as described in section 4.3.3, the data analyses and technical evaluations will be automated as much as possible (Figure 3). This includes the automatic detection of situations and events from logged and situational data. These event and situation detections should be defined and implemented beforehand. Predefined indicators will also be calculated for these detected events and situations. The automatic data analysis will act as a filter for all the data received from the Pilot Sites (Figure 4). The results are provided as objective input for the evaluation tasks, including more detailed technical evaluations.

The main purpose of this approach is to minimise redundant work and inconsistencies between evaluations for example due filtering of the data on events and situations and filter with different criteria.

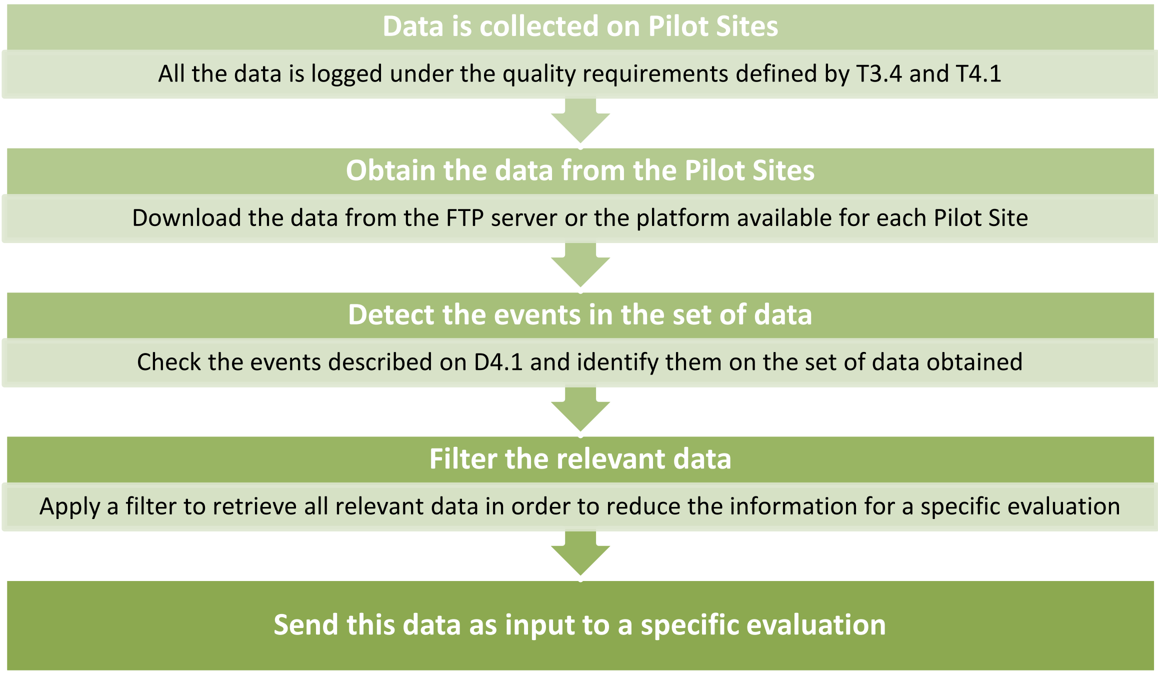

Figure 4: Process for filtering events and situations for evaluation

5.2 Research Questions, Hypotheses and Indicators

The main research question for the added value of IoT for AD is refined in the main concepts for potential improvements of technical functionality and performance. The concepts are defined in the form of first level research questions in Annex 14.1. Each of the next subsections describes a concept. Annex 14.1 also defines sub-level research questions, hypotheses, indicators and measurements. The measurements refer to Annex 14.2 that defines the sources, parameters and data quality criteria for log data (see also section 5.4).

It should be noted that the intention is not to be comprehensive in the research and evaluation of automated driving, or safety, or IoT, or IoT platforms and devices, or security and privacy. After all, AUTOPILOT is not developing or evaluating commercial products. Instead the concepts for research questions are chosen to provide the evidence for the concepts where IoT could provide the *largest* potential for added value to automated driving.

5.2.1 Positioning, localisation, manoeuvring and navigation

The IoT cloud services and data sources identified in section 4.2.2 are essential technical measures for improvement of the internal state, perception systems, motion planning and routing within automated vehicle functions and services. Technical improvements are highly relevant for all automated vehicles and use cases. Examples for improvements are:

- RTK-GPS for accurate positioning with reference signals provided via an IoT platform
- The use of HD-maps, in combination with on-board camera’s, to improve localisation of the relative position on the lane or road

- In-door routing and navigation using IoT devices
- Optimised routes to navigate to an available parking spot using IoT services.

The general hypothesis is that the added value of IoT platform and cloud services should improve perception systems and the accuracy, reliability and geographic areas of positioning, localisation, etc. Performance indicators are defined for accuracy and reliability.

Situations are distinguished by pilot site location, i.e. geographic areas that affect the performance, for example for indoor navigation in Vigo, GPS accuracy in Finland, RTK-GPS services in Brainport and vulnerable road user detection in Versailles and Livorno.

This research question is refined at a second and third level, and in hypotheses to differentiate the specific fusion with IoT data and plausible failure modes.

The research questions related to the Global Positioning System and the Inertial Navigation system, including the positioning data, the data related to the navigation systems and the localisation of the vehicle respect to the other elements of the road. The range and the accuracy with timing references and also the changes with the on-board maps with the IoT will be evaluated. For accurate (lane level) positioning evaluation, measurements are expected in a standard format like NMEA. The NMEA format is by sentences, each one containing a type of information. The essentials are the GGA, RMC, GSA, VTG and ZDA:

- GSA sentence. GPS DOP and active satellites. This sentence provides details on the nature of the fix. It includes the numbers of the satellites being used in the current solution and the DOP. DOP (dilution of precision) is an indication of the effect of the satellite geometry on the accuracy of the fix.
- RMC sentence. NMEA has its own version of essential GPS PVT (position, velocity, time) data. It is called RMC, the Recommended Minimum.
- VTG sentence. Velocity made good. The GPS receiver may use the LC prefix instead of GP if it is emulating Loran output.
- ZDA sentence. Data and time.

Localisation is evaluated on accuracy for determining the relative position on the road; i.e. the longitudinal and lateral position on a road and in a lane. For evaluation of a single localisation system in a vehicle, and external or alternative localisation system is required. Alternatively, an accurate positioning system, such as an RTK GPS signal or road side detectors may be used.

Manoeuvring of automated vehicles, such as following and platooning, lane keeping and lane changing, turning, and parking is evaluated on the accuracy and smoothness of the manoeuvres. Indicators are for example:

- Speed. Speed of the vehicle, road user or device. (m/s).
- Speed limit. Is the maximum legal speed limit. (km/h).
- Percentage speed limit violation. Time and/or distance (or portion of) spend exceeding posted speed limit. (s) or (m), (%), (count).
- Approach speed to events. Speed at ccc seconds or xxx meters before an event. (m/s).
- Acceleration. Referred to the longitudinal, lateral or vertical acceleration of the host or targets vehicles in the road. (m/s^2).
- Brake force. Is the braking power of the vehicle during an event. (-).
- Time gap. The time gap is the value calculated from vehicle speed and gap to a leading vehicle: Time gap = gap / vehicle speed. The time gap to an object, e.g., a lead vehicle (bumper to bumper) or pedestrian, which is travelling in the vehicle's path of travel. (s).

- Probability of following. Reflects the traffic density. (-)
- Time to collision. The mean time required for two vehicles (or a vehicle and an object) to collide if they continue at their present speed and on the same path. Measures a longitudinal margin to lead vehicles or objects. (s).
- Lane Change. Vehicles either must be logged when they change lanes. Alternatively, this can be derived from accurate localisations of vehicle trajectories.
- Route change. Vehicles log a deviation from the previous route. Alternatively, a route change could also be determined as every navigation decision from map matched trajectories. (Number of route change per hour or per kilometre).
- Lane departure. Detect that the vehicle leaves its own lane boundaries. The lane boundaries are defined as the inner edges of the lane markings. The vehicle boundaries are defined as the outer edges of the front wheels.
- Distance between vehicles in a platoon.

A second level research question is for example: *Does IoT improve short range navigation?*

This question is focused on the navigation system of the vehicle and its objective is to identify which benefits IoT can add to the short-range navigation functions and if it also enables new functionalities. Hypotheses are that the range of the short-range navigation can be increased, and the timeliness of received data can be improved by using IoT. It will also prove whether IoT could enable new functionalities to short range navigation such the detection of new types of objects to refine navigation. If the hypothesis is validated, the evaluation will be able to prove that the IoT can enhance or enable the AD functions related to the navigation of the vehicle. Indicators are for example:

- Route changes or number of route changes per hour or per kilometre.
- Travel time uncertainly. The variation on travel time over a certain distance at a specific time. (s)
- Time or frequency in congestion. Driving time or distance spent in congestion relative to total travel time or distance. (%)

More details on motion state measurements and logging are given in section 10.1.5.

5.2.2 Data communication

Communication functionality is provided through alterative communication modes and media. Technical evaluation will focus on the comparison of the performance of alternative communication channels for:

- Ad-hoc V2X communication
- Vehicle – IoT platform communication.

Section 5 of [D1.7] specifies minimum communication performance requirements per use case and device interaction. The objective is to evaluate the realised communication performances in each of these situations and propose feasible performance levels.

V2X communication and communication with IoT platforms is evaluated on the following performance criteria (see also section 5 and Table 20 of [D1.7]):

- End-to-end communication latency; from the generation of a message by the sender, till the reception of the message by receivers.

- Reliability of communication by the packet loss rate or packet delivery ratio of set and received messages.
- Communication range is measured from statistics on and distributions of distances between senders and receivers.

Communication performance is measured for all relevant communication media, speed ranges of devices, and environmental situations experienced at the pilot sites. Communication performance is measured at the facilities or application layers in stations and servers. Note that communication performance indicators for bandwidth and node density may not be evaluated if the node density is too low to experience bandwidth issues during the pilots.

The communication between IoT platforms in the cloud and in vehicles, and between federated IoT platforms are subject of evaluation. The communication between various IoT devices (other than the devices directly participating in the pilots) and IoT platforms is not directly evaluated. The communication for example to road side sensors, drones in 'the cloud', and smartphones of anonymous bystanders will not be evaluated. This communication is indirectly evaluated as it is included in the end-to-end delay from detection time at these IoT devices till the reception of the detections and derived information in the automated vehicles.

On the same note, the communication within a vehicle, and between communication layers within a station, are not evaluated directly either. The net effects of communication performance within and between in-vehicle systems will be evaluated in terms of delays in application decisions and actions, and the overall automated driving performance such as positioning improvements.

To evaluate the performance of communication the locations and timestamps upon sending and reception should be logged. More details and requirements on indicators and measurements for communication are provided in section 10.1.6.

5.2.3 Data management

The concept of data management refers to the capability of IoT devices, such as the automated vehicles being tested, to manage the data needed for the automated driving functions and services. This is the terminology as used for example in [D1.5, section 3.1] as:

> *data management services to applications, including data acquisition, data processing, data fusion at the "edge" and data local storage capabilities to deal with network latency and reliability. Data Management also deals with the collection of information from external elements to the vehicle (i.e. cloud / RSU / other vehicles and infrastructures), exploiting data in order to create services such as planning and control application related to AD system.*

This connotation should be distinguished from the concept of data management to acquire, store and process data for logging and evaluation as used in Task 3.4 and sections 5.4 and 0.

Data management on an in-vehicle IoT platform includes the processes to discovery relevant IoT data sources, to subscribe and process relevant IoT data including the assessment of the quality or the data and fusion with on-board sensor data, and to manage alternative communication channels to search and retrieve required data.

Data management on a cloud-based IoT platform includes device and subscription management, the up and down loading of data from IoT devices, data brokering, discovery services, data aggregation services, (semantic) data transformations to data formats requested by automated vehicles, and the interaction with other IoT cloud services and (federated) platforms.

Technical evaluation evaluates "How the IoT data management adds value to Automated Driving functions?" The added value of IoT platforms and devices is evaluated at the level that a road side or cloud-based IoT platform can provide relevant information to the vehicle IoT platform. IoT functionality and performance is evaluated in the automated vehicles at the interface between the Application layer and the IoT (Capabilities) layer ([D1.7] Figure 34). The evaluations address the following indicators and measures at the in-vehicle IoT platform, at the interface with in-vehicle applications, or at the road side or cloud-based IoT platforms:

- The relevance of IoT data received by the requesting vehicle relative to the (intended) requested data. The relevance is determined by the requesting application in terms of requested data elements, accuracy and timeliness.
- The end-to-end delay between the vehicle request and response reception at the vehicle. This includes tracing the request from the vehicle through IoT platforms and communication networks.
- The end-to-end delay between the detection time of request information, e.g. the detection or generation time of the source IoT device, and the reception time at the requesting vehicle. This includes tracing the source IoT devices through IoT platforms and communication networks.

It is assumed that the automated driving functions and services assess the added value of data, and that the relevant assessments are logged and made available for technical evaluation. The added value can be assessed and logged on the automated vehicles. Alternatively, the added value may also be logged on the IoT platform in the cloud or back end as planned for the Livorno Pilot Site for example. Technical evaluation tools will not be developed to re-engineer or re-evaluate the potential added value for any implementation or an automated driving function or service, or use case.

Note that the technical performance and functionality of platforms like FIWARE, Watson IoT and oneM2M is not evaluated in Task 4.2, nor the communication interfaces with and between these platforms and the semantic data transformations (cf. section 4.3 in [D1.7]). Also, the performance and functionality of the various in-vehicle IoT platforms are not directly evaluated. Verification of these platforms and interfaces is addressed in Task 2.5. Indirectly, the effects are evaluated on the relevance and timeliness of the provided data in the vehicle applications.

An example of a refinement to a third level research question is "*What is the difference in data quality provided via IoT platform versus ITS-G5 (or other advanced V2X communication) in terms of latency, sampling frequency…?"* An hypothesis to test are whether IoT data can enhance the quality of ITS-G5 received information, or whether the IoT data is redundant or provides new information. Another hypothesis to test is whether the vehicle IoT platform and data management functionalities can make efficient use of the redundancy in communication and data sources to increase the reliability of input required by automated driving functions and services, and under which conditions, situations and events. If this hypothesis is validated we could prove that the IoT offers an enhancement on the data management of the system and, so, adds value to the Automated Driving functions.

5.2.4 Security and Privacy

The communication and data both in transition and in rest must be protected against identified threats and fulfil security and privacy requirements defined in dedicated AUTOPILOT document. The final implementation of AUTOPILOT must be reviewed and assessed to verify it is not vulnerable against identified security and privacy threats. In addition, data flows of potentially sensitive information must be reviewed because the solution should be also compliant to GDPR.

Security and privacy requirements of AUTOPILOT solution are defined in [D1.9]. Chapter 5 of [D1.9] summarizes major security threats to the use cases with related risks due to possible vulnerabilities. This chapter should be taken as an initial reference for final assessment and each risk should be assessed for each pilot site to verify it has been well addressed.

Chapter 6 of [D1.9] defines common security and privacy requirements: a minimum that should be followed in each pilot site. This minimum also ensures a compliancy with GDPR regulation in terms of sensitive data handling. Note that this is only a part of the regulation compliancy and the remaining part (such as access to user's information, the right to erase private information, security processes and others) is up to each implementation and will not be assessed during AUTOPILOT evaluation.

Relevant indicators for security evaluation are:

- User interaction timing with and without security.
- Vulnerabilities must be assessed and their severity (Penetration test) must be measured. As reported in [23], the Common Vulnerability Scoring System (CVSS) provides a way to capture the principal characteristics of a vulnerability, and produce a numerical score reflecting its severity, as well as a textual representation of that score. The numerical score can then be translated into a qualitative representation (such as low, medium, high, and critical) to help organisations properly assess and prioritise their vulnerability management processes. Scoring CVSS metrics produces a vector string, a textual representation of the metric values used to score the vulnerability. This vector string is a specifically formatted text string that contains each value assigned to each metric, and should always be displayed with the vulnerability score.

Note that security evaluation will be based on technical documentation and will not require any additional data to be collected during pilot executions.

5.2.5 Environment detections

The environment detections identified in the overall evaluation approach in section 4.2.1 are essential technical measures for improvement of the cooperative and situational awareness of vehicle systems. Environment detections relevant to the AUTOPILOT use cases are for example the detection of:

- Other road users, such as vehicles, pedestrians and cyclists
- Traffic incidents, such as road works and traffic jams
- Traffic information, such as traffic congestion states, traffic light status and signal phases, lane restrictions and usage, maximum speed limits or speed advices
- Road surface conditions, such as pot holes, puddles and speed bumps
- Available parking spots

Potential improvements in detection performance can be measures for example by the type of environmental objects, detection accuracy, rate, and delay, and the geographic position, location and coverage of detections.

Comparison of detection performances of system configurations with and without IoT input should evaluate the added value for example for VRU detection and safety applications, or for traffic state detection, travel time predictions and traffic efficiency. Examples of indicators are:

- Time to collision. The mean time required for two vehicles (or a vehicle and an object) to collide if they continue at their present speed and on the same path. Measures a longitudinal margin to lead vehicles or objects. (s).
- Bluetooth communication could be used for traffic monitoring, also to determine whether the driver has left the vehicle, e.g. through pairing of the driver's smartphone with the in-vehicle OBU. Bluetooth Low Energy is also considered an interesting technology for WSN applications demanding higher data rates, but short range.
- ZigBee is an IEEE 802.15.4 based specification low cost and low power technology, for low data rate and short-range applications.
- Data produced by the virtual devices on the oneM2M platform. Each physical sensor/actuator both in the vehicle and in the infrastructure, may have a "virtual" representation on the oneM2M platform, that dynamically collects the data produced by the "physical things", storing and sharing them for applications. The data published on this IoT Cloud based platform are important to evaluate the usefulness and the quality of the applications.

5.2.6 Safety

Safety of automated driving is addressed as part of approval procedures in pilot site preparations. Automated driving safety will not be evaluated extensively in work package 4, because of the limited scope and scale of piloted situations.. However, the use of IoT data may affect the safety of automated driving and any incidents should be reported and investigated. In addition, an assessment of the impact of IoT on safety can also be made at the end of the project from the previous evaluation results, most notably on environment detection performance, data management, positioning, localisation and manoeuvring.

As a minimum, any safety related incidents under real traffic conditions during piloting should be reported and investigated / evaluated (see [D5.3 section 4.6, Table 12]). Any human intervention, e.g. by a test or co-driver, to disengage an automated driving mode, function or (safety-relevant) service in real-traffic conditions is considered as an incident that should be reported. Causes or situations that caused the incident to report include weather conditions, inattentive road users, unwanted vehicle manoeuvres, and hardware or software failures.

5.2.7 Use Cases

Previous subsections presented the research questions and hypotheses that are common across use cases. In addition, the research questions will also be answered "How can IoT improve a specific use case?" For each use case, this research question is refined in third level research questions for a subset of the above described concepts that are specifically relevant to the use case. Differentiation per use case primarily concerns the situations and events that are expected in the pilot test scenarios. The indicators will be calculated per situation and for the events, hence differentiation is defined from the pilot scenarios that will be briefly described in the next section.

Research questions that are specific to a single use case or implementation may be added, and if necessary the evaluation methodology may be adapted as well, during the technical evaluations later.

5.3 Piloting scenarios

For evaluation of the functionality and performance, the indicators must be determined for similar situations and events in pilot test scenarios. This section briefly describes the most relevant situations and events per use case and the measurements needed to distinguish the situations and events from pilot log data.

Specification of the pilot scenarios is ongoing in work package 3. Pilot scenarios for testing the implemented use cases have been described at a high level already in D1.1. This section gives only a first, high-level, analysis of the scenarios that must be refined in the next phase of pilot preparations.

5.3.1 Automated Valet Parking scenario

The AVP use case has several relevant situations, most of them involving how the test vehicle drives in automated mode from the drop off point to the parking slot and how the vehicles deal with all the obstacles on their way:

- **Drop-off of the vehicle**. The driver will leave the vehicle at the drop-off point. It will connect to the parking infrastructure and it will go from manual mode to AD mode. This is the first important event to log of the use case in order to check timing references and connections and changes checkouts.
- **Routing of vehicle**. The process of looking for a spot available, creating a route inside the parking and driving in automated mode to the destination. It is important to check the route optimisation and evaluate how the system schedules the vehicle.
- **Legacy car or VRU detection**. If during the route a pedestrian, a legacy vehicle or another obstacle appears, the vehicle reaction will be evaluated. For this reason, it is important to test the use case not only in a controlled environment but also in a real world environment.
- **Parking manoeuvres**. The parking process will be evaluated on the parking time needed, the space optimisation on the parking lot, and the number of manoeuvres needed to park.
- **Request and pick-up vehicle**. When the vehicle is requested to be returned to its owner, the parking manoeuvre, routing, and obstacle detection also need to be evaluated till the pick-up zone. Waiting or pick-up duration, disconnection from the parking system and the change from AD to manual mode will also be evaluated.
- **Transition from AD to manual**. Every switch from manual to automated driving mode is be logged with all the information related to the switch including any failure message, transition event, vehicle position and motion state data, vehicle sensor and situational data that triggered the transition. The trigger or transition conditions are the main indicators.

5.3.2 Highway Pilot scenario

The highway pilot use case distinguishes scenarios for automated vehicles and for cooperative vehicles, and in addition for the following situations and events:

1. A test vehicle detects the road hazard and sends out a warning to the driver, other road users and the IoT platform.
2. A test vehicle receives a warning for a road hazard and warns the driver or triggers an automated driving function or service.
3. An automated test vehicle detects the road hazard and reacts automatically as in 1. and 2.

Indicators are defined in general terms for:

- Detection performance of the road hazard or event.
- Location or position accuracy of initial warning or automatic reaction.
- Latency between detection and initial warning or response.

- Distance between event position and initial warning or response.
- Smoothness in longitudinal and lateral manoeuvres of automated responses.
- Occurrence of emergency responses such as hard braking or steering.

In all cases, the detections, communication and response (warning or automatic response) should be logged, together with relevant information of the trace of vehicle positions and road or lane location, additional motion state and manoeuvres, vehicle sensor data, IoT data, and situational data for the trigger or detection conditions.

Following scenarios, events and situations can be distinguished for different types of road hazards:

- **Detection of a road (surface) condition**. The test vehicle should detect the road condition, or receive a warning, with a safety distance enough to react. Detected events and road surface defects are logged. The evasive manoeuvres of a driver or the automated control reactions are logged.
- **Emergency braking / slow vehicle**. In addition to the evaluation of the event detection and response is also the research question why the driver or system had not detected the event earlier to allow smoother reaction.
- **Breakdown or accident**. The detection and warning/response delays are the main indicators of the event.
- **Fast approaching emergency vehicles**. Warning time and distance, and the accuracy of event location relative to the receiver are the main indicators.
- **Traffic jams and queues**. The receiving or host vehicle should adapt its speed and acceleration in relation to the traffic jams. The main indicators for this event will be the warning delay and distance, and the smoothness of the approach (speed, acceleration, time gap, braking).
- **Nearby presence of VRU**. When the vehicle detects the presence of bicycles or pedestrians, or receives a warning, it should respond immediately to avoid any dangerous situation. The vehicle sensors (camera, lidar or radar) and the V2X message are the main indicators for the event detection on the host vehicle.
- **Weather related condition**. When the vehicle receives weather information such as a rain warning it should adapt its speed to maintain the safety. The vehicle should log the estimated safe speed for comparison to the actual speed. The main indicators are the V2X message and the vehicle data and sensors.
- **Transition from AD to manual**. Every switch from manual to automated driving mode is be logged with all the information related to the switch including any failure message, transition event, vehicle position and motion state data, vehicle sensor and situational data that triggered the transition. The trigger or transition conditions are the main indicators.

5.3.3 Platooning scenario

The platooning scenario distinguishes states or situations for creation, management and control of the platoon, interaction with surrounding traffic, and the reallocation of the vehicles.

- **Platoon scheduling and organisation**. Situations are distinguished where vehicles start searching other vehicle to form a platoon, and routing to meeting points. Events are defined by the interaction protocol for brokering and routing. The main indicators for a correct evaluation will be the GPS position, the vehicle data and the IoT messages.
- **Platoon forming process**. Platoon formation starts once platoon members are nearby. Situations and events are defined from interaction protocols to lead, join, merge or leave a platoon. The V2V messages and the vehicle data are the main indicators.

- **Interaction with normal traffic**. Events and situations for interactions with nearby traffic are distinguished, such as merging, entry and exit, cut-in, lane changing, overtaking, breaking, crossing. Safety indicators like time gaps and time-to-collision, in addition to the V2X messages within the platoon, the GPS position and the vehicle data and sensors are the main indicators for assessing safety and smoothness of platoon response.
- **Controlled intersections**. Platoons must cross controlled intersections using C-ITS services for traffic light status, priority requests and speed advice. Situations are distinguished for the approach to red or green light, approval or denial of priority, and traffic queuing. The V2X message, the GPS position and the vehicle data are the main indicators.
- **Dedicated lane use**. Platoons can request the used of dedicated lanes, e.g. the hard shoulder. Situations are defined from the Interaction protocol with traffic control to request, and also incidents on the dedicated lane such as stranded or merging vehicles.
- **Transition from AD to manual**. Every switch from manual to automated driving mode is be logged with all the information related to the switch including any failure message, transition event, vehicle position and motion state data, vehicle sensor and situational data that triggered the transition. The trigger or transition conditions are the main indicators.

5.3.4 Urban driving scenario

The urban driving use case includes the interaction with the traffic lights and traffic signs, vulnerable road user like bicycles or pedestrians, and legacy cars.

- **Single and multiple (uncontrolled) intersections**. The intersections are one of the most relevant situations in the urban driving use case. Situations to distinguish include priority rules, lane restrictions, lane markings, speed limits, potential conflict situations with legacy traffic, and traffic queues and density.
- **VRU interactions**. Each time the vehicle interacts with a VRU needs to be logged to evaluate the reaction to it. The vehicle sensors and data, and IoT data about VRUs are the main indicators.
- **Controlled intersections**. Vehicles must cross controlled intersections using C-ITS services for traffic light status, priority requests and speed advice. Situations are distinguished for the approach to red or green light, speed advice and compliance, and traffic queuing. The V2X message, the GPS position and the vehicle data are the main indicators.
- **Traffic rules and signs**. Vehicles use either a local map or receive IoT data on traffic rules and signs. Evaluation with vehicle data reveals compliance or violations. Vehicle sensor data may also identify any conflicts with neighbouring traffic.
- **Transition from AD to manual**. Every switch from manual to automated driving mode is be logged with all the information related to the switch including any failure message, transition event, vehicle position and motion state data, vehicle sensor and situational data that triggered the transition. The trigger or transition conditions are the main indicators.

5.3.5 Car sharing scenario

The car sharing use case relevant situations involve the requesting and the assignment of the vehicle and also the valuable situations that could occur during the route to the user:

- **Waiting time from request to pick up**. The time between the request and the pick-up must be logged to evaluate the time optimisation and the route alternatives.
- **Car-customer and vehicle-user assignment**. The process of the assignment of the vehicle to the user should also be recorded to evaluate how optimal it is.

- **Events detection during route**. The obstacles and hazards encountered during the route should be logged also. The main indicators are the V2X message and the vehicle sensors.

5.4 Data Collection

Technical evaluation primarily needs log data from the vehicles and IoT platforms and cloud services and situational data from the pilot sites to detect situations and events, and to calculate the indicators. The log and situational data are accessible from the central data management server for evaluation purposes, as described in section 4.4 and Figure 3.

Requests and requirements for data provisioning by the pilot sites through the central data management server, and general data quality requirements are included in section 0.

The requirements for log data to be provided by the central management system, including the initial set of data sources, parameters and quality criteria, are defined in a spreadsheet in Annex 14.2. This is a living document on project place. The current version is an initial version that will be updated during pilot preparations and the refinement of the technical evaluation methodologies.

The log data parameters needed for technical evaluation are organised by data sources. This section identifies the groups of data sources.

- **Vehicle sources.** This group defines the lists of in-vehicle data sources, including the vehicle on-board sensors, and systems connected to the vehicle network (i.e. CAN bus) as data sources:
    - Vehicle dynamics sensors like the rate, speed and acceleration sensors, and on-board GPS.
    - Environment sensors like cameras, LIDAR, radar.
    - Vehicle control systems, such as the location system, navigation system, driver or vehicle interaction control, pedals and steering, and the HMIs.
    - External information systems like a Local Dynamic Map or HD-Maps systems, HVAC.
    - Communication units for V2X and cellular network communication.

  Usually most of this data is accessed and checked by CAN.

  Additional sheets define parameters to be collected for the sources.

- **Vehicle Data**. Defines specific log parameters for the vehicle sources.

- **Derived Data**. Defines parameters that will be derived from the log parameters.

- **Positioning**. If accurate, lane level accurate, positioning needs to be evaluated for a specific pilot test scenario, then we expect to receive GPS information in a standard NMEA format. The NMEA format is by sentences, each one containing a type of information. The essentials are the GGA, RMC, GSA, VTG and ZDA.

  In other pilot test scenarios, it is sufficient to track the vehicles and manoeuvres using WGS84 way points, like the definition of reference positions in ITS-G5 messages (latitude, longitude and heading with confidences).

- **V2X messages**. Defines the mandatory data elements from V2V and V2X messages, including the ITS-G5 messages for CAM, DENM, IVI, SPAT, MAP. For evaluation not all data elements

may be needed, so a subset of the mandatory elements is acceptable under the following conditions:
    - The full message is logged in encoded format (binary UPER or XER), or decoded e.g. in csv or SQL database tables.
    - The decoded messages should be logged as defined in the standards; i.e. maintain the structure of data frames and in the exact data element names.

- **IoT messages**. The IoT messages are still being defined for the use cases, and these have not been included in Annex 14.2. The same approach will be adopted as for the V2X message; i.e. all IoT messages are assumed to be standardised messages.

  Each physical sensor/actuator from an IoT device both in the vehicle and in the infrastructure may have a "virtual" representation on the oneM2M platform, which dynamically collects the data produced by the "physical things", storing and sharing them for applications.

- **Other indicators**. The Wi-Fi or Bluetooth message with information that may not follow any standard is also important to collect.

- **Events.** The relevant events must be specified from the pilot scenarios (section 5.3). Events should be defined for communication, application logic, and user interactions. Events are detected and logged from the applications. Alternatively, events can be defined as functions of log parameters and generated off-line from the log data.

- **Situations.** The relevant situations must be specified from the pilot scenarios (section 5.3). Situations can be defined as functions of the parameters from the previous data sources. Additional situational data must be identified for collection by pilot sites and added later to Annex 14.2.

5.5 How Technical Evaluation differs from Pilot Readiness Verifications and Validations

The Pilot Readiness verification and validation aims to verify the components and solutions developed and integrated in WP2 before the automated vehicles and IoT services are used on the piloting activities. This activity does only cover the verification for pilot readiness and no evaluation activities. These verification activities will include:

- Testing of IoT Platform, IoT device and application functionality and interoperability.
- Testing of communication devices interoperability.
- Automated driving vehicle adaptation verification.

The end-to-end system test ensures the overall readiness in terms of functionality, robustness and performance for piloting.

Otherwise, the Technical Evaluation objective is to evaluate the suitability of IoT technologies applied for automated vehicles on the different pilot sites. This evaluation will include:

- Overall performance and safety assessment of IoT solutions in the connected and automated driving pilots.
- Evaluation of the replicability and sustainability of the implemented architectures
- Interoperability assessment between the IoT technologies and IoT architectures for the required provision of services for connected and automated driving.

- Assessment of security and performance mechanisms provided by IoT solutions in comparison to the required level of security and performance required for Automated Driving.

6 User Acceptance Evaluation Methodology

This section describes the methodology for the user acceptance evaluation, which will be conducted in Task 4.5. The section first outlines the research objectives and the rationales behind them. Second, it formulates an iterative, design-oriented research approach. After defining the underlying concepts, the section formulates research questions and corresponding hypotheses. Then, the section enumerates necessary preconditions for answering the research questions. Last, it lists indicators needed for the evaluation and concludes with an outlook on the envisioned data collection.

6.1 Research Objectives

Evaluating user acceptance in the framework of the AUTOPILOT project is constrained by three central issues. First, the IoT aspect of the project is not only intangible, but also new. It cannot be expected that users have a remote understanding or even an interest in the underlying infrastructure behind their automated driving experience. Second, the automated driving aspect is dynamic. Functionalities in the AUTOPILOT project will be further developed during the piloting phase and are not ready for market penetration. Third, for legal and sometimes practical constraints, most of the AUTOPILOT applications and services will be tested not by “regular” users, but by trained drivers and company employees, thus preventing both a large sample size for each use case per pilot site and a representative sample.

The formulation of the research objectives takes these restrictions into account and carves out a space, in which user acceptance evaluation can still be useful for the project. The two-pronged objectives of the project are to:

- Formulate IoT-related improvements for automated driving functions based on user feedback.
- Determine, whether there are improvements or added value in automated driving functionalities with and without the assistance of the IoT regarding user acceptance.

The first objective is directed rather at the full duration of the project. It is also directed at those AUTOPILOT applications and services that accelerate or enable automated driving. Its underlying conceptual assumption is that users are taken as co-designers, not merely as objects of evaluation. The necessary ramification of this objective is that the evaluation approach takes a strong qualitative turn. In addition, the objective prescribes an iterative research design, which is further described in section 4.3.

The second objective is directed rather at the end of the project. It is also directed at those AUTOPILOT applications and services that enhance automated driving. Its underlying assumption is that the IoT can be switched on and off for some test runs, which also predicates that there is a noticeable difference for the user. The translation of this assumption into piloting requirements is done in section 6.5.

6.2 Research Approach

Based on these objectives, an iterative research approach is sensible for user acceptance evaluation. The basic outline of this approach is presented in Figure 5 below.

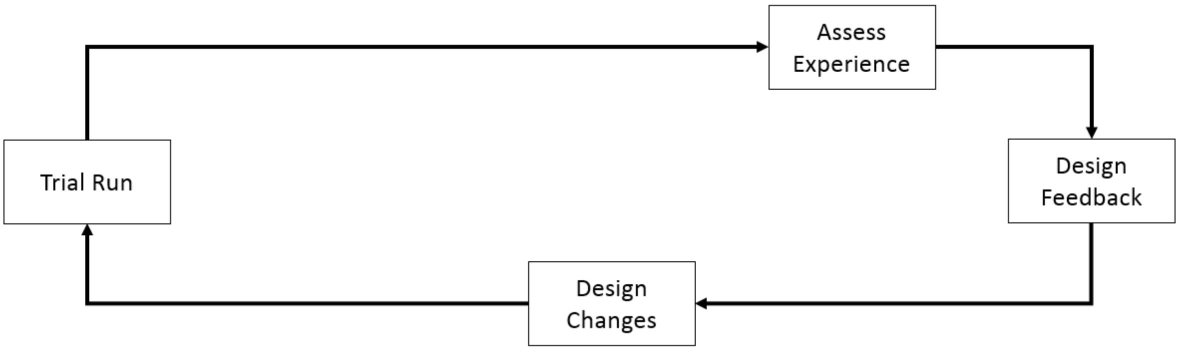

Figure 5: Iterative Research Approach

At the pilot sites, trial runs are conducted over the course of the project. For participants in the trial runs, their experience during the trial is assessed. Their responses take the form of feedback into the design process of the AUTOPILOT applications and services. At the pilot sites, the feedback is looped back into the design process to propel changes in preparation of the next trial run.

This research approach focuses on the user not as object of the evaluation, but of the user as co-subject in a process, which is geared towards formulating improvements on emergent technologies. The experience of users is treated as central to ameliorating technical and non-technical elements of automated driving as experienced in the AUTOPILOT trials. The purpose of user acceptance evaluation, therefore, is to gauge the experience as clearly as possible and to open avenues for users to articulate feedback on their experience.

6.3 Underlying Concepts

Paramount to an understanding of the methodology of user acceptance is the definition of this very concept. Users are defined in line with section 2.3. Acceptance is conceptualised as the "*degree of intention to use or of incorporation of AUTOPILOT services*". This definition follows established work done in [8] and provides an individual driver-centric look on user acceptance in line with [7].

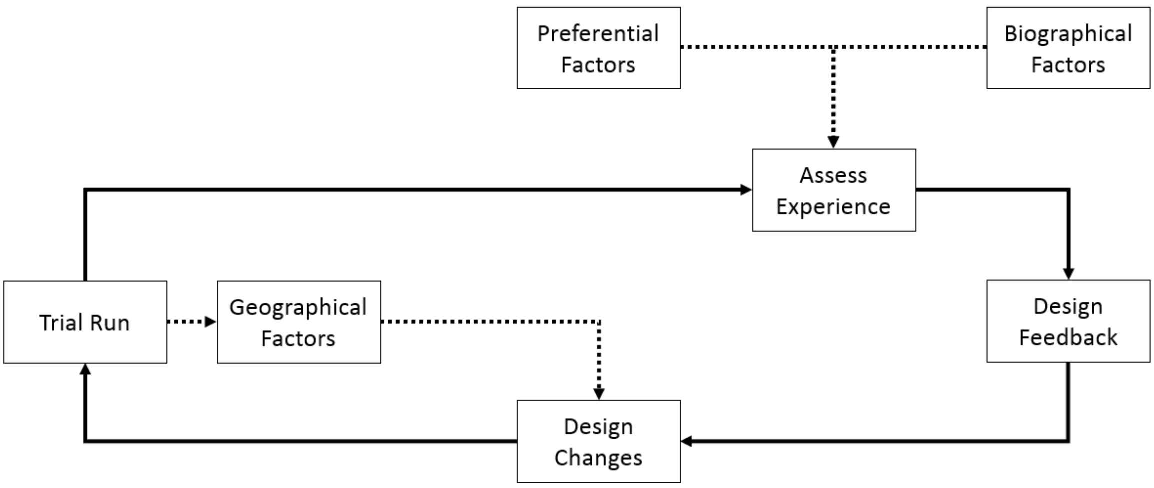

Figure 6: User Acceptance Factors

User acceptance is conceptualised to be dependent on several factors, which are shown in

Figure 6 above. These factors feed into the user experience and are expected to shape the feedback given for design changes.

First, biographical factors are mediating variables on the acceptance on new technologies [10]. In addition to a general list of common biographical factors, the driving experience is assumed to shape user acceptance. Given the restrictions on user groups and the prevalence of trained drivers among the trial participants, previous experience with or exposure to automated driving should also be included among the biographical factors.

- Age
- Gender
- Level of education
- Income
- Legacy driving experience
- Automated driving experience

Second, user acceptance is taken to be a composite concept of a set of preferential factors, which reflect a spectrum of attitudinal stances [8]. It is assumed that not all these preferences are of equal value to all participants, but that some preferential factors are more relevant for technological acceptance. These factors reflect a general distinction in the expectancy regarding performance of the automated driving function and effort in using it [9][20][21]. Preferential factors include:

- Perceived usefulness
- Perceived ease of use
- Perceived control
- Perceived trust
- Perceived security/safety
- Data control and access
- Stress/Mental Workload

Third, geographical factors need to be considered. These factors describe the external characteristics of each individual driving situation that user experience. While it would not be sensible to ask for user acceptance of rainy weather during the test run, awareness of the weather conditions is nonetheless important to contextualise the environment that shaped the user experience of an automated driving function. These factors include:

- Weather condition
- Road condition
- Traffic density
- Time of day
- Time of the year
- IoT Usage

6.4 Research Questions and Hypotheses

Where possible, research questions with focus on pilot sites or use cases are formulated. Research questions are clustered in Table 3 below, which is based on information provided in AUTOPILOT Deliverable D1.1 and section 4.1, with green indicating “enhancement”, orange indicating “acceleration” and blue indicating “enabling”. Dark grey indicates the absence of a function per pilot site and light grey indicates the decision to not test a specific use case per pilot site.

The full list of research questions is provided in Table 4 below. This list of research questions is indicative, but not definitive, as it is expected to evolve dynamically over the course of the iterative evaluation process.

Table 3: Cluster of Research Questions

	Automated Valet Parking	**Highway Driving**	**Platooning**	**Urban Driving**	**Car Sharing**
RQ-UA 1 - 3	RQ-UA 4	RQ-UA 8	RQ-UA 11	RQ-UA 14	RQ-UA 21
Tampere	RQ-UA 5			RQ-UA 15	
Versailles			RQ-UA 12	RQ-UA 16 RQ-UA 17 RQ-UA 18	RQ-UA 22 RQ-UA 23
Livorno		RQ-UA 9		RQ-UA 19	
Brainport	RQ-UA 6	RQ-UA 10	RQ-UA 13		RQ-UA 24
Vigo	RQ-UA 7			RQ-UA 20	

Table 4: List of Research Questions

RQ Code	**Research Question**
RQ-UA 1	Which improvements do users suggest for the AUTOPILOT function?
RQ-UA 2	In how far are users willing to provide their data to the IoT ecosystem?
RQ-UA 3	Does the IoT increase user acceptance of automated driving?
RQ-UA 4	Does provision of information on parking status and position of the vehicle through the IoT enhance acceptance of Automated Valet Parking?
RQ-UA 5	Do weather conditions increase user acceptance of Automated Valet Parking?
RQ-UA 6	Do reduced drop-off and pick-up times increase user acceptance?
RQ-UA 7	Do users perceive the new indoor automated valet parking service as useful?
RQ-UA 8	Do users feel capable of dealing with the mental workload incurred during the highway pilot?
RQ-UA 9	Do users need puddle and roadwork detection on the highway?
RQ-UA 10	Do road hazard warnings decrease perceived control of the driving situation by users?
RQ-UA 11	Does usage of the IoT increase user acceptance of Platooning?
RQ-UA 12	Does the formation of the platoon happen quickly enough for the passengers?
RQ-UA 13	Is perceived comfort of platooning independent from weather conditions?
RQ-UA 14	Do users feel comfortable during their urban driving experiences?
RQ-UA 15	Is the car reaction to IoT signals in the lead-up to traffic lights comfortable for users?
RQ-UA 16	Do tourists consider the AUTOPILOT experience useful?
RQ-UA 17	Which factors impede usage of the point-of-interest notification?
RQ-UA 18	Does usage of the IoT increase user acceptance of level 4 driving in the gardens?
RQ-UA 19	Does the option to engage in non-driving activities increase user acceptance of automated urban driving?
RQ-UA 20	Does usage of the IoT increase user acceptance of urban driving?
RQ-UA 21	Does usage of the IoT increase user acceptance of car sharing?
RQ-UA 22	Which factors influence the perceived ease of use of the car-sharing app?

RQ-UA 23	Is the drop-off scenario convenient for users?
RQ-UA 24	Are users more satisfied with the waiting and travel time than with other car sharing providers in the region?

Research questions fall into three categories, which broadly resemble the multi-layered approach of the FESTA Methodology [1]. The first three research questions (RQ-UA 1, RQ-UA 2 and RQ-UA 3) address all AUTOPILOT pilot sites and use cases in the area of user acceptance, making them the most high-level research questions. Whereas the very first research question translates the first research objective, the other two high-level questions refer to the second research objective.

Five further research questions (RQ-UA 4, RQ-UA 8, RQ-UA 11, RQ-UA 14 and RQ-UA 21) address multiple pilot sites, but one individual use case. The underlying objective for these questions is to be able to qualitatively compare AUTOPILOT solutions and user acceptance thereof across pilot sites. These research questions take a middle-level in the FESTA understanding.

The remaining research questions address individual use cases in individual pilot sites, making them the most low-level research questions in the design. The underlying rationale for these research questions is not to list all possible research questions for each pilot site. Instead, these research questions are formulated with the objective to address interesting, unique or prevalent issues in the pilot sites, to facilitate a design-oriented evaluation process. Consequently, some research questions might be interesting for other pilot sites or use cases as well. For reasons of manageability, these will, however, only be transferred in exceptional cases during the execution of T4.5.

For each research question, one or multiple hypotheses are formulated. The full list of hypotheses is reproduced in the Annex 14.1 to this report. For the research questions of a more exploratory nature (e.g.: RQ-UA 1), the formulated hypotheses are not understood to be the basis for testing, but rather a guidance for expectations of the qualitative co-design process. For research questions of a more explanatory nature, the formulated hypotheses are expected to be tested qualitatively and quantitatively, where possible. Where adequate, a clear null hypothesis is formulated to facilitate this evaluation approach (*e.g. hypothesis H3.2: Usage of the IoT does not increase user acceptance of highway driving*).

6.5 Piloting Requirements

The pilot test scenarios need to be set up in a conducive way to test the hypotheses. This is in line with the scenarios described in sub-section 5.3.

For all use cases, where a comparison of the automated driving function with and without assistance of the IoT is possible, this should be implemented. Back-to-back test runs should be conducted with IoT-related support switched on or off. The order of the test runs should be randomised. The pre-trial briefing should not explain the difference between the two test runs beforehand. This piloting requirement applies for example to:

- Urban Driving in Versailles
- Car Sharing in Brainport
- Platooning in Versailles

For some of the use cases, where a comparison between the IoT-enabled automated driving function and the manual driving baseline is possible, back-to-back test runs should be conducted with IoT-related support switched on or off. The order of the test runs should be manual first, automated second. This requirement applies to:

- Automated Valet Parking in Vigo
- Automated Valet Parking in Tampere
- Automated Valet Parking in Brainport
- Urban Driving in Brainport

For all use cases, where small elements such as artificial road obstacles or increased traffic density can be actively altered by the pilot sites owners, back-to-back tests with the user should be conducted. The order of the test runs should be randomised. The pre-trial briefing should not explain the difference between the two test runs. This piloting requirement applies to:

- Highway Driving in Livorno
- Highway Driving in Brainport
- Automated Valet Parking in Brainport
- Urban Driving in Tampere

Mindful of the legal and practical restrictions on the inclusion of naïve users in the trial runs, it is nonetheless advisable to maximise the diversity of test users on the sites, which also applies to certified drivers and employees of the testing companies. The reason for this is not to achieve a balanced sample for statistical analysis, but rather to include as many different voices and experiences into the user-centric design process, which is at the heart of the evaluation. Pilot sites should be required to actively work towards a spread sample of users to test the AUTOPILOT applications and services.

6.6 Indicators and Metrics

User acceptance is best understood as a scaled quantitative metric, but binary metrics or qualitative assessments are also possible. Indicators will be selected based on pre-existing work on user acceptance.

Weather condition and road condition are categorical indicators, which should be chosen equally across all pilot sites and use cases. Time of the day and time of the year are equally categorical measurements. IoT usage is binary metric. Where possible, trials of the AUTOPILOT service are to be completed with or without assistance of IoT.

Age, income and level of education are categorical indicators. Legacy driving experience and automated driving experience could be treated as categorical indicators or as scaled indicator. All biographical data points are self-reporting variables, which will not be cross-validated.

If pilot sites purposefully alter individual variables in their testing, such as artificial road obstacles, these should equally be recorded in an appropriate format.

6.7 Data Collection

The multi-faceted research approach requires recourse to multiple ways of data collection. Based on the interplay between the iterative research design and the factors of user acceptance, three means of data collection are adequate for the evaluation, as indicated in Figure 7 below.

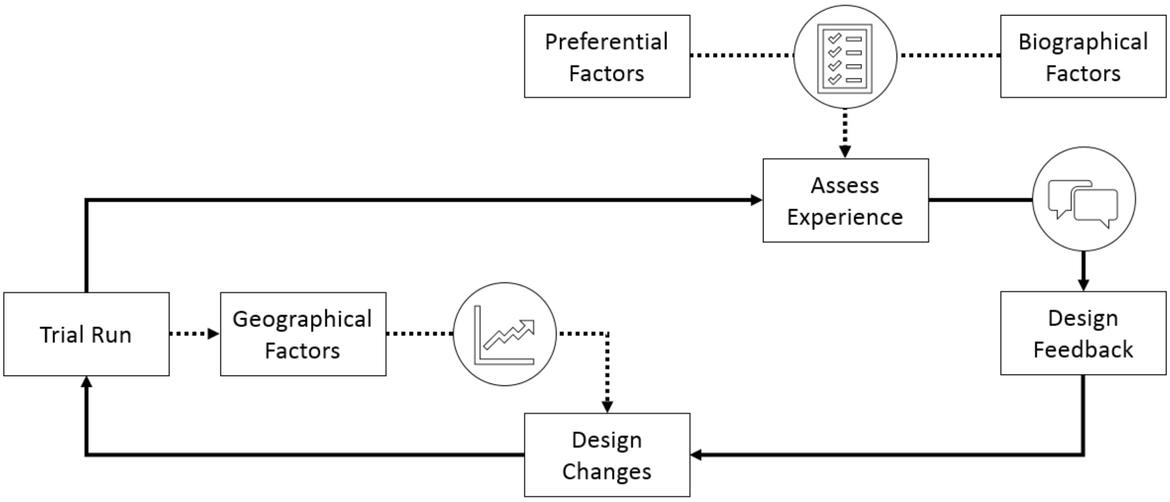

Figure 7: Multi-Method Data Collection

First, biographical factors as well as the interplay between preferential factors and the user experience during the trial run will be collected by means of questionnaires. These questionnaires will be translated in the language of the pilot site. They will include closed items, Likert scales and a limited amount of open ended questions. The exact nature and content of the questionnaire will be determined within T4.5. The reliance of the preferential factors on previously established models allows using items from surveys developed by [21] and [9].

Second, individual interviews and focus group interviews will be organised to link the user experience and the assessment of their acceptance to an improvement-oriented feedback. The interviews will be semi-structured to engage with individual differences while gearing the progress of the interviews towards design-relevant questions. Interviews will also take into consideration the preferences of users as collected in the questionnaires. More information on the scope and frequency of the interview sessions is given in section 9.2 below.

Third, the geographical factors will be collected in log data of the test runs. The data needs to be able to be linked to individual users. If interviews are conducted immediately after a test run, the log data should be made available to the interviewer in order to take it into account for evaluating the responses.

7 Quality of Life Impact Assessment Methodology

This section presents the methodology for the assessment of the Quality of Life (QoL) impacts that will be conducted in Task 4.4. The scope for assessment is limited to the concepts for QoL where most added value of IoT could be realised and can be related to the pilot test scenarios planned in AUTOPILOT. The scope is limited by definition of the objectives and selection of the concepts. Finally, some guidelines are given for piloting scenarios and data collection.

7.1 Objectives and methodological approach

The main objective is to assess the impact of automated driving, progressed by internet of things, on the quality of life. The quality of life assessment includes impacts on personal mobility (travel behaviour), sustainability (traffic safety, transport system efficiency and environment) as well as well-being (e.g. health) (see Figure 8).

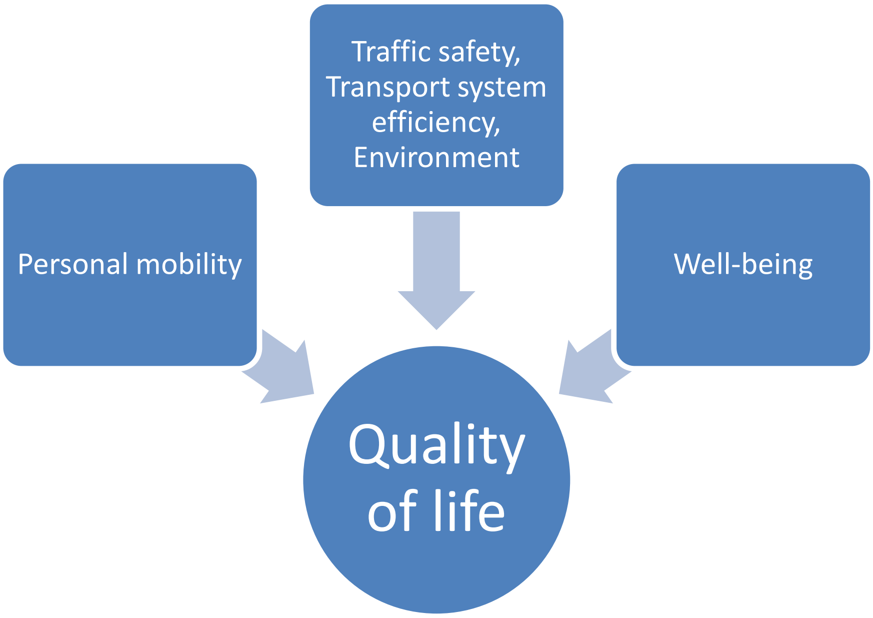

Figure 8. Components of quality of life assessment.

The AUTOPILOT project studies the effects of IoT on automated driving. IoT enables connectivity between vehicles, travellers, the transport network and virtually all other aspects of life. Thus, this task aims to investigate the impacts on society due to connected automated driving: deployment of IoT in combination with automated vehicles.

Introducing automation in transportation is going to fundamentally change not only the transport system, but also mobility behaviour and subsequently people’s daily life. The quality of life impact assessment supports public authorities, businesses and other stakeholders in the decision making concerning the deployment of IoT solutions and automated driving.

The main objectives of this task are to:

- Explore how IoT in automated driving meets personal mobility needs
- Explore the improvements in transport system efficiency with various penetration rates of IoT devices and automated driving vehicles.
- Explore the contribution of IoT to traffic safety improvements
- Explore the contribution of AD and IoT to citizens’ well-being

These effects are mediated via changes in vehicle and travel behaviour induced by AD and IoT (Figure 9). These changes are studied in the pilot tests. Data on first order measures (driving behaviour) is collected within the tests, and second order impacts (such as mobility, safety) are produced with tools such as expert assessment, surveys or modelling. These effects can then be changed to monetary values for socio-economic assessment (section).

Driver and vehicle behaviour is defined in the trilateral impact assessment framework [24] as including acceleration, deceleration, lane keeping, car following, lane changing and gap acceptance. Those are affected by automated driving functions (ADF), which control longitudinal or lateral movement: for example adaptive cruise control (ACC) and lane keeping assistants.

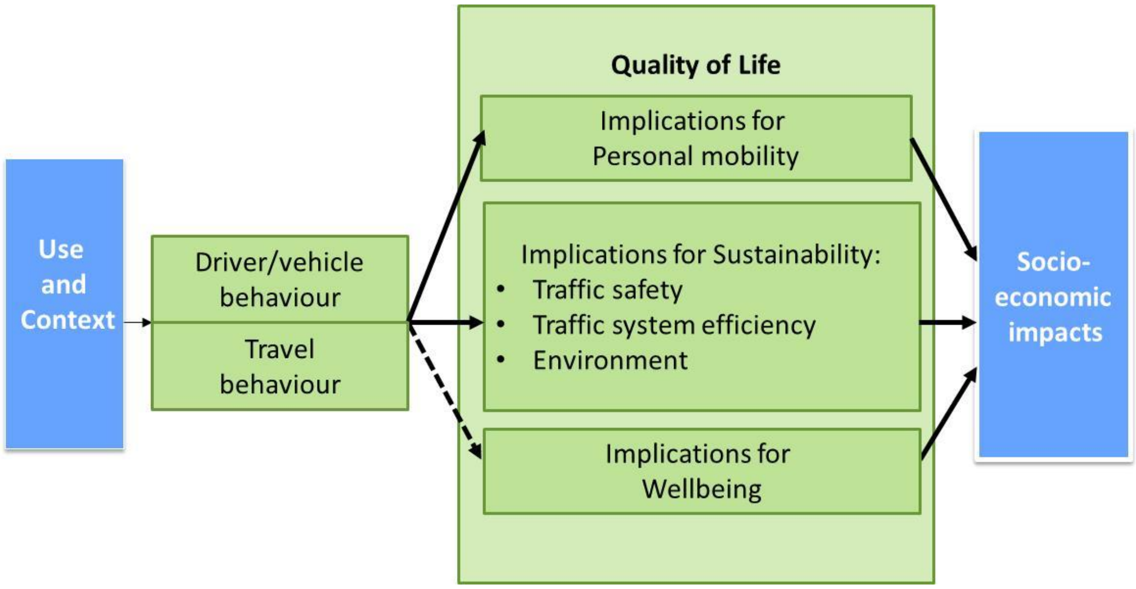

Figure 9 Changes in vehicle and travel behaviour induced by AD and IoT

Road transport automation is expected to have both direct impacts in the short-term as well as indirect impacts, which will take a longer time to form [25].

The focus of the assessment is mainly on general impacts and common road users. Data will be collected from the test vehicles as well as from test users. However, when applicable also impacts on other stakeholders, such as traffic management centres will be estimated. Three main areas of impacts on quality of life are addressed: personal mobility, sustainability (including safety and efficiency) as well as well-being.

As the potential impacts are numerous and wide reaching, and long-term trials are not possible in this pilot project, a scenario-based approach will be used for impact assessment. The new opportunities offered by the AUTOPILOT functions will be presented with examples, which are easy to relate to. Scenario-based analysis is a way to incorporate and generate ideas for new products or services and to identify the possible users and contexts of use for these products.

There is a lot of interest in the media regarding automated driving, but the actual functionalities, which will be available to consumers at the initial phases, are not that familiar to the public. The advantage of using scenarios is that they make ideas more concrete and describe complicated situations and behaviours in meaningful and accessible terms. Scenarios also help different stakeholders understand the implications of new solutions especially from the point of view of the users [44]. Scenarios are not intended to describe the whole functionality of a system, but to describe individual users in individual usage situations.

The functions and use cases will be approached through different scenarios, from a day-to-day perspective, adapted to users' own choices in their daily lives through detailed reflection of their travel behaviour. In a way, scenarios are stories, which make it easier for end users (and other stakeholders) to relate to real-life examples as opposed to abstract system descriptions [44]. This approach makes it easier for the users to assess potential changes and produces more reliable results, although it limits the generalisation of the results [42][43].

The main elements of a scenario are (after [44]):

- user group: who will use the systems
- context: in which situations will the systems be used
- circumstance: under which circumstances will the systems be used
- goals or motivations: why will the systems be used
- interfaces: how will the systems be used
- time frame: when, for long, how often and in which period of life time will the systems be used
- qualities: physical factor of devices (probably not relevant in AUTOPILOT)

The scenarios will be elaborated in the work of task 4.4, in close cooperation with the pilot sites and storyboards. An example of a case would be how commuting trips could change, if for example the use of motorways would be preferred with the platooning function. Another case is presented by the new mobility options enabled for people with reduced mobility (PRM): At present it might be difficult to move from A to B, but AD and IoT could provide a welcome change.

The main method applied in this task is well structured expert assessment based on versatile data obtained through surveys, interviews, pilot tests and simulations. Questionnaires can be complemented with deeper interviews before and after tests. Also focus groups are a possibility for the evaluation. Travel diaries are a suitable tool for assessing potential changes in travel behaviour with Autopilot functions, which trips would be affected etc. They also help the test participants in relating the possibilities offered through AUTOPILOT functions to their daily trips.

The work will start with a meta-analysis of broad impacts of automated driving and IoT from published papers and past projects. Furthermore, a simple but plausible system dynamics based model will be developed to (a) assess the factors and variables that influence the quality of life as a result of automated driving progressed by IoT, (b) test scenarios and (c) evaluate the effectiveness of policy interventions. This model will help understand the underlying structure of relationships producing the observed patterns, which can be obtained from many AD-related or FOT projects.

To reflect the specific interest in IoT from the AUTOPILOT consortium, different adoption and diffusion scenarios will be explored for the IoT technologies tested in the project. This model will then be used to quantify or qualify the important social and environmental factors contributory to the short- and long-term impacts of the combination of AD and IoTs on quality of life defined in the AUTOPILOT project.

The outcome of this work will identify the impact of the IoT solutions on the Quality of Life (QoL) of different travellers' lives across Europe and will offer guidelines and strategies for successful real-life deployment of the project's IoT solutions to Europe and the world.

7.2 Concepts and methods

Michon [28] has defined three levels for the driving task: strategic, tactical and operational level. The strategic (planning) level includes high-level decisions on travelling, such as choice of travel mode and timing of trip. The tactical level (manoeuvring) includes interaction with other road users, turning, overtaking and similar decisions when already on the trip. The operational (control) level includes use of gas and brake, shifting gears and steering.

In the context of AUTOPILOT, the strategic level relates to personal mobility choices. Strategic decisions influence the whole trip and are therefore typically most critical for wide impacts. The operational level, which includes control of the vehicle in concrete driving situations, as well as parts of the tactical level are taken over by the vehicle in automated driving.

The trilateral working group on impact assessment in automated road transport has defined impact paths for automated driving [24]. This framework is used as a basis for mapping the potential impacts of AUTOPILOT functions (Figure 10). Based on the preliminary function descriptions, the factors on which IoT is expected to have an effect are highlighted in yellow. The factors and impact paths will be updated in course of the project, and potential new factors enabled by IoT will be explored.

It can be seen in Figure 10 that the most impacts of IoT on automated driving are expected from vehicle operations: IoT enables connectivity between the vehicles and more precise road conditions and hazard monitoring and optimisation of speed, platoon planning and fleet use.

Changes on the driver/traveller level may occur, if IoT improves the travel experience in a way, which further encourages their use. In all, the figure shows that IoT does have potential in further improving the quality of life provided by automated driving.

It should be noted that in AUTOPILOT the term “Quality of life” is used in a broader sense than in the trilateral working group. It covers safety, network efficiency, emissions, public health, personal mobility and equity impacts, which are separate impact areas in the trilateral group.

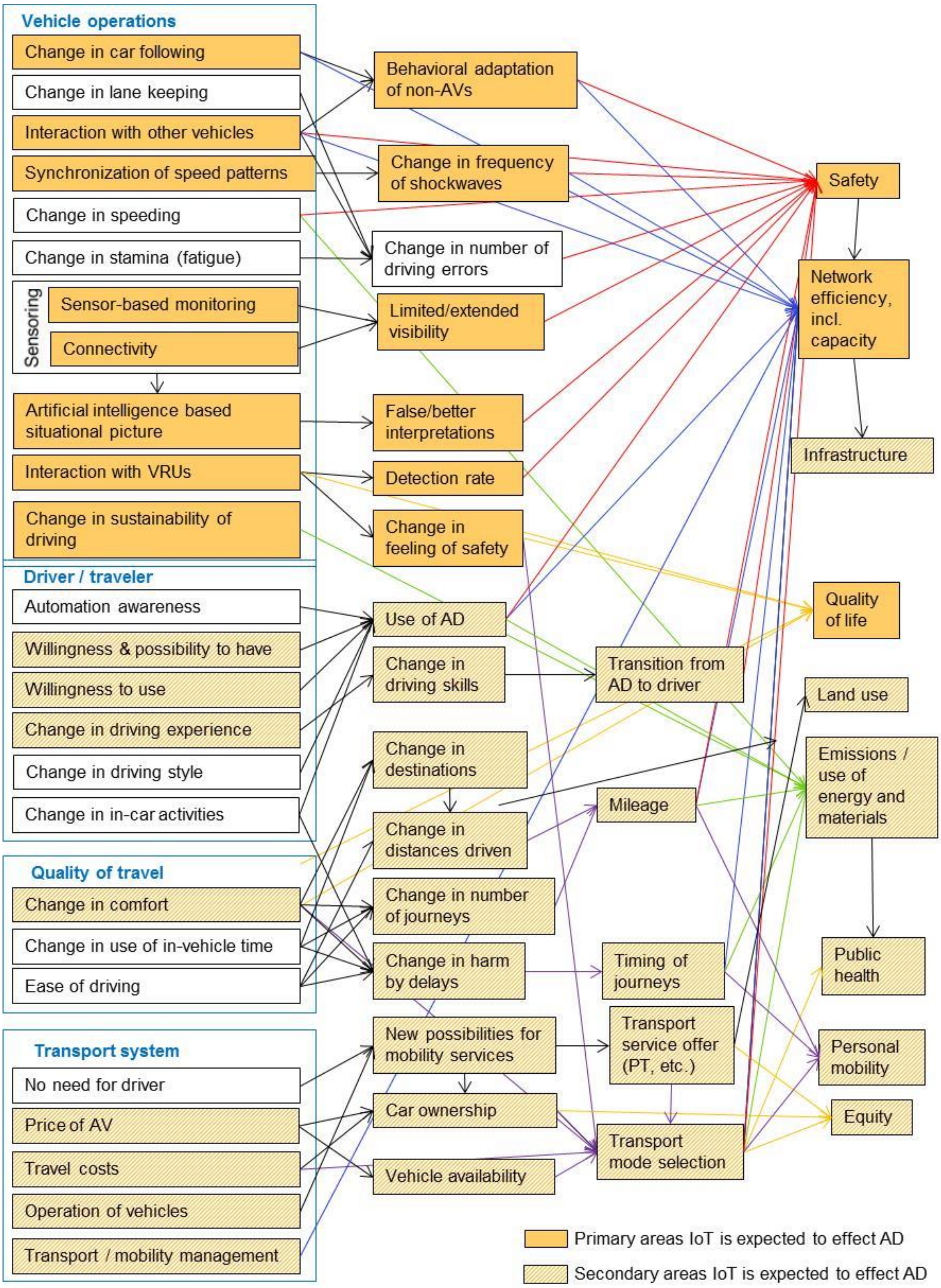

Figure 10 Impact mechanisms for AD and how IoT is expected to affect them (adapted after[24])

The concepts of personal mobility, sustainability and wellbeing are described in the following sub-sections.

7.2.1 Personal mobility

Mobility is typically defined as the potential for movement, in which the realised movement happens. Because this potential is difficult to measure, typically revealed mobility (including amount of travel, travel patterns etc.) is used as an imperfect measure for it [11]. The mobility model (see Figure 11) is a theoretical tool built on scientific literature and specialist interviews to identify the relevant factors and variables related to both potential and revealed mobility.

The model consists of the three main pillars of mobility: amount of travel, travel patterns and journey quality, which are further divided into more specific branches of elements. It specifies personal variables affecting mobility, travel decision-making variables, travel characteristics followed by decisions and their relationships. The tool can be used as a basis in assessing mobility impacts: it helps to clarify which aspects specifically need to be measured or evaluated in order to analyse mobility impacts. Furthermore, the tool can be used in analysing data. The tool was originally developed for the TeleFOT [11], project and used e.g. in the projects DriveC2X [31] and TEAM [32]. It will be enhanced to cover IoT and automated driving and will be used as a basis in the mobility assessment in AUTOPILOT.

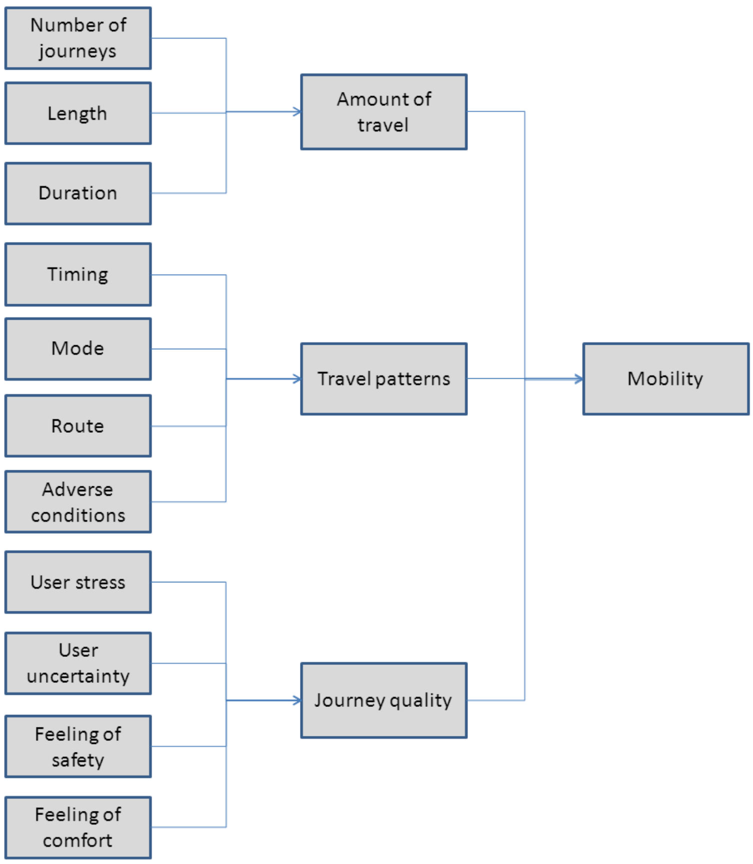

Figure 11: Mobility model for impact assessment (from [11])

The AUTOPILOT functions have potential to influence their users' mobility behaviour in many ways. Users may change the timing of journeys, their destinations, the mode of transport, and the routes selected. The impacts on personal mobility will be assessed in terms of participation in transport, access to destinations, travel times, time lost due to congestion, modes of transport, routes, autonomy, convenience and comfort. Impacts on value of time and travel costs will be assessed when applicable. Travel time reliability, multimodal connectivity and vehicle ownership models are also of interest.

Methods for data collection of the mobility KPIs will be defined. The mobility impacts will be studied through questionnaires and focus groups with test persons or people who have been familiarised with the pilots. For reliable results, the effects are studied by relating questions to the daily trips of the users, e.g. "do your typical trips include driving on motorway" or "is parking a problem in your frequent destinations".

Figure 12 provides a second conceptual framework for the mobility analysis. When analysing the impact of AD with IoT, special focus will lie on exploring the potential of the technology for improving QoL on individual and on societal level, but also on exploring potential conflicts between both levels. For instance, improving QoL on individual level might potentially cause traffic system related issues, such as an increase in VMT and emissions, when using individual motorised modes of transportation becomes more attractive than using other mobility options, such as cycling or walking. On the other hand, the technology might potentially bring improvements on traffic system level but not necessary have directly observable user benefits. Here might be some overlapping topics with the user acceptance evaluation (section 6).

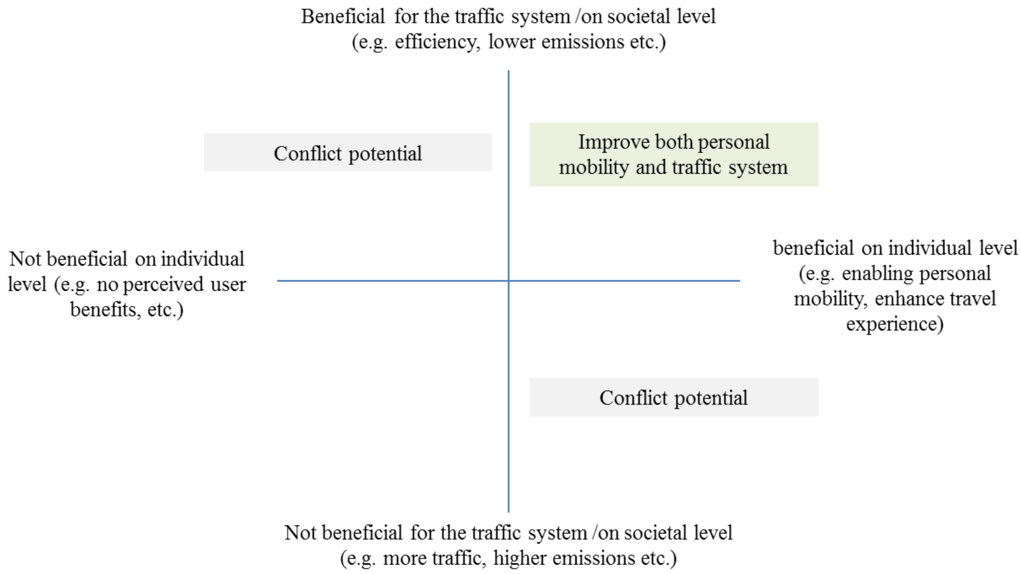

Figure 12: Dimensions for personal and social mobility [45]

As various use cases are considered, an overarching framework will ensure that similar aspects are addressed for all cases but also use case specific characteristics are included. Previous empirical works in the field of acceptance of automated driving have stressed the importance of addressing the specific characteristics of an use case rather than addressing too broad a view on automated driving (see [34][35]).

7.2.2 Sustainability

Impacts on sustainability will be evaluated in terms of safe, green and efficient transport. This includes impacts on traffic safety, efficiency and the environment. These impact areas are introduced below.

7.2.2.1 Safety Impacts

The safety impact assessment approach is based on system nature of transport: when one element of the system is affected, the consequences may appear in several elements and levels of the system, both immediately and in the long term, due to behavioural modification. The assessment follows the generally accepted theoretical background according to which the traffic safety consists of three dimensions, which are (1) exposure, (2) risk of an accident to take place during a trip and (3) consequences (= risk of an accident to result in injuries or death) (Figure 13). These are the three relevant aspects to cover traffic safety, and traffic safety is regarded as a multiplication of these three orthogonal factors [26].

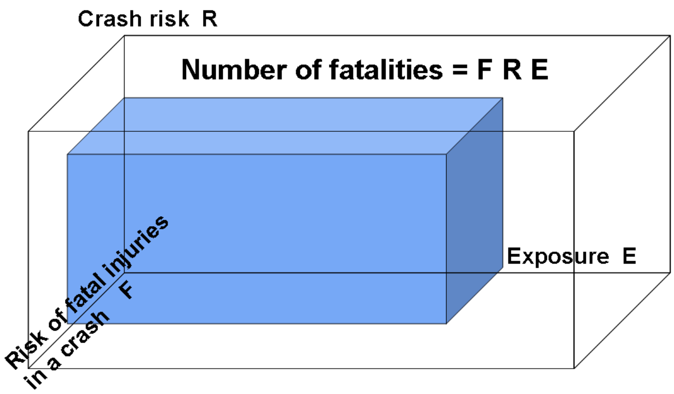

Figure 13. The dimensions of road safety ([26]).

When evaluating safety impacts the challenge is that actual crashes are rare events and proxy measures usually have to be used. Such measures include studying traffic violations, instances where a human driver must take control of the vehicle, exposure to near-crash situations and responses to near-crash situations. Situations reported as uncomfortable or risky by the passengers are also relevant. The assessment will systematically cover the intended and unintended, direct and indirect as well as short-term and long-term impacts of automated vehicle (AV) users and non-users. This analysis covers the three main factors of traffic safety by nine basic impact mechanisms [13], by which ITS can affect road user behaviour and therefore traffic safety:

1. Direct modification of the driving task, drive behaviour or travel experience
2. Direct influence by physical and/or digital infrastructure
3. Indirect modification of AV user behaviour
4. Indirect modification of non-user behaviour
5. Modification of interaction between AVs and other road-users
6. Modification of exposure / amount of travel
7. Modification of modal choice
8. Modification of route choice
9. Modification of consequences due to different vehicle design

The mechanisms were originally formulated by Kulmala [13] for evaluation of ITS systems, adapted from the mechanisms formulated by Draskóczy et al. [12]. The mechanisms 1 to 5 cover the risk of accidents. Mechanism 2 covers the IoT aspects of AUTOPILOT. Indirect modification of AV user behaviour (mechanism 3) refers e.g. to behavioural adaptation. Mechanism 4 covers effects where conventional drivers mimic automated cars, e.g. use too short headways when following vehicles in platoons.

Mechanisms 6 to 8 are related to exposure. Mechanism 9 refers to different vehicle design in automated vehicles (e.g. lighter cars).

The following assumptions are made:

- Safety increases as speed decreases (the so-called power model [26] which describes the relationship between relative mean speed effects and injury accidents)
- Safety increases as standard deviation of speed decreases
- Safety increases as jerk decreases
- Safety increases as speed violations decrease
- Safety increases as following very close decreases (manual driving)
- Safety increases as lateral position is more stable
- Safety increases as vulnerable road users are taken into consideration
- Safety increases as signals are used correctly
- Safety increases as driver condition is not deteriorated (manual driving)
- Safety increases as focus of attention is allocated correctly (manual driving)

The traffic safety impacts will be studied with data collected from the pilot tests, simulation and user questionnaires or focus groups.

7.2.2.2 Efficiency of traffic flow

Traffic efficiency describes how efficiently (in terms of average speed and travel time, number of stops, delay) people and goods can move through the transport network. The primary objectives to study under this topic are whether and to what extent the ADF and IoT have an effect on roadway capacity and traffic flow.

The number of vehicles passing through a cross-section of a road in a certain time constitutes traffic flow (also called traffic volume or throughput). The capacity of a road is defined by the maximum traffic flow, i.e. the maximum number of vehicles that can pass by a point on the road in a period of time (e.g. 1 hour). The capacity is influenced by several factors and their interactions: environmental factors, such as the layout of the road or weather conditions, vehicle factors such as length of vehicles and vehicle composition, and driver behavioural factors, such as preferred safety distance and driver state (see Figure 14). Changes in traffic flow depend on the penetration rate but also on the regulations regarding car following behaviour.

By providing enhanced performance through connectivity and better anticipation of unforeseen events, IoT is expected to affect some of the aspects of capacity and traffic flow and thereby enable, enhance or accelerate perceivable benefits of AV. The areas where improvement is expected at the start of the project are shown in Figure 14. They will be tested and developed during the course of the project.

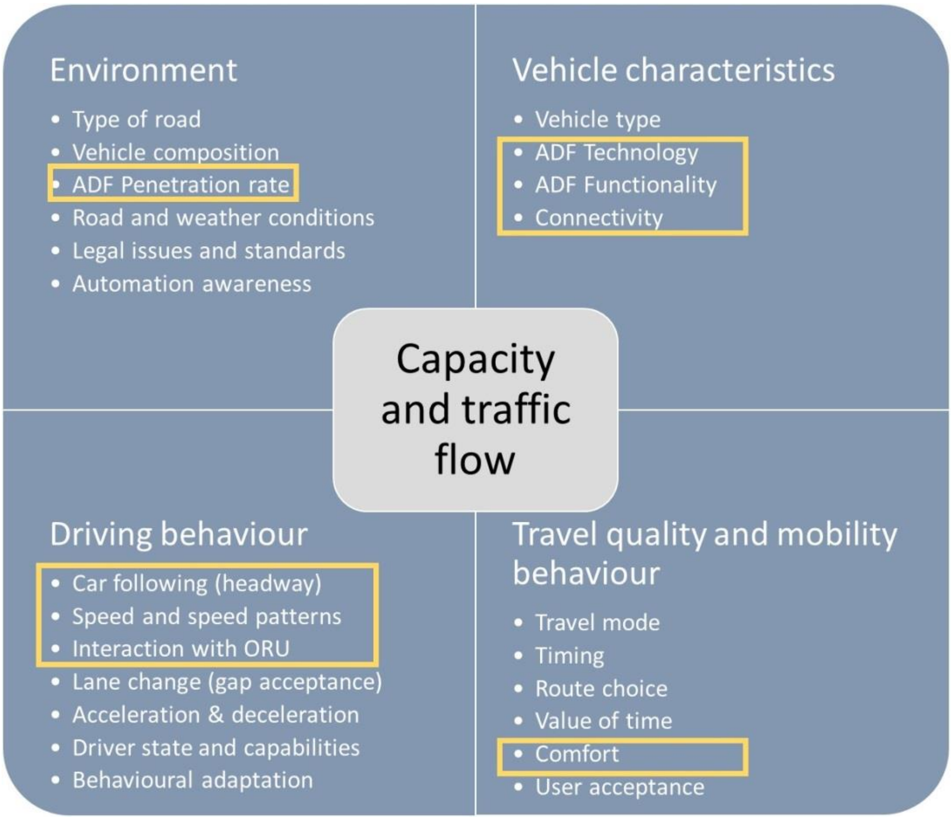

Figure 14. Factors of automated driving affecting capacity and traffic flow (The areas where IoT is expected to play an important role are highlighted with yellow borders).

Methods and tools include direct analysis of the pilot site data and indirect evaluation using traffic microsimulation software. Actual data from pilot sites will be used where available otherwise values found in literature will be applied.

Changes at the tactical driving level, such as changes in speed, acceleration and deceleration as well as headway to other vehicles and gap acceptance affect efficiency of traffic flow. Thus, traffic flow is affected by vehicle operational factors. These are tasks that have typically been carried out by human drivers, but that are being more and more taken over by automated driving functions: acceleration, deceleration, lane keeping, car following, lane changing, gap acceptance [24].

Top down research so far provides mixed conclusions on the effects of automated driving on roadway capacity. Studies indicate a range from reducing capacity over little change to large increases [27]. The results largely depend on the initial assumptions made. Traffic efficiency is dependent on many different but interrelated factors, and the development of those factors is uncertain.

According to Milakis et al. [25], the benefits of automated driving on traffic flow efficiency are highly dependent on the following factors:

- level of automation
- connectivity between vehicles
- penetration rates
- deployment path (dedicated lanes or integrated in mixed traffic)
- human factors (behavioural adaptation)
- change in demand (increased demand possible)

The long term implications are uncertain and largely depend on the development of travel demand as well as the business models adopted: whether the vehicles will be in personal or shared use.

There are many ways, in which automated driving functions may affect driving behaviour. Gaps in knowledge about the potential impacts of connected automated vehicles (CAV) exist as well as a lot of uncertainty around the functionalities [27]. In addition, automated driving functions are expected to change the characteristics of and interactions between the impact areas safety, comfort and efficiency. For example, due to shorter reaction times, automated vehicles may drive with shorter headways. On the other hand, due to safety or comfort reasons, the headways and gaps may also be greater than currently with human drivers. It therefore seems likely that some trade-offs will have to be made between those impact areas.

First studies suggest that in the initial phase, with a low penetration rate of AV, road capacity may decrease due to AV behaving cautiously in the presence of human driven vehicles [25][27].

For the sake of user acceptance of AV and to maximise safety, the vehicles may first be configured to use larger headways and lower acceleration rates than average human drivers do. This has negative effects on roadway capacity. Thus, a vehicle that is designed for the comfort of its users (allowing to use the travel time for other activities) can have unintended impacts on roadway capacity for all road users [27].

It is not yet clear how the parameters used by the ADF will be defined in the long run: by OEM, user preferences, legislation, or some combination of these. Also unclear is how users are willing to trade an increase in travel time against the ability to better use the time [27].

The interaction of automated vehicles and manually driven vehicles is an important topic of interest. Automated vehicles will be introduced gradually into the current transport system, and there will be a period with mixed traffic. AV are expected to confront difficulties in the conventional traffic system in several ways [29]:

- The anticipatory capability of AV is not as good as human capabilities. AV act mainly reactively rather than proactively.
- The behavioural recognition of AV is limited (how to recognise and react to different traffic situations).
- AVs have limited flexibility.
- AVs miss human courtesy and are non-sociable.
- Issue of equality compared to conventional vehicles: should AV and human driven vehicles be treated the same?

On the other hand, also human drivers face challenges with automation of vehicles. These include from [30]:

- overreliance
- behavioural adaptation
- erratic mental workload
- skill degradation
- reduced situation awareness
- inadequate mental model of automation functioning

Several studies suggest that connectivity and cooperation between vehicles and vehicles and infrastructure is essential for achieving benefits for traffic flow. It is expected that traffic throughput will suffer if AV are introduced before sufficient implementation of connectivity [29].

Even with connectivity, a significant penetration rate of CAV (about 40 %) has to be achieved for significant impacts to be seen [25][27]. With a relatively high share of CAV, traffic flow can be improved and shockwave impacts dampened. Unconnected automated functions, such as ACC, can increase capacity if users accept time gap lower than those currently chosen [27], which is not considered likely. The approach in AUTOPILOT, which adds IoT to automated driving, therefore seems promising as it provides possibilities for extensive communication between vehicles and their surroundings.

Traffic efficiency is also linked to environmental impacts: smoother traffic flow with less variance in speed/acceleration and fewer braking actions means also less exhaust emissions (see e.g. [36][33][25]).

7.2.2.3 Environment

The impacts of automated driving on environment depend considerably on travel, driver and vehicle behaviour. Effects on the environment (CO2 emissions) will be derived from the efficiency results (VISSIM and its emission calculation tool Enviver), as emissions are directly related to fuel consumption.

The factors, which are expected to play a role in the impacts of AD and IoT on emissions, noise and surroundings are shown in Figure 15. Advances of technology, different engines etc. will be taken into account where possible.

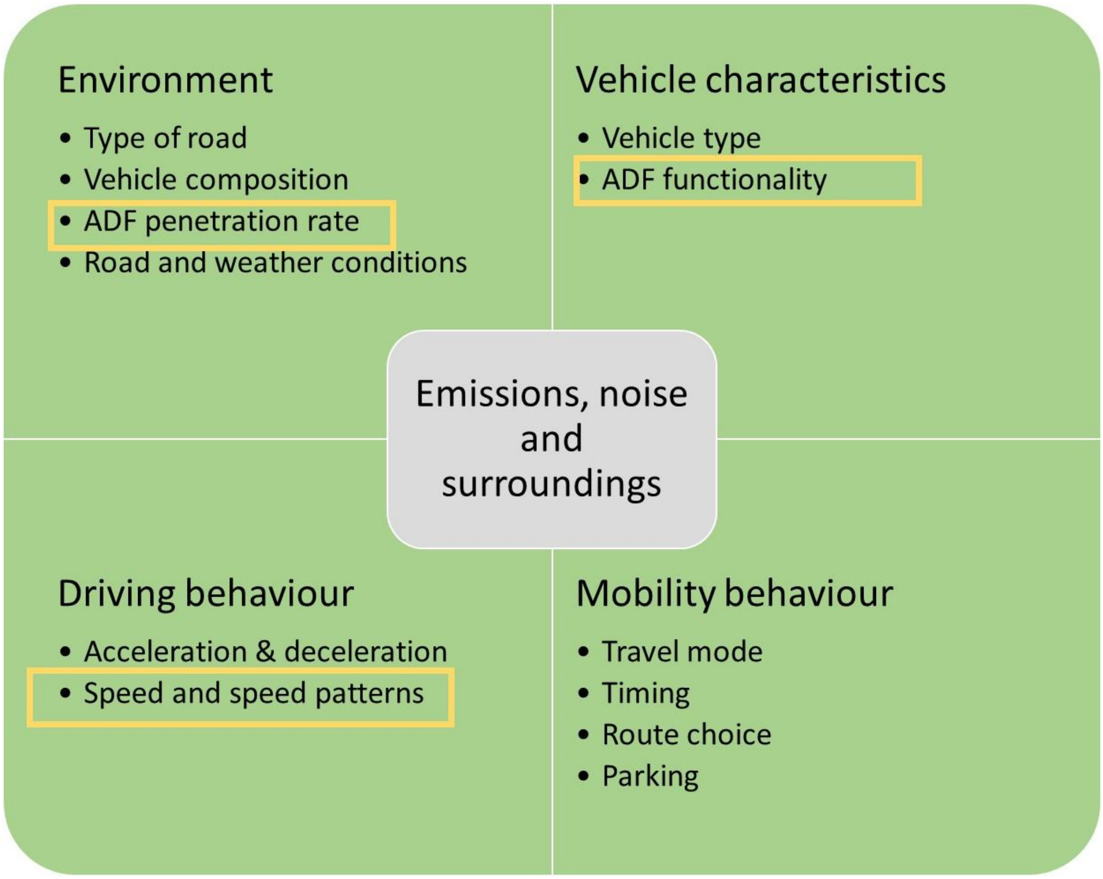

Figure 15 Factors of automated driving affecting the environment (The areas where IoT is expected to play an important role are highlighted with yellow borders).

Noise is a multifactorial effect and depends on the type of road (pavement), type of vehicles, tyres, driving condition (steady state, acceleration). It has been assessed that the impact on noise to be considered based on:

- Number of journeys: there is a negative/positive impact on noise when the number of journeys increases/decreases. Essentially each avoided journey means a reduction of noise.
- Route: changing the route may reduce the negative effect of noise if there is a shift of traffic from rural and city roads to highway. This is not an absolute change but a qualitative one because normally highways are located in areas far enough from residential areas and in many places there are noise barriers when the distance is too low.
- Speed: there is a direct relationship between speed and noise. Noise increases when the speed increases. There is a wide variability and no precise equations which could be valid for any vehicle. However, FHWA developed a Traffic Noise Model [47] which considers speed as a parameter.

Despite the great results achieved in the last years in emission reduction vehicles are still pollutant. There are several factors which influence emissions depending on travel and driver behaviour.

- Number of journeys: there is a direct link between number of journeys and emissions.
- Travel mode: there is a positive effect by using multimodal transport instead of using car only.
- Length/duration: there is a proportional link between travel duration/length and emissions.

- Time budget/ timing: changing travelling time may mean different traffic conditions encountered with impact on emissions.
- Route: changing route may influence emissions. For example the opportunity to use an urban shorter road without traffic jam supported by an intelligent traffic light system may be a positive alternative to a longer high speed road.
- Speed: The relationship between speed and emissions is different for the different pollutants. Nevertheless, considering a vehicle running at constant speed, there is a first range in which the vehicle has high emissions, followed by a range in which the emissions increase in a smooth way and then the emission curve vs. speed become steeper due to aerodynamic effects. On the contrary in real traffic there are many speed variations and the steeper and more frequent the variations are the higher are emissions. For this reason, in order to evaluate the impact due to a warning message it is important to consider not only the mean speed but also the standard deviation or make a comparison of the complete speed profile.
- Pedals: use of pedals is important for emissions. Rapid changes normally imply high emissions.
- Fuel Consumption: there is a linear relationship between Consumption and CO2 emissions. More complex is the relationship with other pollutants which depends on the particular transient condition encountered.

The surroundings are related to environmental impacts via expose. For this reasons research questions related to the road change (urban, rural, highway and residential area) are included.

7.2.3 Well-being

Data for KPI analysis will be collected from the pilot sites and impacts on well-being will be analysed, such as impacts on health and equality, measured with quality of life KPIs.

It is recognised that impaired or mobility-restrained individuals could specifically benefit from automated driving and IoT in terms of access to services, health, reduction of potential social isolation, etc. (see e.g. [33][40]). Therefore, the existing standardised Quality of Life index (e.g. SF 12 and WHOQOL) will be reviewed and adapted, and selected objective and subjective aspects of QoL will be addressed for the specific user groups.

These questionnaires cover different areas of the users' (daily) life and the aim is to identify the ones that will be impacted most by the project's outputs, as the entire process is complex and, by all means, bi-directional (i.e. the person affects the use of the technologies by either being an early or late adopter as much as the technology improves their lives).

QoL assessment will be assessed before any intervention (i.e. riding in the automated vehicle and then after the end of the pilots to assess the QoL after driving experience with the automated vehicle). The measurement of the difference between these two conditions will allow for effect size calculations and, consequently, lead to the actual impact assessment on the QoL of drivers across Europe for diverse and different user groups. The calculations will be based on pre- and post-self-assessment of QoL constructs/ indicators relevant to mobility and driving experience. The user groups addressed may include people that cannot drive or have ceased to drive because of health conditions and other isolated user groups. These groups will not be included in the study but the indirect impact to these groups will be investigated through the responses of the drivers (i.e. direct participants/ users). The well-being of the actual users is of primary importance of the project but the indirect effects to the well-being of other groups could be further explored. The following aspects of QoL will be addressed:

- behavioural competence (e.g. measured by indicators of health, cognition, time use and social behaviour);
- perceived Quality of Life (QoL) (e.g. individual's own perspective);
- psychological well-being (e.g. satisfaction with tested solutions) , positive and negative affect in their lives because of IoT solutions, mental health state, etc.);
- external and physical environment (e.g. self-living rate and economic indicators, car ownership, occupational and socio-economic status (SES)).
- relation of QoL constructs to attitude towards self-driving vehicles, ICT competency, trust and fear as well as other background and demographic variables. This aspect involves the correlation of QoL question items with the participant's pre-conceptions and attitudes towards New Technologies, their use, their penetration in our everyday life activities, their trust of New Technologies and if they fear them (or not). In a nutshell, we will further investigate if these pre-conceptions will affect their acceptance of self-driving vehicles as well as the acceptance of the potential positive effects to their well-being. For example, if a user does not trust or fear New Technologies, it will more difficult to use New Technologies and even more difficult to accept that the use of New Technologies can improve/ affect their well-being.

7.3 Research Questions, Hypotheses and Indicators

The overarching research question is whether IoT brought to automated driving can improve people's quality of life. This question will be studied via the following main research questions and associated topics:

1. How well does IoT in automated driving meet personal mobility needs?
   - Time spent travelling, amount of travel
   - Travel quality incl. perceived travel quality (from the user point of view)
   - Mode of travel
   - Routes
   - Accessibility of locations
   - Health
   - Are there differences between personal and society goals?
   - Does IoT facilitate new modes of travel?
   - Value of time, accepted travel times compared to current situation
2. Are improvements to be expected in transport system efficiency with various penetration rates of IoT and automated driving?
   - Road capacity: speed distribution, headway distribution
   - Interaction between road users (perception of intentions)
3. Can IoT contribute to traffic safety improvements with automated driving?
   - Conflict situations
   - Accident types

From a technical point of view IoT can accelerate, enhance or enable automated driving. Following these main hypotheses, use cases of automation with IoT can accelerate, enhance or enable mobility. These can be seen as different layers of improvement of the personal mobility on the one hand and on transport system on the other.

7.3.1.1 Personal mobility

IoT can (but not necessarily will) change mobility on an individual level in one or in all of the three domains. In the following, the hypotheses on how IoT might improve personal mobility (or the factors influencing mobility behaviour) are summarised following the three categories.

Urban valet parking:

Accelerate: The technology will certainly reduce access and egress time; it is however not clear yet what the perceived usefulness is from the user perspective; an exception might be people with physical constraints who can get out of the car right in front of the door. However, it should be noted that reducing access/egress time using Automated Valet Parking is not necessary related to IoT but more to automated driving. Furthermore this particular use case could bring improvements primarily for parking space and efficiency. Another considerable improvement might be reducing parking costs (due to system and space improvements).

Enhance: On an individual level, providing information or other features using IoT to users of the technology might enhance the user experience. Thus, user requirements on the functions of Automated Valet Parking are addressed in the user acceptance part of the evaluation, but enhancements in user experience have potential to lead to changes in mobility behaviour.

Enable: Depending on the target groups, people with reduced mobility might be enabled to have direct door-to-door access to certain services or places. This is however also more related to automated driving and Automated Valet Parking than to IoT.

Urban driving and car sharing:

Accelerate: Travel time improvements are possible. Also, access to a vehicle on demand can be accelerated with IoT. Support for on-street parking using IoT may improve travel time.

Enhance: User experience is enhanced and changes can occur in the valuation of time spent in the vehicle due to connectivity functions, i.e. IoT. This applies especially to certain user groups, such as tourists. Thus, IoT and automation can be seen as a complete user experience “package”. Parking time can be reduced through more efficient route planning when the actual parking space situation from IoT data is taken into account.

Enable: Some user groups can get the possibility to use a vehicle on demand instead of their own car due to IoT applications. Similarly to Automated Valet Parking, fully automated driving with IoT might enable new user groups to have access to individual motorised modes of transportation.

Highway Pilot:

Accelerate: Speed and respectively trip time can increase. Efficiency not only for the system but also for the user of the system. Also, improving traffic flow efficiency on the highway and preventing traffic accidents will potentially reduce travel time (by preventing congestion).

Enhance: Comfort and traffic flow may be enhanced due to enhanced detection of obstacles, potholes etc.

Enable: In partial automation, the technology might enable more productive or pleasurable use of time: people can engage in other activities, such as reading or texting, as IoT can warn drivers much earlier to take over control of the driving task; AD without IoT will require permanent supervision of the system since the driver might only have few seconds to take over the driving task.

Platooning:

Accelerate: Not directly applicable except for improvements mentioned above for highway pilot. Platooning enabled by IoT can bring improvements on a traffic system level (rather than on individual level).

Enhance: The formation of platoons can be enhanced by IoT. In addition, user experience and value of time spent in the vehicle when platooning vehicles enable people to be engaged in other activities than driving (only the first vehicle of the platoon has to focus on the driving task).

7.3.1.2 Sustainability

Sustainability impacts include improvements in the transport network efficiency, safety and emissions, but also related fields such as city planning, land use, social equity, etc.

Urban valet parking:

Accelerate: The parking time and efficiency of the parking process (such as quicker finding of a free parking place, better parking manoeuvres etc.); parking space use since the vehicle might park also on a parking place far away from the entrance; reduction of emissions in urban areas due to the reduction of searching for on-street parking spaces.

Enhance: Not applicable.

Enable: Changing shape, size, infrastructure of parking places (improving land use); impact on city planning.

Urban driving and Carsharing:

Accelerate: Traffic flow efficiency and emissions by reducing accelerating and breaking when information of traffic light system is available on time; provide more safety by VRU detection; first/ last mile solutions.

Enhance: Not applicable.

Enable: New user groups to use a vehicle; new mobility options and business fields for mobility providers (to be addressed in the business impact evaluation).

Highway Pilot:

Accelerate: Safety by automated driving on the highway (also on lower level of automation since more information on the traffic situation is available much earlier; pre-warning possible.

Enhance: Not applicable.

Enable: Better and more efficient road management by getting information from vehicle and infrastructure on time and sending this information to an IT back-end.

Platooning:

Accelerate: More efficient traffic flow when platooning; improving driving by getting on time information from the lead vehicle.

Enhance: Not applicable; or: formation of platoons can be enhanced by IoT.

Enable: Getting on time information from the lead vehicle.

7.4 Piloting scenarios

As the impacts of automated driving and IoT are numerous, uncertain and difficult to assess, a case-based approach is planned in the assessment. By relating the use cases and functions to their own daily lives, it will be easier for the users to think about potential changes in their behaviour. Specific (mobility) needs and expectation have to be addressed. Also, to which extend the use case addresses, i.e. satisfied, them. There is overlapping in this research topic/field with the user acceptance evaluation. Thus, exchange with and input from the user acceptance task is recommended/ required.

Test users will be asked to view the functions and use cases through their daily mobility habits. A detailed look at their own daily mobility behaviour could be provided by filling out a travel diary for a period of time before taking part in the pilot tests. Afterwards, users are asked to reflect the experiences they gained from the pilot tests to their habitual behaviour and think about potential changes based on the test situations. They will be asked to think how their daily mobility choices could change and what kinds of trips would be affected. Survey answers can be complemented with deeper interviews or focus group discussions.

The opinions of professional drivers, or test drivers of the vehicles, are also of importance: for example they could be asked how they would have behaved in certain traffic situations, e.g. comparing the behaviour of the automated vehicle when encountering vulnerable road users (VRU) in various situations. Would you have behaved differently, slowed down or communicated with the VRU in some way? This group would represent a reference group of experienced drivers.

Observers on the pilot sites could be used to log encounters and conflicts between AV and VRU.

The timing of the questionnaires is important to get valid results. Users should be asked about their experiences with different situations directly after encountering them.

The detailed approach will be elaborated in the work of Task 4.4, and detailed guidelines for the pilot sites will be provided before the user tests start.

7.5 Data collection and analysis

Following data needs to be collected for analysis and evaluation:

- Speed
- Position (latitude and longitude)
- Acceleration, deceleration and rate
- What IoT info was received & used?
- Weather/road conditions
- Fuel consumption
- No of harsh braking
- Distance to vehicle in front
- Look ahead/back distance
- Following variation
- Test user (passenger) questionnaires
- Standardised QoL questionnaires
- focus group discussions

8 Business Impact Assessment Methodology

This section presents the three methodologies that will be applied for the business impact assessment to be conducted in Task 4.3. The approach to define the research questions, hypotheses and indicators is presented as well as guidelines for data collection. The section concludes with the expected output to the business exploitation Task 5.3.

8.1 Objectives and Methodologies

The main objectives for assessing the business impact are to:

- Evaluate the cost-benefit and cost-effectiveness of the AUTOPILOT exploitable results, i.e. the IoT accelerated, enhanced or enabled automated driving systems.
- Evaluate the impact of exploitable results to the market in terms of creating new products and customers, and establishing a new stakeholder ecosystem.

Following subsections briefly discuss the three methodologies that will be used for the business impact assessment in Task 4.3. These will be further described later in deliverable D4.5 “Business Impact Assessment”.

8.1.1 Methodology for Cost-Benefit Analysis (CBA)

Cost-benefit analysis (CBA) on a societal level is based on welfare economics where resource savings make up benefits because of the assumption that they could be used in other parts of the society with at least the same productivity. It is used as a measure to compare the costs of a certain (intelligent transport) investment with the resource savings or benefits to be achieved over a given period of time, for example through increased traffic safety, improved efficiency and mobility reduced travel times, and reduced fuel consumption and pollution. The CBA will use the results of safety, environmental and efficiency impact assessment carried out in Task 4.4 (see section 0). Those impacts can usually be monetised by using standardised cost-unit rates [46].

In a CBA two alternative future scenarios are compared:

- Business-as-usual-scenario which assumes that no services are implemented (without AUTOPILOT)
- With-project-scenario, where the services are implemented (with AUTOPILOT)

The costs of the services must be estimated in terms of investment costs, operation costs and maintenance costs. One or several target years are usually considered, and a benefit-cost ratio is calculated for these target years. The costs are then compared to the target year benefits.

The CBA has five steps:

- Defining scenarios
- Identifying relevant parameters (including target years for the analysis
- Quantifying impacts
- Applying monetary values
- Calculating BCR (benefit-cost-ratio)

As AUTOPILOT is a pilot project and the automated driving functions will be closer to prototypes than final products, estimation of these costs is challenging. The data to be gathered in Market Analysis provides valuable input also for the CBA analysis. If it is not possible to estimate monetary values, the analysis will be carried out qualitatively, taking into account "Go-to-Market" behaviour of stakeholders in the Internet, Automotive OEM and Mobile Communication market by analysing business models of these sectors in recent years. Business model canvases will work as basis for the analysis in case more detailed business plan descriptions are not yet available.

8.1.2 Methodology for Market Analysis

For an evaluation of market analysis, it is first of all necessary to define the means of the term market analysis for the highly innovative market of "IoT assisted Automated Driving".

This complex new market arises from the overlap of several industry lines, mainly automotive (OEM) and suppliers, IT and telecommunication industry including I.T.S. as well as the very dynamic mobility service industry. For all these industries, the strengths, weaknesses, opportunities and threats (SWOT) of a company under examination must be identified. Only with the help of a profound SWOT analysis, adequate business strategies for the industry and service lines mentioned can be derived leading to profitable and sustainable business activity.

Market segmentation is the basis for a differentiated market analysis and differentiation is most important as there is a saturation of consumption bringing increasing competition in offered products. Especially when it comes to the automotive market consumers ask for more individual products and services associated with strong automotive brands. Consequently, market segmentation is necessary including market research and in-depth market knowledge.

Market research about market structures and processes must be done to define the "relevant market". The relevant market is an integral part of the whole market, on which the company focuses its activities. To identify and classify the relevant market, a market classification or segmentation must be done. Nevertheless, it must be mentioned that there is no specific way to segment markets. However, businesses can follow generalised rules like geographic, demographic, psychographic, and behavioural. Good market segmentation should be sustainable, accessible, actionable, measurable, and differentiable [17]. The market can be structured in the following sub-topics:

- Market size (current and future)
- Market trends
- Market growth rate
- Market profitability
- Industry cost structure
- Distribution channels
- Key success factors
- Key success details

Market analysis strives to determine the attractiveness of a IoT based AD products and their market share, currently and in the future. Organisations active in the field of Automated Driving should be able to evaluate such future attractiveness of AD market by understanding evolving opportunities, and threats linked to the organisation's own strengths and weaknesses.

For AUTOPILOT, the implications for business models will be studied along the pilot site stakeholder eco-system and its use cases focusing on:

- Predictions for system uptake
- User expectations for service quality and availability

Pricing models in the extreme price sensitive market of mobility services (see [22]).

8.1.3 Methodology for Multi-Criteria Analysis (MCA)

The framework for multicriteria analysis as defined within ADVISORS (GRD1-1999-10047) project [17] will be adapted for use in the AUTOPILOT project. The multicriteria analysis (MCA) starts with the construction of the so-called evaluation matrix or table and then continues with the aggregation of the information contained in it. The MCA finally yields a ranking of the application areas.

The evaluation table forms the input for the synthetic phase of the multicriteria analysis framework. Generally, this table can be visualised as indicated in Table 5. Each alternative (a_1), is evaluated on each criterion (c_i). The result of each of these partial evaluations is represented in the table by "e".

Table 5: Evaluation table (general case).

Alternative	c_1	c_2	...	c_i	...	c_m
a_1	e_{11}	e_{12}	...	e_{1i}	...	e_{1m}
a_2	e_{21}	e_{22}	...	e_{2i}	...	e_{2m}
...	...	...	...	...	...	...
a_l	e_{l1}	e_{l2}	...	e_{li}	...	e_{lm}
...	...	...	...	...	...	...
a_n	e_{n1}	e_{n2}	...	e_{ni}	...	e_{nm}

Whereby: c_i = a criterion (expected impact, e.g. increased safety) (i = 1, …, m);
m = the total number of criteria;
a_l = an alternative (i.e. application area, in our case the different scenarios deriving from the use cases) (l = 1, …, n);
n = the total number of alternatives (application areas);
e_{li} = the evaluation of alternative application area (l) on numerical criterion i.

When the criteria included in the evaluation table above are constructed for each level of analysis, it is possible to arrange them into different groups so that each specific group corresponds to the objectives of a specific level. Within the evaluation matrix, however, clusters of criteria may be distinguished. One cluster may be related to the effects that can be expressed in monetary units; another cluster may be related to non-monetary units such as safety effects, etc.

All objectives pursued do not have the same importance therefore the criteria included in the evaluation matrix should be weighted. A widely used method for determining weights is the pairwise method, which is used in the Analytic Hierarchy Process (AHP) [18]. In the pairwise method, criteria are compared in pairs. For each pair, the decision maker must state whether the first criterion is as important as the second one or whether the dominance in terms of importance of the first over the second criterion is moderate, strong or "complete".

This information gathered from the corresponding templates is then transformed into a numeric scale. The relative priorities or weights are calculated using the eigenvector method. Since several pairwise comparisons are redundant the overall consistency of the pairwise comparisons can be determined. The implied meaning of weight in the standard AHP procedure is the relative value attached to the scores on the different criteria.

The analytical hierarchy process (AHP) [19] is probably the most widely known and widely used MCA method in decision-making. The AHP method is based on three principles, which form the subsequent steps of the method, namely: (1) construction of the hierarchy, (2) priority setting and (3) logical consistency.

The implementation of this methodology is planned in the form of dedicated workshops, to be organised in each of the pilot sites, in order to secure the participation of related stakeholders (see section 8.4). The first step is to choose the criteria and the scenarios to be used for the evaluation. This first step can also be done remotely. The criteria will be chosen in cooperation with the pilot sites and stakeholders, and the scenarios are derived from the AUTOPILOT use cases as well as input from WP5. During the workshops, participants will be asked first to provide weights to the different criteria and then to evaluate each of the alternatives (scenarios) according to each of the provided criteria. Finally, the feedback from the different evaluation workshops will be calculated by Task 4.3 partners and results will be provided overall and per stakeholder category. These will be further analysed, presented and concluded in D4.5.

8.2 Concepts

8.2.1 Costs – CAPEX/OPEX

Capital expenses (Capex) are used to create benefit in the future by making purchases for physical services or goods. Operating expenses (Opex) consist of the ongoing daily functioning costs of a system.

8.2.2 New products

New products summarise all type of products which can be linked to the definition "single product element within the total set of products targeted to the market of IoT assisted Automated Driving". Product elements might be purely service oriented, e.g. shuttle services, R&D, engineering or ICT services, but they might also include physical products (hardware) such as electronic components and innovative software modules implemented within the products. Considering the wide range of products and industries involved, the total set of products is still under examination, but as products form the basis of commercial activities, properly evaluation is needed to better understand the concerned product groups and elements.

8.2.3 Time to market – TTM

Here Time to Market (TTM) means the standard definition used in commerce. In commerce, TTM stands for the length of time it takes from a product being conceived by a company until its being available for sale. In automotive markets the product development phase is an important time factor as OEM products are outmoded quickly and suffer of heavy competition and market saturation. For ICT companies, TTM must take project staffing into consideration as stakeholders often suffer of delay due to missing project funds for experts needed to implement an innovative product. Compared to this, mobility service companies often must deal with problems of personal staffing and financial funds at the same time and therefore define TTM milestones with early adopters and prototype markets already generating revenues. It will be assumed that TTM for AUTOPILOT products matters most for first-of-a-kind products, where the leader has the luxury of time, while the clock is ticking for the followers.

8.2.4 Penetration rate

Penetration rate means the percentage of automated driving vehicles with IoT services that exist within the overall traffic mix. The concept of penetration rate in the business assessment is based on the assumption that the higher the penetration rate of AD vehicles with IoT the more efficient their operation will be; thus, they will be more acceptable by the users. This implies that the demand for such vehicles and their services will be higher, which broadens the market and increases the business opportunities.

8.2.5 New market & SWOT

As mentioned in section 8.1, complex new markets such as AUTOPILOT mostly arise from the overlap of several industries. In AUTOPILOT, these are mainly automotive (OEM) and suppliers, IT and telecommunication industry including I.T.S. combined with the very dynamic mobility service industry. For all these industries, the strengths, weaknesses, opportunities and threats (SWOT) of the specific industry must be identified. Only with the help of a profound SWOT analysis, adequate business strategies for the industry and service lines mentioned can be derived leading to profitable and sustainable business activity.

8.2.6 New customers

In this chapter, new customers for AUTOPILOT products will focus on existing automotive and mobility service customers as defined in the different pilot sites. All pilot sites refer to specific industries and outlined their target customers. From these customers some of them are existing customers, e.g. tourists for car sharing services, whereas some of them might be totally new customer groups, e.g. logistics service providers using AD car sharing and platooning services for urban logistics and delivery in congested shopping malls.

8.2.7 Benefits (safety, security, mobility, environment, etc.)

The benefits that the AD with IoT would imply to different areas (like safety, security, mobility, environment, etc.) are a crucial parameter to the business assessment, as they should be closely monitored and evaluated, to be used as primary axis for the products' deployment, exploitation, advertising, etc. as well as for defining the framework in which the relevant market should operate, in order to attract customers, create new business and overcome any negative consequences or criticism.

8.2.8 Research and innovation

Research and innovation are two parameters that are highly linked to business. New directions in research, new research topics, innovative concepts and products are only some of the issues that could affect business exploitation and market potential of AD (+IoT)

8.2.9 Automation level and the role of IoT on Automation

The level of automation and the role of IoT on its formulation is crucial to the business assessment, as the different levels of automation may imply different demand, different products and even different customers and, thus, different business models. Moreover, IoT may play a significant role in the formulation of the level of automation, as well as in its introduction to the market.

8.2.10 Stakeholder ecosystem

Here stakeholder analysis means the process of assessing a decision's impact on relevant parties. Stakeholder analysis will be used during the preparation phase of the business impact evaluation to assess the attitudes of the stakeholders regarding the potential changes along the time-to-market phase. In AUTOPILOT, stakeholder analysis will be done on a regular basis to track changes in stakeholder attitudes along the technical implementation of IoT assisted AD functions and services in the different pilot sites.

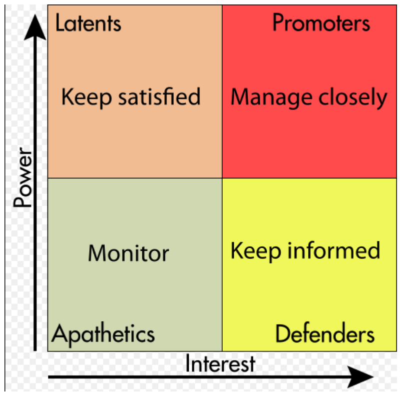

Figure 16 Stakeholder matrix showing the different behaviours towards new markets

8.3 Research questions, Hypotheses and Performance indicators

8.3.1 First-Level Research Questions

Given the definitions described in methodologies and concepts, the focus of business assessment should be put in defining the impact of AD and IoT in the business sector, within the framework of the adopted methodologies. The following first level research questions have been selected:

- What are the new business opportunities and threads (market resistance)?
- Is the demand for AD and/or IoT affected?
- Are the costs affected?
- Is the market affected?
- Is research affected?

8.3.2 Hypotheses

The five first level research questions are relevant for each of the concepts form section 8.2. At a second level, research questions are refined per concept for the first level questions (see Annex 14.1).

The research questions and hypotheses for new markets & SWOT and new customers show quite similar results to the concepts for time to market. Therefore, they are grouped to the list of time-to-market section in Annex 14.1. Especially, the enthusiastic users and ICT market promoters, well-known from Internet business, must be studied and managed carefully as their willingness is a sensitive success factor – of similar importance as is user acceptance and technical reliability.

There is another influence factor for market uptake related to the capability of a stakeholder to cooperate with other market segments, especially if a reluctant defender must cooperate with a fast-moving promoter. Here the experiences of traditional car dealers and high-tech companies promoting navigation and infotainment products can be mentioned as instructive example.

Whenever a car dealer must promote innovative electronic devices, he needs trained personal staff to explain the innovation to the end customer. Therefore, OEMs offer After Sales products within their internet portals and give customers access to these portals to enhance CRM. Drivers can select and choose new products, order them through different distribution channels, and have customer relation with the brand by maintenance of the vehicle.

This way, the dealer participates along the innovation path of new products nevertheless drivers are already aware of the product and know how to use it. More details are provided in Annex 14.1.

8.3.3 Indicators and Metrics

In Market Analysis, SWOT is used as fundamental tool to better understand opportunities and threats of new products appearing in the market. As all indicators (except costs) linked to SWOT are qualitative rather than quantitative, SWOT by itself is the adequate method to be used to create indicators and a metric setting up a relationship between these indicators. But to make the facts and figures of the SWOT analysis reliable, they must be linked to the different market segments (see also section (8.4.1, analysing user groups).

In AUTOPILOT, users cannot be separated from stakeholders as automotive industry goes along all type of users (large enterprises, mid-size companies and SME). The indicators and measures to execute data collection and analysis must be studied under consideration of:

1. Regulatory challenges necessary to overcome key markets (road safety and privacy);
2. Fully developed safe and reliable technical solutions (safety and privacy);
3. Enthusiastic consumers willing to pay for innovative AD services (affordability / ease of use).

This approach is also used in section 6 describing personal factors for user acceptance for safety, privacy, affordability, energy efficiency and ease of use.

It will be necessary to study the different pilot sites' services with the hypotheses listed in Annex 14.1 and underline the data collected from:

- Questionnaires
- SWOT
- Market Research
- Interviews

Due to the overall dependence on AD market deployment, the IoT aspect must be studied, additionally making users aware on interferences and differences of both aspects.

8.4 Data collection

8.4.1 Users

When taking a deeper look into the key users and potential customers which might become enthusiastic using AD functions one can group them into:

- Final users
- B2B users
- Stakeholders
  - Operators
  - SMEs
  - OEMs

Analysing technologies planned to be piloted in AUTOPILOT test sites, one can classify them into the following market segments:

1. In-Vehicle sensor technology for AD functions (OEM & suppliers)

2. External sensor technology shared via V2X or via IoT push mechanism (I.T.S. industry, ICT)
3. IoT based secure data centre technology (ICT)
4. IoT and M2M secure telecommunication technology (Telecom & Suppliers)
5. Trip related pull- and push-update cycles to enable user friendly AD services (Service Industry)

It is remarkable that the innovation mainly comes from in- and out-vehicle-centric sensor technology (1./2.) followed by telecommunication and ICT driven Cloud and IoT stakeholders (3./4./5.). Additionally, there are mobility services related to car-sharing and tourist venues for Urban Driving, e.g. in Versailles.

Based on the identification of these market segments, one finds an industry related Time-to-Market matrix for AD stakeholders (Table 6). As indicated in this table, enterprises will enter the AD market with products addressing various market segments at the same time, but with very different time-to-market scales depending mainly on the type of business and the size of the enterprise under consideration.

Table 6 Stakeholder analysis related to market segments and type of businesses identified

Type of Business	Type of Enterprise	Strength (S) / Weaknesses (W)	Time-to-Market
OEM & Suppliers	Large enterprise	(S) Robust and international manufacturing, worldwide sales and spare parts (W) complex go-to-market strategy due to huge investment risks and global supply chain	>3 years
ICT & I.T.S.	Mixed (large, medium, SME)	(S) Global standards for mobile communication and computing, fast moving global markets with high potential of innovation (W) extremely competitive markets with the need of a large number of users, dominated by US and Asia	>12 months, but <3 years
Service industry	SME/Start-up	(S) highly innovative market players following cost-profit entrepreneur spirit rather than political decisions and constraints of Large Enterprises, fast moving local markets with early adapters breaking into the new market horizons (W) shortage of Cash-Flow or unstable Cash-Inject leading to the “Valley of Death” path often occurring during the start-up phase for new companies	<12 months

Besides these very general factors influencing the product life cycle of the different industries under examination, the AUTOPILOT AD services will be studied in a traditional SWOT analysis in order to identify potential for business exploitation of the different AUTOPILOT AD pilot site services.

To better understand the differences in strengths and weaknesses of OEM industry compared to ICT/IoT and SMEs, a SWOT comparison between the three market segments is presented leading to a prognosis of what might be expected with regards to go-to-market strategies by different stakeholders. For this purpose, we will start to analyse according to SWOT methodology first of all AD services introduced by OEM and/or suppliers.

8.4.2 Data collection Means

Data collection for the needs of the Business analysis will be performed by different means and through different channels. Data from pilot sites will mostly derive through dedicated workshops, during the pilot tests implementation period, in which both pilot participants and related stakeholders will be involved. These Workshops will be interactive, where the Consortium members will present the project and – specifically – the pilots contents and address the procedures, while the pilot participants and stakeholders will be asked to provide feedback by different means, such as answering short questionnaires, participating in the MCA to provide ratings to different alternatives according to pre-set criteria, etc. All these procedures will be defined in detail in deliverable D4.5.

Moreover, the results from the other assessment areas (user acceptance, technical, quality of life) will provide useful input to the overall business assessment, through the investigation of the inferences for business deriving from their results.

8.4.3 Inferences from other assessment areas

As business impact assessment relates to the final product and its potential place in the market, it is essential that it takes into account also the results of the other assessment areas. For example, user acceptance is directly linked to the customers that the product seeks to attract, technical evaluation provides indication of the time that would be needed for the product to reach the market, etc. Some relevant considerations are listed below:

- Technical evaluation gives assessment of technology readiness
- Technical evaluation gives assessment of response time of AD functions
- User Acceptance assess when (response time) users prefer to drive themselves or use AD
- User Acceptance evaluates the user experiences, needed training, etc.
- Quality of Life assessments give input on the potential benefits due to impacts on traffic safety, efficiency and mobility

For this reason, all other assessment areas should keep this in mind and provide, as a conclusion to their analysis (on preliminary and final stage) some indications on the inferences of each to the business sector. This task should be coordinated between the different assessment areas and the results will be communicated to Task 4.3, to be included in the overall business assessment.

8.5 Output to Business Exploitation (Task 5.3)

The objective of the Business Impact Assessment Methodology described here in section ? (in combination with Annex 14.1) is to facilitate the large-scale market uptake of the business solutions developed in the context of the pilot sites. Generic business exploitation models for large-scale deployment can only be based on reliable KPIs reflecting the key stakeholders and success factors for them to push the AUTOPILOT services into the upcoming market for IoT assisted Automated Driving services.

In AUTOPILOT, robust KPIs for the piloted technology can be developed by SWOT analysis and questionnaires with the AUTOPILOT partners from the 3 stakeholder groups mentioned in section 8.4.1. AUTOPILOT representatives from key industry sectors, public authorities and public transport operators will be included as early as possible in order to find out their specific behaviour in the AD market, especially:

- Is a stakeholder a defender, promoter, apathetic or latent (Figure 16)?

- In case of a stakeholder being a promoter, how can he be managed closely to bring market break-through?
- In case of stakeholder being an opposing defender, how should he be informed to avoid severe time-to-market delay?
- What type of apathetic and latent stakeholders were found in the specific pilot sites?
- What are preconditions and key enablers found in the Business Impact Assessment leading to sustainable business exploitation models and roadmaps for large scale deployment?
- What are barriers to short-term and long-term developments found in the Business Impact Assessment leading to sustainable business exploitation models and roadmaps for large scale deployment?
- Are there stakeholders in the pilot site environment, which might have interest to push new products and technologies towards standardisation, e.g. telecommunication industry?
- What kind of members should be invited to join the working group planned to ensure good information exchange between the project and the automation community, e.g. public transport stakeholders from UITP, where also T-Systems has membership status?
- How should pilot sites deliver their road map for the longer-term deployment of the piloted use cases at each pilot location?

Deliverable D4.1 gives Task 4.3 (Business Impact Assessment) the methodology to assess KPIs for business impact which shall lead to the Business Exploitation Plan in deliverable D5.5.

The business exploitation plan will be based on robust KPIs coming out of Research Questions and Hypothesis leading to measurable indicators. Currently, KPIs focus on stakeholder analysis and questionnaires which address the market behaviour towards innovation. As mobility service operators represent a mix of OEM, Public Transport Operators but also Rental Car Companies and Car Sharers, a close look must be taken when it comes to their competitive market behaviour.

For OEMs, robust KPIs are linked to commercial activities in the rental car and car sharing service market, where a sales channel for new models and innovation can be tested bringing innovative functions to market. Questionnaires should be based on user acceptance and end user willingness to pay for IoT-AD services compared to AD services without IOT as well as reluctance towards innovative AD services. Robust KPIs are expected to be derived out of studies related to:

- Mobility behaviour with and without IOT-AD
- Acceptance of Smartphone usage, including user acceptance for Internet booking, P&R or shuttle services in urban driving and tourism, i.e. Automated Valet Parking
- Transaction fee models for Highway Pilot and Platooning services
- Stakeholder behaviour of OEM towards Rental Car and Car Sharers
- Stakeholder behaviour of Public Transport Operators towards innovative services such as urban shuttle operation complementary to established bus lines
- Event management and mobility service operation, i.e. Automated Valet Parking during large scale events

Once the hypothesis and research questions are addressed via Questionnaires to market stakeholders, output will give pilot site partner a suggestion for further business exploitation.

9 Piloting Scenarios

This section collects the input on pilot test scenarios from the evaluation methodologies of the previous sections.

9.1 Pilot test scenarios

The current assumptions on the iterative approach and scale of pilot tests are described in section 4.3. The pilot test scenarios, planning and iterations of pilots, and the incremental updates of automated driving functions and services still must be decided.

Pilot scenarios should be defined from at least two perspectives evaluation perspectives:

- **User focused test sessions**, in which as user experience is created that best reflects the intended end solution of automated driving. This experience is most suited for evaluating the acceptance of users and test drivers. Some guidelines on user selection and pilot setup are described in section 6.5. Ideally, a variety of normal drivers can experience a test run. Alternatively, only test drivers are allowed to operate the automated vehicles, then normal drivers could experience as passengers. Note that legal and ethical issues need to be resolved for taking passengers.

  These test runs are only useful if the automated functions and services are mature. Any incidents or interventions would negatively affect user and stakeholder acceptance during interviews and surveys. A test scenario should comprise:

  - A test run without IoT enabled functionality and services as a baseline, followed by a run with IoT, such that the user can experience the added value of IoT. This scenario is most relevant for IoT enhancing use case.
  - A single test run in which the user can experience automated driving for accelerating or enabling IoT technologies.

  Ideally, a user can experience all use cases in a single session per pilot sites. The users receive a briefing prior to the test session, and a debriefing immediately after the test session to collect their feedback either in the form of a questionnaire or interview. See the next subsection 9.2 for the organisation and preparation. The pilot site should collect any relevant information on events, situations, and road incidents (section 5.2.6). This information is immediate input for interviews with users during the debriefings, or for reference for evaluating the questionnaires filled in by users during the debriefings. Debriefings with interviews are proposed to be organised in conjunction to test site events suggested in section 9.2.2

- **Technical test sessions** are needed in addition to the user-oriented sessions. The objective of technical test sessions is to evaluate the improvements in functionality and technology with alternative system implementations:
  - Test runs for a baseline scenario without new IoT functionality versus a test run with new IoT functionality enabled, in case of accelerating or enhancing IoT.
  - A test run with enabling IoT functionality, for which the functionality and performance is measured and assessed, and possible improvement for system updates of alternative solutions can be assessed between test runs.

  Scenarios for test runs must be defined in more detail. Two sets of test scenarios are proposed:

1. First scenarios are defined for controlled technology tests for the concepts of section 5.2.
2. Once the functionality and performance of the basic technologies are evaluated in 1., then scenarios must be defined to test the added value for the use cases. An initial set of situations and events to define scenarios is sketched in section 5.3.

9.2 Users and stakeholders' activities

The following subsections describe several activities for evaluation purposes to be organised in conjunction with pilot test scenarios for users and stakeholders.

The evaluation team will develop the input for the briefings, debriefings, workshops, questionnaires and surveys in English, and adapt these in collaboration with the pilot sites to their specific situation and test cases. Translation of the material provided by the evaluation team in the local language, adaptation to the local test situations, and organisation of the activities are considered the responsibility of pilot sites.

9.2.1 Briefings and debriefings during pilot test sessions

Ideally, a user can experience all use cases in a single session per pilot sites. The users receive a briefing prior to the test session, and a debriefing immediately after the test session.

Before starting the test session, users are briefed on the scenarios and how to participate. Animations or videos may be used to introduce the concepts, and to put the current pilot setup in perspective of a future vision on automated driving and IoT. Depending on the purpose of the test sessions, the technology and expected situations and events should or should not be introduced to the users. For example, it may be necessary to prepare users how, where and when to observe the effects of IoT on automated driving. It may also be necessary to prepare users for their expected input in the debriefing.

A session is concluded with a debriefing of all users and test drivers collectively or individually. The situations, events and road incidents (section 5.2.6) collected during the test sessions will be explicitly evaluated in the debriefings. The debriefings will focus on the subjective input from the users and test drivers required for assessment of user acceptance as described in section 6.7, quality of life (section 7.5), and business impact (section 8.4), and to note any legal issues. Input is collected in either of two forms:

- The user can fill in a questionnaire that will be processed later by evaluators.
- The evaluators will interview the user.

The latter will be organised in the one or two major events foreseen per pilot site, where evaluators will also participate. Otherwise the pilot sites can use the first option to collect feedback and forward the questionnaires to the evaluators via the central data management server.

9.2.2 Workshops

The purpose of workshops is to organise discussion with both users, pilot site participants, stakeholders, associated partners, and other stakeholders. The objective of the workshops it to collect subjective input from the users and test drivers required for assessment of user acceptance as described in section 6.7, quality of life (section 7.5), and business impact (section 8.4), and to note any legal issues for evaluation Task 4.6. The scale and scope of the workshop is on the future perspective of IoT and AD instead of the single test session, and that input from stakeholders is combined with user experiences.

These Workshops will be interactive, where the Consortium members will present the project and - specifically – the pilots contents and address the procedures, while the previous pilot participants and stakeholders will be asked to provide feedback by different means, such as answering short questionnaires, participating in the MCA to provide ratings to different alternatives according to pre-set criteria, etc.

The workshops can be organised as focus group workshops for specific topics or evaluation objectives related to the specific use cases of the pilot sites.

At least two workshops are proposed to be organised in each pilot site. The timing is determined by the purpose of the two workshops and the maturity of the implemented use cases:

- Early during piloting, when a user can experience the initial automated driving system, and
- At the end of piloting when all IoT enabled functions are operational.

Evaluators will also be available during these events to prepare and participate in the workshops and user interviews. If possible, workshop participants can experience the use case test scenarios or demonstrators as passengers from the general public, and participate in the debriefings, questionnaires, interviews, and focus group workshops. It seems logical to combine these with the public events. It should be noted though that the planning of workshops is determined by the status and maturity of the IoT and AD technology to be demonstrated in order to collect relevant feedback from users and stakeholders instead of publicity for the AUTOPILOT project.

9.2.3 Surveys

In addition to the workshops, surveys are foreseen to collect subjective input from a wider group of stakeholders, including users, required for assessment of user acceptance as described in section 6.7, quality of life (section 7.5), and business impact (section 8.4), and to note any legal issues.

10 Data Provisioning and Quality

This section presents requirements for the provisioning of log data as input for technical evaluation from the central data management server. The process and interface between the central data management server from Task 3.4 and evaluation is described in section 4.4.

Additional situational data is described in section 7.5 and will be added later in Annex 14.2 during the development of the pilot test scenarios. The provisioning of additional questionnaires and survey will be defined later upon development of these artefacts.

Provisioning of log data is subject to some general assumptions and requirements:

- All log data is time synchronised (see section 10.1.3) and time offsets and synchronisation accuracies for data feeds should be verified and specified for evaluation.
- The accuracy, frequency and latency of data feeds, sources or parameters also need to be specified for evaluation.
- Log data is provided in standard formats that can be accessed and processed with standard (preferably open source) tools, such as in SQL databases, CSV, XML or TXT files. For specific data feeds it is appropriate to provide the data in RAW format with its own database.

10.1 Log Data

10.1.1 Timestamp

All data should be timestamped. Preferably a single time reference is adopted for all data provided to and in evaluation to avoid any misinterpretations and conversions between local time zones, sensor systems, and time units.

Requirement 1: Universal Time Coordinated (UTC) is proposed as the single time reference.

UTC is measured in milliseconds since 1 January 1970 and includes leap seconds since then. UTC is the default for system time in Unix and Linux based systems for example.

It should be noted that UTC is not the same as GPS time or TAI time, even though these can be converted unambiguously into UTC (see http://leapsecond.com/java/gpsclock.htm).

TAI is defined for ITS-timestamps in ITS messages such as the CAM, DENM and IVI. Since the TAI ITS-timestamps are essential for uniquely identifying messages, these TAI timestamps should be logged.

Requirement 2: For uniformity in timestamp processing for evaluation, however, it is advised to also include the timestamp converted to UTC.

It should also be noted that CAM messages use the *generationDeltaTime* data element, which cannot be converted into an absolute timestamp except by the sender upon generation of the message.

Requirement 3: ITS-Stations are required to also log the absolute timestamp of the CAM *generationDeltaTime* in UTC when logging a CAM.

10.1.2 Measurement identification

A measurement is defined as the parameter to be logged and provided for analysis, indicator calculation and evaluation. A measurement can be logged frequently as a time series of measurements, such as the position measurements of a GPS sensor. A measurement can also be event-based, such as the reception of a message, a detection or a decision of an application.

To identify measurements, the following simplified system architecture is assumed:

- A station is a logical and physical entity such as a physical road user, on board unit or IoT device.
- An application is a software application or component in a station that provides logging.

A station can have one or more applications that log data. Applications may physically reside on a single hardware or logging unit, or on different units, while each unit has its own system time. An application may generate its own set of log files with specific measurements, may have its own process to collect and store the data.

Alternative applications may log measurements for similar measures. Position data for example can be logged for multiple positioning sensors and systems. Each of these sources could be defined as a separate application and log the same measure. Evaluation will consider every measure from every application and station as a separate measure; i.e. every <station, application, measure> is unique. This allows collecting all data from all stations and pilots in a single data management system.

Requirement 4: Every measurement should be complemented with the attributes from Table 7 to uniquely identify the measurement.

Table 7: Required identification attributes per measurement

Attribute	Description
log_stationid	Unique id of a host station logging the measurement
log_applicationid	Unique id of the application logging the measurement. The application id is at least unique within the station.
log_timestamp	Timestamp at which the application logs the measurement. This is not the timestamp at which the measurement was generated, received or detected.

10.1.3 Time synchronisation

Analysis of time series and events from different data sources requires the time-alignment of the logged data. It is virtually impossible to fix misalignments afterwards during evaluation, and if time synchronisation issues are suspected during evaluation, it is virtually impossible to objectively identify which data sources are time-synchronised and which data sources have a (varying) time offset.

Requirement 5: All stations and applications that log data must be time synchronised within 100 msec accuracy.

Requirement 6: The time synchronisation of all logging units need to be verified (Task 2.5, Task 3.2) before piloting and data collection for evaluation starts. It is also assumed that time synchronisation is monitored and maintained during piloting.

10.1.4 Processing load

Processor overload is a well-known cause for delays in the processing of communication and automated driving functions. To distinguish these effects from communication delays and IoT platform processes, the CPU and internal memory loads of processing nodes (e.g. stations and applications) should be monitored and logged.

Requirement 7: CPU and memory loads should be logged on all station processing units involved in automated driving and communication

10.1.5 Motion State

The motion state of a station is defined by a time series of position measures. All positions are logged in WGS'84 coordinates with a latitude and longitude. The altitude is optional, and deemed irrelevant for current evaluations.

Position measurements may be obtained from various sensors. For the evaluation of positioning accuracy, it is important to log the positions as separate measurements (different parameters) for all sensors. For evaluation of positioning and motion state estimation, it is also relevant to log the measurements for all speed, acceleration and rotation sensors as separate measurements.

Requirement 8: For the evaluation of positioning and motion state estimation accuracies, the position, speed and acceleration sensor measures must be logged as separate measurement parameters. The accuracy measures, such as 95% confidence accuracy measures, should also be logged.

For IoT devices that are not subject to positioning evaluations, it is not necessary to log position or speed measurements from positioning devices or systems, if the positions are already logged in communication messages such as CAM, and if speed and acceleration are also logged or can be derived from these messages. It is assumed here that the IoT device is time synchronized and that the position and speed accuracies in the messages are verified in Task 2.5 or Task 3.2.

10.1.6 Communication messages

Communication performance measurements are collected from the messages at the facilities or application layers in stations and servers. This applies to V2X messages as well as for IoT messages.

The communication between various IoT devices (other than the devices of road users) and IoT platforms is not directly measured and evaluated, such as the communication between road side sensors and drones in the cloud. This communication is indirectly evaluated as it is included in the end-to-end delay from detection time at the IoT device till the reception of the detections and derived information in the automated vehicles.

On the same note, the communication within a vehicle, and between communication layers within a station, are not measured and evaluated directly either. The net effects of communication performance within and between in-vehicle systems will be evaluated in terms of delays in application decisions and actions, and the overall automated driving performance such as positioning improvements.

To evaluate the performance of communication, the locations and timestamps upon sending and reception should be logged. To extract motion states or to evaluate use case related information, (part of) the message contents should also be logged. The following approach is proposed to minimise the required logging resources:

- The relevant contents of messages need only be logged once, typically by the sender.
- Receivers only need to log the message elements to uniquely identify the message. Examples of identification elements are given in Table 8 for ITS-G5 messages. Message identification elements need to be defined for over communication channels and IoT messages in a similar way.

Table 8: Communication message identification

Message Type	Mandatory data elements
CAM	• stationid • generationdeltatime
DENM	• originatingstationid • sequencenumber • referencetime (in TAI)
IVI	• stationid • countrycode • serviceprovideridentifier • iviidentificationnumber • timestamp (in TAI)

The approach assumes the following requirements are met.

Requirement 9: **Every message is logged with a communication action to identify whether the message is 'SENT' or 'Received" by the station**

Requirement 10: **An application logs the timestamp when the message is sent or received.**

Requirement 11: **An application logs the communication medium that is used to send or receive the message, e.g. ITS-G5, LTE, LTE-V2X, NB-IoT, 6LoWPAN, ….**

Requirement 12: **A sender logs all message data elements that are relevant. Relevance is defined by mandatory data elements in the message for a specific station type, application type, use case, or test scenario.**

Requirement 13: **A receiver logs at least the data elements to uniquely identify the message and to unique retrieve the relevant message from the logging of the sender.**

Requirement 14: **The mandatory data of all stations and applications is provided for evaluation, such that all relevant message contents is available at least once, and can be uniquely identified**

Requirement 15: **The messages are logged as specified in the standard, including the structure, data element names and types. Messages may be logged in encoded form, or in decoded form in standard format such as XML, JSON, SQL, and simple messages can be provided in text.**

Data frames and data structures must be normalised when logged in flat structures or SQL.

Data frame and element names have following restrictions:

- **Alphanumeric**
- **Spaces and special characters need to be replaced or omitted**
- **To avoid conversion errors between tools, data element names may be logged in all-small letters (instead of Camel case)**

10.1.7 Events

Event in communication, applications and user interaction, and processes for IoT platforms and cloud services, are specified in models per use case, scenario and situation in pilot test scenarios.

The objective of the model is to avoid implementation specific logging systems and data analysed where possible for evaluation.

The processes for communication, applications, user interaction, and processes for IoT platforms and cloud services, etc. can be generalised to identify:

- Situations or application states
- Events that are detected within these situations
- Actions or decisions made by the application or service logic.

Table 9 gives an example of situations that can be distinguished per use case.

Table 9: Situations in use cases

Use Case	Relevant situations
Automated Valet Parking	• Drop off vehicle • Routing of vehicle • Legacy car or VRU detection • Scheduling of vehicle • Parking manoeuvre • Request and pick up vehicle
Highway Pilot	• Detection of road condition • Emergency braking / slow vehicle • Breakdown or accident • Fast approaching emergency vehicles • Traffic jams and queues • Nearby presence of VRU • Weather related condition • Transition from AD to manual
Platooning	• Platoon scheduling and organisation • Platoon forming process • Interaction with legacy traffic • Lane allocation
Urban Driving	• Single and multiple intersections • VRU interactions • Interaction with traffic lights • Interaction with legacy cars • Interaction with environmental data • Routing to free on-street parking spot
Car Sharing	• Scheduling and sharing car • Waiting time from request to pick up • Car-customer assignment • Assignment of vehicles to users • Events detection during route • Monitoring for parking spot occupancies during route

Within a situation, events, actions and decisions can be identified and predefined as a list or enumeration that is specific to the situation and use case. The events within a situation will occur and be logged with a timestamp and other details as a sequence that corresponds to the concepts in section 5.2 to be evaluated. Event models can be defined for example for:

- Track or path of positions
- Series of way points and route decision points

- Manoeuvring actions
- Driver intervention actions such as steering and braking
- Communication and interaction protocol actions
- Situation specific data management actions, like search actions and the classification of the relevance of data.
- Processing steps for security measures.
- Environment detections
- Safety interventions
- States, events and transitions in application process logic.

10.2 Safety related incidents

Any human intervention, e.g. by a test or co-driver, to disengage an automated driving mode, function or (safety-relevant) service in real-traffic conditions is considered as an incident that should be reported (see [D5.3 section 4.6, Table 12]). An overview of the incidents is specified in Annex 14.1.

10.3 Situational data

Situational data is data acquired about the situations and conditions under which pilot tests are performed, such as traffic and weather conditions. Situational data may be acquired from IoT data sources or from sources external to the pilot tests.

10.4 Surveys and Questionnaires

Surveys and questionnaire will be stored and provided from the central data management server as described in section 4.4.

11 Extending FESTA

FESTA is a methodology that was developed to design and execute FOTs and analyse the results, determining the impact of large-scale deployment of Intelligent Transport Systems. The original methodology focussed on in-vehicle systems, but was later also extended to cooperative systems.

When applying FESTA to road automation, new challenges arise, as no longer a specific system, targeting one function, is evaluated but a whole vehicle, containing a large set of different functions is tested.

AUTOPILOT applied the methodology for developing the field tests and came across the following challenges.

11.1 The context

AUTOPILOT is part of a programme dealing with Internet of Things, applying its concepts and architecture to automated driving. That means that on the one hand the pilot studies can be seen as part of a wider range of FOTs and pilots evaluating automated vehicles, applying the same methodological principles and approach to data collection and user evaluation. On the other hand, the purpose of AUTOPILOT is to evaluate the Internet of Things approach, for which there is no common methodology, as IOT can be used in many very different contexts.

11.2 Function identification and description

In FESTA it is assumed that functions are well-developed and implemented in systems that are near to market. In AUTOPILOT this is not so straightforward. Some of the functions tested in are relatively unambiguous, such as Automated Valet Parking, but others are more complex, such as the Urban Driving function. In D1.1 the specification and description of the functions are provided, but in the first draft it was not yet fully clear how functions were implemented. This is also due to the nature of the IOT approach, as in principle the vehicles can be connected to every other IOT object. This means that no longer just the functions related to the vehicle need to be specified, but also other objects, and leaving open the option that more connections can be made. For example, for Automated Valet Parking cameras from a parking lot can be used to gather information about space and possibilities for safe manoeuvring, but the specification of these cameras will probably become available at a much later stage. Also, iteration comes into play. During the definition of the study design, it may be deemed useful to connect other objects, even if there are not strictly necessary for the functioning of the system.

Another complication is the fact that no longer near to market systems are used, but prototypes that are still under development. By the time the project is finished, probably most of the systems will have been updated, improved, changed etc. This means that the research questions to be answered will need to be on a somewhat higher level. For example, if the evaluation finds out that there is some problem with object detection, this may already be an obsolete result at the end of the project because the vehicle is already equipped with more sensitive sensors.

11.3 Use cases

Use cases are at the heart of the definition of the tests at the pilot sites. AUTOPILOT put a lot of effort in defining the use cases. Several obstacles were encountered in this process:

- Use cases cannot be only defined by what is desirable to evaluate, but also depends on what is allowed to be tested on the real road. For example, sometimes the vehicle is only allowed to drive on a dedicated lane.

- Uses cases are evolving, because of the IOT approach, connections with other objects come into play, and because the functions, and their interactions, are not always fully understood and described.
- Uses cases can have the same name, but may be different in the way they play out. For example, several pilot sites use urban driving, but their use case descriptions are rather different. This means that it may be difficult to compare use cases over different pilot sites.
- In AUTOPILOT it required serious effort and interaction between the pilot sites and the evaluation team in WP4.1 to understand the use cases at a detailed level, and more time was needed for this than originally foreseen.
- Not only the use cases are of interest, but the whole scenario in which the use cases play a role. For example, platooning can be done on a highway for efficiency reasons, but also as a way in a city of rebalancing the shared cars. Although the function is the same, the scenario, and thus the focus of the study, can be very different.

11.4 Research questions and hypotheses

Experience from FOTs shows that defining and selecting research questions is one of the hardest parts of setting up a FOT. AUTOPILOT was no exception, but with an additional complication that IOT was the main focus. WP4.1 started with defining research questions based on the experiences many of us had from earlier FOTs, and focussing on automated driving (e.g. will the average speed go down?). Discussion was needed to find the right focus, namely how IoT could offer potential improvements to automated driving functions or driving modes, and to enable services involving connected and automated vehicles. When we defined the possible ways in which IOT can improve AD, namely by Accelerating, Enhancing or Enabling new services or automated driving functions, it became easier to define research questions. This distinction helps to focus on the future benefits of deploying automation, and steers away from the specific implementation and testing of functions.

In FESTA, the user is the driver of a vehicle with a system, the other people involved are stakeholders. In AUTOPILOT, and in automation pilots, this becomes more complicated. In full automation, the driver becomes a passenger, and in Automated Valet Parking or reshuffling of shared cars, there is no one in the vehicle. For example, for Automated Valet Parking, people may be interested in how the car is parked if they own the car, but not if it is a shared car. However, other people in the parking lot may be very interested, because they may be affected in their parking activity. The same holds for other road users. Stakeholders, such as people who provide a service, or who own a parking house, are also of interest in the evaluation. So, for each research question, it is important to define who the users are, or who the people are that are affected by the automation.

AUTOPILOT has four areas of evaluation: technical evaluation, user acceptance, quality of life and business development. All four bring their specific challenges in defining research questions.

- **Technical evaluation**. Technical evaluation is different from validation and verification, before the evaluation starts, WP 3 will ensure that everything is working as it should. Technical evaluation is focussed on the technical contribution of the IOT approach to automated driving. Its main purpose is to find out what the added value is of the IOT platform, how the communication works out, and whether the data needed for the functioning of the systems are received in-time and of the right quality. As we are dealing with systems that are not always mature, and because of the complex interactions within the IOT paradigm, technical evaluation becomes far more important than in former FOTs.

- **User acceptance**. Traditionally user acceptance is measured by asking a large number of people what they think of the system, whether it is useful, easy-to-use, how much they are willing to pay for it etc. In FOTs drivers usually drive for a long time (months or even years) with a system, so they gain full experience with it and are able to judge it merits. Also, driver behaviour is measured, such as how often they turned the system off, providing a measure for how useful it was for them. During the interactions with the pilot sites, it became clear that the pilot test would be far from the more naturalistic approach in former FOTs. A major reason for this is that most countries/regions do not allow vehicles on the road without a (potential) driver being present, who can takeover control if necessary. This means that most test runs will be performed by engineers from the pilot organisations, and that naïve users from the general public will take more of a passenger role. Also, the driving periods will be short, therefore having more the characteristic of a demonstration than a full-blown FOT. User impression can be measured, but full user acceptance is not possible in these cases. Recommendations to care for this situation is to try to have test vehicles with as many users (passengers) as possible, and have large-scale demonstrations so that at least as many people as possible have an impression of what automated driving looks like. Another solution is having engineers in the back seat where they can exercise control, but giving users the idea that the vehicle drives autonomously.

A difficulty with which we struggled was to define research questions focussing on IOT. Naïve users do not know and do not care whether an automated function is enabled by IOT or by another mechanism. Again, the distinction between Accelerating, Enhancing or Enabling was a good starting point to identify questions relevant for IOT.

Where in FESTA users are seen as subjects, whose behaviour in interaction with the systems is studied, we took a new approach. Instead of seeing users as subjects, we see them as co-designers, who are able to help to improve the systems, and who can contribute to Accelerating, Enhancing or Enabling new functions and services.

- **Quality of Life**. This is usually addressed in FESTA in the impact analysis, looking at the wider implications of ITS deployment on society. AUTOPILOT started with very large and ambitious sets of impacts. In FESTA, impact areas are safety, mobility, efficiency, and environment. Automation, and IOT, will have very wide impacts on society. In AUTOPILOT we have chosen to focus on personal mobility, sustainability (including safety) and well-being. The picture presented in Figure 12 with the dimensions for personal and social mobility illustrated how we can look at the costs and benefits for both individuals and society.

11.4.1 The bottom-up and the top-down approach in defining research questions

In the FESTA handbook, both a top-down approach, starting from the impact areas, and a bottom-up approach starting from scenarios derived from the use-cases. In WP4.1 we started with a top-down approach, the impact areas were already described in the Description of Work. However, research questions remained on a rather abstract level. A much more detailed insight in the uses cases was needed in order to be able to formulate research questions that could be answered. This bottom-up approach was very useful, as use cases are complex. Interaction with the pilot sites was important to be able to understand what scenarios would be possible. In evaluating ADAS systems it is easier to define specific scenarios when the system would be activated (e.g. in lane changes), scenarios in automated driving are far less determined, because so many things can happen in a traffic situation that are hard to foresee beforehand. The pilot-sites developed storyboards, trying to be as precise as possible, making it easier to define research questions. As the impact of automated driving, and the role of IoT, is more far-reaching than that of simple ITS, it is important to iterate between the two-approaches, trying on the one hand to be as precise and detailed as possible, and on the other hand not losing sight of the wider impact automation may have. In the analysis phase it will show whether we will be able to succeed in this.

11.5 Performance indicators and study design

From the research questions and hypotheses, performance indicators were defined.

One of the important tasks in the study design was to define the data that must be collected, as this needed implementation into the vehicles and/or communications systems. With a large variety of vehicle types, sensors, uses cases and communication platforms, it is important that data collected are harmonised, else it will not be possible to do any meaningful analysis on the data over pilot sites and use cases.

In the FESTA methodology, study design is derived from the questions the FOT must answer. In AUTOPILOT, due to the more explorative nature of the use cases, the study design is limited to what is possible. An example is the limitation of people who are allowed to operate the vehicles and the environments in which they are allowed to operate. As we are dealing with prototype systems or experimental vehicles instead of close-to-market systems, and with limitations on when and where vehicles are allowed to operate, it was not possible to put a large number of demands on pilot-sites for implementing an ideal study design. As discussed before, the study design will need to be more flexible, allowing for exploration, and iteration and revision during the project. One of the hardest demands was on trying to involve as many as possible naïve users, even if they are only passengers in a demonstration trial, in order to get as much feedback as possible from non-experts. Another demand was on including as much communication with other IoT objects as possible.

Baseline

The baseline question is a very difficult one in automated driving studies, what do you compare the automated driving with? With no automation, lower levels of automation? Is a baseline useful at all? These questions were discussed in several workshops from the FOT-Net and CARTRE coordination and support actions. However, there is no clear recommendation from FESTA. Ideally in AUTOPILOT a comparison could be made between automation with and without IoT. This is demand WP4.1 started to put on pilot-sites. In some cases, this will be possible, and will provide some interesting comparisons in the technical evaluation, such as the time it takes to get a message to and from the vehicle, the number of message exchanged etc. In other case it is not possible because there is no meaningful implementation available for operating the vehicles without IoT. It is also a consideration that given a limited amount of resources, not too much time should be spend on driving around in a baseline mode, because that would be limiting the number of experimental drives.

12 Conclusions

This deliverable concludes the work done in task 4.1 “Evaluation requirements and methodologies”, or “Evaluation methodologies” for short. This deliverable provides the methodologies for technical evaluation and the assessments of user acceptance and the impact on the quality of life, and the business impact. This provides the starting point for the implementation and execution of these four evaluation tasks 4.2 – 4.5 in the next project phase.

The FESTA methodology is adopted and applied to develop the evaluation methodologies. The deliverables from work packages 1, 2 and 3 provided the input, such as the use cases and storyboards, functional descriptions, architectures and technical specifications for implementing the Internet-of-Things (IoT) technologies and platforms for automated driving (AD) functions and services, communication, and the pilot site descriptions.

The evaluation methodologies are a first step in the extension of FESTA for IoT and AD that will be refined further from the practical knowledge and experience gained from the pilot evaluations in the next two project years.

This deliverable defines the research questions, hypotheses, key performance indicators and measures for each of the four evaluation tasks in sections 0 - �. The detailed specifications are provided in the form of living documents in Annex 14 that will be updated throughout the evaluation phase.

The FESTA methodology is applied both in a top-down approach and in a bottom-up approach. In the top-down approach, the input is used to refine the initial research questions into hypotheses and key performance indicators. Then in a bottom-up approach, additional workshops and storyboard discussions are organised with pilot sites and use case developers to refine the scope and focus of the evaluation methodologies. In this process, the most relevant and common concepts and criteria between pilot sites and use cases are identified, and used to select the most common and relevant hypotheses and research questions.

The focus for evaluation is defined in section 0 as the central research question:

“What is the added value of IoT for AD?”

that will be answered in the remainder of the project from the central hypotheses;

- IoT is *accelerating* the development and deployment of automated driving functions,
- IoT is *enhancing* the functionality or performance of automated driving functions,
- IoT is *enabling* new automated driving functions.

The bi-directional development process provides harmonisation in the evaluation methodologies, and ensures that the most relevant improvements will be evaluated consistently from all four evaluation perspectives. This enables to deliver a final evaluation of the feasibility, suitability and usability of IoT for AD on the most relevant key performance indicators for technical, user acceptance, quality of life and business impact.

To facilitate that the pilot sites can provide the essential input for evaluations, the relevant guidelines, requests and requirements for pilot test scenarios and data provisioning are collected in sections 0 and 0.

The pilot sites, the implementation of the use cases, and pilot test scenarios and plans are still under development. Consequently, the evaluation methodologies are not fully defined either, and must be refined in the evaluation tasks 4.2 – 4.5 and in close collaboration with the other work packages. Most notably, use case specific IoT messages and data flows, situations and events in the pilot test scenarios must be refined in the next project phase and iterations of piloting, as foreseen in the iterative approach to piloting and evaluation (section 4.3).

13 References

[1] FESTA handbook version 6, fot-net.eu/Documents/festa-handbook-version-6/

[2] FOT-NET library with loads of information from seminars, workshops and other documentation, fot-net.eu/library

[3] FESTA webinars and presentations explaining the main steps of the FESTA methodology fot-net.eu/library/?filter=webinars

[4] Report and presentations from workshop on a common methodology for road-automation FOTs and pilots, fot-net.eu/Documents/workshop-on-a-common-methodology-for-road-automation-fots-and-pilots/

[5] Barnard, Y., Innamaa, S., Koskinen, S., Gellerman, H., Svanberg, E., Chen, H., "Methodology for Field Operational Tests of Automated Vehicles", Transport Research Arena TRA, Warsaw, 2016, fot-net.eu/Documents/methodology-for-field-operational-tests-of-automated-vehicles/

[6] FESTA Performance Indicators-Measures-Sensors matrix, http://www.its.leeds.ac.uk/festa/downloads/FESTA%20D2_1_PI_Matrix_Final.xls

[7] CARTRE (2017), "User Awareness, Users and Societal Acceptance and Ethics, Driver Training: Position Paper", CARTRE Thematic Interest Group, unpublished.

[8] Adell, E. (2009), "Driver Experience and Acceptance of Driver Support Systems: A Case of Speed Adaptation", Lund University Department of Technology and Society.

[9] Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D. (2003), "User Acceptance of Information Technology": Towards a Unified View", MIS Quarterly, 27, 425-478.

[10] Madigan, R., Louw, T, Dziennus, M., Graindorge, T., Ortega, E., Graindorge, M., Merat, N. (2016), "Acceptance of Automated Road Transport Systems (ARTS): An Adaptation of the UTAUT Model", Transportation Research Procedia 14, 2217-2226.

[11] Innamaa, S.; Axelson-Fisk, M.; Borgarello, L.; Brignolo, R.; Guidotti, L.; Martin Perez, O.; Morris, A.; Paglé, K.; Rämä, A.; Wallgren, P. & Will, D. (2013). Impacts on Mobility – Results and Implications. Large Scale Collaborative Project, 7th Framework Programme, INFSO-ICT 224067. European Commission 2013 No: Deliverable D4.4.3).

[12] Draskóczy, M., Carsten, O. and Kulmala, R. (1998). Road safety guidelines. Deliverable B5.2 of CODE project (TR1103). Atkins Wootton Jeffreys, Birmingham, UK. http://cordis.europa.eu/telematics/tap_transport/library/code-b5-2.html .

[13] Kulmala, R. (2010). Ex-ante assessment of the safety effects of intelligent transport systems. Accident Analysis and Prevention, 42, 1359–1369.

[14] Bertoncello, M., Wee, D., "Ten ways autonomous driving could redefine the automotive world", http://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world

[15] "Automotive revolution – perspective towards 2030", Advanced Industries, January 2016, https://www.mckinsey.de/files/automotive_revolution_perspective_towards_2030.pdf

[16] "Ebit-Margen der DAX-Unternehmen* im Jahr 2011", Das Statisk-Portal, https://de.statista.com/statistik/daten/studie/223655/umfrage/ebit-margen-der-dax-unternehmen/

[17] De Brucker, K., Macharis, C. (2011), "Best Things First. The Application of Multi-Criteria Analysis to Derive Implementation Priorities for Innovative Road Safety Measures", Infrastructure and Safety in a Collabborative World, pp. 305-325.

[18] Saaty, T.L. (1988), "What is the analytic hierarchy process?", Mathematical models for decision support, pp. 109-121.

[19] Saaty, T.L. (1995), "Decision Making for Leaders", RWS Publications, USA

[20] Vlassenroot, S. & Brookhuis, K. (2014), "Socio-psychological factors that influence acceptability of intelligent transport systems: A model", Ashgate Publishing Ltd.

[21] Vlassenroot, S., Brookhuis, K., Marchau, V. and Witlox, F. (2010), "Towards defining a unified concept for the acceptability of Intelligent Transport Systems (ITS): A conceptual analysis based on the case of Intelligent Speed Adaptation (ISA)", Transportation Research Part F: Traffic Psychology and Behaviour, 13(3), pp. 164-178.

[22] Barnard, Y., Chen, H., Koskinen, S., Innamaa, S., Gellerman, H., Svanberg, E., Zlocki, A., Val, C., Quintero, K., Brizzolara, D., D5.4 Updated Version of FESTA Handbook, FOT-Net Data, "From Functions to Hypotheses", 31.1.2017, section 4.8.3, https://connectedautomateddriving.eu/wp-content/uploads/2017/08/2017-01-31_FOT-NET_FESTA-Handbook.pdf

[23] Common Vulnerability Scoring System (CVSS), https://www.first.org/cvss/specification-document

[24] Innamaa, S., Smith S., Barnard, Y., Rainville, L., Rakoff, H., Horiguchi, R. & Gellerman, H. (2017). Trilateral impact assessment framework for automation in road transportation. Draft 4 January 2017. ART WG.

[25] Milakis, D.; van Arem, B. & van Wee, B. (2017). Policy and society related implications of automated driving: A review of literate and directions for future research. Journal of Intelligent Transportation systems, 21 (4), 324–348.

[26] Nilsson, G. (2004). Traffic safety dimensions and the power model to describe the effect of speed on safety. Bulletin 221. Lund Institute of Technology, Lund University.

[27] Department for Transport (2016). Research on the impacts of Connected and Autonomous Vehicles (CAVs) on Traffic Flow. Stage 1: Evidence Review. Department for Transport, March 2016.

[28] Michon, J. A. (1985). A critical view of driver behaviour models: what do we know, what should we do? In: L. Evans and R. C. Schwing (Eds.). Human Behaviour and Traffic Safety, pp. 485-524. New York: Plenum Press.

[29] Calvert, S.; Wilmink, I.; Soekroella A.; van Arem B. (2016). Considering knowledge gaps for automated driving in conventional traffic. 4th International conference on advances in civil, structural and environmental engineering. Rome, Italy.

[30] Saffarian, M., de Winter, J., & Happee, R. (2012). Automated Driving: Human-Factors Issues and Design Solutions. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56(1), 2296-2300

[31] Malone, K.; Rech, J.; Hogema, J.; Innamaa, S.; Hausberger, S.; Dippold, M.; van Noort, M.; de Feijter, E. Rämä, P.; Aittoniemi, E.; Benz, T.; Burckert, A.; Enigk, H.; Giosan, I. Gotschol, C.; Gustafsson, D.; Heinig, I.; Katsaros, K.; Neef, D.; Ojeda, L.; Schindhelm, R.; Sütterlin, C. & Visintainer, F. (2014). Impact Assessment and User Perception of Cooperative Systems. Large-scale intergrating project, 7th Framework programme. Objective ICT-2009.6.2: ICT for Mobility of the Future. Deliverable D11.4.

[32] Aittoniemi, E.; Rämä, P.; Kuisma, S.; Malin, F.; Kahilaniemi, S.; Mustaniemi, A.; Coconea, L.; Schünemann, B.; Häusler, F.; Ordinez, R. (2016). Impacts of TEAM applications. Large-scale integrating project, 7th Framework programme. FP7-ICT-2011-8, objective ICT-2011.6.7.a Deliverable D5.5.1.

[33] Anderson, J. M., Nidhi, K., Stanley, K. D., Sorensen, P., Samaras, C. & Oluwatola, O. A. (2014). Autonomous vehicle technology: A guide for policymakers, Rand Corporation

[34] Cyganski, R., Fraedrich, E. & Lenz, B. (2015) Travel-time valuation for automated driving: A use-case-driven study. Proceedings of the 94th Annual Meeting of the TRB.

[35] Fraedrich, E., Lenz, B., and Wolf, I. (2016). User perspective on autonomous driving. A use-case-driven study in Germany. In : Arbeitsberichte. Geographisches Institut, Humboldt-Universität zu Berlin. Heft 187.

[36] Fagnant, D. J. & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. In: Transportation Research Part A, 77 167-181

[37] Lutin, J.M. (2015). The Challenge of Planning for a World of Automated Vehicles – Definitions, Characteristics, and Modal Use Considerations. Transportation Research Board 94th Annual Meeting.

[38] Milakis, D., Snelder, M., van Arem, B., Correia, G.H.s.A. (2016). Development and transport implications of automated vehicles in the Netherlands: scenarios for 2030 and 2050. In: European Journal of Transport and Infrastructure research, Vol. 17 (1), ISSN: 1567-7141.

[39] Saffarian, M., de Winter, J.C. F., and Happee, R. Automated driving: Human-factors issues and design solutions. In: Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting.

[40] Trommer, S., Kolarova, V., Fraedrich, E., Kröger, L., Kickhöfer, B., Kuhnimhof, T., Lenz, B. & Phleps, P. 2016. Autonomous Driving-The Impact of Vehicle Automation on Mobility Behaviour.

[41] Zmud, J., Tooley, M., Baker, R., and Wagner, J. (2015). Paths of Automated and Connected Vehicle Deployment: Strategic Roadmaps for State and Local Transportation Agencies. Texas A&M Transportation Institute, College Station, TX.

[42] Visser, F., Stappers, P., Van der Lugt, R., Sanders, E.: 'Contextmapping: experiences from practice', CoDesign, 2005, 1, (2), pp. 119–149

[43] Bødker, S.: 'Scenarios in user-centred design – setting the stage for reflection and action', Interact. Comput., 2000, 13, (1), pp. 61–75

[44] Heinilä, J., Strömberg, H., Leikas, J., Ikonen, V., Iivari, N., Jokela, T., Aikio, K., Jounila, I., Hoonhout, J. & Leurs, N. (2005). User-Centered Design. Guidelines for Methods and Tools. The Nomadic Media consortium, November 2005. http://www.vtt.fi/inf/julkaisut/muut/2005/UCD_Guidelines.pdf

[45] Kolarova, V., Müller, S., Koller, F. and Bahamonde-Birke, F. (2018). Potenziale, Herausforderungen und Wirkungen des automatisierten und vernetzten Fahren. In: Arbeitsberichte zur Verkehrsforschung. ISSN: 2513-1699

[46] Stevens, A. (2004). The application and limitations of cost-benefit assessment (CBA) for intelligent transport systems, in Economic impacts of Intelligent transport systems: Innovations and Case studies in Transportation Economics, Volume 8, 91-111. Elsevier

[47] Highway Traffic Noise: Analysis and Abatement Guidance, (2011), FHWA, https://www.fhwa.dot.gov/ENVIRonment/noise/regulations_and_guidance/analysis_and_abatement_guidance/revguidance.pdf

AUTOPILOT Deliverables:

[D1.1] Initial specification of IoT-enabled Autonomous Driving use cases, version 2.1, 30 June 2017

[D1.3] Initial IoT Self-organizing Platform for Self-driving Vehicles, version 0.8, 30 June 2017

[D1.5] Initial open IoT Vehicle Platform Specification, version 3.1, 3 October 2017

[D1.7] Initial specification of communication system for IoT enhanced AD, versions 1.0, 30/09/2017

[D1.9] Initial Specification of Security and Privacy for IoT-enhanced AD, version 1.0, 29 September 2017

[D5.3] Performance and KPIs for autonomous vehicles and IoT pilot impact measurement, version 0.7, 16 November 2017

[Storyboards] Presentations on storyboards for piloting use cases, on Project Place | Pilot Sites

14 Annexes

14.1 Annex 1 – Research Questions, Hypothesis, and Indicators

The research questions, hypotheses and indicators from sections 0 to ⍰ are defined and maintained in a spreadsheet on Project Place with filename “AUTOPILOT_WP4_RQ_HY_KPI_<version>.xlsx".

AUTOPILOT_WP4_RQ_HY_KPI_0.5.xlsx

14.2 Annex 2 – Data Requirements

The requirements for input data for evaluation from sections 0 to ⍰, and section 0 are defined and maintained in a spreadsheet on Project Place with filename “AUTOPILOT_WP4_DataReqs_<version>.xlsx”.

AUTOPILOT_WP4_DataReqs_0.5.xlsx